Plasticité moléculaire de deux écotypes de pin maritime soumis à un stress osmotique

Philippe Chaumeil

▶ **To cite this version:**

Philippe Chaumeil. Plasticité moléculaire de deux écotypes de pin maritime soumis à un stress osmotique. Biochimie [q-bio.BM]. Université Henri Poincaré - Nancy I, 2006. Français. NNT : . tel-00115272

HAL Id: tel-00115272
https://theses.hal.science/tel-00115272

Submitted on 20 Nov 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

FACULTE des SCIENCES et TECHNIQUES
U.F.R. Sciences et Techniques Biologiques
Ecole Doctorale : Ingénieries des Ressources, Procédés, Produits et Environnement
Département de Formation Doctorale : Biologie Végétale et Forestière

Thèse

Présentée pour l'obtention du titre de

Docteur de l'Université Henri Poincaré, Nancy I

en Biologie Végétale et Forestière

par **Philippe CHAUMEIL**

Plasticité moléculaire de deux écotypes de pin maritime soumis à un stress osmotique

Soutenance publique le 13 Avril 2006

Membres du jury :

Président :	P. DIZENGREMEL	Professeur des Universités, U.H.P., Nancy I
Rapporteurs :	B. GARBAY	Professeur des Universités, ESTBB, Bordeaux
	M. ZIVY	Chargé de Recherche, CNRS, Gif sur Yvette
Examinateurs :	O. BRENDEL	Chargé de Recherche, INRA Nancy
	P. LABEL	Directeur de Recherche, INRA Orléans
Directeur de thèse :	C. PLOMION	Directeur de recherche, INRA Bordeaux

INRA-UMR BIOGECO – Equipe de Génétique – 33612 CESTAS Cedex – France

Remerciements

~

A la fois introduction et conclusion de la thèse, ils permettent avant tout de mettre l'accent sur l'aspect humain de tout projet de recherche qui est l'aboutissement de nombreuses contributions en particulier dans le cadre de cette thèse. Je tiens à remercier ici toutes les personnes qui ont participé à son élaboration.

- Antoine Kremer qui m'a accueilli dans son laboratoire où j'ai eu la chance d'effectuer mon DEA et de poursuivre en thèse. Tous ceux qui sont un jour passés à «Pierroton » ont remarqué son mode de fonctionnement ouvert aux échanges entre personnes et entre équipes qui est si enrichissant.
- Christophe Plomion qui a dirigé cette thèse, pour la confiance qu'il m'a accordée et la liberté dans le choix d'orientation et de réalisation de mes axes de recherche. Je tiens à souligner la façon dont Christophe sait valoriser, motiver et développer les compétences des membres de son équipe. Merci pour ta réactivité, la disponibilité dans les moments délicats et surtout d'avoir rendu possible de mener à son terme cette thèse. Merci également de savoir rester proche des étudiants et pour les bons moments passés à Salles.
- Concernant ce mémoire, je tiens à remercier Bertrand Garbay et Michel Zivy d'avoir accepté d'en être les rapporteurs.
- Oliver Brendel, Philippe Label et Pierre Dizengremel d'avoir accepté de faire partie de ce jury.
- Sébastien Duplessis, Virginie Garcia, Annick Moing, Christophe Rothan pour leur participation active aux comités de thèse.

~

Pour leur contribution directe à la réalisation de ce travail, je voudrais remercier :

- Patrick Léger pour toutes les discussions constructives, les améliorations dans le fonctionnement du labo et sans qui nous n'aurions pas pu générer les ESTs de racines.
- Aurélien Barré, Nacer Mohellibi et l'équipe du CBiB qui ont réalisé le pipeline d'annotation et ont répondus à mes nombreux coups de téléphones.
- Christophe Boury aka « Le Jeune » pour les heures passées à découvrir les joies de l'extraction, la qPCR et surtout l'annotation! Je conserverai également un bon souvenir des soirées « barbec » et de tes talents artistiques de montage photo qui font encore partie des murs ou plutôt des portes !
- Marie Foulongne avec qui j'ai eu plaisir à travailler et à discuter … des extractions d'ARN, des puces à ADN et du reste !
- Virginie Garcia et Johan Petit pour avoir partager leur bureau, leur bonne humeur, leurs conseils et leur expertise sur la réalisation et le traitement des données des puces.
- Manon Moreau qui a réalisé avec Céline Lalanne l'analyse du protéome (alors, ça polymérise jamais dans les tuyaux ? ;-)) et qui est devenue accro à la qPCR.
- Silvia Fluch (société PICME) qui a réalisé les puces.
- Jean-Marc Frigério qui a réalisé les banques de racines et d'aiguilles m'a convertit au Perl ("=~('(?{'.('+[@.)}'^'])@]]').'"'.('+(@,/^}'^'[@)@@+_').',$/})')) mais pas encore au Pingouin…
- L'équipe américaine de Ron Sederoff qui m'a accueilli chaleureusement.

Je n'oublie pas non plus tous ceux qui ont su répondre présents pour m'aider tout au long de la thèse et ceux qui sont devenus plus que des collègues de travail.

En particulier Céline qui doit toujours se demander si je ne vais pas tomber dans les pommes ;-) ; Manon pour sa gentillesse et la bonne humeur qu'elle sait apporter autour d'elle ; Agnès et David Chagné pour les fous rires et l'ambiance inimitable qu'ils mettaient dans les labos ; MH qui m'a ouvert les portes de son club ldp VIP et m'a soutenu dans les moments difficiles ; HCl puis Franck Salin avec qui j'ai partagé mon goût pour la bidouille informatique ce qui nous a valu le surnom de « Megabytes » … ; Fred qui a le mérite de supporter mes visites inopinées dans son bureau, la petite tape dans le dos et qui m'a permis de lâcher la pression ; Greg qui m'a initié à l'ARN et m'a fait découvrir le surfcasting et Jeannot pour ses blagues et la pêche sur le bassin.

Je remercie également Jeremy avec qui j'ai partagé le bureau et qui a supporté ma vision radicalement différente d'un bureau de travail. Je m'incline devant ta capacité à conserver un bureau parfaitement rangé et classé. Merci aussi pour ta bonne humeur et ton sens de l'humour. Et n'oublies pas, selon le vent, la voile :-D.

Mes remerciements iront également à Jérôme et à Jorge pour le soutient moral entre étudiants en fin de thèse, Manuela pour les discussions tardives, Carlos et sa petite famille pour les soirées Pisco. Je remercie aussi les « secrétaires » Annie, Corinne et Florence qui tentent de nous décoder les mystères administratifs. Merci à Annie et Valérie pour leur maîtrise de l'orthographe et de la grammaire. Merci à tous les permanents, non permanents et visiteurs qui ont contribué à ce que je ne vois pas passer ces 5 années à pierroton. Il m'est impossible de les citer tous, j'espère que les bretons et non bretons (hein JMG !) se reconnaîtront, de même que les argentins, australiens, brésiliens, chiliens, espagnols, italiens, japonais, mexicains (de souche ou de cœur) et portugais.

Une dédicace spéciale à Dje et Florence de Nancy car ce sont de vrais Amis que je me suis fait là bas et avec qui j'ai partagé des moments forts.

Un clin d'œil de PKM à la team oncle Albert, je remercie Cyborg pour avoir réussi à rapprocher deux électrons libres, Pikatchu pour ses connaissances de la plateforme de Toulouse ;-) et LoL qui est allé se perdre au pays du burger.

Je terminerai en remerciant mes parents et ma famille qui m'ont permis d'arriver jusqu'ici et surtout Gwen, Sterenn a Garantez, qui a supporté au quotidien les hauts comme les bas depuis maintenant 4 ans et sans qui je ne conçois plus mon avenir.

Abréviations

~

Liste des principales abréviations utilisées :

ABA : Acide abscissique
ACC : 1-aminocyclo-propane-1-carboxylic acid
ADN / DNA: Acide Desoxyribonucléique
ADNc : ADN complémentaire
AGP : Arabinogalactan proteins
ARN / RNA : Acide Ribonucléique
ARNm : ARN messagers
ARNt : ARN totaux
ASR : ABA-, Stress-, and Ripening-induced protein
Blast : Basic Local Alignment Search Tool
CCoAOMT : Caffeoyl Coenzyme A 3-O-methyltransferase
CN: Consensus
CL : Cluster
CT : Contig
COMT : Catechol-O-methyltransferase
EF : Ecotype Français
EM : Ecotype Marocain
EST : Expressed Sequences Tags
FT : Facteur de Transcription
GA : Gibbérelline(s)
Hsp / HSP : Heat-shock proteins
LEA : Late embryogenesis Abundant protein
miRNA : microARN
PA : Polyamine(s)
PAL : Phenylalanine ammonia-lyase
PCR : Polymersae Chain Reaction
PEG : Polyéthylène glycol
QTL : Qunatitative Trait Loci
qPCR : PCR quantitative en temps réel
RFO : Raffinose family oligosaccharide(s)
SC : Expérience « Stress Court »
SL : Expérience « Stress Long »
SNP : Single Nucleotide Polymorphism
SOD : Superoxyde Dismutase
TC : Tentative Consensus
XTH / XET : Xyloglucan endotransglucosylase/hydrolases
4CL : 4-coumarate-CoA ligase

Sommaire

A INTRODUCTION ---------- ***2***

A.1 Un climat en cours de modification ---------- **2**

A.2 Hydrologie et géologie de la région landaise ---------- **3**

A.3 Des marais à la sécheresse, un équilibre fragile ---------- **4**

A.4 Des incertitudes sur la durabilité du massif forestier aquitain ---------- **5**

A.5 Objectifs ---------- **6**

A.6 Plan de la thèse ---------- **8**

B SYNTHESE BIBLIOGRAPHIQUE ---------- ***12***

B.1 Le Pin maritime ---------- **12**

B.1.1 Botanique et Taxonomie ---------- 12

B.1.2 L'appareil racinaire ---------- 14

B.1.3 Répartition géographique ---------- 15

B.1.4 Habitat ---------- 16

B.2 Les plantes et la sécheresse ---------- **17**

B.2.1 L'eau et la plante ---------- 17

B.2.1.1 L'eau du sol ---------- 17

B.2.1.2 L'eau dans la plante ---------- 18

B.2.1.3 L'eau et les racines ---------- 19

B.2.2 Le stress hydrique ---------- 22

B.2.2.1 Les stratégies de réponses ---------- 22

B.2.2.2 Les systèmes modèles ---------- 23

B.2.2.2.1 Systèmes tolérants ---------- 23

B.2.2.2.2 Modèles génétiques ---------- 25

B.2.2.2.3 Plantes de grande culture ---------- 26

B.2.2.3 Réponses morphologiques et physiologiques ---------- 26

B.2.2.3.1 Régulation de la croissance ---------- 26

B.2.2.3.2 Croissance racinaire ---------- 28

B.2.2.3.3 Les hormones ---------- 30

B.2.2.3.3.1 Acide abscissique ---------- 31

B.2.2.3.3.2 Ethylène ---------- 34

B.2.2.3.3.3 Polyamines (PA) ---------- 36

B.2.2.3.3.4 Les auxines ---------- 38

B.2.2.3.3.5 Les Gibbérellines ---------- 38

B.2.2.3.3.6 Cytokinines ---------- 39

B.2.2.4 Taille du génome et adaptation au stress hydrique ---------- 39

B.2.2.5 Gènes et voies métaboliques régulées au cours du stress ---------- 40

B.2.2.5.1 Méthodes de détection des gènes régulés ---- 41
B.2.2.5.2 Détoxication ---- 42
B.2.2.5.3 Fonctions chaperonne ---- 43
B.2.2.5.3.1 Les Hsp (Heat-shock proteins) ---- 43
B.2.2.5.3.2 Les LEA (Late Embryogenesis Abundant Proteins) ---- 45
B.2.2.5.4 Ajustement osmotique ---- 46
B.2.2.5.4.1 Proline ---- 47
B.2.2.5.4.2 Glycine bétaïne ---- 48
B.2.2.5.4.3 Sucres et dérivés ---- 48
B.2.2.5.5 Transport de l'eau et des solutés ---- 52
B.2.2.5.6 Modification structurale de la cellule ---- 54
B.2.2.5.6.1 Modification de la paroi ---- 55
B.2.2.5.6.2 Dégradation de la paroi ---- 56
B.2.2.5.6.3 Biosynthèse de la paroi ---- 56
B.2.2.5.6.4 Cytosquelette ---- 57
B.2.2.5.7 Régulation de la transcription ---- 57
B.2.2.5.7.1 proteines (b-ZIP) ---- 58
B.2.2.5.7.2 Protéines Myb-like et Myc-like bHLH ---- 58
B.2.2.5.7.3 Protéines AP2(APETALA2)/EREBP ---- 59
B.2.2.5.7.4 Protéines WRKY ---- 60
B.2.2.5.7.5 Autres Protéines ---- 61
B.2.2.5.7.6 Histones et Méthylations ---- 61
B.2.2.5.7.7 Autres mécanismes de régulation ---- 62
B.2.2.5.8 Perception et transduction du signal ---- 63
B.2.2.5.8.1 Perception du signal ---- 63
B.2.2.5.8.2 Calcium et second messagers ---- 64
B.2.2.5.8.3 Cascade des MAPKs ---- 65
B.2.2.5.8.4 Protéines régulées par le Calcium ---- 65

B.3 Le stress hydrique chez les arbres ---- 65

B.3.1 Une problématique spécifique ---- 65
B.3.1.1 Particularités physiologiques ---- 65
B.3.1.2 Ressources moléculaires ---- 66
B.3.2 Le stress hydrique chez le pin maritime ---- 68
B.3.2.1 Les approches physiologiques ---- 68
B.3.2.2 Les Approches génétiques et moléculaires ---- 71
B.3.2.2.1 Cartographie et détection de QTL ---- 71
B.3.2.2.2 Identification des protéines impliquées dans la réponse au stress hydrique ---- 72
B.3.2.2.3 Identification des transcrits impliqués dans la réponse au stress hydrique ---- 73

B.4 Stratégies et techniques pour la détection des gènes impliqués dans l'adaptation du pin maritime à la sécheresse. ---------- 74

C MATERIEL ET METHODES ---------- 78

C.1 Matériel végétal et conditions de culture ---------- 78

C.1.1 Germination ---------- 78

C.1.2 Culture Hydroponique ---------- 79

C.1.3 Contrôle des pathogènes ---------- 79

C.1.4 Eclairage ---------- 80

C.1.4.1 Photosynthèse ---------- 80

C.1.4.2 Action Phytochromique ---------- 81

C.1.4.3 Sources ---------- 81

C.1.5 Application du stress ---------- 82

C.1.5.1 Stress court ---------- 82

C.1.5.2 Stress long ---------- 82

C.1.6 Récolte du matériel végétal ---------- 83

C.2 Mesure du niveau de stress et paramètres physiologiques ---------- 83

C.2.1 Mesure du potentiel de base ---------- 83

C.2.2 Dosage des terpènes ---------- 84

C.3 Analyse du transcriptome et caractérisation des EST ---------- 84

C.3.1 Extraction des ARN ---------- 84

C.3.2 Banques d'ADNc ---------- 85

C.3.2.1 Banque de xylème ---------- 85

C.3.2.2 Banques de racines et aiguilles ---------- 86

C.3.3 Obtention des séquences ---------- 86

C.4 Analyse d'expression ---------- 87

C.4.1 Synthèse des ADNc ---------- 87

C.4.2 Microréseaux à ADNc ---------- 88

C.4.2.1 Réalisation de l'Unigène ---------- 88

C.4.2.2 Dépôt des sondes sur lame de verre ---------- 90

C.4.2.3 Plan d'expérience ---------- 91

C.4.2.4 Synthèse des cibles ---------- 91

C.4.2.5 Préparation, hybridation et lecture des lames ---------- 92

C.4.2.6 Traitement des données ---------- 93

C.4.2.7 Identification des gènes différentiellement exprimés ---------- 94

C.4.2.8 Agrégation des données (Clustering) ---------- 95

C.4.3 PCR quantitative en temps réel ---------- 97

C.4.3.1 Experimentation ---------- 97

C.4.3.2 Choix des amorces ---------- 98

C.4.3.3 Choix des gènes de référence ---- 99
C.4.3.4 Traitement des données de fluorescence ---- 99

C.5 Analyses Protéiques ---- 100
C.5.1 Extraction des protéines et Electrophorèse bidimensionnelle ---- 100
C.5.2 Analyse des gels ---- 101
C.5.3 Analyse quantitative ---- 101
C.5.4 Choix des spots à séquencer ---- 101
C.5.5 Séquençage et identification des protéines ---- 102

C.6 Analyses Bioinformatiques ---- 103
C.6.1 Analyse et Assemblage des séquences ---- 103
C.6.2 Annotation des Consensus et Singletons ---- 105
C.6.3 Niveau d'expression *in silico* des EST (Northern electronique) ---- 106
C.6.4 Annotation des spots protéiques ---- 106

D RESULTATS ET DISCUSSION ---- 109

D.1 Matériel végétal et paramètres physiologiques ---- 109
D.1.1 Culture Hydroponique ---- 109
D.1.2 Potentiel de base ---- 109
D.1.3 Dosage des terpènes ---- 110
D.1.4 Discussion ---- 111

D.2 Analyse et annotation des séquences ---- 115
D.2.1 Séquençage ---- 115
D.2.2 Assemblage des séquences ---- 117
D.2.3 Annotation des séquences ---- 120
D.2.3.1 Annotation des consensus et des singletons ---- 120
D.2.3.2 Assignation des catégories fonctionnelles ---- 124
D.2.4 Northern électronique ---- 125
D.2.5 Unigène ---- 129
D.2.6 Discussion ---- 131

D.3 Analyse du transcriptome ---- 134
D.3.1 Extraction des ARN totaux ---- 134
D.3.2 Puces à ADNc ---- 134
D.3.2.1 Caractéristiques des lames produites et mise en œuvre ---- 135
D.3.2.1.1 Qualité des dépôts d'ADNc ---- 135
D.3.2.1.2 Synthèse des cibles et hybridation ---- 136
D.3.2.1.3 Acquisition des données ---- 136
D.3.2.2 Traitement des données brutes de fluorescence ---- 137
D.3.2.2.1 Préparation et étude descriptive des données ---- 137
D.3.2.2.2 Correction des données ---- 139

D.3.2.2.3 Ajustement des données ---- 140
D.3.2.3 Analyse de l'expression différentielle ---- 142
D.3.2.3.1 Choix des paramètres statistiques ---- 142
D.3.2.3.2 Quels sont les gènes différentiellement exprimés ? ---- 143
D.3.2.3.3 Quelles sont les fonctions des gènes détectés ? ---- 144
D.3.2.3.3.1 Communication / Transduction du signal ---- 145
D.3.2.3.3.2 Croissance cellulaire, division ---- 145
D.3.2.3.3.3 Devenir des protéines ---- 146
D.3.2.3.3.4 Energie ---- 147
D.3.2.3.3.5 Métabolisme ---- 147
D.3.2.3.3.6 Organisation cellulaire ---- 149
D.3.2.3.3.7 Réponse au Stress ---- 150
D.3.2.3.3.8 Synthèse de protéines ---- 152
D.3.2.3.3.9 Trafic intracellulaire ---- 153
D.3.2.3.3.10 Transcription ---- 153
D.3.2.3.3.11 Transport ---- 154
D.3.2.3.3.12 Non-classé ---- 155
D.3.2.3.4 Bilan des principales fonctions régulées ---- 155
D.3.2.3.5 Les gènes détectés sont-ils les même dans les deux expériences ? ---- 157
D.3.2.4 Regroupement et classification des données d'expression ---- 158
D.3.2.4.1 Organisation et profil d'expression des groupes générés ---- 158
D.3.2.4.2 Plasticité moléculaire et gènes d'intérêt ---- 161
D.3.2.4.3 Profils d'expression et fonctions associées ---- 163
D.3.2.4.3.1 Gènes dont l'expression diffère entre écotypes ---- 163
D.3.2.4.3.2 Gènes dont l'expression dépend de la condition ---- 165
D.3.2.4.3.3 Gènes répondant différemment entre écotypes lors du stress ---- 166
D.3.2.4.3.4 Gènes dont le profil d'expression est atypique ---- 169
D.3.3 PCR quantitative en temps réel ---- 170
D.3.3.1 Choix des gènes ---- 170
D.3.3.2 Niveau d'expression des gènes ---- 172
D.3.3.2.1 Gènes vérifiés conjointement en stress court et stress long ---- 172
D.3.3.2.2 Gènes vérifiés en stress long ---- 176
D.3.3.2.3 Gènes vérifiés en stress court ---- 177
D.3.3.2.4 Gènes de la biosynthèse de l'éthylène ---- 178
D.3.4 Discussion ---- 179

D.4 Analyse du protéome ---- 183

D.4.1 Détection des protéines ---- 183
D.4.2 Identification des spots protéiques différentiellement exprimés ---- 183
D.4.3 Annotation des spots protéiques ---- 184
D.4.4 Expression du protéome et du transcriptome ---- 185

D.4.4.1 « Métabolisme des acides aminés » ---- 185
D.4.4.2 « Biosynthèse des métabolites secondaires »---- 186
D.4.4.3 « Métabolisme » ---- 187
D.4.4.4 « Cytosquelette »---- 188
D.4.4.5 « Réponse au stress »---- 188
D.4.4.6 Autres catégories---- 189
D.4.5 De la protéine au gène : cas de l'ACC désaminase ---- 189
D.4.6 Discussion---- 191

E CONCLUSIONS ET PERSPECTIVES---------- 194

E.1 Développement et utilisation d'une approche à haut débit pour l'étude de la réponse au stress hydrique.---------- 194

E.2 Perspectives de cette étude ---------- 197

E.2.1 Exploitation des ressources mises en place ---- 197
E.2.2 Potentiel adaptatif des gènes candidats---- 198
E.2.3 De l'organe à la cellule---- 199
E.2.4 Une autre voie de régulation : les micros ARN ---- 199
E.2.5 Contrôle épigénétique ---- 199

BIBLIOGRAPHIE----------201

ANNEXES----------217

Introduction

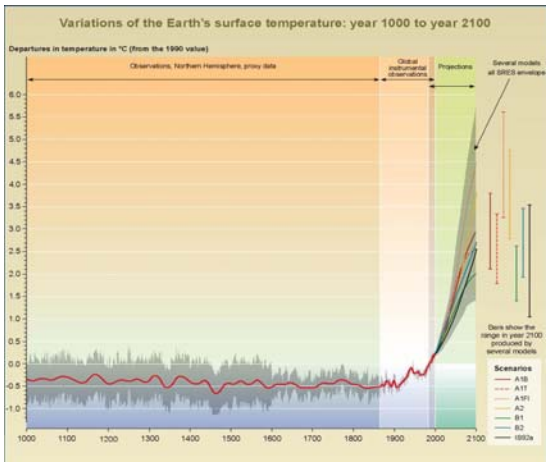
A INTRODUCTION

A.1 Un climat en cours de modification

Depuis quelques années, des études ont été menées pour vérifier qu'une évolution rapide du climat était en cours au niveau mondial. Que ce soit par l'enregistrement et l'analyse de relevés de précipitations, de température, de concentration atmosphérique en différentes molécules ou par des relevés et modélisations météorologiques, l'ensemble des facteurs étudiés confirme une évolution rapide et globale du climat. En 1998, l'organisation mondiale de météorologie (WMO) et le programme environnement des Nations Unies ont constitué une structure intergouvernementale pour étudier les changements climatiques (IPCC : http://www.ipcc.ch/).

Si au cours du dernier millénaire les températures globales de la surface terrestre sont restées stables (Fig. A-1), on assiste à une augmentation nette au cours des 100 dernières années. La même tendance est observée pour la concentration atmosphérique en CO_2, N_2O, CH_4 et SO_4^{2-} produits par l'activité humaine (Fig. A-2). En France, l'augmentation constatée des températures moyennes pour la région Aquitaine est de 1.1°C et celle des maximales de 0.9°C (données Météo France Fig. A-3 ; A-4). Les années 2001 à 2005 ont été marquées par des déficits de précipitations, des sécheresses estivales et des températures au-dessus de la normale (Fig. A-5). L'année 2003 a été particulièrement chaude en Europe, surtout dans le Sud-ouest. Des études récentes (Schar and Jendritzky 2004; Schar *et al.* 2004) essayant de déterminer si un tel phénomène est de l'ordre de l'exceptionnel tendent à montrer que plus de la moitié du risque d'un tel évènement est dû à un changement anthropique du climat.

Les modèles de prévisions climatiques pour le siècle à venir proposés par le GIEC (Groupe International d'Experts sur le Climat : http://www.ipcc.ch/) prévoient une augmentation de température de l'ordre de 2-3 °C à 3-4 °C pour la France avec une augmentation nettement plus marquée pour la saison chaude. En ce qui concerne les précipitations, aucun changement n'est prévu pour la saison hivernale, mais la saison estivale devrait connaître des précipitations nettement plus faibles ou incertaines selon les scénarios (Fig. A-6). Si les modélisations à l'échelle mondiale sont incertaines, c'est d'autant plus le cas à une échelle régionale. Nous pouvons toutefois nous attendre à une augmentation des températures avec des saisons plus contrastées, ainsi qu'une augmentation de la variabilité du climat et de certains phénomènes extrêmes (GIEC 2001).



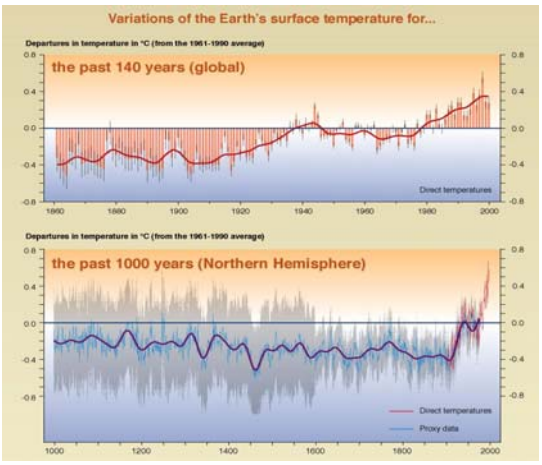


Figure A-1 : Evolution de la température de la surface terrestre, source IPCC.

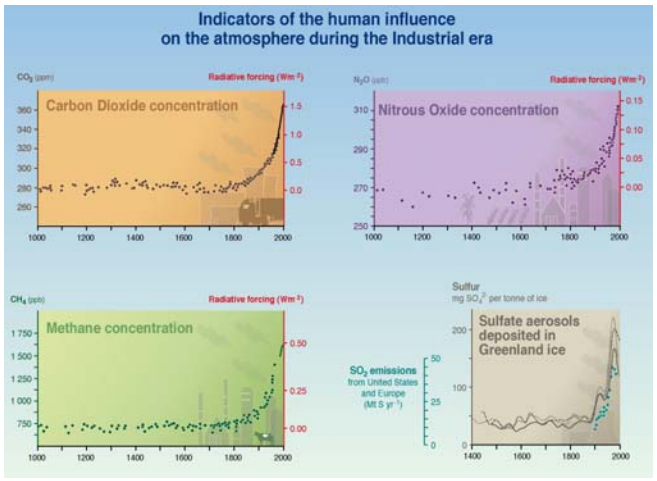


Figure A-2 : Concentration en CO_2, N_2O, CH_4 et SO_4^{2-}, source IPCC.

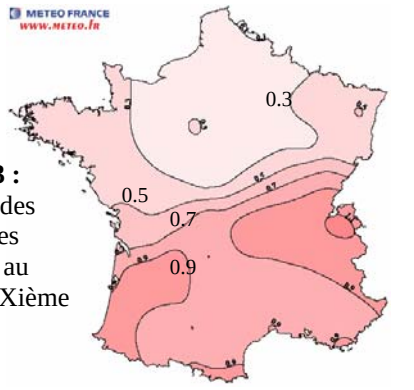


Figure A-3 : Evolution des températures maximales au cours du XXième siècle.

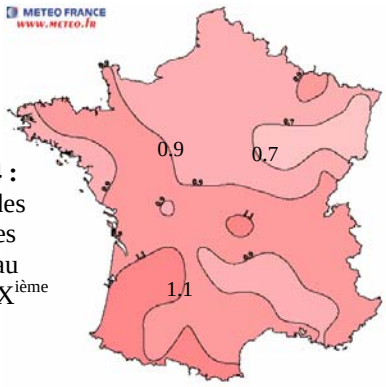


Figure A-4 : Evolution des températures moyennes au cours du XX^{ième} siècle.

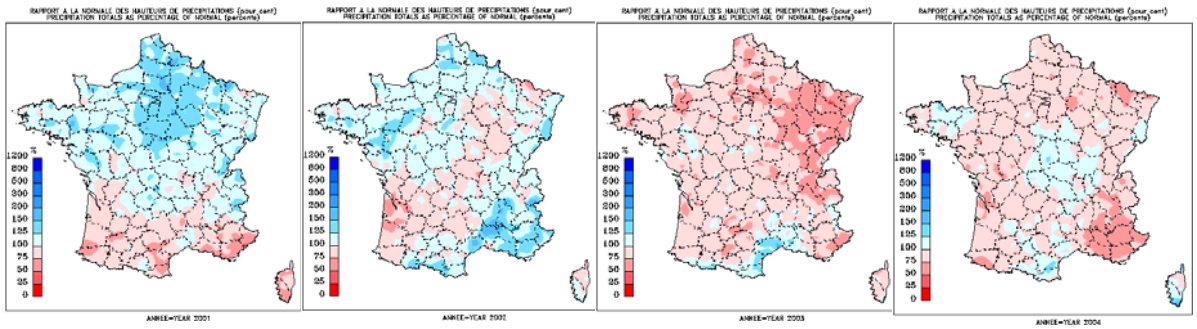
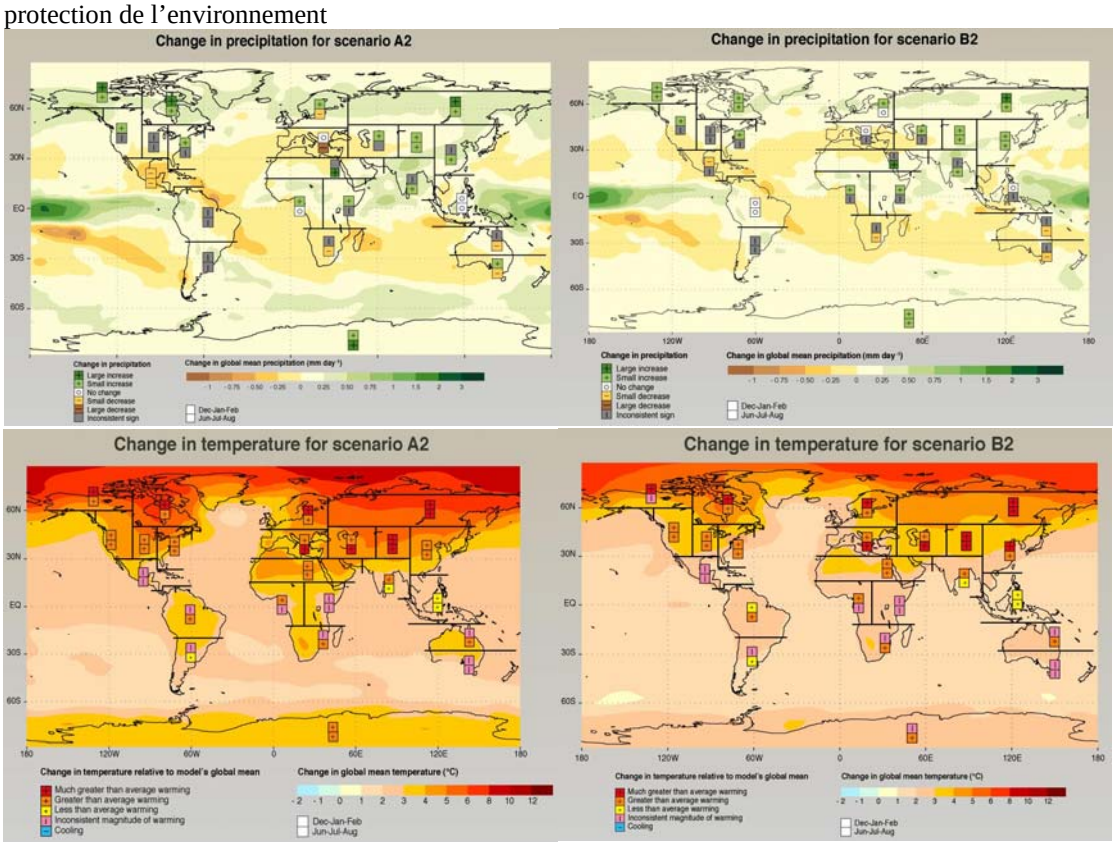


Figure A-5 : Evolution des précipitations en France au cours des années 2001 à 2004 (données Météo-France).

Figure A-6 : Prévisions de l'évolution des précipitations et températures au XXIième siècle :
Scenario A2 = monde hétérogène, démographie converge progressivement, accroissement continu de la population, développement économique plus lent que dans les autres scénarios et régionalisé.
Scenario B2 = monde à fonctionnement régionalisé, démographie à accroissement continu de la population plus faible que A2, développement économique régional hétérogène, évolution technologique plus faible que A2, protection de l'environnement



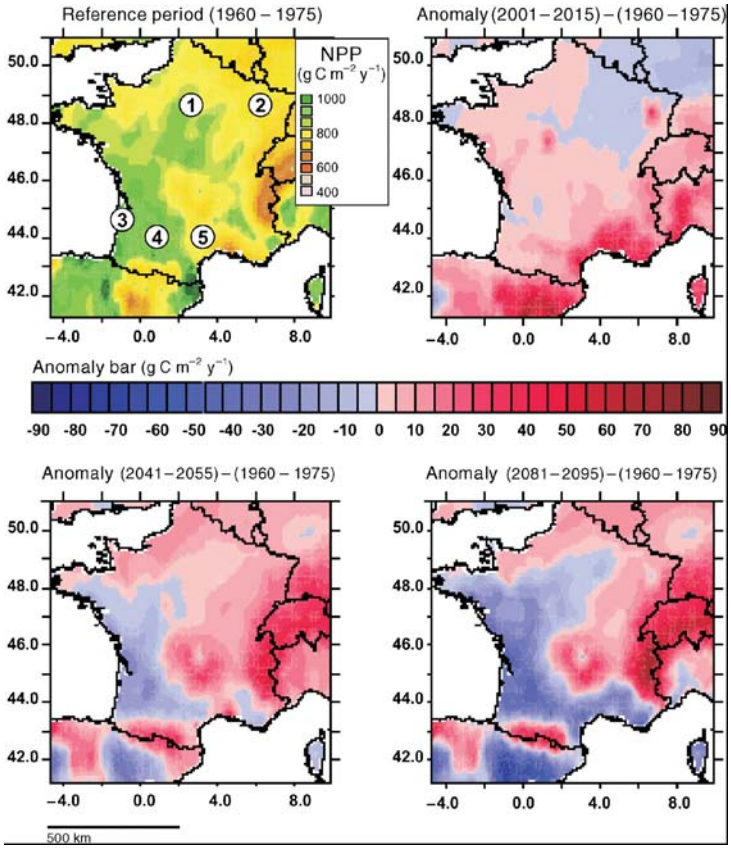


Figure A-7 : Productivité Primaire Annuelle Nette (NPP) des feuilles et aiguilles du système tempéré en France (en haut à gauche) et les prévisions d'évolution de la NPP en 2008, 2048 et 2088 (respectivement haut-droit, bas-gauche et bas-droit). D'après Lousteau *et al.* (2005).

L'évolution des conditions climatiques aura des répercussions directes sur le développement et la croissance des végétaux et en particulier sur la productivité des forêts cultivées (Fig. A-7) (Loustau *et al.* 2005). La question est d'autant plus sensible en sylviculture que les semences plantées aujourd'hui donneront des arbres qui - contrairement aux cultures annuelles - devront faire face aux évolutions climatiques des 50 à 100 prochaines années.

Face à des changements climatiques lents, les écosystèmes forestiers peuvent évoluer grâce à leur plasticité et à leur diversité génétique. Cependant, leur adaptation à des perturbations brutales n'est pas du tout aussi certaine. Si les teneurs supérieures en CO_2 et le réchauffement peuvent stimuler la croissance dans certaines régions, l'instabilité du microclimat et la sécheresse peuvent en annuler ces effets pouvant, à l'extrême, amener à planter des espèces plus méridionales pour maintenir un couvert forestier (Broadmeadow *et al.* 2005).

A.2 Hydrologie et géologie de la région landaise

La région Aquitaine regroupe à elle seule près d'un million d'hectares de pins maritimes sur les 1.357 millions plantés en France. L'aire d'exploitation française du pin maritime est donc centrée sur cette région, plus particulièrement le département des Landes, qui présente des spécificités pédologiques particulières.
La capacité de croissance des végétaux dépend directement des caractéristiques du sol et du sous-sol dans lequel les racines se développent. Le sol landais actuel se singularise par une vaste zone sableuse triangulaire de près d'un million d'hectares.

Cette vaste plaine s'est formée durant le Jurassique à partir d'un vaste golfe qui s'est comblé progressivement par des dépôts d'épaisses séries sédimentaires sur plusieurs milliers de mètres. A la suite de mouvements tectoniques au tertiaire, certaines parties du triangle landais se sont enfoncées ou soulevées et le comblement s'est poursuivi. Au quaternaire, le comblement s'est poursuivi par d'abondants apports fluviatiles où se mêlent de façon variable des sables, argiles, graviers et galets. La présence des niveaux argileux dans la partie supérieure de la formation est particulièrement importante. Sur l'axe landais, ils constituent souvent le soubassement direct du sable des Landes et ont localement un grand rôle sur le régime de l'eau dans le sol. C'est pendant le quaternaire que les Landes vont devenir la plus grande plaine sableuse de France. Avec les changements climatiques et la fin des grandes glaciations, on observe une élévation du niveau de la mer qui rapproche progressivement le

rivage de sa position actuelle et remanie constamment les sédiments de la plateforme continentale. Les réseaux hydrographiques actuels se mettent en place et le vent alimente en sable l'intérieur des terres (Montané 1994).

Il se forme alors un constituant essentiel du sol landais : l'alios. Le lessivage des sables forme un horizon supérieur gris clair (podzol) de quelques décimètres d'épaisseur. Au-dessous se forme l'alios, qui est un grès tendre où les grains de sable sont cimentés par les acides humiques et l'oxyde de fer donnant une teinte brune à rougeâtre. Sa constitution interdit tout passage des racines vers l'eau de la nappe. L'épaisseur et la compacité de l'alios vont déterminer la vitesse de pénétration des eaux pluviales vers la nappe phréatique qui, associée aux très faibles pentes, va créer un écotype de lande sèche ou de lande humide.

On distingue ainsi différents types de landes correspondant à des écosystèmes différents fonctions de la profondeur de la nappe d'eau et de la fertilité du sol :

- En lande humide, la Molinie et l'Ajonc nain côtoient la Bruyère à quatre angles et la Bruyère ciliée. L'alios ainsi que les nappes sont très proches de la surface à 30-50cm.
- La lande mésophile fait place à la Fougère aigle, à l'Ajonc, au Genêt et à la Brande.
- Sur le sol sec de la lande sèche se trouvent la Callune, la Bruyère cendrée, l'Hélianthème. Les nappes sont plus profondes et l'alios plus éloigné de la surface.

A.3 Des marais à la sécheresse, un équilibre fragile

Jusqu'au milieu du XIX^{e} siècle, l'eau apportée par les pluies hivernales ne pouvant pas traverser rapidement l'alios, et le système hydrique de cette vaste étendue sans pente ne pouvant évacuer l'eau, les landes de Gascogne se transforment en zone marécageuse. Les parties racinaires sont donc noyées et seules les espèces adaptées à l'ennoyage peuvent subsister. L'eau, croupie par décomposition des végétaux, s'évacue progressivement à l'approche de la saison estivale puis l'eau disponible aux racines dans l'horizon supérieur est rapidement épuisée. De telles conditions ne permettant pas la culture ou la sylviculture, l'élevage et l'apiculture sont pratiquement les seules activités. Il existe cependant quelques bois de pin maritime fournissant le bois, le charbon et la gemme. A cela s'ajoute la progression vers l'intérieur des terres des dunes littorales non stabilisées.

A la fin du XVIII^{e} siècle, Brémontier réussit à fixer les dunes par des plantations de végétation dunaire protégée du vent par des branchages. Puis Chambrelent mit en place un plan d'assainissement des Landes de Gascogne basé sur la réalisation d'un réseau de canaux collecteurs à l'origine de la loi du 19 juin 1857 obligeant les communes à assainir et à

ensemencer leurs landes. Les terres drainées ont dès lors pu être plantées et cultivées. Progressivement les forêts de pin maritime ont remplacé les anciens marais.

Aujourd'hui nous assistons à un déficit hydrique marqué entraînant des dépérissements massifs des peuplements de Pin maritime. D'après les informations relevées par les professionnels de la sylviculture et l'équipe d'écophysiologie de l'INRA de Pierroton, près de 30000 hectares sont touchés dans la région de Solferino et Escource au cœur des Landes (communication personnelle Alexandre Bosc, UMR EPHYSE). Considérant l'étendue des pertes, la cause hydrique peut être mise au premier plan, suivie par d'autres stress comme des attaques de pathogènes profitant d'arbres affaiblis. Il s'agit d'une zone de landes humides avec un alios situé à 40-50cm qui stoppe la pénétration des racines. L'horizon accessible par le système racinaire est donc très limité mais la nappe des landes humides est proche de la surface. En temps normal, l'eau liquide atteint les racines en dehors de la saison estivale avec l'accumulation des précipitations. Un important système hydraulique avec de gros fossés permet de drainer l'eau en hiver pour éviter l'ennoyage. Cependant, les prélèvements conséquents dans les nappes pour l'irrigation des terres agricoles, combinés à 3 années de déficit de précipitations et de sécheresse, ont provoqué une dépression du niveau de la nappe qui, à son niveau supérieur, ne doit plus dépasser l'alios et donc atteindre les racines.

A.4 Des incertitudes sur la durabilité du massif forestier aquitain

La forêt aquitaine a un double rôle à jouer : économique et environnemental. Avec 2.5 milliards d'euros de chiffre d'affaires, elle génère 21000 emplois directs auxquels s'ajoutent 7000 emplois en menuiserie, meubles et commerce (www.mediaforest.net). De sa bonne gestion dépend la préservation de la santé de l'économie et de l'écosystème d'une région entière. Les choix politiques et économiques récents soulignent la nécessité de préserver ce moteur économique essentiel en soutenant la création du pôle de compétitivité « Industries et Pin maritime du Futur » qui allie la recherche publique et privée.

En 2003, le pin maritime représentait 26% (2.9Mm3) du bois de trituration et 36% (5.4Mm3) du bois d'œuvre résineux au plan national.

Les utilisations du pin maritime sont multiples. Son bois peut être utilisé sous sa forme native en construction (bois de charpente, bardage, lambris, planches, poteaux, etc.), en ameublement, en emballage (palettes, caisses et coffrage) ou transformé en pâte à papier et panneaux de particules. D'autres usages sont plus spécifiques comme l'écorce pour le jardin ou la résine et autres extraits (picnogenol, parfums).

Face à la demande croissante en bois, cette forêt cultivée, dévolue à une fonction de production, aura un rôle de plus en plus important à jouer. L'amélioration des pratiques sylvicoles associées à un programme d'amélioration génétique initié dans les années 1960 à l'INRA (Baradat and Pastuszka 1992) visant à augmenter la production et la qualité du bois, a permis d'augmenter la productivité de façon significative. En effet, le gain attendu sur le volume de bois est de l'ordre de 30% pour les variétés actuellement plantées en Aquitaine (Alazard and Raffin 2002).

Cependant, dans un milieu changeant, on peut se poser la question de savoir si les gains espérés seront réalisés. Une estimation indique actuellement que si l'on considère les précipitations, l'évapotranspiration et la production de bois à l'hectare, 850 litres d'eau sont nécessaires pour produire 1kg de bois. Les variétés à croissance rapide auront probablement des exigences accrues en eau dans un avenir qui s'annonce peu favorable. Il est donc essentiel de prendre en compte d'autres critères que la croissance et la qualité du bois pour permettre un développement durable de ce massif forestier. Ainsi, il apparaît essentiel de prendre en compte l'adaptabilité des variétés futures à la modification de l'écosystème. Cela peut passer par l'augmentation de la diversité génétique des peuplements voire le mélange d'espèces et/ou l'amélioration génétique de l'écotype landais (sur lequel est essentiellement basé le programme d'amélioration) pour la résistance à la sécheresse. Des études récentes ont en effet montré qu'il existait au sein des populations aquitaines une forte variabilité pour l'efficience d'utilisation de l'eau (Brendel *et al.* 2002), un des caractères liés à l'adaptation à un déficit hydrique.

A.5 Objectifs

Mon travail de thèse s'inscrit dans le cadre de la poursuite des travaux initiés par l'INRA sur le déterminisme génétique (Costa 1999; Brendel *et al.* 2002), moléculaire (Costa 1999) et physiologique (Nguyen and Lamant 1988; 1989b; 1989a; Costa 1999; Nguyen-Queyrens *et al.* 2002) de la réponse du pin maritime à un déficit d'alimentation en eau. Ces études ont pour but d'améliorer nos connaissances sur le sujet et de permettre d'identifier des critères de sélection pour le choix de génotypes plus performants et mieux adaptés à leur environnement. Si les études physiologiques ont permis de mettre en évidence une variabilité de comportement entre écotypes de pin maritime, les QTL d'efficience d'utilisation de l'eau et d'autres caractères physiologiques soulignent l'existence d'un contrôle génétique de la

réponse au déficit hydrique. Il reste néanmoins à identifier les gènes impliqués dans l'adaptation à la sécheresse chez cette espèce.

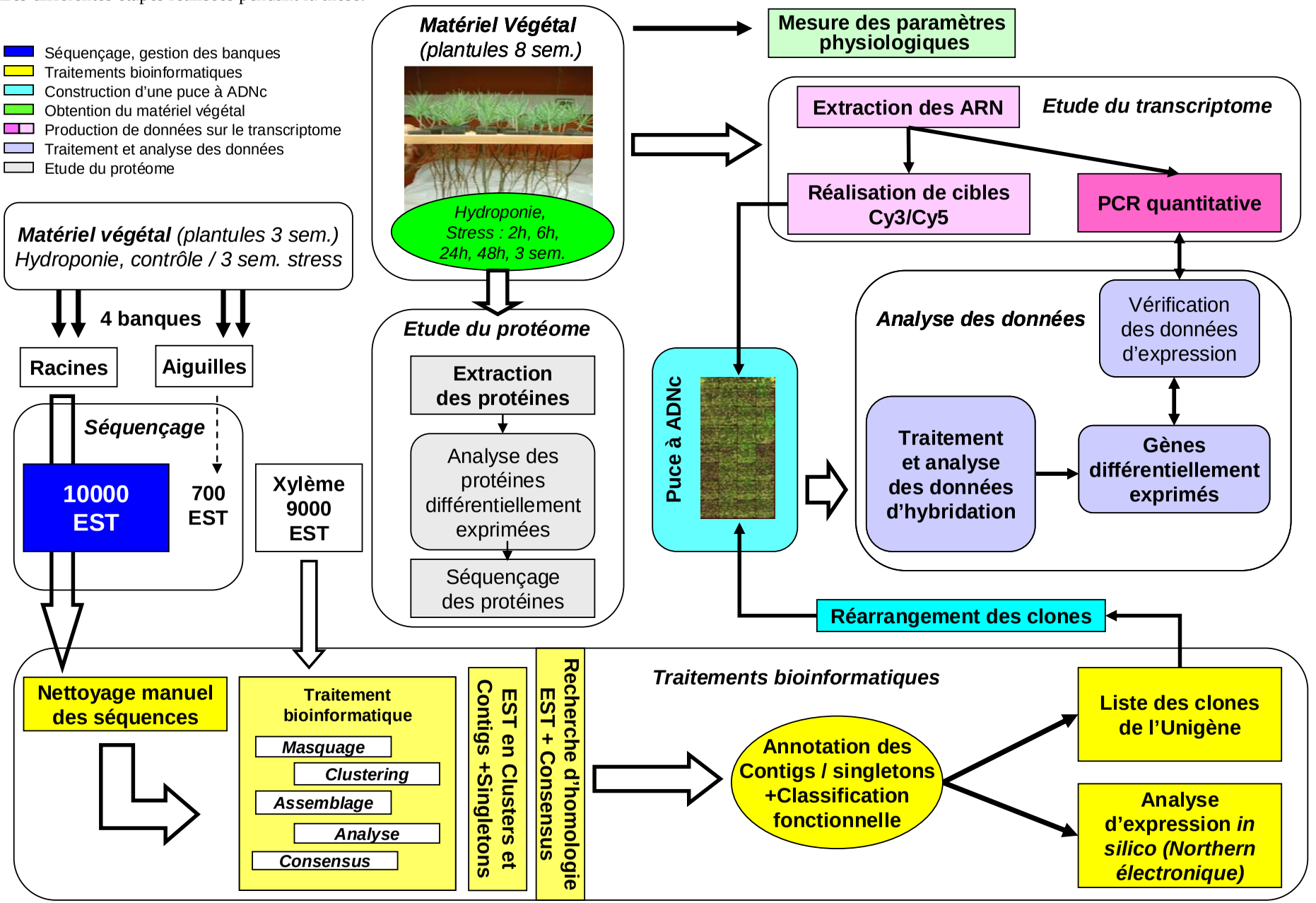
Mon travail se focalise sur ce dernier aspect. Il concerne l'identification de gènes régulés en condition de stress hydrique édaphique et qui présentent une variation d'expression chez des écotypes contrastés quant à leur réponse adaptative à la contrainte hydrique. Les recherches se concentreront sur les gènes exprimés dans les racines ; organe clé impliqué dans le captage de l'eau et qui a jusqu'ici fait l'objet de peu d'études chez les plantes en général.

Parallèlement à cet objectif scientifique, mon travail comprenait également des défis technologiques. Il me fallait en effet développer une approche « génomique » en mettant en place de nouveaux outils et méthodes d'étude de l'expression du génome : le séquençage massif d'étiquettes de séquences exprimées (EST) pour obtenir une base de ressources génomiques, le développement d'outils bioinformatiques pour les analyser et les rendre accessibles à la communauté scientifique, la création et l'utilisation des premiers réseaux d'ADNc (microarrays) chez le pin maritime.

La figure A-8 synthétise la démarche adoptée au cours de ce travail. Un code couleur permet de visualiser les différentes étapes suivies. La stratégie s'articule autour de 7 grands axes : 1) Production de séquences, 2) Traitements bioinformatiques, 3) Construction d'une puce à ADNc, 4) Obtention du matériel végétal, 5) Production de données sur le transcriptome, 6) Analyse de données d'expression au niveau transcriptomique 7) Etude du protéome. Ces sept étapes sont ici brièvement présentées :

1. Le projet a débuté par la production de séquences issues d'ARN extraits de racines (indiqué dans la partie gauche de la figure A-8). Ce programme de séquençage d'EST portait sur deux banques de racines obtenues sur de jeunes plants (3 semaines) de pins maritimes élevés en hydroponie : l'une correspondant à des plants témoins et l'autre à des plants soumis à 3 semaines de stress osmotique (PEG, -0.45MPa). Je disposais des banques et j'ai pris en charge les aspects séquençages avec l'aide d'un technicien du laboratoire.
2. Ces séquences ont été injectées dans une chaîne d'analyse bioinformatique pour l'assemblage et la recherche d'homologie contre les banques publiques. J'ai pris en charge la partie des traitements bioinformatiques concernant l'annotation, la

Figure A-8 :
Les différentes étapes réalisées pendant la thèse.



classification fonctionnelle, l'analyse d'expression *in silico* et l'obtention d'un Unigène pour la construction d'une puce à ADNc.

3. Après réarrangement et duplication des banques, que j'ai transmises à notre partenaire autrichien PICME (http://www.picme.at), une puce à ADNc a été construite.
4. Dans le même temps j'ai mis en place le dispositif pour obtenir le matériel végétal utilisé pour les études d'expression. Il s'agissait de plants âgés de 8 semaines, cultivés en hydroponie et soumis à un stress osmotique par ajout de PEG.
5. Ce matériel a fait l'objet de mesures physiologiques et d'une étude du transcriptome. Après extraction des ARN, j'ai synthétisé les cibles et hybridé les lames produites.
6. Les données générées ont alors été traitées et analysées afin d'identifier des gènes différentiellement exprimés lors du stress hydriques chez deux écotypes, français et marocain.
7. Le protéome du même matériel végétal a également été analysé.

A.6 Plan de la thèse

Ma thèse est structurée en quatre parties principales :

La **partie 1** est une synthèse bibliographique. J'ai choisi de focaliser cette étude sur les mécanismes de réponses moléculaires et physiologiques des plantes au stress hydrique. Je traiterai aussi, brièvement, des stratégies possibles pour la découverte des gènes majeurs pour l'adaptation à la sécheresse chez le pin maritime. Cette partie bibliographique est divisée en quatre chapitres.

- Le premier chapitre présentera les caractéristiques botaniques de l'espèce étudiée, les écotypes et son aire de répartition géographique.
- Le second chapitre traitera des réponses des plantes aux stress. Après avoir décrit les relations de l'eau et de la plante, nous aborderons les mécanismes de régulation physiologique que la plante met en œuvre au cours du stress. Puis, nous passerons en revue les voies métaboliques et les gènes connus dont l'expression est régulée au cours du stress.
- Le troisième chapitre sera consacré au stress hydrique chez les arbres. Nous découvrirons les travaux de recherches menés jusqu'à présent sur cette thématique chez le pin maritime. Les approches physiologiques seront présentées, puis nous passerons au déterminisme génétique avec les études de détection de QTL et de

génétique quantitative, puis aux mécanismes moléculaires avec l'étude du protéome et du transcriptome.

- La dernière partie abordera les aspects techniques et les stratégies pouvant être mises en œuvre pour la détection des gènes impliqués dans l'adaptation du pin maritime à la sécheresse.

La **partie 2** détaille les principales méthodologies utilisées. Je présenterai successivement le matériel végétal et le dispositif expérimental, les techniques de biologie moléculaire utilisées pour obtenir les données d'expression transcriptomique et protéomique, puis les méthodes bioinformatiques et statistiques mises en œuvre pour traiter les données.

La **partie 3** se structure également en 4 chapitres. Chaque chapitre présente les résultats produits et leur interprétation.

- Le chapitre I concerne l'obtention du matériel végétal et les paramètres physiologiques mesurés au sein du dispositif expérimental que j'ai mis en œuvre.
- Le chapitre II concerne l'analyse et le traitement des séquences. Nous nous intéresserons tout d'abord au prétraitement et à l'assemblage des séquences. Ces étapes ont été réalisées en faisant appel à un nouveau système d'interface web associé à un pipeline bioinformatique auquel j'ai participé pendant 2 ans en tant que beta testeur. Ce nouvel outil fait l'objet d'une note dont je suis co-auteur (article soumis, Annexe Px-1). Les séquences produites dans ce projet pin maritime ont également été le support pour la recherche *in silico* de SNP (Single Nucleotide Polymorphism) par une chaîne d'analyse automatique présentée dans une publication de Plant Molecular Biology dont je suis également co-auteur, annexe Px-2 (Le Dantec *et al.* 2004). Nous présenterons ensuite le travail d'annotation fonctionnelle réalisé sur les séquences puis leur utilisation pour l'approche par « Northern électronique » dont l'objectif était de détecter les variations d'expression des transcrits par comptage du nombre de séquences présentes dans les bases de données d'EST. Cet aspect a fait l'objet du stage de DESS de Christophe Boury que j'ai co-encadré. La fin de ce chapitre donnera les informations sur l'Unigène que j'ai généré et sera suivie d'une discussion des résultats observés.
- Le chapitre III est dédié à l'analyse du transcriptome. Il traitera dans un premier temps de l'ensemble des données d'expression que j'ai obtenues en utilisant la technologie

des puces à ADNc. Nous passerons en revue les résultats des analyses statistiques et des méthodes de classification. Dans un second temps, les niveaux d'expression obtenus par PCR quantitative sur un ensemble de gènes sélectionnés seront présentés. Les résultats de ces deux approches complémentaires seront interprétés et discutés. Une étudiante de DES que j'ai co-encadrée (Manon Moreau) a également participé à l'acquisition de ces résultats.

- Dans le chapitre IV nous synthétiserons les données obtenues par l'analyse du protéome (Stage de Master 2 de Manon Moreau) et les comparerons aux données du transcriptome.

Enfin la **partie 4** est une conclusion générale de l'ensemble des résultats obtenus dans le cadre de cette thèse et propose des perspectives au vu des informations recueillies.

Synthèse Bibliographique

B SYNTHESE BIBLIOGRAPHIQUE

Comme tout organisme vivant, la plante est essentiellement constituée d'eau. La plupart des végétaux sont soumis à des variations de la disponibilité en eau du milieu. Ils ont donc dû développer au cours de l'évolution des mécanismes de réponse adaptative pour faire face à des périodes de déficit en eau. De nombreux travaux ont permis de décrire les réponses des plantes en état de stress. Les effets sont multiples et vont de l'altération de l'expression des gènes et du métabolisme cellulaire, à la modification de la croissance avec des conséquences sur le rendement des cultures. Les réponses peuvent être directes, initiées par le stress et facilement identifiables, ou indirectes pour palier aux conséquences du stress. Au niveau moléculaire, cela implique la régulation d'un nombre conséquent de voies métaboliques. Après la présentation de l'espèce, nous ferons dans cette synthèse l'état des connaissances sur les mécanismes majeurs impliqués lors d'un déficit en eau. Nous décrirons également les voies métaboliques permettant la synthèse des métabolites impliqués dans la réponse au stress hydrique, la régulation de leur production étant déterminée par les enzymes des voies de biosynthèse et des facteurs de transcription.

B.1 Le Pin maritime

B.1.1 Botanique et Taxonomie

Les conifères appelés également résineux sont les représentants actuels des gymnospermes apparus à la fin de l'ère Primaire. Ces espèces ont colonisé tout l'hémisphère nord au cours de l'ère secondaire. Les traces fossiles du genre *Pinus* sont nombreuses et remontent au début du crétacé (Wang *et al.* 2000) soit il y a environ 140 millions d'années. Les changements climatiques avec les périodes glaciaires et interglaciaires du quaternaire ont eu pour effet la colonisation de territoires ou la création de zones refuges à l'origine de spéciations. La figure B-1a indique la chronologie de la différentiation des gymnospermes.

Les conifères forment l'ordre des *Coniférales* ou *Pinales*, divisé classiquement en six familles (www.species2000.org) : les *Araucariaceae* (araucarias, Agathis), les *Céphalotaxaceae* (Cephalotaxus), les *Cupressaceae* (cyprès, thuyas, genévrier), les *Pinaceae* (mélèzes, sapin, pins, cèdres épicéas), les *Podocarpaceae* (Podcarpus) et les *Taxodiaceae* (séquoias).

La famille des *Pinaceae* est divisée en sept genres : *Abies*, *Cedrus*, *Larix*, *Picea*, *Pinus*, *Pseudotsuga*, et *Tsuga*. Trois autres genres peuvent être distingués : *Cathay*, *Keteleeria* et *Pseudolarix* (http://tolweb.org).

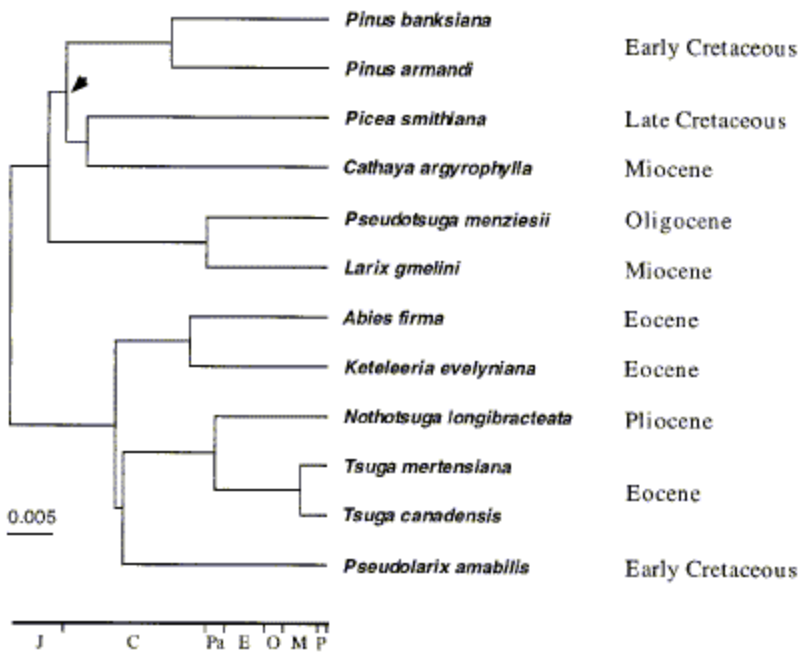
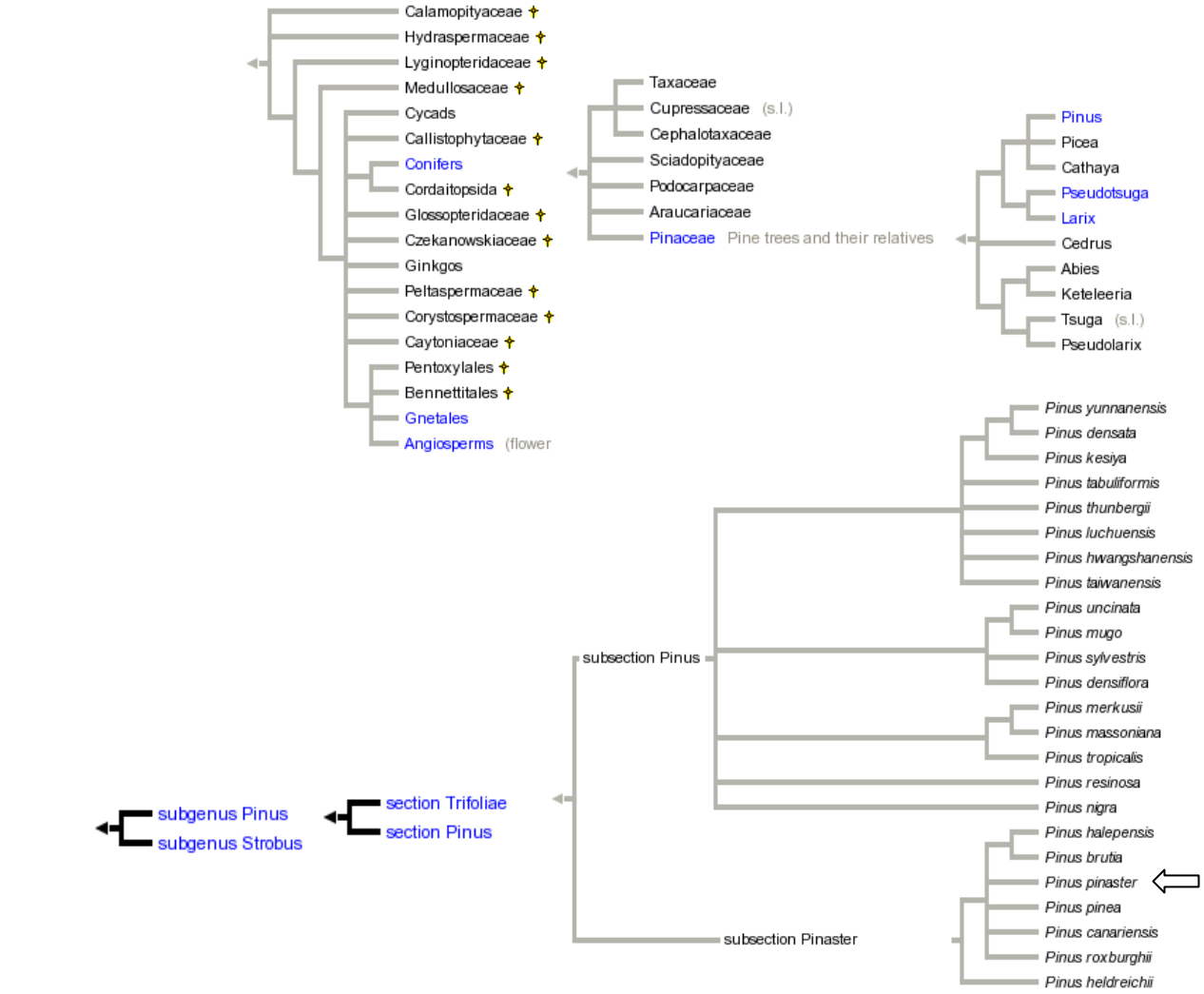


Figure B-1a : Arbre phylogénétique et origine de la différentiation des gymnospermes

Figure B-1b : Arbre phylogénétique et classification botanique des espèces du genre Pinus



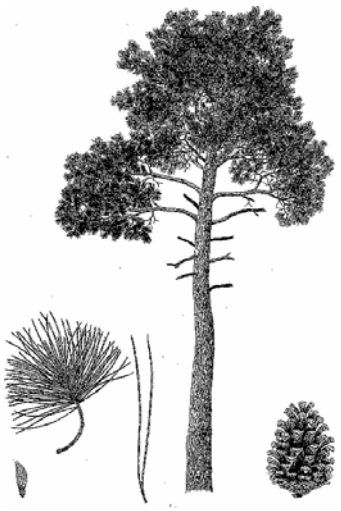


Figure B-2 : Pin maritime
(Pinus pinaster Ait.)

Le genre *Pinus* est séparé en deux sous-genres : *Pinus* (2-3 aiguilles) et *Strobus* (5 aiguilles) qui regroupent un peu plus d'une soixantaine d'espèces. Ces sous genres sont divisés en sections, subdivisées elles-mêmes en sous-sections (Fig. B-1b).

Division	*Coniferophyta*
Classe	*Pinopsida / Coniferopsida*
Ordre	*Pinales / Coniférales*
Famille	*Pinaceae*
Genre	***Pinus***
Sous-Genre	*Pinus*
Section	*Pinaster / Pinus*
Sous-Section	*Australes / Pinaster*
Espèce	***Pinaster***

D'après http://www.itis.usda.gov/index.html

Les pins ont pour caractéristique commune la présence d'un appareil foliaire sous forme d'aiguilles pointues, longues ou courtes et réunies en groupes de 2, 3, 4 ou 5. Ce sont des plantes monoïques, dont l'inflorescence femelle ou cône mûrit en deux (rarement trois) ans après fécondation.

Pinus pinaster (Fig. B-2) est couramment appelé Pin maritime, Pin des Landes, Pin de Bordeaux, Pinastre, Pin de Corte pour la variété corse ou encore « Cluster Pine » en anglais. La floraison a lieu en avril/mai avec une production massive de pollen transporté par le vent. Les graines sont souvent ailées, ce qui facilite leur dissémination par le vent et l'extension de leur aire de distribution.

C'est une espèce allogame, sempervirente, à croissance initiale rapide, qui arrive à maturité à 40-50 ans en atteignant en général une taille de 20 à 35 m. Son espérance de vie peut cependant atteindre les 200 ans. Le tronc, flexueux à la base, est recouvert d'une écorce abondante (Taux d'écorce en volume = 25%) brun violet, rouge sombre en profondeur très épaisse et fortement crevassée. Le bois a une densité moyenne (Masse volumique brute : 880kg/m^3, masse volumique sans écorce = 700kg/m^3), rougeâtre foncé, et à aubier nettement plus clair ; les canaux résinifères sont nombreux, gros, et visibles à l'œil nu ; la sécrétion de résine est abondante. Son houppier est conique, étalé et peu compact. Les aiguilles sont rigides et présentent des stomates sur les deux faces. Elles sont regroupées en fascicules de deux aiguilles pouvant dépasser 15-20cm de long mais dans le très jeune âge, les sujets vigoureux en ont souvent trois, au moins sur la pousse principale. On distingue les euphylles

ou feuilles aciculaires qui apparaissent aussitôt après les cotylédons sur les tiges et rameaux du semis de l'année. Puis les aiguilles ou pseudophylles se développent à la place des euphylles qui sont très vite réduites à l'état d'écailles brunes (Maugé 1987).

B.1.2 L'appareil racinaire

Il a pour fonction l'ancrage de l'arbre et son alimentation en eau. C'est l'organe directement au contact de la ressource hydrique et de son développement va dépendre les capacités de réponses aux périodes de sécheresse.

Le système racinaire du Pin maritime est souvent qualifié de pivotant ou plongeant. En réalité, le système est double ; constitué d'un système plongeant très développé et d'un système traçant (Maugé 1987; Danjon *et al.* 1999).

Le système plongeant est le premier à se développer sur les jeunes plants à partir d'un axe unique qui se ramifie ensuite, donnant d'autres racines à développement vertical (Fig. B-3). Lorsque l'horizon atteint est riche en eau, le développement du système plongeant s'arrête au profit d'un pinceau racinaire constitué de nombreuses racines fines dont la taille peut atteindre 1m (dunes) ou sera réduite dans le cas d'une nappe. Cependant, cette organisation du système plongeant est modifiée selon la nature du sol et les possibilités de pénétration des racines en profondeur (Fig. B-5). La plasticité du système racinaire permet de modifier considérablement l'axe de développement préférentiellement vertical.

Si un obstacle infranchissable - comme l'alios d'une lande humide dont nous reparlerons dans les chapitres suivants- est rencontré, les racines verticales s'arrêtent et forment un moignon et un réseau racinaire enchevêtré circulaire d'environ 2 mètres de diamètre (Danjon *et al.* 2005)(Fig. B-6) offrant une résistance d'arrachement très limitée en cas de forts vents d'où les fréquents chablis rencontrés lors des tempêtes.

Si l'obstacle peut être contourné – comme l'alios très fissuré des landes sèches -, les racines verticales modifient leur progression en le longeant. Dès que le point de franchissement est atteint, les racines adoptent leur mode de développement initial (Fig. B-4). L'ancrage est alors très bon et les dégâts observés en cas de tempête sont plutôt des volis.

Le système traçant se développe juste sous la base du collet en émettant 5 à 7 racines d'ordre 1 à développement radial et parfaitement horizontal desquelles partent des racines d'ordre 2, 3, etc. conservant toujours la direction centrifuge des racines d'ordre 1. Elles se développent dans l'horizon, superficiel bien aéré et constitué de matières organiques en décomposition. Tout obstacle stoppe la croissance de la racine radiale qui est donc impossible

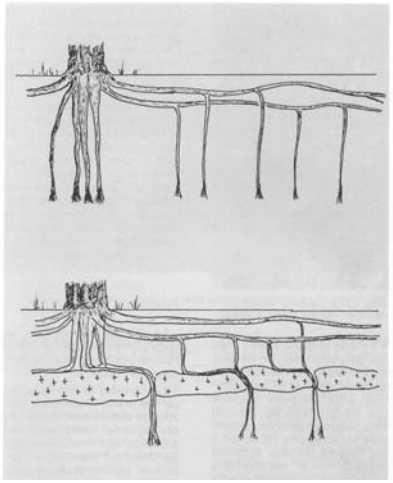


Figure B-3 : Développement du système racinaire plongeant et traçant dans un sol idéal chez le pin maritime

Figure B-4 : Développement du système racinaire plongeant et traçant en présence d'alios fissuré du type lande sèche chez le pin maritime (Maugé (1987))

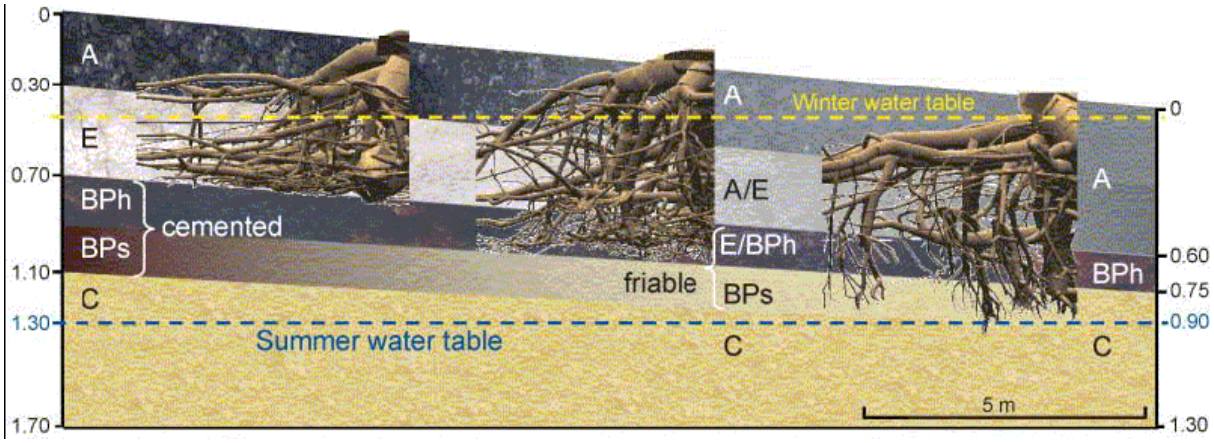


Figure B-5 : Nature de l'enracinement de pins maritimes (reconstitution 3D) en fonction de la nature du sol. (Danjon (2005)). A gauche : sol bien drainé, l'alios (BP) stoppe la progression des racines. A droite, sol mal drainé, l'alios friable permet un passage des racines.

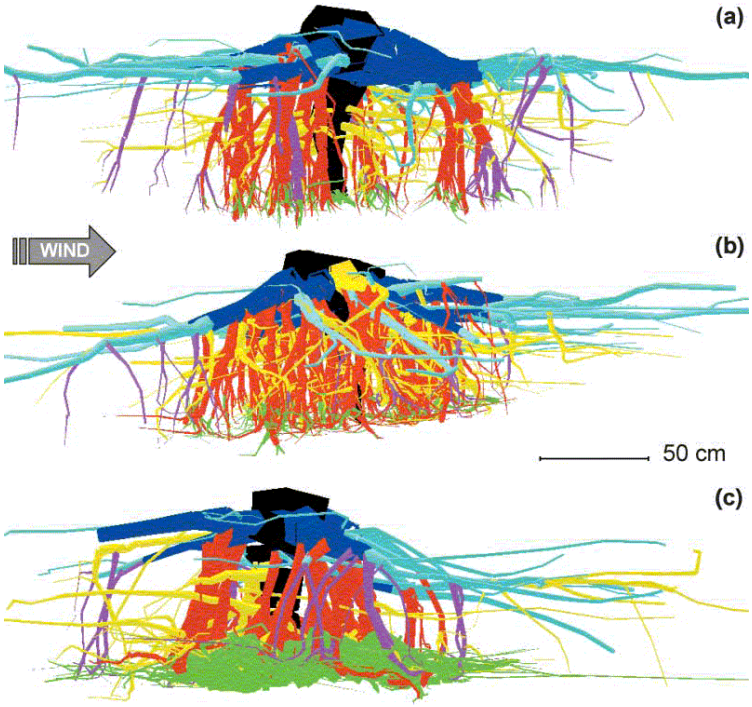


Figure B-6 : Reconstitution 3D du système racinaire « en plateau » d'un pin maritime des landes. L'enracinement diffère (a,b,c) en fonction du vent et de la nature du sol. (Danjon(2005))

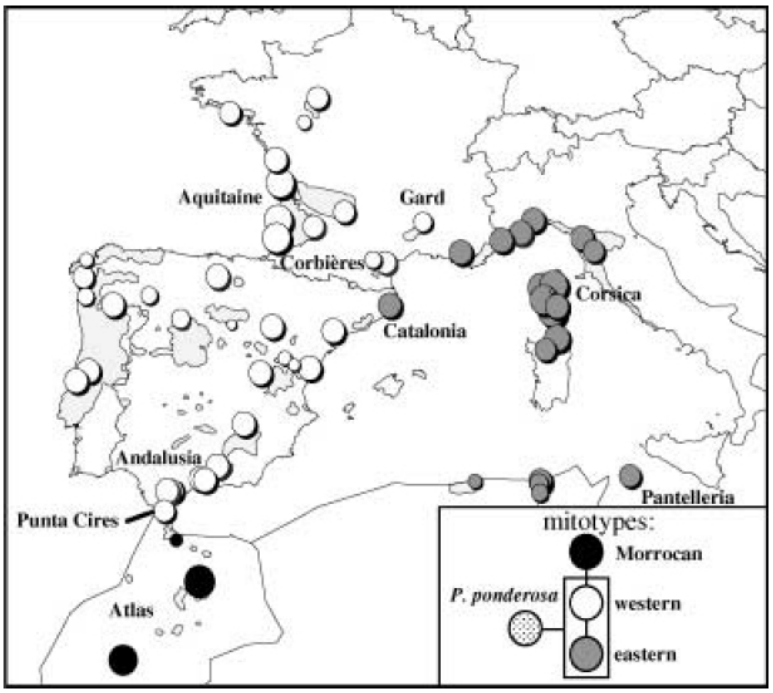


Figure B-7 : Distribution des mitotypes de pin maritime et lien de phylogénie basé sur les variations de nad1 intron 2 d'après Burban et *al.* (2003).

dans un sol un peu compact ou occupé par le chevelu racinaire de graminées comme la molinie. De ce réseau de racines principales partent des racines fines d'une taille de 10 à 15 cm pour un diamètre proche du millimètre formant une sorte de manchon. Ces racines, souvent mycorhizées, présentent des excroissances couvertes de poils absorbants. Elles ont une durée de vie brève de 1 à 3 ans puis se décomposent. Leur fonction principale est de capter activement les éléments nutritifs ce qui nécessite de l'oxygène. Le sol doit pour cela être humide sans toutefois être gorgé d'eau ce qui provoquerait une anoxie. Le système traçant a donc un rôle d'ancrage et d'alimentation de l'arbre en éléments nutritifs. C'est sur cet organe qu'ont porté nos investigations.

B.1.3 Répartition géographique

Dans une publication récente, les auteurs Burban et Petit (2003) décrivent des données de phylogéographie obtenues sur le Pin maritime en utilisant des marqueurs chloroplastiques. En se servant des marqueurs à hérédité uniparentale, ils ont ainsi pu préciser la structuration génétique de l'espèce décrite dans la littérature et basée sur d'autres marqueurs (terpènes, protéines totales, isoenzymes, microsatellites nucléaires et AFLP). Le pin maritime est séparé en 3 groupes : un groupe atlantique comprenant les populations de l'ouest de la France et de la péninsule ibérique, un groupe méditerranéen constitué de toutes les populations de l'est de l'Europe, de la partie nord africaine et des populations de Catalognes, et un groupe marocain (Fig. B-7). Nous nous intéresserons plus particulièrement à l'écotype « Tamjout » originaire du Maroc. Cette station est la plus orientale des provenances marocaines. Localisée au sud de Taza à 1650 m d'altitude en exposition nord ouest, le climat se situe dans l'étage méditerranéen subhumide. Les précipitations atteignent 650 mm par an répartis en 250 mm l'hiver, 210 mm au printemps, 40 mm en été et 150 mm en automne. Le sol est calcaire avec une forte proportion de carbonate et de calcaire actif (Guyon 1980).

En France, l'aire naturelle du pin maritime se situe à faible distance de la mer dans les Landes de Gascogne, la Corse, les Maures, l'Estérel et les Corbières. Par comparaison avec la station marocaine, les précipitations annuelles dans le département des Landes atteignent environ 980 mm avec un minimum de 200 mm pour la saison estivale. Le pin maritime a aussi été introduit dans d'autres régions, notamment en Bretagne, Pays de Loire, Centre, Charente, Aquitaine et Cévennes. Dans le reste du monde, le pin maritime a été introduit au Chili (100 000 ha), en Australie occidentale (500 000 ha en cours de plantation), en Afrique du Sud (40 000 ha), en Argentine, en Nouvelle-Zélande (3 000 ha) et en Grèce (10 000 ha).

B.1.4 Habitat

Le pin maritime est bien adapté aux climats maritimes très tempérés, à température douce et régulière. Cette espèce exige une légère humidité de l'air, mais supporte la sécheresse estivale.

Il tolère des températures de 40°C voire plus quelque soit la provenance. Par contre, la variabilité est importante pour la résistance au gel. En France, les arbres originaires du Portugal ou de Galice ont montré une forte sensibilité lors des hivers de 1956 et 1985 tandis que les variétés landaises ont bien supporté des températures de -25°C (Maugé 1987).

Le niveau des précipitations importe peu si l'accès à une ressource en eau est possible, mais comme nous l'avons vu auparavant, l'humidité de l'horizon superficiel aura beaucoup d'importance sur l'assimilation des éléments nutritifs et donc la production de bois.

Cette essence héliophile exige beaucoup de lumière. C'est une espèce pionnière qui supporte mal la compétition avec une autre végétation, surtout lors de son installation.

Le pin maritime préfère les stations riches, avec des sols légers et profonds, bien pourvus en eau tant qu'il n'y a pas de compétition. Mais il se contente très bien le plus souvent de sols acides et pauvres (essence frugale) qui constituent la majorité de son aire de répartition. Ces sols présentent toutefois régulièrement des carences en phosphore.

Il se développe également bien sur des sols très pauvres en éléments nutritifs -comme les dunes- grâce à son système traçant et aux mycorhizes avec lesquelles il entretient des relations symbiotiques.

Cette espèce est souvent dite calcifuge en raison de son affinité pour les pH acides. Or à ce pH, les argiles sont dispersées, peu pénétrables et le sol compact, ce qui ne permet pas au système traçant de se développer.

B.2 Les plantes et la sécheresse

B.2.1 L'eau et la plante

B.2.1.1 L'eau du sol

La plante tire l'eau et les éléments nutritifs du sol.

Un sol est un milieu complexe, constitué de particules solides (pierres, graviers, sables ou limons), de colloïdes et d'une phase liquide. Les colloïdes sont des macromolécules très hydrophiles qui peuvent être d'origine minérale –ce sont des argiles– ou d'origine organique –les humus-. La combinaison de ces colloïdes forme un complexe argilo-humique. La phase liquide est directement accessible par les racines et constituée d'eau et de substances dissoutes. Elle n'est cependant pas libre et plus le sol se dessèche, plus il est difficile d'en extraire de l'eau qui est liée aux constituants du sol par les forces osmotiques et les forces matricielles.

Les forces osmotiques ont pour origine la concentration en ions et molécules organiques des solutions. Les solutés de deux solutions séparées par une membrane perméable –comme ceux du sol et ceux des racines- iront spontanément de la solution la plus concentrée vers la solution la moins concentrée. De la même manière les molécules d'eau iront de la solution où le potentiel chimique de l'eau est élevé vers celle où il est le plus bas. Le potentiel chimique représente l'enthalpie libre d'une mole d'eau et dépend de la « concentration de l'eau ». En cas de sécheresse, la « concentration en eau » étant plus faible, celle des ions et molécules organiques est donc plus élevée. Les forces osmotiques peuvent alors atteindre 20 bars.

Les forces matricielles englobent les forces d'imbibition et les forces capillaires. Les forces d'imbibition ont pour origine les attractions électrostatiques entre les charges négatives des colloïdes et les charges positives des atomes d'oxygène des molécules d'eau. Lorsque le sol se dessèche les forces mises en œuvre atteignent plusieurs centaines de bar. Les forces capillaires sont liées à des phénomènes de tension superficielle et retiennent l'eau dans les interstices fins.

La notion de potentiel hydrique Ψw est couramment employée pour évaluer les capacités d'une plante à prélever de l'eau du sol.

Pour la plante, le potentiel hydrique dépend essentiellement du bilan de deux forces opposées : le potentiel osmotique Ψs qui dépend de l'osmolarité de la solution baignant les

cellules des racines et la turgescence Ψp exercée par les parois cellulaires. Les flux hydriques résulteront de cet équilibre entre forces osmotiques et mécaniques.

$$\psi_W = \psi_S + \psi_P$$

Le potentiel hydrique d'un sol est égal mais de signe opposé, à l'énergie qu'il faut lui appliquer pour en libérer un gramme d'eau.

Pour maintenir un flux d'eau du sol vers les racines, la plante doit maintenir le potentiel hydrique des racines inférieur à celui du sol. Dans le cas contraire, il y a perte de turgescence et la plante perdra de l'eau. En situation de stress hydrique, le potentiel hydrique du sol s'abaisse, la plante doit alors répondre en abaissant également le potentiel hydrique des racines et peut pour cela, diminuer la pression mécanique exercée par les parois (début de plasmolyse) ou moduler la pression osmotique en augmentant l'osmolarité de la solution cellulaire par accumulation de solutés ou osmoticum, ce mécanisme est appelé osmorégulation.

B.2.1.2 L'eau dans la plante

L'eau captée par les racines va occuper l'ensemble des cellules et des espaces intercellulaires du végétal. La mesure d'un autre paramètre, la teneur en eau relative, permet de connaître le niveau de saturation en eau ou de turgescence de la plante. Il est déterminé de la façon suivante : $RWC = \left[(MF - MS)/(MT - MS)\right] \times 100$ RWC=Teneur Relative en Eau

MF, MS, MT= Masse de Matière Fraiche, Sèche et Turgescente

Au niveau des racines, l'eau est rapidement canalisée dans le système conducteur xylémien constitué de cellules différentiées en vaisseaux conducteurs et en trachéides s'arrêtant dans les organes foliaires. La circulation de l'eau s'effectue du sol vers les feuilles où elle va être évaporée sous l'action du rayonnement solaire. Les mouvements d'eau dans la plante sont donc globalement axiaux du sol à l'atmosphère. Un arbre évapore ainsi plus de 200 litres d'eau par jour (Wullschleger *et al.* 1998). Cependant, sous l'action de la photosynthèse, les chloroplastes produisent des hydrates de carbones comme le saccharose. Ces assimilas gagnent un système conducteur phloémien formé de tubes criblés où circule la sève élaborée qui va alimenter l'ensemble de la plante. Le sens de circulation est principalement opposé à celui de la sève brute bien que la sève alimente également les zones de croissance active que sont les bourgeons.

La montée de l'eau depuis le sol jusqu'aux feuilles à travers le système vasculaire nécessite un système de forces. Outre les pressions osmotiques déjà citées, la principale est expliquée

par la théorie de la tension-cohésion foliaire et la transpiration énoncée par Dixon et Joly en 1894 et reformulée par Zimmermann en 1983. Elle repose d'une part sur l'existence d'une continuité hydraulique entre le sol, la plante et l'atmosphère et d'autre part sur la forte cohésion des molécules d'eau entre-elles et avec les parois des trachéides et des vaisseaux. La plante peut être comparée à une colonne d'eau, l'évaporation de l'eau au niveau des feuilles provoque un appel d'eau qui se transmet à toute la colonne et se traduit par une force de succion au niveau racinaire. Les parois des éléments vasculaires doivent donc supporter des tensions fortes et d'autant plus importantes que la disponibilité en eau du sol est faible. La résistance et les caractéristiques hydrauliques de ces parois auront une influence sur les propriétés de réponse à des situations de stress hydrique (Steudle 2001).
Les forces de cohésions supportent des tensions pouvant atteindre 20 bars mais la présence d'une rupture dans la colonne comme la présence d'une bulle d'air provoque une cavitation. L'élément conducteur se remplit d'air puis de vapeur d'eau et la circulation d'eau est stoppée.

La transpiration est observée sur toute la surface de l'appareil foliaire mais elle est localisée essentiellement au niveau des stomates sur la face inférieure des feuilles ou sur les deux faces dans le cas des aiguilles. Il s'agit de dispositifs formés par deux cellules de garde contrôlant la fermeture d'un espace appelé chambre sous-stomatique qui communique avec l'extérieur par l'ostiole. Ils sont le lieu d'échange du CO_2 et de l'H_2O nécessaire à la photosynthèse. Chez les arbres, le CO_2 diffuse vers les cellules du mésophile où il se combine sous l'action de la Rubisco au Ribulose 1,5 bisphosphate pour donner l'Acide 3-phosphoglycérique qui sera utilisé dans le cycle de Calvin. Si l'apport de ces gaz et de l'eau est indispensable pour permettre la réaction chimique et donc la croissance, les stomates laissent également passer l'eau sous forme de vapeur. La plante module l'ouverture des stomates en fonction des conditions de lumière, de l'état hydrique de la plante et du déficit de pression de vapeur (VPD). Le système stomatique a été largement étudié et est un des principaux acteurs de régulation de l'état hydrique de la plante (Li *et al.* 2000; Schroeder *et al.* 2001; Laporte *et al.* 2002; Luan 2002). Les stomates ont donc un rôle majeur dans la réponse au stress hydrique (Buckley 2005). Quant à la transpiration de l'épiderme, elle va dépendre de la nature de celui-ci et la présence d'une cuticule (Burghardt and Riederer 2003).

B.2.1.3 L'eau et les racines

L'appareil racinaire est l'organe spécialisé dans l'absorption de l'eau et des substances minérales en régulant comme nous venons de le voir les propriétés osmotiques de ses tissus. Il

apparaît donc comme une cible d'étude particulièrement intéressante dans le cadre de réponses à des déficits hydriques. Cependant sa morphologie complexe et ramifiée rend les observations difficiles en terre. Les expériences en hydroponie permettent de simuler l'état de sécheresse d'un sol en modulant l'osmolarité de la solution par l'ajout de macromolécules comme le polyéthylène glycol.

Nous avons vu que l'eau était canalisée dans des éléments vasculaires. Avant de parvenir à ces structures différentiées, l'eau est absorbée par les cellules épidermiques puis doit franchir l'endoderme. Le transfert passif des solutés de cellule à cellule reste faible à cause de la faible perméabilité aux solutés des membranes. L'eau par contre circule plus facilement au travers des membranes. Elle peut emprunter différentes voies de circulation (Steudle 2000) (Fig. B-8):

- L'eau peut circuler dans le symplasme, c'est-à-dire de cytoplasmes en cytoplasme sans transit par la vacuole, par l'intermédiaire des plasmodesmes qui sont des communications intercellulaires.
- Elle peut également circuler dans l'apoplasme qui englobe l'ensemble des parois pectocellulosiques, lacunes et méats, en communication directe avec le milieu extérieur.
- Lorsque le transport s'effectue au travers des membranes plasmiques, il peut être facilité par des aquaporines. Il s'agit de la voie transcellulaire.

Dans les deux premiers cas, aucune membrane n'est traversée alors que dans la circulation transcellulaire, deux membranes doivent être traversées par couche cellulaire. La régulation de la perméabilité membranaire par la présence d'aquaporines va pouvoir moduler la circulation hydrique. Les flux apoplasmiques vont dépendre de la composition et de l'épaisseur des parois et de la présence de barrières imperméables telles que la bande de Caspari ou une lamelle de subérine que l'on retrouve au niveau de l'endoderme ou de l'exoderme (Fig. B-8). La présence de ces deux structures va modifier considérablement les propriétés hydrauliques du tissu considéré.

Le trajet de l'eau va pouvoir être combinée entre les voies décrites. La circulation symplasmique et transmembranaire ne vont pas pouvoir être distinguées en pratique et forment la circulation de cellule à cellule.

Les racines sont donc des structures composites et ne peuvent pas être résumées à la simple membrane parfaite d'un osmomètre.

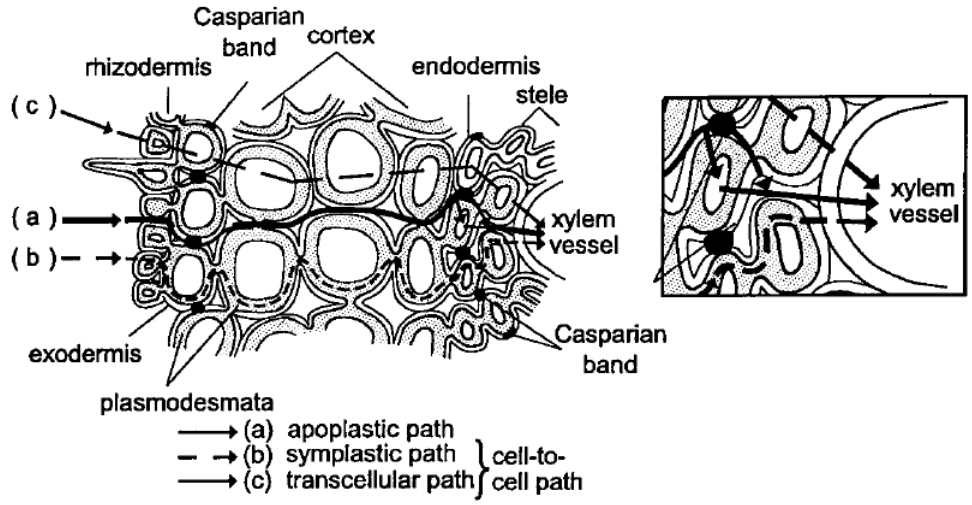


Figure B-8 : Trajet de l'eau dans les racines : a) voie apoplasmique, b) voie symplasmique, c) voie transcellulaire de vacuole à vacuole. La voie apoplasmique est interrompue par le cadre subérifié (Bande de Caspari) d'après Steudle *et al.* (2000)

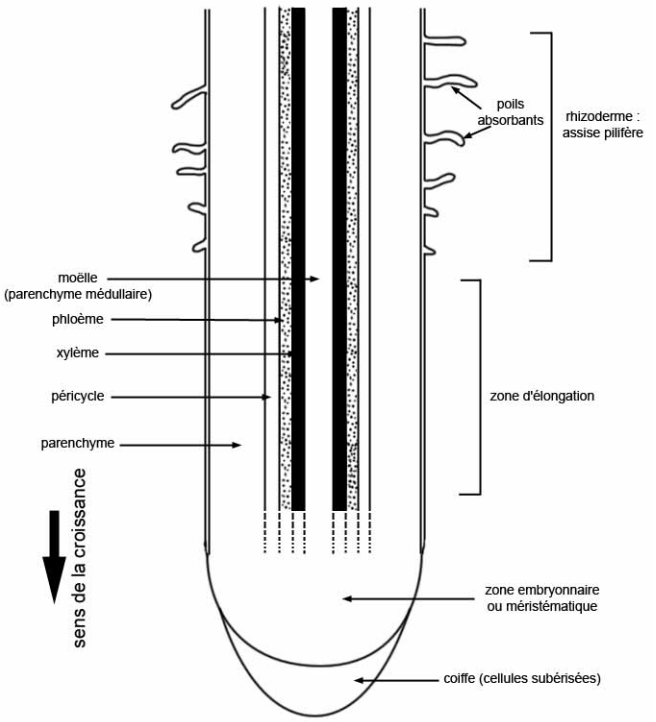


Figure B-9a :
Schéma d'une coupe longitudinale de la partie distale d'une racine.

http://fr.wikipedia.org/wiki/Racine_(botanique)

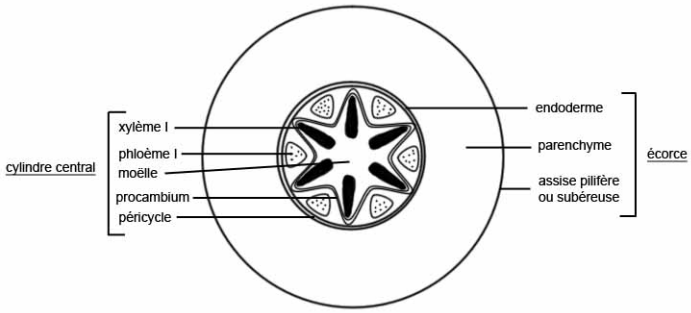


Figure B-9b :
Schéma d'une coupe transversale de la zone subéreuse d'une racine.

http://fr.wikipedia.org/wiki/Racine_(botanique)

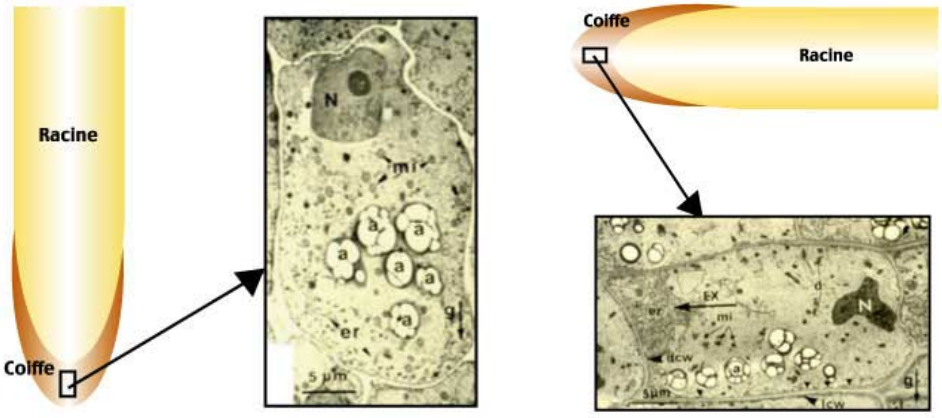


Figure B-10 : Perception de la gravité dans une racine de Lens culinaris.
A Gauche : Photographie d'un statocyte lorsque la racine est en position verticale
A Droite : Photographie d'un statocyte lorsque la racine est en position horizontale
N = Noyau, a = Amyloplastes. (D'après http://www.snv.jussieu.fr/bmedia/gravitropismeBM/perception.html)

Une racine peut être divisée en quatre zones (Fig. B-9a) :

- La zone subéreuse qui correspond à la partie la plus âgée de la racine. Elle porte les racines secondaires ou radicelles.
- La zone pilifère qui émet des poils absorbants. Ce sont eux qui permettent l'absorption de l'eau et des sels minéraux.
- La zone d'accroissement située derrière la coiffe. Elle est responsable de la multiplication cellulaire.
- La coiffe qui termine et protège la racine. Elle permet la pénétration dans le sol.

Au niveau anatomique, la zone subéreuse des racines est constituée de plusieurs tissus différenciés (Fig. B-9b). Au centre, une racine comporte une couche cellulaire appelée endoderme enfermant un cylindre central constitué du péricycle et des tissus conducteurs du phloème et du xylème transportant la sève brute et élaborée. Cette structure ou stèle est entourée par le parenchyme cortical et l'épiderme formant le cortex.
La zone d'accroissement, suivie de la coiffe entourant l'extrémité de la racine, est constituée de cellules en division au métabolisme très actif.
La coiffe protège le méristème apical où se déroulent les divisions cellulaires qui permettent la croissance de la racine. Des cellules centrales ou centre quiescent sont le siège de divisions cellulaires à l'origine des cellules initiales de l'endoderme et des cellules de la coiffe. Elles assurent le renouvellement des cellules pendant que les cellules apicales les plus externes se desquament progressivement. Les parois de ces cellules se transforment en un mucilage qui permet un glissement doux de la jeune racine. La coiffe joue un rôle majeur dans l'orientation de la croissance de la racine. Dans la région centrale de la coiffe (Columelle), les cellules riches en amidon servent de capteur de la gravité par la position des amyloplastes et permettent de modifier le sens de croissance de la racine en fonction de sa position (Fig B-10).

Une extraction des ARN de racines entières regroupera donc les niveaux d'expression des transcrits correspondant à différents types tissulaires. Les variations d'expression observées pourront donc correspondre entre autre, à des modifications de croissance, d'initiation de racines secondaire ou de poils absorbants, de division cellulaire ou encore de régulation de l'osmolarité. Les nouvelles méthodes de microdissection cellulaire comme celles utilisant des lasers permettent de focaliser l'étude sur un tissu ou un type cellulaire précis (Nakazono *et al.* 2003; Casson *et al.* 2005; Day *et al.* 2005). Cette approche est en

cours de développement au laboratoire et devrait permettre, conjointement à des méthodes d'amplification linéaire des ARNm extrait, d'affiner les études d'expression à l'échelle cellulaire.

B.2.2 Le stress hydrique

B.2.2.1 Les stratégies de réponses

Les végétaux rencontrent fréquemment des variations de leur environnement. Dès lors que les facteurs externes sortent des conditions optimales de croissance de la plante, ils provoquent un stress qui a des effets sur la croissance, le développement et la productivité. Ces stress peuvent être provoqués par des organismes vivants –on parle de stress biotique– ou par des excès physiques ou chimiques de l'environnement –on parle de stress abiotique-.

Parmi les stress abiotiques, le stress hydrique édaphique peut se définir comme un déficit de disponibilité en eau pour la plante. L'origine de ce déficit peut venir d'une salinité excessive du sol, d'une sécheresse ou du gel qui par cristallisation des molécules d'eau diminue sa disponibilité. Mais en général, différents stress sont combinés et un stress ayant pour origine la sécheresse peut également devenir un stress thermique (par augmentation de la température des organes) et un stress oxydatif (par altération des structures cellulaires). Les réponses de la plante sont multiples et si certaines sont spécifiques d'un stress, d'autres sont des réponses communes aux agressions extérieures (Fig. B-11).

Chaque espèce va répondre au stress par des stratégies de **résistance** différentes lui permettant de survivre. Chaves *et al*. (2003) détaillent les mécanismes de résistance à la sécheresse. Ils sont groupés en trois catégories même si ces mécanismes ne sont pas exclusifs et qu'une plante combine toute une gamme de types de réponses:

- L'**échappement** concerne les espèces capables d'accomplir leur cycle de vie avant que le déficit physiologique en eau n'intervienne. C'est le cas des plantes des régions arides effectuant leur cycle reproductif après une pluie et dont les graines supportent de longues périodes de sécheresse. Cette stratégie ne s'applique pas aux plantes longévives comme les arbres.
- L'**évitement** consiste à maintenir le potentiel hydrique et à limiter au maximum les pertes d'eau par transpiration pour se prémunir de l'exposition au stress. Certaines plantes ferment très rapidement leurs stomates, d'autres limitent la masse foliaire, présentent des feuilles modifiées ou conservent les feuilles âgées pour se protéger du rayonnement solaire. Les feuilles sont aussi parfois munies de cuticules épaisses. Une

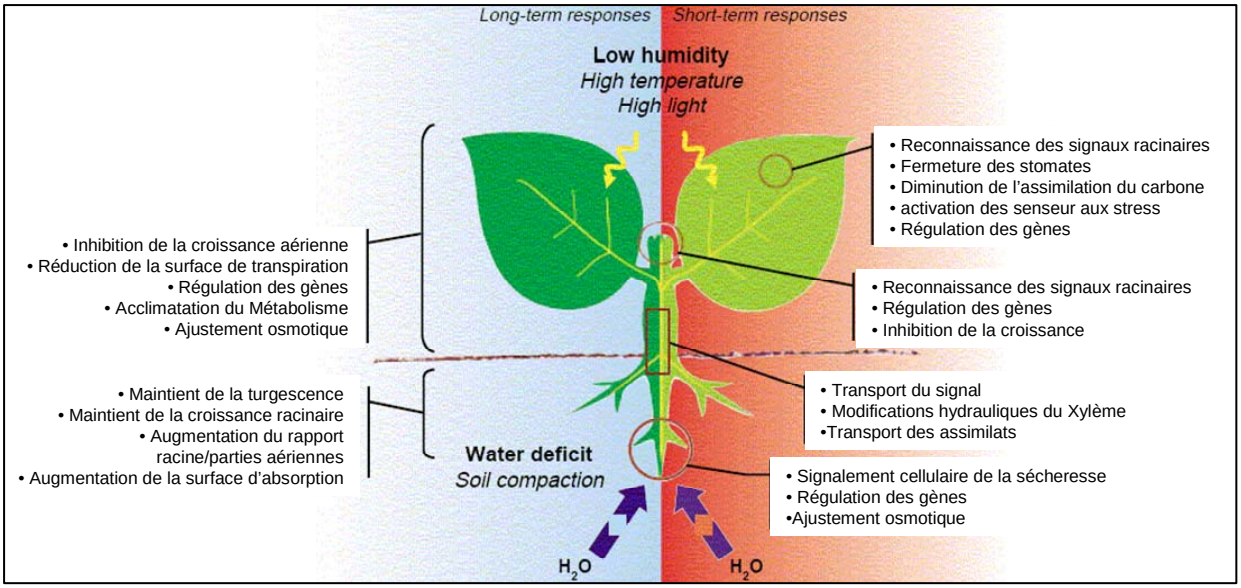


Figure B-11 : Réponses de la plante à la sécheresse. A gauche, réponse à long terme ou acclimatation ; A droite, réponses à court terme. D'après Chaves *et al.* (2003)

allocation du carbone privilégiant un système racinaire profond et développé permet également d'optimiser l'accès aux ressources en eau.

- Les plantes **tolérantes** mettent en œuvre des mécanismes leur permettant de supporter le stress. En diminuant le potentiel hydrique elles maintiennent la turgescence par l'ajustement du potentiel osmotique, obtenu par accumulation d'ions minéraux et/ou de composés organiques.

Mais les plantes ont la possibilité de répondre au stress par une gamme importante de processus que nous détaillerons ci-après.

Les espèces ne présentent pas toutes la même plasticité de réponse au stress. Certaines caractéristiques sont déterminées génétiquement qu'il y ait présence du stress ou non, comme un système racinaire profond, un système foliaire constitué d'aiguilles ou un fonctionnement photosynthétique du type CAM. De telles **adaptations** morphologiques et physiologiques donnent un avantage sélectif à une population au cours de l'évolution.
D'autres mécanismes de résistance concernent les réponses d'un individu aux changements de son environnement. Une telle **acclimatation** va permettre à la plante d'ajuster son homéostasie pour établir un nouvel équilibre compatible à la survie dans ces nouvelles conditions. La résistance d'une plante à un stress sera une combinaison des caractères constitutifs de l'espèce et de sa plasticité propre. Une plante pourra donc être considérée comme appartenant à une espèce évitante (ex : morphologie constitutive) mais d'une variété tolérante (ex : capacité induite d'ajustement du potentiel osmotique).

B.2.2.2 Les systèmes modèles

Pour répondre à une question de recherche et identifier les processus sous-jacent, il est préférable de travailler sur des organismes présentant des aptitudes particulières quant au caractère étudié. Dans le cas de la résistance au stress hydrique, les espèces considérées peuvent se répartir en trois groupes. Nous verrons tout d'abord le cas des systèmes tolérants, puis les modèles génétiques et enfin les plantes de grande culture.

B.2.2.2.1 Systèmes tolérants

Avec une teneur en eau de 10%, la graine est une structure bien adaptée au manque d'eau. Elle permet à la plupart des espèces de survivre à la dessiccation et aux basses températures mais seulement sous cette forme. Cette dessiccation est indispensable au cycle reproductif. Elle est suivit par une réhydratation massive lors de la germination et par le

développement des cellules embryonnaires. Une étude portant sur le pois (*Pisum sativum*), le maïs (*Zea mays*) et le soja (*Glycine max*) a montré que cette tolérance à la dessiccation était progressivement perdue en cas d'imbibition des graines (Koster and Leopold 1988). Les auteurs enregistrent une corrélation entre le pourcentage de levée des graines après traitement (imbibition suivie d'une dessiccation) et la teneur en sucre en fonction du temps d'imbibition. Ingram et Bartel (1996) rapportent que les graines de nombreuses espèces ont fait l'objet d'études approfondies sur le transcriptome et le protéome, telles que *Arabidopsis thaliana*, le coton (*Gossypium spp.*), l'orge (*Hordeum vulgare*), le maïs (*Zea mays*) et le riz (*Oriza sativa*). Un des problèmes majeurs de ce modèle d'étude est de dissocier les gènes impliqués dans la tolérance de ceux impliqués dans le développement de l'embryon. Il a néanmoins permis d'identifier une famille de protéine, les LEA (Late embryogenesis Abundant), que l'on retrouve également fortement exprimée lors de la réponse au stress hydrique. Les LEA sont des protéines hydrophiles qui s'accumulent dans les stades avancés de la maturation de la graine puis disparaissent après la germination (Galau *et al.* 1986; Roberts *et al.* 1993). Le maintien de l'état dormant est fortement lié à la présence d'Acide abscissique (ABA) dont le rôle d'inhibiteur de la germination (Bewley 1997) a été démontré chez le soja, la tomate ou le blé mais pas chez le tournesol.

Le pollen est également une structure à très faible teneur en eau. Comme les graines, les mécanismes du stress sont confondus avec ceux du développement. Mais l'étude des transcrits de pollen a permis d'identifier des gènes impliqués dans la dessiccation. C'est le cas du transcrit LLA23 cytoplasmique identifié sur *Lilium longiflorum* présent uniquement dans le pollen et dont la séquence est proche de celle d'autres protéines impliquées dans le stress (Lp-3, DS2 et ASR) (Huang *et al.* 2000).

Un autre système d'étude regroupe certaines espèces remarquables, capables de tolérer des dessiccations presque totales et de reprendre une turgescence après réhydratation. Ce sont les plantes reviviscentes ou poïkilohydriques parmi lesquelles on peut citer *Xerophyta viscosa* (Baker) et *Craterostigma plantagineum* (Bartels and Salamini 2001) pour laquelle Peter Scott (2000) souligne les modifications du métabolisme intervenant au cours du cycle déshydratation-réhydratation. Les feuilles réduisent leur surface de 85% avec une teneur en eau comprise entre 5 et 10%, les cellules sont fortement concentrées en sucres et des protéines LEA sont fortement exprimées. L'auteur souligne aussi que très peu d'études ont été réalisées sur les racines. Ces systèmes possèdent des mécanismes de protection des cellules très développés et permettent de découvrir et d'étudier de nombreuses protéines du stress. C'est le

cas de *Xerophyta viscosa* (Mowla *et al.* 2002), pour laquelle ces auteurs décrivent une nouvelle enzyme XvPer1 homologue à des antioxydants, ou encore de *Craterostigma Plantagineum* qui a récemment permis de découvrir une nouvelle famille de gènes de réponse à la déshydratation et à l'ABA nommée PTP (Plastid Targeted Protein) (Phillips *et al.* 2002) ainsi que des facteurs de transcription régulant cette réponse tels que CpR18 (Hilbricht *et al.* 2002).

D'autres espèces sont capables de croître à la fois sur des sols à très forte teneur en sel et de survivre à des températures très basses comme *Mesembryanthemum crystallinum*. Elles constituent également de bons modèles d'étude des systèmes de tolérance à la dessiccation (Barkla *et al.* 1999).

Si les mécanismes de réponse sont exacerbés chez ces plantes, ils sont cependant communs aux autres végétaux. Contrairement au système « graines », ils permettent de s'affranchir des effets développementaux.

B.2.2.2.2 Modèles génétiques

Une autre approche porte sur l'utilisation des plantes modèles. L'effort de recherche de la communauté scientifique sur ces espèces a permis de générer une base de connaissances conséquente en physiologie et en génomique. La séquence complète du génome est désormais disponible chez *Arabidopsis thaliana* ("A. thaliana genome" 2000; Initiative 2000), le riz (Yu *et al.* 2002) et le peuplier. Pour d'autres espèces comme le blé, le maïs et la tomate, le séquençage du génome est en cours mais des cartes génétiques fines et des bases de données d'EST et de BAC sont disponibles, rendant possible les approches de clonage positionnel et l'étude de l'expression quasi exhaustive de l'ensemble des gènes. Des collections de mutants ou de variétés permettent d'identifier des individus sensibles, résistants et d'identifier les gènes et voies métaboliques concernées. En 2003, sur *Arabidopsis*, plus de 88'000 insertions par *Agrobacterium* avaient été localisées précisément sur le génome ce qui représentait plus de 21'700 gènes pour lesquels un mutant était désormais disponible (Alonso *et al.* 2003). Hirochika *et al.* (2004) ont recensé les ressources en mutant disponibles chez le riz dont le nombre dépasse le million. Ils soulignent l'intérêt de mettre à disposition de la communauté une banque de séquences bordant les mutations (FST, Flanking Sequence Tag) dont 1373 ont été identifiées en 2005 (Enckevort *et al.* 2005). En effet, la connaissance de la séquence des régions flanquant le transposon permet d'effectuer des recherches électroniques d'homologie pour un gène donné, d'identifier un mutant pour ce gène et de commander les semences pour étudier l'impact phénotypique de cette mutation.

Des mutants ultra sensible au sel (mutants sos) ont permis de déterminer les gènes régulés spécifiquement lors du stress (Gong *et al.* 2001). Les auteurs ont identifié une liste de 89 gènes régulés par soustraction de banques d'ADNc d'*Arabidopsis* sauvage contrôle vs stressé, puis ont comparé leur expression avec les mutants sos. Six gènes ont pu alors être identifiés.
De nombreux autres mutants ont servi à identifier des gènes et protéines impliqués dans la réponse à la sécheresse (Chen *et al.* 2005). Notamment les gènes régulé par les hormones et en particulier l'ABA et les gènes ABI (ABA-insensitive) associés à la réponse à cette hormone (Campalans *et al.* 1999; Bianchi *et al.* 2002; Dong *et al.* 2005).

B.2.2.2.3 Plantes de grande culture

Avec l'agriculture intensive, la question de l'alimentation en eau est devenue primordiale. Les exigences en eau des cultures rendent souvent l'irrigation incontournable alors que les ressources du milieu sont limitées. Les semenciers proposent une large gamme de variétés dont la résistance à la sécheresse est variable. Les surfaces plantées, les climats différents et les degrés de tolérances disponibles font des plantes de grandes cultures un objet d'étude intéressant pour des études de génétique quantitative en conditions naturelles. Les programmes de recherche des grands groupes agroindustriels ont d'ores et déjà brevetés des variétés résistantes ou modifiées génétiquement. Leurs recherches s'orientent actuellement sur des variétés résistantes à la sécheresse (ex : http://www.monsanto.com/monsanto/layout/sci_tech/prod_pipeline/productpipeline.asp).

B.2.2.3 Réponses morphologiques et physiologiques

B.2.2.3.1 Régulation de la croissance

L'impact d'un déficit en eau se traduit par une diminution de la croissance de la plante. Les rendements au champ peuvent chuter très rapidement. Une analyse des pertes de rendement par application de stress hydriques chez le riz a été réalisée par Pantuwan (2000). Selon l'intensité du traitement, le stade de développement et la modalité d'application, des chutes de 18 à 81% ont été enregistrées par rapport au témoin irrigué de façon optimale. L'eau est indispensable à la photosynthèse et à la régulation thermique par évaporation de l'appareil foliaire. En situation de stress, le défaut d'alimentation en eau entraîne une chute de la photosynthèse. En outre, l'eau du sol en solubilisant les éléments nutritifs facilite leur transport jusqu'aux racines et leur assimilation. En situation de sécheresse les flux hydriques sont plus faibles et la nutrition diminuée.

Au niveau cellulaire, le grandissement commence par un relâchement des parois accompagné par une absorption de solutés permettant de maintenir le potentiel osmotique. Sous l'action de la pression osmotique, et du relâchement des tensions des parois, l'augmentation de volume est initiée. Nous le verrons par la suite, la modification des propriétés physiques des parois est sous le contrôle de gènes spécifiques. Ces phénomènes de croissances cellulaires ont été étudiés par Triboulot (1996) dans le cadre de sa thèse.

La relation entre la consommation en eau et la production de composés organiques lors de la photosynthèse est évaluée par l'efficience d'utilisation de l'eau : rapport entre la quantité de biomasse produite et l'eau utilisée par la transpiration. Ce paramètre est pour une part déterminée génétiquement et constitue une caractéristique de l'espèce ou de l'individu et pour une part sous contrôle environnemental (Brendel *et al.* 2002). La perte en eau et l'assimilation du CO_2 sont régulés par la conductance stomatique à la vapeur d'eau et au CO_2, ainsi que par le gradient de concentration de ces deux éléments entre l'intérieur (w_i, c_i) et l'extérieur (w_a, c_a) de la feuille. Une plante va pouvoir augmenter son efficience d'utilisation de l'eau en limitant les pertes en eau par fermeture des stomates et en augmentant sa capacité photosynthétique pour le CO_2 ce qui signifie une diminution de la concentration sous stomatique en CO_2 (c_i bas). Une telle plante consommera moins d'eau à masse de composés organiques produite égale.

La discrimination isotopique du carbone est une mesure indirecte de cette efficience d'utilisation de l'eau. En effet, le carbone existe sous la forme ^{12}C ou ^{13}C dans le CO_2 atmosphérique. Le ^{13}C, de part sa masse plus élevée, diffusera plus difficilement à travers les stomates et sera plus difficilement incorporé lors des réactions biochimiques. Son incorporation finale dans les composés organiques dépendra donc de la conductance stomatique et de la concentration en CO_2 de la feuille, et par conséquent à l'efficience d'utilisation de l'eau. Mais ce ne sont pas les seuls paramètres à intervenir (état hydrique du sol, appareil racinaire, discrimination du carbone d'autres sources..) et la relation directe n'est pas forcément mise en évidence entre espèces chez les conifères (Froux 2002).

En terme de performances, l'avantage d'une efficience d'utilisation de l'eau élevée vis-à-vis de la sécheresse est très discutable, et détaillée dans l'ouvrage édité par Bacon (2004). Les variétés à forte efficacité d'utilisation de l'eau ne pourront pas toujours tirer avantage de cette caractéristique en conditions d'alimentation en eau non limitante et la biomasse produite pourra être, dans ces conditions, inférieure à celle d'une variété à faible efficacité d'utilisation de l'eau. C'est le cas des écotypes landais et marocains. Sur les tests de provenance plantés dans les Landes, la provenance marocaine se caractérise par une vigueur moindre en hauteur

et diamètre par rapport à la provenance landaise. Les mesures de $\delta^{13}C$ sur des cernes de bois ont révélées une valeur plus basse pour la provenance « Tamjout » qui peut être considérée comme évitante en comparaison à la provenance landaise (Nguyen-Queyrens *et al.* 1998).

La sensibilité à la sécheresse en conditions extrêmes peut aussi provenir d'un risque élevé à l'embolie. Il s'agit d'une cavitation des vaisseaux conducteurs qui se remplissent d'air ou de vapeur d'eau. La sève circule sous tension, l'introduction accidentelle d'une bulle d'air –par un insecte, l'abscission d'une feuille, une blessure ou lors d'un gel/dégel- ou une tension trop forte, remplit l'élément conducteur et provoque l'arrêt de la circulation. La propagation du phénomène entre éléments conducteurs voisin va dépendre de la taille et de la morphologie des pores de communication entre vaisseaux. Des pores larges laisseront passer une bulle d'air sous des tensions plus faibles. Les pores des gymnospermes sont protégés par des torus qui fonctionneraient à la manière d'une soupape si une différence de pression est établie de part et d'autre de la ponctuation (Fig. B-12). Il existe une variabilité inter et intra-spécifique à l'embolie. F. Froux (2002) observe des réponses et sensibilités différentes entre *C. sempervirens*, *C. atlantica*, *P. halepensis* et *P. nigra* où *P. nigra* est l'espèce la plus vulnérable et *P. halepensis* la moins vulnérable. Une régulation stomatique efficace semble jouer un rôle en permettant de ne pas atteindre des tensions critiques à l'origine de l'embolie dans une stratégie d'évitement. Cependant, pour les deux espèces du genre *Pinus*, la marge de sécurité entre le potentiel hydrique de fermeture des stomates et celui d'initiation de la cavitation, était pratiquement nul alors que ces deux espèces ont des vulnérabilités différentes. D'autres facteurs interviennent et l'auteur souligne l'importance des parties racinaires où l'embolie peut se produire mais où des mécanismes de « désembolisation » -poussée racinaire nocturne- et de pousse de nouvelles racines latérales est possible.

La conductance hydraulique peut également chuter si les parois des éléments conducteurs ne résistent pas aux pressions négatives élevées. Les parois du xylème peuvent se collapser et bloquer la circulation. Ce phénomène réversible lors de la réhydratation a été mis en évidence par Cochard *et al.* (2004) sur des aiguilles de pins soumises à des déshydratations poussées. L'implication de ce nouveau système de réponse reste à étudier plus précisément.

Le métabolisme et l'expression des gènes spécifiquement impliqués dans ces mécanismes seront une autre source de distinction des variétés face à la résistance au stress hydrique.

B.2.2.3.2 Croissance racinaire

Comme évoqué précédemment, pour une station donnée, la morphologie du système racinaire va déterminer la disponibilité en eau accessible par la plante. Lors d'un stress

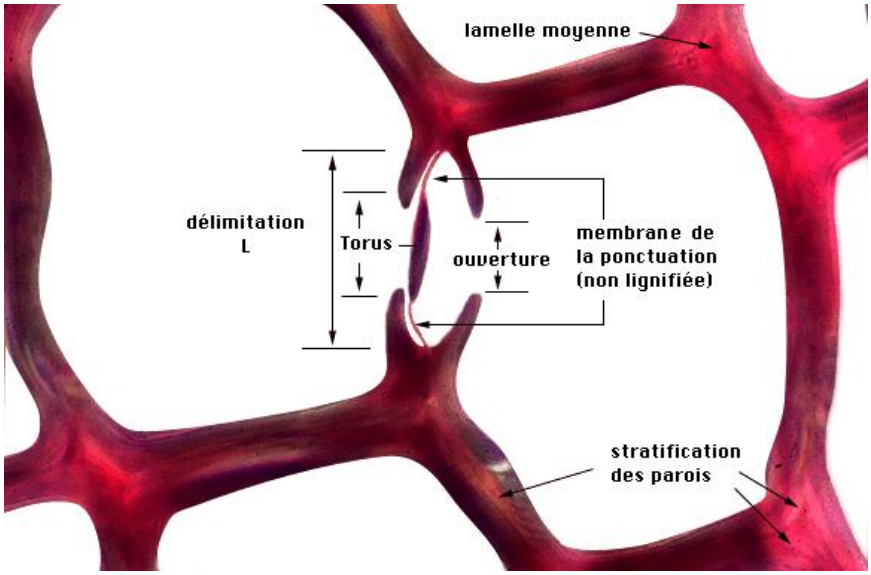


Figure B-12 : Trachéides du xylème secondaire d'une tige de pin blanc (*Pinus strobus*) : détail d'une ponctuation aréolée avec torus. (http://sylva.for.ulaval.ca/foret/anatomie/types_cell/trach_transv_torus.html)

hydrique, la croissance des parties aériennes est généralement réduite mais la plante peut opter pour une stratégie de croissance racinaire. La réallocation du carbone dans cette voie peut lui permettre d'atteindre de nouvelles ressources en eau. Des expériences menées sur *Arabidopsis* confirment cette stratégie avec une augmentation de la biomasse racinaire lors de stress modérés (van der Weele *et al.* 2000). C'est aussi le cas chez le pin maritime pour lequel une valeur plus élevée du ratio de biomasse racinaire sur parties aériennes est observé pour l'écotype marocain (Sarrauste 1982; Loustau *et al.* 1995; Nguyen-Queyrens *et al.* 1998).

Dans une revue récente (Ueda *et al.* 2005) sur le développement des apex racinaires, les auteurs ont classé les gènes connus selon leur rôle dans la destinée et la morphologie du chevelu racinaire, la séparation de l'endoderme et du cortex, ou la spécification des cellules quiescentes et celles de la coiffe. La première catégorie regroupe la majorité des gènes découverts avec par exemple GLABRA2(GL2) qui est une protéine à domaine homéobox jouant un rôle clé dans l'initiation d'un chevelu racinaire et régulée par une liste de facteurs de transcription favorisant sa régulation (WER, TTG1, GLABRA3(GL3), EGL3) ou au contraire l'inhibant (CPC, TRY, ETC1). Les auteurs rapportent également d'autres gènes impliqués comme ACTIN2(ACT2), KOJAK(KjK), OXI1 ou RabA4b, et citent de nombreux gènes dont l'expression est liées à la présence d'une hormone végétale dont nous parlerons ci-après : l'Auxine. Les gènes impliqués dans les mécanismes des deux autres catégories sont pratiquement inconnus à ce jour mais le fonctionnement des apex racinaires est primordial puisque c'est à ce niveau que se décide l'axe de croissance de la future racine.

Jusqu'à présent, très peu d'études sur l'hydrotropisme ont été réalisées. L'orientation de la croissance est en effet également fortement dépendante du gravitropisme et les effets se confondent. Eapen *et al.* (2005) rapportent les résultats obtenus lors de cultures horizontales d'*Arabidopsis* sur des milieux gélosés avec un gradient de potentiel hydrique. La croissance des racines des phénotypes sauvages présente un changement brutal de direction pour éviter la région à potentiel hydrique bas ce qui n'est pas le cas du mutant nhr1 caractérisé par un gravitropisme exacerbé et de volumineux amyloplastes (Eapen *et al.* 2005). Les auteurs citent également d'autres expériences sur des mutants touchant le transport polarisé basipète de l'auxine et insensibles au gravitropisme. Les résultats indiquent qu'une part du gravitropisme aurait pour origine le gradient en auxine de la coiffe racinaire à la zone d'élongation. Ils avancent l'hypothèse que l'auxine interviendrait dans l'hydrotropisme avec une voie similaire à celle du gravitropisme. L'acide abscissique (ABA) pourrait jouer également un rôle au vu de la sensibilité à l'hydrotropisme de mutants aba1-1 et abi2-1 décrit par McCourt (1999). Le comportement de ces mutants montre qu'il y a bien un contrôle génétique de l'hydrotropisme

mais les gènes impliqués ne sont pas encore connus et rien ne nous renseigne sur la nature des senseurs hydriques de cette région. Cependant, les hormones contrôlent fortement le développement racinaire.

L'architecture du système racinaire est basée sur le développement de racines latérales sur des racines existantes. L'application de ce modèle réitératif aboutit à une morphologie complexe. Pour observer l'effet d'un stress osmotique sur le développement des racines latérales, Deak et Malamy (2005) ont placé des plantules d'*Arabidopsis* dans des milieux gélosés contenant des sels, du Chlorure de Potassium ou du Mannitol. Ils ont pu observer une inhibition presque totale du développement de racines latérales. L'initiation des divisions du péricycle n'était pas altérée mais le développement des primordiums était stoppé. Le stress osmotique régule donc le développement des racines latérales. L'Acide abscissique étant connu pour avoir un rôle clé dans les réponses osmotiques et pour bloquer le développement latéral des racines, les auteurs ont testé le comportement des mutants ABA-déficients aba2-1 et aba3-1. Ils ont pu observer un développement du système latéral réduit de 52.2 et 28.4% par rapport au type sauvage présentant une réduction de 97.7% dans un milieu additionné de mannitol ce qui confirme le rôle de l'hormone dans le contrôle de la croissance des racines latérales. Les auteurs ont également analysé un autre mutant –lrd2– leur permettant de découvrir un gène de régulation du développement des racines latérales nommé LRD2 et rapportent d'autres gènes impliqués dans ce phénomène et décrits dans la littérature comme ARN1, LIN1, PDR2 ou NIT3.

A l'inverse, certains facteurs comme l'auxine ou l'oxyde d'azote (NO), favorisent le développement des racines latérales. Sur des plants de tomates, Correa-Aragunde *et al.* (2004) ont étudié le rôle de l'oxyde d'azote sur la présence de racines latérales. Ils montrent qu'un traitement avec un donneur de NO oriente le développement racinaire vers un système latéral tandis qu'un traitement avec un inhibiteur de NO induit un phénotype sans racine latérale. De plus, le dosage de la teneur en NO endogène leur a permis de détecter une accumulation dans les cellules lors du développement de primordiums de racines latérales. Ils confirment ainsi l'implication du NO déjà identifié dans les réponses aux hormones végétales.

B.2.2.3.3 Les hormones

L'acide abscissique et l'éthylène sont couramment qualifiés d'hormones du stress de par leur implication dans les réponses de la plante aux perturbations de son environnement. Si de nombreux composés ont une action physiologique au niveau cellulaire ou tissulaire, une catégorie de molécules, les hormones végétales présentes en très faible quantité ont un effet

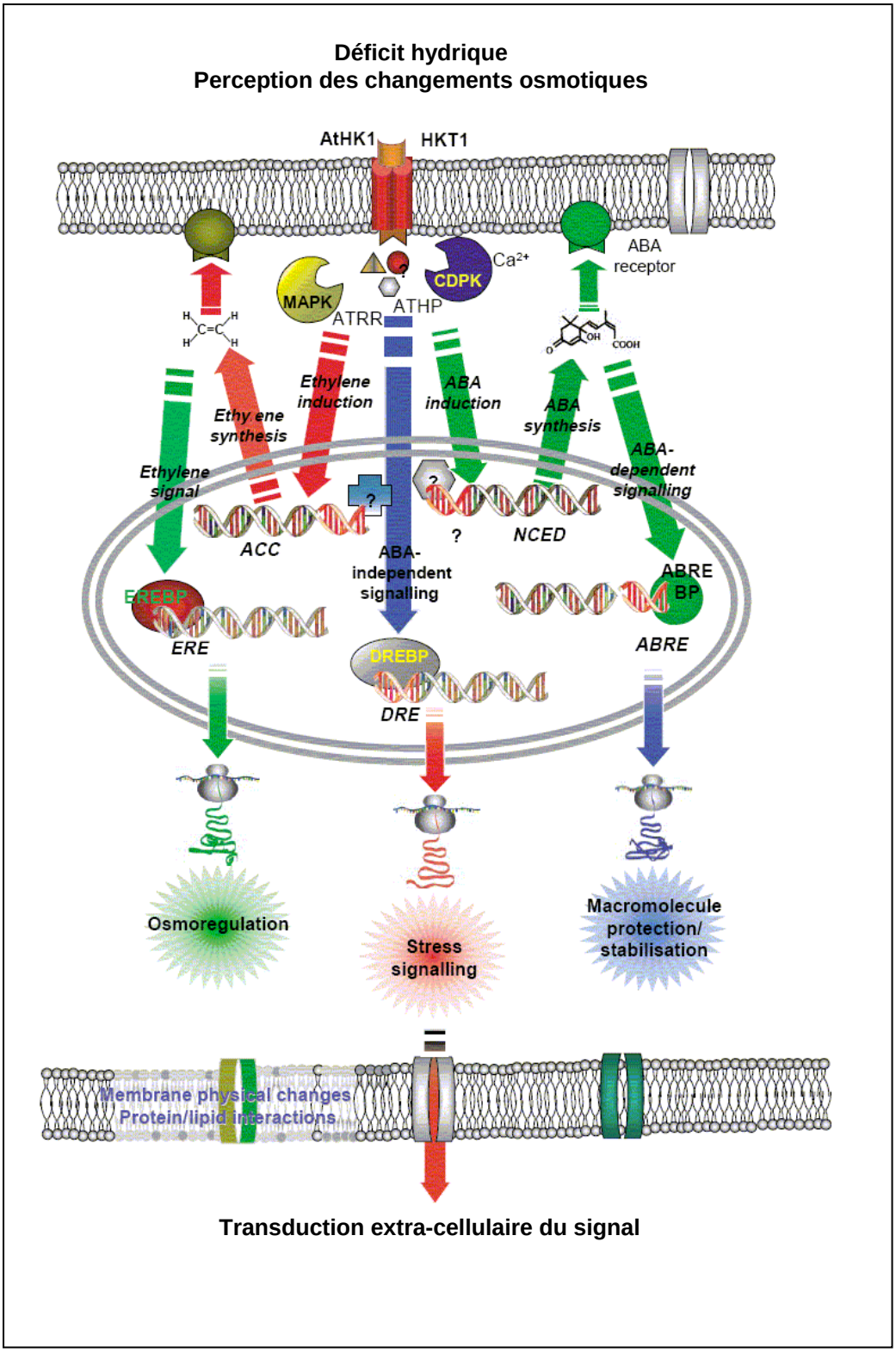


Figure B-13 : Détection, transmission du signal et réponses à la sécheresse au niveau cellulaire (d'après Chaves *et al.* (2003)). Abréviations : ABRE : ABA-responsive element ; ABRE BP : ABRE binding protein ; ACC : 1-aminocyclopropane-1-carboxylic acid ; AtHK1 : Arabidopsis thaliana histidine kinase; ATHP : un phospho-relais intermédiaire; ATRR : un régulateur de la réponse; CDPK : Ca2+-dependent kinases; DRE : dehydratation responsive element; DREBP : DRE binding protein; ERE : ethylene-responsive element; EREBP : ERE binding protein; HKT1 : un transporteur à forte affinité au K+; MAPK : mitogen-activated protein kinase.

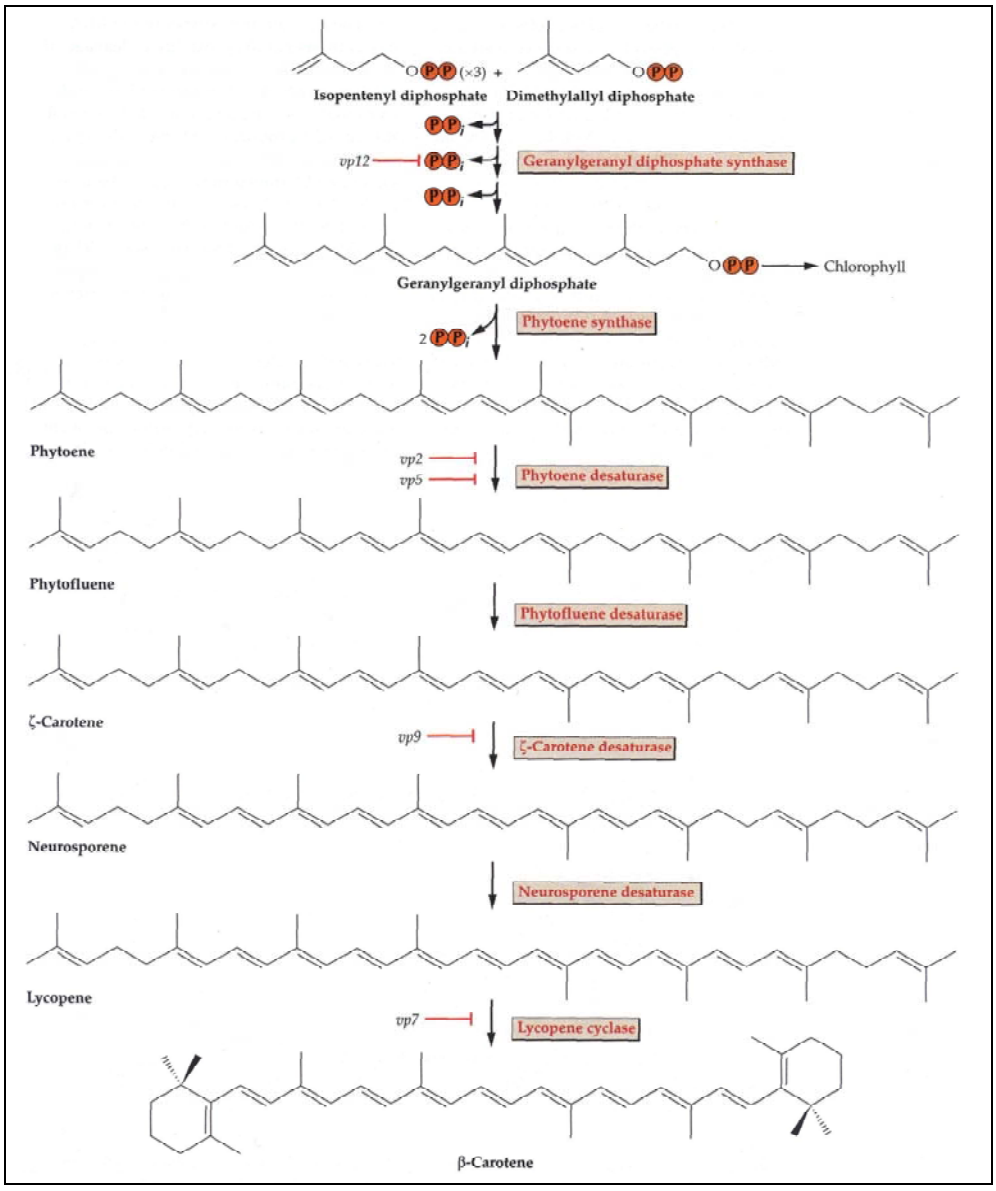


Figure B-14 : Etapes précoces de la voie de biosynthèse indirecte C40 de l'ABA. Production de GPP et synthèse du β-carotène. Les enzymes déficients des mutants vp sont indiqués. (Crozier *et al.*, 2000)

plus global et peuvent circuler au sein de la plante pour modifier l'état physiologique de celle-ci. Elles vont : propager une information du lieu de perception d'un stimulus extérieur vers les zones de réponse de la plante à ce signal, réguler la physiologie des cellules cibles et déclencher des voies métaboliques entières. Chaves *et al.* (2003) structurent ces voies en trois groupes : celles à médiation éthylène contrôlant l'osmorégulation, les ABA dépendantes à l'origine de macromolécules de protection/stabilisation et les ABA indépendantes impliquées dans le signalement du stress (Fig. B-13). Nous passerons en revue dans les paragraphes suivants les caractéristiques et les rôles des différentes hormones végétales.

B.2.2.3.3.1 Acide abscissique

Nous avons déjà évoqué l'implication de l'acide abscissique (ABA) dans la dormance des graines, l'hydrotropisme et l'inhibition du développement des racines latérales. Cette hormone a pour rôle physiologique principal communément admis l'**inhibition de la croissance** et du **développement** des parties aériennes sur des plants bien alimentés en eau, elle agit comme une anti-gibbérelline. Mais de récentes études sur le maïs en condition de stress hydrique indiquerait que ces caractères serait en fait sous le contrôle d'une **balance** entre l'ABA et l'éthylène (Sharp and LeNoble 2002). Elle a également un rôle important en cas de **déficit hydrique** en agissant sur la **fermeture des stomates**.

Sur le plan biochimique, il fait partie des isoprènes aussi appelés terpènes qui dérivent d'un précurseur : l'Isopentenyl pyrophosphate (IPP). La voie de synthèse est bien connue chez les bactéries mais diffère de celle des plantes qui a été récemment corrigée. En effet, pendant longtemps, il était admis que tous les terpènes dérivaient de l'acide mévalonique (MVA) mais une autre voie –voie MVA indépendante ou voie MEP- permettant de synthétiser l'IPP a été découverte chez les plantes supérieures.

Nambara *et al.* (2005) décrivent dans une revue détaillée la voie de synthèse de l'ABA qui est en fait formé par clivage de caroténoïdes (C_{40}) et non à partir du Farnesyl diphosphate (C_{15}). Le β-Carotène est obtenu à partir du Geranylgeranyl diphosphate (GGPP), lui-même issu de la condensation d'IPP et de Dimethylallyl diphosphate sous l'action d'une enzyme clé : la Geranylgeranyl diphosphate synthase (GGPP synthase) (Fig. B-14). Des mutants vp2, vp5, vp9 et vp7 touchant des enzymes de cette voie ont été identifiés chez *Arabidopsis*. Après métabolisme, le β-Carotène est d'abord transformé en Zeaxanthin puis successivement en Antheraxanthin et Violaxanthin sous l'action d'une zeaxanthin epoxidase (ZEP) pour laquelle des mutants caractérisés par une réduction en ABA et une accumulation en Zeaxanthin sont également disponibles. Le clivage de la Violaxanthin et de la Neoxanthin en Xanthoxin est

sous le contrôle d'une famille de 9 enzymes appelées NCED identifiées chez *Arabidopsis* (Bray 2002). La Xanthoxin donne finalement un aldéhyde abscissique et enfin l'ABA sous l'action de 2 enzymes produites par les gènes AtABA2 et AAO3. (Fig B-15). Une liste des gènes connus chez *Arabidopsis* est fournit par Bray (2002).

Une liste des mutants identifiés et des enzymes touchés a été établie par Taylor *et al.* (2000) et Leung *et al.* (1998).Les auteurs décrivent les phénotypes des plants ainsi que l'impact en condition de stress hydrique. D'après la littérature citée par les auteurs, des études sur *Nicotiana sp*. mettent en évidence une augmentation de la concentration en ZEP des racines de plants exposés à la sécheresse. En ce qui concerne les NCED, les racines de plants de tomate et de haricot en condition de stress hydrique présentent de fortes augmentations de concentrations suivies par une teneur en ABA augmentée. Chez *Arabidopsis*, les gènes AtNCED3, AAO3, AtABA3 et AtZEP sont induit lors de la déshydratation mais le gène AtABA2 n'est pas affecté et son expression est constitutive (Taylor *et al.* 2000).

La teneur en ABA, comme celle de tout composé, n'est pas déterminée uniquement par les gènes de sa biosynthèse mais par l'équilibre entre les voies de biosynthèse et de dégradation. Nambara *et al.* (2005) mettent l'accent sur ce point en rapportant le rôle de l'ABA 8'-hydroxylase codée par les gènes CYP707A régulant une des 3 voies de dégradation de l'ABA et induit lors de stress osmotiques ou de la réhydratation (Fig. B-16).

L'ABA est synthétisée dans les racines et dans les feuilles. Sauter *et al.* (2001) reprennent les lieux de synthèse, de circulation et les effets de l'acide abscissique endogène sur des plantes en situation de stress hydrique (Fig. B-17). Lors d'un stress hydrique, l'ABA s'accumule en grandes quantités dans les racines. Il gagne le xylème et circule ensuite jusqu'à l'appareil foliaire où il modifie les flux d'ions K^+ et Cl^- de la vacuole vers le cytoplasme des cellules de garde (Leung and Giraudat 1998). Les stomates se ferment et la conductance stomatique chute. L'acide abscissique peut diffuser hors du système conducteur en empruntant les voies symplasmiques ou apoplasmiques mais la présence d'obstacles comme les bandes de Caspari vont freiner la diffusion. La synthèse d'ABA dans les feuilles apparaît lorsque le potentiel hydrique de la feuille n'assure plus la turgescence mais l'ABA présent dans la sève peut diffuser jusqu'aux stomates avant d'atteindre ce point critique. Une expérience sur le passage de l'ABA dans les tiges tend à soutenir l'hypothèse selon laquelle le transport de l'ABA serait modifié en fonction du pH du milieu cellulaire qui varie si la plante est stressée ou non (Sauter *et al.* 2001; Wilkinson and Davies 2002). Un pH de 6.3, rencontré en conditions normales, favoriserait le passage dans les cellules et donc le captage de l'ABA par les tissus ce qui baisserait la concentration dans le xylème. A l'opposé, un pH de 7.2

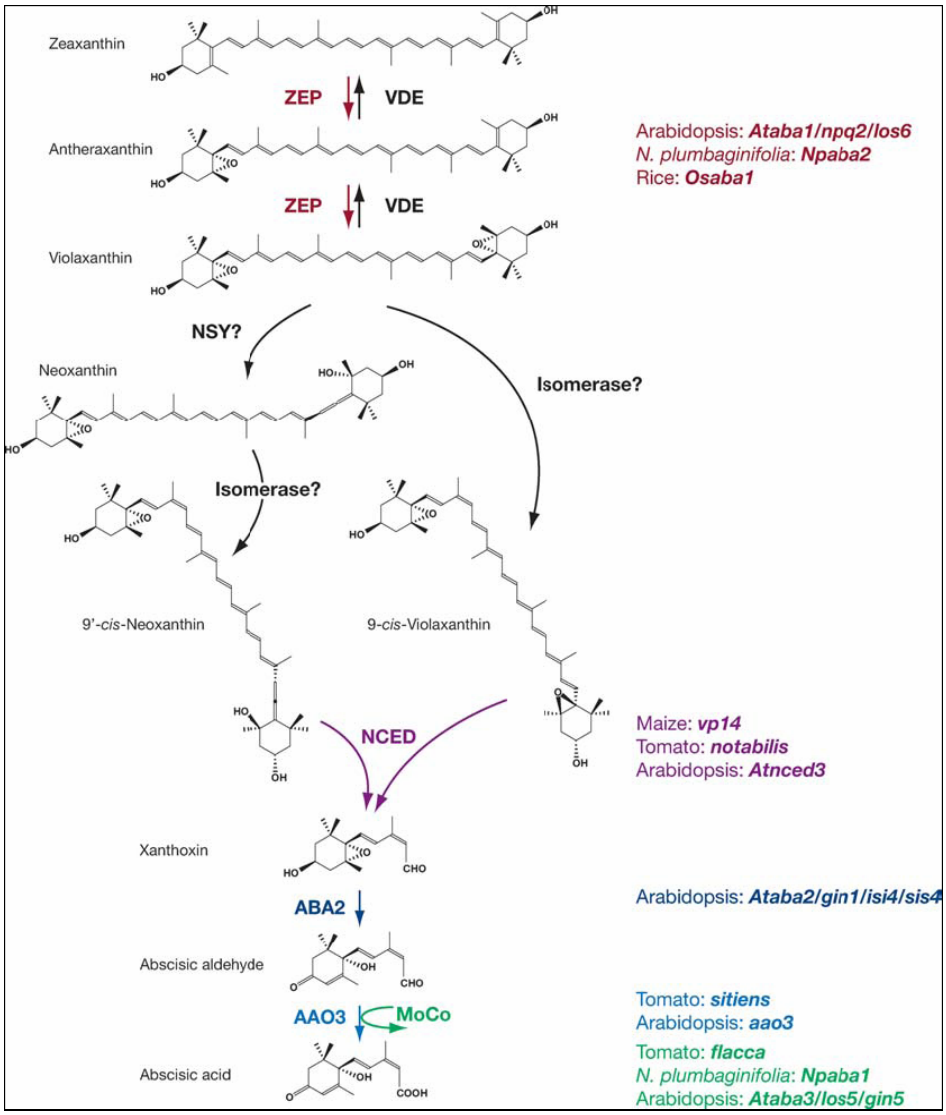


Figure B-15 : Voie de biosynthèse de l'ABA. Les enzymes déficientes pour des mutants de différentes espèces sont indiquées. Abréviations : ZEP : zeaxanthin epoxidase ; VDE : violaxanthin de-epoxidase ; NSY : neoxanthin synthase ; NCED : 9-cis-epoxycarotenoid dioxygenases ; ABA2 : alcool déshydrogénase de chaînes courtes ; AAO3 : abscissic aldehyde oxidase ; MoCo : Molybdenum cofactor sulfurase. (Nambara *et al.*, 2005)

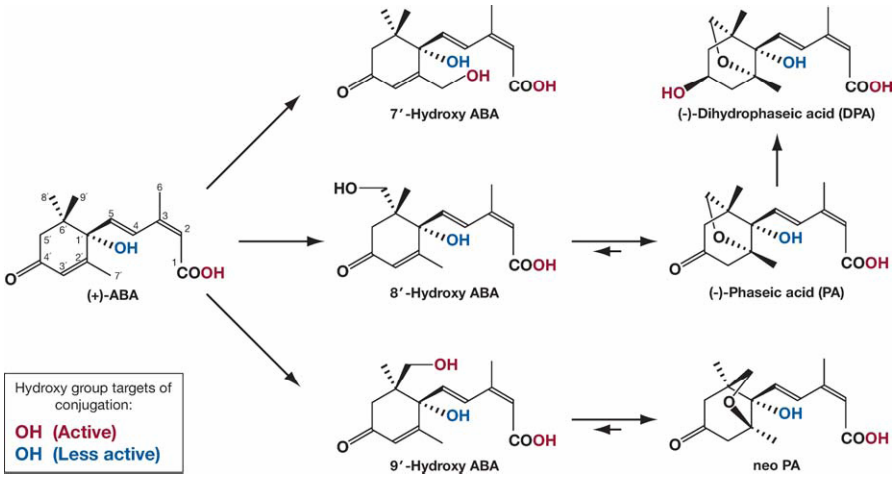


Figure B-16 : Voies de dégradation de l'ABA. Trois voies différentes d'hydroxylation sont présentées. La voie de la 8'-Hydroxylation semble être la voie prédominante de dégradation. (Nambara *et al.*, 2005)

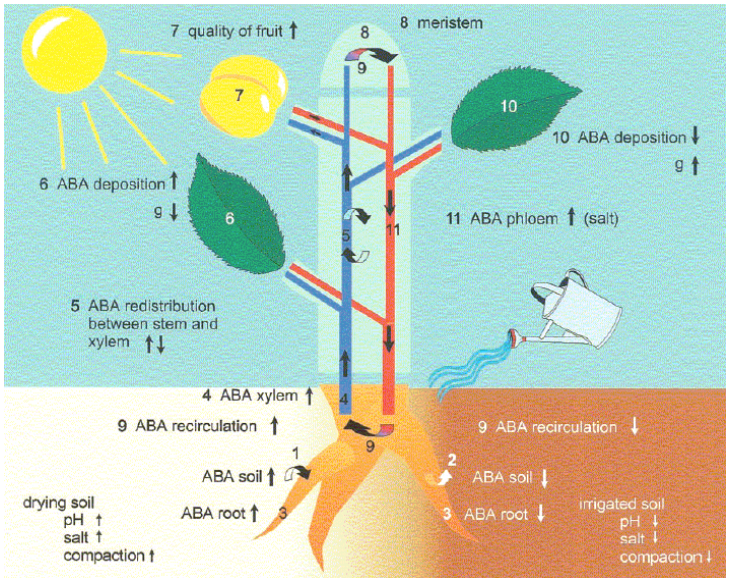


Figure B-17 : Effet de l'état d'hydratation sur l'ABA dans la plante. (Sauter *et al.*, 2001)
1-ABA dans la rhizosphère, 2-Perte d'ABA dans le milieu environnant, 3-Accumulation d'ABA dans les racines par augmentation de la biosynthèse, 4-Transport radial de l'ABA dans les racines, 5-Redistribution de l'ABA entre les tissus de la tige et les vaisseaux du xylème, 6-10-ABA dans les feuilles, 7-ABA dans les fruits, 8-ABA dans les méristèmes, 9-Recirculation de l'ABA.

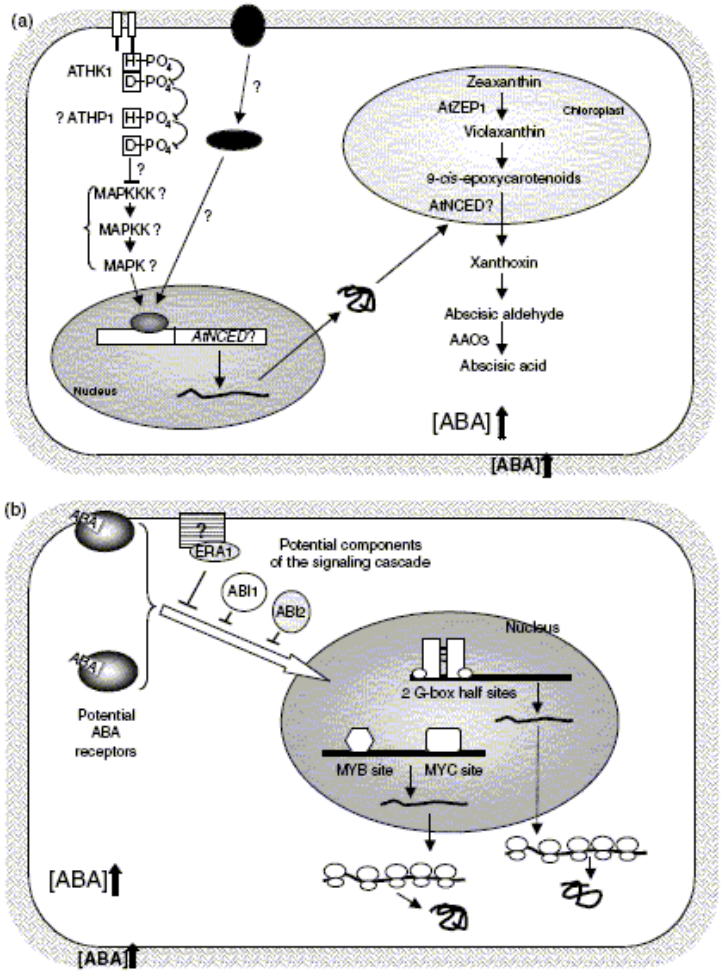


Figure B-18 : Modèle de réponse d'une cellule d'*A. thaliana* soumise à un stress hydrique. a) mécanisme possible de perception du stress avec accumulation d'ABA. b) proposition d'une voie de transmission du signal avec reconnaissance de l'ABA et régulation de l'expression de gènes. (Bray *et al.*, 2002)

correspondant à une plante stressée aurait un captage moindre, une concentration supérieure dans le xylème et l'apoplasme ce qui favoriserait l'accès de l'ABA aux stomates. En diffusant, l'ABA peut circuler également dans le phloème qui va transporter l'ABA synthétisé au niveau des feuilles et des racines. En arrivant dans les racines, il peut re-circuler et représenter une part non négligeable du flux racines vers feuilles.

La perception du signal permettant l'initiation de la biosynthèse de l'ABA est mal connue mais certains mécanismes sont proposés au vu des mécanismes observés chez *E. coli* ou *S. cerevisiae* (Bray 2002). Un ARNm d'*Arabidopsis* codant pour une protéine ATHK1 similaire aux protéines du complexe de perception des stress osmotiques des bactéries ou levures pourrait faire partie d'une cascade de perception du signal. Le senseur activerait cette ATHK1 qui pourrait interagir avec une autre protéine ATHP1 puis une cascade de type MAP kinases par l'intermédiaire de phosphorylations. Il y aurait alors activation de la transcription des ARNm codant pour les enzymes de la voie de biosynthèse de l'ABA telle que les AtNCED déjà citées.

Les mécanismes de régulation des gènes par l'ABA restent également flous. Les récepteurs à cette hormone ne sont pas encore identifiés mais les mutants Abi1 et Abi2 évoqués plus haut supposent l'existence d'une cascade de transmission du signal reliant des récepteurs à la régulation de la transcription par l'intermédiaire d'Era1 (farnésyl-transférase), Abi1 et Abi2 (Fig. B-18). L'hypothèse de l'existence de cette cascade de signalisation semble corroborée par des résultats obtenus en exposant des graines d'*Arabidopsis* à la protéine HrpN d'un pathogène connue pour stimuler la croissance et la résistance aux pathogènes. Sur le génotype sauvage, cette exposition stimule la résistance à la sécheresse de plants en situation de stress hydrique en activant la synthèse d'ABA qui provoque une fermeture des stomates. Par contre, les mutants Abi1-1 et Abi2-1 ne présentent pas ces caractéristiques. L'expression des gènes ABI1 et ABI2 répond de la même façon que le génotype sauvage mais l'expression du gène ABH1 n'est pas stimulée. ABH1 interviendrait après ABI1 et ABI2 dans la cascade de régulation (Dong *et al.* 2005).

Les motifs cibles de la régulation de la transcription par l'ABA sont par contre mieux identifiés. Il existe des éléments promoteurs de régulation en CIS ; nous citerons les ABRE, les motifs MYC et MYB. Les ABRE (ABA responsive element) sont des séquences particulières –« ACGT containing G-box » (Leung and Giraudat 1998)- sur lesquelles se fixent les facteurs de transcription bZIP qui sont des protéines constituées d'un domaine de fixation à l'ADN suivit d'une « leucine zipper ». Les motifs MYC -ACACATGT- et MYB –YAAC(G/T)G- se retrouvent chez les gènes induits lors de stress sévères (Bray 2002). La

recherche des motifs possibles se poursuit et de nouveaux motifs sont identifiés, ils sont regroupés sous le nom de ABFs (ABRE binding factors) (Choi *et al.* 2000).
Les gènes régulés spécifiquement par l'ABA sont difficilement identifiables car confondus avec des voies de régulations indépendantes de l'ABA ou déclenchées par d'autres hormones comme l'éthylène lors du stress. Néanmoins, une approche utilisant des mutants déficients pour la voie de l'ABA en condition de stress a permis à Chen *et al.* (2000) d'identifier par électrophorèse bidimensionnelle des polypeptides différentiellement exprimés dans les racines d'*Arabidopsis* mais le séquençage des polypeptides et leur identification n'a pas été réalisé (Chen *et al.* 2005).

Le nom de l'hormone vient de sa découverte lors de recherches sur l'abscission des fruits du cotonnier par Addicott *et al.* (1963). Mais l'abscission est maintenant attribuée à l'action de l'éthylène. Lors d'une étude sur l'exposition simultanée à un stress hydrique (sécheresse) et à l'ABA de deux génotypes contrastés de peuplier, Chen *et al.* (2002) ont observé une abscission des feuilles marquée chez le génotype sensible contrairement au génotype tolérant. Le suivi des concentrations en ABA, en éthylène et en polyamines –une autre classe d'hormones– révèle que la concentration foliaire en ABA enregistrée chez le génotype sensible est accompagnée d'une augmentation de la teneur en éthylène et d'une baisse en polyamines. Ces voies sont bien corrélées par une compétition notamment pour un précurseur commun (S-adenosylmethionine), mais c'est l'éthylène qui est à l'origine du phénomène d'abscission (Chen *et al.* 2002).

B.2.2.3.3.2 Ethylène

Cette hormone a tout d'abord une action sur les graines en étant capable de **lever la dormance**. Elle favorise également la **maturation des fruits** et déclenche **l'abscission des feuilles** qui est contrôlée par une balance Auxine/Ethylène (Heller *et al.* 1993). La migration de l'auxine dans les tissus serait ralentie par l'éthylène. Son action sur la croissance se traduit par une inhibition de l'**élongation racinaire** au profit de la croissance radiale (Crozier *et al.* 2000). Sur les racines, l'application d'éthylène induit la formation d'un **chevelu racinaire**. Des expériences de transformation génétique permettant de diminuer la teneur en éthylène des plants ont permis de mettre en évidence d'autres propriétés de cette hormone. La diminution de la teneur a pu être réalisée en dégradant le précurseur direct à l'aide de l'ACC désaminase. Ainsi, une tolérance au nickel a été observée chez des plants d'*Arabidopsis* transformés par ce gène ce qui souligne sont implication dans la réponse aux stress (Stearns *et al.* 2005). La transformation de bactéries n'exprimant pas initialement ce gène a permis de favoriser la

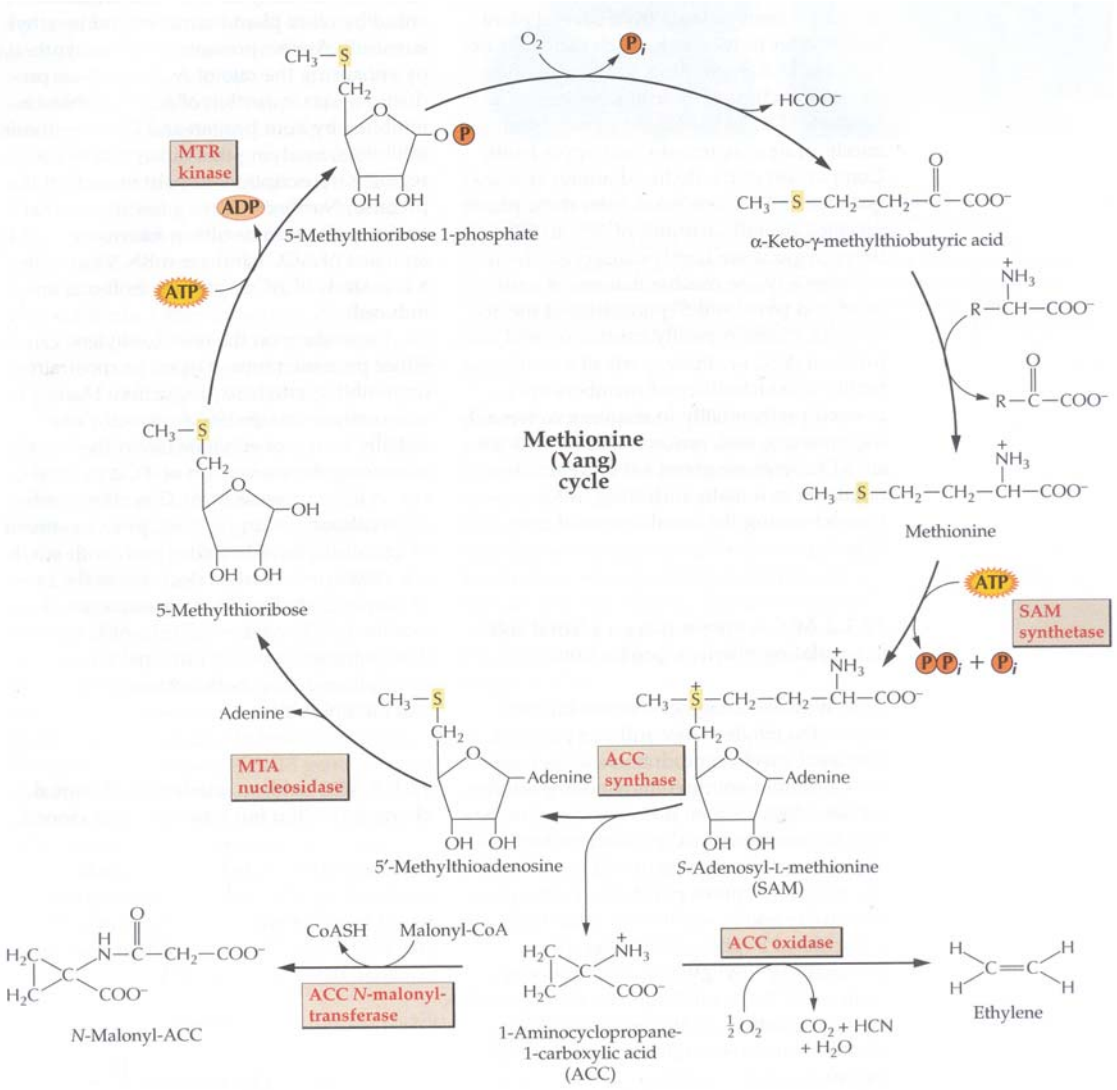


Figure B-19 : Cycle de la méthionine et biosynthèse de l'éthylène. L'éthylène est synthétisé à partir de la méthionine par l'intermédiaire de la SAM et de l'ACC. Les enzymes qui catalysent ces trois étapes sont l'ATP methionine S-adenosyl-transférase (SAM synthase), la S-adenosyl-L-methyl-thioadenosine-lyase (ACC synthase) et l'ACC oxydase. D'après Crozier *et al.* (2000).

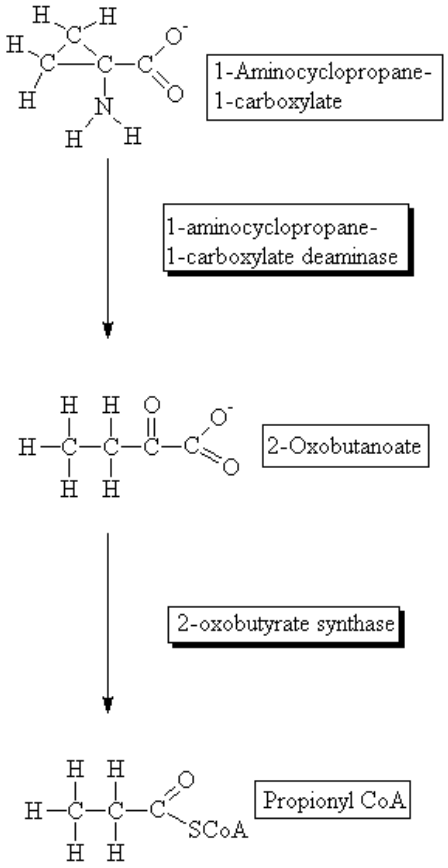


Figure B-20 : Action de l'enzyme 1-Aminocyclopropane-1-carboxylate désaminase (ACC deaminase) sur le précurseur (ACC) direct de l'éthylène.

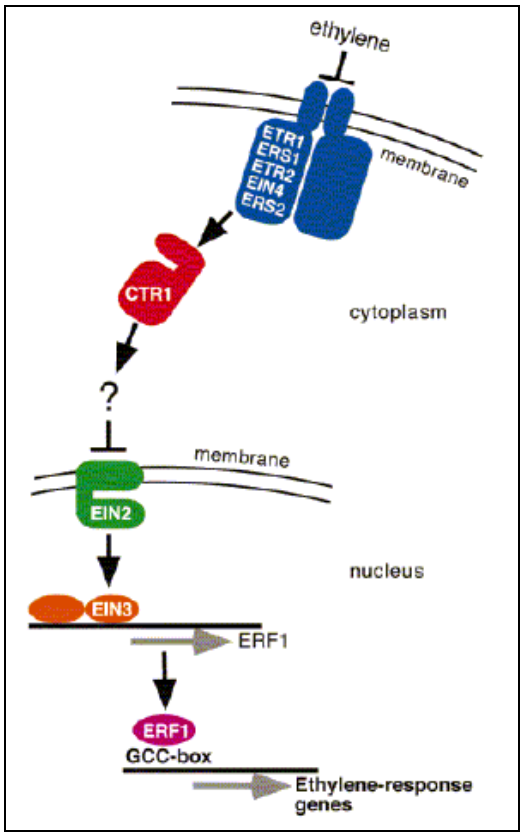


Figure B-21 : Schéma de la voie de transduction du signal. Cinq récepteurs à l'éthylène (ETR1, ERS1, ETR2, EIN4 et ERS2) transmettent le signal éthylène à la protéine kinase CTR1 qui présente des similarités avec la famille des MAPKKKs. EIN2 agit en tant que domaine membranaire. La localisation spécifique des récepteurs à l'éthylène et de EIN2 sont inconnus. Le facteur de transcription EIN3 favorise la transcription de ERF1 qui code pour un membre de la famille des facteurs de transcription EREBP. ERF1 va se lier aux promoteurs à boîte GCC pour activer la transcription des gènes de réponse à l'éthylène. (Chang *et al.*, 2001)

nodulation de plants de pois et de trèfle par respectivement *Rhizobium leguminosarum* et *Sinorhizobium meliloti* (Ma *et al.* 2003; Ma *et al.* 2004). Le rôle de l'éthylène dans la réponse au stress n'est pas clairement mis en évidence car les effets de l'ABA et de l'éthylène sont étroitement liés. Les expériences de Sharp et LeNoble (2002) soulignent ce phénomène et suggèrent une inhibition de la croissance des pousses et des racines de maïs en condition de stress. Elle aurait pour origine une concentration élevée en éthylène induite par des promoteurs de croissance de ces jeunes plants et une concentration relative en ABA faible.

L'éthylène est un gaz, sa formule à 2 carbones est très simple C_2H_4. Il est synthétisé à partir de la méthionine (Fig. B-19) qui sous l'action de la SAM synthase donne de la **S-Adenosyl-L-méthionine** ou **SAM.** Ce composé intervient dans la synthèse des polyamines mais permet aussi d'obtenir de l'**ACC** (acide 1-Aminopcyclopropane-1-carboxylique) par une réaction catalysée par l'**ACC synthase** qui fait partie du cycle de la méthionine, appelé également cycle de Yang. L'ACC produit peut alors donner différents composés dont l'éthylène sous l'action d'une **ACC oxydase,** mais aussi du 2-Oxobutanoate ou α-ketobutyrate puis du Propionyl CoA dans une réaction catalysée par l'**ACC désaminase** (Fig. B-20), ou encore du N-Malonyl-ACC avec l'aide d'une ACC N-malonyl-transférase. La régulation de la production de l'éthylène est principalement dépendante de l'ACC synthase. Des régulations peuvent cependant intervenir en aval par compétition sur l'ACC qui est le précurseur direct de l'éthylène en donnant l'α-ketobutyrate et l'ammonium par l'ACC désaminase ou pour la synthèse des polyamines. Cette voie est bien identifiée chez les bactéries du sol capables d'exprimer une ACC désaminase pour contrôler la production d'éthylène et favoriser leur interaction avec la plante (Ma *et al.* 2003; Ma *et al.* 2004). Cet enzyme est souvent détecté lors de l'étude des promoteurs de croissance des rhizobactéries (PGPR) (Penrose and Glick 2003). De nombreuses expériences de transformations génétiques de plantes exposées à des stress et exprimant une ACC désaminase d'origine bactérienne démontrent la diminution de la production d'éthylène (Hontzeas *et al.* 2004; Stearns *et al.* 2005). Si la plupart des hormones nécessitent une voie de dégradation, l'éthylène peut être éliminé simplement sous forme de gaz. Les gènes codant pour les enzymes de la voie de biosynthèse sont bien décrits dans la littérature et Johnson *et al.* (1998) fournissent une liste des gènes identifiés à ce jour.

L'éthylène agit sur des récepteurs membranaires dérivés de la famille Histidine-Protéine kinases qui par phosphorylation transmettent le signal à un aspartate (Chang and Stadler 2001) (Fig. B-21). Parmi les récepteurs d'*Arabidopsis* classés en sous familles, on citera les

récepteurs ETR1 (ethylene resistant), ERS1, ETR2, EIN4 et ERS2 (Chang and Bleecker 2004). Le mutant etr1 ne présente pas un certain nombre de réponses à l'éthylène comme l'activation de la germination, la stimulation de l'activité peroxydase ou l'accélération de la sénescence (Kieber 1997). Chang *et al.* (2004) rapportent que les études récentes indiquent que des facteurs de transcription EIN3 et EIL1 sont nécessaires à la transmission du signal capté par les récepteurs dans la phase lente de réponse à l'éthylène (par opposition aux réponses initiées dans les premières minutes suivant exposition au gaz). Une protéine membranaire EIN2 découverte récemment agirait en amont de EIN3. EIN2 serait un élément clé, les mutations de ce gène bloquent totalement les réponses à l'éthylène. Les mutants etr1 ou ein2 ont une teneur basale en éthylène 5 à 10 fois supérieure ce qui suggère qu'une voie de régulation négative de la biosynthèse serait déficiente. La liste des gènes identifiés et leur position supposée dans la cascade de transduction du signal est représentée dans la figure B-22. Les mutants se répartissent en deux classes, la première regroupe ceux qui surproduisent de l'éthylène (Eto) et la seconde regroupe les mutants conservant leur réponse en présence d'inhibiteurs de la synthèse et de la fixation de l'éthylène ce qui suggère plutôt une perturbation de la transduction du signal (Kieber 1997). Un élément clé de cette transduction est la protéine kinase CTR1 (Johnson and Ecker 1998).

A l'issue de cette cascade de transmission du signal, les gènes régulés possèdent des séquences particulières identifiées en étudiant les gènes PR (Pathogenesis-related). Ces séquences ERE possèdent un motif AGCCGCC ou « GCC box » nécessaire et suffisant à la régulation des gènes porteurs. Puis un domaine de fixation à l'ADN appelé domaine AP2 a été identifié. Les protéines capables de reconnaître ces séquences sont appelées EREBPs (ERE-binding proteins) et pourraient être les cibles de régulations initiales de l'éthylène.

Au niveau racinaire le gène EIR1, présent chez des mutants insensibles à l'éthylène, détermine l'agravitropisme qui est un caractère rencontré chez des mutants pour la sensibilité à l'Auxine. Cela suggère un rôle carrefour de EIR1 entre les voies de l'éthylène et de l'auxine.

B.2.2.3.3.3 Polyamines (PA)

Cette catégorie d'hormones est plus abondante dans la plante que les gibbérellines ou les cytokinines. Ce sont des polycations de faible masse moléculaire. Les membres les plus couramment rencontrés sont la putrescine, la spermidine et la spermine (Fig. B-23). La putrescine est synthétisée à partir de l'arginine par deux voies, l'une faisant intervenir la L-Ornithine et l'Ornithine décarboxylase, et l'autre catalysée par l'Arginine décarboxylase passe

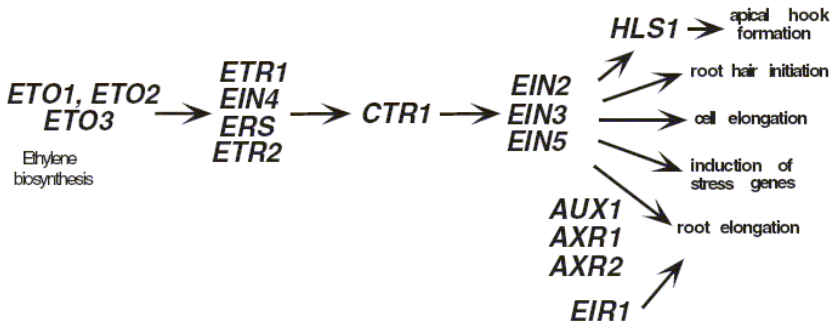


Figure B-22 : Proposition de l'organisation des gènes impliqués dans la production et la perception de l'éthylène. A droite, différentes réponses possible à l'éthylène sont indiquées.

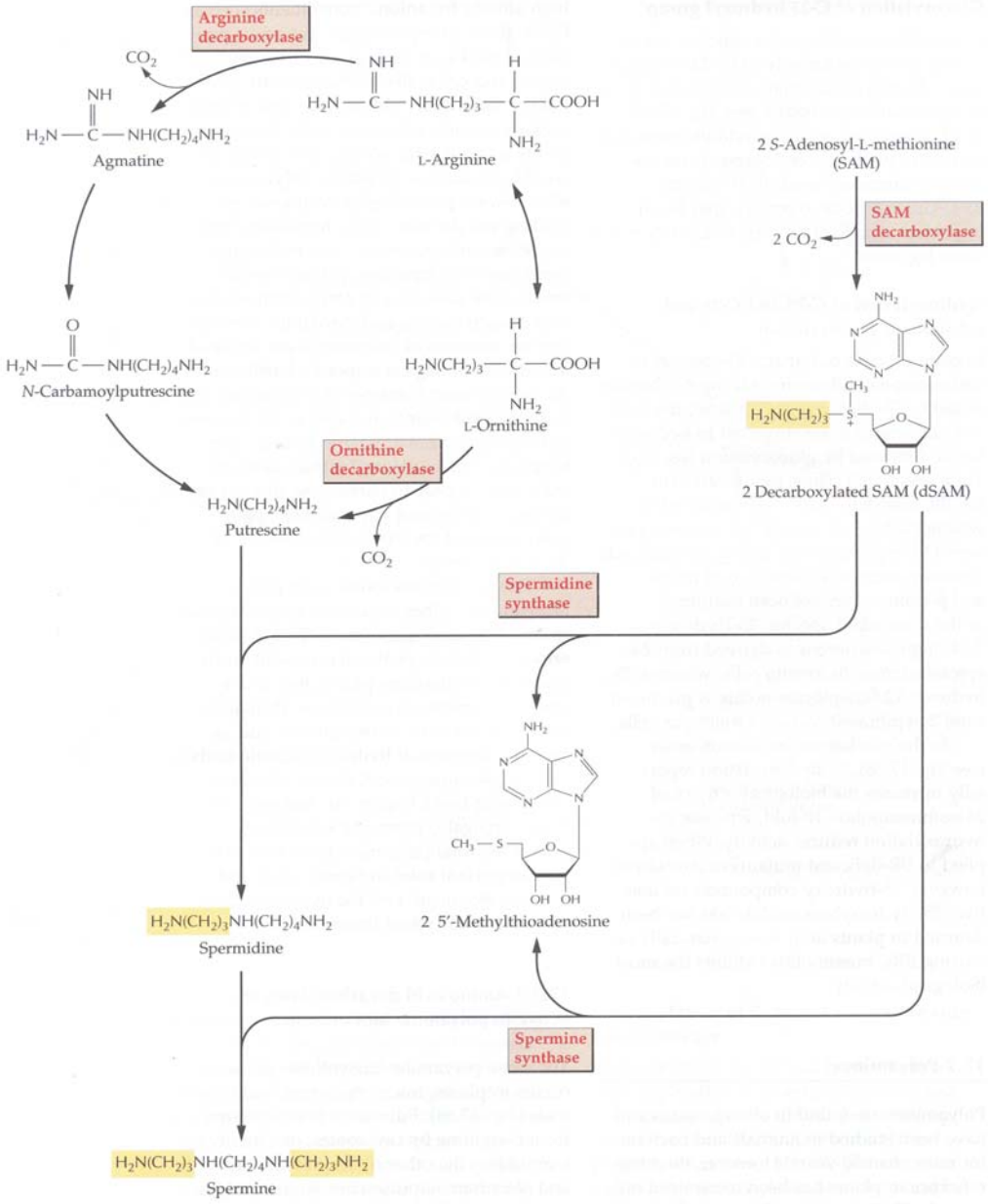


Figure B-23 : Biosynthèse de la putrescine, spermidine et spermine. Chez la plupart des plantes, l'arginine décarboxylase est plus active que l'ornithine décarboxylase. (Crozier *et al.*, 2000)

par l'Agmatine et le N-Carbamoylputrescine. La Spermidine et la spermine dérivent de la SAM par action de la SAM décarboxylase puis de la Spermidine synthase ou de la Spermine synthase (Martin-Tanguy 2001). Leur rôle exact n'est pas complètement connu. Cependant, l'utilisation de mutants des gènes de la voie de biosynthèse a permis d'obtenir de nombreux phénotypes tels que : croissance ralentie, entre-nœuds courts, branchaison des tiges ou petites feuilles. Les mutants présentaient également une teneur en éthylène fortement augmentée ce qui confirme la compétition pour un précurseur SAM commun (Walden *et al.* 1997). Il a été montré que lors de stress osmotiques, les plants tolérants ont des concentrations en polyamines supérieures aux plants sensibles (Tiburcio *et al.* 1994). Une des hypothèses rapportées par Gallardo *et al.* (2003) quant à leur rôle serait que leur biosynthèse pourrait être un mécanisme de détoxification de l'ammonium en conditions de stress. La nature polycationique de ces molécules leurs confère un fort pouvoir de liaison aux molécules chargées négativement au pH physiologique telles que les acides nucléiques, les protéines ou les membranes phospholipidiques. En condition de stress, ils interagissent avec les membranes et les stabilisent (Roberts *et al.* 1986) et pourraient aussi agir comme destructeur des radicaux libres. L'exposition de plants de pois chiches à des stress osmotiques et à des stress froid a permis à Nayyar et Chander (2004) d'observer une augmentation de la teneur en PA lors du stress. Par contre la nature des PA n'était pas la même pour les deux stress. En stress froid il s'agissait de putrescine tandis qu'en stress osmotique cela concernait la spermidine, ce qui confirme la littérature. Les auteurs ont également observé que l'application exogène de PA -plus particulièrement la putrescine et la spermidine, dans une moindre mesure la spermine- favorisait la croissance des plants stressés. L'application d'inhibiteurs des enzymes de biosynthèse de ces PA s'est traduit par des dommages marqués lors du stress (Nayyar and Chander 2004). Une revue de Couée *et al.* (2004) met l'accent sur l'implication des PA dans le développement racinaire. Les auteurs nous renseignent sur la nature des PA détectées dans les racines. En plus de la putrescine, de la spermidine et de l'Agmatine, il existe des formes conjuguées de PA comme la Caffeoylputrescine et la Feruloylputrescine, et des formes liées insolubles. Les PA jouent également un rôle dans le développement racinaire primaire, latéral et adventif. Couée *et al.* (2004) rapportent des études montrant une corrélation entre la quantité de PA et la formation des racines adventives, la croissance des racines primaires, la prolifération et le cycle cellulaire.

Après avoir revu en détail les hormones les plus impliquées lors d'un stress hydrique, nous allons décrire succinctement les trois hormones végétales majeures impliquées dans un

grand nombre de régulations physiologiques. Elles ne sont pas en soit caractéristiques de la réponse au stress mais interviendront dans la balance hormonale de la plante.

B.2.2.3.3.4 Les auxines

Synthétisées dans les apex des tiges (ou dans les tiges chez certains ligneux comme le Frêne), elles migrent dans la plante jusque dans les racines. Leur action principale est de favoriser la **croissance cellulaire** en stimulant l'élongation des cellules. Cependant, au niveau racinaire l'action est opposée, avec une **inhibition de la croissance**. L'action mitogène est particulièrement marquée dans les zones génératrices libéro-ligneuses c'est-à-dire les cellules cambiales. Elles ont également un **pouvoir rhizogène** fort dans le cas d'une action limitée dans le temps car l'effet inhibiteur sur la croissance racinaire gène le développement des jeunes ébauches. Elles interviennent également sur le développement des bougeons et sont responsables du phénomène de dominance apicale.

L'auxine joue également un rôle de stimulation de la production d'éthylène qui lui-même réduit la migration de l'auxine. L'importance de cette balance Auxine/Ethylène se retrouve au niveau de la zone d'abscission des feuilles des espèces caduques. La chute des feuilles est une des stratégies de réponse au déficit hydrique prononcé. La réduction de la surface foliaire permet de limiter les pertes par évapotranspiration et de réduire les risques de cavitation. La biosynthèse de l'AIA (acide indole-3-acétique) s'effectue à partir du Tryptophane et plusieurs voies sont possibles. La principale passe par la synthèse d'acide indole-pyruvique et d'indole acétaldéhyde. Dans une revue, Bartel (1997) reprend en détail la biosynthèse de l'auxine.

B.2.2.3.3.5 Les Gibbérellines

L'effet le plus remarquable de ces hormones est le gigantisme induit lors de son application. Les gibbérellines (GA) provoquent l'**élongation des entre-nœuds** du à une élongation et une **prolifération des cellules** de la tige. A forte concentration on peut également observer une augmentation de la **taille** des feuilles. Elles induisent la formation des cônes chez les conifères. Elles ont aussi une action opposée à celle de l'ABA sur la dormance des graines et des bourgeons. Les GA peuvent donc être qualifiées de **régulateur de croissance**.

Comme l'ABA, ce sont des composés terpéniques mais à 19 ou 20 carbones. La synthèse est très active dans les apex des jeunes pousses et les racines. Elle s'effectue à partir d'Isopentenyl-pyrophosphate (IPP) qui par polymérisation donne finalement du Geranylgeranyl-pyrophosphate (GGPP) (Fig. B-24). Le GGPP se cyclise en Kaurène, sous

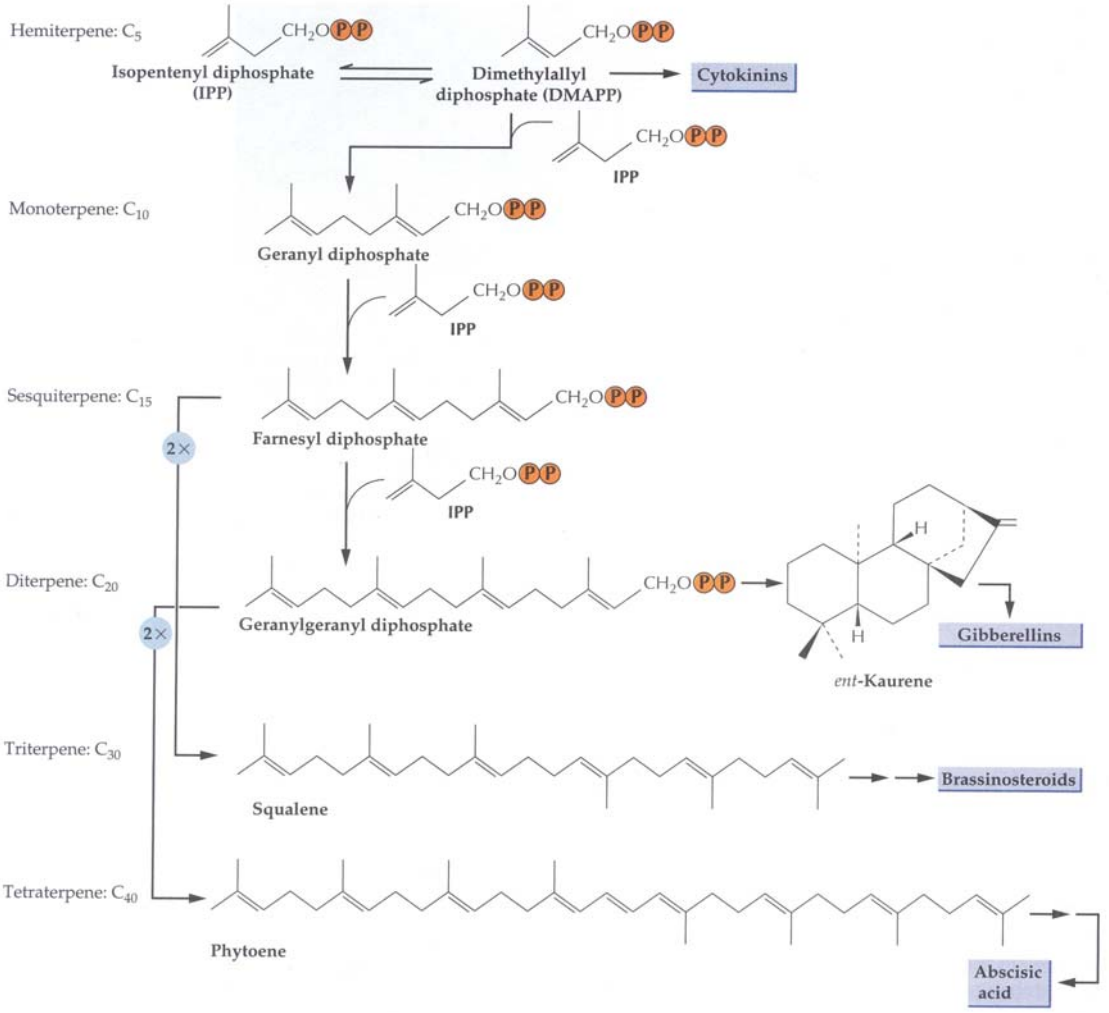


Figure B-24 : Voie de biosynthèse des terpènes montrant l'origine terpénique de la biosynthèse des Gibbérellines, Cytokinines et de l'ABA. (Crozier *et al.*, 2000)

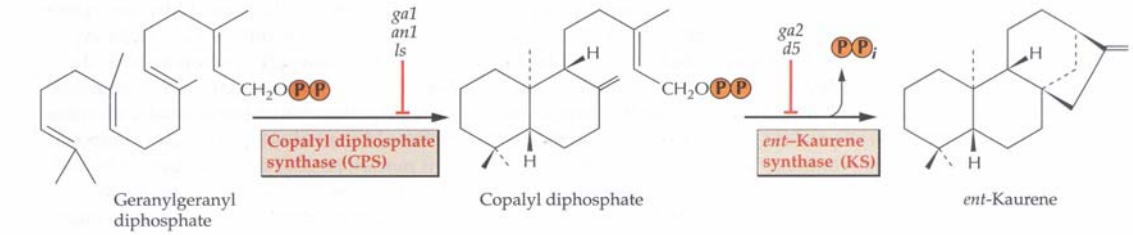


Figure B-25 : Biosynthèse du Kaurene, précurseur de synthèse des Gibbérellines, à partir du GGPP. Les mutants d'Arabidopsis et du pois pour les enzymes catalysant les réactions sont indiqués. (Crozier *et al.*, 2000)

l'action de la Copalyl diphosphate synthase (CPS) et de l'ent-Kaurène synthase (KS) (Fig. B-25). Le Kaurène est le point de départ de toutes les gibbérellines. Pour une revue plus précise des voies de biosynthèse, des récepteurs et mutants identifiés et des effets physiologique, se référer à la publication de Richards *et al.* (2001).

B.2.2.3.3.6 Cytokinines

Les cytokinines regroupent un ensemble de molécules qui provoquent entre autres réactions, la division des cellules et la néoformation des bourgeons conjointement à l'auxine. Le ratio Cytokinine/Auxine va déterminer le développement des pousses –ratio élevé- ou le développement racinaire –auxine seule-. Elles se retrouvent à l'état libre dans les tissus ou fréquemment associées à du ribose pour former des ribosides. Leur lieu de synthèse de prédilection semble être la région apicale des racines. Aloni *et al.* (2004) ont pu montrer par immunolocalisation et gène rapporteur GUS cette localisation apicale. De plus, ils ont mis en évidence que le gravitropisme des racines serait sous le contrôle des auxines et des cytokinines. En effet, la distribution asymétrique de la zéatine dans la zone apicale correspond à la direction de croissance adoptée par la racine ce qui s'explique par l'action de régulation négative sur la croissance racinaire par les cytokinines (Aloni *et al.* 2004). En agissant sur la croissance racinaire, les cytokinines peuvent jouer un rôle dans le développement des racines en réponse au stress hydrique. La kinétine (Amasino 2005) et la zéatine sont des cytokinines naturelles. Biochimiquement, elles dérivent de l'adénine avec une base isopentenique attachée à l'azote N_6 du groupement aminé.

D'autres molécules que nous n'aborderons pas ici ont également un rôle de régulation dans la plante comme les brassinostéroïdes, l'acide jasmonique et l'acide salicyliques (Yang *et al.* 2004).

B.2.2.4 Taille du génome et adaptation au stress hydrique

Avec 2n = 2x = 24 chromosomes, le caryotype est remarquablement uniforme à ceux des autres pins du genre Pinus, tant sur le plan de la morphologie que du nombre (Murray 1998). Les chromosomes sont de grande taille mais en nombre limité et représentent env 50pg/2C pour *Pinus pinaster* (Murray 1998; Chagne 2004). La totalité du génome n'est pas codante et la proportion d'ADN en simple copie représente uniquement 14-24%. Une étude réalisée sur *Pinus strobus* par Elsik *et al.* (2000), divise son génome en fonction du niveau de redondance. Les séquences simple copie constituent 14% soit plus de 3000 Mb, ce qui en

taille dépasse déjà le génome des céréales. Les régions répétées sont classées en zones faiblement (20%), moyennement (45%) et hautement (14%) répétées. Une très grande partie du génome est donc constituée de séquences répétées comme par exemple des retro-transposons. Le rôle de cet excès d'ADN répété et un éventuel avantage sélectif pour l'espèce n'est pas encore très bien expliqué. Quelques études ont mis en avant une relation entre la taille des génomes et les caractéristiques adaptatives des plantes et en particulier les gymnospermes à leur habitat (précipitation et température) (Wakamiya *et al.* 1993). La théorie se base sur une relation entre la taille du génome, celle des cellules et celle de la paroi cellulaire qui impliquerait au niveau des cellules des trachéides une différence de conductivité et caractéristique hydraulique. La taille des génomes des pins semble en effet être corrélée à des paramètres physiologiques comme le potentiel hydrique minimum avant flétrissement (Wakamiya *et al.* 1996) qui est directement relié à la capacité de résistance à des potentiels osmotiques du sol bas en cas de sécheresse. Cependant, si une étude à plus grande échelle, portant sur 117 gymnospermes (Murray 1998) confirme ces observations pour certaines espèces, dans d'autres cas l'hypothèse est rejetée. La taille du génome est variable au sein du genre Pinus mais selon l'espèce, des variations intra-spécifiques ou au contraire une taille stable comme chez *Pinus taeda* peuvent être observées. Aucune conclusion générale ne peut donc actuellement être avancée quant à l'implication de la taille du génome sur une aptitude à résister à la sécheresse.

Une étude exhaustive et systématique du génome tel que les programmes de recherche conduit sur *Arabidopsis thaliana* ou *Populus tremuloides* ne sont pas applicable à l'espèce étudiée. Si l'on regarde la taille du génome du genre Pinus (Fig. B-26) et en particulier de *Pinus pinaster*, associé à la forte proportion d'ADN répété, on se rend compte qu'il n'est actuellement pas envisageable d'effectuer un séquençage systématique du génome. Il est toutefois aisément possible d'accéder au génome exprimé dans une condition et pour un organe donné par une extraction des ARN transcrits.

B.2.2.5 Gènes et voies métaboliques régulées au cours du stress

Les techniques de biologie moléculaire et de biotechnologie ont permis ces deux dernières décennies d'identifier de nombreux gènes régulés lors du stress hydrique. Face au stress, nous avons vu que la plante répond en régulant sa physiologie par l'intermédiaire de nombreux mécanismes à médiation hormonale notamment. Au niveau cellulaire, les signaux de stress doivent en premier lieu être perçus par des récepteurs spécifiques. Ils vont permettre d'activer des cascades de transduction du signal qui entraîneront la transcription et

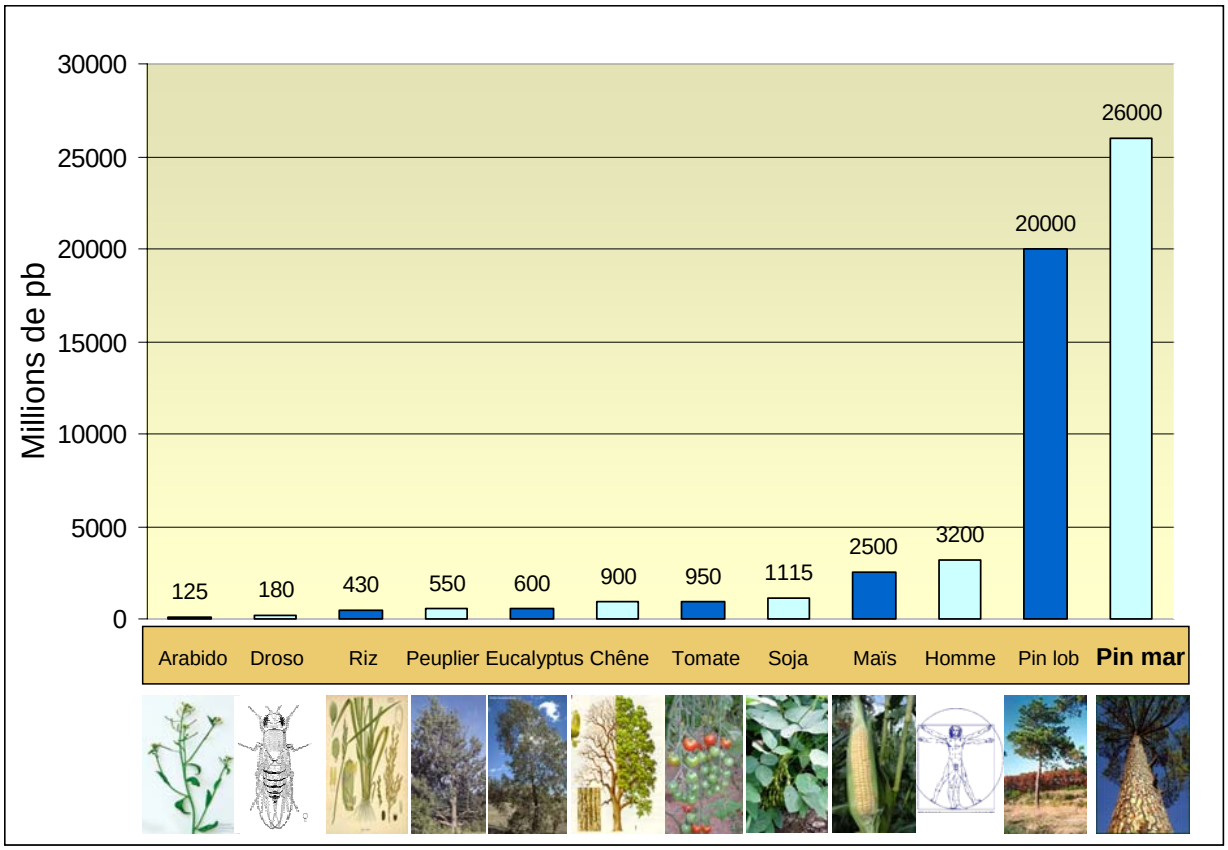


Figure B-26 : Schéma comparatif de la taille physique respective du génome diploïde de *A. thaliana* (125 Mb/C, 0.35pg/2C1, 5 chromosomes), le riz (430Mb/C, 1pg/2C1, 12 chromosomes), le chêne (900Mb/C, 1.85pg/2C, 12 chromosomes)[1], l'homme (3200Mb/C, 6.4pg/2C, 23 chromosomes), le pin maritime (25700Mb/C, 51pg/2C, 12 chromosomes)2.

1: Royal Botanic Gardens, Kew (www.rbgkew.org.uk).
2: Chagné *et al.* (2002)

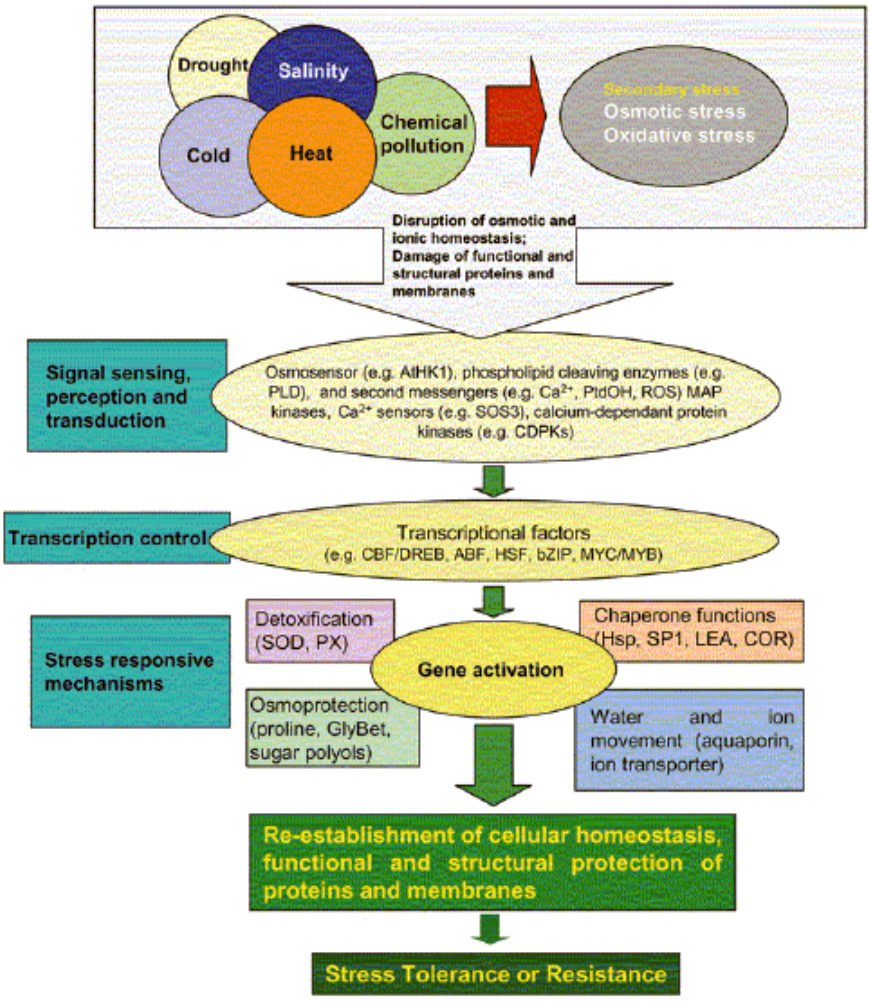


Figure B-27 : Processus de la perception du signal à la mise en place des mécanismes de résistance et/ou tolérance par la régulation de l'expression des gènes impliqués dans la réponse de la plante à des stress abiotiques. Les gènes régulés peuvent intervenir dans la détoxification, l'osmoprotection, les fonctions chaperonnes et les flux d'eau et d'ions. (Wang *et al.*, 2003)

l'expression de gènes spécifiques. La figure B-27 synthétise ce processus et regroupe les gènes régulés en 4 catégories : détoxification, osmoprotection, fonctions chaperonnes ou contrôle des flux ioniques et hydriques. Dans ce chapitre, nous allons reprendre les gènes identifiés et impliqués dans ces différentes étapes ou fonctions. Les gènes cités ne sont pas tous induits spécifiquement par un stress hydrique mais peuvent intervenir pour rétablir l'homéostasie cellulaire quelque soit la source de perturbation.

B.2.2.5.1 Méthodes de détection des gènes régulés

L'identification des gènes régulés lors d'un stress peut se faire à l'échelle d'un gène ou sur un ensemble de gènes.

La première approche implique généralement une connaissance à priori du gène étudié. C'est le cas lorsque la teneur d'un composé est régulée au cours des conditions expérimentales et que la voie de synthèse biochimique et les enzymes concernés sont connus. Il est alors possible de rechercher le(s) gène(s) codant(s) pour les enzymes et d'utiliser des mutants déficients pour le gène d'intérêt. En comparant les phénotypes du mutant et du sauvage, il sera possible de confirmer ou d'infirmer l'implication du gène dans le caractère étudié. Le recours à la transgénèse est une autre option. La sur ou sous-expression du transgène choisi va alors permettre de vérifier l'hypothèse de son rôle dans la réponse au stress appliqué.

Si aucune connaissance sur la fonction n'est disponible sur un gène, des études du niveau d'expression dans les conditions expérimentales sont possibles. La PCR quantitative ou le Northern par exemple, vont permettre de déterminer le niveau d'expression du gène dans chaque condition.

La deuxième approche est plus adaptée si aucun gène n'est privilégié. Il s'agit d'identifier des gènes dont l'expression est régulée et/ou corrélée avec les conditions expérimentales appliquées. De nombreuses techniques permettent de repérer les transcrits régulés au cours du stress au sein de l'ensemble des ARN totaux. Les principales sont les techniques SSH d'hybridations suppressives soustractives –qui permettent d'éliminer les transcrits communs aux différents échantillons-, les techniques d'identification de polymorphisme sur les ADNc (ADNc-AFLP) issus des différentes conditions, et les techniques d'hybridation où le niveau d'expression des gènes dans les différentes conditions est obtenu par hybridation des ADNc sur des séquences d'ADN choisies et fixées sur des membranes (Filtres) ou des lames de verre (puces à ADN). Chaque séquence correspond à un

gène donné. Il est possible par cette méthode d'étudier simultanément plusieurs milliers de gènes.

Les différentes approches adoptées en génomiques pour étudier la tolérance d'une plante au stress ont été reprises dans une revue de Cushman *et al.* (2000).

B.2.2.5.2 Détoxication

Les cellules stressées comme les cellules non stressées produisent des composés oxydants. Les plantes disposent de systèmes de défense pour les éliminer. La production et la dégradation sont en équilibre dans les cellules non stressées. Lorsqu'une plante est exposée à de basses températures, à la sécheresse ou à d'autres stress, la quantité de composés oxydants augmente. Ce sont des produits de réactions impliquant le peroxyde d'hydrogène (H_2O_2) et les ions superoxyde (O_2^-) qui sont regroupés sous le nom d'AOS (active oxygen species). Ces **AOS** ont un pourvoir oxydatif très fort capable de causer des dommages aux cellules qui en réponse synthétisent des composés spécifiques chargés d'éliminer ces AOS. Les principaux acteurs de cette détoxification sont les superoxyde dismutases (**SOD**), les ascorbate peroxydases (**APX**), les catalases (**CAT**), les monodehydroascorbate réductases (**MDAR**), les dehydroascorbate réductases (**DHAR**) et les glutathione réductases (**GR**) (Iba 2002). La SOD va permettre de transformer les radicaux superoxydes en O_2 et H_2O_2 qui sera pris en charge par l'APX et la CAT. Les MDAR, DHAR et GR vont permettre la régénération de l'acide ascorbique utilisé par l'APX.

La famille des SOD est divisée en trois grands groupes en fonction du cofacteur utilisé. On distingue les SOD à Manganèse (Mn SOD) –présentes dans les mitochondries et les peroxisomes-, les SOD à Fer (Fe SOD) –localisées dans les chloroplastes- et les SOD à Cuivre et Zinc (Cu-Zn SOD) qui sont divisées en deux groupes, l'un représenté dans le cytoplasme et les peroxysomes, et l'autre chloroplastique et extracellulaire (Alscher *et al.* 2002). L'augmentation de la SOD lors d'un stress va permettre d'éviter une dégradation des constituants cellulaires par les **ROS** (Reactive Oxygen Species), mais de nombreux gènes codent pour des SOD et tous les gènes ne seront pas régulés par le stress. Chez le maïs il a été identifié 4 Cu-Zn SODs cytosoliques (SOD-2, SOD-4, SOD-4A et SOD-5), 4 Mn SODs mitochondriales et une Cu-Zn SOD chloroplastique. L'étude des gènes SOD4 et SOD4A chez le maïs par Guan *et al.* (1998) montre que ces deux gènes sont similaires en tout point, excepté pour la région promotrice. Les expériences réalisées sur des embryons, des mutants et des feuilles de jeunes plants permettent d'avancer que les deux gènes répondent différemment à l'ABA aux différents stades de développement étudiés. Le gène SOD4 est surexprimé à 8h

et 12h dans les feuilles des plants exposés à un stress osmotique, et il s'agit d'une réponse sous le contrôle de l'ABA. L'expression du gène SOD4A sera dépendante du tissus et du stade de développement (Guan and Scandalios 1998).

Les autres acteurs de la détoxification sont également codés par des familles de gènes complexes régulées par des voies métaboliques qui peuvent être reliées ou non à une réponse au stress. Ainsi, chez *Arabidopsis*, Jespersen *et al.* (1997) distinguent 52 séquences réparties en 7 groupes codant pour des ascorbate peroxydases. Les Catalases sont des assemblages de monomères. Un nombre de gènes faible permet d'obtenir un nombre de combinaison important et l'identification de la régulation de l'activité enzymatique va dépendre de la combinaison des régulations individuelles des gènes codant chaque unité (Frugoli *et al.* 1996).

Les mécanismes de protection de la cellule aux ROS sont interconnectés et difficiles à comprendre. En dehors des composés déjà cités, ils font appel au cycle des xanthophyles qui fait intervenir des agents complexant, des pigments et des composés facilement oxydables comme la glutathionne (GSH) et des enzymes tels que la Glutathionne S-transférase (GST) (Light *et al.* 2005).

B.2.2.5.3 Fonctions chaperonne

Les stress abiotiques sont à l'origine de dysfonctionnements protéiques. La plante va activer de nombreux gènes permettant d'accumuler des protéines parmi lesquelles des **HSP** (Heat-shock protéins), des **LEA** (Late embryogenesis abundant proteins) et autres chaperonnes qui vont jouer un rôle de protection de la cellule. Elles sont responsables du repliement protéique, de l'assemblage, de la translocation et de la dégradation dans de nombreux processus cellulaires. Elles vont stabiliser les protéines et les membranes et permettre de retrouver les conformations protéiques actives en condition de stress (Wang *et al.* 2003).

B.2.2.5.3.1 Les Hsp (Heat-shock proteins)

Les Hsp sont divisées en cinq familles en fonction de la taille des protéines produites : les Hsp70 (DnaK), les Hsp60, les Hsp90, les Hsp100 (Clp) et les sHsp (small HSP) dont la taille varie de 12 à 40 kDa. Si elles ont été mises en évidence lors d'expositions à des températures élevées, on les trouve également surexprimées lors de l'exposition des plants à la sécheresse, aux fortes teneurs en sel, à des agressions chimiques ou des stress oxydatifs. Les Hsp60, Hsp70 et les Hsp90 vont interagir avec de nombreuses protéines co-chaperonnes qui vont réguler leur activité ou faciliter leur action de repliement protéique. Les Hsp vont

également interagir entre elles pour former un véritable réseau de protéines chaperonnes et agir en synergie avec les autres mécanismes de réponse au stress (Wang *et al.* 2004). Wang *et al.* (2004) ont réalisé une revue très complète sur les fonctions connues des différentes familles d'Hsp.

Les Hsp70 vont empêcher la formation d'agrégats protéiques et permettre aux protéines de se replier pour adopter une conformation active. Elles sont aussi impliquées dans les transports protéiques, la transduction du signal et l'activation de la transcription. Les Hsp70 répondent également aux attaques de pathogènes. L'infection de plants d'*Arabidopsis* par le virus TuMV ou le virus TCV induit l'expression des gènes AtHSC70-1, -2, -3 et AtHSP70 mais pas AtHSP70B (Aparicio *et al.* 2005). Les auteurs montrent également que la quantité de protéine HSP70 cytosolique en se liant à des HSF (heat shock factor) exercerait un rétrocontrôle sur l'expression du gène HSP70. Ce type de régulation concernerait aussi les sHsp (Rhoads *et al.* 2005). La surexpression du facteur HSF3 chez *Arabidopsis* se traduit par une accumulation de Hsp en l'absence de choc thermique et confère aux plants une tolérance accrue à la chaleur (Prandl *et al.* 1998).

Les Hsp60 dont le rôle essentiel est de faciliter le repliement protéique, sont essentiellement chloroplastiques et mitochondriales, même si certaines formes se retrouvent dans le cytoplasme.

En plus du rôle de contrôle du repliement, les Hsp90 (Milioni and Hatzopoulos 1997) se distinguent en intervenant dans les réseaux de transduction du signal, le contrôle du cycle cellulaire, la dégradation des protéines et la destination des protéines.

Les Hsp100/clp vont interagir avec les Hsp70 et les sHsp pour désagréger et déplier les complexes protéiques et permettre leur solubilisation puis leur repliement dans la conformation correcte. L'étude de plusieurs mutants d'*Arabidopsis* thermosensibles portant des mutations pour le gène HOT1 codant pour la protéine AtHsp101 et la recherche de suppresseurs de mutations a permis aux auteurs de restaurer la solubilité d'agrégats de sHsp après un stress thermique (Lee *et al.* 2005). Ce travail révèle une interaction entre la famille des Hsp100/Clp et les sHsp.

Les sHsp ont en commun un domaine conservé de 90 acides aminés appelé α-crystallin domain (ACD). C'est la famille la plus exprimée et répandue chez les plantes. Les sHsp ont une très grande affinité pour se lier aux protéines dans leur forme non native et vont empêcher leur agrégation. Elles interagissent alors avec les protéines chaperonnes ATP dépendantes du complexe DnaK pour replier les protéines sous leur forme native. Chez la tomate, la Hsp21 est induite dans les feuilles et les fruits par un stress thermique. L'expression constitutive de

cette protéine chez la tomate protège le Photosystème II des stress oxydatifs combinant l'exposition au froid et à de fortes lumières appliquées lors de l'expérience (Neta-Sharir *et al.* 2005). Les auteurs ont également observé une accumulation de caroténoïdes dans les fruits des lignées transgéniques.

B.2.2.5.3.2 Les LEA (Late Embryogenesis Abundant Proteins)

Les LEA sont des protéines hydrophiles qui s'accumulent très fortement dans les tissus tolérants à la dessiccation -comme les graines– ou en réponse à la déshydratation, les basses températures, la salinité et l'application d'ABA. Distribuées très largement chez les mono- et dicotylédones, c'est initialement sur le coton qu'elles ont été découvertes. Plusieurs groupes -nommés d'après l'identifiant des clones d'ADNc- ont été formés sur la base de leur similarité de séquence. Le Groupe 1 (D19) est caractérisé par un motif commun de 20 acides aminés. Les LEA du Groupe 2 (D11) aussi appelé **Déhydrines** ou **RAB** partage 3 domaines hautement conservés : le segment-K qui est un motif de 15 acides aminés riches en lysine, une chaîne de sérines contigües ou segment-S, le segment-Y qui est un motif DEYGNP. La protéine adopterait également une conformation en hélice α. Les LEA du Groupe 3 (D7) partages un motif de 11 acides aminés permettant une conformation en hélice α. A ces trois groupes ce sont ajoutés : le Groupe 4 (D113), le Groupe 5 (D29) et le Groupe 6 (D34) ayant également des propriétés communes (Ramanjulu and Bartels 2002). Cette classification a été récemment remise en question par Wise (2003). L'auteur a réalisé une étude bioinformatique en prenant en compte la séquence, la distribution en acides aminés, la présence de motifs, domaines ou conformations spécifiques ainsi que les fonctions supposées associées. Les résultats proposent une nouvelle distribution des groupes existants avec le rapprochement possible de certains d'entre eux (Wise 2003). Toutes les LEA, excepté celles du groupe 5, sont très hydrophiles et stable à haute température souvent même après ébullition. Les fonctions des LEA restent largement méconnues. Mais si l'on considère leur haute solubilité associée à de fortes concentrations, une distribution en acides aminés biaisée et la présence de conformation en hélice α, il est peut probable que leur fonction soit enzymatique. Par contre leur thermo stabilité, leur structure, leur accumulation dans les graines et lors de stress hydriques suggèrent que les LEA jouent un rôle fondamental dans la tolérance à la dessiccation. C'est ce que laisse supposer les observations sur la quantité uniformément répartie dans l'embryon et la localisation cytosolique de deux LEA chez le coton (Roberts *et al.* 1993). Des hypothèses ont été avancées sur leur rôle possible de soluté permettant de conserver une hydratation minimum des cellules et de protéger les structures cytoplasmiques.

La fonction de protection des protéines par les LEA a été montrée chez le pois. L'accumulation des transcrits de la PsLEAm -une LEA du groupe 3 localisée dans la matrice mitochondriale– s'effectue dans les dernières étapes de maturation de la graine puis disparaissent avec la réhydratation lors de la germination. L'induction de cette LEA peut être obtenue dans les feuilles en soumettant les plants à un stress hydrique violent. L'expression a pu être maintenue plusieurs heures après germination par un traitement des graines à l'ABA (Grelet *et al.* 2005). Les auteurs ont montré que la PsLEAm protégeait des effets de la dessiccation deux enzymes mitochondriales (Rhodanese et Fumarase) et augmentait même l'activité de la Rhodanese. La régulation de l'expression des LEA est fréquemment mise en évidence sur des plants exposés à la sécheresse ou à des stress osmotiques. Chez le peuplier, l'expression d'une Déhydrine a été mesurée lors d'un arrêt d'arrosage, de l'application de PEG 6000, d'un milieu concentré en sel et de l'exposition aux basses températures (Caruso *et al.* 2002). Le transcrit peudhn1 était induit dans chaque condition. La tolérance au stress n'est pas toujours améliorée par la surexpression d'une LEA. Cependant chez *Arabidopsis*, l'obtention de plants transgéniques surexprimant simultanément 2 Déhydrines (RAB18 et COR47 ou LT129 et LT130) a permis d'obtenir des plants avec une tolérance accrue au gel. Tous les types de LEA ne sont pas encore connus. Récemment, le gène CpEdi-9 codant pour une protéine hydrophile a été découvert chez *C. plantagineum*. La séquence ne donnait aucune homologie nucléotidique ou protéique avec des gènes connus mais la structure de la protéine et certains motifs ont permis de la rapprocher des Déhydrines ou des LEA du groupe 3. Le transcrit de cette nouvelle LEA est fortement exprimé dans les graines matures et dans les cellules des vaisseaux conducteurs du phloème de plants exposés à la déshydratation, au sel ou à l'ABA. Il est également rapidement accumulé dans les feuilles et racines en déshydratation (Rodrigo *et al.* 2004). Les auteurs ont également étudié la partie promotrice du gène. Ils ont trouvé des motifs ABRE, CE1-like, Myb-like, ASCE-like, PB et un élément spécifique du phloème. La construction de gènes rapporteurs GUS chez des plants d'*Arabidopsis* et de tabac leur a permis de confirmer que ce promoteur répondait à la dessiccation, à l'ABA et s'accumulait au niveau des stomates, du système vasculaire et lors de la formation du pollen. Nous verrons plus en détail la régulation de la transcription dans un prochain paragraphe.

B.2.2.5.4 Ajustement osmotique

Nous avons vu que l'ajustement osmotique était un paramètre essentiel de la résistance au stress. L'augmentation du potentiel osmotique peut être obtenue par accumulation d'ions

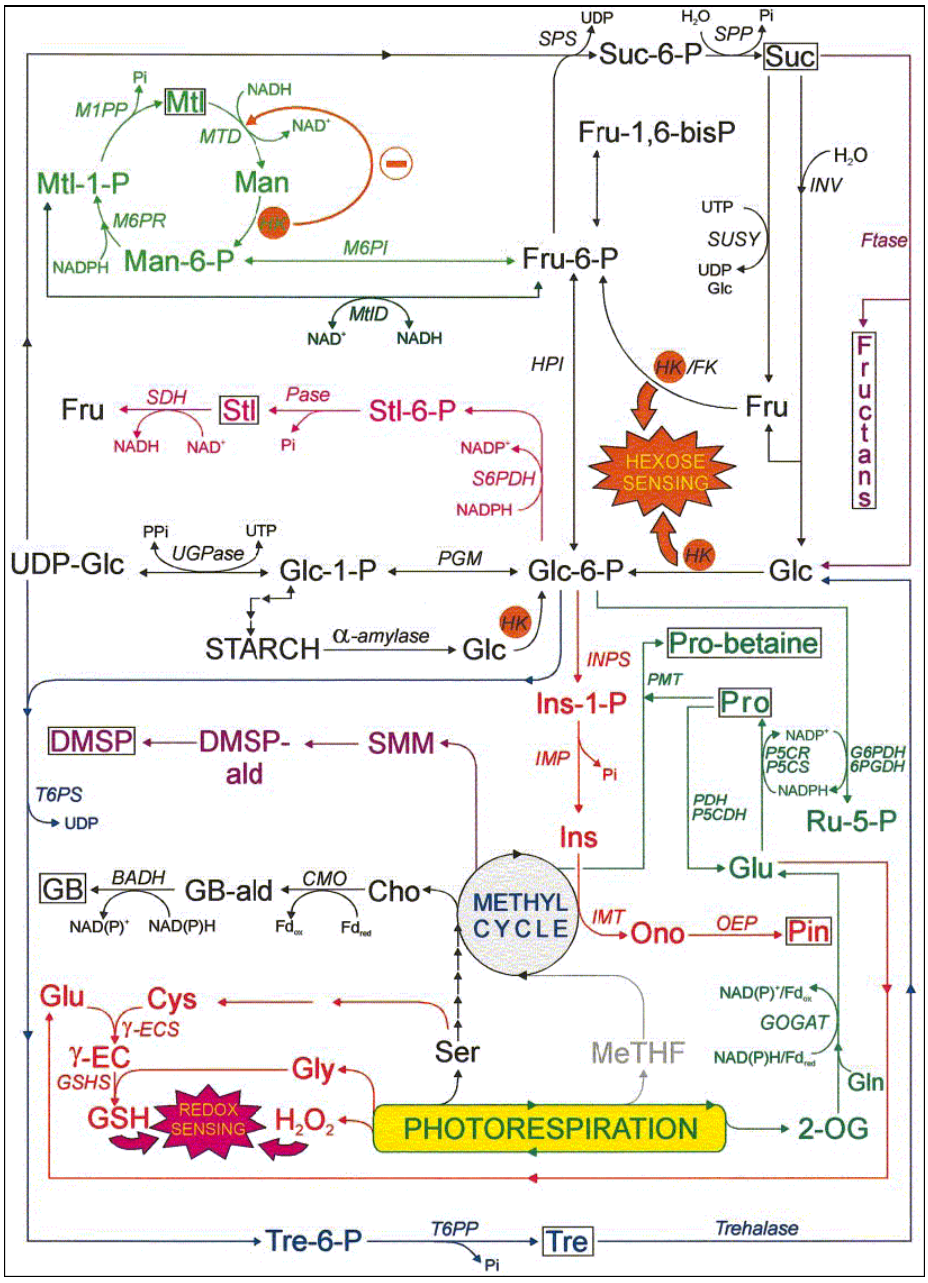


Figure B-28 : Osmolytes et leur métabolisme. Les lignes interrompues indiquent de multiples étapes intermédiaires. (Hare *et al.*, 1998)

Abréviations des composés : Cho, choline; DMSP, 3-dimethylsulphoniopropionate; DMSP-ald, DMSP aldehyde; g-EC, g-glutamylcysteine; Fd, ferredoxin; Fru, fructose; Fru-6-P, fructose-6-phosphate; Fru-1,6-bisP, fructose-1,6-bisphosphate; Glc-1-P, glucose-1-phosphate; Glc-6-P, glucose-6-phosphate; GB, glycine bétaine; GB-ald, glycine bétaine aldehyde; Gln, glutamine; Glu,Glutamate; GSH, reduced glutathione; Ins-1-P, myo-inositol-1-phosphate; Ins, myo-inositol; Man, mannose; Man-6-P, mannose-6-phosphate; Met, methionine; MeTHF, N5-methylenetetrahydrofolate; Mtl-1-P, mannitol-1-phosphate; Ono, ononitol; 2-OG, 2-oxoglutarate; Pin, pinitol; Pro, proline; Ru-5-P, ribulose-5-phosphate; Ser, serine; SMM, S-methylmethionine; Suc, sucrose; Suc-6-P, sucrose-6-phosphate; Stl-1-P, sorbitol-1-phosphate; Tre, trehalose; Tre-6-P, trehalose-6-phosphate; UDPGlc, uridine 5¢-diphosphoglucose.

Abréviations des enzymes (en italique) : BADH, betaine aldehyde dehydrogenase; CMO, choline monooxygenase; g-ECS, g-EC synthetase; FK, fructokinase; Ftase, fructosyltransferase (levansucrase); G6PDH, glucose-6-phosphate dehydrogenase; GOGAT, glutamate synthase; GSHS, GSH synthetase; HK, hexokinase; HPI, hexose phosphate isomerase; IMP, myo-inositol-1-phosphate monophosphatase; IMT, myo-inositol O-methyltransferase; INPS, myo-inositol-1-phosphate synthase; INV, invertase; M6PI, mannose-6-phosphate isomerase; M6PR, mannose-6-phosphate reductase; M1PP, mannitol-1-phosphate phosphatase; mtlD, bacterial mannitol-1-phosphate 5-dehydrogenase; MTD, plant mannitol dehydrogenase; OEP, ononitol epimerase; Pase, phosphatase; PDH, proline dehydrogenase; P5CDH, P5C dehydrogenase; P5CR, P5C reductase; P5CS, P5C synthetase; PMT, proline methyltransferase(s); 6PGDH, 6-phosphogluconate dehydrogenase; PGM, phosphoglucomutase; SPS, sucrosephosphate synthase; SPP, sucrose-6-phosphatase; SDH, sorbitol dehydrogenase; S6PDH, aldose-6-phosphate reductase; SUSY, sucrose synthase; T6PP, trehalose-6-phosphate phosphatase; T6PS, trehalose-6-phosphate synthase; UGPase, UDPGlc pyrophosphorylase.

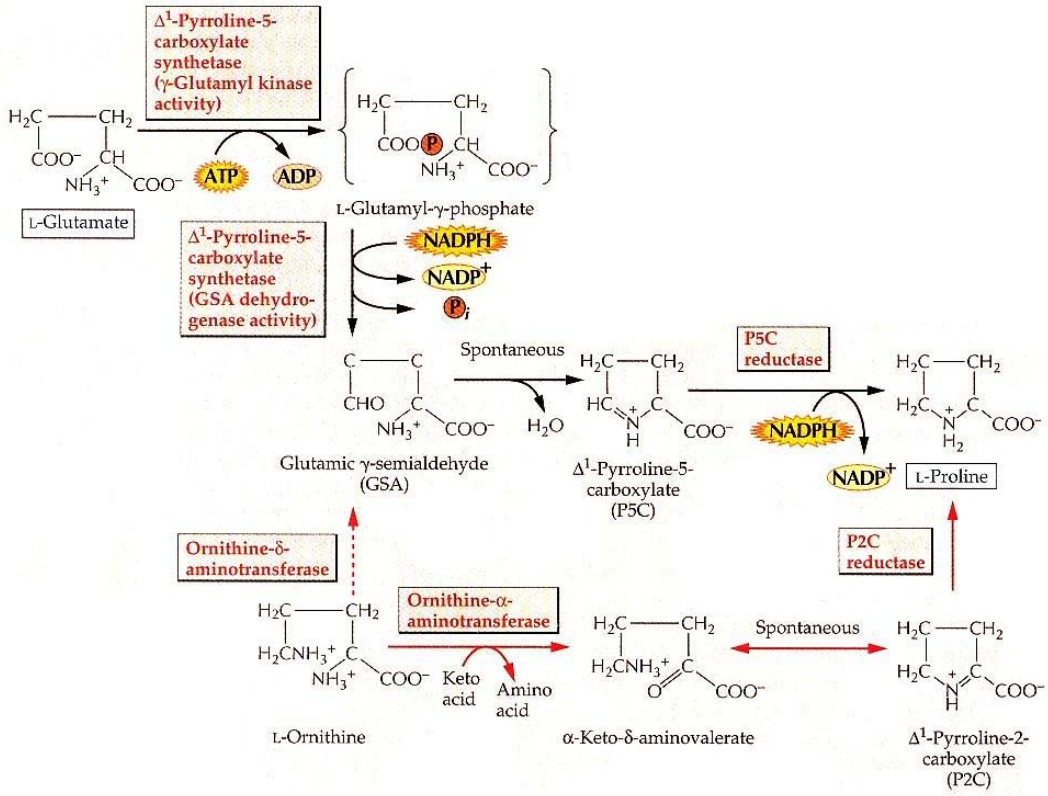


Figure B-29 : Voie de biosynthèse de la proline chez les plantes. La voie dérivée du glutamate est bien établie (flèches noires). En rouge la voie de l'ornithine dont l'importance reste à déterminer. (Bray *et al.*, 2000)

dans la cellule. Cependant, à forte concentration, les charges électriques vont interagir et altérer les propriétés des cofacteurs, substrats, membranes et enzymes. La cellule peut compartimenter ces ions dans la vacuole mais aussi produire des composés organiques qui en plus du rôle d'osmolyte, vont pouvoir offrir une osmoprotection en stabilisant les protéines.
Les principales molécules identifiées sont la proline, la glycine bétaïne et les sucres alcool. Hare *et al.* (1998) ont fait une synthèse sur les osmolytes, leur accumulation et leurs rôles au cours du stress. Il synthétise dans la figure B-28 les principales voies métaboliques et les enzymes impliquées.

B.2.2.5.4.1 Proline

Un des osmolytes le plus répandu est certainement la **proline**. La synthèse de cet acide aminé à partir du **L-Glutamate** est catalysée par un enzyme : le Δ^1-Pyrroline-5-carboxylate synthétase (**P5C synthétase**) dont la fonction est double. Il permet d'abord la conversion en présence d'ATP du L-Glutamate en L-Glutamyl-γ-phosphate qui lui-même est convertit en présence de NADPH en Glutamic γ-semialdehyde (GSA) donnant le **P5C** (Δ^1-Pyrroline-5-carboxylate). Un autre enzyme, le **P5C réductase** permet d'obtenir la proline en présence de NADPH (Fig. B-29). Une autre voie encore mal connue serait possible à partir de l'Ornithine. La proline se retrouve également dans les cellules sous forme d'analogues parmi lesquels le N-methyl-L-proline (MP), le trans-4-hydroxy-Nmethyl-L-proline (MHP), le trans-4-hydroxy-N-dimethyl-L-proline (DHP), ou la stachydrine (proline bétaïne).

L'importance de la proline dans la tolérance au stress a orienté nombre de recherches sur les gènes impliqués dans sa synthèse. Une équipe japonaise a étudié la réponse des deux enzymes cités lors de l'exposition de plants d'*Arabidopsis* à différents traitements (Yoshiba *et al.* 1995). Les résultats de Northern blot montrent que la P5C synthétase est induite par la sécheresse, le stress salin et l'exposition à l'ABA mais pas les stress thermiques. Ils ont pu vérifier que l'expression du gène lors du stress osmotique était accompagnée par une accumulation de proline. Par contre, la P5C réductase n'était pas régulée lors du stress. Dans l'hypothèse que les autres voies de biosynthèse seraient minoritaires, la synthèse de la proline serait donc chez *Arabidopsis* sous le contrôle d'un seul enzyme, la P5C réductase, codé par un gène unique (Savoure *et al.* 1995). Lors de la disparition du stress, la proline fortement accumulée doit être dégradée. Plusieurs voies de dégradations ont été proposées et étudiées. La proline peut être oxydée en glutamate en faisant intervenir la **proline oxydase** ou **proline déshydrogénase** (ProDH). Verbruggen *et al.* (1996) ont isolé un ARNm, At-POX, codant pour cet enzyme chez *Arabidopsis*. Les auteurs ont montré que l'expression de At-POX est

régulée négativement lors d'un stress osmotique et précède l'accumulation de proline. Lors de la levée du stress son expression augmente. Nakashima *et al.* (1998) ont confirmé ces résultats en mesurant une accumulation d'un transcrit ERD5 codant pour un précurseur de ProDH lors de la réhydratation, d'un stress hypo-osmotique ou de l'exposition à de la proline chez *Arabidopsis*. La création d'un système de gène rapporteur GUS avec le promoteur de la ProDH montre une expression dans l'ensemble de la plante avec une coloration plus marquée à l'extrémité des racines (Nakashima *et al.* 1998). L'expression d'un ARN antisens de la ProDH, réalisé sur *Arabidopsis* par Nanjo *et al.* (1999), confirme les résultats précédents en améliorant la tolérance au stress froid et salin des lignées transformées.

Les osmoticums, pour jouer leur rôle physiologique, doivent être compartimentés. La proline synthétisée par la plante fait l'objet d'un transport actif. Ueda *et al.* (2001) ont identifié un ADNc codant pour un transporteur de Proline (HvProT) dont l'expression est fortement induite au niveau des racines mais pas des feuilles lors du stress salin. La spécificité du transporteur pour la proline en présence de compétiteurs tels que d'autres acides aminés, la glycine bétaïne ou la choline, a été vérifiée (Ueda *et al.* 2001). Des ADNc codant pour des transporteurs spécifiques de la proline et induits lors de stress osmotiques ont été identifiés chez *Arabidopsis* (ProT2, ProT1, famille des AAP), la tomate et le riz (Rentsch *et al.* 1996; Igarashi *et al.* 2000).

B.2.2.5.4.2 Glycine bétaïne

En tant qu'osmolyte, la glycine bétaïne ne se retrouve accumulée principalement que chez les plantes hautement tolérantes au sel ou au froid. Elle a pu être détectée chez l'épinard, la betterave à sucre ou encore le maïs. Elle est synthétisée en deux étapes à partir de la choline dans une réaction catalysée par la choline monooxygénase et l'aldéhyde bétaïne déshydrogénase. Une expérience de résistance au stress froid a été réalisée sur le tabac (Parvanova *et al.* 2004). Les auteurs ont utilisé des lignées transgéniques GB9 exprimant le gène bactérien codA qui permet l'accumulation de glycine bétaïne. La mesure de marqueurs du stress tels que l' « electrolyte leakage », la teneur en Malondialdehyde (MDA) –produit final de la peroxydation des lipides-, l'H_2O_2, les catalases et les peroxydases, indique que les plants riches en glycine bétaïne sont plus tolérants au stress froid que le phénotype sauvage.

B.2.2.5.4.3 Sucres et dérivés

Si les sucres sont les éléments carbonés primaires synthétisés et exportés dans toute la plante lors de la photosynthèse, ils ont également un rôle non énergétique d'osmorégulateur et

d'osmoprotectant. L'augmentation de la teneur en sucres ou dérivés, et leur rôle dans la réponse aux stress hydrique a été montré depuis longtemps chez de nombreuses espèces (Gorham *et al.* 1981). L'accumulation de saccharose chez les plantes reviviscentes est particulièrement nette et prendrait une large part dans l'acquisition de la tolérance à la dessiccation. *C. plantagineum* accumule de grandes quantités de 2-octulose qui servirait de source à la conversion en saccharose pour une part supérieure à 80% (Scott 2000). Lors d'une dessiccation sévère une saturation en sucres permet une vitrification du cytoplasme de la cellule et évite une cristallisation des solutés qui causerait des dommages cellulaires irréversibles. Les propriétés physiques obtenues sont proches de celles d'un solide et sont à l'origine de la tolérance à la dessiccation des graines. Koster *et al.* (1988) ont rapporté une corrélation entre la teneur en sucres solubles –stachyose et raffinose- des graines et la capacité germinative du soja, du pois et du maïs.

Le sucre principal intervenant dans la croissance est certainement le saccharose mais d'autres rôles lui sont attribués dans les mécanismes de régulation de l'expression des gènes. Il peut être synthétisé à partir d'UDP-glucose et de Fructose-6-Phosphate ou de Fructose. Dans le premier cas l'enzyme impliqué est la sucrose-phosphate synthase (SPS), dans le second, il s'agit du sucrose synthase (SuSy) mais in vivo cette enzyme catalyse plutôt la réaction inverse de dégradation du saccharose. Dans une synthèse sur le rôle et la régulation de la SPS chez les plantes supérieures, Huber *et al.* (1996) rapportent que cet enzyme est activé par les stress osmotiques chez l'épinard et les tubercules de pomme de terre.

L'osmoprotectant le plus efficace en termes de concentrations est sans doute le **Tréhalose**. Ce disaccharide non réducteur a la particularité d'être particulièrement stable à la température, au pH et aux réactions avec d'autres molécules comme les acides aminés. Présent chez de nombreuses bactéries, champignons et quelques plantes tolérantes à la dessiccation, il n'est pratiquement pas détecté chez les plantes supérieures. Cependant des séquences homologues aux enzymes de sa voie de biosynthèse ont été découvertes chez *Arabidopsis* (Goddijn and Dun 1999). Il s'agit de la tréhalose-6-phosphate synthase (TPS) qui permet de convertir l'UDP-Glucose et le Glucose-6-P en Tréhalose-6-P, et de la tréhalose-6-phosphate phosphatase (TPP) qui convertit le Tréhalose-6-P en Tréhalose (Fig B-30). Des essais de transformation de plants de tabac avec des TPS de levures ou de bactéries ont permis d'obtenir des plants avec une meilleure tolérance à la sécheresse, une capacité photosynthétiques sous stress améliorée, une baisse de la teneur en saccharose et une

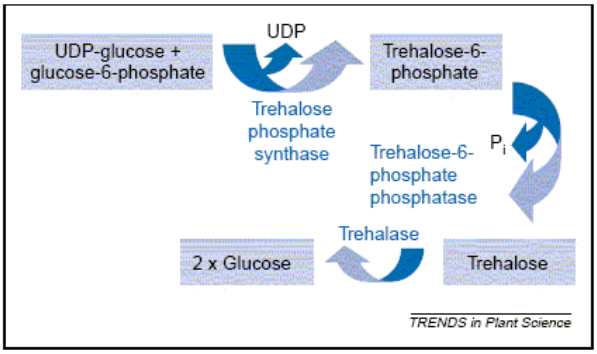


Figure B-30 : Biosynthèse et métabolisme du tréhalose chez les plantes. (Penna *et al.*, 2003)

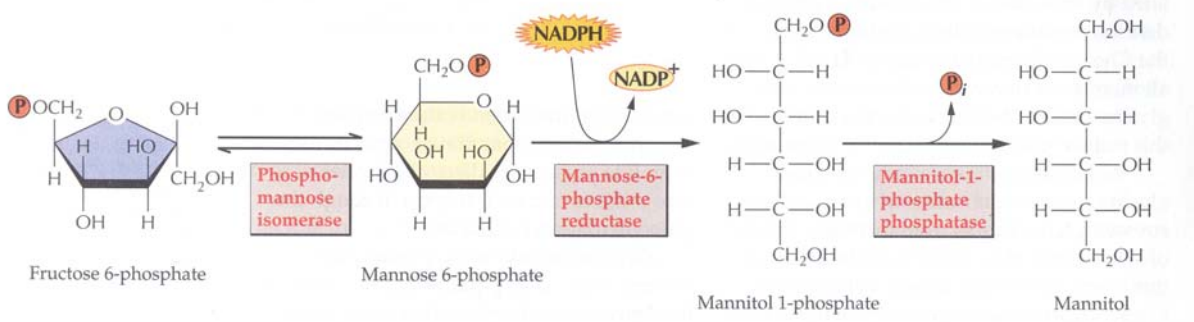


Figure 22.10
The biosynthetic pathway for mannitol.

Figure B-31 : Voie de biosynthèse du Mannitol. (Bray *et al.*, 2000)

croissance retardée. Les quantités de tréhalose accumulées restaient cependant faibles (≤ 0.5 $\mu molg^{-1}$) (Penna 2003).

Le **Mannitol** est la forme réduite du mannose. Ce sucre alcool se retrouve très largement distribué chez les plantes. Son accumulation est souvent mise en évidence lors de stress salins. Le mannitol est synthétisé à partir de Fructose-6-phosphate (Fig. B-31). Ce F6P est convertit en Mannose-6-phosphate par la Mannose-6-phosphate isomérase puis une réaction catalysée par la Mannose-6-phosphate réductase en présence de NADPH permet d'obtenir du Mannitol-1-phosphate qui donne le Mannitol sous l'action de la Mannitol-1-phosphate phosphatase ou Mannitol-1-phosphatase. Chez les bactéries, il existe un gène Mt1D -codant pour une Mannitol-1-Phosphate Déshydrogénase– qui est nécessaire à l'assimilation du Mannitol et du Glucitol par la bactérie (Watanabe *et al.* 2003). L'enzyme produite, catalyse la réaction de conversion entre le D-fructose 6-phosphate et le D-mannitol 1-phosphate en présence de NADH. L'expression combinée de ce transgène Mt1D et d'un autre gène GutD –codant une glucitol-6-phosphate déshydrogénase– chez le pin *taeda* s'est traduite par une surexpression des transcrits correspondants et une accumulation de mannitol et glucitol. Les plants obtenus présentaient une meilleure tolérance au stress salin (Tang *et al.* 2005). Le **Glucitol** est aussi appelé sorbitol, c'est un sucre présent à de très fortes concentrations dans les fruits des *Rosaceae* où il peut représenter 60 à 90% des composés carbonés provenant des feuilles. On le retrouve également impliqués dans la régulation osmotique (Lo Bianco *et al.* 2000). L'expression d'un transgène de la pomme codant pour une sorbitol-6-phosphate déshydrogénase chez le tabac permet de mesurer une accumulation de sorbitol. Les teneurs enregistrées seraient suffisantes pour avoir un effet sur le potentiel osmotique. Mais les concentrations élevées de ce composé chez une plante dont le métabolisme n'y est pas adapté, ont entraîné des perturbations comme un développement racinaire réduit, des nécroses des feuilles ou une croissance réduite. Seul les plants ayant des teneurs faibles en sorbitol ont conservé le phénotype sauvage et l'impact sur la tolérance au stress était difficile à interpréter (Sheveleva *et al.* 1998). Cette expérience met en évidence les difficultés à estimer l'implication d'un composé ou d'un gène dans la réponse à un stress. Les voies de biosynthèse sont interconnectées et l'augmentation de l'activité d'un enzyme, en convertissant de façon accrue un composé va perturber les voies de biosynthèse pour lesquelles ce composé est une source. Dans le cas des sucres cités jusqu'ici, les sources sont des sucres intervenant directement dans la glycolyse qui est la voie directe de production d'énergie dans la cellule. De plus l'accumulation d'osmoprotectant, en stabilisant les protéines peut gêner les réactions biochimiques.

Il existe une autre catégorie d'osmoticums regroupés sous le nom **d'Oligosaccharides de la famille du Raffinose (RFO family)**. Leur rôle essentiel dans le processus de dessiccation des graines a été rapporté chez de nombreuses espèces. La synthèse du **raffinose** et du **stachyose** nécessite du galactinol qui est produit par une galactinol synthase (**GoIS**) à partir d'UDP-galactose et de myo-inositol (Fig. B-32). La raffinose synthase produit le raffinose à partir de galactinol et de saccharose tandis que la stachyose synthase catalyse la synthèse de stachyose à partir de galactinol et de raffinose. La GoIS va donc réguler l'étape initiale de biosynthèse des RFO et son activité ainsi que l'expression des gènes codant la GoIS sont induits lors du stress froid. Taji *et al.* (2002) ont étudié l'accumulation du Galactinol et des RFO chez des plants d'*Arabidopsis* exposés au froid, au sel et à la déshydratation. Les teneurs en galactinol et raffinose mesurées étaient fortement augmentées. Les auteurs ont identifié 7 gènes codant la GoIS chez *Arabidopsis*. Trois des 7 gènes étaient surexprimés pendant le stress. Les gènes AtGoIS1 et 2 répondaient au stress salin et à la sécheresse ; ils possèdent des promoteurs DRE ou DRE-like et ABRE-like et sont légèrement induits par l'ABA. AtGoIS3 répondait au froid ; il possède des promoteurs DRE, DRE-like et ABRE mais n'est pas induit par l'ABA. L'utilisation de lignées transgéniques a permis de confirmer d'une part, que seul AtGoIS3 était induit par des facteurs de transcription DREB1A et d'autre part que la surexpression d'AtGoIS conférait au plant transgénique une nette tolérance à la sécheresse et une transpiration réduite de leurs feuilles (Taji *et al.* 2002). Les auteurs émettent l'hypothèse d'une induction de la production d'ABA par le galactinol et le raffinose et ont mis en évidence le rôle d'osmoprotectant sur les membranes des chloroplastes.

Les *Pinaceae*, les *Fabaceae* et les *Caryophylaceae* accumulent du **Pinitol**. Les concentrations de ce sucre alcool sont plus fortes chez les espèces halophytes ou adaptées à la sécheresse. Chez *Mesembryanthemum crystallinum* exposés à de fortes concentrations en sel, le Pinitol peut constituer plus de 70% des sucres solubles (Nelson *et al.* 1999). Ce constituant majeur des plants de soja est répartit dans la plante selon un gradient de concentration ; les teneurs les plus élevées correspondant aux feuilles des nœuds supérieurs les plus exposées aux rayonnements solaires (Streeter *et al.* 2001). Une étude sur des variétés introduites en chine a mis en évidence que les plants présentant les concentrations les plus fortes correspondent aux variétés sélectionnées pour leur rendement dans les régions les plus sèches (Streeter *et al.* 2001). Ces résultats soulignent l'implication du Pinitol dans la tolérance à la sécheresse chez le soja. Le Pinitol, le **myo-Inositol** et l'**Ononitol** font partie de la famille des cyclitols caractérisés par un cycle à 6 carbones. Les enzymes intervenant dans sa voie de biosynthèse à

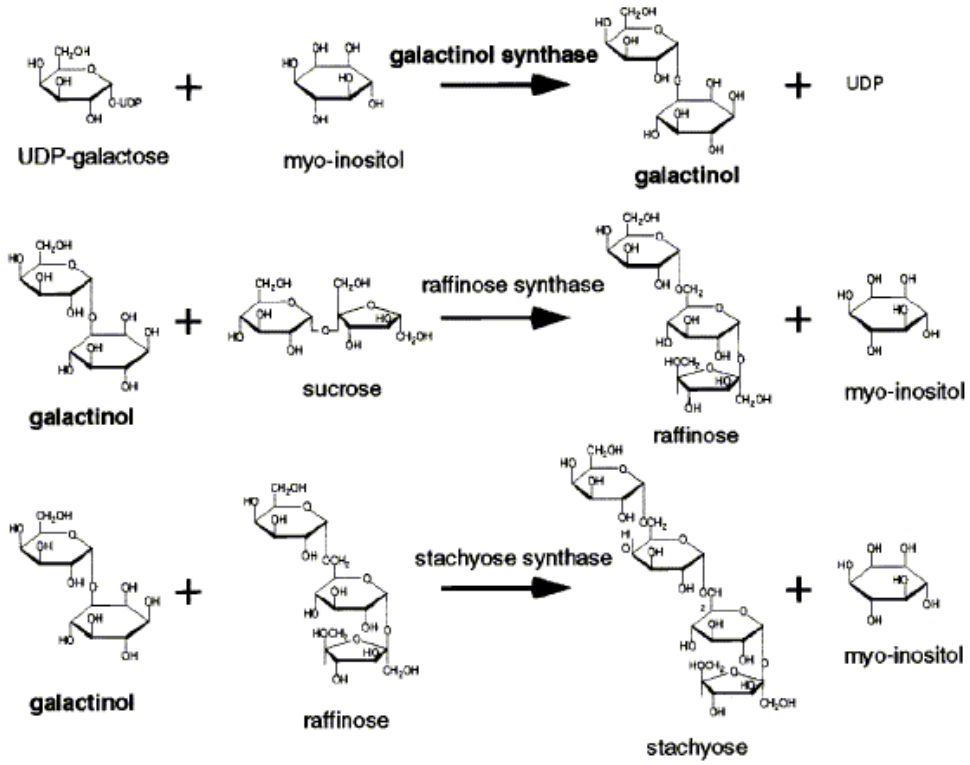


Figure B-32 : Voies de biosynthèse des oligosaccharides de la famille du Raffinose (RFO family) (Taji *et al.*, 2002)

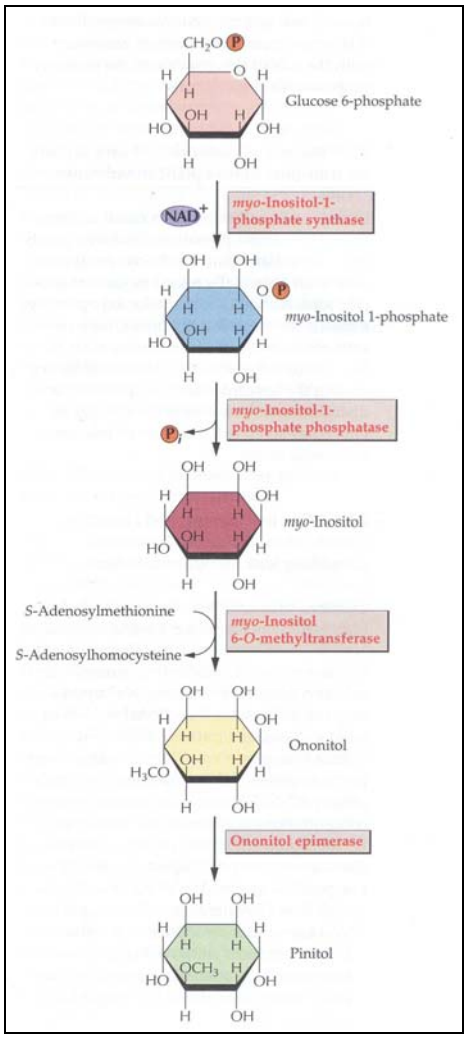


Figure B-33 : Voie de biosynthèse du pinitol. (Bray *et al.*, 2000)

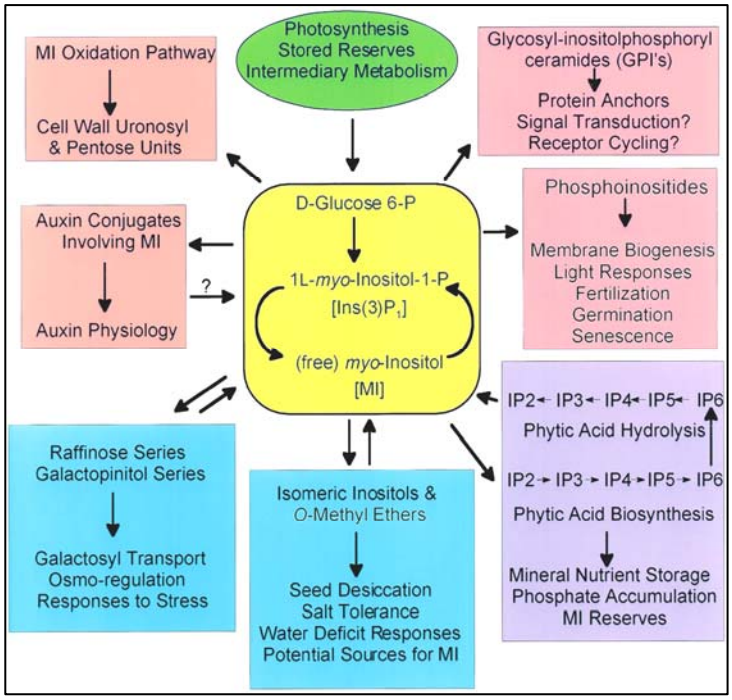


Figure B-34 : Rôle central de l'Inositol et voie de biosynthèses associées. (Loewus *et al.*, 2000)

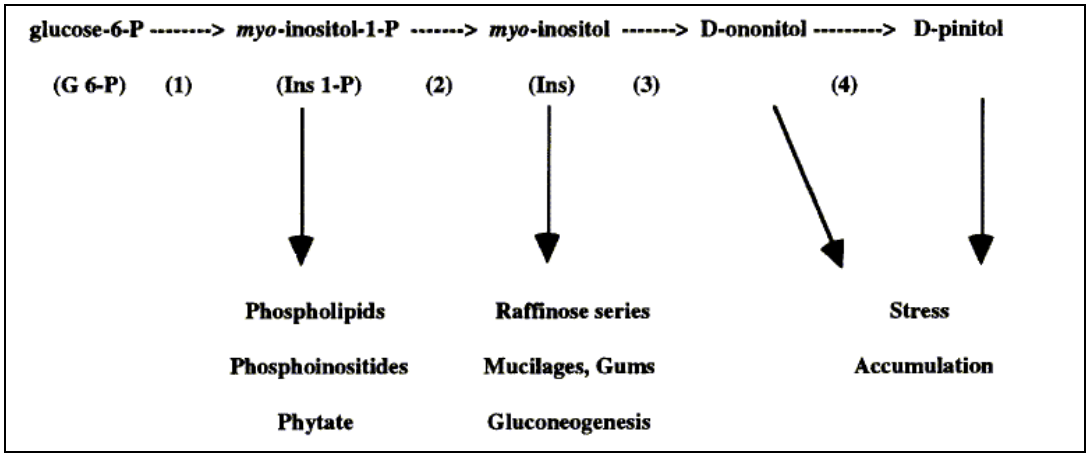


Figure B-35 : Voie de Biosynthèse de l'Inositol en relation avec la biosynthèse des osmolytes. (1) myo-Inositol-1-phosphate synthase (INPS1) ; (2) myo-Inositol-1-phosphate-phosphatase (IMP1) ; (3) myo-Inositol-O-methyltransferase (IMT1) ; (4) Ononitol épimérase (OEP1) (Ishitani *et al.*, 1996)

partir du glucose-6-phosphate sont au nombre de quatre (Fig B-33). La première, la myo-Inositol-1-phosphate synthase (INPS1), convertit le G6P en myo-Inositol 1-phosphate (Ins 1-P) qui sert de précurseur à la synthèse des phospholipides, phosphoinositides et phytates. L'Ins 1-P est déphosphorylé par la myo-Inositol-1-phosphate phosphatase (IMP1) et donne le myo-Inositol (Ins) qui occupe un rôle clé pour de nombreuses voies métaboliques comme la synthèse des dérivés du Raffinose (Fig. B-34). La méthylation du myo-inositol par la myo-inositol-O-methyltransferase (IMT1) aboutit à l'Ononitol dont l'épimérisation par l'Ononitol épimérase (OEP1) donne le Pinitol (Fig.B-35). Plusieurs études ont montré que les transcrits de l'INPS1 et de l'IMT1 étaient fortement induits lors de l'exposition de *Mesembryanthemum crystallinum* au sel (Ishitani *et al.* 1996). Dans cette plante, l'accumulation du myo-inositol, de l'ononitol et du pinitol se fait dans le cytosol et des mécanismes de transport actif d'inositol sont mis en place pour accumuler le composé au niveau des feuilles au détriment des racines (Nelson *et al.* 1999). Le rôle du myo-Inositol dans la plante a été développé de façon très complète dans une revue de Loewus *et al.* (2000). L'auteur souligne que la seule voie de synthèse de novo du myo-Inositol repose sur l'IMP1 et décrit les principales réactions dans lesquelles il est impliqué. La tolérance à la sécheresse et au milieu salin peut être apportée par l'Ononitol. La surexpression de l'IMT1 par des plants de tabac leur permet d'accumuler l'Ononitol et de limiter l'inhibition de l'assimilation photosynthétique lors du stress, les plants obtenus sont plus tolérants au stress (Sheveleva *et al.* 1997). Nous verrons plus loin l'importance du pinitol dans la réponse au stress hydrique chez le pin maritime.

B.2.2.5.5 Transport de l'eau et des solutés

Les flux d'eau dans les cellules vont dépendre de la perméabilité des membranes cellulaires à l'eau mais aussi à des ions et autres molécules dont le gradient de concentration entre les compartiments intra et extra cellulaires influe sur les potentiels osmotiques.

Une famille de protéines intra membranaires, les **MIP** (Major Intrinsic Protein) va jouer un rôle important dans la conductance de ces membranes. Les **aquaporines** (**AQP**) font partie de ces MIP et ont été localisées dans la membrane plasmique (**PIP** /Plasma Intrinsic Protein) et dans la membrane de la vacuole (Tonoplaste)(**TIP** /Tonoplast Intrinsic Protein). Ces protéines de 23 à 29 kDa ont 6 domaines transmembranaires en hélice α et de nombreuses régions conservées. Les AQP laissent passer sélectivement l'eau et les gaz dissous. Le niveau d'expression des gènes codant pour des aquaporines n'est pas similaire quelque soit le tissus. Les racines et les fleurs sont en général les deux organes où l'expression est la plus forte. Chez *Nicotiana glauca*, les ARNm codant pour des MIP s'accumulent dans les racines ou les

tiges et leur expression est faible dans les cellules du mésophile, les feuilles entières ou les cellules de garde (Smart *et al.* 2001). On observe également une régulation du niveau d'expression des gènes codant pour des aquaporines. Une répression forte de l'accumulation des ARNm de NgMIP2, NgMIP3 et NgMIP4 a été observée chez *N. glauca*. L'étude de 5 TIP chez le tournesol a montré que lors d'une déshydratation, la quantité de transcrits dans les racines augmentait pour SunTIP7, diminuait pour SunTIP18, augmentait transitoirement pour SunTIp20 et ne fluctuait pas pour SunRb7 et SunγTIP (Sarda *et al.* 1999). La régulation de l'expression des aquaporines va influer sur les propriétés de perméabilité à l'eau des membranes. Leur fonction a été étudiée plus avant chez le tabac en se servant d'une technique d'ARNm antisens. Siefritz *et al.* (2002) ont comparé le phénotype de plants transgéniques sous exprimant le gène NtAQP1 codant pour une PIP1 avec le phénotype sauvage. Les auteurs ont observé une perméabilité plus forte des membranes des cellules exprimant le gène. Les plants transgéniques présentaient une conductivité hydraulique réduite et une résistance plus basse lors de l'exposition à des stress osmotiques (PEG) violents. Chez *Arabidopsis*, le génome contient 35 gènes codant des aquaporines dont 13 sont homologues à des PIP. Deux études récentes ont porté sur l'étude de l'expression de ces gènes. Il apparaît que l'expression des AQP est dépendante de l'organe. AtPIP1;1 et AtPIP2;2 sont exprimés préférentiellement dans les racines ce qui est le cas de la plupart des PIP étudiées. La plupart des gènes sont régulés négativement dans les feuilles en condition de sécheresse excepté AtPIP1;4, AtPIP2;5 surexprimés et AtPIP2;6, AtSIP1;1 non régulés (Alexandersson *et al.* 2005). Ces auteurs ont pu vérifier que la variation de quantité de transcrits des PIP suivait bien celle des protéines. La réponse à différents types de stress abiotiques –osmotique, froid, haute salinité- et à l'ABA a été testée par Jang *et al.* (2004). Le traitement froid provoque une sous expression de pratiquement toutes les PIP. Le stress osmotique induit des régulations positives ou négatives marquées. Le stress salin est à l'origine d'une légère surexpression concernant pratiquement toutes les PIP. L'application d'ABA induit des réponses variables. Cependant certaines PIP ont des régulations opposées entre feuilles et racines. Les auteurs proposent un schéma des traitements régulant les différentes PIP (Fig. B-36) (Jang *et al.* 2004).

La sélectivité des aquaporines pour l'eau n'est pas absolue. Tyerman *et al.* (2002) décrivent les autres rôles de ces canaux. Les AQP peuvent être perméables au CO_2, à l'ammoniaque, l'ammonium, à l'acide borique et au peroxyde d'hydrogène.

Le contrôle des concentrations en Na^+ et K^+ est indispensable au fonctionnement de la cellule. Des protéines spécialisées vont assurer le transport membranaire de ces ions. Il a été montré que l'expression des gènes codant pour ces canaux membranaires est régulée lors des stress.

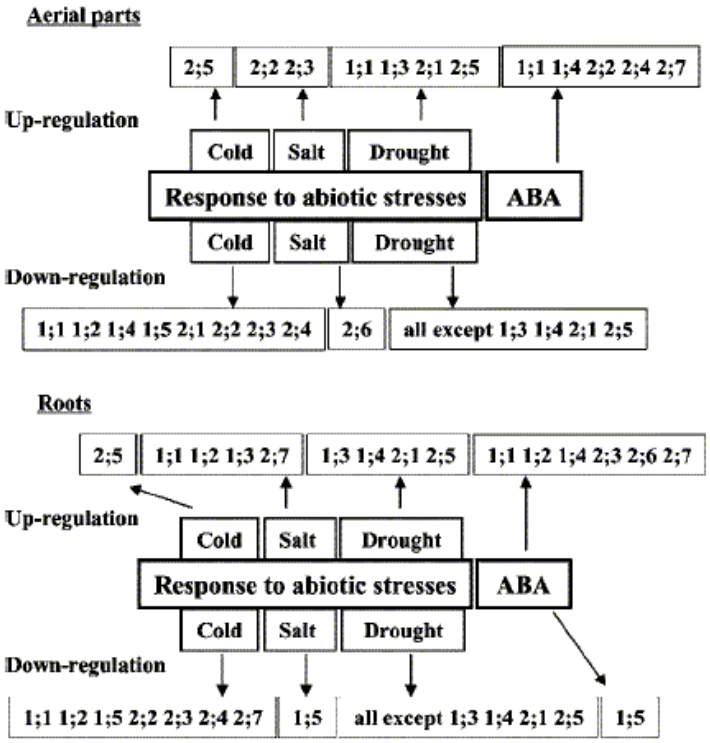


Figure B-36 : Schéma de l'implication des gènes codant des PIP (indiqués par es chiffres) dans les réponses aux stress abiotiques ou à l'ABA. Les gènes indiqués sont régulés positivement ou négativement au moins par un facteur 3 dans les racines ou les parties aériennes chez Arabidopsis lors des traitements indiqués. (Jang *et al.*, 2004)

Shi *et al.* ont montré que le gène SOS1, codant pour un antiporteur membranaire Na+/H+ d'*Arabidopsis*, était essentiel à la tolérance aux fortes salinités (Shi *et al.* 2000). L'expression constitutive de SOS1 chez des plants transgéniques permet une meilleure croissance du système racinaire en condition de stress salin (Shi *et al.* 2003). Ces plants, ont une accumulation de Na^+ réduite car l'efflux de Na^+ est augmenté. L'accumulation de solutés dans les vacuoles des cellules peut être obtenue par la surexpression de gènes codant pour des pompes à proton telles que AVP1. Des plants transgéniques exprimant cette H^+-pyrophosphatase vacuolaire permet aux plants d'acquérir une tolérance à la sécheresse et aux stress salin (Gaxiola *et al.* 2001).

Il existe d'autres type de transporteurs spécialisés dans le transport d'autres composés comme les acides aminés (Chen and Bush 1997; Igarashi *et al.* 2000)ou les monosaccharides (Truernit *et al.* 1996; Weber *et al.* 2000). L'expression des gènes codant ces protéines peut être régulée et tissus dépendante (Kwart *et al.* 1993; Truernit *et al.* 1996). Le produit transporté peut avoir un rôle dans la réponse au stress hydrique. Il a été montré que les transporteurs pour la proline pouvaient être induit par les stress (Rentsch *et al.* 1996; Ueda *et al.* 2001).

B.2.2.5.6 Modification structurale de la cellule

Au cours de la croissance, du développement et de stress biotiques et abiotiques, la structure de la paroi cellulaire est continuellement modifiée sous l'action d'enzymes. Lors d'un stress hydrique, le maintien d'une croissance racinaire accompagné d'un arrêt de croissance des parties aériennes est souvent observé. La croissance nécessite l'expansion et/ou la division cellulaire. La paroi joue un rôle essentiel dans ce mécanisme.

Basiquement, les parois des cellules végétales sont constituées d'un réseau de microfibrilles de cellulose reliées entre elles par des glycans. Mais les parois sont des entités complexes faites d'un amalgame d'eau, de polysaccharides, de protéines, de lignine et de substances telles que cutine, subérine, cires auxquelles s'ajoutent le Bore, le Calcium et des enzymes. Les protéines localisées dans la paroi se distinguent par la richesse en un ou deux acides aminés, la présence de domaines hautement répétitifs dans leur séquence et leur forte ou faible glycosylation. Les cinq classes principales sont les **hydroxyproline-rich glycoproteins** (HRGPs) ou **extensines**, les **glycine-rich proteins** (GRPs), les **proline-rich proteins** (PRPs), les **arabinogalactane proteins** (AGPs) et les lectines (Showalter 1993). La structure, la localisation et la composition de ces protéines a été développée par Cassab (1998). Les gènes codant pour ces protéines appartiennent à de grandes familles multigéniques (Imoto *et al.*

2005). En dehors des protéines citées, on retrouve également toute une série d'enzymes intervenant dans les modifications, la dégradation et la synthèse des parois.

Les **GRPs** contiennent jusqu'à 70% de glycine avec une glycosylation faible à modérée. Elles pourraient se diviser en deux groupes distincts suivant la présence ou non d'un peptide signal qui déterminerait leur localisation membranaire ou non.

Les **PRPs** présentent de grandes homologies avec les extensines et sont identifiées par leurs répétitions Pro-Pro. Leur rôle dans la structure de la paroi reste encore très peu connu mais la présence des GRP et PRP dans les parois (Ryser 2003) et leur régulation dans différentes conditions (Showalter 1993; Akiyama and Pillai 2003) laisse supposer un rôle possible dans la réponse au stress hydrique.

Les **extensines** sont des protéines nettement plus étudiées que l'on retrouve dans différents tissus dont les racines. Leur motif caractéristique est la répétition d'un pentapeptide $Ser(Hyp)_4$ constitué de Sérine galactosylé et d'Hydroxyproline qui porte de 1 à 4 arabinose. Ces extensines sont sécrétées par la cellule au niveau de la paroi où elles se fixent par des liaisons inter et intra moléculaires. Leur expression est régulée par le stress et tissu spécifique. Le gène AtExt1 est ainsi exprimé dans le chevelu racinaire et les zones d'élongation des racines primaires (Merkouropoulos and Shirsat 2003).

Les **AGPs** sont des protéoglycanes, caractérisée par leur très forte glycosylation qui peut atteindre 90% de la masse du produit formé. Elles ne sont pas liées de façon covalente à la paroi et se retrouvent dans des vésicules dérivant du Golgi, dans la membrane plasmique et la paroi. Leur composition suggère des rôles possibles de glue, lubrifiant ou humectant, mais mis à part leur composition commune, aucune fonction certaine ne leur est encore attribuée. Les transcrits codant des arabinogalactanes sont régulièrement mis en évidence dans les réponses à divers stress. Chez le pin *taeda*, l'analyse de l'expression de 11 transcrits d'AGPs a montré que toutes les AGPs étudiées, exceptée ptaAGP5, étaient réprimées dans les racines de plants exposés à la sécheresse (Yang *et al.* 2005).

Harrak *et al.* (1999) ont identifié sur la tomate une Proline- Thréonine- et Glycine-rich Protéine (PTGRP) dont le gène est réprimé lors de la sécheresse et activé lors de la réhydratation. Les auteurs avancent que cette protéine répondant à la sécheresse aurait une localisation dans la paroi.

B.2.2.5.6.1 Modification de la paroi

L'expansion de la cellule nécessite un relâchement de la paroi. Les microfibrilles de cellulose ne sont pas extensibles. La structure de la paroi peut être modifiée par l'action de

complexes enzymatiques. Cela nécessite de couper les liaisons entre les microfibrilles de cellulose et les glycanes puis de rétablir les liaisons après glissement des fibres les unes par rapport aux autres. Les **expansines**, les **Xyloglucan endotransglucosylase/hydrolases** (XTH) et les **Pectines estérases** vont intervenir dans ce processus. Les expansines induisent directement l'extension de la paroi tandis que les XTH, les pectinases, et les glycanases ne le font généralement pas. L'analyse des génomes du riz et de *Arabidopsis* a permis d'identifier respectivement 80 et 38 ORF répartis en 4 classes d'expansines (Li *et al.* 2003). Sur des plants de maïs en condition de stress osmotique, le maintien de l'élongation des cellules de la région apicale des racines primaire et des feuilles est associé à l'augmentation de l'activité et des transcrits des expansines (Wu *et al.* 2001; Sabirzhanova *et al.* 2005). Les XTH vont cliver et reformer les liaisons entre les chaines de xyloglucan qui forment l'hémicellulose majoritaire de la paroi primaire. On dénombre 33 ORF chez *Arabidopsis* et 29 ORF chez le riz (http://labs.plantbio.cornell.edu/XTH/overview.htm). L'appellation de XTH est une nouvelle nomenclature remplaçant les anciennes appellations XET, XEH et EXGT. Au vu de leur rôle supposé, l'activité des XTH serait associée à l'élongation des parois. Si cette hypothèse n'est pas toujours validée, de nombreuses études tendent à la confirmer. L'analyse de l'expression des 33 gènes d'*Arabidopsis* montre que les membres ont des profils d'expression différents selon le tissu et les hormones appliquées (Yokoyama and Nishitani 2001). Chez des plants de soja exposés à un faible potentiel hydrique, l'activité en XET dans la zone d'élongation de l'hypocotyle était associée aux variations d'extensibilité mesurées (Wu *et al.* 2005).

B.2.2.5.6.2 Dégradation de la paroi

Les principales enzymes intervenant dans ce mécanisme sont les **Pectate lyases** -qui vont cliver la pectine pour donner des oligosaccharides-, les **endo-** et **exo-polygalacturonases**, et les **β-galactosidases** auxquelles se rajoutent les **Endo-1,4-β-glucanase** ou **cellulases.** L'expression de ces enzymes est également affectée par le stress hydrique (Bray 2004).

B.2.2.5.6.3 Biosynthèse de la paroi

La synthèse de paroi nécessite l'intervention de nombreux composés. Les protéines et enzymes déjà citées vont intervenir dans la formation d'une nouvelle paroi. Deux enzymes du complexe de synthèse de la cellulose seront indispensables à la synthèse de l'hémicellulose : la **Cellulose synthase** et les **Galactosyl/xylosyl transférases**. L'étude de mutants d'*Arabidopsis* a permis de mettre en évidence qu'un gène de la cellulose synthase jouait un

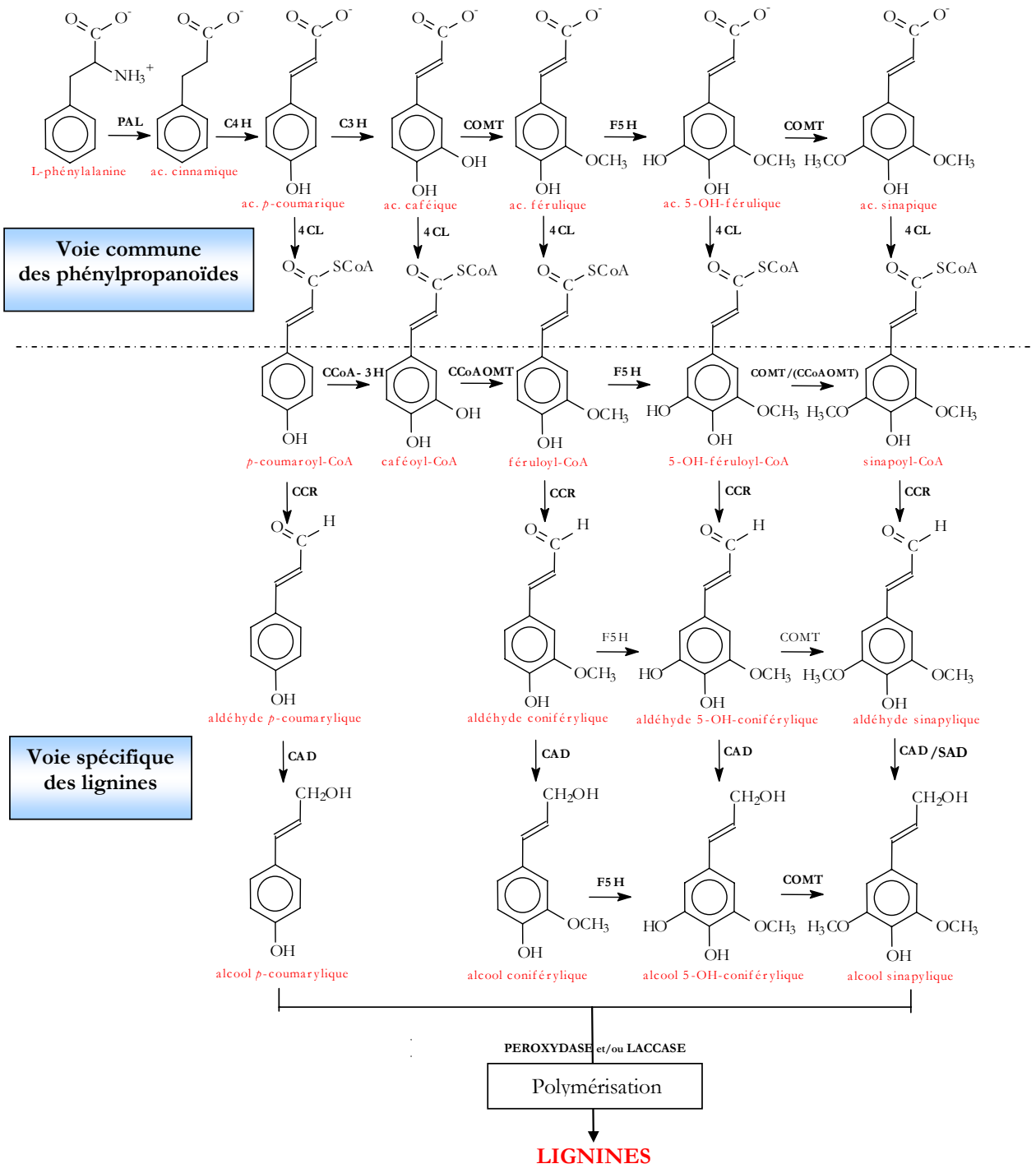


Figure B-37 : Voie de biosynthèse des lignines d'après Rech (2002).
Les enzymes mentionnées sont : la phénylalanine ammonia-lyase (PAL), la cinnamate–4-hydroxylase (C4H), la coumarate-3-hydroxylase (C3H), l'acide caféique/acide 5-hydroxyférulique 3-O-méthyltransférase (COMT), la férulate-5-hydroxylase (F5H), la 4-hydroxycinnamate CoA ligase (4CL), la p-coumaroyl-CoA 3-hydroxylase (CCoA-3H), la caféoyl-CoA 3-O-méthyltransférase (CCoAOMT), la cinnamoyl CoA réductase (CCR), l'alcool cinnamylique déshydrogénase (CAD) et l'alcool sinapylique déshydrogénase (SAD) .

rôle important dans la réponse à la sécheresse. En effet, les mutants allèliques lew2-1 et lew2-2 sont plus résistants aux stress hydriques (sécheresse, sel, osmotique) que le phénotype sauvage. Le gène touché est AtCesA8/IRX1 qui encode pour une sous unité du complexe de synthèse de la cellulose (Chen *et al.* 2005). La lignine est également un composé essentiel des parois des cellules végétales. Ce polymère modifie les propriétés physiques en apportant hydrophobicité et résistance mécanique. La voie de biosynthèse (Fig. B-37) a pour précurseur la L-phénylalanine. Les principaux enzymes sont la PAL, la C4H et la 4CL qui interviennent également dans la synthèse des composés phénoliques, et les C-COMT, CCoAOMT, C3H, F5H, CAD et CCR impliqués plus spécifiquement dans la régulation de la synthèse des polymères de lignine (cf. Fig. B-37 pour les abréviations). Dans une étude sur le maïs, Vincent *et al.* (2005) ont montré, en étudiant en particulier l'accumulation des protéines de **COMT**, que dans la feuille, la zone de lignification se situe immédiatement sous la zone d'élongation. En réponse au stress hydrique, cette zone se décale vers la base de la feuille et le niveau de lignine diminue. Les auteurs avancent que la diminution de la lignification serait une réponse adaptative à la sécheresse pour éviter une lignification des zones d'élongation. Dans sa thèse sur le pin maritime, Costa (1999) avait détecté l'accumulation de CCoAOMT dans les aiguilles lors du stress. C. Dubos (2001) avait détecté lors de sa thèse une diminution de l'expression d'un transcrit homologue à la COMT tandis que les transcrits de la PAL, la 4CL et la CAD étaient induits lors du stress.

B.2.2.5.6.4 Cytosquelette

Les microtubules interviennent entre autre dans la morphologie, la mobilité de la cellule ainsi que dans les transports intracellulaires de molécules. Une étude de Nyporko *et al.* (2003) souligne l'importance des microtubules dans la résistance au froid. L'exposition de cellules apicales de racines de blé à un stress osmotique violent provoque leur désintégration. Il semble qu'ils soient remplacés par des macrotubules de tubuline et des paracristaux de tubuline (Komis *et al.* 2002).

B.2.2.5.7 Régulation de la transcription

La recherche de gènes régulés permet d'améliorer nos connaissances sur les fonctions et voies métaboliques impliquées dans la réponse de la plante à son environnement. Chaque gène ne participe toutefois que pour une part limitée dans le processus de tolérance au stress. Certains gènes vont produire des facteurs de transcription (FT) qui vont activer un ensemble de gène. L'identification de gènes qui codent pour ces facteurs de transcription permet de

découvrir des réseaux de régulation et les gènes associés. Plusieurs familles de FT induits lors de la déshydratation ont déjà été identifiées. Ils sont classés en fonction de leur structure protéique, du type de stress les induisant et des éléments auxquels ils vont se lier sur le promoteur des gènes cibles (Ramanjulu and Bartels 2002).

B.2.2.5.7.1 proteines (b-ZIP)

Les facteurs b-ZIP (Basic leucine zipper) ont en commun un domaine de liaison à l'ADN riche en acides aminés basiques et adjacent à un domaine leucine zipper. La famille des **ABF** (ABRE binding factors) appartient à cette catégorie. Ils sont capables de reconnaître des éléments de régulation spécifique, les **ABRE** (ABA-responsive elements), situé en cis sur le promoteur du gène. Il existe deux motifs pour les éléments ABRE. Les **G/ABRE** partagent une séquence consensus (C/T)ACGTGGC que l'on retrouve dans l'élément Em1a (GGACACGTGGC) du gène Em du blé. L'expression du gène Em découvert au niveau de l'embryon a été étudié sur plusieurs espèces (Marcotte Jr *et al.* 1989; Bostock and Quatrano 1992; Vicient *et al.* 1998). Il est induit par l'ABA, les stress osmotiques, et sert de modèle dans les études de régulation de l'expression (Guan and Scandalios 1998). Les **C/ABRE** ont en commun la séquence CGCGTG. Les voies de régulations faisant intervenir ces deux types d'éléments ne sont pas les mêmes. Une revue détaillée sur les différents membres de la famille des ABF a été présentée par Choi *et al.* (2000). Parmi les facteurs de transcription capables de reconnaître les éléments G/ABRE, on peut citer EmBP-1, TAF-1 et GBF3.

B.2.2.5.7.2 Protéines Myb-like et Myc-like bHLH

Les FT de ces familles présentent une structure à deux hélices séparées par une boucle pour les Myc et un coude pour les Myb. La structure des **Myc** comportent une région hélice-boucle-hélice basique à motif leucine zipper. Le site de reconnaissance est du type CACATG (Abe *et al.* 1997). La séquence à l'origine de cette conformation pour les **Myb** est formée par la répétition imparfaite de trois motifs avec conservation de la position d'un tryptophane tous les 18-19 acides aminés. Il existe un site reconnaissance conservé dont la séquence est (T/C)AACTG (Urao *et al.* 1993). L'expression de ces facteurs est liée à la déshydratation. Chez *Arabidopsis* le gène Atmyb2 est induit par la déshydratation et les transcrits disparaissent avec la réhydratation. L'exposition au stress salin et à l'ABA provoque également l'expression de ce gène présentant une très forte homologie avec les protéines Myb (Urao *et al.* 1993). Il est possible de vérifier l'affinité d'une protéine à se fixer sur une séquence d'acides nucléiques spécifique. Dans le cas de Atmyb2, les auteurs ont utilisé un

vecteur d'expression pour produire la protéine et tester sa spécificité à se lier sur des acides nucléiques contenant des séquences de reconnaissance pour Myb. La découverte de gènes codant pour des FT spécifiques d'un type d'élément est également possible. Sur *Arabidopsis*, Abe *et al.* (1997) ont étudié le gène rd22 induit lors de la déshydratation ou de l'application d'ABA. Leur but était de vérifier qu'une région de 67 pb du promoteur, comportant un site de reconnaissance Myb et un site Myc, suffisait à la régulation du gène. Si un gène codant pour un facteur de transcription Myb était connu (ATMYB2), un gène codant pour une protéine Myc était nécessaire. Les auteurs ont utilisé une banque d'expression d'ADNc qu'ils ont criblé en utilisant la séquence du site de reconnaissance Myc. Ils ont alors pu démontrer par des constructions promoteur-GUS que le gène rd22 mis en évidence dans la réponse au stress hydrique pouvait être régulé à la fois par des FT des familles Myc et Myb.

B.2.2.5.7.3 Protéines AP2(APETALA2)/EREBP

Ce domaine est une caractéristique des facteurs du type **DREB** (Dehydratation-responsive element binding factor) et **CBF** (C-repeat binding factor)**.** Ceux ci vont pouvoir reconnaître et se lier à un motif « C-repeat » (CCGAC) des éléments de régulation **CRT** (C-repeat) ou DRE (Dehydratation-responsive element)(A/GCCGAC). On retrouve la présence de ces éléments dans les promoteurs des gènes COR (cold-regulated) induits par le froid.

Les éléments DRE ont été mis en évidence par Yamaguchi-Shinosaki *et al.* (1994) sur *Arabidopsis*. Deux gènes rd29A et rd29B, très similaires et proches sur le génome, répondaient différemment à la déshydratation, aux basses températures, à la salinité et à l'ABA. Les résultats laissaient supposer la présence de 2 éléments de régulation chez rd29A et un seul chez rd29B (Yamaguchi-Shinozaki and Shinozaki 1994). En étudiant les séquences promotrices et des plants transformés de tabac, les auteurs ont identifié un élément répondant à la déshydratation et contenant le motif TACCGACAT. Des protéines capables de se lier à ces éléments DRE sont connues chez *Arabidopsis*. Qiang *et al.* (1998) ont isolé les transcrits des gènes DREB1A et DREB2A dans une banque d'ADNc. L'expression constitutive de ces gènes chez des plants transgéniques de tabac a induit la transcription du gène rd29 ce qui confirme leur rôle de FT. Les auteurs ont également constaté que la tolérance au froid et à la dessiccation des plants transgéniques était améliorée. Par contre, si les deux gènes contrôlent l'expression de rd29, ils ne sont pas induits par les mêmes stress. DREB1A répond aux basses températures tandis que DREB2A est activé par la déshydratation (Qiang *et al.* 1998). Les FT n'induisent pas forcément la transcription d'un gène. Sur le maïs, Kizis *et al.* (2002) ont isolé des gènes codant pour les facteurs DBF1 et DBF2 du type AP2/EREBP. Contrairement à

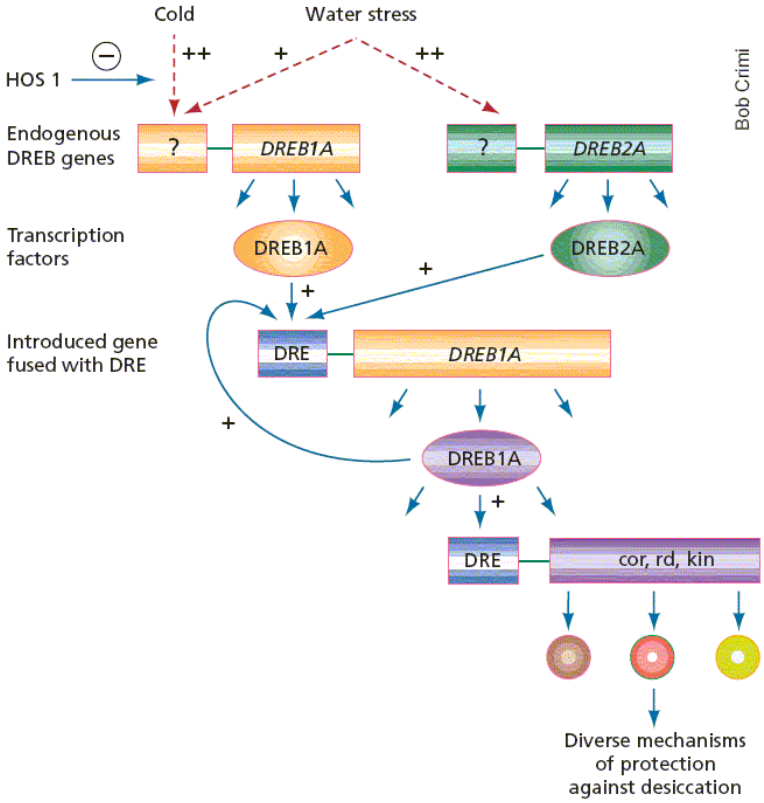


Figure B-38 : La surexpression d'un FT améliore la tolérance à la déshydratation, au froid et au sel chez Arabidopsis. La méthode présentée permet une surexpression d'un FT codant DREB1A en utilisant une construction où le promoteur du transgène est induit par le produit du transgène. (Smirnoff *et al.*, 1999)

DBF1, DBF2 n'est pas induit lors des stress couramment testés et son niveau de transcrits reste très faible. La surexpression de DBF2 a même inhibé l'activité du promoteur du gène rab17 porteur des éléments DRE2. Les auteurs avancent l'hypothèse d'une action de DBF2 sur une autre partie du promoteur (Kizis and Pages 2002). L'expression des FT dépend des conditions de stress appliquées à la plante mais aussi du tissu étudié. Ainsi lors de stress salins, DREB1D et DREB1F sont induit dans les racines, mais DREB2C, DREB2D et DREB2E ne sont induits que dans les feuilles. L'induction de l'expression de DREB2E n'est détectable que dans les racines de plants exposés à l'ABA (Sakuma *et al.* 2002).
En connaissant les FT et les séquences promotrices, il est possible de développer une nouvelle approche de construction de plants transgéniques. Si les mécanismes d'activation des gènes codant pour les FT ne sont pas connus, il est possible d'amplifier le pouvoir de régulation d'un facteur en construisant un transgène constitué du FT précédé d'un promoteur activé par le même facteur. Une telle boucle d'auto amplification a été proposée sur le modèle des DREB par Smirnoff *et al.* (1999) (Fig. B-38).

B.2.2.5.7.4 Protéines WRKY

Cette famille de gènes est l'un des groupes majeurs de régulateurs de transcription spécifiques des plantes. Elles se distinguent par la présence d'un domaine d'un 60^{aine} d'acides aminés hautement conservés parmi lesquels la séquence WRKYGQK suivie du motif « zinc binding » CCHH ou CCHC. Les éléments sur lesquels ces facteurs vont se fixer sont du type « W-box » dont la séquence est (T)TGCA(C/T) (Marè *et al.* 2004). Ces facteurs jouent un rôle important dans les réponses aux pathogènes, aux blessures, lors de la sénescence ou la germination, mais également dans d'autres stress (Seki *et al.* 2002). Dans le génome d'*Arabidopsis*, 72 gènes codant pour des protéines de la famille des **WRKY** ont été découverts par Dong *et al.* (Dong *et al.* 2003) et 49 d'entre eux sont régulés lors d'une attaque de pathogène. Les recherches sur ces facteurs n'en sont qu'à leur début. La fondation nationale pour la recherche (NSF) supporte un nouveau projet coordonné par Zhixiang Chen de l'Université de Purdue (http://www.btny.purdue.edu/WRKY/index.html). Ce projet a pour but de permettre : premièrement l'analyse de l'expression de ces gènes au cours du développement et de la réponse de la plante aux stress biotiques et abiotiques, deuxièmement de déterminer le rôle de ces gènes en observant le phénotype de mutants ne les exprimant pas ou les exprimant de façon constitutive, et troisièmement en analysant les interactions protéiques réalisées par ces WRKY.

B.2.2.5.7.5 Autres Protéines

D'autres protéines à fonction de FT impliquées dans les réponses aux stress ont été décrites dans la littérature.

Les facteurs **HD-ZIP** n'ont été identifiés que chez les plantes. Il s'agit de protéines avec un homéodomaine de fixation à l'ADN associé à un motif leucine zipper. Ils sont induits par la déshydratation mais réagissent différemment à l'ABA. Quelques représentants ont été identifiés sur *C. plantagineum* (CPHB-1 et -2) et d'autres sur *Arabidopsis* (ATHB-7, -6 et -12). Les gènes régulés par ce type de facteur ne sont pas encore identifiés.

Nous avons vu le rôle de protéines chaperonnes des Hsp. La séquence promotrice de ces protéines induites par des stress thermiques contient des éléments **HSE** sur lesquels peuvent se lier des facteurs d'activation en trans **HSF** (Heat Shock Factors). Les éléments HSE ont en commun une séquence consensus « nGAAnnTTCnnGAAn ». Chez *Arabidopsis*, Prändl *et al.* (1998) ont étudié le HSF3 et le HSF4 ayant chacun les caractéristiques de cette famille (régions DBD, HR-A et HR-B) mais dont le reste de la séquence ne présentait pas d'homologie. Le niveau de transcrit du HSF3 n'est pas affecté par les chocs thermiques tandis que le HSF4 est induit lors d'un stress de 2h à 37°C. Les auteurs ont testé si l'expression de ces HSF suffisait à induire des HSP de classe I. Seule la surexpression du HSF3 dans des plants transgéniques a permis d'induire ce type de HSP (Prandl *et al.* 1998). Il existe donc différents HSF répondant à des stress différents et activant des classes de HSP distinctes.

B.2.2.5.7.6 Histones et Méthylations

Jusqu'ici, nous avons parlé de régulations des gènes par des facteurs de transcription mais la régulation est possible en amont, par l'accessibilité au gène dans le noyau.

Le noyau est constitué par de la chromatine formée d'ADN et de protéines associées. Cette chromatine adopte plusieurs états selon le niveau de compaction de la molécule d'ADN. A ces états correspond un niveau d'accessibilité. L'ADN super enroulé d'une chromatine condensée ne pourra pas être transcrit. L'enroulement se fait en premier lieu autour du nucléosome qui est un octamère d'Histones nucléosomiques H4, H3, H2A et H2B ou « core Histones ». Une Histone H1 extranucléosomique permet la superspiralisation de l'enroulement initial (Maillet 1995). Il existe également des variants de ces histones (Huh *et al.* 1995) et d'autres protéines appelées « linker Histones » pouvant agir sur la structure de la chromatine et souvent induits lors de stress hydriques. L'expression d'un variant his1-s d'histone H1 a ainsi été étudié chez la tomate. Ce gène est induit précocement dans les feuilles lors d'un stress hydrique ou de l'application d'ABA. L'accumulation de la protéine de 31 kDa est également observée avec

une localisation nucléaire et dans la chromatine. L'intensité du stress appliqué n'a pas d'incidence sur le niveau d'expression enregistré (Scippa *et al.* 2000). Au vu des résultats et des connaissances sur les histones, les auteurs discutent d'un rôle structural possible de his1-s dans la protection de l'ADN lors du déficit hydrique ou un rôle de régulation des gènes. Ascenzi *et al.* (1997) ont isolé une « linked-Histone » His1-3 de 19 kDa qui diffère des histones connues His1-1 et His1-2 et présente des homologies avec une autre protéine H1-D. L'His1-3 est surexprimée spécifiquement lors du stress hydrique dans les feuilles et les parties aériennes d'*Arabidopsis* contrairement aux His-1,-2 et His4. Les auteurs n'ont par contre observé aucun changement du niveau de transcrits par les basses températures, les hautes températures ou les fortes concentrations en sel. Ce gène semble répondre spécifiquement au stress hydrique.

Les propriétés physico-chimiques des histones peuvent être changées par des modifications post traductionnelles. Les deux types de modifications les plus connues sont la méthylation et l'acétylation des queues N-terminales des H3 et H4, et N- et C-terminales des H2A et H2B. Les **HMT** (histones methyltransférases) et les **PRMT** (protein arginine methyl-transferases) sont des enzymes capables de méthyler les H3 et H4. L'acétylation est effectuée par des protéines de la famille des **HATs** (histones acetyl transferases), tandis que la réaction inverse est catalysée par les **HDAC** (histones déacétylases). Cet aspect de la régulation des histones par modification post-traductionnelle est reprise dans une revue de Peter Loidl (2004).

Les mécanismes de régulation par méthylation concernent également l'ADN. Chez *Arabidopsis*, l'action de la cytosine méthyltransférase **CMT3** et des autres cytosine méthylases joue un rôle essentiel dans l'inhibition de l'expression de régions chromosomiques (Finnegan *et al.* 1996; Finnegan and Kovac 2000).

B.2.2.5.7.7 Autres mécanismes de régulation

L'expression d'un gène et la production de la protéine sont déconnectés. Différentes isoformes peuvent être produites à partir d'un gène par épissage alternatif des exons. La quantité de protéines produites va dépendre du niveau de traduction des ARN messagers et de leur cycle de vie. L'activité de la protéine va ensuite dépendre des modifications post-traductionnelle des protéines qui peuvent être glycosylées ou phosphorylées.

Depuis peu de temps, les études sur les **microARN** suivent un engouement croissant. Les microARN (miRNA) sont de courtes séquences nucléotidiques capables comme les siRNA (Small interfering RNAs) de réguler négativement les ARNm cibles. Les miRNA sont codés par des gènes distincts du gène codant l'ARNm cible et sont obtenus à partir de précurseurs

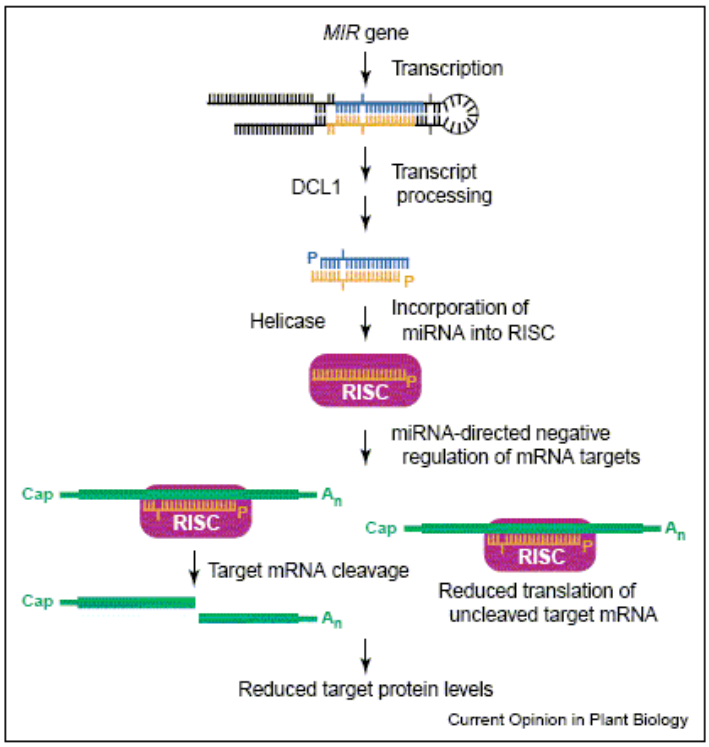


Figure B-39 : Représentation schématique de la formation des microARNs et leur fonction chez les plantes. (Dugas *et al.*, 2004)

dont la séquence forme une boucle en épingle à cheveux imparfaite (imperfectly paired hairpin precursors) (Fig. B-39). Sous l'action d'une Dicer ribonucléase, un fragment d'ARN double brin (dsRNA) est généré à partir du précurseur. Un des brins du miRNA mature est incorporé dans un complexe d'inhibition RISC (RNA-induced silencing complex). Lorsque ce complexe se fixe sur l'ARN messager cible, ce dernier est clivé, ce qui aboutit à une régulation négative de la traduction (Dugas and Bartel 2004; Kidner and Martienssen 2005). Dugas *et al.* (2004) rapportent une liste de miRNA identifiés ainsi que leur cible. Il apparaît que les cibles sont souvent des facteurs de transcription connus.

B.2.2.5.8 Perception et transduction du signal

Avant l'activation de gènes par des facteurs de transcription, les voies de transduction du signal commencent par la perception du signal suivi par la génération de seconds messagers comme l'Inositol phosphate et les ROS (reactive oxygen species). Classiquement, ces seconds messagers vont entraîner une variation de la teneur en Ca++ qui à son tour déclenche des cascades de phosphorylation jusqu'à l'activation finale de facteurs de transcription. Les gènes régulés par ces facteurs peuvent être impliqués dans la biosynthèse de molécules comme l'ABA qui agira comme signal sur de nouvelles voies de transduction. Plusieurs cascades de transduction du signal peuvent donc se succéder ou coexister en réponse à l'évènement initial –par exemple la baisse de disponibilité en eau-. Les études récentes ont montré qu'entre le stimulus initial et la réponse physiologique, la transduction du signal tient plus d'un réseau complexe d'activation et de répression que d'une cascade linéaire (Knight and Knight 2001). Un stress spécifique va donc pouvoir activer des cascades de transduction communes à différents stress et d'autres spécifiques, et beaucoup de gènes pourront être induit par plus d'un stimulus particulier. La revue de Xiong *et al.* (2002) reprend l'état des connaissances sur la perception et la transduction du signal au cours des stress froid, sécheresse et salin. Ces trois types de stress vont engendrer une déshydratation nécessitant des réponses communes mais ils vont aussi engendrer des perturbations plus spécifiques. On peut donc s'attendre à ce que la perception du stimulus initial se fasse par un réseau de senseurs.

B.2.2.5.8.1 Perception du signal

Les récepteurs primaires du stress restent difficiles à identifier. Les différents aspects du stress que la cellule serait susceptible de reconnaître par l'intermédiaire de capteurs regroupent la diminution du potentiel hydrique, la diminution de la pression de turgescence, la concentration de petites molécules, le changement du volume cellulaire ou des changements

de conformation de macromolécules (Bray 2002). Dans le cas des stress osmotiques, le gène **AtHK1** de *Arabidopsis* semble être le composant initial permettant de relier un changement de l'osmolarité à la chaîne de transduction du signal pour la régulation des gènes inductibles par la sécheresse (Urao *et al.* 1999). Ce gène est homologue au gène SLN1 de la levure codant pour une **histidine kinase** transmembranaire qui joue le rôle d'osmosenseur. La voie de perception et de transduction du signal pour l'osmolarité est connue pour cet organisme. L'osmosenseur permet d'activer la cascade des MAPK par l'intermédiaire de deux molécules phospho-relais YPD1 et SSK1 (Boudsocq and Lauriere 2005).

Des récepteurs aux hormones végétales ont également été mis en évidence. Nous avons déjà décrit la cascade de transduction suivant la fixation de l'éthylène sur son récepteur ETR1. Il existe également les récepteurs ERS1, ETR2, EIN4 et ERS2. Cette cascade de transduction a été décrite récemment par Chang *et al.* (2001). Inoue *et al.* (2001) ont découvert un récepteur histidine kinase CRE1 pour la cytokinine. Les senseurs primaires déclenchent souvent un influx massif de Ca^{++} dans la cellule. Les **canaux ioniques** membranaires pourraient aussi faire office de récepteur primaire. Il a été montré que l'ABA pouvait réguler l'activité de canaux potassiques et le potentiel de membrane de cellules racinaires (Roberts and Snowman 2000). La fermeture des stomates ferait intervenir la fixation de l'ABA sur des complexes **protéine-G** (GPCRs : G-protein coupling receptor) contenant des GTPases (Jones 2002).

B.2.2.5.8.2 Calcium et second messagers

A la suite de la perception initiale d'un des facteurs du stress par le récepteur, des messagers secondaires vont être les médiateurs de l'information. Ils vont transporter le signal du lieu de perception au lieu d'activation de la cascade de phosphorylation. Les principaux seconds messagers sont le Calcium, les diacylglycerols (DAG), l'inositol-1,4,5-trisphosphate (IP_3), les ROS et l'ABA (Xiong *et al.* 2002). L'inositol et les diacylglycerols sont les produits du clivage de la phosphatidylinositol bisphosphate par la phospholipase C. Au cours du stress hydrique, les plantes comme *Arabidopsis* peuvent augmenter la production d'IP_3 en induisant l'expression du gène PI5K qui code une phosphatidylinositol 4-phosphate 5-kinase catalysant la phosphorylation d'IPP (Mikami *et al.* 1998). Hirayama *et al.* (1995) ont montré que l'expression d'un gène codant une phospholipase-C était induite par la déshydratation et le stress salin. Les changements de concentration en Ca^{++} vont permettre l'intégration des différentes voies de transduction.

B.2.2.5.8.3 Cascade des MAPKs

Les cascades faisant intervenir les « mitogen-activated protein kinases » (MAPK)s sont probablement le mécanisme de transduction majeur chez les plantes. Une MAPK kinase kinase (**MAPKKK**) phosphoryle une MAPK kinase (**MAPKK**) qui à son tour phosphoryle une **MAPK**. Plusieurs gènes codant pour les MAPKKK (AtMEKK1-4), pour les MAPKK (AtMKK) et pour les MAPK (AtMPK1-6) (Mizoguchi *et al.* 1993) ont été isolés chez *Arabidopsis* (Ichimura *et al.* 1998). Asai *et al.* (2002) ont réussi à mettre en évidence une cascade complète de MAPKs intervenant dans la réponse de *Arabidopsis* aux pathogènes depuis les récepteurs LRR aux éliciteurs jusqu'à l'expression de facteurs de transcription. Les auteurs ont utilisé des mutants présentant des déficiences de cette voie de transduction. L'expression de ces MAPKs est régulée par les stress abiotiques (Mizoguchi *et al.* 1996).

B.2.2.5.8.4 Protéines régulées par le Calcium

Cette catégories regroupe la classe des calmodulines, des « calcium-dependent protein kinases (CDPKs) et des phosphatases régulées par le calcium. Il existe plusieurs isoformes pouvant avoir des spécificités pour un stress comme les AtCDPK1 et AtCDPK2 impliquées dans les stress salins et la sécheresse chez *Arabidopsis* (Urao *et al.* 1994). Chez le riz, les transcrits du gène OsCDPK7 qui code pour une CDPK sont accumulés dans les tissus vasculaires des racines lors du stress froid, salin et de la sécheresse. L'expression constitutive chez des plants transgéniques aboutit à une accumulation de transcrits dans les mêmes régions (Saijo *et al.* 2001).

B.3 Le stress hydrique chez les arbres

Les arbres, composante centrale de l'écosystème terrestre, sont des organismes ligneux longévifs pouvant atteindre plusieurs dizaines de mètres.

B.3.1 Une problématique spécifique

B.3.1.1 Particularités physiologiques

Ces organismes sont caractérisés par une croissance secondaire exacerbée au niveau du cambium où les cellules se différentient pour donner le phloème et le xylème secondaires. Ce dernier est lignifié et forme le bois de cœur et l'aubier. Du fait de leur grande taille et des temps de génération long, les arbres sont difficiles à étudier avec les approches modernes utilisées en génétique. Ils posent des questions de recherche spécifiques et leurs

caractéristiques différentes des autres plantes supérieures suggèrent qu'ils apportent des réponses physiologiques différentes. La hauteur de l'arbre est un des premiers facteurs influant sur les mécanismes physiologiques mis en œuvre. La sève doit parvenir à des hauteurs dépassant les 40 mètres et pouvant atteindre 110 mètres chez le Séquoia. A ces hauteurs, les organes foliaires subissent un stress hydrique et doivent répondre en abaissant leur potentiel osmotique et en fermant leurs stomates. Les valeurs de $\delta^{13}C$ et de potentiel atteignent chez cet arbre les valeurs les plus extrême mesurées chez les plantes (Koch *et al.* 2004). Nous avions expliqué (cf. B.2.1.1) que le potentiel hydrique Ψ_w dépendait du bilan de la pression osmotique Ψ_s et de la pression mécanique Ψ_p exercée par les parois cellulaires qui tend à faire sortir l'eau des cellules. D'un point de vue global de l'arbre par rapport au sol, la force de gravité s'exerçant sur une colonne d'eau de plusieurs dizaines de mètres ne peut plus être négligée et l'équation simplifiée du potentiel hydrique doit tenir compte de ce paramètre Ψ_g (Kramer 1983).

$$\psi_W = \psi_S + \psi_P + \psi_g$$

La taille importante du réseau vasculaire chez les arbres associé aux forces de tensions exercées sur la colonne d'eau et sur les parois des vaisseaux conducteurs souligne la nécessité d'une bonne régulation des pertes en eau et de réponses adéquates pour répondre au stress hydrique et éviter la cavitation (Cruiziat and Tyree 1990).

Une autre particularité est l'importance des mycorhizes chez les arbres qui forment une association symbiotique entre le mycélium d'un champignon et les racines de la plante. Elles vont jouer le rôle des poils absorbants et permettre à l'arbre d'accéder à l'eau présente dans le volume de sol occupé par le mycélium (Heller *et al.* 1993). Cette symbiose qui joue un rôle majeur dans les mécanismes de réponse de l'arbre à la sécheresse a été encore peu étudiée (Nouaim and Chaussod 1996; Garbaye 2000).

B.3.1.2 Ressources moléculaires

Le développement de recherches sur ces organismes est complémentaire des approches réalisées sur les plantes modèles et de grande culture. Le peuplier est venu rejoindre la liste des plantes modèles déjà citées. Les ressources génétiques disponibles se sont grandement enrichies avec le séquençage du génome du peuplier (Brunner *et al.* 2004) (http://genome.jgi-psf.org/Poptr1/Poptr1.home.html). Pour les résineux, si aucune espèce n'a le statut de plante modèle, les ressources génétiques se sont largement développées pour les espèces d'intérêt économique du genre Pinus comme le pin loblolly (*Pinus taeda*), le pin de Monterey / pin à encens (*Pinus radiata*) et le pin maritime (*Pinus pinaster*) ou l'épicéa (*Picea abies*).

Les projets de séquençage d'EST ont permis de construire une base de données conséquente de séquences de gènes exprimés. Les données publiées ont été regroupées par le TIGR qui donne accès aux séquences ainsi qu'à leur assemblage bioinformatique et leur annotation à l'adresse http://www.tigr.org/tdb/tgi/plant.shtml. Des bases dédiées sont également accessibles pour le peuplier http://poppel.fysbot.umu.se/, le pin *taeda* http://www.ccgb.umn.edu/biodata/ et le pin maritime http://cbi.labri.fr/outils/SAM2/COMPLETE/index.php.

La bibliographie sur l'étude de la sécheresse chez le peuplier est riche de références. La problématique a été abordée du point de vue physiologique (Chen *et al.* 2002; Johnson *et al.* 2002; Hukin *et al.* 2005; Marron *et al.* 2005; Monclus *et al.* 2005) comme moléculaire (Caruso *et al.* 2002; el-Khatib *et al.* 2004; Marjanovic *et al.* 2005). Cette espèce présente de nombreux avantages par rapport aux autres ligneux. Outre le fait que nous disposions de la séquence de son génome et de séquences d'EST, cette espèce a une croissance très rapide (certaines espèces exploitables à 15 ans) comparativement au pin (35-40 ans) ou au chêne (100-140 ans). L'obtention de plants transgéniques est également une technique bien maîtrisée sur le peuplier et les approches de biologie moléculaire adoptées chez *Arabidopsis* sont transférables à cet arbre (Strauss *et al.* 2004). Cette technique est aussi réalisable chez le pin maritime ou chez le pin *taeda* (Tang *et al.* 2005). Chez les conifères nous ne disposons pas de génome entièrement séquencé mais de banques d'EST et des cartes génétiques saturées permettant la comparaison des génomes de nombreuses espèces du genre Pinus (ex : (Chagne 2004)). Les grandes thématiques de recherche chez les résineux et les arbres en général se rapportent à la formation et la qualité du bois, la résistance aux pathogènes et la réponse au stress hydrique. Différents projets de séquençage d'EST se sont succédés sur le pin *taeda*. Un premier projet financé par la NSF avait pour objectif d'identifier des gènes impliqués dans la formation du bois (http://pinetree.ccgb.umn.edu/). Un nouveau projet en cours a pour but d'identifier des allèles pour des gènes candidats d'intérêt économique (http://dendrome.ucdavis.edu/adept/index.html). Il existe également un projet axé sur les réponses du transcriptome racinaire aux conditions environnementales (http://fungen.botany.uga.edu/Projects/Pine/PineInfo.htm) et un projet de la NSF sur les embryons somatiques et zygotiques (http://www.tigr.org/tdb/e2k1/pine/index.shtml).

Ces projets ont permis de démarrer des études du transcriptome en utilisant des microarrays sur la réponse au stress hydrique (Heath *et al.* 2002). Dans la bibliographie, les recherches sur la réponse au stress hydrique chez les conifères couvrent la plupart des paramètres physiologiques et catégories de gènes exposés précédemment. Sur le plan physiologique, on pourra citer par exemple une étude sur la comparaison de la résistance au stress osmotique de

huit provenances de *Pinus halepensis* (Calamassi *et al.* 2001) et la thèse de Froux (2002) sur les caractéristiques hydrauliques, la régulation stomatique et l'efficience d'utilisation de l'eau de quatre espèces.

Sur le plan moléculaire, Sathyan (2004) a combiné les techniques de PCR quantitative et d'hybridation par Northern blot pour identifier des gènes différentiellement exprimés lors de la sécheresse entre deux populations de pin d'Alep et entre deux populations de pin *taeda* (Sathyan 2004). L'auteur a notamment mis en évidence la régulation de gènes codant pour des facteurs MYB, des LEA ou la cyclophiline.

B.3.2 Le stress hydrique chez le pin maritime

Si le programme d'amélioration du pin maritime a débuté à l'INRA dans les années 60 sur des critères de rendement et de rectitude, l'étude de la réponse aux contraintes hydriques a débuté plus récemment. Différentes études, que nous avons résumées dans ce chapitre, ont permis d'aborder cette question sous différents aspects : physiologique/écophysiologique, génétique et moléculaire.

B.3.2.1 Les approches physiologiques

Les travaux précurseurs de Guyon (1980 et 1982) ont confirmé la présence d'une variabilité des individus au sein de cette espèce. Les données de mesures morphologiques sur 18 provenances ont permis de distinguer les groupes : Atlantique, Corse, Maroc-Espagne et Tunisie. Les mesures de pression de sève, de transpiration et de teneur en eau des provenances sur des stations différentes indiquent que la variabilité concerne également le comportement hydrique. Les variétés les plus opposées appartiennent aux groupes atlantiques, tunisiens et marocains qui diffèrent sur leurs stratégies de régulation de la transpiration et leur potentiel de base en condition de sécheresse (Guyon 1980; 1982). La sélection naturelle a favorisé les individus les plus adaptés aux conditions climatiques de leur région d'origine en privilégiant la limitation des pertes en eau au détriment de la croissance pour les régions les plus xériques et les meilleurs compétiteurs (les plus efficaces pour l'utilisation de l'eau) dans les zones humides (Alia *et al.* 1993).

La compréhension du comportement écophysiologique du pin maritime landais vis-à-vis de la disponibilité en eau a été améliorée par le travail de El Hadj Moussa (1989) dans le cadre de sa thèse. L'estimation des apports externes, de l'interception par le couvert, de la transpiration de la strate arborescente et celle du sous-bois, et l'évolution du stock hydrique du sol, ont permis d'établir un bilan hydrique à l'échelle de la parcelle et de l'arbre. Les résultats

montrent une régulation de la conductivité hydraulique et de la conductance stomatique avec une modulation de la transpiration suivant les disponibilités en eau. L'auteur indique que le comportement du pin des landes est du type « évitement ».

Les travaux se sont ensuite répartis entre l'aspect régulation stomatique et l'aspect régulation osmotique.

L'ouverture des stomates va dépendre de nombreux facteurs et aura un impact sur l'assimilation photosynthétique du carbone. La sensibilité stomatique à la sécheresse est très élevée chez le pin maritime comparée à d'autres espèces forestières comme le chêne (Guehl *et al.* 1994). En condition de déficit hydrique, la conductance stomatique chute et l'efficience d'utilisation de l'eau augmente (Picon *et al.* 1996). Cependant la conductance stomatique va également dépendre de la conductance du système conducteur des branches et des tiges qui sera fonction de la longueur de ces organes (Warren and Adams 2000). La conductance stomatique diminue avec l'accroissement de la taille ou de l'âge de l'arbre (Delzon *et al.* 2004). La mesure de la discrimination isotopique du ^{13}C ($\delta^{13}C$) dans les tissus reflète les conditions hydriques et les propriétés hydrauliques du continuum sol-plante-atmosphère lors de l'assimilation du carbone. Nguyen-Queyrens *et al.* (1998) ont mesuré le $\delta^{13}C$ de la cellulose du bois des cernes compris entre 1972 et 1991 sur des pins de 26 ans de provenance landaise et marocaine. Les valeurs de $\delta^{13}C$ enregistrées pour les cernes successives ont des variations suivant globalement les évolutions climatiques. Les périodes de sécheresses correspondaient à une augmentation du $\delta^{13}C$. La provenance marocaine avait des valeurs plus négatives de $\delta^{13}C$ quelque soit l'âge des cernes. Comme nous l'avons déjà signalé, ce ne sont donc pas forcément les provenances des régions les plus arides qui utilisent au mieux l'eau disponible. Les auteurs ont avancé les hypothèses que la provenance marocaine dont le système racinaire est plus développé serait plus évitante. Brendel *et al.* (2002) ont identifié des QTL (Quantitative trait Loci) pour le $\delta^{13}C$ et ont montré qu'il s'agissait d'un paramètre ayant une héritabilité significative. Il existe donc un déterminisme génétique de la discrimination isotopique du ^{13}C.

Le potentiel osmotique est régulé lors du stress hydrique et est variable entre génotypes issus du croisement des provenances corse et landaise (Nguyen-Queyrens *et al.* 2002). Un potentiel osmotique bas va permettre le maintien de la turgescence à des potentiels hydriques bas. L'application d'un stress osmotique par ajout de PEG a induit une diminution du potentiel osmotique des racines de plants marocains et landais. Le stress appliqué était de

courte durée (6 jours maximum) pour une intensité comprise entre -0.03MPa et -0.8MPa. Les provenances marocaines ont maintenu une croissance racinaire plus élevée et ont présenté un ajustement osmotique plus efficace que les provenances landaises (Nguyen and Lamant 1989b). Cette dernière caractéristique a été confirmée par Loustau *et al.* (1995) en exposant des plants d'origines landaises, espagnoles et marocaines à des conditions salines. L'impact du stress osmotique sur la croissance racinaire et le grandissement cellulaire a été étudié en détail par Triboulot (1996) dans le cadre de sa thèse.

Sur les arbres adultes des Landes de Gascogne, il semble que le potentiel hydrique foliaire ne dépasse jamais une valeur seuil de -2.0MPa indépendamment de la gravité et de l'intensité du stress hydrique (Delzon *et al.* 2004).

Nguyen et Lamant ont analysé des osmolytes pouvant être responsables de cet ajustement osmotique. Les auteurs ont mesuré la concentration en K^+ de plants soumis à un stress osmotique PEG. Les variations en concentration observées (+30mM) dans les racines de la provenance marocaine par rapport à la provenance landaise peuvent s'expliquer par la simple déshydratation des cellules. Le K^+ n'est pas responsable des variations de potentiel hydrique observées (Nguyen and Lamant 1989a). Les auteurs ont également conduit une expérience pour mesurer l'accumulation du D-pinitol et du myo-inositol entre ces deux provenances. Les auteurs ont montré tout d'abord que le niveau de pinitol était 2 à 3 fois plus élevé dans des conditions de bonne alimentation en eau pour la variété marocaine que pour la landaise, tant au niveau des racines qu'au niveau des parties aériennes. Le pinitol était accumulé de façon similaire entre les deux provenances lors du stress. Les concentrations de Myo-inositol semblaient augmenter uniquement dans les racines des plants marocains (Nguyen and Lamant 1988).

Une étude de Lamhamedi *et al.* (1992) sur la colonisation du système racinaire du pin maritime par des champignons ectomycorhiziens et l'effet sur le potentiel hydrique à l'interface sol-racine souligne l'impact de telles symbioses sur l'amélioration de la résistance de l'arbre à des conditions sèches.

Les travaux les plus récents de l'équipe d'écophysiologie INRA-EPHYSE, concernent la modélisation des effets des changements climatiques globaux sur la production des forêts françaises (Loustau *et al.* 2005). Si les modèles prévoient une augmentation de la production sur la période 2030-2050 suivie d'un plateau ou d'un déclin en 2070-2100, les auteurs soulignent aussi l'importance des variations régionales et des espèces cultivées. L'élévation de température et l'allongement de la saison de végétation qui en découlent, profiteront essentiellement aux espèces à feuilles caduques et non au pin maritime. Les espèces ont des

réponses différentes aux augmentations de la teneur en CO_2 et au déficit hydrique, le chêne ferme ses stomates si les concentrations augmentent ce qui n'est pas le cas du pin maritime. Finalement, les modèles étudiés prévoient que l'effet positif de l'accroissement en [CO_2] sur la production sera inhibé par la sécheresse dans le sud de la France.

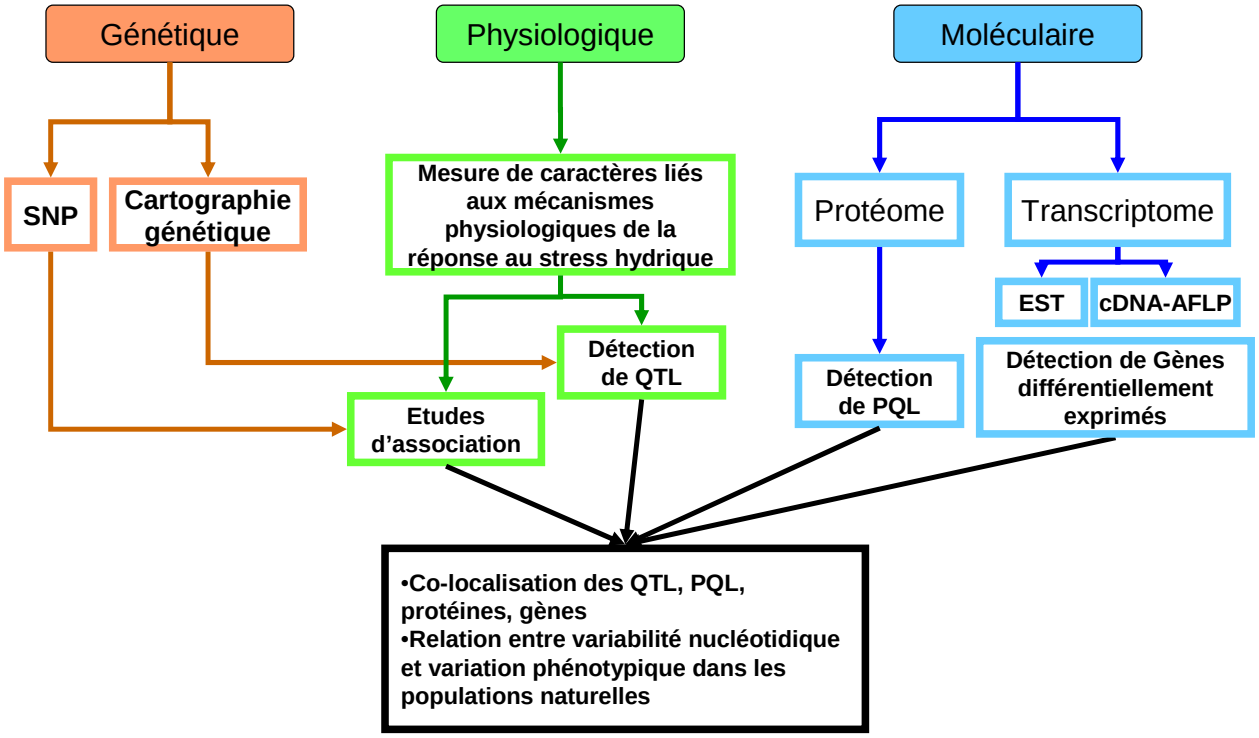
B.3.2.2 Les Approches génétiques et moléculaires

Les résultats des recherches citées ci-dessus ont montré qu'il existe une variabilité de réponse au stress hydrique entre les différentes provenances. De plus, il apparaît qu'une part de cette variabilité a un déterminisme génétique. L'UMR BioGEco a donc lancé des programmes de recherche dont le but est d'identifier les gènes impliqués dans la réponse à la sécheresse. Cela permettrait dans le futur de : 1) sélectionner des variétés sur leur capacité à résister à la sécheresse tout en conservant une bonne productivité et une bonne qualité du bois. 2) fournir des critères diagnostiques (diversité génétique dans ces gènes) pour évaluer la potentialité d'évolution des populations d'amélioration et des populations naturelles aux changements climatiques. La démarche qui a été adoptée pour la recherche du déterminisme moléculaire de la réponse au stress hydrique est schématisée à la figure B-40.

B.3.2.2.1 Cartographie et détection de QTL

Une carte génétique a été construite par Costa (1999) au cours de son travail de thèse et a servi de support pour la localisation de QTL. La descendance utilisée était constituée de 200 individus F2 issus de l'autofécondation d'un hybride entre 2 arbres élites Corse et Landais. Les plants âgés de 2 ans ont subi un stress hydrique progressif par simple arrêt d'arrosage pendant 6 semaines. Costa (1999) a recherché des QTL pour un ensemble de caractères décrivant les comportements hydriques (potentiel hydrique de base (Ψ_w), consommation en eau (C), potentiel osmotique (Ψ_s), teneur relative en eau (TRE)), les échanges gazeux (conductance stomatique (g_s), transpiration (E), photosynthèse maximale (A_{max}), $\delta^{13}C$) et la croissance (croissance en hauteur, biomasses aériennes et racinaires, analyses minérales (N, P, K)) avant et toutes les semaines pendant le stress. Les QTL détectés avaient des effets relativement faibles et se concentraient essentiellement sur les groupes de liaison 6, 8 et 12. Une des régions génomiques co-localisant des QTL pour des caractères mesurés au cours du stress (Ψ_w, TRE, g_s, A_{max} et E) semblait être particulièrement importante dans le contrôle génétique de la réponse au stress hydrique chez le pin maritime. Deux QTL pour le $\delta^{13}C$ (mesuré sur les aiguilles) ont été détectés. Le QTL majeur expliquait 12.3% de la variabilité du caractère et se positionnait sur le groupe 6.

Figure B-40 : Connexions entre les approches génétiques, physiologiques et moléculaires utilisées pour étudier la réponse à un stress hydrique chez le pin maritime.



Une étude complémentaire portant sur la discrimination isotopique du ^{13}C dans le bois et la largeur des cernes (croissance radiale) a été menée par Brendel *et al.* (2002). Un premier dispositif expérimental a été utilisé pour estimer la variabilité, l'héritabilité et les corrélations génétiques des caractères étudiés. Il s'agissait d'un demi-diallèle 12x12 de 564 arbres dont les parents avaient été sélectionnés pour leur croissance et leur rectitude. Les auteurs ont mis en évidence une héritabilité significative des deux caractères, ce qui laisse envisager des gains possible en croissance et en efficacité d'utilisation de l'eau (WUE) par croisement de génotypes favorables. Pour les auteurs, d'après le modèle de Farquhar, la corrélation positive constatée entre le $\delta^{13}C$ et la croissance suggère que la variation de WUE entre les arbres aurait pour origine une différence d'assimilation photosynthétique plutôt qu'une régulation stomatique différente. Une famille de plein-frères, constituée de 202 individus a permis d'étudier l'architecture génétique de ces caractère complexes. Quatre QTL ont été détectés pour le $\delta^{13}C$ expliquant près d'un tiers de la variation phénotypique observée. En revanche, aucune co-localisation entre les QTL de $\delta^{13}C$ et ceux de la largeur moyenne des cernes n'a pu être mise en évidence, ce qui suggère l'absence de contrôle génétique commun aux deux caractères.

B.3.2.2.2 Identification des protéines impliquées dans la réponse au stress hydrique

La première étude réalisée par Costa (1999) a porté sur les individus F2 utilisés pour la cartographie. Le protéome de 114 de ces plants a été analysé avant et après application du stress par électrophorèse bidimensionnelle des protéines extraites sur les aiguilles. L'objectif de cette expérience était d'identifier et de caractériser des protéines candidates pour la réponse au stress hydrique. Il a étudié 163 protéines (spots) et caractérisé 43 d'entre elles par microséquençage interne (Costa *et al.* 1998). Sur les 163 spots, 120 spots (dont 75 sous-exprimés et 45 surexprimés) ont montré une réponse au stress. La caractérisation des protéines d'aiguille du pin maritime qui répondent au stress hydrique, a montré que différents métabolismes cellulaires étaient impliqués : métabolisme du carbone, de l'azote, détoxication de la cellule, métabolisme de la paroi cellulaire, protéines chaperonnes, de choc thermique ou actine.

Parallèlement, le regroupement des spots en fonction de leur profil d'expression a permis d'obtenir un groupe de protéines accumulées et un autre de protéines dont la quantité diminue entre les deux conditions. Chaque groupe est lui-même divisé en sous-groupes de protéines dont l'évolution de la quantité évolue de façon similaire. Un sous-groupe intéressant

comprenait six variables physiologiques (Ψ_w, TRE, g_s, A_{max}, E, C) auxquelles se rattachaient des protéines de fonction inconnue et dont la quantité diminuait dans la condition stressée, montrant une implication possible de ces protéines dans la réponse au stress.
La disponibilité d'une carte génétique du pin maritime a rendu possible l'identification de PQL (Protein Quantity Loci) dont certains co-localisaient avec des QTL des caractères physiologiques tel que celui correspondant à la discrimination isotopique du carbone ($\delta^{13}C$).
Parmi les protéines identifiées et dont la quantité est corrélée négativement ou positivement aux mesures physiologiques on trouve des HSPs, des Cu/Zn SOD, la disulfide isomérase, les SAM synthases, l'isoflavone réductase, la glutathione peroxydase, la CCoA-OMT et des protéines de la photosynthèse. La conductance stomatique (g_s) était le paramètre présentant le plus de relations significatives avec l'intensité des spots dans la condition contrôle, tandis que dans la condition stressée, le paramètre interférant le plus était la Teneur Relative en Eau (TRE).

Une autre étude, réalisée par Dubos (2001), portait sur de jeunes plants cultivés en hydroponie et soumis à un stress PEG. Les protéines ont été extraites et des variations significatives de quantité ont été trouvées sur les aiguilles (34 spots (8%)), les tiges (51 spots (14%)) et les racines (53 spots (14%)). Dans chaque tissu, des spots présentaient des régulations positives (Aiguilles : 16, Tiges : 32, Racines : 32) lors du stress et d'autres une régulation négative. Certaines variations étaient d'ordre qualitatif avec apparition ou disparition lors du stress. Parmi ces protéines certaines étaient communes à celles identifiées par Costa comme la rubisco activase, une Hsp de faible poids moléculaire, des SAM synthase et une protéine chaperonne Hsp70.

B.3.2.2.3 Identification des transcrits impliqués dans la réponse au stress hydrique

A l'issu de l'approche protéomique, Dubos (2001) a démarré une thèse sur l'étude de la réponse au stress hydrique de jeunes plants de pin maritime en menant une investigation sur la régulation du transcriptome. Si l'arrêt d'arrosage de l'expérience précédente reproduisait assez bien la situation des plants lors d'une sécheresse estivale, il avait pour inconvénient de ne pas maîtriser de façon précise l'intensité du stress et les conditions de culture appliqués. De plus, l'accès au système racinaire n'est pas aisé et l'absence de mycorhizes n'est pas garantie. Dubos a donc opté pour une culture hydroponique et l'application d'un stress osmotique par adjonction de Polyéthylène Glycol. Dans sa thèse, Dubos a étudié un factoriel croisant intensité (-0.08 MPa ; -0.15 MPa ; -0.30 MPa ; -0.45 MPa ; -0.60 MPa ; Terre) et durée (0 ; 1 ;

10 ; 21 ; 45 jours) de stress hydrique sur des plantules d'origine landaise âgées de 3 semaines. Afin d'identifier les variations d'expression des transcrits dans les parties aériennes et les parties racinaires, une technique d'ADNc-AFLP a été adoptée. Elle a permis la détection de 81 gènes régulés au cours du stress. Des analyses d'expression par Northern et RT-PCR ont été réalisées sur 40% des gènes et ont confirmé leur expression différentielle. Les gènes de fonction connue impliqués dans la réponse au stress hydrique font partie du métabolisme des sucres, de la photosynthèse, de la synthèse des parois et des mécanismes de défense comme la PR-10 décrite comme régulée lors des stress biotiques et abiotiques (Dubos and Plomion 2001; Dubos *et al.* 2003). Au niveau des racines, des transcrits accumulés lors du stress sont homologues à l'aldéhyde déshydrogénase (métabolisme des lipides), la pLP3 (protéine induite lors du stress chez *P. teada*), une glycolate oxidase, une ATPase, une ACC oxydase (éthylène), une Xyloglucan Endotransglycosylase (Parois) ou une S-adenosyl-L-homocysteine hydrolase (SHH déjà décrite lors du stress)(Dubos and Plomion 2003). Dans les gènes réprimés lors du stress, on trouve des homologues à la COMT (caffeic acid ortho-methyl transferase)(synthèse lignine), une protéine putative RING zing finger (facteur de transcription), une pre-protéine translocase SecA, une protéine de la famille des LEA hydrophobes ou encore une histone H2B.

Au niveau des aiguilles, Dubos *et al.* (2003) ont détectés 25 transcrits présentant des différences significatives d'expression. Les amplitudes de régulation mesurées allaient de 1.18 à 2.14 pour les 11 transcrits induits, et de 1.49 à 16.6 pour les 14 transcrits dont l'expression diminuait lors du stress. Les fonctions identifiées avaient trait aux mécanismes de défense de la plante (flavonoid 3′,5′-hydroxylase (F3′H), dihydroflavol 4-reductase (DFR), pathogenesis-related protein, allene oxide synthase (cytochrome P450 protein), glucan endo-1,3-β-glucosidase) et étaient induits lors du stress. Tandis que les gènes réprimés étaient plutôt reliés aux gènes de la photosynthèse (preprotein translocase SecA, hlorophyll a/b-binding proteins type 4 et type 1, small subunit of the ribulose-biphosphate carboxylase).

B.4 Stratégies et techniques pour la détection des gènes impliqués dans l'adaptation du pin maritime à la sécheresse.

Pour la découverte des gènes dont le rôle est majeur dans la relation de la plante à son environnement, l'approche expérimentale se base sur la comparaison de génotypes contrastés placés dans des conditions où l'expression des gènes d'intérêts a un impacte sur le phénotype observé.

Lorsqu'ils sont disponibles, l'utilisation de mutants est une possibilité (Nadeau and Frank 2000). Nous avons cité plusieurs travaux adoptant cette démarche pour l'identification de gènes spécifiques. Cependant, cela ne permet la découverte que d'un petit nombre de gènes. Une autre option consiste à utiliser la variabilité des populations naturelles pour le caractère étudié. Dans le cas du pin maritime, nous disposons de variétés très contrastées vis-à-vis de leur réaction à un déficit en eau. L'étude de ces écotypes en situation de stress hydrique devrait permettre d'observer une plasticité moléculaire différente entre provenances. La notion de plasticité phénotypique se définit comme l'aptitude d'un génotype à donner différents phénotypes en fonction de l'environnement. Par analogie, nous définissons la plasticité moléculaire comme la propriété d'un génotype à modifier l'expression de son génome en réponse à des environnements distincts.

L'utilisation des QTL (Quantitative Trait Loci) pour la détection des régions du génome expliquant une partie de la variabilité du caractère étudié, est aussi une alternative séduisante (Borevitz and Chory 2004). Une des difficultés de ce type d'étude tient d'une part à la transférabilité des QTL détectés d'un dispositif expérimental à un autre, et d'autre part à l'identification d'un gène précis, étant donné l'intervalle de confiance autour de la position la plus vraisemblable d'un QTL. Les études d'associations permettent d'étudier de façon plus fine la corrélation entre un polymorphisme nucléotidique d'un gène candidat et la variation du caractère étudié (Cardon and Bell 2001). Ce type d'étude est actuellement en cours sur le pin maritime et le stress hydrique dans le cadre de la thèse d'Emmanuelle Eveno.

La découverte des gènes peut se limiter pour un tissu donné au génome exprimé. La transcriptomique regroupe un ensemble de techniques d'étude des transcrits. La comparaison des niveaux d'expression d'un gène pour les écotypes et dans les conditions testées permet de déterminer si une variabilité d'expression existe pour ce gène. L'analyse du profil d'expression permettra d'identifier des gènes candidats expressionnels. Certaines techniques comme la PCR quantitative vont permettre d'étudier précisément un gène mais ne sont pas adaptée à l'étude à de milliers de gènes exprimés dans un tissu. La réalisation de banques d'ADNc offre la possibilité d'insérer des ADNc dans des vecteurs puis de les cloner dans des bactéries. Une approche particulièrement intéressante est la création de banques suppressives soustractives par hybridation (SSH) (Diatchenko *et al.* 1996). Cette méthode permet d'éliminer la redondance et donc d'enrichir la banque obtenue en transcrits différentiellement exprimés. Un séquençage des clones permet d'identifier les gènes plus spécifiquement exprimés dans un écotype et/ou une condition donnée.

Les analyses d'expression par hybridation de cibles sur des sondes constituent une autre possibilité. Elles permettent l'étude simultanée de l'expression de milliers de gènes. Quelque soit le type de filtre (dépôt bactérien, ADNc ou oligomères), il faut néanmoins avoir accès à la séquence des sondes déposées. La génération d'EST s'est positionnée depuis une dizaine d'années à la fois comme une alternative et un complément au séquençage complet de génomes (Rudd 2003). Les projets de séquençage se sont développés avec l'utilisation des microarrays. La production massive de ce type de séquences courtes de basse qualité nécessite l'utilisation de moyens bioinformatiques de traitement des données. Parallèlement au développement de ces techniques, les outils bioinformatiques ont également fortement progressés pour répondre aux besoins croissant de traitement des données générées. Cela concerne le traitement des séquences, leur assemblage pour reconstituer la séquence complète des gènes, l'identification de leur fonction mais aussi l'analyse des données d'expression et la détection de gènes ou réseaux de gènes d'intérêt. Les bases de données telles que dbEST (http://www.ncbi.nlm.nih.gov/dbEST/), PlantGDB (http://www.plantgdb.org/) et TIGR PGI (http://www.tigr.org/tdb/tgi/plant.shtml) ont ouvert de nouvelles possibilités d'étude sur le génome exprimé. Mais les caractéristiques de ces séquences et leur traitement informatique soulève un certain nombre de questions que je développerai dans la partie résultat de cette thèse.

Nous disposons donc de nombreuses informations physiologiques et moléculaires sur la réponse du pin maritime au stress hydrique. Il nous reste cependant à identifier de façon plus exhaustive les gènes dont l'expression est régulée au cours du stress. Cette thèse a pour but d'utiliser de nouvelles approches pour pouvoir analyser l'expression et l'implication potentielle du plus grand nombre de gènes afin d'identifier les gènes candidats pour les mécanismes de résistance au stress hydrique.

Matériel & Méthodes

C MATERIEL ET METHODES

C.1 Matériel végétal et conditions de culture

Deux écotypes (provenances) de pin maritime (*Pinus pinaster* Ait.) –Aquitaine (France) et Tamjout (Maroc) – ont été utilisés pour réaliser cette étude.

Les graines d'origine Aquitaine ont été récoltées lors d'une campagne de ramassage réalisée par le CEMAGREF (Centre de Machinisme Agricole et Génie Rural des Eaux et Forêts) dans des peuplements classés du Médoc. Ce lot de graines avait également été utilisé dans le cadre de la thèse de C. Dubos (http://www.pierroton.inra.fr/biogeco/genetique/theses.html) (Dubos 2001).

La provenance marocaine est issue du peuplement de Tamjout (Fig. C-1), massif de 500 ha caractérisé par une proportion importante de carbonate et de calcaire actif dans le sol à partir d'un mètre de profondeur. Les pins y sont très droits, à branches fines régulièrement réparties le long du tronc.

Les graines ont été conservées en chambre froide à 4°C et maintiennent une très bonne capacité germinative.

C.1.1 Germination

Les conditions décrites par Dubos (2001) exigeaient une stérilisation des graines par de l'H_2O_2 30% puis une germination sur bande de papier Wattman (3M) stériles humides traitées par 3 antifongiques dans des germoirs placés à 18°C. Les graines ainsi traitées se sont révélées sensibilisées à la présence de pathogènes.

Une technique plus classique a donc été adoptée. Des germoirs désinfectés à la Javel sont remplis d'une couche de Vermiculite stérilisée à l'autoclave (121°C, 20min.) (Fig. C-2). Les graines sont ensuite réparties dans les germoirs puis recouvertes d'une seconde couche de Vermiculite. Le substrat est largement humidifié et la germination s'effectue à température ambiante sous éclairage fluorescent. La vermiculite offre l'avantage par rapport au papier humide de ne pas être un substrat aussi favorable au développement bactérien ou fongique. Les plantules doivent rester 1 semaine maximum dans les germoirs et les mégagamétophytes doivent être retirés des cotylédons pour éviter une « fonte du semis ». Aucun traitement sanitaire n'a été nécessaire.

Afin d'obtenir un nombre suffisant de plantules levant au même moment, la quantité de graines a été multipliée par 4 par rapport aux besoins expérimentaux.

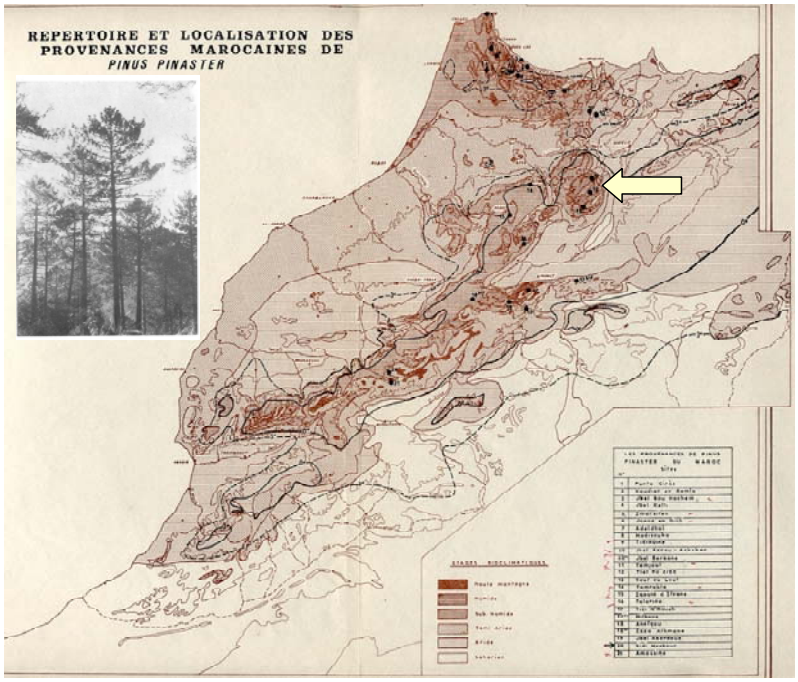


Figure C-1 : Situation géographique et photographie de l'écotype marocain « Tamjout ».

Figure C-2 : Dispositif utilisé pour la germination des graines.

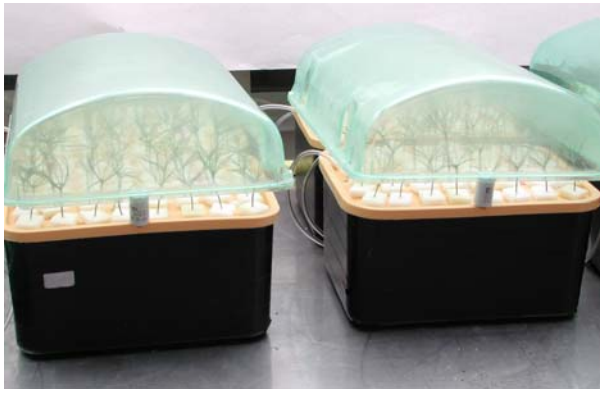


Figure C-3 : Etape de transfert des plantules depuis les germoires vers les conditions hydroponiques.

Figure C-4 : Détail du système de maintient des plants par des carrés de mousse et opacification des bacs par un film plastique.

Figure C-5 : Système d'oxygénation de la solution nutritive par injection d'air à travers des diffuseurs pour aquarium.

Tableau C-1 : Composition du milieu de Seillac (1960)

Solution 1X	[]° milieu final (µM)
$MgSO_4$ $7H_2O$	500
$CaCl_2$ $2H_2O$	500
KH_2PO_4	1000
NH_4NO_3	4000
H_3BO_3	16,7
$MnCl_2$ $4H_2O$	3
$ZnCl_2$	0,3
$CuCl_2$ $2H_2O$	0,3
Na_2MoO_4	0,07
EDTA	100
$FeSO_4$ $7H_20$	100

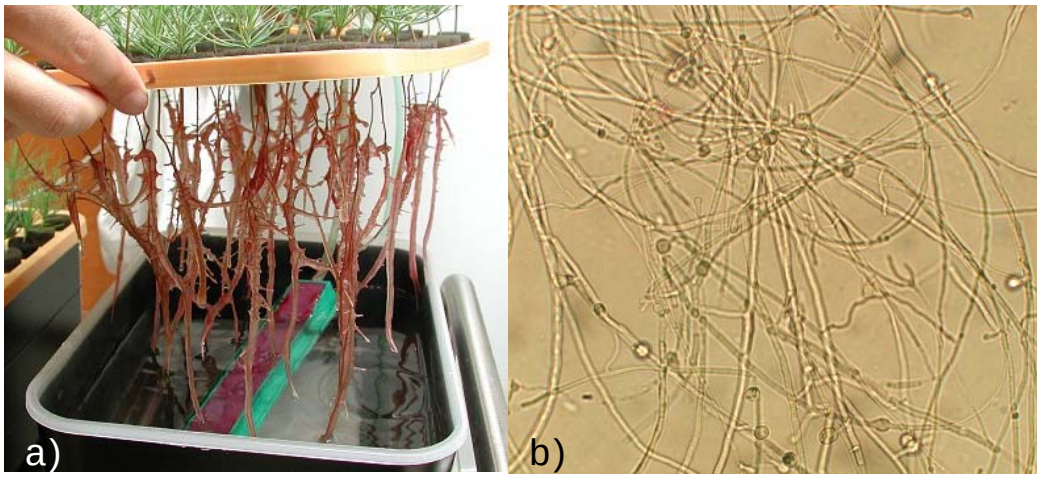


Figure C-6 : a) Mucilage de couleur rose provoqué par l'attaque d'un pathogène.
b) Vue microscopique d'un prélèvement du mucilage.

C.1.2 Culture Hydroponique

Ce mode de culture offre l'avantage de faciliter l'accès aux parties racinaires (objet de notre étude), d'éviter les mycorhisations qui perturberaient l'analyse de la réponse propre des plants à la sécheresse et de contrôler et d'homogénéiser l'apport en éléments nutritifs. Cependant, l'application d'un stress hydrique ne peut se faire que par un stress osmotique et le dispositif ne permet l'étude des plants qu'à un stade de développement précoce.

Après l'apparition des euphylles, les plantules ont été débarrassées des téguments pouvant rester au niveau du collet (source de développement de pathogènes), puis transférés en milieu hydroponique. Pour éviter un desséchement trop violent dû au transfert des bacs de germination (humidité 100%) vers les conditions hydroponiques, des couvercles de germoirs ont été placés sur les bacs de culture pendant 48h (Fig. C-3). Les bacs ont été remplis d'eau le temps de transférer la totalité des plants et de mettre en route le dispositif en présence de solution nutritive.

Les bacs de culture d'une contenance de 10L ont été opacifiés à l'aide de film plastique noir opaque. Leurs couvercles ont été percés et les plantules sont maintenues par des carrés de mousse fendue permettant la croissance de la tige (Fig. C-4). La solution nutritive, basée sur celle décrite par Seillac (1960) (Tab. C-1) est autoclavée et renouvelée une fois par semaine pour éviter un appauvrissement en éléments nutritifs, une variation du pH ou le développement de pathogènes. L'utilisation d'H_2O miliQ permet de s'affranchir d'une variation dans les teneurs en micro-éléments et de l'apport de pathogènes. A chaque renouvellement, les bacs sont lavés et javellisés. L'oxygénation du milieu est assurée par un compresseur relié à des diffuseurs d'aquarium de 25 cm. Le débit est réglé individuellement pour chaque bac via un robinet pour obtenir un bullage optimal tout en évitant des turbulences trop prononcées au niveau des racines (Fig. C-5).

C.1.3 Contrôle des pathogènes

Deux séries de cultures successives ont été mises en œuvre en utilisant du milieu de Seillac 1x puis 0.2x et ont subi une attaque de pathogène dans les jours suivant l'application du stress. L'agent pathogène a un développement fulgurant sur 48h. Les racines se recouvrent d'un mucilage rose puis se nécrosent très rapidement (Fig. C-6a). Une préparation microscopique a permis d'écarter l'hypothèse d'une attaque bactérienne et mettre en évidence une origine fongique (Fig. C-6b). Des isolements réalisés par l'unité de santé végétale de l'INRA Bordeaux ont permis d'identifier un *Fusarium oxysporum*. Ce champignon est couramment rencontré sur les cultures de tomate hors sol. Bien qu'un traitement

phytosanitaire ne soit pas souhaitable, il apparaît comme incontournable. Le choix s'est porté sur le Tachigaren 360 qui est un fongicide systémique utilisé pour lutter contre la nécrose du collet et la pourriture des racines : *Fusarium oxysporum f. sp. radicis-lycopersici* (F.O.R.L.) en cultures de tomates. Sous forme liquide, il est particulièrement adapté à une utilisation dans le système d'irrigation. Un traitement préventif à 5ppm une semaine avant l'application et pendant toute la durée du stress a été retenu.

C.1.4 Eclairage

Le pin maritime est une espèce héliophile, la quantité et la qualité de la lumière apportée doivent donc être suffisantes et non limitantes pour la croissance.

Les expériences déjà mises en œuvre dans l'Unité avaient été réalisées dans une chambre de culture climatisée (24/18°C) et équipée de lampes MAZDA aux halogénures métalliques de 250W /4200K à raison d'environ 200W/bac apportant 260µmol.m^{-2}.s^{-1} (16/8h). Cette salle n'étant plus disponible, deux solutions se présentaient :

- la culture en serre où la quantité de lumière n'est pas limitante mais non contrôlée et où la régulation de la température devient un facteur critique essentiellement l'été où la température atteint 35°C sous cooling et peut atteindre des températures létales pour les plants en cas de panne. Cette solution, quoique envisageable pour l'automne est à éviter durant la saison estivale.
- La réalisation d'une seconde « chambre de culture » sans climatisation possible. Le maintient d'une température homogène dépend alors étroitement de la chaleur dégagée par le système d'éclairage. Cela implique l'utilisation d'un éclairage par tubes fluorescents avec déport des ballasts, et extraction de la chaleur qu'ils émettent.

C'est cette seconde solution qui a été retenue. Deux tables métalliques sur roulettes de 2m*80cm ont ainsi été construites et équipées de rampes d'éclairage (Fig. C-7).

C.1.4.1 Photosynthèse

La quantité d'énergie lumineuse est un paramètre important. L'éclairement solaire direct maximum est de 95000 lux (unité se rapportant à la vision humaine) soit 950W/m² dont 4% dans l'UV, 42% dans l'IR et 54% dans le visible =515W/m² dont 410W/m² dans la PAR (photosynthetically active radiation : 400-700) ce qui équivaut à un flux de photons d'environ 1800µmol/m²s (Heller *et al.* 1993).

Figure C-7 : Vue générale des tables métalliques équipées de leur système d'éclairage à hauteur réglable et avec déport des ballasts dans des caissons ventilés.

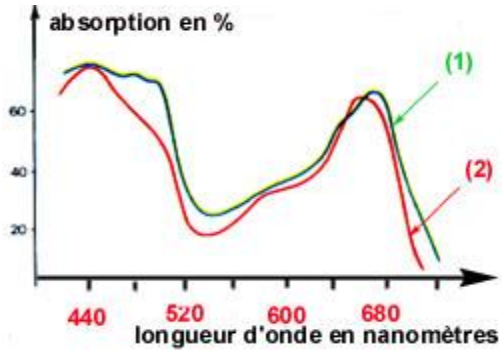


Figure C-8 : Spectres d'absorption (1) et d'activité photosynthétique (2) de l'algue verte Ulva taeniata.

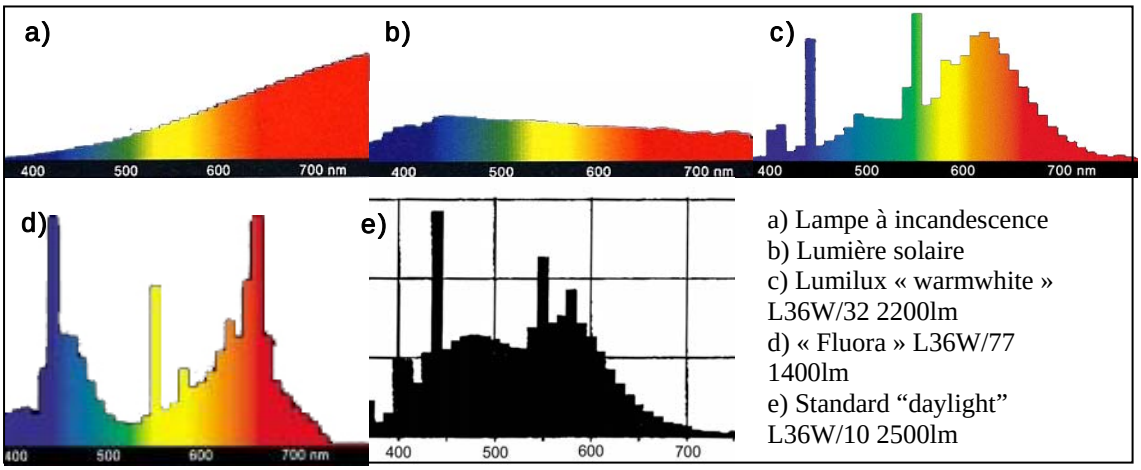


Figure C-9 : Spectres d'émission entre 400 et 700 nm de différentes sources lumineuses. (http://www.osram.com)

Comme on peut le constater chez *Ulva taeniata,* les pics d'efficacité photosynthétique et d'absorption de la lumière par la chlorophylle coïncident (Fig. C-8). Les photons utilisable par la plante pour la photosynthèse seront compris dans une gamme de longueur d'onde de 400 à 700nm (PAR). Le risque avec un éclairage fluorescent est d'avoir un éclairage insuffisant, trop proche du point de compensation (point critique où le carbone fixé lors de la photosynthèse équilibre celui rejeté par la respiration). Le rayonnement variant en fonction de l'inverse du carré de la distance à la source (Gosse, 1984 doc INRA), il est intéressant de réduire la distance entre les plantules et les tubes dont le dégagement calorifique est faible. Dans notre cas, la distance à été réduite à environ 15 cm.

C.1.4.2 Action Phytochromique

Si la croissance de la plante dépend directement de la quantité de lumière disponible, elle est également sensible à l'équilibre des couleurs émises. La morphologie de la plante est sous le contrôle du phytochrome. Ce pigment existe sous deux formes différentes, convertibles l'une en l'autre. La forme P660 absorbe la lumière dans le rouge clair (RC), l'autre, P730, dans le rouge lointain (RL). Sous l'effet de radiations RC la forme P660 se convertit en sa forme active P730 et inversement sous des radiations RL. Le phytochrome contrôle entre autres la longueur de l'entrenœud, la floraison et la germination.

Le facteur zeta qui est le rapport des photons 660nm/730nm détermine un équilibre « phytochrome actif » (phi) / « phytochrome total ». Une valeur de phi faible (RC/RL faible) favorise l'élongation des entrenœuds, une valeur de phi élevée (RC/RL important) produit des plantes très ramassées, à entrenœuds courts. En conditions naturelles, phi varie entre 0.5 (soleil) et 0.20 (sous couvert végétal), des tubes « cool white » donnent un phi de 0.7 et des tubes « Warm white » 0.65 (Gaudillere and Chasles 1986). Les sources incandescentes donneront donc des plantes étiolées à petites feuilles tandis que les sources fluorescentes classiques impliqueront des plantes trapues et ramifiées.

Avant expérimentation, il est cependant difficile de déterminer l'équilibre RC/RL obtenu.

C.1.4.3 Sources

Les courbes de distribution spectrale de l'énergie sont très variables selon la source lumineuse artificielle et diffèrent beaucoup de celles de la lumière solaire (Fig. C-9b). Les sources incandescentes ont des spectres très riches dans le rouge et l'infra- rouge (Fig. C-9a). Les tubes fluorescents classiques de type « daylight » ont une distribution assez homogène sur la gamme 400 à 700nm (Fig. C-9e) tandis que les tubes "Warmwhite" ont une distribution

centrée sur 620nm (Fig. C-9c). Les tubes spécifiques comme le « Fluora » (OSRAM) ont été développés spécialement pour un rendement élevé autour des pics d'absorption des chlorophylles a et b dans le bleu 440 / 480 nm et dans le rouge 650 / 680 nm (Fig. C-9d).
Le choix a donc porté sur l'utilisation de tubes standard « daylight », d'un tube « Fluora » à haut rendement photosynthétique centré au dessus de chaque rangée de bacs et d'un tube « warmwhite » pour rééquilibré le spectre vers le rouge.
Les mesures de PAR réalisée avec un capteur Li-COR Quantum (Fig. C-10) ont données des valeurs comprises entre 360µmol/m²s à 15 cm et 380µmol/m²s à 10cm. Les tubes sont espacés de 5 cm soit une puissance de 400W/m^2.
Une photopériode classique de 16h jour et 8h nuit a été appliquée.

C.1.5 Application du stress

Le stress hydrique est obtenu par ajout de Polyéthylène glycol (PEG) à la solution nutritive afin d'en augmenter son osmolarité. Le PEG utilisé (Sigma, réf. P3640) ne nécessite pas de purification avant utilisation pour éliminer les résidus d'aluminium ou de magnésium qui à forte concentration présentent une toxicité pour les plantules (Triboulot, 1996). La masse molaire de 3350g/mol est suffisamment élevée pour limiter l'absorption racinaire tout en préservant la fluidité de la solution nutritive. Les courbes de correspondance du potentiel osmotique de la solution de Seillac en fonction de la concentration en PEG (Dubos 2000) nous donnent –0.45MPa pour 180g/L *(milieu M3)* et –0.60Mpa pour 210g/L *(milieu M4)*.
L'apparition des premières aiguilles secondaires ou pseudophylles (Fig. C-11) a été choisie comme témoin du stade de développement des plantules pour l'application du stress (Fig. C-12).

C.1.5.1 Stress court

Pour l'étude de la réponse des plantules à court terme, la solution nutritive de 4 bacs est remplacée par une solution à -0.60MPa. Un bac est également récolté pour servir de témoin à l'expérience « Stress court ». Les autres bacs sont échantillonnés à T0+2h, T0+6h, T0+24h et T0+48h.

C.1.5.2 Stress long

Pour chaque provenance, un bac est placé dans une solution à -0.45MPa pour une durée de 3 semaines et un bac est conservé dans une solution nutritive de contrôle. A l'issue de la période de « Stress Long », les échantillons témoins et stressés sont récoltés.

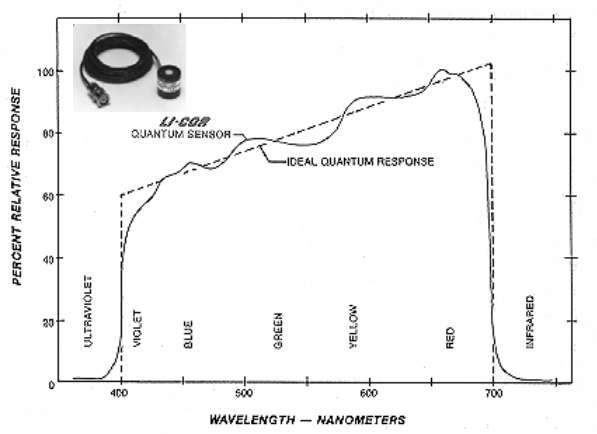


Figure C-10 : Réponse spectrale type des détecteurs « Li-COR Quantum » en fonction de la longueur d'onde et la réponse quantique idéale (réponse identique à tous les photons dans la gamma 400-700nm).

Figure C-11 : Développement des premières aiguilles secondaires ou pseudophylles.

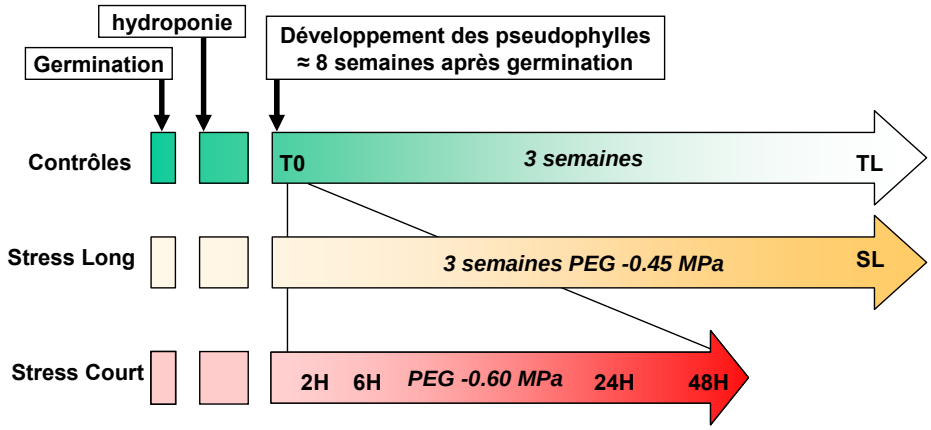


Figure C-12 : Schéma chronologique situant les prélèvements effectués à partir du développement des premières aiguilles secondaires. 7 bacs de cultures par provenance. T0 = Témoin du stress court, TL = Témoin du stress long, SL = Stress Long, 2H/6H/24H/48H = échantillons prélevés 2/6/24/48 heures après application du stress.

Figure C-13 : Enchevêtrement des racines rendant l'individualisation des plants impossible. Système racinaire des plants après 7 semaines de croissances dans la solution nutritive.

a)	b)	c)
		d)

Figure C-14 : Mesure du potentiel de base sur une plantule en utilisant la chambre de Schölander. De gauche à droite, dispositif avec chambre et loupe binoculaire (a), couvercle de la chambre et maintient de la plantule au niveau du collet (b), vue de l'intérieur de la chambre avec les arrivées et sortie d'azote comprimé (c), vue de la tige avec début d'exsudation de la sève sous l'effet de la pression (d).

Dans la suite de ce mémoire, une combinaison Provenance x Condition sera nommée un « échantillon ».

Etant donné que nous avons 5 conditions pour le stress court (T0, T0+2h, T0+6h, T0+24h et T0+48h), 2 conditions pour le stress long (TL, SL), et 2 provenances (France, Maroc), nous parlerons de 10 échantillons pour l'étude stress court et 4 échantillons pour l'étude stress long. Les échantillons T0+2h et T0+24h, n'ont pas été utilisés dans cette thèse pour la réalisation des études d'expression.

C.1.6 Récolte du matériel végétal

Les acides nucléiques et en particulier les ARN sont particulièrement sensibles à la dégradation par les enzymes du type ARNase. Certains gènes peuvent également répondre rapidement à un changement de l'environnement. Les parties racinaires formant un enchevêtrement inextricable (Fig. C-13), les plantules ont donc été sectionnées au niveau du collet, puis la masse racinaire coupée par lots d'une dizaine d'individus. Après un égouttage rapide, les tissus sont plongés immédiatement dans un bain d'azote liquide, emballés puis stockés à -80°C. L'option d'un rinçage des racines pour éliminer le PEG a été écartée pour préserver les pousses secondaires fragiles. Les parties aériennes (aiguilles + tige) sont traitées de la même manière et stockées à -80°C. Elles n'ont pas été analysées dans cette étude.

C.2 Mesure du niveau de stress et paramètres physiologiques

C.2.1 Mesure du potentiel de base

Lors de l'application d'un stress de type osmotique par ajout d'un osmoticum à la solution, il est important d'évaluer le niveau de stress réellement provoqué chez la plante.
Un paramètre couramment utilisé en écophysiologie est la mesure du potentiel hydrique de base, estimée par la mesure du potentiel hydrique foliaire à l'aube et exprimée en unité de pression. Nous avons mesuré ce potentiel avant la remise en route de l'éclairage. La pièce étant totalement dans l'obscurité pendant toute la durée des mesures. Dans notre cas, la taille de nos plants a permis de mesurer le potentiel directement sur la tige en plaçant l'ensemble des parties aériennes dans une chambre de Schölander (Fig. C-14). Des mesures ont été réalisées sur les stades T0, T0+48h, TL et SL.

C.2.2 Dosage des terpènes

La teneur de ces composés volatiles dans les parties aériennes est connue pour varier au cours de stress abiotiques (Penuelas and Llusia 1999) (Llusia and Penuelas 2000) et biotiques (Faldt *et al.* 2003) (Tiberi *et al.* 1999). Elle permet également de différentier les provenances Aquitaines et Marocaines. Au sein du laboratoire, des études d'association entre caractères phénotypiques dont la teneur en terpènes et variants nucléotidiques de gènes candidats sont actuellement développées dans le cadre d'une thèse. L'étude porte sur les variations de la résistance au stress hydrique de 20 populations de pin maritime distribuées selon un cline de précipitation. Différentes familles de France, Espagne, Corse, Tunisie et Maroc ont été échantillonnées et plantée en Espagne et en France dans un dispositif expérimental comprenant 9000 plants. Parmi les caractères mesurés, un dosage des terpènes doit permettre de caractériser des différences de comportement entre ses familles. Nous avons donc fait le choix de mesurer également les terpènes dans le cadre de cette thèse.

Des aiguilles ont été prélevées sur plusieurs plantules des échantillons de l'expérience « stress long ». Chaque prélèvement a été pesé, découpé et conservé dans des piluliers remplis de pentane. Les échantillons ont été analysés au laboratoire Chimie Analyse de l'Institut du Pin (Patrick PARDON). La teneur en terpène est mesurée par chromatographie en phase gazeuse (Trace GC, Thermo Quest ; injecteur AS 2000, Thermo Finningan ; colonne, SGE 25 QC2/BP21 0.25) et la détection se fait par un détecteur à ionisation de flamme (FID) intégré à l'appareil.

C.3 Analyse du transcriptome et caractérisation des EST

C.3.1 Extraction des ARN

Les échantillons racinaires sont broyés manuellement sous azote liquide. Cette étape est déterminante et le matériel doit être broyé très finement.

Les ARN totaux sont obtenus selon le protocole de Chang et al. (Chang *et al.* 1993) modifié pour améliorer les rendements et la qualité des extractions sur notre matériel d'étude :

- Le SDS a été retiré du tampon d'extraction et le volume a été ramené à 15mL pour 1g de matière fraîche.
- Une extraction Chloroforme/Isoamyl alcool (24 :1) avec rinçage de la phase aqueuse a été ajoutée au protocole. Cette étape permet d'augmenter le rendement jusqu'à 30%.

Les acides nucléiques sont précipités en présence d'Acétate de sodium à 0.3M et 2 volumes d'éthanol à -20°C pour une durée de 2 heures. Le culot obtenu après centrifugation est repris dans de l'eau exempte de nucléases.
Un traitement à la DNAse suivit d'une nouvelle extraction Chloroforme/Isoamyl alcool (24 :1), précipitation puis solubilisation et purification est effectuée pour éliminer la DNAse, le tampon et les produits de dégradation.

La qualité des ARN est ensuite contrôlée par électrophorèse sur gel d'agarose permettant de détecter les dégradations éventuelles et la présence d'ADN. Un spectre d'absorption entre 220 nm et 300nm nous permet d'estimer la pureté de l'extrait en calculant les rapports de D.O. (DO_{260nm}/DO_{280nm} et DO_{260nm}/DO_{230nm}) et de déterminer la quantité d'ARN (DO_{260nm}).
Le rendement varie entre 70 et 160 µg ARNt/g MF suivant la nature de l'échantillon. Le rapport DO_{260nm}/DO_{280nm} était voisin de 2.

Lors de chaque série d'extraction, 5-6 g de matière fraiche sont traitées dans 6 tubes indépendants.
Pour chaque traitement, un minimum de 3 séries d'extractions d'ARNs distinctes ont été réalisées jusqu'à obtention d'une quantité suffisante pour les études d'expression.
Les extraits de bonne qualité ont été « poolés » - ce qui permet d'éviter les biais possibles d'une seule extraction - et dosés. Ils ont alors été précipités et repris dans de l'eau DEPC afin d'obtenir une solution stock de concentration élevée comprise entre 6 et 10 µg/µL et de rapport DO_{260nm}/DO_{280nm} proche de 3, indispensable à la synthèse des cibles à hybrider sur les microarrays.

C.3.2 Banques d'ADNc

Les études d'expression à haut débit telles que les puces à ADNc nécessitent le développement préalable de ressources génomiques et en particulier de banques d'ADNc. Dans le cas présent, deux banques distinctes ont été construites : l'une à partir de xylème (tissus conducteur) en différentiation et l'autre à partir de tissus racinaires.

C.3.2.1 Banque de xylème

Dans le cadre du projet européen GEMINI (http://www.pierroton.inra.fr/Gemini/), l'équipe du Professeur Canovas de l'Université de Malaga a produit une banque d'ADNc en utilisant

du xylème en différentiation associé à 4 types de bois (Xylème associé à la face comprimée, à la face opposée, à du bois initial et à du bois final) fournit par notre unité.
Les ARN totaux des différents types de xylème ont été obtenus selon la méthode de Chang et al. (1993) et mélangés en quantité égale. Les ADNc et les banques ont été synthétisées en utilisant le ZAP-cDNA synthesis kit (Stratagene, La Jolla, CA, USA). Environ 10000 clones ont été repiqués et expédiés à la société Genome Express pour le séquençage d'EST en 5'. Une analyse détaillée et l'exploitation de cette ressource a été menée dans le cadre de la thèse de Jorge Paiva (2006).

C.3.2.2 Banques de racines et aiguilles

Ces banques ont été construites par J.M. Frigério (INRA de Pierroton) sur du matériel végétal issu d'une expérience de culture hydroponique de plantules de pin maritime soumis à un stress osmotique par ajout de PEG. Les plantules de 2 provenances (Landes et Maroc) âgées de 3 semaines sont soumises à un stress de -0.45MPa sur une durée de 3 semaines. Les ARN totaux des parties aériennes (tige et appareil foliaire) et des parties racinaires ont été extraits séparément selon la méthode de Chang et al. (1993). Les ADNc et les banques ont été synthétisés en utilisant le ZAP-cDNA synthesis kit (Stratagene, La Jolla, CA, USA) et nommés RN (Root Normal), RS (Root Stress), AN (Aerial part Normal) et AS (Aerial part Stress).
Dans le cadre de cette thèse, nous avons étalé des aliquotes sur milieu Luria Bertani (LB)-agar à 50 µg/ml d'ampicilline et repiqué en milieu liquide LB-ampicilline au format 96 puits un total d'environ 15000 clones. Après une phase de croissance « overnight » sous agitation à 37°C, les plaques sont stockées à -80°C en présence de glycérol à une concentration de 25%.

La duplication des plaques des différentes banques a toujours été réalisée selon les conditions décrites ci-dessus.

C.3.3 Obtention des séquences

Au début de la thèse (en 2002), nous avons mis en place une chaîne semi-automatique de séquençage dans le cadre du pôle Génotypage-Séquençage de la plateforme Génomique fonctionnelle de Bordeaux. Les étapes de séquençage ont été possibles grâce aux outils nouvellement disponibles. Les PCR sur colonies ont été conduites en plaque 384 sur un robocycleur GeneAmp PCR system 9700 (Perkin Elmer, Foster City, CA, USA). L'amplification de l'insert se fait dans 20 µL contenant du tampon 1X, 0.2mM dNTPs, 0.2

µM d'amorce M13F et M13R, 2 mM $MgCl_2$ et 0.3 U *Taq* DNA Polymerase (Invitrogen, Carlsbad, CA, USA). La matrice est apportée en piquant des cultures fraîches. Les paramètres sont : 95°C, 10 min, 95°C, 1 min, 55°C, 1 min, 72°C for 1 min pour 40 cycles et une élongation finale à 72°C pour 10 min. Le produit PCR est purifié avec le système « Montage PCR_{96} Cleanup Kit (Millipore, Billerica, MA, USA) » selon le protocole du fabriquant. Un robot pipeteur Genesis RSP 100 (Tecan, Maennedorf, Switzerland) prend en charge le transfert de liquide des plaques 384 puits vers 96 puits. Les produits PCR purifiés servent de matrice (2.5µL) pour la réaction de séquence utilisant 2.5µL de mixe inclus dans le « « DYEnamic™ ET Dye Terminator Kit (MegaBACE™) » (Amersham Biosciences, Buckinghamshire, UK).

Nous avons choisi un séquençage en simple passe avec une amorce T3 pour séquencer l'extrémité 5' des gènes. L'avantage est d'augmenter la probabilité de trouver des homologies avec des séquences connues et donc d'attribuer une fonction au gène. Le robocycleur GeneAmp system 9700 384 puits est réglé comme suit : 94°C, 20 sec, 50°C, 15 sec, 60°C, 1min et 35 cycles. La totalité du produit de la réaction de séquence est purifiée par centrifugation (1000xg, 5min) sur plaques (Millipore Multiscreen MAHVN45) contenant du Sephadex G-50 (Amersham Biosciences).

Les données brutes de séquences sont détectées sur un séquenceur 96 capillaires MEGABACE™ 1000 (Amersham Biosciences).

C.4 Analyse d'expression

C.4.1 Synthèse des ADNc

Pour la fabrication des sondes, la rétrotranscription en ADNc à partir des ARN totaux « microarrays » a été réalisée selon le protocole du kit de marquage CyScribe Post-Labeling kit (Amersham) en présence d'amino allyl-dUTP. Le kit étant prévu pour 500 ng d'ARN messager et considérant un rapport ARNm/ARNt de 0.5 à 1%, il faut une quantité de l'ordre de 50 à 100 µg d'ARN totaux. Le volume maximum d'ARN totaux concentrés utilisable dans le kit, soit 8 µL, a donc été retenu. Le temps de rétrotranscription de 1.5 heure prévu a été modifié. Après 1.5 h d'incubation à 42°C, nous avons ajouté 1µL de reverse transcriptase, et la réaction d'extension a été prolongée de 1.5 h pour s'assurer d'une rétrotranscription optimale. Les ADNc de l'expérience « stress court » ont été obtenus en utilisant l'enzyme CyScript reverse transcriptase fournie tandis que les ADNc de l'expérience « stress long » ont été synthétisés avec de la Superscript II kit (Invitrogen Life Technologies, Grand Island, NY).

Pour les expériences de PCR quantitative, les ADNc ont été synthétisés avec de l'enzyme Promega ImProm II (Promega, Madison, WI) à ARN totaux constant. Les solutions d'ARN totaux sont ramenées à une concentration d'environ 300 ng/µL puis dosées précisément par spectrophotométrie. Le volume exact nécessaire pour obtenir 1 µg d'ARN totaux est prélevé pour conduire la rétrotranscription selon le protocole du fabricant.

C.4.2 Microréseaux à ADNc

C.4.2.1 Réalisation de l'Unigène

L'Unigène représente l'ensemble non-redondant des gènes contenus dans les banques d'ADNc. Chaque membre peut se définir comme la sélection d'une (ou plusieurs) séquence représentative d'un gène unique. Les possibilités techniques des microarrays utilisées sont de l'ordre de 20000 dépôts par lame. L'objectif était de réduire de façon conséquente le coût et le nombre de sondes à déposer tout en conservant le maximum d'information biologique.
Nous avons exploité les données bioinformatiques disponibles dans nos bases de séquences de pin maritime à la date du 22 septembre 2003. Les résultats d'assemblage (Fig. C-18) semblaient montrer que les séquences d'un même gène appartenaient au même consensus. Nous avons donc considéré que chaque consensus représentait un gène différent et choisi un représentant unique par consensus (cf. C.6.1).

Dans le cas des singletons, chaque séquence doit, par définition, correspondre à une séquence unique. Cependant, certains singletons peuvent appartenir au même gène, mais n'ont pu être regroupés. Nous avons fait le choix d'écarter ces "doublons" de l'Unigène. Pour cela, les singletons ont été comparés sur la base du résultat de la recherche d'homologie obtenue par la stratégie multiBlast du pipeline (cf chapitres suivants) avec un seuil de significativité placé à 10^{-7}. Les singletons ne comportant aucune homologie avec un score inférieur à cette valeur seuil sont intégrés directement à l'Unigène. Dans le cas contraire, les singletons ayant au moins un résultat commun entre eux parmi les 5 premiers résultats (Blast) inférieurs au seuil, sont considérés comme similaires et un seul représentant du groupe est retenu.

Pour choisir un représentant de chaque consensus, nous avons essayé d'homogénéiser la taille des fragments déposés. En effet, à nombre de copies et teneur en GC identique, la fluorescence produite par la cible dépend de la longueur du fragment. Un fragment long incorpore plus de dNTP fluorescents qu'un fragment court. Cependant, nous ne disposions pas de la longueur réelle des inserts correspondant aux EST. Nous avons donc calculé une

longueur théorique à partir des données de bioinformatique. Les séquences ont été obtenues par séquençage d'EST en 5'. Les EST sont des ADNc issus d'ARNm, or les EST correspondent rarement à l'ARNm pleine longueur (Fig. C-15). En effet, la transcription inverse produit un fragment d'ADNc débutant en région 3' de l'ARNm (queue polyA) qui se termine souvent avant la fin de la région 5'. De plus, le fragment d'ADNc peut se casser avant ou pendant son insertion dans le vecteur de clonage. Il résulte de ces événements un fragment partiel de l'ARNm de départ.

Ainsi, après l'assemblage des EST en contigs, permettant d'en déduire des séquences consensus (CN), on peut calculer une longueur théorique pour chaque insert (Fig. C-15):

Longueur Théorique de l'insert = Longueur du Consensus – nb bases entre l'extrémité 5' de l'EST et l'extrémité 5'du Consensus

Dans le cas des contigs, nous avons adopté la stratégie suivante pour choisir le clone représentatif du contig (Fig. C-16) :

- Choix des inserts situés au plus proche de la région 3' du gène, afin de garantir une spécificité accrue. En effet, la différenciation au niveau des membres d'une famille multigénique se fait généralement dans la partie 3' de la séquence. C'est la zone la plus variable.
- Choix des inserts les plus proches d'une longueur théorique définie (borne), garantissant une hybridation correcte des cibles sur les sondes, en termes de spécificité et d'homogénéité de quantité de cibles fixées. La longueur théorique doit également être comprise entre un seuil minimum et un seuil maximum. Nous avons opté pour une longueur théorique d'environ 500 pb (paires de bases), un seuil minimum de 200 pb et un seuil maximum de 800 pb.

Ces deux critères de choix engendrent plusieurs cas possibles (Figure C-16):

1) la longueur théorique de l'insert est comprise entre 500 pb et 650 pb.

Il s'agit du cas le plus intéressant pour notre stratégie (le meilleur intervalle de longueur théorique), c'est pourquoi seulement dans ces cas nous avons choisi l'insert dont l'EST était le plus proche de la queue polyA.

2) la longueur théorique de l'insert est comprise entre 450 pb et 500 pb.

Choix de l'EST avec le plus faible écart (Δ) entre la longueur théorique de l'insert et la valeur 500 pb.

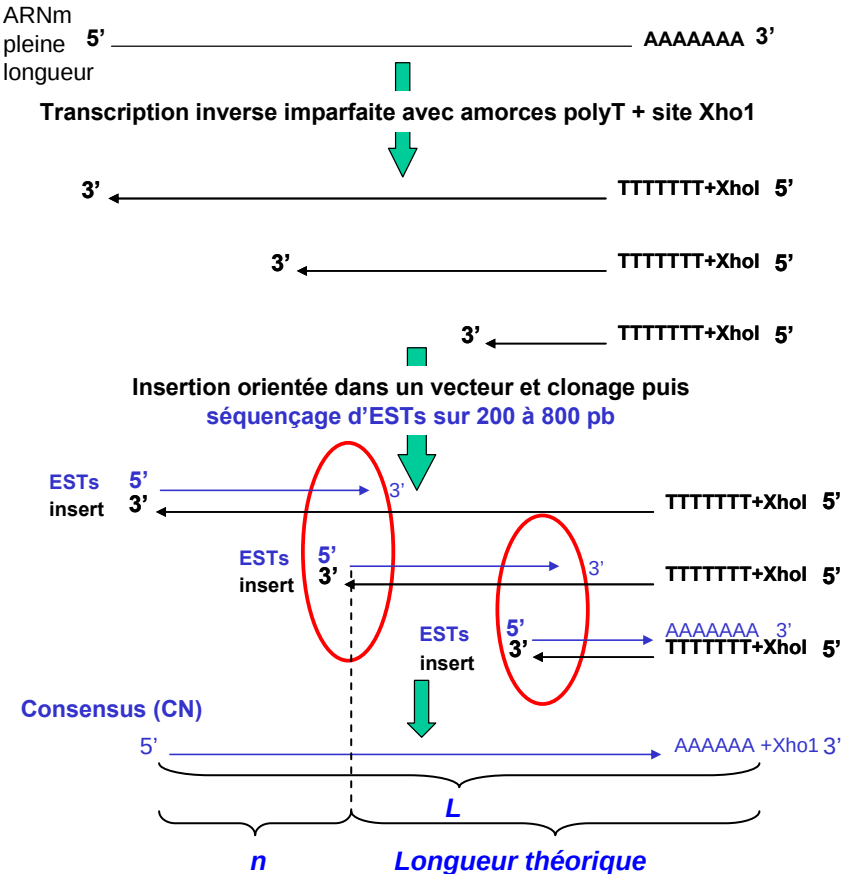


Figure C-15 : Conséquences de la rétrotranscription imparfaite et du clonage sur les ESTs générés en 5'. Les ESTs obtenus (en bleu) ne commencent pas tous à l'extrémité 5' correspondant à l'ARNm pleine longueur. Les zones de recouvrement entre ESTs (entourées) permettent par assemblage bioinformatique de retrouver l'intégralité de la séquence du transcrit (consensus). Nous avons défini la longueur théorique d'un EST comme étant la longueur du consensus (L) moins le nombre de paires de bases (n) entre l'extrémité 5' du consensus et 5' de l'EST.

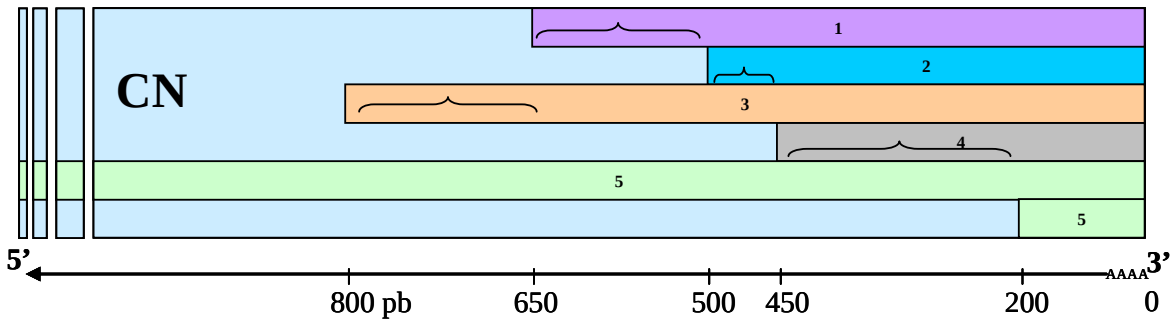


Figure C-16 : Sélection des clones pour l'Unigène. Stratégie adoptée pour le choix du clone représentatif d'un contig sur la base de la longueur théorique de l'insert.
Dans l'ordre de priorité :

1. S'il existe un insert de longueur théorique comprise entre 500 et 650 pb, choix du clone pour lequel l'EST est le plus proche de l'extrémité 3'
2. Insert compris entre 450 et 500 pb, choix du clone avec l'insert le plus proche de 500 pb
3. Insert entre 650 et 800 pb, choix du clone avec l'insert le plus proche de 650 pb
4. Insert entre 200 et 450 pb, choix du clone avec l'insert le plus proche de 450 pb
5. Choix du clone dont l'insert est le plus proche de 500 pb

3) la longueur théorique de l'insert est comprise entre 650 pb et 800 pb.

Choix de l'EST avec le plus faible Δ.

4) la longueur théorique de l'insert est comprise entre 200 pb et 450 pb.

Choix de l'EST avec le plus faible Δ.

5) la longueur théorique de l'insert est inférieur à 200 pb ou supérieur à 800 pb.

Choix de l'EST avec le plus faible Δ.

Nous avons également ajouté à cet Unigène, une sélection de clones issus des banques américaines de *Pinus taeda* dont les inserts codaient des enzymes impliqués dans le métabolisme des sucres et de l'azote et pour lesquels nous ne disposions pas de séquences de pin maritime. L'effet du PEG sur l'expression de ces gènes avait été étudié lors du développement d'embryons somatiques de *Picea glauca* (Stasolla *et al.* 2003). Des contrôles positifs et négatifs issus des études d'expressions sur la formation du bois ont aussi été inclus. Nous avons déposé des séquences de : 4CL (4-coumarate-CoA ligase (E.C. 2.1.1.14)(AL751056)), 5P, Cytochrome P450 (BX255398), Desmine, GP (glycine rich protein (BX000600)), G3PDH (Glycéraldéhyde 3-phosphate déshydrogénase), HSP (Class I LMW heat shock protein (BX000656)), Luciférase (1,3,5,12), Ras, Thioredoxin H (BX000658), Tubuline alpha chain (BX000608), Ai6V(gène d'aiguille différentiellement exprimé lors d'une expérience d'exposition au vent), X42 et X46 (gènes différentiellement exprimés lors d'études sur la formation du bois en ADNc-AFLP).

C.4.2.2 Dépôt des sondes sur lame de verre

Cette étape a été réalisée en partenariat avec la société autrichienne PICME (http://www.picme.at).

La première étape de réarrangement des clones sélectionnés pour l'Unigène a été effectuée au laboratoire en utilisant un robot pipeteur Tecan Genesis 100 adapté pour cette utilisation spécifique. Un double de l'ensemble des clones de pin maritime ainsi que les plaques de l'Unigène ont été envoyé à cet organisme qui est appelé à jouer le rôle de centre de ressource génomique pour la communauté scientifique forestière européenne dans le cadre du réseau d'excellence EVOLTREE. Après extraction des plasmides et amplification par PCR, les produits ont été déposés dans du tampon SSC additionné de Bétaïne sur des lames de verre de marque Corning, modèle GAP.

C.4.2.3 Plan d'expérience

Pour les études d'expression, nous avons décidé de n'étudier que les stades Contrôle, Stress 6h et Stress 48h pour l'expérience « stress court » ; Contrôle long et stress long pour l'expérience « stress long ».

Comme il le sera discuté plus loin, le choix des cibles à co-hybrider sur une même lame détermine la puissance de l'analyse statistique et le nombre de lames à hybrider.

Dans notre cas, pour réduire les variabilités expérimentales et le coût, nous avons choisi de conduire séparément les expériences de stress court et de stress long. Le plan d'expérience est basé sur un dessin en boucle (loop design). L'expérience « stress long » consiste initialement en une boucle simple de 8 lames tandis que l'expérience « stress court » présente 3 boucles imbriquées pour un total de 18 lames (Fig. C-17). Au total, 26 lames sont donc nécessaires pour étudier l'expression de 10 échantillons (2 provenances dans 5 conditions).

C.4.2.4 Synthèse des cibles

Il existe deux méthodes principales pour le marquage des cibles avec des fluorochromes (Cy3 et Cy5 dans le cas des puces à ADNc).

La première, dite marquage direct, utilise l'incorporation de nucléotides liés à un fluorochrome lors de la rétrotranscription. La rétrotranscription et le marquage se font donc en une seule et même étape. L'inconvénient de cette méthode vient de l'encombrement de la molécule fluorescente qui peut gêner l'incorporation du nucléotide lors de la transcription et faire chuter le rendement de la réaction.

La seconde méthode utilise un marquage indirect. Le brin d'ADNc est synthétisé en présence d'un amino allyl-dNTP dont l'encombrement diffère peu de celui des dNTP et par conséquent ne modifie pas le rendement de la transcription. Le marquage est obtenu après dégradation des ARN et purification des ADNc par réaction chimique entre l'amino allyl-dNTP et un NHS-ester du fluorochrome. C'est cette seconde approche que nous avons utilisé en suivant le protocole du kit CyScribe Post-Labeling (cat: RPN5660, Amersham) comprenant les étapes de synthèse des ADNc, dégradation par lyse alcaline des ARN, purification sur colonne, marquage indirect (Allyl-dUTP) et purification finale sur colonne.

Les lames sont hybridées avec une quantité d'ADNc constante. Les cibles ont donc été dosées par spectrophotométrie à 260 nm et pour chaque série d'hybridation, une quantité identique d'ADNc a été pipetée en s'alignant sur l'échantillon contenant la quantité d'ADNc la plus faible. Le volume des cibles prélevées est ramené à une valeur inférieure à 5 µL en utilisant des colonnes Microcon YM-30 (cat : 42410, Millipore).

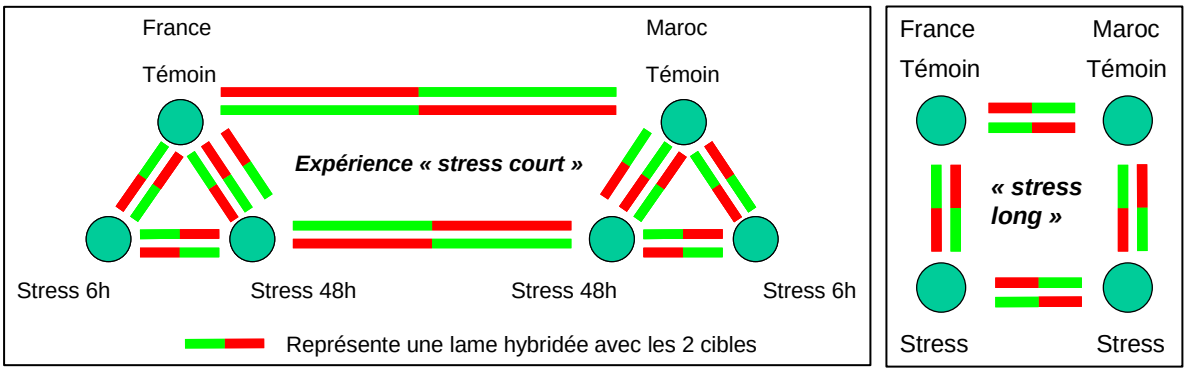


Figure C-17 : Plan d'expérience en « loop design » choisi pour les hybridations des lames. A gauche, l'expérience « stress court » nécessite 18 lames. A droite, 8 lames sont utilisées pour l'expérience « stress long ».

Les deux cibles marquées de fluorochromes différents et devant être hybridées sur un même lame sont mélangées. Le volume de la cible composite est ramené à un volume de 100 µL avec de la solution d'hybridation composée de 58 µL de Formamide, 12 µL de Denhardt 50x, 29 µL de SSC 20x, 2 µL H_20, 2 µL SDS 10% et 1 µL d'ADN de sperme de saumon 10 mg/ml dénaturé 5 minutes à 100°C avant utilisation. La solution cible composite est alors dénaturée 1 minute à 100°C puis placée sur glace 1 minute et préchauffer à 42°C avant hybridation.

C.4.2.5 Préparation, hybridation et lecture des lames

L'ensemble de ces étapes a été réalisé avec les équipements du pôle Transcriptome de la Plateforme Génomique Fonctionnelle de Bordeaux installée pour partie au sein de l'Institut de Biologie Végétale Intégrative IFR 103 (IBVI) sur le site de l'INRA de Villenave d'Ornon.

Pour une meilleure conservation, les lames sont stockées à température ambiante sans autre traitement après dépôt que l'exposition aux UV pour fixer les sondes sur le support. Pour de meilleurs résultats, il est préférable d'attendre au minimum 2 semaines (3 mois dans notre cas) avant d'effectuer un traitement de préparation (Post-processing) des lames qui doivent dans ce cas être hybridées rapidement (http://microarrays.nki.nl/download/protocols/Post-Processing%20cDNA%20arrays.pdf).

Les lames sont préparées manuellement par trempage dans différent bains de solutions filtrées à 0.45µm et à température ambiante : 2 lavages successifs dans des solutions de SDS à 0.2% pendant 2 minutes, 2 lavages successifs dans de l'eau miliQ pendant 2 minutes, 2 minutes dans de l'eau miliQ bouillante, 5 minutes de séchage en agitant manuellement à température ambiante, 3 lavages successifs dans des solutions de SDS à 0.2% pendant 1 minute, 1 minute dans un bain d'eau miliQ, 10 secondes dans de l'eau miliQ bouillante, séchage pendant 5 minutes et élimination éventuelle de l'humidité résiduelle à l'azote.

Les lames ainsi préparées sont placées dans une station d'hybridation automatique Tecan HS 4800, ce qui permet de réduire les variations entre les différentes expériences d'hybridations. L'ensemble des solutions utilisées ont été filtrées à 0.45µm. La station gère les températures, l'agitation des cibles complexes fluorescentes, la fluidique et le séchage des lames. A la sortie de la station, les lames sont prêtes à être scannées.

Le programme adopté commence par un lavage à 42°C avec une solution de SSC 1X et SDS 0.1X sous agitation pendant 30s. Les solutions cibles préchauffées à 42°C sont introduites dans les chambres d'hybridation dont le volume est de 100µL. L'étape d'hybridation se

déroule à 42°C sous agitation moyenne pendant 16h. Les lavages se font ensuite à 30°C et commencent par une solution SSC 1X et SDS 0.1X pendant 1 min et 30s d'agitation, suivi d'une solution SSC 0.1X et SDS 0.1X pendant 1 min et 30s d'agitation, puis un lavage SSC 0.1X pendant 30s et 30s d'agitation. Un séchage final à l'azote est effectué à 30°C pendant 30s.

Les lames sont conservées à l'abri de la lumière et scannées immédiatement après hybridation sur un scanneur Axon Instruments Genepix 4000 B. La lecture a été réalisée avec un pas de 5µm. La détection des spots s'est faite à l'aide du logiciel Genepix Pro qui permet d'aligner les grilles de « spottage », d'ajuster automatiquement et manuellement la taille des zones de détection, de labelliser les spots de mauvaise qualité et d'extraire les données numériques de fluorescence en générant des fichiers exportables vers un tableur.

C.4.2.6 Traitement des données

La lecture des lames génère des fichiers de taille conséquente, difficiles à exploiter directement sous des tableurs classiques. Pour effectuer l'ensemble des étapes de traitement et manipulation des données il m'a fallu générer un ensemble de scripts sous langage Perl (ActivePerl, http://www.activesate.com). Les analyses statistiques ont été conduites sous l'environnement R de calculs statistiques (http://www.r-project.org/).

Les données brutes de fluorescence sont filtrées, nettoyées et formatées pour permettre l'application de la fonction « rlowess » de la bibliothèque « MAANOVA » (Wu 2003).

Cette fonction va normaliser les données lame par lame par une transformation « LOWESS ou LOESS » (Draghici 2003) en tenant compte de l'intensité et de la position des spots. Elle a pour effet de redresser le nuage de point d'un « scatter plot » représentant les ratios du signal Rouge/Vert en fonction de l'intensité Rouge*Vert.

Les données normalisées sont récupérées et les gènes filtrés sur le nombre de données manquantes. Considérant le nombre de données manquantes pour chaque facteur : condition, provenance et fluorochrome, seuls les gènes ayant moins de 1/3 de données manquantes quelque soit le facteur, sont retenus. Les données sont alors réorganisées et transformées en log de base 2 pour l'analyse statistique.

Nous adoptons une approche similaire à celle développée par le groupe de Gary Churchill (http://www.jax.org/staff/churchill/labsite/software/anova/) et présenté dans la bibliothèque MAANOVA (Wu 2003). Deux analyses de variance consécutives sont appliquées.

Le premier modèle effectue une normalisation de l'ensemble des données. Dans ce cas, les gènes sont considérés comme des répétitions :

Modèle C-1 : $Y_{ijk\lg} = \mu + A_i + D_j + AD_{ij} + Pt_k + Ct_l + PtCt_{kl} + \varepsilon_{ijk\lg}$

Ce modèle permet de fixer les effets globaux techniques de lame (A), du fluorochrome (D) et d'interaction (AD), mais aussi des effets techniques confondus à un effet global Provenance (Pt), Condition (Ct) et interaction (PtCt). On retire donc l'influence de la Provenance et de la Condition sur l'ensemble des gènes. Cette approche permet d'éliminer un effet global de ces facteurs sur l'ensemble des gènes afin de générer des données corrigées en récupérant l'erreur résiduelle (ε) et la moyenne (μ). Ces données corrigées vont nous permettre d'identifier les effets significatifs dans un second modèle.

Le second modèle consiste cette fois en une analyse gène par gène (g) en appliquant le modèle :

Modèle C-2 : $\mu + \varepsilon_{ijk\lg} = A_i + D_j + AD_{ij} + R_r + P_k + C_l + PC_{kl} + e_{ijklr}$

Dans ce modèle, R symbolise les répétitions de dépôts sur la lame (au nombre de 2).
Un des problèmes majeur de ce type d'analyse de données d'expression provient de la multiplicité des tests. Le test d'analyse de variance est en effet répété autant de fois qu'il y a de gènes (dépôts d'ADNc distincts). Le risque α fixé pour un test unique peut amener à un nombre de faux positifs important lorsque le test est répéter des milliers de fois. Les valeurs des tests générées par les ANOVA ont donc été corrigées par la méthode « FDR » (False Discovery Rate) de Benjamini et Hochberg incluse dans la fonction « p.adjust » de R (Benjamini and Hochberg 1995). Le FDR se définit comme la proportion attendue d'hypothèses nulles – ici H0 « il n'y a pas de différences entre les niveaux des facteurs » - rejetées à tort (faux positifs) parmi les hypothèses rejetées. Cette méthode contrôle le taux de faux positifs tout en conservant la proportion attendue de faux positifs (carence de type I) selon l'hypothèse nulle.
Les résultats sont exportés vers un tableur Excel pour réaliser des tris sur les significativités des tests et la part de variation expliquée par le modèle afin de déterminer les gènes régulés par les facteurs Provenance, Condition et Provenance x Condition.

C.4.2.7 Identification des gènes différentiellement exprimés

Les données de l'expérience « stress court » et celles de l'expérience « stress long » sont traitées séparément. La détection des gènes différentiellement exprimés est réalisée en tenant compte des valeurs des probabilités associées au test et de la part de variation expliquée par le modèle et par les facteurs provenance (P), condition (C) et interaction (I). La première

étape de filtrage consiste à retenir uniquement les gènes pour lesquels un effet significatif d'au moins un des trois facteurs est détecté. Cela est obtenu classiquement en fixant arbitrairement un seuil de significativité en dessous duquel l'effet est considéré significatif. Le choix de ce seuil, fixé à 0.01, sera discuté dans la partie « Résultats et Discussion ». Le second filtrage porte sur la part de variabilité expliquée représentée par le coefficient de détermination « r^2 ». La part de variabilité expliquée par les trois facteurs cités plus haut doit être au minimum de 40 % ou au moins un des facteurs explique plus de 33 % de la variabilité. Cela permet de s'assurer que les gènes détectés présentent une part non négligeable de la variabilité expliquée par les facteurs d'intérêt. Comme il le sera discuté lors de la présentation des résultats, ces valeurs ont été fixées en fonction du jeu de données analysé. Aucun filtrage n'a été appliqué concernant un niveau de régulation minimum étant donné les faibles amplitudes de variations observées sur les données normalisées.

C.4.2.8 Agrégation des données (Clustering)

Les données d'expression d'un gène (g) pour une série d'échantillons (n) peuvent être représentées par des vecteurs à n dimensions : $g_g = (e_1, e_2, \ldots, e_n)$

Le clustering se définit comme un processus de regroupement d'entités similaires reposant sur la mesure de distances entre vecteurs à n dimension. Un algorithme regroupe en différentes classes de similarité les entités sur la base de leurs distances respectives. Cet algorithme est donc très sensible à la distance utilisée. Les classes peuvent alors être organisées sous la forme de dendrogramme suivant différentes méthodes comme la classification hiérarchique qui est agglomérative, les k-means qui vont regrouper les données selon un nombre de classes défini par l'utilisateur et le « Self-Organizing Map » (SOM) qui est une méthode non supervisée, basée sur les réseaux neuronaux.

De nombreuses mesures peuvent être utilisées pour évaluer cette similarité. S. Draghci (2003) décrit et compare de nombreuses distances métrique dont les plus courantes sont la distance euclidienne et le coefficient de corrélation de Pearson.

- La mesure de distance euclidienne est la distance classique entre 2 vecteurs à n dimensions basée sur le théorème de Pythagore.
- La distance de corrélation de Pearson mesure la corrélation entre 2 vecteurs à n dimension. Si les profils des gènes sont clustérisés, ceux ayant des niveaux d'expression entre échantillons évoluant dans le même sens seront similaires mais l'influence du niveau d'intensité ne sera pas prise en compte. Un gène fortement exprimé quelque soit les conditions sera similaire à un gène faiblement exprimé tant

que les variations entre échantillon restent semblables. Cependant la méthode est sensible aux entités sortant de la distribution moyenne (« outliers »).

Le choix de la méthode de calcul de la distance permet donc d'orienter le regroupement des données en fonction des caractéristiques que l'on veut mettre en évidence. Pour l'analyse de nos données, nous avons choisi d'utiliser la distance euclidienne.

Le clustering peut être réalisé sur l'ensemble des données. Dans notre cas, nous nous intéresserons uniquement au regroupement des gènes dont les variations d'expression auront été détectées significatives par l'ANOVA.

Nous avons pris comme valeur pour chaque échantillon - c'est-à-dire pour une provenance et une condition donnée -, la moyenne des données corrigées issues du modèle d'ANOVA pour cet échantillon. Pour chaque gène, les valeurs des échantillons sont normalisées en les divisant par la moyenne des données pour ce gène. Cela permet d'éviter un regroupement sur le niveau d'expression global d'un gène tout en conservant la possibilité de regroupement sur l'amplitude de la régulation et sur la norme de réaction. Le calcul d'un ratio d'expression par rapport à un échantillon contrôle normalisera également le niveau d'expression mais sera plus sensible à la valeur du contrôle, ce qui risque de regrouper les gènes en fonction du niveau d'expression du contrôle.

Pour réaliser le clustering, nous avons utilisé le logiciel en ligne EPCLUST (http://ep.ebi.ac.uk/EP/EPCLUST/). Les valeurs centrées sur la moyenne du gène sont passées en log 2 pour obtenir une représentation graphique avec des valeurs autour de 0. Les données sont alors regroupées par similarité des gènes ou des échantillons en utilisant une distance euclidienne et la méthode de classification hiérarchique UPGMA de calcul de la distance moyenne. Nous avons également utilisé le module intégré SOTArray disponible au sein de la suite d'analyse GEPAS (http://gepas.bioinfo.cnio.es/cgi-bin/sotarray) toujours en distance euclidienne et en conservant le seuil de variabilité proposé par défaut pour générer des classes de similarité plus synthétiques.

Une approche complémentaire a été utilisée pour regrouper les données par similarité des profils d'expression en utilisant le module Click de l'outil en ligne Expander de l'université de Tel-Aviv (http://www.cs.tau.ac.il/~rshamir/expander/expander.html). L'algorithme utilise des techniques statistiques et graphiques sans a priori sur la structure ou le nombre de groupes

attendus (Sharan *et al.* 2003). Un paramètre d'homogénéité intra-classe est modulé pour affiner la composition des groupes.

C.4.3 PCR quantitative en temps réel

Le principe de base de cette technique réside en la détection et l'enregistrement de la fluorescence au cours de la réaction PCR ; ce qui permet de calculer la quantité initiale de matrice dans différents échantillons. Cela est possible en comptant le nombre de cycles PCR nécessaires pour atteindre un seuil donné de fluorescence. Dans notre cas, la fluorescence, produite par du SybrGreen™, est proportionnelle à la quantité d'ADN double brin. Le nombre de cycles (Ct) nécessaires pour atteindre le seuil sera d'autant plus faible que le nombre de copies du transcrit étudié sera élevé dans l'échantillon.

L'utilisation de cette technique pour l'estimation de la quantité de transcrits présents dans les échantillons englobe deux approches principales, la quantification absolue et la quantification relative. La première permet d'estimer le nombre de copies d'un transcrit dans un échantillon en se basant sur des courbes de calibration obtenues à partir d'échantillons contenant un nombre de copies connu du transcrit ou ADNc étudié. La seconde, utilisée dans notre cas, permet de déterminer des changements du niveau de transcrits d'un gène entre de multiples échantillons et de l'exprimer par rapport au transcrit d'un gène de référence. En effet, l'utilisation d'un contrôle endogène permet de corriger les variations inter échantillons dues entre autre à des erreurs de quantification des ARNt. Le choix de ce gène de référence est délicat, toute variation de son niveau d'expression atténuera voire masquera les changements réels du transcrit étudié et génèrera des artéfacts d'expression (Bustin 2000).

C.4.3.1 Experimentation

Pour prendre en compte les variations expérimentales lors de la synthèse d'ADNc et réduire les variations intra échantillon, 3 synthèses indépendantes ont été réalisées pour chaque échantillon La méthode adoptée est similaire à celle décrite en C.4.1 avec utilisation d'oligo dT. Les ADNc sont dilués au 10ème avant utilisation. Les réactions PCR sont conduites sur un robocycleur Chromo4 (MJ Geneworks, Inc) avec les paramètres suivant : 95°C pendant 3 min, puis 40 cycles à 95°C pendant 15s et 60°C pendant 45s et une courbe de fusion finale de 65°C à 95°C par 0.5°C d'intervalle. Le volume réactionnel de 20 µL contenait du « iQ™ SYBR Green Supermix » (Bio-Rad Laboratories, Inc, CA) à concentration 1X, les amorces du gène d'intérêt à 150nM et 3µL d'ADNc. La fluorescence est enregistrée à la fin de chaque

cycle par le logiciel Opticon Monitor 3.1.32 (Bio-Rad Laboratories, Inc, CA) qui permet également de déterminer la valeur seuil de Ct et l'exportation des données de fluorescence.

L'estimation de l'efficacité de la réaction est faite selon la méthode classique des courbes de dilution et du calcul de la pente (http://www.gene-quantification.de/). Pour chaque rétrotranscription, un « pool » des échantillons est utilisé dilué au $1/10^{éme}$, $1/100^{éme}$, $1/1000^{éme}$ et $1/10000^{éme}$. La représentation graphique du nombre de cycles (Ct) en fonction du logarithme de base 10 de la quantité initiale permet de déterminer une droite de régression dont la pente nous informe sur l'efficacité de la réaction en pourcentage selon la formule : $10^{-((1/pente)-1)}$.

Les données brutes sont corrigées en choisissant de leur soustraire la valeur de la ligne de base déterminée par la moyenne de la fluorescence comprise entre le $3^{ème}$ cycle et le $14\text{-}16^{ème}$ selon le gène étudié. La valeur du seuil de fluorescence est déterminée manuellement en se plaçant dans la partie linéaire de la courbe du logarithme de base 10 de fluorescence en fonction du nombre de cycles.

C.4.3.2 Choix des amorces

Le choix des amorces est particulièrement important en PCR quantitative. Dans notre cas, nous ne connaissons pas a priori la position des introns et exons. Il est donc indispensable de travailler sur des ARN traités à la DNAse pour éviter toute amplification sur l'ADN génomique. Il est préférable de rechercher à dessiner les amorces dans la région 3' des gènes et de lancer une recherche d'homologie sur le fragment étudié pour déterminer les régions les plus spécifiques. Le risque est de co-amplifier différents membres de familles multigéniques régulées de manière différentes.

Pour chaque gène étudié, les amorces ont été définies à l'aide du logiciel en ligne Primer3 http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi.

Les principales conditions à respecter sont :

- Amplification de fragments courts, taille recommandée 100-150pb
- Teneur en G/C basse préférable, tolérée entre 20% et 80% mais 50-60% recommandée
- Température de fusion (Tm) des paires d'amorces proche de 60°C
- Différence entre Tm ≤ 0.5°C si possible
- Pas de formation de dimères
- Pas plus de 4 nucléotides identiques consécutifs, spécialement les G
- Pas plus de 2 G ou C dans les 5 derniers nucléotides (extrémité 3')

Dans cette étude, nous avions choisi les amorces afin d'amplifier des fragments d'une longueur comprise entre 250 et 350 pb ce qui nous laissait éventuellement la possibilité d'un marquage radioactif et leur utilisation pour des études d'hybridation sur filtres.

C.4.3.3 Choix des gènes de référence

Parmi les gènes les plus utilisés pour normaliser les données d'expression que l'on rencontre dans la littérature, on peut citer l'isoforme beta de l'actine, la Glyceraldehyde-3-phosphate déshydrogénase, l'Ubiquitine C ou certaines protéines ribosomales comme la 60S P0, L13a, L10 ou 28S et 18S (Vandesompele *et al.* 2002). Le gène idéal serait exprimé à un niveau constant dans tous les tissus d'un organisme, à tous les stades de développement, et ne serait pas affecté par les traitements appliqués. En pratique, un gène adapté dans un cas n'est pas forcément transférable à une autre expérimentation.

Pour notre étude, nous avons recherché dans les données d'expression des microarrays un gène dont l'expression dans les racines ne variait statistiquement pas quelque soient les conditions et la provenance. Après sélection de 3 gènes, un couple d'amorce a permis de détecter par PCR quantitative des niveaux d'expression parfaitement stables dans toutes les conditions. Il s'agit d'un ARNm codant pour une protéine ribosomale 40S S27.

C.4.3.4 Traitement des données de fluorescence

Les valeurs sont exportées vers une macro Excel nommée GENEX – Gene Expression Macro (v1.10–2004, Bio-Rad) accessible à l'adresse http://www.gene-quantification.de/download.html#genex.

La méthode utilisée dans cette macro est dérivée de celle décrite par Vandesompele (Vandesompele *et al.* 2002) et très proche de celle proposée par Pfaffl (Pfaffl 2001). Elle permet de prendre en compte les données d'un gène de référence ainsi que l'efficacité de la réaction contrairement à la méthode classique du ΔΔCt où :

$$\Delta\Delta Ct = \Delta Ct_{\text{échantillon 1}} - \Delta Ct_{\text{échantillon 2}}$$

avec $\Delta Ct = Ct_{\text{gène}} - Ct_{\text{gène référence}}$ pour un échantillon donné.

Le gène de référence, d'expression supposée constante dans l'ensemble des échantillons étudiés, a pour but de corriger les variations expérimentales reliées à l'échantillon.

Dans le cas de la macro, une quantité relative (QR) pour un gène (j) est calculée à partir des Ct moyens de chaque échantillon (i) et d'un échantillon contrôle pour tenir compte de l'efficacité :

$$QR_{ij} = \text{Efficacité}_{j}^{\left(Ct_{j} \text{ de l'échantillon contrôle} - Ct_{ij} \text{ de l'échantillon X}\right)}$$

L'échantillon contrôle peut être choisi arbitrairement ou l'échantillon avec le Ct moyen le plus faible.

La normalisation par un ou plusieurs gènes de référence (k) est effectuée en calculant un facteur de normalisation pour chaque échantillon (i) :

$$\text{Facteur de Normalisation}_{i} = \left(QR_{i1} \times QR_{i2} \times \ldots \times QR_{ik}\right)^{\left(\frac{1}{k}\right)}$$

qui est la moyenne géométrique des quantités relative des (k) différents gènes de référence pour l'échantillon (i).

Le niveau d'expression du gène (j) pour l'échantillon (i) s'exprime alors sous la forme :

$$\text{Niveau d'Expression}_{ij} = \frac{\text{Quantité relative}_{ij}}{\text{Facteur de Normalisation}_{i}}$$

Les données aberrantes et isolées, c'est-à-dire ayant une valeur de Ct totalement différente des deux autres répétitions et des autres échantillons, ont été omises.

La macro nous donne alors la possibilité de représenter graphiquement le niveau d'expression des différents échantillons et l'écart-type associé aux 3 synthèses d'ADNc (Fig. Ax-3).

C.5 Analyses Protéiques

Les tissus analysés correspondent aux parties racinaires de l'expérience « Stress Long » et sont identiques à ceux utilisés pour les analyses transcriptomiques. L'étude du protéome a été réalisée par Manon Moreau (formateur : Céline Lalanne) dans le cadre de son Master.

C.5.1 Extraction des protéines et Electrophorèse bidimensionnelle

Le protocole utilisé est inspiré de celui de (Damerval *et al.* 1986), disponible à l'adresse http://www.pierroton.inra.fr/genetics/2D/Proteomevert/Protocoles/protobidinew.pdf .

Pour chaque échantillon de l'expérience « Stress Long », 4 répétitions ont été effectuées, soit un total de 16 gels.

C.5.2 Analyse des gels

Les gels colorés au bleu de coomassie sont scannés et analysés en utilisant la suite logicielle et matérielle Amersham Biosciences (Labscan et ImageMaster Platinum v5.0) (Gion *et al.* 2005). Les paramètres utilisés lors de l'acquisition sont : mode transparent, filtre rouge et sensibilité 300 dpi. Après détection automatique des spots, les gels sont repris manuellement et nettoyés pour éliminer les spots artéfactuels. Seuls les spots détectés sur au moins deux gels du même échantillon sont conservés. Le volume de chaque spot est calculé sur la base de sa surface et des niveaux de gris enregistrés. Un gel de référence créé à partir de l'information de l'ensemble des gels permet de construire une carte générale de tous les spots détectés. Les gels sont comparés à cette carte de référence afin de donner la même identité au même spot présent sur plusieurs gels.

C.5.3 Analyse quantitative

Une analyse de variance (ANOVA) avec les facteurs Provenance (*P*) et Condition (*C*) comme variable dépendante a été réalisée selon le modèle :

$$Y_{ijk} = \mu + P_i + C_j + P_i C_j + \varepsilon_{ijk}$$

Ce modèle permet d'identifier les spots protéiques significativement spécifiques d'une provenance, d'une condition ou de l'interaction entre provenances et conditions. Le seuil de significativité choisi a été placé à 0.001 pour tenir compte de la multiplicité des tests (120 spots détectés).

C.5.4 Choix des spots à séquencer

Etant donné le nombre conséquent de spots présentant au moins un effet significatif, seuls ceux présentant un effet d'interaction ont été prélevés dans un premier temps. Ces spots peuvent correspondre à des stratégies différentes de réponse au stress entre les deux provenances. Dans un second temps, quelques spots montrant uniquement un effet provenance ont été prélevés ainsi que des spots situés relativement proches les uns des autres sur gel (variation qualitative de type présence/absence) et présentant au moins un effet provenance, dans le but d'identifier des formes allèliques d'une même protéine. Quelques spots montrant un effet traitement ont également été prélevés dans le but d'identifier des

mécanismes généraux de réponse à la sécheresse, indépendamment de la provenance. Enfin, des spots spécifiques des marocains stressés ont été retenus.

C.5.5 Séquençage et identification des protéines

Digestion trypsique et Spectrométrie de masse

Les spots prélevés sont lavés par une solution de bicarbonate d'ammonium 25 mM et d'acétonitrile (ACN) 50% pendant 2H. La solution est alors remplacée par de l'ACN et incubée 10 min pour déshydrater les spots. Après 5 min de séchage au speedvac, les protéines sont digérées une nuit à 37°C par la trypsine, qui coupe au niveau des lysines et des arginines, en présence d'HCl 1 mM et de bicarbonate d'ammonium 50 mM.

Les peptides hydrophiles sont extrait après ajout de 10 µL de bicarbonate d'ammonium 50 mM, incubation 10 min sous agitation, centrifugation 30 s, et récupération du surnageant.

Les peptides hydrophobes sont récupérés par 3 extractions successives commençant par une incubation de 10 min des spots précédents en présence de 10 µL de tampon d'extraction (eau/ACN/Ac acétique 47.5/47.5/5) puis récupération du surnageant.

Les différentes extractions d'un même spot sont réunies puis le volume est ramené à 25 µL environ par évaporation au speedvac à température ambiante avant ajout de 0.006 µL d'acide acétique par µL d'extrait.

La partie séquençage protéique a pu être effectuée en utilisant les équipements disponibles au sein du pôle protéomique de la Plate-forme de Génomique Fonctionnelle de Bordeaux

Les mélanges peptidiques sont analysés par spectrométrie de masse en tandem comme décrit dans l'article de Jean-Marc Gion *et al.* (2005) sur le protéome de la formation du bois chez le pin maritime. Le principe de la technique utilisée se base sur la séparation des peptides par chromatographie liquide en phase inverse en fonction de leur hydrophobicité puis analyse des peptides ionisés par un spectromètre de masse produisant un spectre MS. Certains de ces peptides sont sélectionnés, fragmentés en sous peptides et le plus souvent en acides aminés et ré-analysés pour obtenir des spectres MS/MS. On dispose donc pour chaque spot protéique d'un nombre variable de peptides de séquence connue avec une précision variable. Une analyse informatique est alors nécessaire pour identifier une protéine à partir des séquences peptidiques disponibles. Les spectres MS/MS des peptides sont comparés de façon automatique (programmes Turbo Sequest et Sequest Browser) aux spectres théoriques générés par digestion *in silico* d'une banque d'EST traduits de pin et de la base protéique SwissProt.

C.6 Analyses Bioinformatiques

C.6.1 Analyse et Assemblage des séquences

Les électrophorégrammes ont été analysés par le logiciel « Phred » qui détecte la séquence en acides nucléiques et génère un score de qualité global de la séquence et un score qualité individuel par base. Un outil de comparaison rapide de deux lots de séquence, inclus dans un script Perl écrit par J-M. Frigerio, permet d'enlever les séquences du vecteur pour éditer rapidement et manuellement chaque séquence et obtenir un score Phred 20 de qualité global de la séquence supérieur à 85%. Dans une séquence, Phred attribue à chaque pic de l'électrophorégrammes une probabilité que la base nucléotidique détectée soit réellement la bonne. Le score est généralement un nombre entre 1 et 40 où le nombre représente la puissance de 10. Un score Phred 20 signifie qu'une base n'est pas la bonne une fois sur 100 (10^2) et un score Phred moyen de 20 signifie que la séquence est mauvaise une fois sur 100.

Le principe de l'assemblage bioinformatique d'EST est basé sur le fait que nous ne disposons que de courts fragments de séquences du transcrit d'un gène. En effet, comme nous l'avons déjà évoqué, si l'on considère les transcrits d'un gène disponibles après extraction, tous ne seront pas pleine longueur et donneront donc des ADNc partiels. De plus, la transcription inverse utilise comme amorces des polyT qui vont se fixer sur la queue polyA des ARNm, mais la réaction n'est pas parfaite et produit des ADNc incomplets dans la partie 5' (Fig. C-15). Etant donné que nous avons choisi un séquençage d'EST en 5', nous obtiendrons en général les 600-700 premières paires de bases en 5'. Il en résulte une série de fragment de séquences réparties le long du gène. L'objectif de l'assemblage est de rechercher les homologies de séquences entre EST pour reconstituer la séquence complète du gène à partir des fragments de séquences. Cela est rendu possible de proche en proche grâce aux zones de recouvrements des différents EST. Si des séquences se recouvrent en 5' et d'autres en 3' mais qu'aucune séquence n'est disponible pour la partie centrale, alors deux groupes d'EST seront générés pour un même gène. Plus le nombre de séquences disponibles pour un gène est élevé, plus on a de chance de reconstituer l'intégralité du gène. Les séquences qui n'ont pas de régions suffisamment spécifiques homologues à d'autres EST donneront des singletons. Pour éviter de regrouper les EST sur la base de régions de faibles complexité (ex : régions répétées comme les microsatellites), il est indispensable d'effectuer un masquage de ces régions.

Les séquences de bonne qualité sont injectées dans une suite logicielle nommée EPA, développée et testée en collaboration avec le CBiB (Centre de Bioinformatique de Bordeaux).

Cet outil est construit autour d'une base de donnée MySQL et du module StackPack™ (Electric Genetics www.egenetics.com) qui permet un assemblage rapide et un alignement des séquences. Les séquences sont analysées par le programme RepeatMasker (http://www.repeatmasker.org/) pour masquer les zones de faible complexité comme par exemple les queues polyA ou les microsatellites. Les séquences sont tout d'abord regroupées en **cluster** (CL) par le module D2cluster (Burke *et al.* 1999) qui découpe les séquences en motifs hexanucléotidiques et recherche une suite de motifs identiques entre les séquences (Fig. C-18). Ce premier regroupement permet un assemblage plus fin en **contigs** (CT) par le programme Phrap (http://www.phrap.org/). Des séquences **consensus** (CN) sont alors générées par le module CRAW (Burke *et al.* 1998) qui va analyser les alignements produits par Phrap. Seul le consensus primaire (CN primaire) est conservé pour l'étape d'annotation.
Les EST qui n'ont pu être rapproché d'un contig sont appelés **singletons**.
A titre de comparaison, la stratégie adoptée par le TIGR (http://www.tigr.org/tgi/faq2.shtml) commence par un contrôle et nettoyage de la présence de séquences de vecteur, *E. coli* et/ou poly A/T. Les séquences de taille inférieure à 100 pb ou comportant plus de 3% de N ne sont pas prises en compte. Pour l'assemblage, les EST sont clustérisés sur la base d'un minimum de 40 bases homologues, plus de 94 % d'identité dans la zone de recouvrement et une région d'un maximum de 30 paires de bases n'alignant pas avec les autres EST clustérisés. Le programme utilisé pour la création des séquences TC (Tentative Consensus) est « Paracel Transcript Assembler »

La structure de la base de données donne la possibilité de travailler sur des lots de séquences organisés en projets. Dans notre cas, un premier projet regroupe l'ensemble des séquences de la banque « xylème » (projet « Pinus Pinaster xylem EST ») et un second projet regroupe celles de racines (projet « Pinus Pinaster root EST »). Puis un projet global rassemble l'intégralité des séquences (« Pinus Pinaster Total »).

Avec le CBiB nous avons développé une interface web disponible pour les projets publics à l'adresse (http://cbi.labri.fr/outils/SPAM/COMPLETE/). Elle offre un accès commun aux requêtes et résultats d'homologie (Blast), aux informations de séquences des EST et consensus, ainsi qu'aux annotations fonctionnelles comme KEGG (Kyoto Encyclopedia of Genes and Genomes), FunCat (The Functional Catalogue) et GO (Gene Ontology).

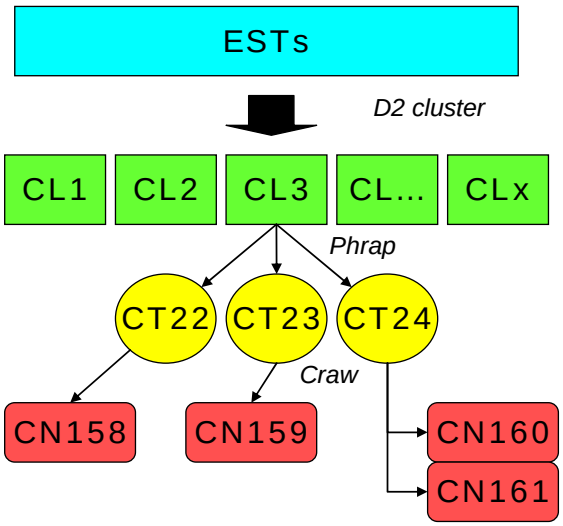


Figure C-18 : Différents niveaux de regroupement des ESTs lors de l'assemblage bioinformatique. Les ESTs sont regroupés en « clusters » (CL) par le programme D2 cluster. L'alignement des ESTs de chaque CL est analysé par le programme Phrap qui génère des « contigs » (CT). Le programme Craw va déterminer une séquence consensus (CN) à partir des ESTs de chaque contig. Si une seule séquence consensus ne peut être déduite, des consensus alternatif sont générés.

Les homologies de séquence des EST et des séquences consensus ont été recherchées dans les différentes bases nucléiques et protéiques publiques. Une stratégie en cascade a été utilisée pour diminuer les temps de calcul. Une première requête BlastX contre la base SwissProt est testée et les meilleurs résultats sont conservés. Si aucune homologie n'est trouvée avec un indice de confiance (E-Value) $\leq 10^{-5}$, la même requête est testée su la base TrEMBL et si toujours aucune homologie n'est trouvée avec une E-Value $\leq 10^{-5}$, le résultat de la requête Blastn contre la base EMBL est retenu.

C.6.2 Annotation des Consensus et Singletons

Chaque singleton et consensus a été annoté en fonction du résultat des recherches d'homologie (BLAST). La concordance de l'annotation choisie pour la séquence consensus avec les homologies trouvées pour les EST composant le contig est vérifiée. Si besoin, une nouvelle recherche d'homologie sur d'autres bases de donné est lancée. Elle doit permettre de déterminer si le conflit d'annotation observé entre l'annotation attribuée à partir du consensus et celle de l'EST vient du défaut de séquence informative dans l'EST par rapport au consensus ou bien de l'assemblage au sein d'un même consensus d'EST correspondant à des gènes différents.

Une fois les EST annotées, il est intéressant de les classer par catégories fonctionnelles. Une annotation fonctionnelle a été entrée chaque fois que possible dans la base en suivant la classification de « Gene Ontology ».
Par ailleurs, nous avons regroupé les informations disponibles sur les séquences du projet LIGNOME dans un tableur Excel pour affiner l'annotation et leur attribuer une catégorie fonctionnelle. Ce travail s'est donc focalisé uniquement sur les informations disponibles pour les tissus racinaires, objet d'étude de cette thèse.
Une même enzyme peut porter des appellations synonymes. Pour éviter les confusions et considérer deux gènes comme différents et non comme membre d'une même famille multigénique, nous avons systématiquement recherché la dénomination officielle des enzymes sur le site Expasy (http://www.expasy.org)». Nous avons choisi de suivre la classification fonctionnelle utilisée par le MIPS (Munich Information center for Protein Sequences) mais sans utiliser les classes « cellular biogenesis », « ionic homeostasis » et « development » pour limiter des erreurs d'affectation de notre part. Nous avons ainsi réparti les EST dans les catégories suivantes : croissance cellulaire/division, organisation cellulaire, énergie,

métabolisme, communication / transduction du signal, transport, trafic intracellulaire, transcription, synthèse de protéines, devenir des protéines, réponse au stress, non classée.

C.6.3 Niveau d'expression *in silico* des EST (Northern electronique)

Le séquençage systématique d'EST de banques non normalisées permet d'accéder à une information sur l'expression des gènes séquencés. En effet, en comptant le nombre de fois que l'ARNm d'un gène donné est séquencé dans une banque, il est possible d'estimer le niveau d'expression de ce gène par rapport à la totalité des EST séquencés. Si plus d'une banque est séquencée, il est alors théoriquement possible de comparer les niveaux d'expression d'un gène entre les différentes banques. Dans notre cas, nous disposions d'une banque de racines témoin et d'une banque de racines stressée. Une analyse statistique de l'abondance des transcrits des différents gènes est donc possible. Ce Northern électronique, encore appelé Northern *in silico*, e-Northern ou digital Northern, permet de détecter les gènes significativement différentiellement exprimés entre les deux conditions. Nous avons utilisé la suite logicielle « IDEG6 » (Identification Differentially Expressed Genes 6 test statistics) spécifiquement développée pour ce type d'analyses et accessible en ligne à l'adresse http://telethon.bio.unipd.it/bioinfo/IDEG6/ (Romualdi *et al.* 2003). Cette suite tient compte du nombre total de séquences de chaque banque pour effectuer une normalisation du nombre d'EST par contig. Les tests qui ont été utilisés sont les suivants : test d'Audic & Claverie, test exact de Fisher, test du Chi², test R de Stekel & Falciani, test du Chi² multiple.

C.6.4 Annotation des spots protéiques

Les spectres MS/MS des peptides sélectionnés sont comparés aux spectres théoriques générés par digestion *in silico* de banques d'EST traduits (Pinus Gene Index du TIGR) et de banques protéiques (Swissprot). La comparaison est automatisée à l'aide des modules de recherche Turbo Sequest et Sequest Browser. Une liste de propositions d'assignations de séquences est générée et différents paramètres vont permettre de sélectionner les propositions les plus pertinentes :

- Le Xcor traduit la corrélation entre les spectres MS/MS expérimental et théoriques et doit supérieur à 1.9 pour un ion monochargé, 2.2 pour un dichargé et 3.75 pour un trichargé.
- Le ΔCn représente la différence de score trouvée entre les candidats consécutifs lors de la recherche d'homologie avec des séquences connues. Il doit être supérieur à 0.1 entre les deux premiers candidats.

- Le nombre d'ions fragments trouvés par rapport au nombre d'ions fragments attendus. Il faut retrouver au moins 2 peptides pour que l'assignement soit considéré comme valable.

L'annotation est alors attribuée sur la base d'une recherche d'homologie (Blast) de la séquence obtenue précédemment sur l'identité des spectres MS/MS avec les bases de données publiques protéiques. L'utilisation des bases de données *Pinus* du TIGR, qui contiennent 320000 EST dont les EST de pin maritime, est utile à l'identification de protéines spécifiques au genre *Pinus*. Toutes les données protéiques des échantillons, aux spectres, en passant par la quantité des protéines, seront prochainement disponibles à l'URL (http://moulon.inra.fr/~langella/proticdemo/Protic/home/index.php?menu=&file=credits) grâce à la base de données PROTICdb (Ferry-Dumazet *et al.* 2005) développées par l'équipe Bioinformatique du Laboratoire Commun de Biologie Moléculaire et de Bioinformatique de l'INRA du moulon en collaboration avec le CBiB et notre unité.

Résultats et discussion

D RESULTATS ET DISCUSSION

D.1 Matériel végétal et paramètres physiologiques

D.1.1 Culture Hydroponique

Les premières levées ont eu lieu 3 semaines après la mise en germination et les plantules de la provenance marocaine ont atteint une taille suffisante pour être transférées en hydroponie au bout de 4 semaines. La taille minimale est celle de la radicelle qui doit être d'environ 5 cm pour atteindre le milieu liquide dans notre dispositif. Les plants d'origine landaise ont levé quelques jours plus tard et ont pu être transférés en hydroponie 5 semaines après la mise en germination. Au moment du transfert, les plants présentaient un flétrissement marqué sur une période de 24h. L'éclairage a donc été coupé pour éviter tout échauffement. Une cinquantaine de plants ont été placés par bac de culture.

Lors des deux premières séries de culture, l'attaque par le *Fusarium sp.* s'est déclarée durant la première semaine d'exposition au stress. La troisième mise en culture sur solution nutritive diluée au 1/5^{e} en présence d'antifongique a permis d'éviter une autre attaque.

Après 8 semaines de culture, les jeunes plants sont très vigoureux. Les feuilles aciculaires (euphylles) d'un vert sombre et intense ne présentent aucun défaut de pigmentation signe d'une chlorose ou autre carence. Les premières pseudophylles se sont développées sur certains individus (Fig. D-1a). Les racines sont très bien développées (Fig. D-1b). Le port du plant diffère par contre nettement de celui des semis de pin maritime en conditions naturelles qui sont élancés, peu ramifiés et à entrenœuds longs (Fig. D-2a). Les plantules sont dans le cas présent très ramassées, à entrenœuds très courts et portent de nombreuses ramifications (Fig. D-2b). Au moment du transfert du milieu nutritif vers le milieu hyper osmotique, les plants présentent un flétrissement rapide qui disparaît progressivement au cours des premières heures.

D.1.2 Potentiel de base

La mesure du potentiel de base étant destructive, cinq individus ont été prélevés pour chacun des écotypes et conditions analysés. Les échantillons ont été prélevés dans les conditions « témoin court » et « 48 heures de stress » pour l'étude de la réponse à court terme et dans les conditions « témoin long » et « stress long » pour l'étude de la réponse à plus long terme. L'obscurité totale a pu être maintenue le temps des prélèvements et des mesures, de 4

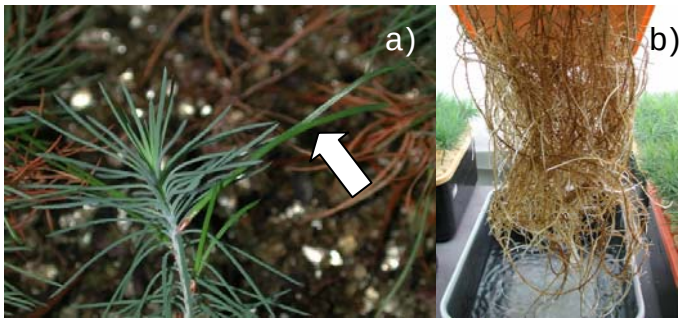


Figure D-1 :
a- Développement des aiguilles secondaires ou pseudophylles
b- Développement du système racinaire de 48 plantules âgées de 7 semaines

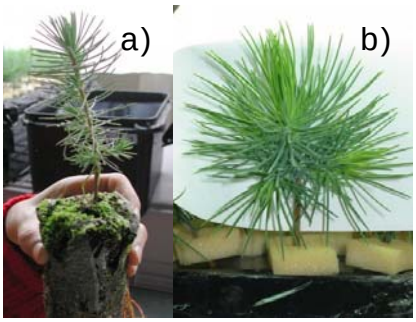


Figure D-2 :
a- Morphologie des jeunes plants de pin maritime élevés en conditions naturelles
b- Développement à entrenœuds courts des plants élevés en hydroponie

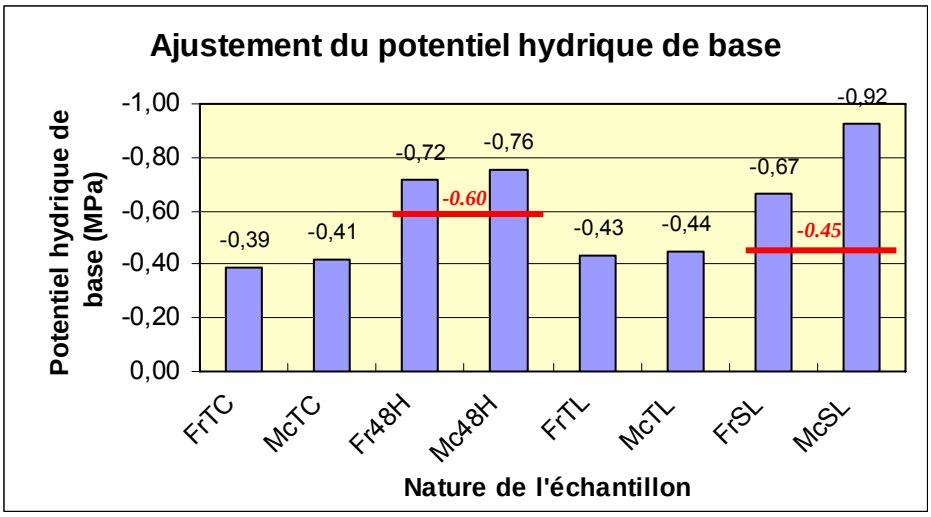


Figure D-3 : Valeurs moyennes et écart-types des potentiels hydriques de base en MPa pour les différents échantillons analysés. (Fr = Provenance Française, Mc = Provenance Marocaine, TC = Témoin Court, TL = Témoin Long, 48H = 48H de Stress, SL = Stress Long)

facteur	Stress Court	Stress Long
Provenance	*0,2712*	**8,95e-07**
Condition	**1,96e-09**	**7,48e-13**
Interaction	*0,8545*	**2,77e-06**

Tableau D-1 : Résultats des analyses de variance (P-Value) pour tester l'hypothèse nulle H0 = « Les potentiels hydriques de base (Ψh) des échantillons ne sont pas différents » sur les mesures de l'expérience « Stress court » et de l'expérience « Stress Long ». Les facteurs testés dans le modèle sont la Provenance (P), la Condition (C) et l'Interaction. (Ψh = P + C + PC + ε)

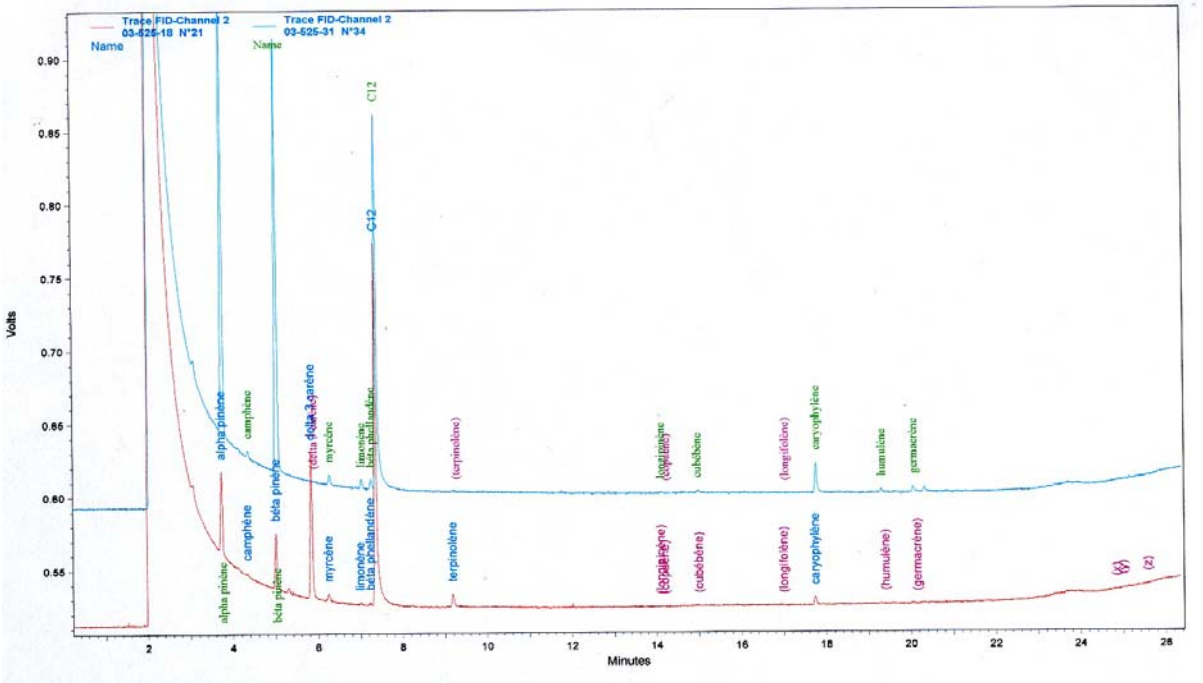


Figure D-4 : Chromatogramme de détection des composés terpéniques obtenus par chromatographie en phase gazeuse. Intensité du signal au niveau du détecteur (en ordonnée) en fonction du temps de rétention (en abscisse).

heures à 10 heures du matin. Les valeurs moyennes obtenues et les écart-types sont représentés dans le graphique (Fig. D-3). Le potentiel hydrique de base reste stable autour de -0.4MPa (soit -4 bars) pour les deux provenances en l'absence de stress osmotique. En situation de stress osmotique violent lors de l'expérience stress court, le potentiel osmotique de la solution nutritive atteint -0.6MPa et les deux provenances répondent en ramenant leur potentiel hydrique de base à une valeur comprise entre -0.72 à -0.75MPa (soit -7.2 et -7.5 bars). L'exposition à un stress de 3 semaines dans une solution nutritive à -0.45MPa permet de différencier les deux provenances quant à leur réponse au stress. Pour les plants d'origine marocaine, le potentiel hydrique passe de -0.44MPa en l'absence de stress à -0.92MPa contre -0.66MPa pour la provenance landaise. L'écotype marocain aura donc plus de facilité pour capter l'eau du milieu. Une analyse de variance sur les valeurs de potentiel hydrique de base avec comme variables explicatives la provenance, la condition et l'effet d'interaction a été conduite séparément sur les données de l'expérience « Stress court » et « Stress long ». Les probabilités de rejet de l'hypothèse nulle sont rapportées dans le tableau D-1. Dans le cas du stress court, seul l'effet condition est significatif. Pour l'expérience « Stress Long », on observe que les effets provenance, condition et interaction sont hautement significatifs. Dans ce cas, le potentiel hydrique de base diffère entre provenances en condition de stress (Test de Student : P-value=1.23^{-7}) mais pas en condition témoin (Test de Student : P-value=0.72). Les deux provenances ne différent donc que pour l'exposition au stress long.

D.1.3 Dosage des terpènes

Des études sur des plantes méditerranéennes, ont montré que la sécheresse / température avait une influence sur l'évolution des concentrations en terpènes (Llusià and Penuelas 1998; Llusia and Penuelas 2000). Une étude sur *Q. ilex* et *P. halepensis* a montré une relation entre la teneur en terpènes émis et respectivement le taux de photosynthèse (PFD) / la température pour le chêne / le pin (Penuelas and Llusia 1999). Une autre étude portant sur *Pinus sylvestris* et *Picea abies* soumis à la sécheresse a montré une augmentation en monoterpène chez ces deux espèces (Turtola *et al.* 2003). Les terpènes sont donc des marqueurs potentiels du stress hydrique.(Llusià and Penuelas 1998)

La technique de détection des terpènes par chromatographie en phase gazeuse permet l'identification et la quantification de 16 composés terpéniques : des monoterpènes $C_{10}H_{16}$ (alpha pinène, beta pinène, delta 3 carène, camphène, limonène, myrcène, beta phellandrène, terpinolène), des sesquiterpènes $C_{15}H_{24}$ (caryophyllène, copaène, cubébène, germacrène, humulène, longifolène, longipinène) et les C12 (Fig. D-4). L'extraction des terpènes par le

pentane concernait uniquement les parties aériennes des plants contrôles et 3 semaines de stress de l'expérience « Stress long » pour laquelle nous avions détecté une différence sur l'ajustement du potentiel de base. Les terpènes en C12, le longipinène et le copaène n'ont pas été détectés dans ces échantillons. Les terpènes les plus abondants sont l'alpha pinène, le beta pinène et le delta-3 carène. Pour chaque condition de chaque provenance, de 5 à 6 plants ont été prélevés et les dosages ont été dupliqués. Les quantités de terpènes mesurées dans les prélèvements ont été ramenées à une teneur par gramme de matière fraîche. Le tableau D-2 regroupe les valeurs de probabilité associées au test F des analyses de variance. Le modèle testé prend en compte un effet de la provenance, de la condition et de l'interaction.

En considérant un seuil de significativité à 5%, l'effet provenance est significatif pour l'alpha pinène, le camphène, le limonène le caryophyllène et l'humulène. L'effet condition n'est mis en évidence pour le delta 3-carène, le myrcène, le terpinolène, le caryophyllène et l'humulène.

Sur la figure D-5a, il apparaît que la teneur en **alpha pinène** est légèrement supérieure dans la provenance marocaine mais c'est surtout le rapport alpha pinène sur béta pinène, nettement plus élevé dans la provenance marocaine qui différencie les deux écotypes (Fig. D-5b).

Le **delta 3-carène** est pratiquement indétectable dans la provenance marocaine contrairement aux plants d'origine landaise qui ont tendance à accumuler ce composé au cours du stress même si une forte variation de concentration est observable (Fig. D-5c). Ce type de profil se retrouve également pour le **terpinolène** (Fig. D-5d).

La teneur en **caryophyllène** est globalement plus élevée dans la provenance marocaine et augmente au cours du stress.

Le **cubébène** et le **germacrène** ne semblent s'accumuler que chez les plants stressés d'origine landaise.

D.1.4 Discussion

La culture de plants de pin maritime en condition hydroponique a été problématique en raison des attaques fongiques rencontrées au cours des deux premières expérimentations. Le maximum de précautions possibles avait pourtant été mis en œuvre pour éviter ce type de contaminations. Les deux premières cultures se sont déroulées dans une salle climatisée et équipée d'un éclairage horticole dont l'accès était restreint. Le nettoyage complet à l'eau de javel de la salle et du matériel, et l'utilisation d'eau distillée associée à l'autoclavage des milieux n'ont pas suffit à éviter le développement de *Fusariu*m. Le fait que l'apparition du pathogène coïncide avec l'application du stress laisse supposer que le *Fusarium* profite d'un affaiblissement du système de défense des plants provoqué par le stress osmotique. Les

Composé	Modèle avec Interaction		
	Provenance	Condition	Interaction
rapport alpha/beta	5,00e-03	0,98	0,46
alpha pinène (C10)	8,00e-04	0,38	0,57
Camphène (C10)	9,00e-04	0,34	0,14
béta pinène (C10)	0,79	0,52	0,38
delta 3 carène (C10)	0,28	0,04	0,14
Myrcène (C10)	0,84	0,03	0,13
Limonène (C10)	0,03	0,24	0,54
béta phellandrène (C10)	0,7	0,15	0,37
Terpinolène (C10)	0,37	0,02	0,09
Longifolène (C15)	0,16	0,14	0,1
Caryophyllène (C15)	0,02	0,04	0,86
Humulène (C15)	0,017	0,027	0,17
Germacrène (C15)	0,51	0	1,00e-04
Cubébène (C15)	*1*	*1,40e-03*	*0,015*

Tableau D-2 : Valeurs de probabilité (P-value) des analyses de variance sur les dosages des composés terpéniques extraits des parties aériennes. Les valeurs significatives à 5% sont indiquées en couleur foncée. Pour le germacrène et le cubébène, très peu de valeurs étaient disponibles.

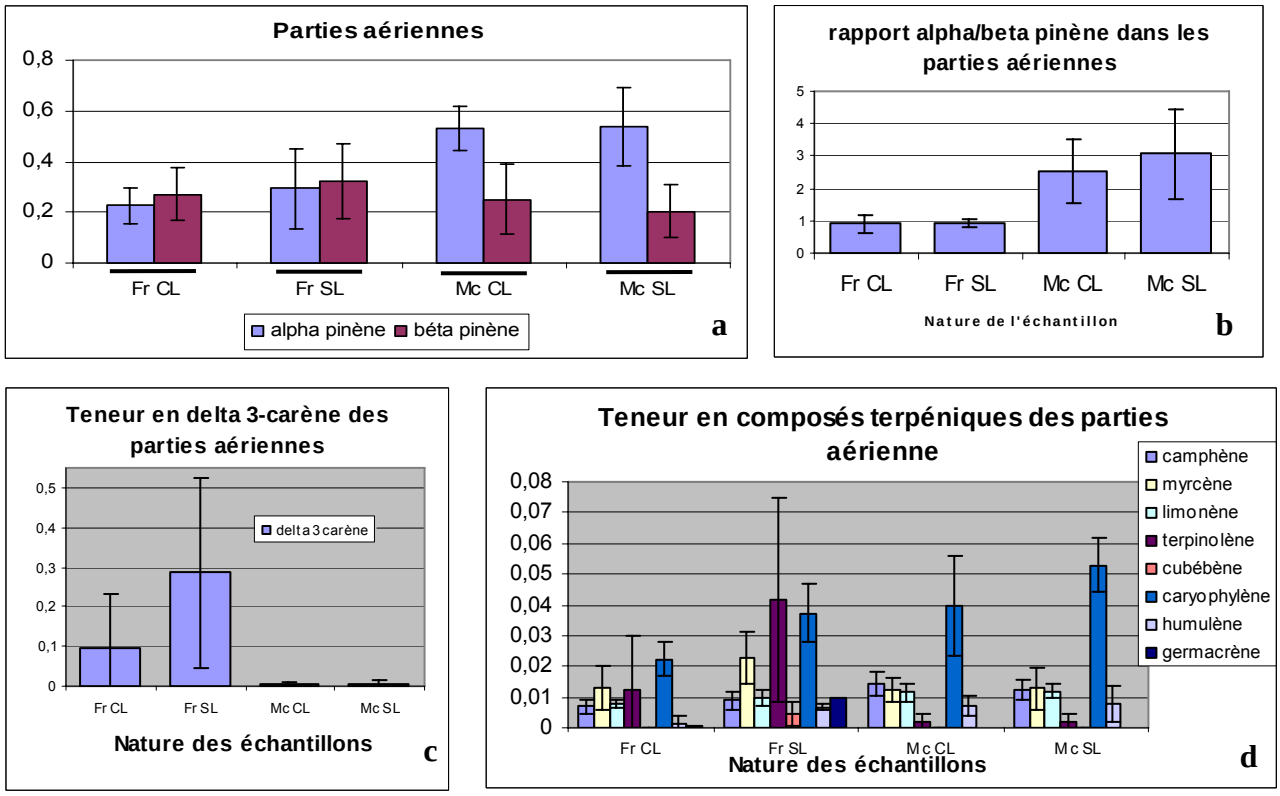


Figure D-5 : Teneur en composés terpéniques pour les provenances Françaises (Fr) et Marocaines (Mc) en condition contrôle (CL) et stress long (SL).

transcrits codants pour des Défensines ou des « pathogenesis related protein (PR)» sont d'ailleurs très souvent surexprimés chez les plants soumis à des stress hydriques. Dubos *et al.*. avaient, par exemple, détectés une PR-10 sur des plants de pin maritime élevés en hydroponie (Dubos and Plomion 2001). Il est également intéressant de noter que ce champignon à l'origine de fusariose par *Fusarium oxysporum lycopersici* (FOL) est un pathogène rencontré couramment dans les cultures de tomates hors sol qui provoque la pourriture de la racine et du collet. Après deux expérimentations exposées au pathogène, la seule alternative restante était d'utiliser un antifongique en traitement préventif. L'ajout de tout produit phytosanitaire peut entraîner une modification de l'expression de certains gènes ou inhiber certaines voies métaboliques impliquées dans la réponse au caractère étudié. On peut cependant penser que l'impact de l'antifongique sera similaire chez les plants contrôles et les plants stressés.

Concernant les conditions d'éclairage utilisées pour cette expérience, la croissance des plantes indique que du point de vue quantitatif, nous répondons aux besoins des plants. Par contre, du point de vue qualitatif, le port des plants élevés en hydroponie diffère de celui des plants en conditions naturelles par la présence d'entrenœuds courts. Cela confirme l'hypothèse d'un éclairage à valeur de phi élevée (RC/RL important). Il faudrait donc prévoir un apport de rayonnement RL par des sources incandescentes à raison de 10%-20% de l'énergie totale utilisée en prêtant attention à l'apport calorifique engendré.

Le flétrissement observé visuellement au moment de l'application du stress osmotique suivi d'une récupération rapide indique que le stress est suffisant pour induire une réponse de la plante. L'évolution des valeurs de potentiel hydrique de base nous confirme que les plants subissent effectivement un stress osmotique et répondent en adaptant de façon significative leur potentiel hydrique. En l'absence de stress, les plants ont une turgescence optimale. Le potentiel hydrique mesuré sur l'ensemble de l'appareil aérien et de la tige est proche de -0.4MPa pour une solution nutritive dont le potentiel osmotique est de -0.08MPa ((Dubos 2001)), le potentiel hydrique est donc nettement inférieur au potentiel osmotique. La solution hyper osmotique provoque bien une diminution du potentiel hydrique mais qui est de l'ordre de 0.33MPa au bout de 48H ce qui est inférieur à la variation du potentiel osmotique de la solution qui est de 0.60MPa. L'intensité du stress ou le temps d'application ne permettent peut être pas aux plants de mettre en place l'ensemble des mécanismes de réponse. Le potentiel hydrique des plantes reste cependant inférieur au potentiel osmotique de la solution, ce qui contribue à la disparition du flétrissement. Dans le cas de l'expérience « stress long », l'état d'équilibre doit être atteint et la faible mortalité (moins de 5%) semble indiquer que

l'intensité du stress est supportée par la plante. L'abaissement de potentiel hydrique n'est que de 0.23MPa pour l'écotype landais alors que la variation de potentiel osmotique de la solution atteint 0.45MPa. Si le flétrissement est absent, l'abaissement du potentiel hydrique ne suit pas celui de la solution. Par contre, l'écotype marocain abaisse son potentiel hydrique dans les mêmes proportions que l'abaissement du potentiel osmotique de la solution (0.47MPa). Ces mesures mettent en évidence que les deux provenances régulent leur potentiel hydrique de façon significativement différente. L'écotype marocain a clairement un ajustement osmotique plus efficace que la provenance landaise, ce qui confirme les résultats de Nguyen (1989b). Les conditions expérimentales appliquées devraient donc permettre de mettre en évidence des réponses au stress chez les deux provenances.

Les monoterpènes sont des composés très volatils, il est même possible de mesurer et suivre l'évolution des teneurs atmosphériques en alpha-pinène, beta-pinène, delta-3-carène ou limonène émis par les pins maritimes (Simon *et al.* 1994). Les teneurs atmosphériques mesurées par Simon *et al.* (1994) montrent des variations journalières corrélées à de nombreux facteurs tels que la température, l'intensité lumineuse ou la pression de vapeur d'eau. Une part de la variabilité intra-échantillon observée peut provenir d'une hétérogénéité lors du prélèvement, de la volatilité des composants et des faibles teneurs en terpènes des échantillons. Les analyses statistiques permettent cependant de détecter un effet provenance pour plusieurs composés. La variabilité génétique de teneur en monoterpènes entre différents écotypes de *Pinus pinaster* a déjà été constatée (Tognetti *et al.* 2000). A ce titre, le rapport de concentration entre l'alpha-pinène et le beta-pinène est un très bon marqueur de distinction des provenances landaises et marocaines (P. Pardon, communication personnelle). La régulation des teneurs par le stress hydrique est moins nette. Mis à part le caryophyllène dont les quantités diffèrent à la fois pour l'effet provenance et pour l'effet condition, l'analyse statistique avec prise en compte de l'effet d'interaction a permis de détecter des effets conditions. Il semble que la production de plusieurs composés terpéniques soit modifiée lors de la réponse au stress hydrique, notamment le delta-3-carène, le myrcène, l'humulène et le caryophyllène dont la production augmente au cours du stress. Les monoterpènes (C_{10}) dérivent du GPP tandis que les sesquiterpènes (C_{15}) sont synthétisés à partir du FPP (Fig. D-6). Nous avons déjà indiqué dans la synthèse bibliographique que la synthèse de l'ABA et des GA dérivait d'isoprènes comme le GGPP ou les carotènes. On peut donc s'attendre à observer une régulation de certaines enzymes de la voie de biosynthèse des terpènes. Comme nous le verrons lors de la présentation des résultats des transcrits, nous avons détecté des gènes

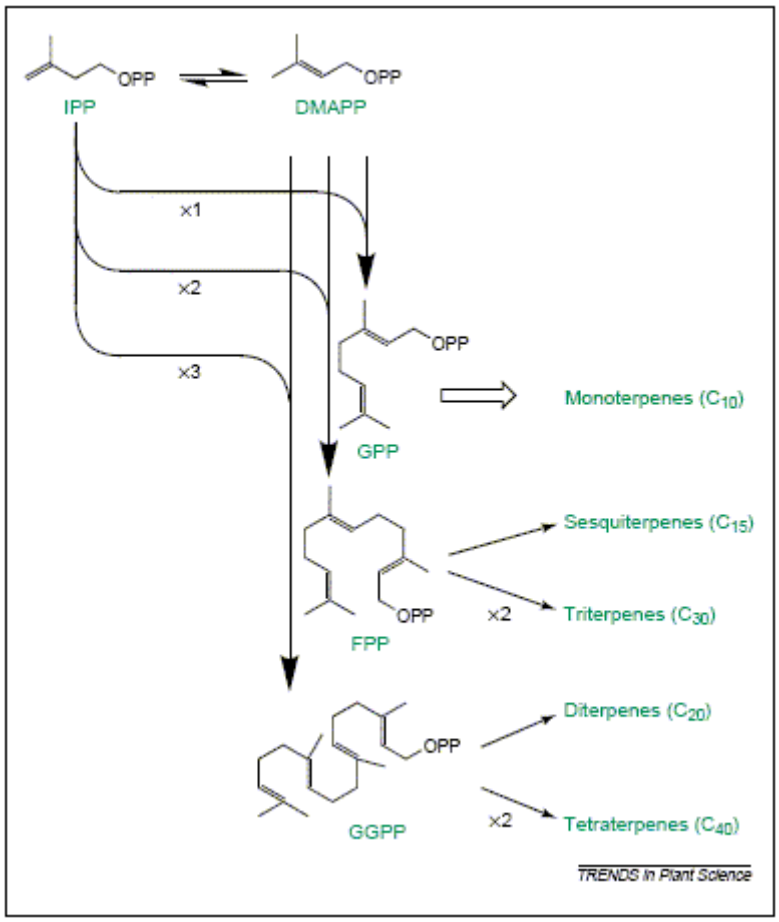


Figure D-6 : Voie de biosynthèse des isoprènes. La condensation avec le DMAPP d'un IPP permet la synthèse de GPP, de 2 IPP donne le FPP et de 3 IPP donne le GGPP. Le GGPP est un précurseur de la synthèse de l'Acide Abscissique et des Gibbérellines (Mahmoud, 2002). Abréviations : IPP : Isopentenyl diphosphate, DMAP : Dimethylallyl monophosphate, GPP : Geranyl Pyrophosphate, FPP : Farnesyl Pyrophosphate, GGPP : Geranyl Geranyl Pyrophosphate.

présentant des homologies avec la pinène synthase et la GPP synthase. De plus les composés terpéniques peuvent jouer un rôle dans la résistance aux pathogènes. Le limonène présent à des concentrations légèrement supérieures chez la variété marocaine aurait un pouvoir répulsif sur des insectes pathogènes comme la chenille processionnaire (Tiberi *et al.* 1999). Les terpènes sont également à l'origine des caroténoïdes, leur implication dans la réponse au stress osmotique n'est donc pas étonnante. Ce type de stress se traduit également par un stress oxydatif face auquel la plante peut répondre par la production de composés antioxydants tels que les caroténoïdes. Pour confirmer cette hypothèse, il serait nécessaire de réaliser un nouveau plan d'expérience dans ce sens.

D.2 Analyse et annotation des séquences

Ce chapitre présentera les caractéristiques des séquences générées et les résultats des traitements bioinformatiques qui ont été appliqués. Nous commencerons par l'assemblage des séquences puis leur annotation et leur répartition en catégories fonctionnelles. Nous continuerons par la présentation des données d'expression in silico. Nous terminerons par la réalisation de l'Unigène et la discussion des résultats obtenus.

D.2.1 Séquençage

Au début de la thèse, nous disposions de 6984 séquences issues d'une banque d'ADNc de xylème ainsi que 721 séquences de banques de parties aériennes de plantules élevées en hydroponie dont 242 provenaient de plants contrôles et 479 de plants stressés. Le projet de séquençage d'EST des banques de racines que nous avons mené au cours de la première année de cette thèse a aboutit à l'étalement de plus de 14000 clones. Après élimination des séquences de moins de 100 paires de bases, plus de 9500 séquences ont été générées. La première phase de traitement qui a suivi consistait à nettoyer individuellement chaque séquence en fonction de l'indice de qualité global et par base pour obtenir un score Phred20 supérieur à 85%. Cette démarche permet de s'assurer de la qualité des séquences avant injection dans le pipeline bioinformatique (cf Matériel et méthodes) mais prend un temps considérable. En analysant a posteriori nos critères de décision et la logique appliquée sur ce lot de données, cette étape manuelle a pu être entièrement automatisée. Un script Perl se charge de détecter et d'éliminer alternativement les zones de basse qualité aux extrémités de la séquence en prenant en compte le score moyen et le nombre de bases de basse qualité consécutives comme nous le faisions manuellement. La présence de séquences correspondant au vecteur de clonage est finalement recherchée en vu de leur élimination.

Deux projets distincts, l'un d'EST issus d'une banque d'ADNc construite à partir de xylème, l'autre construite à partir des racines et parties aériennes, et un projet global (ensemble des EST de pin maritime) ont été construits et injectés dans le pipeline. Dans le cadre de cette thèse nous nous sommes intéressés plus particulièrement aux séquences issues de racines (Fig. D-7).

Ce projet comprend 10238 séquences obtenues sur 9711 clones différents (plusieurs passes pour certains clones) dont 8996 ont pour origine les banques de racines et 715 sont issus des parties aériennes. Au total, plus de 4.4 Mb ont été séquencées. Les EST des racines stressées représentent 4324 séquences contre 5193 séquences pour les EST des racines contrôles.

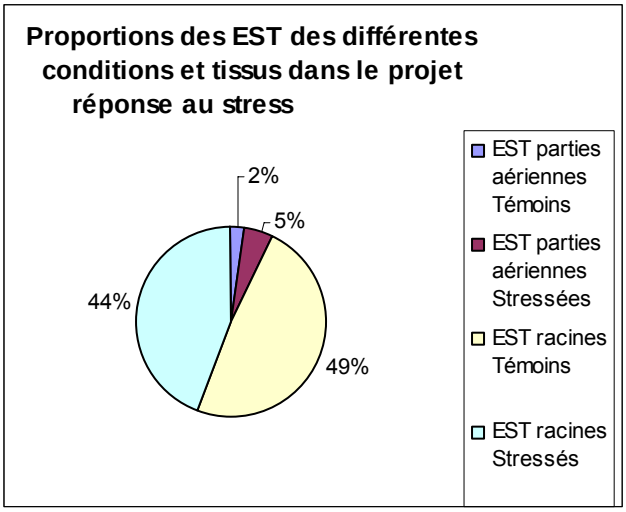


Figure D-7 : Représentativité des différents tissus et conditions dans le projet réponse au stress.

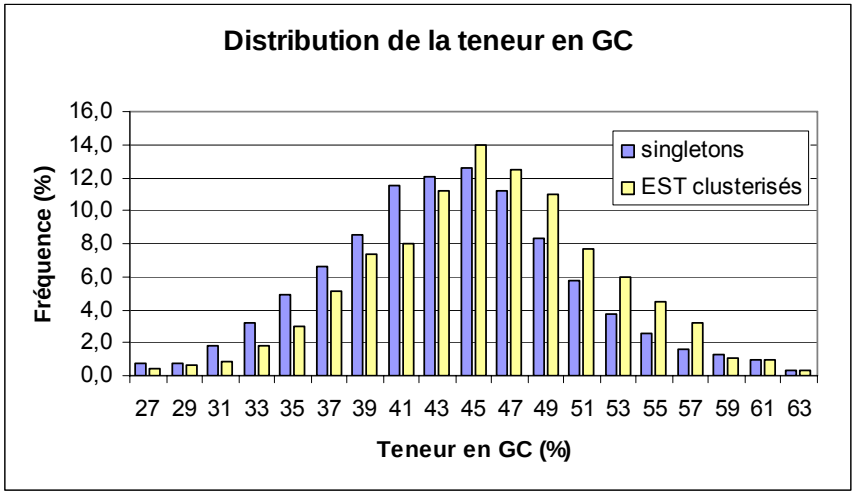


Figure D-8 : Distribution des concentrations en bases G et C des ESTs pour les singletons et les séquences clustérisées.

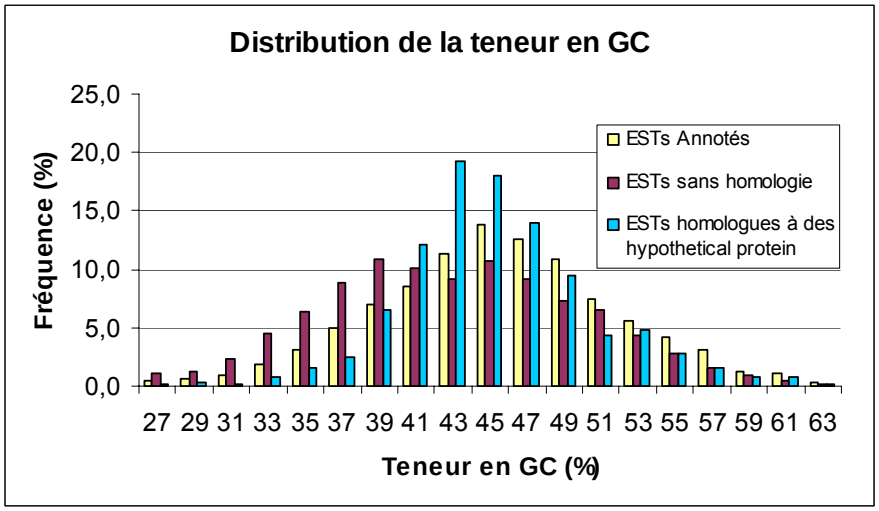


Figure D-9 : Distribution des concentrations en bases G et C des ESTs pour les séquences annotées, les homologues à des « hypothetical proteins » et les séquences sans homologies.

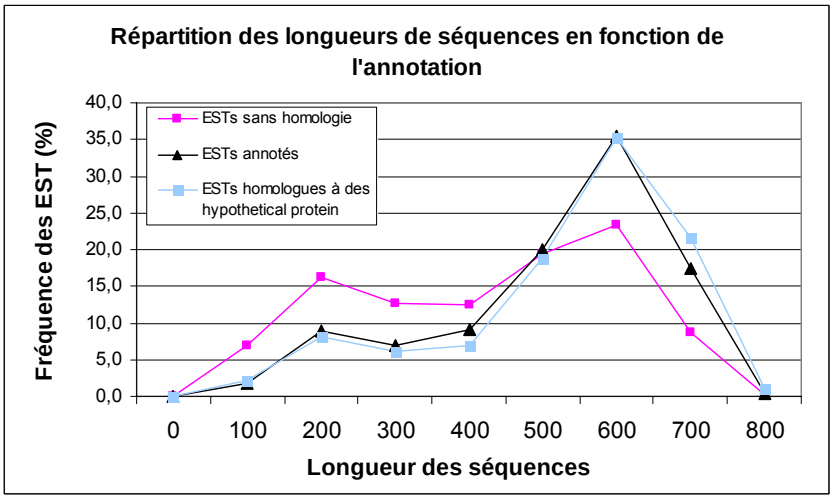


Figure D-10 : Répartition des longueurs des séquences en fonctions de l'annotation.

Si l'on considère les résultats de l'assemblage et de l'annotation que nous aborderons dans les chapitres suivants, nous pouvons calculer un indice de redondance définit comme :

$$\text{Re}\,dondance = \frac{\text{nombre d'EST clustérisés - nombre de consensus}}{\text{nombre total d'EST}}$$

Ce projet étant constitué de 7355 EST clustérisés dans 2261 consensus, la redondance est donc de 49,7 % ce qui correspond à 5144 clones différents (2883 singletons et 2261 consensus) pour un total de 2.5 Mb non redondantes.

La teneur en GC est globalement de 44.9% mais n'est pas distribuée de façon homogène selon le type de séquence étudié. La figure D-8 présente les distributions des teneurs en GC des singletons et des EST clustérisés. Il apparaît que la distribution des singletons est centrée sur une valeur de GC plus faible (43.7 %) que celle des EST clustérisés (45.3 %). L'hétérogénéité de la distribution est aussi constatée en distinguant les séquences selon leur annotation (Fig. D-9). La distribution de la concentration en GC pour les séquences annotées est similaire à celle des EST clustérisés de la figure D-8. La distribution des EST présentant une homologie avec des « hypothetical proteins » est asymétrique. Les faibles teneurs en GC sont sous représentées mais la classe la plus abondante se situe autour de 43-45 %. La distribution des EST sans homologie avec des séquences connues est bimodale, la partie droite suit la distribution des séquences annotées et la partie gauche est centrée sur la classe 39-41 %.

La longueur moyenne des séquences obtenues après nettoyage et élimination du vecteur est de 445 paires de bases. Comme pour la teneur en GC, les distributions varient en fonction du type d'annotation rencontrée. La figure D-10, montre que la répartition des longueurs est similaire pour les EST annotés et les EST homologues à des « hypothetical proteins ». Le profil est bimodal, mais plus de 75 % des séquences ont une taille supérieure à 400 pb, la classe 600-700 pb étant la plus abondante. La répartition des longueurs pour les EST sans homologie montre un profil bimodal plus marqué, la partie droite correspondant aux courbes que nous venons de décrire, la partie gauche étant centrée sur la classe 200 à 300 pb. Une analyse statistique (ANOVA et test-T) indique que la courbe de longueur des EST sans homologie diffère significativement des deux autres.

Le profil bimodal des courbes de longueur des EST reflète les difficultés rencontrées pour optimiser les performances de la réaction de séquence avec la chimie utilisée. Lorsque les conditions sortent légèrement des conditions optimales, cela se traduit par la génération de séquences courtes avec une chute de la qualité. Nous obtenons soit des séquences longues soit des séquences courtes en fonctions de la réaction de séquence ou des paramètres d'injection.

Nous venons de voir que les courbes de distribution de longueur et de teneur en GC différaient en fonction du type d'annotation assigné aux séquences. Les séquences courtes sont plus abondantes dans les EST sans homologie et leur teneur en GC est plus basse que celle des autres EST. Nous avons observé une distribution bimodale pour la teneur en GC des EST sans homologies avec un pic à 39% et un pic à 45 %. Des données similaires ont été obtenues sur les EST de *Solanum tuberosum* avec une teneur de GC inférieure de 4% pour les EST dont les résultats Blast $<10^{-2}$ (Crookshanks *et al.* 2001). Si l'on s'intéresse à la teneur en GC calculée à partir des génomes d'*A. thaliana* et *O. sativa* (Yu *et al.* 2002), les valeurs de GC sont plus basses pour les introns que pour les exons, respectivement (0.324 – 0.370) et (0.432 – 0.514), et la teneur en GC décroit de l'extrémité 5' vers l'extrémité 3' dans les gènes du riz. Ainsi, de faibles valeurs de GC pour les EST sans homologies pourraient être reliées avec une grande proportion de séquences de la région 3'. Cette hypothèse concorderait avec la stratégie de séquençage adoptée. La région 3' étant connue pour déterminer la spécificité des membres d'une famille multigénique plutôt que la fonction de la protéine. Les recherches d'homologie sont plus difficiles en 3'. Une autre interprétation serait une plus forte occurrence de séquences de la région UTR (UnTranslated Region) que l'on suppose plus proches des régions introniques pour leur teneur en GC. Les séquences courtes rendent également plus difficile la recherche d'homologie et l'assemblage, ce qui peut également expliquer leur plus forte proportion dans la catégorie des EST sans homologie.

D.2.2 Assemblage des séquences

L'assemblage bioinformatique des séquences permet d'identifier et d'aligner les séquences des EST issus d'un même gène (cf. Matériel et méthode). Les trois assemblages bioinformatiques, accessibles à l'URL : http://cbi.labri.fr/outils/SAM2/COMPLETE/, sont les suivants :

- L'injection des 10238 séquences du projet « réponse au stress » dans le pipeline a conduit à leur regroupement en 2118 contigs contenus dans 1719 clusters et 2883 singletons. Des consensus primaires ont été générés pour chaque contig. Des consensus supplémentaires, au nombre de 143, ont également été générés pour 58 contigs (Fig. C-18). Ce projet est nommé « Pinus Pinaster root EST ».
- Le projet EST de xylème, accessible sous le nom « Pinus pinaster xylem EST », est organisé en 995 clusters contenant 1226 contigs et 2948 singletons.
- Le projet « Pinus pinaster total » comprenais l'ensemble des EST soit 18498 séquences réparties en 2697 clusters, 2893 contigs et 5051 singletons. Pour ce projet,

4003 consensus dont 2893 consensus primaires ont été générés. C'est à partir de ces 2893 contigs et 5051 singletons que nous avons choisi les séquences de l'Unigène dont nous parlerons plus loin. Une nouvelle version du projet a été mise récemment en ligne et contient maintenant 17770 séquences.

Pour le projet « racines », 72 % des séquences sont clustérisés contre 73 % pour le projet total. La fusion des deux projets xylème et racines n'a donc pas permis de réduire significativement la proportion de singletons. Cela peut signifier qu'il existe un grand nombre de gènes faiblement exprimés pour lesquelles nous ne disposons que de séquences sans recouvrement. Les deux banques concernent cependant des tissus très différents et l'ajout de séquences ne provoque donc pas une simple augmentation du nombre d'EST par gène ce qui permettrait de réduire le nombre de singletons, mais l'apport d'EST correspond à des gènes ou membres de familles « spécifiquement » représentés dans chacun des tissus. L'ajout des séquences de xylème permet de passer de 2118 à 2893 contigs. De façon simplifiée, cela signifie que 775 nouveaux « gènes » sont détectés par ajout des séquences de xylème. Dans le même temps, si la proportion de consensus alternatifs restait très faible au sein du projet « racines », dans le projet total, nous avons 1110 consensus alternatifs pour 2893 consensus primaires. Les consensus alternatifs sont générés dans le cas où l'alignement des EST constitutifs d'un contig présente des régions polymorphes. Les EST sont alors classés en plusieurs groupes de séquences similaires et un consensus est généré pour chaque sous-groupe. Une région est considérée polymorphe si elle atteint plusieurs dizaines de bases (selon paramètres de l'algorithme). Cela ne concerne donc pas les erreurs ponctuelles de séquençage ou des polymorphismes de quelques bases isolées (SNP). La forte proportion de consensus alternatifs indique donc qu'un même gène présente un polymorphisme sur une région de quelques dizaines de paires de bases entre les transcrits des racines des plants landais utilisés en hydroponie et les transcrits des arbres corses utilisés pour l'étude du « xylème ». Ce polymorphisme peut correspondre à différents membres de familles multigéniques ou des formes allèliques très différentes.

Le tableau D-3 présente les annotations des contigs constitués de plus de 20 EST dans le projet « racines ». Le contig le plus représenté correspond à la sous unité 25S ribosomique qui regroupe 157 EST soit un peu plus de 1.5 % des transcrits séquencés. Les ARN ribosomaux font partie des transcrits les plus abondants de la cellule. Certains membres de cette famille sont exprimés de façon constitutive et stable dans la cellule tel que l'ARN 18S

Tableau D-3 : Nombre d'ESTs de racines et Annotation des Contigs comportant plus de 20 ESTs dans le projet « racines ». Seules les 8996 ESTs issus de racines sont pris en considération.

Nombre d'ESTs	Annotation
157 (1.75 %)	25S ribosomal subunit (AJ271027)
58	Phosphopyruvate hydratase (EC 4.2.1.11)
42	Naringenin-chalcone synthase (EC 2.3.1.74)
34	Elongation factor 1-alpha
30	Avr9/Cf-9 rapidly elicited protein 146 (Q9FQZ6)
30	Methionine adenosyltransferase (EC 2.5.1.6)
30	Ubiquitin
28	Translationally controled tumor protein
28	Late embryogenesis abundant protein
26	60S ribosomal protein L41 (U26255)
23	Tonoplast intrinsic protein (P24422)
23	Proline-rich protein (Q9LLZ6)
21	ADP-ribosylation factor (P51823)
20	Metallothionein-like protein EMB30 (Q40854)
20	Polcalcin (O64943)
20	Defensin (O65740)
20	Catalase (EC 1.11.1.6)
20 (0.2%)	putative protein

dans les racines de pin (Sathyan 2004) tandis que d'autres ont leur expression régulée au cours du stress comme des 40S ou 60S (Kawasaki *et al.* 2001; Ozturk *et al.* 2002).
On rencontre également des transcrits de la Phosphopyruvate hydratase (EC 4.2.1.11) ou enolase qui intervient dans la glycolyse en catalysant la réaction de déphosphorylation du 2-phosphoglycérate au Phosphoenol-pyruvate. Il est donc confortant de retrouver cet enzyme qui intervient dans la synthèse d'un composé clé de la production d'énergie pour la cellule et qui est indispensable à la production de l'Acetyl-CoA qui intervient dans de nombreuses voies métaboliques.
Les EST homologues à la Naringenin-chalcone synthase (EC 2.3.1.74) sont rencontrés 42 fois. Cet enzyme est nécessaire à la biosynthèse des flavonoïdes. Ces composés jouent le rôle d'antioxydant pour la cellule et évite les dégradations des constituants cellulaires. La catalase (EC 1.11.1.6) présente dans la liste avec 20 EST va aussi intervenir dans les phénomènes de prévention de l'oxydation. Cet enzyme catalyse la réaction de conversion du peroxyde d'hydrogène en molécules d'oxygène et d'eau.
Nous trouvons en quatrième position un contig contenant 34 EST homologues au facteur d'élongation 1-alpha. Ce type de transcrits faisait également parti des 6 transcrits les plus abondant dans le transcriptome des tubercules de pomme de terre avec une abondance atteignant 2.6 % des transcrits (Crookshanks *et al.* 2001).
Parmi ces contigs, il y a des fonctions pouvant intervenir dans la réponse au stress hydrique comme une AVr9/Cf-9, une LEA, une protéine riche en proline et une Polcalcine (Tab. D-3).

Les contigs contenant le plus d'EST concernent donc le métabolisme cellulaire (ex : constituants des ribosomes, enzymes de la glycolyse), la protection contre l'oxydation des composant cellulaires (flavonoïdes, catalase) et des enzymes de réponse au stress (LEA, Protéine Riche en Proline, Polcalcine, Défensine).

Un nouveau projet de séquençage du pin maritime vient juste d'être terminé. Il s'agit de 8131 séquences de bourgeons réalisées en partenariat avec le Centre National de Séquençage (CNS). Nous disposons donc depuis l'automne 2005 de 26629 séquences. L'assemblage de ces nouveaux EST avec ceux de racines et du xylème s'organise en 7358 singletons et 3874 Clusters regroupant les 4182 Contigs. Le nombre de consensus passe à 4640 dont 458 sont des consensus alternatifs. Par rapport aux résultats de l'assemblage « Pinaster total », le nombre de consensus alternatifs a fortement diminué et rejoint les proportions du projet racine initial. Le nombre de contig augmente par contre de 2893 à 4182.

L'ajout de séquences de bourgeons a vraisemblablement permis d'éclater les contigs contenant des formes alternatives d'un gène –représentés par les consensus alternatifs– en contigs différents. La proportion de singletons reste stable avec 27 % des séquences.

D.2.3 Annotation des séquences

Le rôle de cette étape est d'attribuer une fonction aux singletons et aux gènes dont la séquence consensus partielle a été obtenue par l'assemblage informatique. L'annotation a été réalisée pour les séquences des projets « xylème » et « racines ». Une attention particulière a été portée au projet « racines » en attribuant aux annotations les noms officiels des enzymes.

D.2.3.1 Annotation des consensus et des singletons

La première étape d'annotation a été réalisée en se servant de l'interface d'annotation du pipeline. Notre travail d'annotation sur les séquences de pin maritime a permis de faire évoluer les fonctionnalités et l'ergonomie de cette interface web entre la version de développement mise en place au début de la thèse et la version actuellement disponible à l'URL (http://cbi.labri.fr/outils/SAM2/). La figure D-12, montre le résultat de cette collaboration avec le Centre de Bioinformatique de Bordeaux. L'annotation se fait, pour les EST clustérisés, au niveau de la séquence consensus puis l'annotation est héritée sur les EST composant le contig. Les résultats de recherche d'homologie sont indiqués en bleu dans la partie inférieure droite de l'interface et leur intitulé alimente le champ « proposition » en haut à gauche. L'annotateur peut sélectionner l'annotation choisie et modifier l'intitulé dans le champ « Annotation ». La seule prise en considération des résultats du Blast et des valeurs de probabilité associées est insuffisante pour assigner une annotation correcte. Il faut également tenir compte de l'étendue de l'homologie trouvée. Une forte homologie avec une séquence connue mais sur quelques dizaines de paires de bases uniquement ne signifie pas que la fonction du gène est celle de l'homologue mais uniquement que les deux gènes ont un domaine commun qui peut éventuellement correspondre à un site actif mais aussi à un domaine sans rapport avec la fonction réelle du gène.

Cela soulève un des risques majeurs de l'annotation par inférence électronique. Une séquence peut présenter une forte homologie avec la séquence d'un gène codant une protéine de fonction connue et vérifiée biochimiquement. Si l'homologie concerne la région codant le site actif de la protéine, l'héritage de l'annotation du gène homologue sur le gène recherché ne pose pas de problème. Dans le cas contraire, si une forte homologie est rencontrée, rien de garantit que la fonction biochimique du gène recherché soit la même que celle du gène

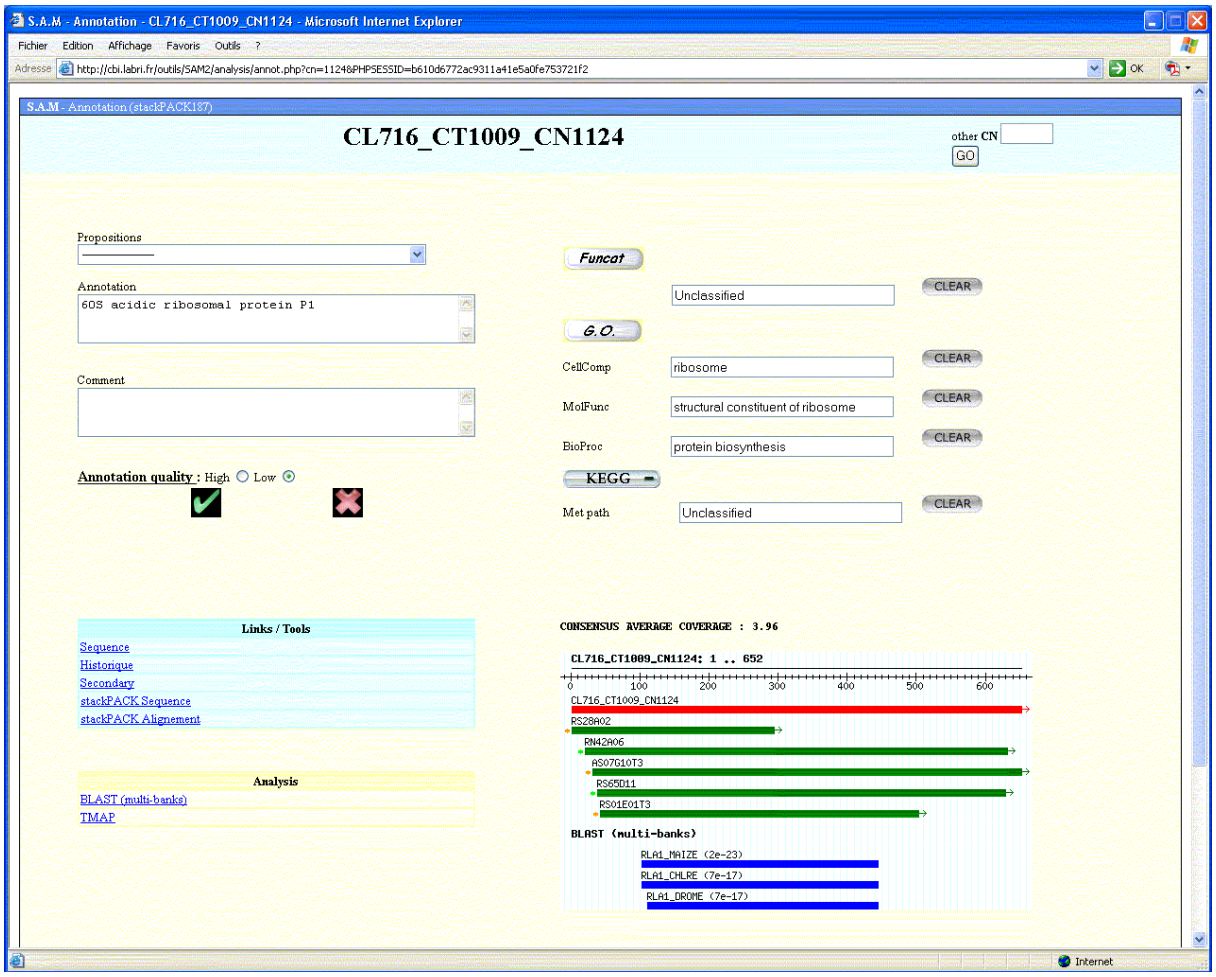


Figure D-12 : Dernière version de l'interface d'annotation du pipeline d'annotation des ESTs.

		Confiance	Quantité info
Blast Prot	**SwissProt**	++	-
	Trembl	+	++
Blast Nucl	**TIGR (2005)**	+	+/-
	EMBL	-	++

Figure D-13 : Bases de données publiques utilisées pour la recherche d'homologie. La qualité des annotations, la quantité d'information disponible et le type données (protéique/nucléique) sont indiqués.

homologue et il peut en résulter une annotation erronée. Une autre séquence pourra présenter une forte homologie avec ce gène mal annoté. Dans le cas où l'homologie trouvée ne concerne pas la région qui était homologue à la séquence du gène de départ de fonction vérifié biochimiquement, l'annotation peut également être erronée. De proche en proche, la fonction attribuée au gène recherché peut n'avoir rien de commun avec la fonction réelle de la protéine codée. Une annotation est toujours délicate à réaliser et le maximum de précautions doit être pris pour s'assurer de son exactitude avant de la publier sur les bases de données publiques, ce qui est loin d'être toujours le cas.

La figure D-13 synthétise les bases qui ont été utilisées pour les recherches d'homologie et leurs caractéristiques en termes de confiance dans les annotations indiquées, quantité d'information disponible et type de données proposées. La stratégie de recherche d'homologie en cascade sur les bases SwissProt, TrEMBL puis EMBL (cf. Matériel et méthodes) donne une priorité à la base protéique SwissProt dont la qualité d'annotation est la plus élevée, les fonctions des protéines ayant le plus souvent été vérifiées biochimiquement. Le nombre de séquences disponibles et le nombre d'espèces représentées est cependant plus faible que dans les autres banques (http://www.expasy.org/sprot/). La banque TrEMBL est une traduction de la banque EMBL. La quantité d'information est donc très élevée. Les annotations sont par contre moins fiables que celles de SwissProt. La recherche d'homologie sur des séquences traduite en acides aminés permet de s'affranchir des variations nucléotidiques silencieuses inter espèce et de considérer la fonction du gène. La base EMBL permet la recherche d'homologie contre un maximum de séquences disponibles sur la base de séquences nucléotidiques.

La stratégie en cascade, si elle permet de réduire considérablement le temps nécessaire à la recherche d'homologie, a toutefois un inconvénient en privilégiant essentiellement les résultats obtenus sur les premières bases testées. La valeur de la p-value seuil doit être bien choisie. Dans notre cas, fixée à 10^{-5}, elle s'est avérée trop lâche. Cela se traduisant parfois par un résultat d'homologie faible (ex : 10^{-6}) sur SwissProt car la protéine codée par le gène n'est pas présente dans la base mais possède des similarités faibles avec une autre protéine. Une recherche sur TrEMBL peut par contre renvoyer une très forte homologie avec des séquences connues. Ce type de situation se rencontre notamment pour des gènes déjà séquencés chez *Pinus taeda* qui diffèrent fortement des autres espèces et qui par conséquent ne sont pas présents dans la base SwissProt. Un autre cas de figure concerne une différence de résultats entre le consensus et les EST qui le composent. Cela s'explique par la quantité d'information disponible. Une séquence consensus est plus informative qu'un EST. Plus la longueur de la séquence recherchée est importante, plus précise sera la recherche. Une homologie avec la

séquence consensus peut être trouvée alors qu'aucune homologie correcte n'est observée pour les EST pris individuellement. Comme dans le cas expliqué précédemment, un consensus peut présenter une homologie faible avec des accessions SwissProt alors qu'il existe une forte homologie avec une accession de la base TrEMBL. Dans ce cas, nous avons fréquemment observé que les EST composant le contig présentaient l'homologie avec l'accession de TrEMBL.

C'est pour cela qu'une vérification systématique de la concordance des homologies trouvées pour les EST d'un contig et celles du consensus a été faite. Un système de point de couleur verte ou orange devant la barre verte foncée symbolisant les EST du contig indique si le résultat du Blast de l'EST concorde avec celui du consensus (Fig. D-12). Il est possible de relancer spécifiquement un Blast pour le consensus ou un EST sur une base en particulier.

L'assemblage bioinformatique peut être à l'origine de chimères. Dans le cas du projet « racines », nous n'avons pas été confrontés à ce problème. Ce phénomène intervient lorsqu'un insert est constitué par la juxtaposition de deux EST de gènes différents. Cette séquence chimérique effectue un lien entre l'assemblage des deux gènes. Sur la représentation graphique de l'interface, il apparaîtrait deux blocs d'EST alignés reliés par une seule séquence. Dans ce cas, les résultats de Blast montreraient des homologies avec des gènes différents pour les deux blocs.

Le TC57901 du TIGR est un exemple de ce type de configuration. Il est consultable à l'URL http://www.tigr.org/tigr-scripts/tgi/tc_report.pl?tc=TC57901&species=pinus. Les 80 accessions sont organisées en deux blocs. L'alignement des séquences du premier bloc qui comprend la plupart des accessions couvre les 675 premières paires de bases du consensus. Le second alignement couvre la région des bases 700 à 2263. Une seule séquence permet de relier ces deux alignements. Les séquences du premier bloc ont été inversées pour l'alignement tandis que l'orientation de celles du second bloc a été conservée. La recherche d'homologie donne une Déhydrine pour le second bloc, la base 807 correspondant à la première base du match. Le premier bloc présente une homologie avec une « unknown protein » d'*Arabidopsis* en sens complémentaire, la base 576 correspondant à la première base de la protéine homologue. L'annotation des EST de ce TC a été donnée à partir du résultat d'homologie du second bloc qui ne regroupe que 33 des 80 accessions.

Pour procéder à l'annotation, le résultat du Blast doit donc être analysé en tenant compte de la base de données d'où il provient, l'étendue de l'homologie doit être notée ainsi que la concordance des résultats de Blast des EST et leur position dans l'alignement. Si les informations disponibles ne sont pas suffisantes, une nouvelle recherche d'homologie peut

être lancée. Le champ d'annotation peut alors être alimenté. Avec l'intégration de nos séquences de pin maritime dans le projet d'assemblage « *Pinus* » du TIGR, une annotation de nos séquences intégrerait aujourd'hui les bases de données *Pinus* du TIGR.

Au cours de la phase d'annotation, il arrive fréquemment que des résultats de blast ayant de bons scores correspondent à des fonctions d'intitulés différents alors qu'il s'agit en réalité de la même enzyme. Cela vient du fait que les dénominations des enzymes ont évolué au cours des années. A titre d'exemple, la SAM synthase (EC 2.5.1.6) ou S-adenosylmethionine synthase est aussi appelée AdoMet synthetase et son nom officiel est Methionine adenosyltransferase. Nous avons donc exporté l'ensemble des annotations réalisées dans un tableur Excel et recherché le nom officiel des enzymes sur le site web http://www.expasy.org afin d'uniformiser les annotations. Cette étape très laborieuse s'est avérée extrêmement utile pour les études d'expression *in silico*. Nous avons également pu vérifier la cohérence des annotations des contigs d'un même cluster.

Nous avons vu dans les chapitres précédents que certaines séquences ne présentaient pas d'homologie avec des gènes de fonctions connues. Dans les bases de données nous pouvons rencontrer les annotations « putative protein », « unknown protein » ou encore « hypothetical protein ». L'annotation « putative protein » est généralement rencontrée pour des séquences de transcrits (ex : EST) mais pour lesquelles la traduction en protéine n'est pas montrée. Les séquences annotées « unknown protein » correspondent normalement à des transcrits traduits en protéines mais dont la fonction est inconnue. Le terme « hypothetical protein » est plus ambiguë et peut parfois prendre la définition donnée pour « putative protein » ou parfois celle de « unknown protein ». Certaines accessions sont ainsi notées « Hypothetical 10.4kDa protein (Expressed protein) » ce qui indique que la protéine était bien exprimée lors de sa découverte mais de fonction inconnue et d'autres accessions « Hypothetical 85.6 kDa protein » peuvent provenir d'une modélisation de la protéine potentiellement produite.

Nous avons donc choisi dans notre cas de désigner par « putative protein » ou « no hit » les séquences ne présentant pas d'homologie (> 1e-05) à l'issue de la stratégie en cascade. Nous sommes en présence de séquences d'EST, il s'agit donc de séquences de transcrits. Il est probable mais pas certain que la séquence soit traduite en protéine. Dans le cas où une homologie avec d'autres transcrits annotés « hypothetical protein » est trouvée, nous avons considéré deux situations. Si une précision était indiquée sur la nature de la protéine hypothétique (ex : 28.2kDa) nous avons conservé l'annotation de l'homologue. Dans le cas

contraire, nous désignerons les séquences comme « homologue to hypothetical protein » ou « unknown protein ». Cette notation est toutefois discutable et dans le cas où une publication des annotations sur les bases publiques était envisagée, il serait préférable de donner une annotation du type : « similar to » suivi de l'annotation de la séquence homologue pour lever toute ambiguïté. Il est également possible de suivre le TIGR qui a choisit d'ajouter un qualificatif avant l'annotation pour préciser la qualité de l'homologie obtenue (Tab. D-4).

La répartition des annotations des EST (Fig. D-14) montre que la proportion d'EST annotés, sans homologie et homologues à des protéines hypothétiques est similaire pour les deux conditions au sein d'un même tissu. Bien que la proportion des EST issus des parties aérienne soit très faible (Fig. D-7), on peut toutefois remarquer la plus forte proportion de séquences ne présentant pas d'homologie dans les EST de racines. Une hypothèse pour expliquer cette différence serait que la quantité d'EST générés étant beaucoup plus faible pour les parties aériennes, la probabilité de séquencer des transcrits abondant donc mieux connus est plus forte. Il se peut également que les transcrits de racines soient moins bien décrits que ceux des parties aériennes.

D.2.3.2 Assignation des catégories fonctionnelles

L'objectif de la classification est d'organiser les transcrits des gènes identifiés dans un nombre limité de catégories fonctionnelles et de suivre un vocabulaire contrôlé pour les termes utilisés dans la description des fonctions des gènes. Les trois ressources principales, faisant référence et intégrées à l'interface d'annotation (Fig. D-12), sont les bases FunCat (The Functional Catalogue), GO (Gene Ontology) et KEGG (Kyoto Encyclopedia of Genes and Genomes). Les bases FunCat et GO ne sont pas des bases de séquences, il s'agit de bases permettant de rechercher une terminologie associée aux termes recherchés suivant celle du MIPS (Munich Information Center for Protein Sequences). La base KEGG en revanche repose sur le principe du Blast. Une recherche est lancée sur une base protéique contenant uniquement des protéines très bien identifiées, annotées dont les fonctions et les positions dans les voies métaboliques sont connues. L'utilisation du KEGG peut être automatisé mais le nombre de protéines dans la base est restreint et ne permet d'assigner une classe fonctionnelle qu'aux enzymes les plus connues. Notre choix s'est donc porté sur la classification GO qui s'articule autour de 3 niveaux d'organisation : la fonction moléculaire, le processus biologique et la localisation cellulaire. A chaque fois que cela était possible, les champs ont été renseignés mais une même enzyme peut avoir des fonctions multiples dans la cellule et

Tableau D-4 : Qualificatifs utilisés par le TIGR pour l'annotation.

critère	**qualificatif**
100% > % d'identité >= 90%	*homologue to*
90% > % d'identité >= 70%	*similar to*
% d'identité < 70%	*weakly similar to*
couverture protéine > 98%	*complete*
couverture protéine <= 98%	*partial*

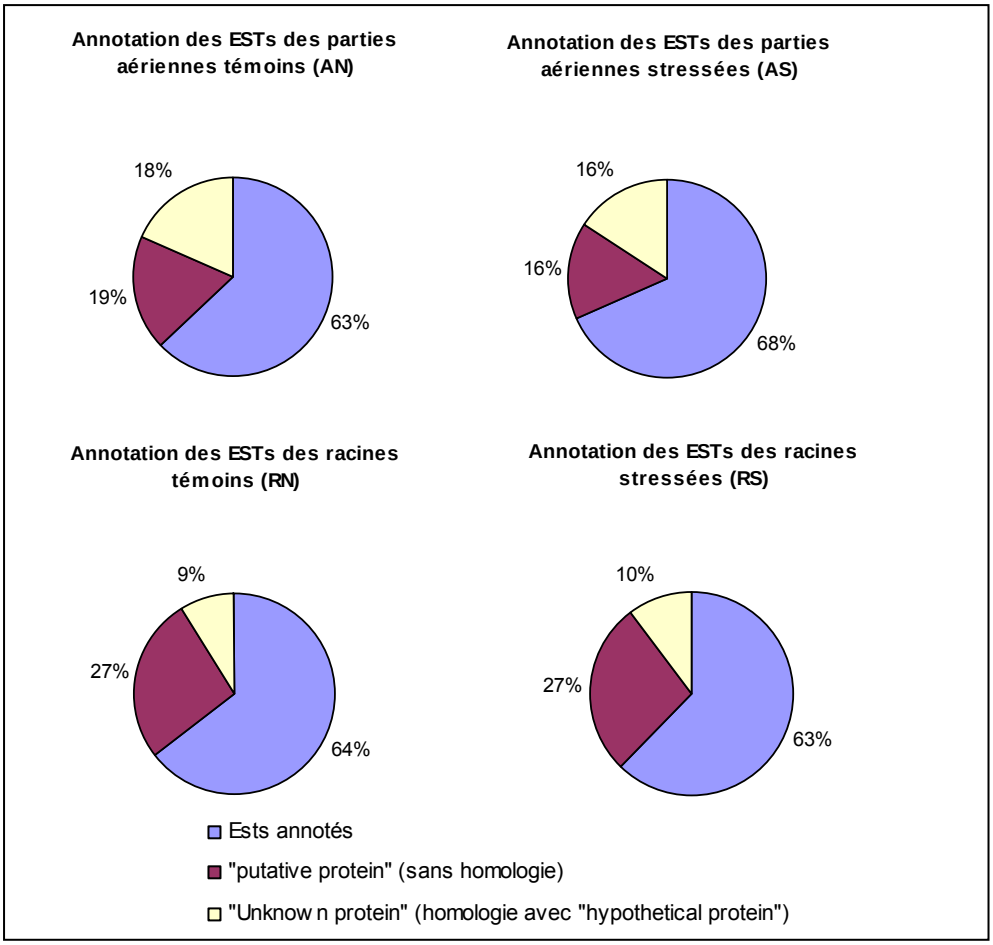


Figure D-14 : Proportion d'ESTs annotés, sans homologie et homologues à des protéines hypothétiques pour les deux tissus et les deux conditions dans le projet « réponse au stress ».

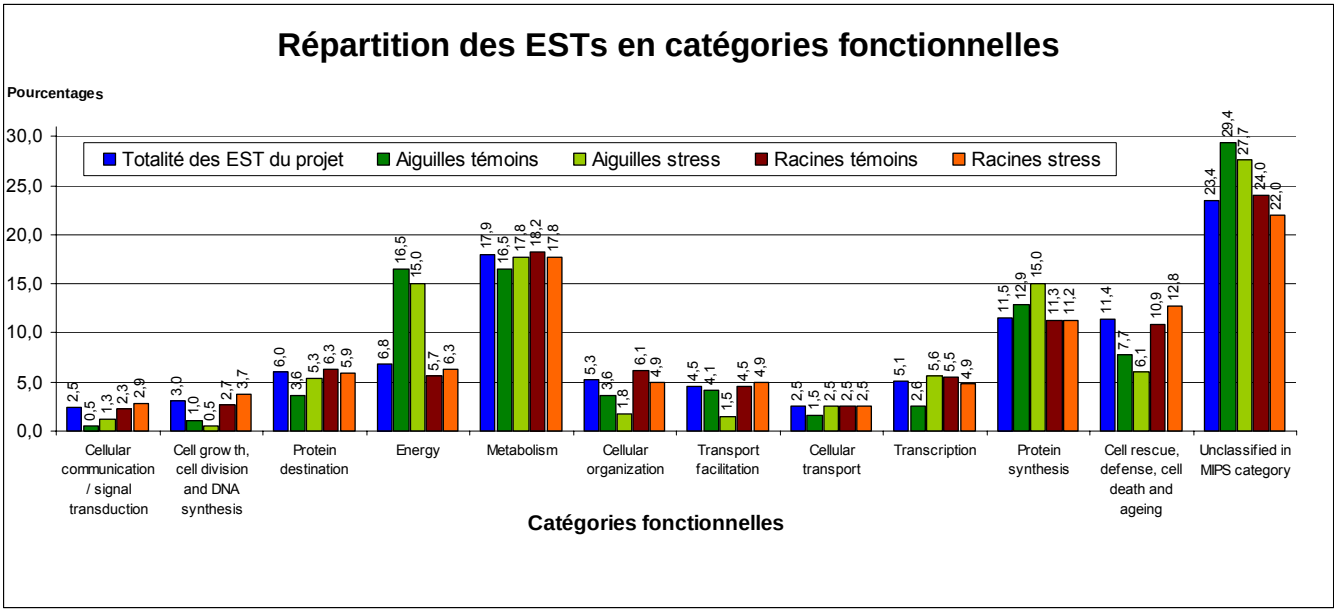


Figure D-15 : Répartition des ESTs en catégories fonctionnelles.

l'assignation à une classe particulière ne peut pas toujours être tranchée. Il reste donc une part de subjectivité de l'annotateur dans cet exercice.

La répartition des annotations de GO en 12 grandes catégories de l'ensemble des EST pour lesquels une annotation a pu être donnée, –c'est-à-dire mise à part les EST sans homologie et homologues à des « hypothetical protein »- est indiquée à la figure D-15. Comme pour la proportion des EST annotés « putative protein » ou « unknown » (Fig. D-14), la répartition des EST dans les différentes catégories fonctionnelles est très proche pour les deux conditions d'un même tissu. Par ailleurs, la différence de répartition dans les catégories est nette entre les deux tissus. Les EST de la classe énergie sont beaucoup plus abondants dans les parties aériennes. Cela est dû à l'abondance des transcrits homologues à des protéines du photosystème, de la chlorophylle et de la Rubisco. La sur-représentativité de cette classe se fait au détriment des classes « communication cellulaire et transduction du signal », « Organisation cellulaire », « Transport » et « Trafic intracellulaire ».

Pour les racines, la plus grande part des EST soit près de 23% n'ont pu être affectés à une classe spécifique. La classe la plus abondante est ensuite celle du métabolisme, suivie de la classe correspondant à la synthèse des protéines. La classe regroupant les gènes de mort cellulaire, de défense et de réponse au stress atteint plus de 11% des EST de racines. Les classes les moins représentées concernent la communication cellulaire et la transduction du signal, la croissance et division cellulaire et le trafic intracellulaire.

D.2.4 Northern électronique

Cette méthode d'estimation du niveau d'expression des transcrits entre plusieurs banques a été appliquée aux banques de racines stressées et contrôles. Cette approche repose sur le nombre d'EST séquencés pour un gène donné. Pour estimer correctement l'expression d'un gène, il faudra être en mesure d'identifier quels sont les EST issus des transcrits de ce gène. Une première hypothèse possible est de considérer que chaque contig constitue un gène différent, ce qui soulève trois incertitudes majeures :

- Nous avons vu que l'assemblage repose sur la similarité des séquences et les zones de recouvrement entre EST pour la génération des clusters et des contigs. Il est donc possible que les transcrits d'un même gène soient répartis dans deux contigs distincts si aucune séquence ne permet de relier les deux alignements.
- Concernant les singletons, une situation similaire peut se présenter. Une séquence est classée dans les singletons si elle n'a pu être alignée avec d'autres séquences. Cela ne signifie pas pour autant que nous soyons en présence d'un gène supplémentaire. Les

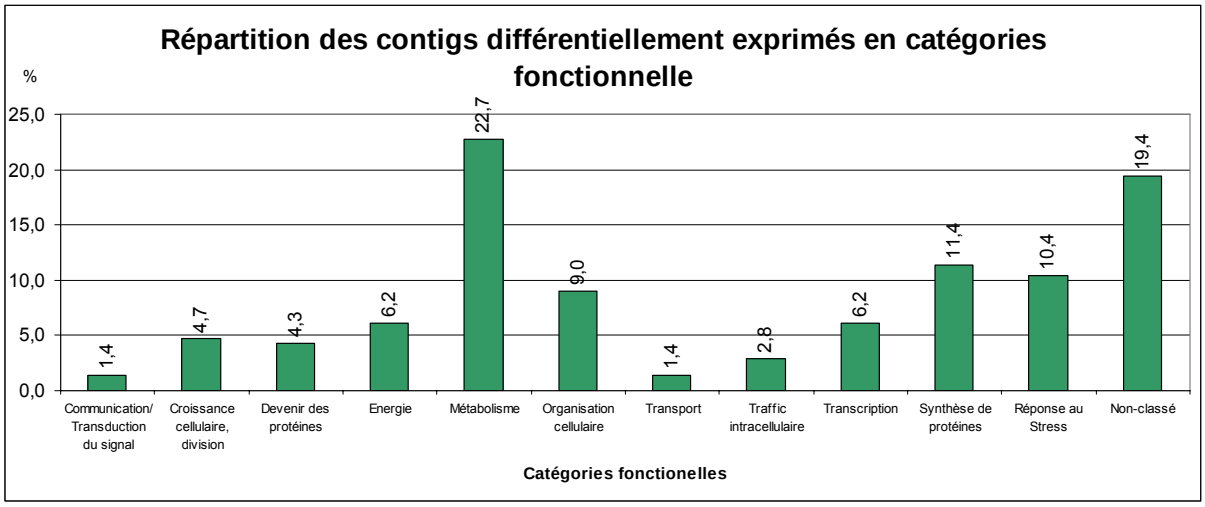


Figure D-16 : Répartition en catégories fonctionnelles des contigs détectés comme différentiellement exprimés *in silico*.

EST des transcrits d'un gène peuvent être regroupés au sein d'un contig mais ne couvrir que sa partie 5'. Un transcrit pour lequel l'EST est positionné en 3' sera alors considéré comme un singleton.

- La situation inverse est également possible, les EST de différents membres d'une famille multigénique peuvent être groupés sous le même contig avec d'autant plus de risques que la séquence en 3' n'est pas connue. Dans ces situations, l'estimation de l'expression du gène par comptage sera biaisée.

L'analyse statistique par IDEG6 a été réalisée en utilisant comme jeu de données les EST issus de racines en éliminant l'intégralité des singletons et des éventuelles séquences multiples d'un même clone. Le comptage des EST de chaque banque (stressé vs contrôle) a été fait en se plaçant au niveau du contig. Dans ce cas, nous considérons qu'un contig correspond à un gène, et aucune distinction n'est faite entre les EST du consensus primaire et ceux des éventuels consensus alternatifs. Le tableau Ax-1 (voir en annexe) donne la liste des 279 contigs détectés comme différentiellement exprimés par au moins un des tests statistiques sur les 2022 contigs analysés. Ils sont classés des plus fortes surexpressions dans la librairie 1 (racines contrôle) aux plus fortes sous expressions. La première constatation concerne le faible nombre d'EST par banque constituant les contigs. On peut se poser la question de la représentativité des données pour un décompte qui dépasse rarement 6 EST. Lorsque la normalisation (sur le nombre d'EST de chaque banque) est appliquée et qu'aucun EST n'a été obtenu dans une des banques, les différences d'expression sont artificiellement très élevées. Il faut donc prendre les résultats avancés avec précaution.

La figure D-16 nous présente les catégories fonctionnelles auxquelles les contigs détectés appartiennent. Par comparaison avec la répartition en catégories fonctionnelles de l'ensemble des EST (Fig. D-14), il apparaît que la proportion de « gènes » des classes « métabolisme », « organisation cellulaire » et « croissance cellulaire/ division » est plus élevée dans les « gènes » détectés comme différentiellement exprimés. A l'inverse, les classes « devenir des protéines », « transport » et « non-classé » sont sous-représentées.

Plusieurs contigs ont des annotations pouvant avoir un rôle dans la réponse au stress hydrique. Cependant, des transcrits de fonctions généralement mises en évidence lors du stress hydrique sont ici détectés comme plus abondants dans la banque de racines contrôles.

Nous avons l'Inositol 1-alpha-galactosyltransférase (EC 2.4.1.123) (CT469) avec un ratio de plus de 25 fois supérieur dans le contrôle. Cet enzyme permet la conversion du myo-

Inositol en Galactinol qui est un précurseur de la voie des RFO qui, comme nous l'avons vu, peuvent être accumulés lors du stress. La surexpression chez les racines contrôles ne va donc pas dans ce sens. C'est aussi le cas des contigs (CT477) et (CT219), avec respectivement des facteurs de 14 et 7 fois, homologues à des facteurs de liaison aux éléments de réponse à l'éthylène (EREBF). Une LEA (CT276) semble aussi régulée par un facteur 2.45 en faveur des contrôles. La prévention des oxydations est représentée par plusieurs peroxydases (CT168, CT1176 et CT235) et une L-ascorbate peroxydase (CT529) ainsi qu'une Flavonoide 3',5'-hydroxylase (CT1039) et « Metallothionein-like protein EMB30 » (CT431). Si ces fonctions ne sont pas des marqueurs spécifiques de la réponse au stress hydrique, il est courant de les mettre en évidence en situation de stress hydrique qui engendre aussi un stress oxydatif (Rabbani *et al.* 2003). Les peroxydases, comme les laccases peuvent également intervenir dans la polymérisation des monolignols en lignine et donc avoir un rôle dans la synthèse des parois. La sous-expression lors du stress d'un transcrit homologue à une « Metallothionein-like protein EMB30 » conforte les résultats de Dubos *et al.* sur les racines de pins maritime en stress osmotique (Dubos and Plomion 2003). Nous pourrons enfin citer une Polcalcin (T56), un précurseur d'Alpha-expansin (CT859) et une Défensine (CT166) qui sont des acteurs de la réponse au stress pour respectivement la transduction du signal, l'extensibilité des parois et le système de défense contre les pathogènes.

A l'opposé, parmi les contigs comportant plus de séquences de la banque stressée, nous trouvons également une liste de fonctions intéressantes concordant plus avec le sens de variation attendu en condition de stress hydrique. Parmi les plus forts ratios d'expression, nous retrouvons une LEA (CT275), des protéines de liaison au calcium (CT59)(CT60) et une Flavonoide 3',5'-hydroxylase (CT346). Les contigs en rapport avec la réponse à l'oxydation et à la détoxication sont plus abondants ce qui concorderait avec l'hypothèse d'un stress oxydatif plus fort chez les plants en stress hydrique. Les représentants sont des « Metallothionein-like protein » (CT312 et CT314), des sous-unités du complexe endopeptidase du Proteasome (CT1596 et CT1609) ainsi qu'un précurseur de « cathepsin B-like cysteine proteinase » (CT1841) qui est une protéine localisée dans le lysosome. Des enzymes intervenant à plusieurs niveaux dans la voie de synthèse des Flavonoïdes sont également détectés avec une 3',5'-hydroxylase (EC 1.14.13.88) (CT346, CT342, CT344) et une Flavonol 3-O-glucosyltransférase 5 (EC 2.4.1.91) (CT1670) (http://www.genome.ad.jp/dbget-bin/show_pathway?map00941+1.14.13.88+2.4.1.91). Dubos *et al.* avaient également trouvé une accumulation de transcrits homologues à la Flavonoïde 3',5'-hydroxylase (EC 1.14.13.21). Nous trouvons aussi des fonctions en rapport avec la structure des parois, comme une

Xyloglucan:xyloglucosyl transferase (CT1600) et un précurseur de la Glucan endo-1,3-beta-glucosidase (CT1704), ce qui concorderait avec une modification des parois cellulaires lors d'un stress hydrique. Le contig CT1153 homologue à une sous unité zeta-1 du Coatamer pourrait indiquer une activité plus importante de glycosylation des protéines au niveau du Golgi, se répercutant par un nombre de vésicule plus important en condition de stress. On pourra également noter que les transcrits de l'histone H1 (CT1938 et CT994) ainsi que l'histone H3 (CT1610) seraient plus abondant en condition de stress alors que l'Histone H4 (CT309) était détectée plus exprimé dans la banque contrôle. Une différence sur le niveau de compaction de l'ADN est donc possible entre la condition de stress et la condition contrôle. Nous terminerons en constatant que des facteurs de transcription, des protéines de réponse aux hormones et des protéines impliqués dans la transduction du signal, sont détectés dans la banque stressée. La liste comporte une protéine en doigt de zinc (CT934 et CT1822), un facteur de transcription WRKY (CT1598), une protéine homéobox-leucine zipper (CT1556), une « ABA stress ripening protein » (CT132), une protéine de membrane induite par l'ABA (CT1850) et une MAP kinase (CT1805). Ces résultats souligneraient une régulation de la transcription plus active dans les plantes exposées au stress hydrique.

Nous avons évoqué la problématique associée à l'identification des EST des transcrits d'un gène et le possible regroupement des EST d'un même gène dans plus d'un contig, ou à l'inverse, la distinction possible des membres de familles multigéniques à l'échelle des consensus. Nous avons donc essayé d'identifier les EST issus d'un gène à partir du regroupement le plus strict représenté par les consensus. Puis nous avons regroupé les EST de plusieurs CN en tenant compte de leur appartenance à un CT, à un CL et à leurs annotations. Le résultat du Northern électronique des 2097 « gènes » obtenus est donné dans le tableau Ax-2. 247 gènes sont détectés. Si les valeurs des tests changent, dans l'ensemble nous retrouvons le même type de fonctions que dans l'analyse précédente. Les gènes connus pour leur implication dans le stress hydrique que l'on trouvait plus abondants dans la banque contrôle ne sont pratiquement plus détectés tandis que ceux de la banque stressée le sont toujours. Cette seconde approche à l'échelle des consensus avec regroupement sur les annotations, quoi que plus subjective donne des résultats plus en accord avec les résultats attendus sur la réponse au stress hydrique.

L'analyse des résultats de ce Northern *in silico* a mis en évidence des contigs homologues à des gènes connus pour leur implication dans la réponse au stress hydrique.

Toutefois, certains d'entre eux sont détectés pour leur abondance dans les EST de la banque contrôle. L'explication de ce résultat peut tenir aux biais possibles de l'approche technique. Nous avons évoqué les biais bioinformatiques, mais la création même de la banque n'est pas exempte de biais. En effet, un biais peut exister à la création de la banque au moment de l'extraction et de la transcription inverse, mais aussi la facilité d'insertion de l'insert dans le vecteur qui dépendra de la taille et de l'abondance des transcrits, et de la réussite du clonage. La taille des inserts influe sur la facilité de transformation des bactéries, leur croissance et l'amplification des inserts par PCR. En terme d'abondance relative des transcrits, la représentativité du transcriptome par les séquences finalement obtenues n'est pas garantie. Les plants en hydroponie utilisés pour la création de la banque peuvent également présenter un stress en condition de contrôle différent (ex : anoxie, oxydatif, pathogène …) de celui appliqué par ajout de PEG.

D.2.5 Unigène

L'objectif d'un Unigène est de limiter au maximum la redondance du lot de séquences qui sera utilisé pour les études d'expression. Le principe – exposé dans la partie Matériels et Méthodes – est de sélectionner une séquence représentative de chaque gène. Dans la pratique cela se traduit par la sélection d'un clone pour chaque contig ainsi que l'ensemble des singletons. Dans notre cas nous avons essayé d'éliminer la redondance possible au sein des singletons sur la base des résultats de Blast. Le projet qui nous a servi de base de travail est le projet « Pinaster total » qui regroupait l'ensemble des séquences de pin maritime.

La méthode de recherche d'homologie appliquée aux singletons a permis d'éliminer 815 singletons sur les 5051. Le nombre d'accessions issues de la sélection de singletons est donc de 4236 auxquelles viennent s'ajouter les accessions représentatives de chaque contig.

Pour les représentants de contig, nous avons voulu homogénéiser la longueur des fragments amplifiés. Même si les sondes sont déposées en excès par rapport à la cible, l'intérêt de cette démarche est de limiter la variation du nombre de copies de chaque sonde déposée sur la lame. Des fragments de tailles très différentes auront des rendements PCR variables et le nombre de copies à concentration d'acides nucléiques constante sera d'autant plus faible que la taille du fragment est élevée. La figure D-17 montre que la distribution des longueurs théoriques des EST retenus est bien centrée sur 600 paires de bases et que la plus grande partie des longueurs sont comprises entre 450 et 700 paires de bases. La stratégie de sélection développée dans le cadre de cette thèse a donc permis d'atteindre notre objectif, qui était d'homogénéiser les longueurs théoriques des EST retenus.

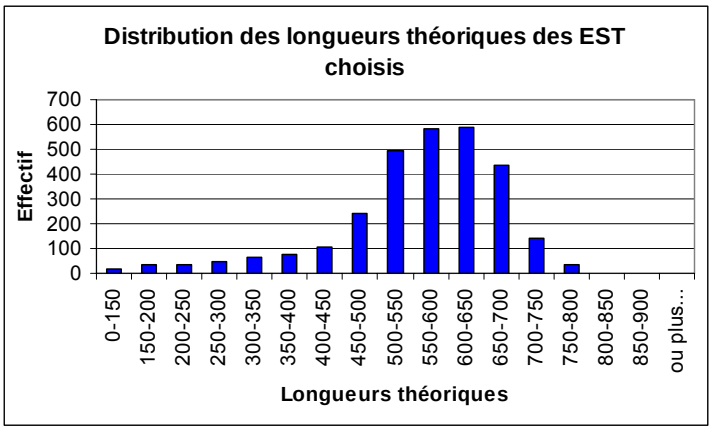


Figure D-17 : Distribution des longueurs théoriques des ESTs appartenant aux contigs sélectionnés pour l'Unigène

Après amplification de l'ensemble des clones de l'Unigène par la société PICME, nous avons eu accès à la taille réelle des fragments amplifiés. La figure D-18 montre la distribution des écarts entre la longueur théorique que nous avons calculée et la longueur réelle des inserts estimée par lecture sur gel d'agarose. La distribution n'est pas centrée sur zéro comme nous l'espérions. Une des explications sur l'origine de ce décalage tient à l'amplification des régions du vecteur flanquant l'insert et à la lecture sur gel. Nous pouvons voir que la distribution présente toutefois une classe majoritaire. Si l'on déduit des longueurs réelles la valeur du mode – c'est-à-dire la valeur la plus représentée-, soit 208, il est possible de représenter graphiquement la distribution des longueurs théoriques et des longueurs réelles (Fig. D-19). Nous pouvons voir sur la figure D-19 qu'une relation linéaire existe pour les fragments dont la longueur réelle est inférieure à 700 paires de bases. Au-delà, la longueur des inserts était sous-estimée. Ces résultats montrent les limites de notre méthode d'estimation bioinformatique de la longueur des inserts. Elle repose sur l'hypothèse que la séquence consensus couvre l'intégralité de la séquence du transcrit en 3'. Nous pouvons déduire des résultats que ce n'est pas le cas et qu'il nous manque la partie 3' pour un grand nombre de gènes.

La figure D-20 montre que la longueur réelle de l'insert a été estimée avec une erreur de moins de 20 % pour la moitié des clones. Si l'on considère que l'erreur vient de l'absence de séquence en 3', alors la figure D-21 indique que pour 37 % des EST considérés plus de 250 paires de bases ne sont pas séquencées en 3' et pour 21 % des EST la partie non connue du contig est supérieure à 500 pb. Il serait donc intéressant de procéder à un séquençage en 3' des clones appartenant aux contigs concernés. Ainsi, la sélection d'un tiers des contigs, (soit approximativement 800 contigs) et le séquençage de 2 ou 3 clones par contigs enrichirait considérablement la quantité d'information disponible. La partie 3' est la région la plus spécifique des membres de familles multigéniques. La disponibilité d'informations de séquences dans cette zone permettrait : 1) d'améliorer la spécificité d'amplification lors d'études d'expression par PCR quantitative et la sélection d'oligomères pour la réalisation éventuelle de « puces oligo » en dessinant les amorces en 3' ; 2) d'augmenter la probabilité de détection de polymorphismes SNP.

La réalisation de l'Unigène nécessite également un réarrangement des clones bactériens sélectionnés. Cette étape a été réalisée au laboratoire. Cela nécessite le repiquage de milliers de clones et les risques principaux sont les contaminations et les erreurs de repiquage. Une automatisation de ces transferts de clones permet d'éviter les erreurs et de

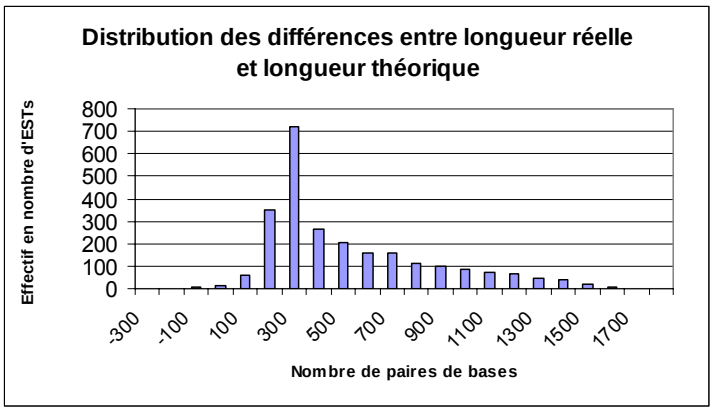


Figure D-18 : Distribution des différences entre la longueur réelle et la longueur théorique des inserts des clones sélectionnés pour l'Unigène.

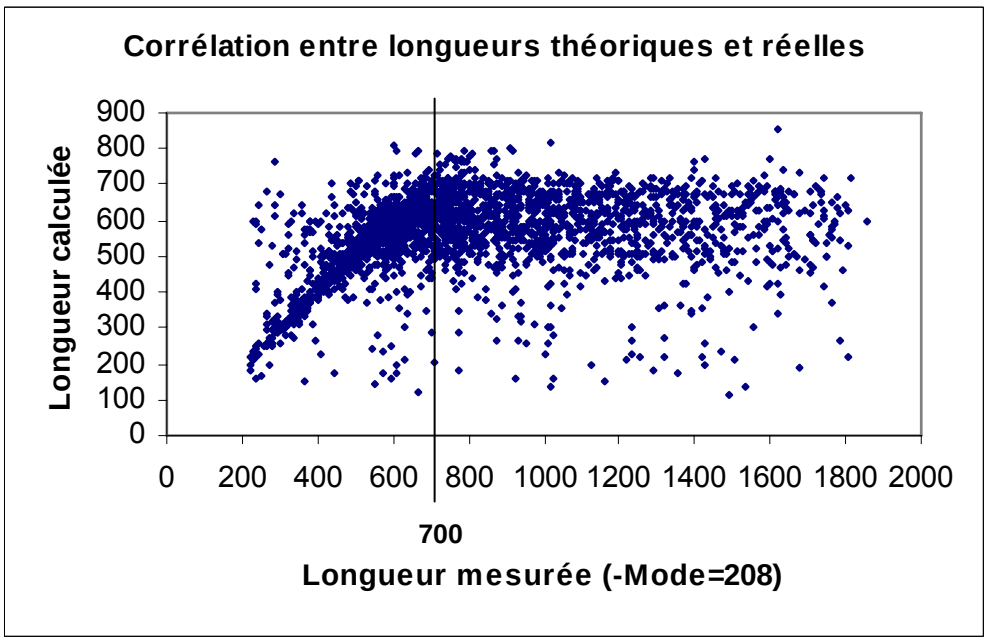


Figure D-19 : Représentation graphique des longueurs calculées en fonction des longueurs mesurées pour les inserts sélectionnés lors de la création de l'Unigène.

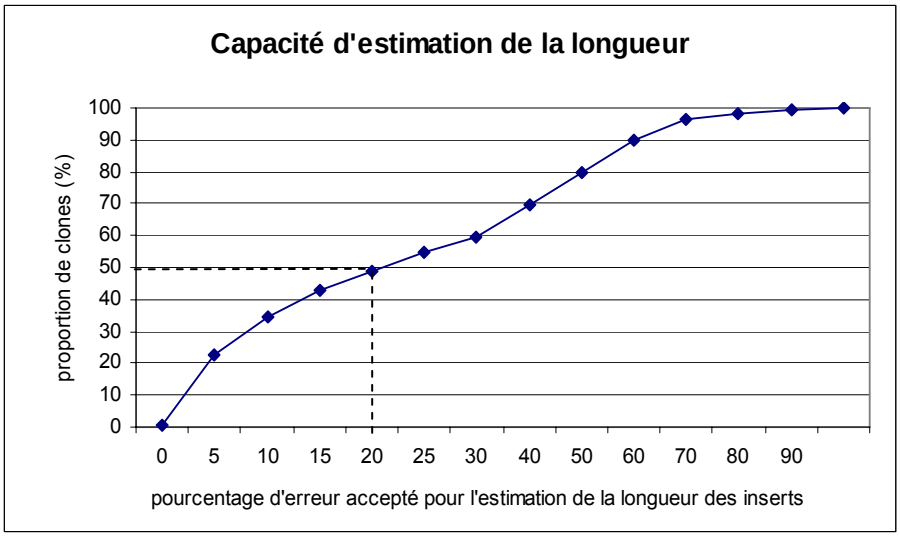


Figure D-20 : Capacité d'estimation de la longueur des inserts contenus dans les clones de nos banques de pin maritime.

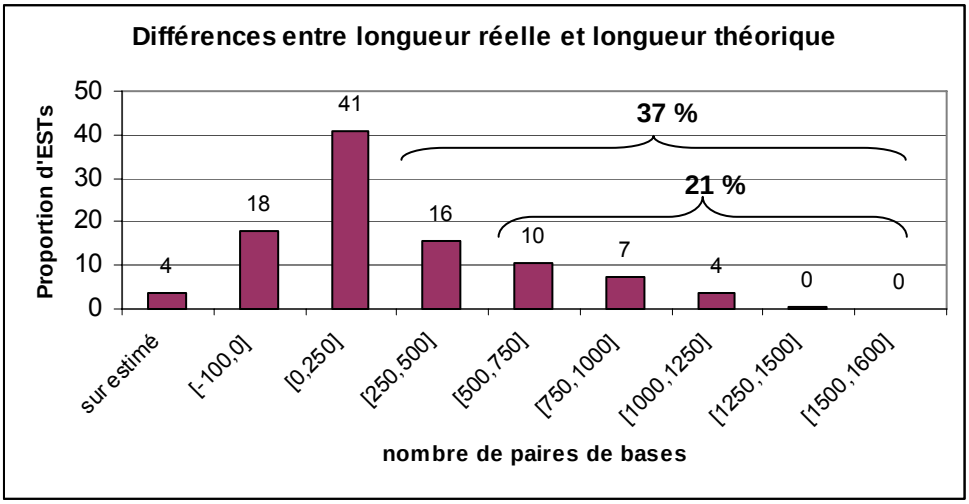


Figure D-21 : Proportion d'ESTs et différence entre la longueur réelle de l'insert (valeur du mode déduit) et sa longueur théorique.

s'assurer d'une traçabilité des échantillons. Dans ce but, j'ai adapté et optimisé une station robotisée de pipetage de type Tecan Genesis pour cette application. J'ai également réalisé Une interface graphique pour permettre la génération des fichiers de commande du robot à partir de la liste des clones de l'Unigène et des fichiers d'identification des banques tout en conservant une traçabilité des échantillons sous la forme de fichiers et de plans de plaques compatibles avec les fichiers d'entrée utilisés par le séquenceur de l'UMR. Ce système s'est avéré efficace et a permis d'éviter les erreurs de repiquages et les contaminations pour repiquer les 7118 clones de l'Unigène. A l'issu du réarrangement, un double de l'ensemble des banques ainsi que l'Unigène a été expédié à la société PICME pour la réalisation des puces.

D.2.6 Discussion

En 1999, pratiquement aucune séquence n'était disponible chez le pin maritime. En 2000, les séquences de xylème ont constitué une première base d'informations. Au début de cette thèse, les données de séquences disponibles pour *Pinus pinaster* dans les bases de données publiques concernaient essentiellement les séquences de xylème, les séquences des parties aériennes et quelques séquences de racines générées au cours de projets cités au paragraphe D.3.2. En ajoutant près de 9000 EST à ce catalogue, le travail réalisé au cours de cette thèse a permis d'augmenter significativement les ressources en EST disponibles sur cette espèce. A ce jour, 19201 séquences sont disponibles dont 19047 ont été produites dans des projets impliquant notre laboratoire. Aux 18254 séquences d'EST déjà disponibles viennent de s'ajouter les 8131 séquences du projet « bourgeon ». En quelques années, les ressources génomiques de pin mises à la disposition de la communauté scientifique ont considérablement augmentées et constituent une source d'information permettant d'envisager de nouvelles approches comme la recherche de polymorphisme de type SNP dans les séquences. Il y a actuellement dans dbEST près de 329469 EST de *Pinus taeda*, 18254 EST de *Pinus pinaster*, 1663 EST de *Pinus sylvestris* et seulement 69 EST de *Pinus radiata* (1914 séquences dans EMBL) (http://www.ncbi.nlm.nih.gov/dbEST/index.html). L'assemblage bioinformatique permet d'accéder à un autre niveau d'information que les simples données de séquences. La compilation des données d'espèces du même genre comme le projet Pinus du TIGR offre la possibilité de combiner les informations disponibles et d'améliorer les possibilités d'annotation. L'intérêt est toujours de créer un assemblage sur le maximum de séquences disponible pour obtenir la couverture la plus élevée du génome exprimé et permettre un assemblage de qualité. Nous avons vu comment chez le pin maritime, un

séquençage en 3' de séquences bien choisies permettrait d'améliorer considérablement la couverture en 3'.

Nous avons réalisé une annotation la plus précise possible de nos séquences du projet « racines »avec les informations disponibles à un moment donné. La démarche adoptée nécessite d'y consacrer un temps considérable estimé à 5 mois dans le cadre de cette thèse. Une des difficultés principales de l'annotation est de maintenir les données à jour avec l'enrichissement quotidien des bases de données. Les systèmes d'annotation automatique ne font que transférer l'annotation du résultat sur la requête. Une autre possibilité vient d'être développée sur le pipeline que nous avons développé avec le Centre de Bioinformatique de Bordeaux. Une fois les annotations assignées, la mise à jour des bases de données permet de relancer les Blast. Seules les accessions pour lesquelles une homologie avec un score supérieur aux précédents hits sont identifiées et seront annotées de nouveau.

L'approche de Northern *in silico* permet de donner une indication sur les niveaux d'expression entre banques. Elle a mis en évidence des gènes distribués différentiellement en rapport avec la réponse au stress hydrique. Cette méthode est toutefois très sensible à la composition des banques et à l'assemblage informatique. L'analyse d'expression ne peut pas se limiter à ces approches *in silico* et ne saurait remplacer une analyse par hybridation ou qPCR. Nous en retiendrons parmi les fonctions les plus régulées la détection de facteurs de transcription comme WRKY, protéine en doigt de zinc et protéine homéobox-leucine zipper, qui mériteraient une étude plus approfondie.

Les données obtenues en réalisant l'Unigène ont mis en évidence que les données de séquences d'un programme de séquençage d'EST en 5' ne permettent pas de couvrir la totalité de l'extrémité 3'. La détermination d'un Unigène sur la seule base de ces données ne permet pas de choisir des clones de taille homogène dont la position couvre uniquement l'extrémité 3' gage de spécificité. Il n'est pas non plus raisonnable d'envisager la sélection d'oligomères pour l'étude d'expression sans avoir connaissance de la région 3', l'idéal étant de disposer d'un signal d'expression associé à la fonction donc en 5' et un signal plus spécifique donc en 3'. Il serait judicieux de compléter les informations de séquence par un séquençage en 3' pour un sous ensemble de contigs dont au moins 500 paires de bases ne sont pas connues en 3'. Cela va dépendre de la qualité de la banque obtenue. La position et la répartition des EST le long du transcrit pleine longueur pourrait nous informer sur la richesse de la banque en ARNm pleine longueur et si l'étape de transcription inverse a été performante. Sur une banque répondant à ces deux critères, on peut s'attendre à obtenir des

EST ayant une bonne couverture des 800 paires de bases en 5' au détriment de la partie 3'. Dans ce type de projets de séquençage massif d'EST, une stratégie serait de commencer par un séquençage en 5' qui a l'avantage de fournir des séquences pour lesquelles l'attribution d'une fonction est facilitée. Puis, après analyse bioinformatique et estimation de la longueur réelle de chaque contig, de concentrer l'effort de séquençage sur la région 3' des contigs en fonction de l'information déjà disponible. L'accès à la séquence en 3', outre les avantages déjà cités, permet par dessin d'une amorce spécifique et utilisation d'une technique de 5' RACE d'accéder à la séquence complète du transcrit. Ce travail aura donc permis d'identifier un sous ensemble de clones pour lequel un séquençage en 3' sera réalisé.

D.3 Analyse du transcriptome

Ce chapitre présentera les résultats générés et les analyses réalisées pour les deux techniques d'étude du niveau d'expression des transcrits que nous avons utilisées. Nous verrons tout d'abord les résultats d'hybridation avec la technique des microarrays puis la détection du niveau d'expression de gènes spécifiques par l'approche de PCR quantitative.
L'URL : http://www.pierroton.inra.fr/biogeco/genetique/theses.html donne accès en ligne à une partie des résultats.

D.3.1 Extraction des ARN totaux

Le protocole d'extraction des ARN totaux de racines de pin maritimes combiné à une récolte rapide des tissus sous azote liquide et un stockage à -80°C a permis d'obtenir des ARN de très bonne qualité sans dégradation (Fig. D-22). La synthèse des cibles fluorescentes selon le protocole de la société Amersham nécessite cependant de grandes quantités d'ARN à des concentrations élevées. Selon ce fabricant, la synthèse d'une sonde nécessite 500ng d'ARN messagers. En considérant que la proportion en ARNm est proche de 0.5 à 1% des ARN totaux, chaque réaction nécessite entre 50 et 100 µg d'ARN totaux. Dans notre cas nous avons décidé d'utiliser une quantité d'ARN totaux d'environ 500ng par synthèse. Les solutions d'ARN devaient avoir une concentration supérieure à 5µg/µL. Des solutions mères ont été obtenues avec une concentration comprise entre 6 et 10 µg/µL. Vu le plan d'expérience retenu (loop design), jusqu'à 7 synthèses distinctes sur le même échantillon étaient requises, ce qui au final représente un minimum de 350µg d'ARN total. L'optimisation du protocole d'extraction m'a permis d'augmenter le rendement de près de 30%. Pour chaque échantillon (1 écotype dans une condition), plusieurs extractions d'ARN indépendantes ont été réalisées. Vu le développement des racines très enchevêtrées, l'individualisation des plants était impossible au moment de la récolte. La masse racinaire a donc été divisée en plusieurs lots sur lesquels les prélèvements pour l'extraction des ARN ont été répartis.

D.3.2 Puces à ADNc

Dans cette partie, les accessions et les identifiants indiqués dans les tableaux font référence à l'insert complet présent dans le clone dont l'identifiant est donné. Les données d'expression associées à ces accessions font donc référence au signal obtenu par hybridation sur la totalité de l'insert et non à la seule séquence accessible par le numéro d'accession.

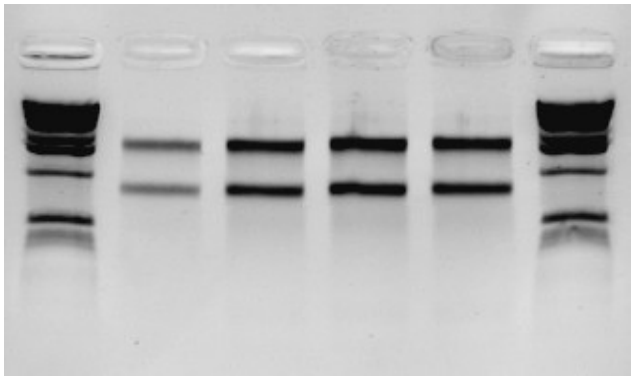


Figure D-22 : Dépôt sur gel d'agarose des ARN totaux de racines de pin maritime. On note la présence des 2 bandes majeures correspondant aux ARN ribosomaux 18S et 28S.

Figure D-23 :
Vue d'une des lames après hybridation, le joint recouvre en partie la première ligne de 4 blocs. (Image obtenue par combinaison des couleurs artificielles verte et rouge des deux canaux utilisés pour la lecture.)

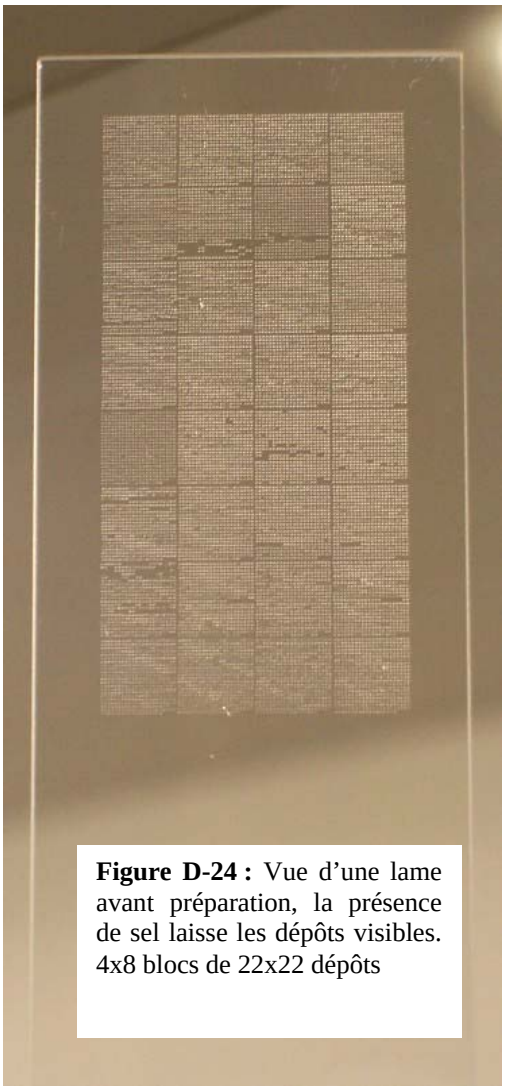


Figure D-24 : Vue d'une lame avant préparation, la présence de sel laisse les dépôts visibles. 4x8 blocs de 22x22 dépôts

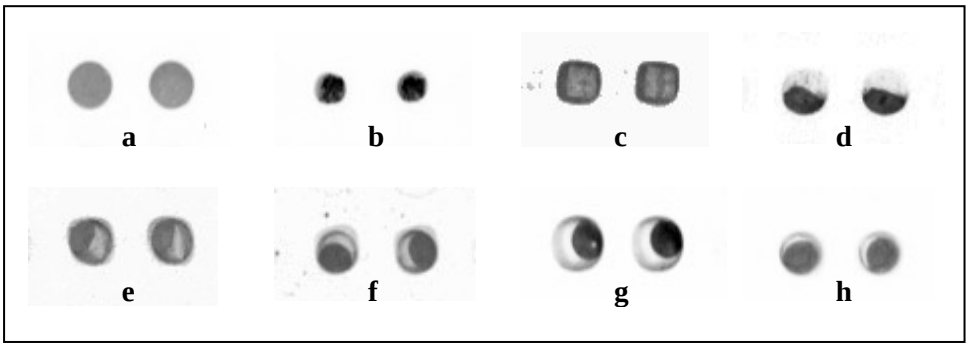


Figure D-25 : Formes de spots anormales rencontrées sur les lames utilisées

D.3.2.1 Caractéristiques des lames produites et mise en œuvre

D.3.2.1.1 Qualité des dépôts d'ADNc

Les lames produites par notre partenaire autrichien (www.picme.at) comportent 32 blocs organisés en 8 lignes et 4 colonnes (Fig. D-23). Chaque bloc est constitué de 22 lignes et 22 colonnes pour un total de 484 dépôts/Bloc (Fig. D-24). Le nombre de dépôts sur ces lames atteint 15488 spots. Si l'on ne considère pas les « blancs » et les « contrôles » (4CL, 5P, G3PDH, Desmine, Nebuline, Luciferases, Tubulines, UDPG, GP, P450, HSP, RAS, Sinucléines, Oligos, etc.), les lames obtenues comportent 7265 accessions différentes, chacune déposée en double. Pour des raisons de facilité de production, les dupliquas ont été déposés côte à côte par la société.

Nous avons rencontré un problème de compatibilité entre l'emplacement de la zone de dépôt et la position de la chambre d'hybridation équipant la station d'hybridation Tecan du pôle Transcriptomique de Bordeaux. Sur la figure D-23, il apparaît clairement que le joint d'étanchéité de la chambre d'hybridation (en vert) recouvre en partie les 4 premiers blocs. L'information des 9 premières lignes est inexploitables soit un total d'environ 800 dépôts. A cela se rajoute des phénomènes de perturbation de l'hybridation à l'endroit de l'injection (ici dans l'angle en haut et à droite) et aux extrémités supérieure et inférieures de la chambre d'hybridation où se trouve une petite dépression dont le rôle est de retenir les bulles pouvant se former. Nous pouvons également observer dans le bas de la lame, heureusement non occupé par les dépôts, des traces dont l'origine probable est une perturbation des flux provoquée par l'étiquette du code à barres.

Dans l'ensemble, la qualité des dépôts était satisfaisante, toutefois, une partie des dépôts (<5%) présentaient des formes irrégulières. La figure (Fig. D-25a) représente la forme idéale attendue, mais certains spots avaient un diamètre plus faible (Fig. D-25b) ce qui pouvait parfois se traduire par une saturation du signal. D'autres spots anormaux n'étaient pas circulaires (Fig. D-25c et d) mais le plus souvent avec des formes complexes (Fig. D-25e à h). Pratiquement aucun spot en forme d'anneau (« donuts ») ou de comète n'a été rencontré. Le problème posé par ces spots non circulaires est la délimitation du spot par le logiciel d'analyse d'image dont la grille de détection est basée sur une forme circulaire à diamètre ajustable. Aucun des paramètres automatiques testés ne s'est avéré satisfaisant. Une relecture et une délimitation manuelle après la détection automatique a été réalisée pour éliminer les spots aberrants ou affiner la délimitation des spots.

D.3.2.1.2 Synthèse des cibles et hybridation

La station d'hybridation peu accueillir un nombre maximum d'hybridations simultanées de 12 lames. Cependant, une étape du protocole de purification ne devant pas excéder 10 minutes, il est difficile d'excéder l'hybridation de 8 lames simultanément. Pour réduire l'hétérogénéité de chaque série d'hybridation, nous avons divisé les hybridations en plusieurs séries correspondant chacune à une des boucles du « loop design » (Fig. D-26). Concernant l'expérience « stress court », une première série de 6 lames regroupe les hybridations de l'écotype landais (triangle), une seconde pour l'écotype marocain (triangle) et une troisième pour les lames restantes, ce qui fait un total de 18 lames. Les lames de l'expérience « stress long », au nombre de 8, ont pu être traitées en une seule fois (Fig. D-27). Quelques lames supplémentaires ont été hybridées à la fin de l'expérience pour compléter les informations des lames pour lesquelles le signal était de moins bonne qualité.

L'hybridation des lames se fait à quantité d'ADNc constante. Afin de réduire l'hétérogénéité au sein d'une même série, la quantité d'ADNc hybridée a été ramenée à celle de l'échantillon le plus faible. Les quantités d'ADNc dosées après marquage et purification des cibles variaient d'un facteur 1 à 3 pour une même série. Il n'était pas envisageable d'effectuer une nouvelle synthèse pour les cibles dont la réaction avait moins bien fonctionnée, étant donné que les cibles fluorescentes ne sont pas stables dans le temps.

D.3.2.1.3 Acquisition des données

Le niveau de fluorescence détecté par le scanner dépend de la puissance du laser et du gain du photomultiplicateur (PMT). Nous avons choisi de conserver des valeurs fixes de ces paramètres pour l'ensemble des lames, pour ne pas introduire un facteur supplémentaire de variation dans les résultats. Les gains ont été fixés à 530 Volts pour le canal à 532nm et à 680V pour le canal à 635nm d'après la lecture de deux lames de test. Après avoir effectué la lecture des premières lames, il est apparu que le signal représenté en rouge (Cy5) correspondant à la longueur d'onde 635nm était inférieur au signal vert (Cy3). Nous avons décidé de ne pas modifier les paramètres choisit et de corriger les données par une normalisation statistique. Comme évoqué plus haut, la détection des spots a été réalisée de façon automatique puis manuellement par relecture de l'intégralité des spots. Le bruit de fond des lames était globalement faible. Tous les spots dont le signal était perturbé par la présence de poussières, d'un bruit de fond trop important ou de zones de mauvaise hybridation ont été éliminés.

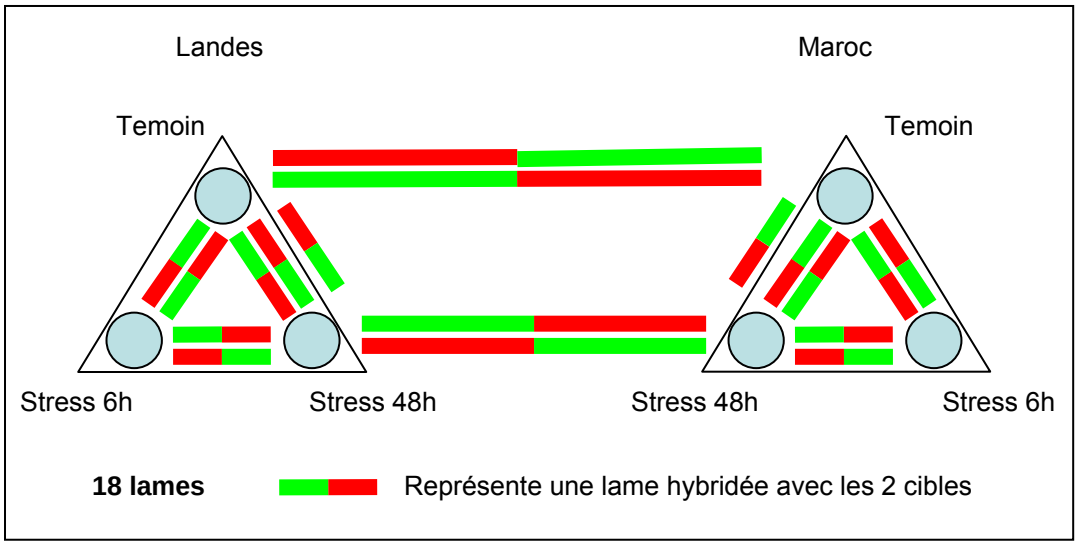


Figure D-26 : Schéma des hybridations réalisées pour l'expérience « Stress court ».

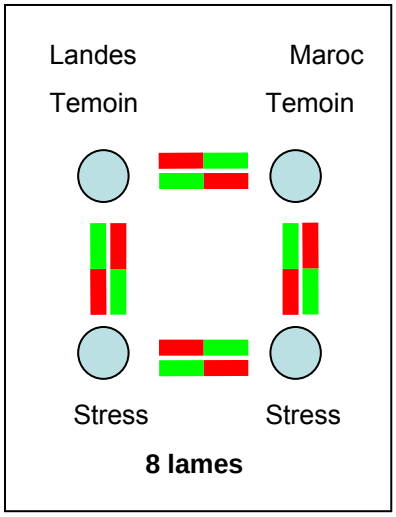


Figure D-27 : Schéma des hybridations réalisées pour l'expérience « Stress Long ».

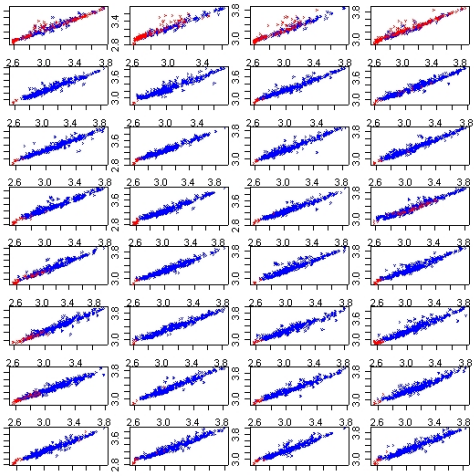


Figure D-28 : « Gridcheck », représente le signal Cy3 et Cy5 pour chaque bloc (ici 32 bloc)

D.3.2.2 Traitement des données brutes de fluorescence

D.3.2.2.1 Préparation et étude descriptive des données

Le logiciel d'analyse d'image Genepix Pro utilise un fichier de type « .gal » pour l'identification des spots sur la lame. Ce fichier, transmit par la société Picme, utilisait un système de « code barre » pour la traçabilité et l'identification des échantillons déposés. Le premier traitement des fichiers de données de fluorescence a donc consisté à changer ces codification par l'identifiant des clones qui nous servira pour désigner les transcrits, et à simplifier les fichiers en éliminant les informations non utilisées. Par la suite, l'utilisation de différents outils d'analyse et les libraires de fonctions du logiciel R ont exigé la présentation des données sous des formats spécifiques. Les fichiers requis pouvait dépasser 570000 lignes ce qui n'est pas pris en charge par les tableurs classiques comme Excel. J'ai donc utilisé le langage de programmation Perl pour effectuer les manipulations de fichiers.

La première transformation appliquée est le passage en logarithme de base 2 des intensités de fluorescence. Cela a permis d'obtenir une distribution avec moins de valeurs extrêmes (données non présentées) et d'être plus cohérente avec les données d'intensités des pixels qui sont comprises entre 0 et 2^{16}-1 (Martin 2002).

Les données de fluorescence comportent de nombreux biais dont les sources vont de la technique utilisée aux variations expérimentales. Nous avons voulu utiliser le module « MAANOVA » sous R dédié à l'analyse des données issues de microarrays. Ce module tient compte des mauvais spots (marqués « mauvais » par l'utilisateur) et des spots non détectés (indiquées par NA, la valeur étant celle du bruit de fond à l'emplacement attendu) mais de ne prend pas en charge (génération d'une erreur logiciel) les spots non-détectés dont l'intensité a une valeur nulle ce qui était le cas des spots situés en dehors de la chambre d'hybridation. Tous les spots dont au moins une valeur était nulle sur une des lames ont été éliminés des fichiers de données. L'utilisation des fonctions « gridcheck », « riplot » et « arrayview » de la librairie « MAANOVA » sous R a permis de générer pour chaque lame un ensemble de représentations graphique dont un exemple est présenté dans les figures (Fig. D-28 ; Fig. D-29 ; Fig. D-30). Ces outils descriptifs permettent d'estimer de façon approximative la qualité des hybridations. Sur la figure D-28, le nuage de points doit se rapprocher d'une droite ce qui est ici le cas. Les spots marqués comme mauvais apparaissent en rouge. Nous remarquons que la distribution des points rouges est concentrée sur les 4 premiers blocs qui comme nous l'avons signalé regroupent les dépôts situés hors de la chambre d'hybridation. Sur les autres blocs, les spots rouges concernent essentiellement des valeurs très faibles d'intensité. Les

nuages de points se répartissent dans une gamme d'intensité (indiquée par l'échelle) proche pour l'ensemble des blocs. Le dépôt et l'hybridation sont donc relativement homogènes sur l'ensemble de la surface de la lame. La figure D-29 permet d'accéder au même type d'information sous une autre forme, les premières lignes de dépôt sont cette fois en bas de la figure. Le ratio d'intensité entre les deux canaux est représenté par des couleurs artificielles allant du rouge au vert. La lame présentée ici permet également de mettre en évidence le problème de positionnement du joint de la chambre d'hybridation. Aucun signal n'est détecté pour les spots hors de la chambre (bande noire en bas) puis les spots situés sous le joint saturent en intensité (ligne blanche). La zone d'injection est également visible par une zone verte en bas à gauche (en haut à droite sur la lame). On peut également se rendre compte que l'hybridation n'est pas homogène sur la partie supérieure droite de la lame (zone en bas à gauche de la figure). Une attention particulière devra être portée pour éliminer les spots concernés. L'intensité du canal représenté en vert est, comme nous l'avons déjà signalé, globalement supérieure au canal représenté en rouge

Le « RI plot » est une représentation graphique classique permettant de donner le ratio des signaux rouges sur vert ($\log_2(R/G)$) en fonction du produit des intensités ($\log_2(RG)$). Idéalement, sur l'ensemble des données de fluorescence de la lame, le signal rouge et le signal vert devraient être équilibré quelque soit l'intensité de fluorescence. Le nuage de points obtenus devrait être horizontal avec un ratio de 1 (soit centré sur 0 en log2) et distribué de façon Normale en intensité. Sur la figure D-30, ce n'est pas le cas. Les données nécessitent donc une correction. Ce type de distribution, rencontrée pour l'ensemble des lames, résulte d'une part au déséquilibre de signal entre le rouge et le vert et d'autre part aux mauvais spots (chambre d'hybridation) ou à la présence de dépôts de sondes fluorescentes et autres contrôles (série de points à gauche et en bas).

Ces représentations graphiques qui permettent de valider une expérience, peuvent donc guider les analyses et mettre en évidence des problèmes spécifiques. Sur l'ensemble des lames hybridées, au vu des résultats des représentations graphiques présentées, une seule lame de l'expérience « stress court » a dû être écartée. Elle concernait une hybridation des cibles de l'écotype marocain en condition contrôle et en condition 6 heures de stress. Un swap de deux lames a donc été réalisé pour ces deux échantillons amenant le nombre de lames de l'expérience « stress court » à 19. Pour l'expérience « stress long », quelques problèmes de traces résiduelles après hybridation et la qualité des données obtenues, nous ont incités à effectuer un swap de deux lames supplémentaires entre les échantillons contrôles et stressés de l'écotype landais. Dix lames sont donc disponibles pour cette expérience.

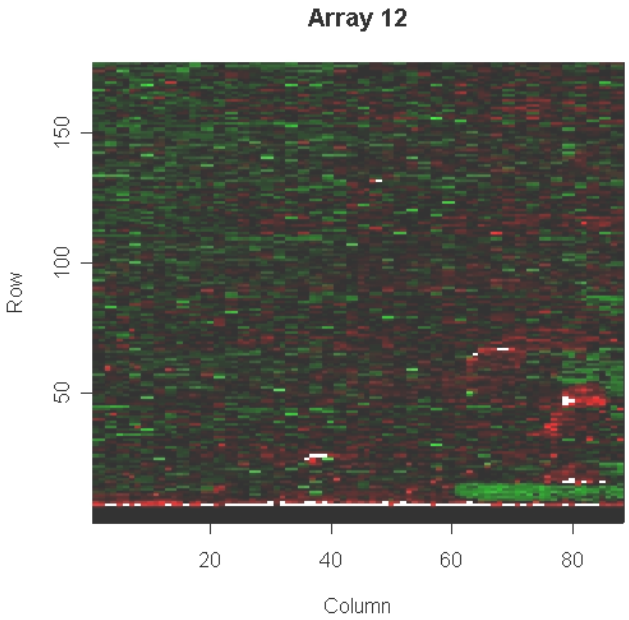


Figure D-29 : « Arrayview », représente l'intensité relative du signal Cy3 par rapport à Cy5 selon le schéma de dépôt.

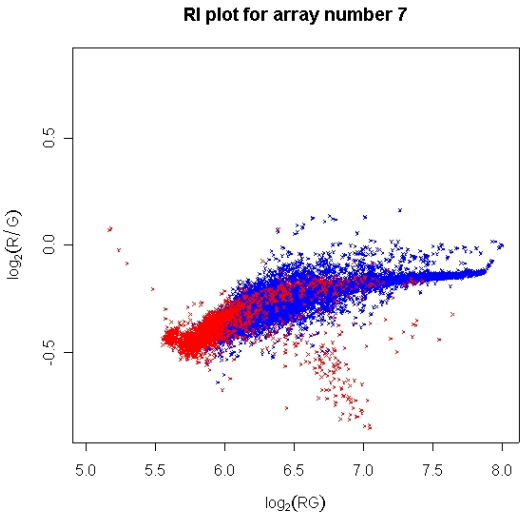


Figure D-30 : « RI plot », représente le logarithme du rapport du signal Cy5 sur Cy3 en fonction du logarithme du produit des intensités.

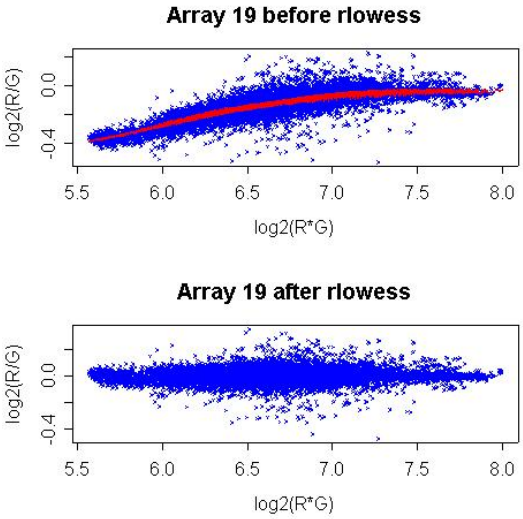


Figure D-31 : « rlowess», représente le « RIplot » avant et après la correction des données par une fonction de lowess régional.

D.3.2.2.2 Correction des données

Le bruit de fond n'a pas été pris en compte pour corriger les données. En effet, dans notre cas, le bruit de fond était globalement bas et uniforme. De plus, l'intensité de fluorescence des spots n'est pas très élevée. Martin (2002) discute dans sa thèse de la prise en considération du bruit de fond. D'un point de vue technique, rien ne permet d'affirmer que le signal de bruit de fond des zones inter-spot et intra-spot est similaire étant donné que les acides nucléiques associés aux solutions de dépôt vont dans une certaine mesure perturber la fixation de la fluorescence à l'endroit des spots (ex : effet de spot d'intensité négative par décrochage des acides nucléiques). Sur le plan statistique, la correction d'un bruit de fond estimé localement peut introduire : 1) une source de bruit supplémentaire en surestimant la part du bruit de fond dans le cas d'un signal faible ou 2) influer de façon importante sur les ratios des valeurs d'intensité rendues très faibles par soustraction du bruit de fond. Nous avons donc travaillé directement sur les intensités de fluorescence.

Nous avons transformé les données en appliquant la méthode rlowess de la fonction « transform.madata » de la librairie MAANOVA. Cette méthode de « lowess régional » va permettre de ramener le nuage de point des ratios en fonction de l'intensité (Fig. D-30) vers sa position idéale comme cité précédemment en prenant en compte la position du spot (Ligne et Colonne)(ratio ~ intensité + ligne + colonne). Cela permet une première normalisation. Les données sont rééquilibrées entre les signaux rouge et vert, la répartition du nuage de point selon l'intensité est normalisée et les effets d'hétérogénéité locale de l'hybridation sont réduits. Le résultat obtenu pour une lame est représenté graphiquement à la figure D-31. Comme nous pouvons le voir, cette première correction est très efficace mais se limite à une correction des données intra-lame.

Pour ne pas perturber l'équilibre du nuage de points et la correction rlowess, les spots correspondant à des dépôts d'oligonucléotides fluorescents ont été supprimés des fichiers.

Les données corrigées ont ensuite été filtrées sur le nombre de données manquantes. Pour chaque accession, la proportion de données manquantes a été calculée pour chaque facteur intervenant dans le modèle d'analyse de variance globale (Modèle C-1, cf. D.4.2.6). Un nombre maximum de 30 % de données manquantes a été toléré. Pour l'expérience « stress court », les données manquantes ont été calculées pour l'écotype landais, marocain, la condition contrôle, stress 6H, stress 48H, les fluorochromes Cy3 et Cy5. Au total, les données de 555 accessions, avec plus de 30% de NA dans un des cas, ont été écartées. Pour l'expérience « stress long », le seuil de 30% a été testé pour la provenance landaise,

marocaine, la condition contrôle, stressée et les fluorochromes Cy3 et Cy5. Au final, 1481 accessions ne répondaient pas aux critères de sélection. La plus forte proportion d'accessions écartées dans l'expérience « stress long » peut s'expliquer par le plus faible nombre de lames, le plan d'expérience en boucle simple et par conséquent la plus faible redondance des données disponibles pour un échantillon. La perte des données d'une lame pour une accession a immédiatement un fort impact sur le nombre de données disponibles pour l'une des conditions ou l'un des écotypes. Cette étape permet de s'assurer d'un minimum de 13 valeurs pour le SC et 12 pour le SL afin de tester l'effet des facteurs dans l'analyse de variance.

D.3.2.2.3 Ajustement des données

Les suites d'analyse MAANOVA et LIMMA ainsi que d'autres outils d'analyse de données d'expression de microarrays disponibles en ligne qui ont été testés n'ont pas permis de prendre en compte le plan d'expérience en « loop design ». Dans le cas de MAANOVA, si les premières étapes de traitement lame par lame étaient très satisfaisantes, l'analyse inter-lame ne fonctionnait pas avec nos données. Le reste des analyses a donc été entièrement réalisé en se servant des fonctions de statistique classiques sous R.

Si l'on considère chaque dépôt comme une répétition, il est possible de représenter par des boîtes à moustache la distribution des données en fonction de différents facteurs pour chaque expérience (Fig. D-32). Dans le cas du « stress long », il persiste une hétérogénéité des distributions entre lames, entre conditions et entre fluorochromes après correction par rlowess (Fig. D-32a). On constate également une dissymétrie, la médiane étant décalée vert la valeur de premier quartile et les valeurs atypiques que l'on retrouve au-delà du troisième quartile concernent les fortes intensités. La distribution des données est donc caractérisée par une plus forte proportion de faibles valeurs. Pour le « stress court », une hétérogénéité est présente également pour les facteurs lame et condition mais également pour la provenance. Il n'y a par contre pratiquement pas d'hétérogénéité pour les fluorochromes (Fig. D-32c). Dans les deux expériences, la proportion des interquartiles Q1-Q3 et Q1-Q2 est conservé au sein du facteur étudié. Les valeurs adjacentes basses sont également conservées. L'hétérogénéité vient essentiellement d'un décalage de la distribution des valeurs qui peut être corrigé par une normalisation en appliquant un modèle d'analyse de variance global où les gènes/accessions sont considérés comme des répétitions. L'objectif de cette normalisation (Modèle C-1) est de réduire la variabilité inter-lames des données corrigées sur la variation intra-lame dans l'étape précédente. Le modèle global appliqué devrait permettre de fixer les variations observées dues au facteur lame (A), fluorochrome (D), interaction lame-fluo (AD), Provenance (P), Condition

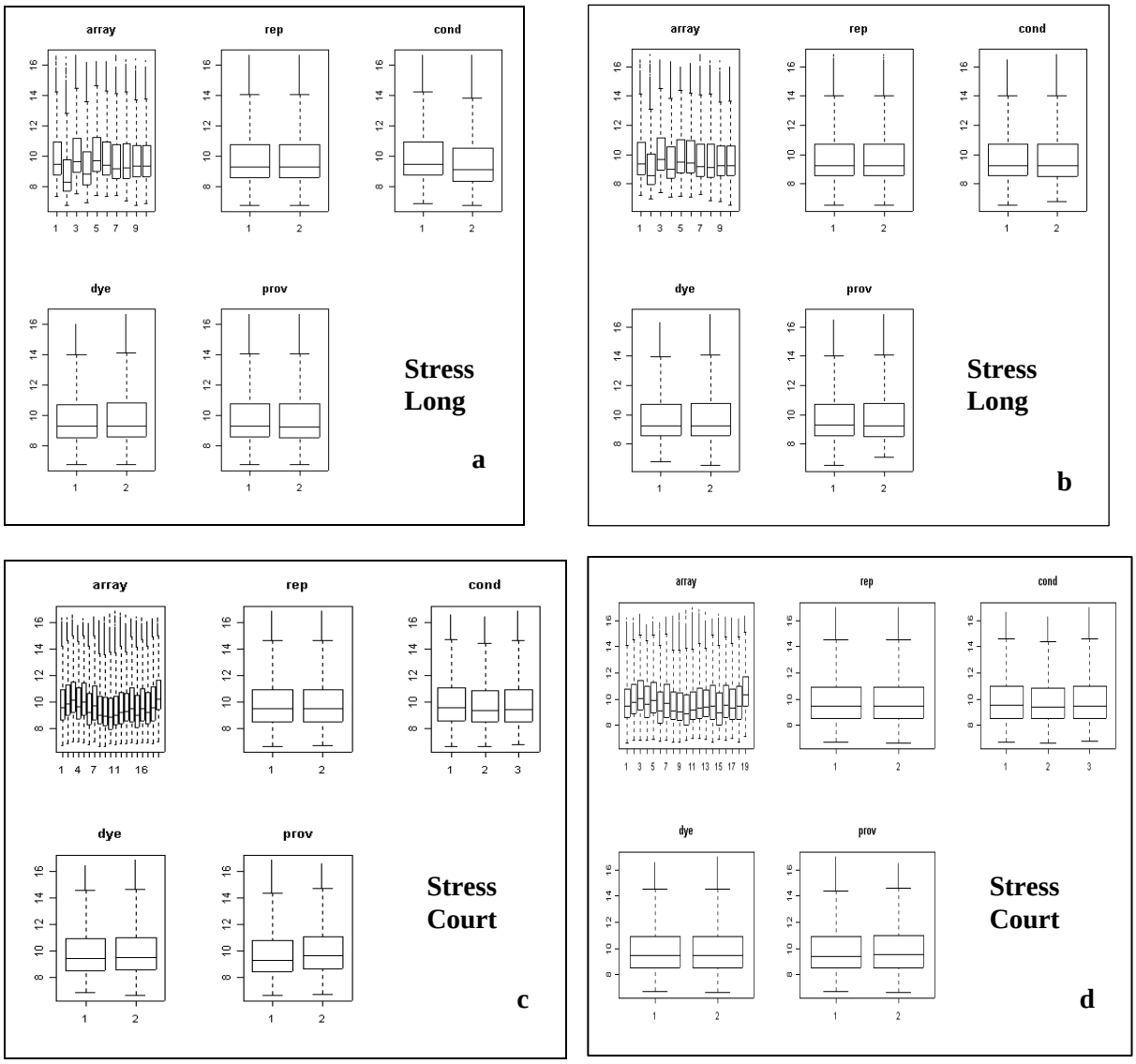


Figure D-32: Distribution des données en fonction des facteurs de l'analyse de variance globale représentée par des boîtes à moustache (par lame, par répétition, par condition, par fluorochrome et par provenance). Les données sont présentées avant (a et c) et après (b et d) l'ajustement des données par le modèle. (Facteurs : « array » = Lame, « rep » = répétition, « cond » = condition, « dye » = fluorochrome, « prov » = provenance)

Tableau D-5 : Bilan des données après correction par rlowess et ajustement par le modèle d'analyse de variance global.

	Stress Court	Stress Long
Minimum (log)	6,636	6,735
1er Quartile (log)	8,526	8,556
Médiane (log)	9,466	9,269
Moyenne (log)	9,903	9,83
3eme Quartile (log)	10,955	10,752
Maximum (log)	16,768	16,578
Nb Total Données	492632	223520
NA.	12364	5084
Nb de Lames	19	10

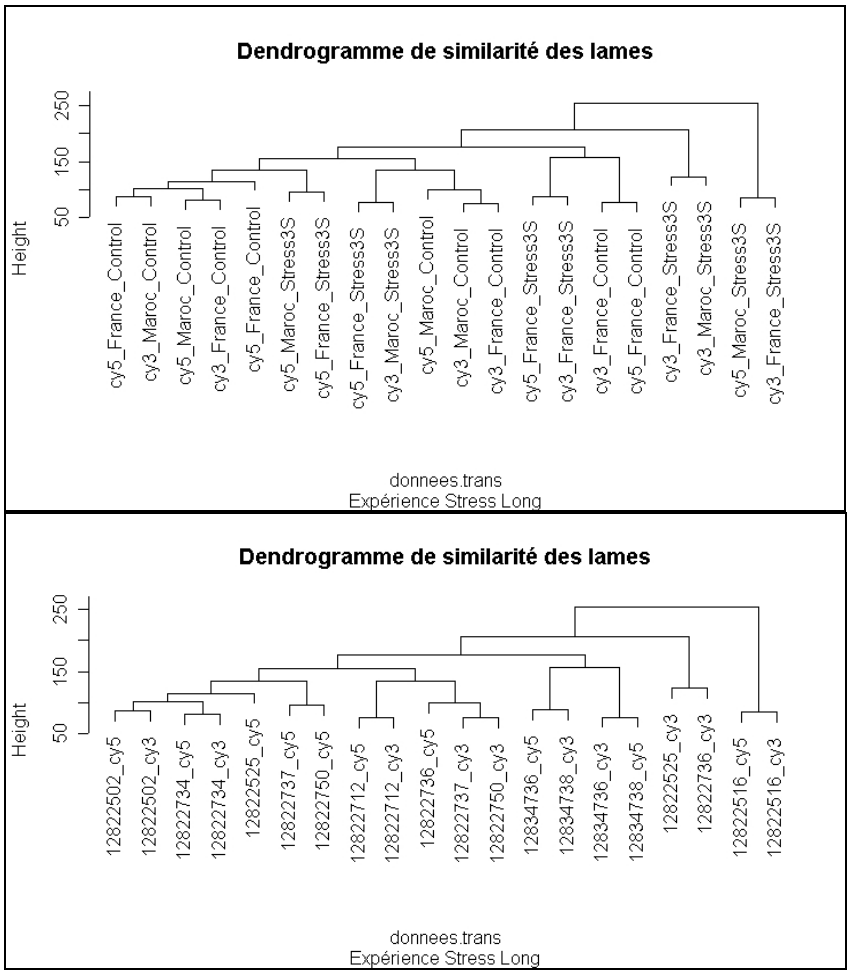


Figure D-33 :
Dendrogramme de similarité des cibles en considérant les valeurs d'intensité de fluorescence pour l'expérience « stress long ». (Fluorochrome, Provenance et condition en haut)(numéro de lame et fluorochrome en bas)

(C) et interaction prov-cond (PC). Nous avons utilisé la fonction « aov » sous R. Les valeurs ajustées ont été calculées et correspondent à la moyenne des intensités mesurées (μ) et à la variance résiduelle du spot (ε_{ijklg}). La valeur moyenne est présentée dans le résumé des données introduites dans le modèle (Tab. D-5). Les valeurs indiquées correspondent à des logarithmes de base 2 d'intensités de fluorescences corrigées. Les valeurs moyennes sont respectivement 9.903 et 9.83 pour le « stress court » et le « stress long ». L'amplitude des valeurs est plus faible dans le « stress long » que dans le stress court aussi bien pour les valeurs extrêmes (min, max) (9.843 vs 10.132) que pour 75% des données comprises dans un intervalle (Q3-Q1) de 2.196 (« stress long ») et 2.429 (« stress court »). On notera aussi entre les deux expériences que les intensités de fluorescence corrigées sont plus basses dans le « stress long ». La proportion de données manquantes est de 2.5 % dans l'expérience « stress court » contre 2.27% pour le stress long. Si l'on se rapporte à la distribution des données (Fig. D-32b et d) après normalisation par ce modèle, on constate que les distributions des valeurs par facteur répétition, condition, fluorochrome et provenance sont beaucoup plus homogènes. Une légère différence persiste entre les différentes conditions du « stress court », où le troisième quartile est légèrement plus bas pour le stress à 6 heures. Par contre, si la variabilité du facteur lame a été réduite, elle reste importante et n'a pu être fixée entièrement par le modèle. L'hétérogénéité entre lame peut provenir entre autres de la nature de la lame, du niveau global du signal, de l'hybridation localement moins bonne, de la présence de poussières ou de traces de sels, etc.

Le regroupement des données sur la similarité des cibles hybridées par une méthode de classification hiérarchique de type UPGMA aboutit aux dendrogrammes des figures D-33 et D-34. Pour l'expérience stress long (Fig. D-33), on peut remarquer que les cibles sont regroupées par paires et correspondent à une lame, excepté pour les lames 12822525,737 et 750 pour lesquelles le facteur condition a semble-t-il joué. Une certaine structuration des données semble persister dans les données corrigées et ajustées de l'expérience stress long avec un effet du facteur lame puis condition et enfin provenance. L'effet du type de fluorochrome n'est pas présent. Pour l'expérience stress court (Fig. D-34), l'effet du facteur lame reste très marqué. Les cibles sont groupées par paire d'une même lame excepté les lames 12822503 et 506. Le second niveau de regroupement également bien marqué concerne le facteur provenance avec sur le dendrogramme D-34, la gauche de l'arbre groupant l'écotype marocain et la droite l'écotype landais. Cela peut indiquer une réalité biologique, les données des deux écotypes différant au niveau transcriptionnel, ou d'un biais de l'expérience. Dans notre cas, les deux hypothèses peuvent se combiner étant donné que l'expérience stress court

Figure D-34 :
Dendrogramme de similarité des cibles en considérant les valeurs d'intensité de fluorescence pour l'expérience stress court. (Fluorochrome, Provenance et condition en haut) (numéro de lame et fluorochrome en bas)

a été réalisée en trois lots d'hybridation successifs (boucle landais, boucle marocain puis lames boucle centrale). Un effet « série d'hybridation » semble être confirmé en observant l'extrémité droite de l'arbre qui n'a pas de structuration par la provenance mais dont les lames 12822515,714 et 12834760 font partie de la dernière série.
Nous considérerons toutefois comme acceptables ces données corrigées et ajustées pour la détection des gènes différentiellement exprimés et la suite des analyses.

D.3.2.3 Analyse de l'expression différentielle

D.3.2.3.1 Choix des paramètres statistiques

L'analyse de variance du modèle décrit par le modèle C-2 (cf. C.4.2.6) permet de tester l'existence d'un effet provenance (P), d'un effet condition (C) ou d'un effet d'interaction entre provenance et condition (I). Pour chaque expérience, l'analyse de variance a été effectuée gène par gène et les valeurs de probabilité associées aux tests ont été corrigées par la méthode du FDR. Les valeurs des probabilités et les r^2 associés aux tests sont indiqués dans les tableaux Ax-3 et Ax-4 en annexe. Le pourcentage de variabilité expliqué par le modèle (r^2 total) et la part de la variation expliquée par les facteurs cités ci-dessus sont également indiqués. De plus, pour chaque gène, les valeurs moyennes des valeurs corrigées sont indiquées pour chaque facteur. Lorsqu'un effet significatif est détecté, nous avons déterminé les ratios d'expression qui sont calculés sur les valeurs ajustées de fluorescence.
La sélection des gènes différentiellement exprimés se fait d'une part sur la base des probabilités associées au test et d'autre part sur la significativité du modèle représentée par les r^2. Les seuils limites de tolérance pour ces paramètres ont été déterminés en fonction du lot de données étudiées. Notre objectif était de conserver une liste de gènes différentiellement exprimés n'excédant pas un maximum de 600 entrées et constituée des candidats les plus significativement liés aux facteurs provenance et/ou condition.
Considérant nos données, une exigence commune aux deux expériences a été de rejeter les gènes pour lesquels la part de variation totale expliquée par les facteurs provenance, condition et interaction (r^2 PCI) du modèle était inférieure à 40 % ou qu'au minimum un des facteurs d'intérêt soit responsable d'au moins un tiers (33%) de la variation observée. Cette approche a permis de ne pas prendre en considération les gènes pour lesquels les facteurs ne sont responsables que pour une faible part dans la régulation de leur expression.
Ce premier filtre appliqué, le choix des gènes à conserver comme étant significativement régulés entre provenances ou conditions s'est fait en appliquant un seuil d'acceptation des

probabilités associées au test pour ces facteurs de 0.01. Pour être retenu, le gène doit avoir une p-value inférieure au seuil fixé pour au moins un des facteurs (effet Provenance ou Condition ou Interaction). Cela signifie que nous acceptons une erreur de type 1 de 1 chance sur 100 de détecter un faux positif. Il faut se rappeler que les p-value sur lesquelles le filtre est appliqué ont préalablement été corrigée par la méthode FDR pour tenir compte de la répétition des tests.

D.3.2.3.2 Quels sont les gènes différentiellement exprimés ?

Sur ces critères, nous avons obtenu une liste de 458 gènes dans l'expérience « stress court » (SC) et 612 gènes pour le « stress long » (SL). L'ensemble de ces gènes et des paramètres statistiques associés est accessible en annexe dans les tableaux Ax-3 et Ax-4. Ces tableaux ainsi que les profils d'expression de l'ensemble des accessions, sont également accessibles en ligne par l'URL http://www.pierroton.inra.fr/biogeco/genetique/theses.html.

Le tableau D-6 synthétise les sources de variations significatives détectées qui sont à l'origine de la sélection des gènes. On constate que pour le SC (« stress court »), la majorité des gènes retenus présentent un effet provenance seul tandis que pour le SL (« stress long ») les gènes détectés présentent à plus de 80% un effet condition seul. Les autres effets majeurs détectés dans le SC sont l'effet condition seul qui concerne 15% des gènes sélectionnés et l'effet cumulé de la provenance et de la condition représentant 19% des gènes. Ces deux effets se partagent respectivement 4 et 8 % des gènes statistiquement significatifs de l'expérience SL. Nous remarquerons que l'effet d'interaction entre une provenance et une condition est mis en évidence de façon très mineure dans les deux types d'expériences. L'interaction pure ne concerne qu'un seul gène (BX252971), annoté « putative protein » (Tab. Ax-3), qui est détecté dans le SC. Ces données indiquent que les gènes que nous étudierons dans l'expérience SC seront majoritairement des gènes dont l'expression est différente entre les deux provenances indépendamment de la condition. Leur expression constitutive différente entre les deux écotypes peut toutefois participer à l'établissement d'un avantage constitutif dans la réponse de la plante au stress hydrique. Dans le cas du SL, les gènes étudiés seront probablement des gènes dont la régulation de l'expression est une réponse au stress hydrique.

Les données du tableau D-7 nous renseignent sur le degré de variation expliquée par les facteurs provenance (P), condition (C) et interaction PC (I) du modèle pour l'ensemble des gènes retenus. Ces trois facteurs peuvent expliquer jusqu'à près de 79% (SC) et 98% (SL) de la variation observée pour certains gènes. En moyenne, ces facteurs expliquent 45% (SC) et 56% (SL) de la variabilité avec un minimum proche de 34%.

Tableau D-6 : Distribution des gènes différentiellement exprimés en fonction des facteurs significatifs.

	Stress Court		*Stress Long*	
effet significatif	Nb gènes	%	Nb gènes	%
Condition (C)	70	15	504	82
Condition+Interaction (CI)	3	1	15	2
Interaction (I)	1	0,2	0	0
Provenance (P)	292	64	25	4
Provenance+Interaction (PI)	4	1	9	1
Provenance+Condition (PC)	87	19	52	8
Provenance+Condition+Interaction (PCI)	1	0,2	7	1

Tableau D-7 : Valeurs minimales, maximales, médiane et moyenne du r2 PCI.
(niveau de variation expliqué par les facteurs provenance (P), condition (C) et Interaction (I) du modèle)

r^2 PCI	*Stress Court*	*Stress Long*
Minimum	33,8	34,3
Maximum	78,8	97,8
Médiane	42	53,8
Moyenne	44,6	56,5

Tableau D-8 : Fréquences des ratios d'expression des gènes pour les facteurs statistiquement significatifs.

Fréquences de ratios	*Stress Court*		*Stress Long*	
Classes	*Facteur provenance*	*Facteur Condition*	*Facteur provenance*	*Facteur Condition*
ratio ≤ 1,25	3	0	1	20
1,25 < ratio ≤ 1,5	64	5	24	220
1,5 < ratio ≤ 2	256	64	36	224
2 < ratio ≤ 3	58	68	11	82
3 < ratio ≤ 4	3	14	3	20
4 < ratio ≤ 5	0	6	0	4
ou plus...	0	4	0	4

Tableau D-9 : Distribution des ratios d'expression des gènes pour les facteurs statistiquement significatifs.

ratios d'expression	*Stress Court*		*Stress Long*	
	Facteur provenance	*Facteur Condition*	*Facteur provenance*	*Facteur Condition*
Minimum	1,03	1,4	1,23	1,15
Maximum	3,69	6,62	3,39	7,56
Médiane	1,71	2,07	1,65	1,56
Moyenne	1,76	2,36	1,74	1,75

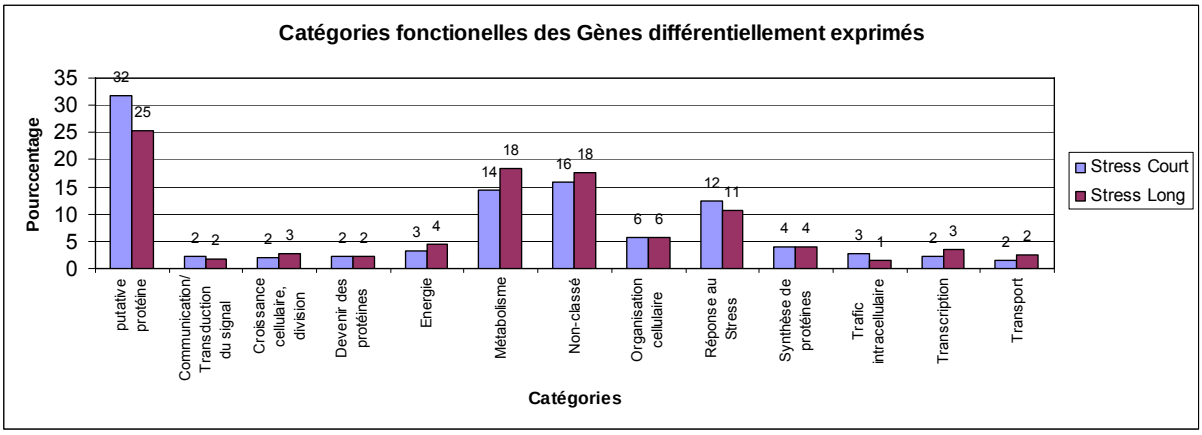


Figure D-35 : Distribution des gènes détectés en fonction des catégories fonctionnelles pour l'expérience stress court et l'expérience stress long.

Une façon simpliste - couramment employée il y a quelques années dans la littérature – pour sélectionner les gènes différentiellement exprimés est de conserver uniquement les gènes dont le ratio d'expression est supérieur à 2. Dans notre cas, cela concerne relativement peu de gènes (Tab. D-8). Les ratios observés (Tab. D-9) sont globalement plus élevés pour les données du SC où les différences d'expression sont en moyenne proches de 2.36 pour un effet condition mais pour le facteur provenance et les ratios de l'expérience SL les ratios sont de l'ordre de 1.75. Les ratios les plus élevés sont mesurés pour l'effet condition avec 7.56 pour un gène de l'expérience SL et 6.62 pour un gène de l'expérience SC. A l'opposé, un effet provenance significatif a été détecté en SC pour un gène présentant un ratio de 1.03. Même si le test est significatif, une différence d'expression si faible doit attirer l'attention sur toute interprétation biologique. Ces données ont cependant été conservées mais doivent être interprétées avec le recul nécessaire. Dans ce cas précis le gène (BX676790), annoté « putative protéine » présente un effet significatif à la fois pour la provenance (ratio provenance = 1.03) et pour la condition (ratio conditions = 2.53). Le gène présente donc un effet condition marqué cumulé à un effet provenance de très faible amplitude mais qui peut souligner une différence d'expression entre les écotypes pour un gène impliqué dans la réponse au stress hydrique. Les ratios inférieurs à 1.25 ne représentent pas une grande proportion des gènes détectés (Tab. D-9) mais lorsqu'un effet condition est détecté dans l'expérience SL, dans 42% des cas le ratio est inférieur à 1.5. Si ces valeurs semblent faibles, elles peuvent toutefois être dues à une normalisation poussée qui peut avoir pour effet d'écraser les variations d'expression. Un ratio de faible amplitude ne devra pas être considéré de la même manière dans l'expérience SC et l'expérience SL.

D.3.2.3.3 Quelles sont les fonctions des gènes détectés ?

Les fonctions des gènes différentiellement exprimés que nous citerons dans ce chapitre sont issues des annotations fines et classifications fonctionnelles réalisées sur les séquences de racines décrites à la section « D.2.3 Annotation des séquences » pour les séquences déposées dont le clone est originaire de la banque « racine ». La fonction des séquences originaires de la banque de « xylème » a été déterminée en se servant des données d'assemblage de l'ensemble des séquences de pin maritime (projet « Pinus pinaster total » (StackPack 191)) et des annotations du TIGR (PGI, release 6.0). Lorsque l'EST de xylème appartenait à un consensus du projet StackPack 191 constitué également d'EST de racines, l'annotation des EST de racine a été héritée. Dans le cas contraire, une nouvelle annotation a été attribuée à partir des annotations du TIGR

La répartition des gènes différentiellement exprimés dans les catégories fonctionnelles est indiquée pour les deux expériences dans la figure D-35. La proportion de gènes n'ayant pas d'homologie avec des séquences connues est plus élevée dans le SC que dans le SL. Les classes majoritaires correspondent dans les deux cas au métabolisme, aux gènes non classés et à la réponse au stress. Les gènes non classés et du métabolisme occupent une part plus importante dans le SL que dans le SC. La classe correspondant au trafic intracellulaire est sous représentée dans l'expérience SL. L'affectation d'un gène à une classe est toutefois difficile et restrictive. Beaucoup de gènes ont des fonctions qui peuvent intervenir au croisement de plusieurs voies métaboliques. Certains membres de familles multigéniques sont spécifiques d'une réponse au stress mais l'indication de leur fonction ne permet pas forcément de les affecter à la catégorie réponse au stress.
Le tableau Ax-5 disponible en annexe regroupe l'ensemble des fonctions des gènes par catégorie fonctionnelle. Un code à une lettre indique également quel effet significatif est à l'origine de la sélection du gène.

D.3.2.3.3.1 Communication / Transduction du signal

Dans cette catégorie, nous trouvons plusieurs gènes codant pour des récepteurs protéine kinase ou des précurseurs de ces récepteurs et une protéine avec domaine à sept hélices transmembranaires. Des gènes intervenant dans la transduction du signal comme des phospholipases, l'IMP et des calmodulines sont également différentiellement exprimés.
La calmoduline (CaM) est la principale cible du Ca^{2+} dans les cellules eucaryotes (Bouche *et al.* 2005). Le calcium est, comme nous l'avons vu, un des seconds messagers majeurs dans les cascades de régulation. Bouche *et al.* (2005) rapportent la mise en évidence et le rôle des senseurs au calcium cytosolique dans la réponse aux stress biotiques et abiotiques. D'autres protéines (Calmodulin Binding protein) vont pouvoir se lier à la calmoduline et être activées. Kudla (1999) montre que la calcineurine (Ca^{2+}/calmoduline-dependent protein phosphatase) est différentiellement régulée par le stress. Il est intéressant de constater que la calmoduline comme le recepteur proteine kinase CLAVATA1 sont détectés dans les deux expériences mais pour les effets provenance (SC) vs condition (SL).

D.3.2.3.3.2 Croissance cellulaire, division

Au cours du stress, des gènes impliqués dans l'élongation cellulaire sont régulés. L'expression du gène codant pour la protéine Diminuto est par exemple inhibée par le stress. On pourra également citer la prohibitin dont l'expression est inhibée au cours du stress. Il

s'agit d'une protéine associée à une activité antiproliférative découverte lors de l'étude de la sénescence sur des cellules de mammifères. Une régulation négative à également été constatée dans une étude sur l'orge (Ozturk *et al.* 2002). Des gènes homologues ont depuis été identifiés chez *Arabidopsis* et le tabac (Snedden and Fromm 1997). Cet enzyme agirait en inhibant la synthèse d'ADN.

Nous avons aussi détecté des gènes en rapport avec la composition des parois cellulaires tel que les gènes codant la protéine MERI-5 à activité xyloglucosyl transferase (EC 2.4.1.207) qui sont régulés négativement avec le stress et un gène surexprimé, homologue à l'expansine, détecté dans l'expérience SL. Nous avons discuté dans la partie bibliographie du rôle de ces gènes dans la réponse au stress hydrique.

Deux gènes, codant des « Fasciclin-like arabinogalactan protein » présentent une diminution de leur expression en SL. Johnson *et al.* (2003) discutent du rôle possible de cette famille de protéines dans l'adhésion cellulaire. Les auteurs soulignent l'implication de ces gènes dans la réponse au stress et montrent que les Fasciclines sont régulées au cours du développement et par l'ABA qui diminue l'expression de leurs gènes (Johnson *et al.* 2003).

D.3.2.3.3.3 Devenir des protéines

Parmi les gènes détectés, nous retrouvons des gènes qui concernent la dégradation des protéines avec des représentants du complexe protéosomique codant les sous-unités 26S ou 20S. Leur expression est stimulée lors du SL. Une autre série de gène concerne la modification post-traductionnelle des peptides avec la « signal peptidase I » (EC 3.4.21.89) qui clive les parties « N-teminal signal » ou peptide signal des protéines sécrétées et périplasmiques, l'endopeptidase Clp (EC 3.4.21.92), une alpha galactosyltransferase ou la thimet oligopeptidase (EC 3.4.24.15) qui va cliver de façon préférentielle les liaisons avec des résidus hydrophobiques.

La régulation des gènes impliqués dans les mécanismes de dégradation des produits cellulaires est représentée par un gène de « cathepsin B-like cysteine proteinase » mis en évidence dans le SC avec un effet provenance et condition. Son expression est plus forte pour l'écotype marocain (EM) et augmente avec le stress. La localisation cellulaire de la cathepsine est lysosomale.

L'augmentation des activités d'hydrolyse ou de peroxydation peut permettre à la cellule de faire face aux altérations des composants cellulaires au cours du stress.

D.3.2.3.3.4 Energie

La mise en évidence de la PEP carboxylase –connue pour son rôle dans la photosynthèse- surexprimée lors du stress long peut paraître étrange sur une étude de transcrits racinaires. Toutefois cet enzyme catalyse la réaction du phosphoenolpyruvate à l'oxaloacétate et n'a pas une localisation exclusive dans les parties aériennes. Le PEP et l'oxaloacétate font partie de la voie de la glycolyse et du cycle de Krebs. La régulation de cet enzyme aura sans doute un effet direct sur la voie majeure de production d'énergie de la cellule en évitant le passage par le pyruvate et l'acetyl-CoA qui sont des points clés de nombreuses autres voies métaboliques comme la synthèse des acides gras ou les caroténoïdes. Des gènes codant pour des pyruvate kinase (PEP <-> pyruvate) et pyruvate décarboxylase sont sous exprimés lors du stress ce qui confirme une inhibition de la voie classique du phosphoenolpyruvate vers le cycle de Krebs au profit d'une voie PEP vers oxaloacétate.
Nous avons mis aussi en évidence des gènes codant des protéines impliquées dans le transport des électrons et les réactions d'oxydo-réduction. Nous citerons des NADPH dehydrogénases régulées négativement, des NADPH-quinone reductase régulées positivement, une flavoproteine-ubiquinone et différentes thioredoxines dont le type de régulation et l'effet détecté est variable.
Des plantacyanines comme des gènes homologues à la « basic blue protein » ou à une phytocyanine sont différentiellement exprimés entre provenances ou conditions selon les gènes. L'accession BX678708 a une expression significative en SL à la fois pour l'effet condition et pour l'effet d'interaction. Le profil d'expression correspond à un maintien de l'expression au cours du stress pour l'écotype marocain (EF) dont l'expression est plus élevée que l'EF (écotype français) en condition contrôle mais plus basse en condition de stress. Ce gène semble régulé négativement pour l'EM et non régulé pour l'EF.

D.3.2.3.3.5 Métabolisme

Avec la réponse au stress et les non-classés, cette catégorie regroupe le plus grand nombre de gènes. L'interprétation de la régulation des gènes de cette classe et leur implication dans la réponse au stress est difficile. Le nombre et le manque de spécificité des voies métaboliques susceptibles d'être touchées ne permettent généralement pas de tirer des conclusions définitives.

Parmi les gènes surexprimés, on pourra citer la Glutathione transférase pour laquelle 4 accessions sur 5 indiquent une accumulation des transcrits lors du stress. Il y a aussi la Flavonol 3-O-glucosyltransférase qui intervient dans la voie des flavonoïdes, la 4-

hydroxyphenylpyruvate dioxygenase (EC 1.13.11.27) - qui agit sur le precurseur de synthèse de la tyrosine et peut aboutir à la synthèse de fumarate, acetoacetate ou plastoquinone -, une alpha galactosidase 2 ou une beta fructofuranosidase qui peuvent avoir une action sur la paroi cellulaire.

Nous retrouvons également des gènes codant des enzymes pouvant intervenir dans le métabolisme de l'éthylène au niveau de la S-adenosylmethionine (SAM). L'accession (BE643924) homologue à la SAM synthase a une expression diminuée lors du SL. En SC, l'analyse statistique indique que l'expression est plus élevée pour l'EM que pour l'EF. L'expression de la SAM synthase et de la SAM décarboxylase – qui peut favoriser la synthèse de spermine, spermidine ou putrescine - représentée par l'accession (BG039488) est significativement plus élevée chez les marocain dans l'expérience SC. Deux autres accessions (AL751120 et CR393742) codent pour des methyltransferases utilisant la SAM et sont surexprimées lors du stress.

L'expression des gènes codant pour la Caffeate O-methyltransferase et la Caffeoyl-CoA O-methyltransferase est diminuée lors du SL. Ces deux enzymes interviennent au niveau de la biosynthèse de la lignine et ont donc un impact sur les propriétés chimiques et physiques des parois cellulaires. Trois autres gènes en rapport avec cette voie codent pour la Phenylalanine ammonia-lyase (PAL)(EC 4.3.1.5). Deux d'entre eux ont des régulations opposées au cours du SL, le dernier est détecté dans l'expérience SC et a une expression plus élevée chez les landais. La diminution de la COMT confirme les résultats des études précédentes de Dubos (2001). Les expressions opposées des deux transcrits codant la PAL pourraient indiquer des gènes différents intervenant dans des voies métaboliques distinctes. La surexpression, montrée par Dubos en RT-PCR, associée à la régulation d'autres gènes de la voie de biosynthèse des monolignol suggérait une modification possible dans la composition de la lignine au court du stress.

Dans l'expérience SC, des méthallothionéines sont différentiellement exprimées entre provenances et entre conditions. Les transcrits sont plus abondant pour l'EM et sont accumulés au cours du stress avec cependant une légère diminution de leur quantité à 6 heures pour l'EF.

Nous mettrons enfin l'accent sur des gènes impliqués dans la synthèse des terpènes. Tout d'abord, les gènes codant pour des pinènes synthases ont été identifiés. Leur expression est réduite lors du stress. Dans le stress court l'effet provenance montre un niveau plus élevé chez l'EM. Cela correspond à ce qui a pu être observé sur les parties aériennes lors du dosage de l'alpha-pinène. Nous avons aussi l'expression d'un gène codant une géranyl diphosphate

synthase qui diminue avec le stress et qui dans l'expérience SC est significativement plus élevées dans l'EM.

D.3.2.3.3.6 Organisation cellulaire

Cette catégorie regroupe les gènes intervenant dans la structure de la cellule, de la paroi et du noyau.

Nous trouvons l'extensine dont expression est réduite lors du stress et qu'il ne faut pas confondre avec l'expansin citée dans la catégorie croissance cellulaire. L'extensine est un constituant de la paroi cellulaire au même titre que les microfibrilles de cellulose, les xyloglucanes ou les pectines. Concernant la cellulose, deux gènes codant des celluloses synthases sont régulés négativement de même que l'expression d'un gène homologue à une endoglucanase ou cellulase. Plusieurs gènes homologues à un précurseur de Glucan endo-1,3-beta-D-glucosidase ont leur expression significativement diminuée en SL tandis que les transcrits du gène (AL751122), homologue à une Glucan endo-1,3-beta-D-glucosidase, sont augmentés. En SC, un des deux gènes est commun à ceux détectés en SL et présente un effet provenance. Le gène (AL751122) codant un précurseur de l'enzyme a une expression supérieure dans l'EF tandis que celui homologue à l'enzyme (BX249875) a une expression supérieure dans l'EM. Cela peut indiquer la présence de plusieurs gènes régulés de façon différente pour cet enzyme. Des régulations différentes sont également détectées en SC pour des gènes codant des pectate lyases. Deux de ces gènes présentent un effet provenance seul et sont plus exprimés dans l'EF. Un autre gène est détecté avec un effet provenance ainsi qu'un effet condition. Cette fois l'expression est plus forte pour l'EM, mais le profil d'expression montre une régulation différente entre les deux provenances. L'expression en condition contrôle est plus forte chez l'EM, à 6 heures de stress l'expression diminue chez l'EF et augmente chez l'EM puis à 48 heures les niveaux d'expression sont identiques entre les deux provenances. Les enzymes et constituants de la paroi, bien que classés ici dans l'organisation cellulaire, peuvent avoir comme nous l'avons vu dans la partie bibliographie une grande influence sur la résistance au stress hydrique en modifiant les propriétés des parois.

Le cytosquelette est représenté ici par des gènes homologues à la tubuline et à l'actine. Leur expression est ici toujours régulée négativement lors du stress.

Nous avons aussi une série de gènes codant des « Ubiquitin-conjugating enzyme » détectés lors du SC avec un effet provenance et dont le niveau de transcrits est plus élevé pour l'EM. Deux gènes homologues à l'Ubiquitine sont sous exprimés en stress long et un gène

(BX250835) ayant des homologies avec une Ubiquitine carboxyl-terminal hydrolase family 2 est surexprimé lors du stress.

Au niveau nucléaire, des gènes homologues à des histones H2A, H2B et H4 sont détectés avec des effets condition en SL et provenance ou provenance et condition en SC. Une surexpression est également mise en évidence pour un gène homologue à une protéine « hétérochromatin-like ». Tous les gènes codant les histones ont dans notre cas leur expression diminuée au cours du stress excepté pour l'accession (BX682172) homologue à une histone H2A qui est surexprimée. Dubos (2001) avait étudié l'expression de 7 histones différentes au court du stress dans les aiguilles, les tiges et les racines. Il avait constaté également une diminution de l'expression d'un transcrit homologue à une histone H2B dans les racines stressées et deux types d'histone H2A montraient des régulations opposées. La H2A de type 1 était surexprimée dans les racines lors du stress tandis que la H2A de type 2 était sous exprimée. L'expression de l'histone H4 était diminuée en stress. Les histones et les gènes intervenant dans la structure de la chromatine peuvent jouer un rôle dans la régulation épigénétique lors du stress. Le niveau d'enroulement de la chromatine va en effet pouvoir déterminer l'accessibilité et par conséquent la possible d'expression des gènes localisés dans la région concernée.

D.3.2.3.3.7 Réponse au Stress

L'identification et la régulation des gènes de cette catégorie est très instructive. Nous retrouvons ici de nombreux gènes connus pour leur implication dans la réponse aux stress. Plusieurs gènes présentent des effets provenance et condition simultanés voire des effets d'interactions.

Plusieurs accessions correspondant à des protéines de choc thermique (HSP) habituellement identifiées lors de stress hydriques sont surexprimées à la fois lors du SC et du SL. Notre séquence de HSP déposée comme contrôle positif est très fortement régulée (x 6.62) et comme l'accession (BX681237) présente en SC un effet provenance significatif avec une expression plus élevée pour l'EM quelque soit la condition. L'accession (BX252065) est homologue à une HSP de classe II de 17.9 kDa. Son expression est plus élevée chez l'EM en SC comme en SL, et lors du SL, un effet provenance et interaction est mis en évidence. Les plants de l'EM conservent une forte expression similaire à celle de la condition contrôle tandis que les plants de l'EF voient leur expression diminuer avec le stress. Une régulation différente est observée pour un autre gène (BX252366) codant une HSP de classe I. Les transcrits de cette HSP sont exprimés de façon identique dans les deux provenances avec une

surexpression lors du SL. Par contre, en SC, l'accumulation des transcrits avec le stress est beaucoup plus rapide chez les marocains, ce qui est à l'origine de l'effet provenance détecté. Un comportement proche est observé pour un gène (BX250781) codant une HSP de faible poids moléculaire dont l'expression est régulée positivement dans les deux provenances au cours du SC comme du SL, mais avec une induction plus forte pour l'EM.

Des gènes de réponse aux hormones du stress sont également différentiellement exprimés. Un gène codant une protéine membranaire induite par l'ABA (BX679279) est surexprimé lors du stress. En SC, l'induction est plus forte dans l'EM. Des gènes homologues à des « Abscissic-stress ripening protein » sont présents. L'accession (BX666054) est surexprimée lors du SL. En SC, un effet provenance est significatif pour les accessions (BX682324 et BX251094) qui ont un niveau d'expression supérieur pour l'EM. Un gène codant pour une protéine de réponse à l'éthylène (CR394461) est surexprimé en SC. La quantité de transcrits est supérieure chez les landais mais les marocains induisent très fortement l'expression à 48 heures diminuant l'écart d'expression en situation de stress. Le gène (BX251044), impliqué dans la synthèse d'éthylène, code pour une ACC oxydase. Il est légèrement surexprimé lors du SL.

Les gènes codant des LEA ont des régulations différentes. Les accessions (BX681356 et BX253089) présentent une augmentation de leur expression lors du SL et également lors du SC pour (BX253089) qui de plus est accumulé plus fortement chez l'EM. A l'opposé, les accessions (BX678224 et BX677964) qui ont des profils similaires sont réprimées lors du SC. La diminution du niveau de transcrit est brutale dès 6 heures de stress chez l'EF pour l'accession (BX677964) et plus progressive chez l'EM. Un effet provenance est également présent pour les accessions (BX666039, BX681471 et BX678224) pour lesquelles le niveau de transcrit est supérieur pour l'EF. Le cas du gène (BX678224) identifié en SL est plus particulier car il présente une forte interaction en plus de l'effet provenance. Le gène est surexprimé pour l'EF et sous exprimé pour l'EM, révélant des stratégies de réponse différentes au niveau moléculaire.

D'autres gènes connus comme étant reliés à la réponse au stress sont détectés. Nous trouvons des gènes homologues à des protéines induites par un stress aluminium, à des peptides antimicrobiens, à des Chitinases, des catalases, une protéine chaperonne pour la SOD et à une Déhydrine. Leur expression augmente lors du stress. La Déhydrine représentée par l'accession (BX249564) est détectée à la fois dans l'expérience SC et l'expérience SL. Son expression est fortement induite avec des ratios de 3.21 (SC) et 2.72 (SL).

Par contre, nous trouvons aussi des gènes dont l'expression est diminuée par le stress. Cela concerne par exemple les Arabinogalactanes ou les Défensines dont la régulation négative a été mise en évidence dans les deux expériences. C'est également le cas de l'accession (BX681136) homologue à une protéine induite par la déshydratation. Son expression diminue fortement lors du SC (ratio de 4.66) et du SL (ratio 2.70). Un effet provenance est également présent dans les deux cas mais le niveau est plus élevé chez l'EF en SL vs EM en SC. En fait, l'étude du profil d'expression en stress court montre que le niveau plus élevé chez l'EM tient à une diminution moins rapide du niveau de transcrit dans l'EM. Nous reparlerons de ce gène lors de l'étude par PCR quantitative.

La réduction de l'expression concerne aussi un gène codant pour une protéine de résistance aux pathogènes détectée dans les deux expériences. L'accession (BX678491) également homologue à une protéine reliée à la réponse aux pathogènes (PR-5) a une diminution de son expression en SC et un niveau plus élevé chez l'EM. Par contre les accessions (BX680166 et BX681832) détectées respectivement dans les expériences SC et SL sont surexprimées au cours du stress. La Snakin-1 est un peptide conférant une résistance aux pathogènes qui à été mis en évidence chez la pomme de terre (Segura *et al.* 1999). Les auteurs n'avaient pas observé d'induction par les stress abiotiques appliqués. Dans notre cas, l'expression de l'accession (BX253632) homologue à ce peptide diminue significativement lors du stress.

Nous citerons pour terminer les nombreux gènes homologues à un précurseur de la peroxydase (EC 1.11.1.7) qui ont été détectés dans les deux expériences. Les accessions qui ont un effet condition significatif montrent une diminution de l'expression au cours du stress, excepté pour l'accession (BX680225). Les effets provenance détectés correspondent à un niveau des transcrits supérieur pour l'EM. Certaines accessions ont été détectées dans les deux types de stress mais d'autres ne sont mises en évidences que dans une des deux expériences.

D.3.2.3.3.8 Synthèse de protéines

Les gènes différentiellement exprimés de cette classe sont des gènes intervenant dans la synthèse des sous-unités ribosomales. Les sous-unités concernées sont la 60S, la 40S, la 30S et la 25S. Lorsqu'un effet condition est présent, l'expression est inhibée lors du stress, excepté pour les gènes codant la sous-unité 40S-S8, 30S, 26S et 25S pour lesquels le stress induit l'expression des transcrits. Un effet provenance est mis en évidence en SC avec un niveau de transcrit plus élevé chez l'EM, excepté pour les sous unités 60S-L41, 30S et 40S. La régulation des gènes des protéines ribosomales a été observée dans d'autres études

(Kawasaki *et al.* 2001; Ozturk *et al.* 2002) dont celle de Buchanan *et al.* (2005) sur la réponse du transcriptome du sorgho à des stress salins, osmotiques (PEG) et à l'application d'ABA. Les auteurs montrent une forte régulation positive des transcrits de la protéine ribosomale 60S L10 dans les racines en présence de PEG. La régulation est donc opposée à celle observée dans notre étude. Nous trouvons également le facteur d'élongation EF-1 gamma détecté à la fois dans le SL où il est surexprimé lors du stress et dans le SC où un double effet provenance et condition est détecté. Pour le SC, les transcrits sont plus exprimés dans l'EM mais les accessions (BX682634) et (BX252622) sont respectivement sous et surexprimés lors du stress. Le facteur EF-1 est un complexe formé de plusieurs sous-unités dont la sous-unité gamma. Il intervient dans les mécanismes de traduction des ARNm en protéines. Une différence d'expression entre les écotypes et une régulation positive (BX252622) lors du stress suggère un rôle possible dans les stratégies différentes de régulation de la synthèse des protéines de réponse au stress.

D.3.2.3.3.9 Trafic intracellulaire

Au niveau de l'appareil de Golgi où les protéines subissent des glycosylations, le transport d'une citerne à une autre est assuré par des vésicules recouvertes d'un manteau dont la sous-unité est le coatomère (Maillet 1995). Des gènes codant pour ces coatomères ont été mis en évidence dans l'expérience SC avec un effet provenance. Les transcrits de la sous-unité alpha sont plus abondant dans l'EM tandis que ceux de la sous-unité beta sont plus abondant chez l'EF. Dans les données du SL, un effet provenance et interaction est détecté pour la sous-unité alpha. La quantité de transcrits augmente légèrement lors du stress chez l'EM alors qu'elle chute chez l'EF. Les autres fonctions identifiées concernent le transport nucléaire (Nuclear transport factor 2 (NTF-2)), le transport des protéines (chaperonine-60 ss unité beta, Sec22, Sec13p) ou le cytosquelette (profilin, myosin, kinesin).

D.3.2.3.3.10 Transcription

Par rapport aux autres classes, si un effet provenance seul a permis de détecter quelques gènes, un effet condition ou provenance plus condition concerne la plus grande partie des gènes. Nous retrouvons les facteurs de transcription évoqués dans la synthèse bibliographique. Il s'agit par exemple des protéines en doigt de zinc, dont l'expression est induite lors du stress (BX681072, BX253595, BX682945) ou réprimée (BX251702). L'expression de BX679178 est caractérisée par une régulation négative lors du stress pour l'EF et une régulation positive pour l'EM. Le profil d'expression de l'accession (BX253595)

correspond à un niveau global des transcrits plus élevé pour l'EM et une induction au cours du SC et du SL, excepté une légère diminution à 6 heures de stress. Nous avons également le facteur Rof1 (réprimé lors du stress), HBp-1b (induction), homéobox leucine zipper (induction), ABI5 (répression), RING-finger 26 (induction) et G-box binding (induction). En outre, nous avons un gène codant un facteur EREBP (BX680424) dont l'expression est inhibée au cours du stress. Dans l'étude de Buchanan *et al.* (2005), un facteur EREBP présentait le même comportement au niveau des racines exposées au PEG.
Ces facteurs de transcription sont régulés aussi bien dans le SC que dans le SL et présentent pour certains des profils d'expression ou de répression transitoire à 6 heures de stress.

D.3.2.3.3.11 Transport

Dans cette classe nous retrouvons des protéines membranaires. Cependant l'incertitude de l'annotation peut amener à y placer des protéines dont le rôle n'est pas le transport mais la réception des signaux. Pour ce qui est des fonctions, nous avons des gènes codant pour des protéines du tonoplaste (BX679636, BX681365), de la coiffe racinaire (BX679628), une protéine mitochondriale de transport de l'oxaloacetate (BX252185) et une protéine de la membrane plasmique à rôle d'aquaporine (BX249892). Les aquaporines sont généralement mises en évidence lors de la réponse au stress hydrique. Dans notre étude, seulement deux accessions (BX254875, BX249892) correspondant à des aquaporines ont été identifiées et voient leur expression réprimée au court du stress. Cependant, l'accession (BX254875) est intéressante car elle présente un effet provenance. Son expression est plus élevée dans l'EM que dans l'EF ce qui peut donner un avantage à l'EM pour la réponse au stress hydrique. Les gènes identifiés codant pour des transporteurs d'acides aminés (BX667565) ou des peptides (CR394220) sont fortement induits. C'est également le cas d'un transporteur membranaire (BX680183) et d'une pompe à protons (AL750557). Le transport actif d'acides aminés peut par accumulation participer à la régulation de l'osmolarité cellulaire et permettre répondre efficacement à un stress osmotique. Le profil d'expression en SC pour le gène codant l'H+ ATPase diffère très fortement entre les deux provenances. L'EF présente une inhibition transitoire à 6 heures tandis que l'EM induit transitoirement son expression à 6 heures. Le niveau d'expression est plus élevé chez l'EM.
La régulation des gènes codant pour des aquaporines n'a pas été mise en évidence dans cette étude, par contre le pin maritime semble répondre au stress par un transport actif d'acides aminés et de polypeptides ainsi que par des flux d'ions H+.

D.3.2.3.3.12 Non-classé

Cette catégorie est essentiellement occupée par les gènes dont seule une homologie avec des protéines hypothétiques (hypothetical protein) ou inconnues (Unknown protein) a pu être trouvée. Leur annotation est notée « Unknown ». Les gènes correspondants ont été détectés avec des effets provenance comme condition. Pour l'expérience SL, l'effet condition a été mis en évidence pour 8 gènes de fonction inconnue sous-exprimés en stress et 29 gènes surexprimés. L'effet provenance concerne 1 gène qui est surexprimé dans l'EM et 3 dans l'EF. Pour l'expérience SC, les gènes significatifs pour la condition sont surexprimés lors du stress et pour l'effet provenance, ils sont tous, sauf 2 gènes, plus exprimés chez l'EF.

Nous trouvons également des protéines dont la fonction n'a pas été établie et que nous avons choisi de placer dans la catégorie « Non-classés ». Certains gènes ont été rapportés dans la littérature comme régulés lors du stress. C'est le cas des gènes homologues au clone PtaAGP5 de *Pinus taeda* qui pourrait coder une arabinogalactane qui sont dans notre cas régulées négativement lors du SC comme du SL. Dans cette hypothèse, ce comportement confirmerait les résultats de Yang *et al.* (2005) qui ont mesuré par PCR quantitative le niveau de transcrits de 11 arabinogalactanes sur des plants soumis à la sécheresse. Le niveau de transcrits était significativement abaissé lors du stress pour toutes les arabinogalactanes sauf une.

Des gènes codant des protéines riches en proline sont détectés dans les deux expériences et présentent une régulation négative. Nous avons également des protéines associées à l'embryogenèse comme EMB3 induite lors du SL ou EMB7 détectée en SC avec un effet provenance et un niveau de transcrit supérieur pour l'EM.

On pourra enfin citer un gène homologue à une protéine de liaison au Calcium qui est caractérisée par un triple effet provenance, condition et interaction en SL. Ce gène a une expression similaire entre les deux écotypes en condition contrôle mais l'expression chute au cours du stress a l'EF tandis qu'elle se maintien pour l'EM. Ce type de profil est particulièrement intéressant car il montre un gène régulé différemment par les provenances lors du stress.

D.3.2.3.4 Bilan des principales fonctions régulées

Les fonctions des gènes évoqués dans le paragraphe précédent nous permettent d'entrevoir les principales régulations mises en jeux lors du stress et chez les deux écotypes.

A l'échelle de la cellule, nous avons une régulation de la croissance par des gènes tels que Diminuto et celui de la prohibitin dont l'inhibition favoriserait la division et

l'allongement cellulaire. La composition de la paroi cellulaire semble également être l'objet de modifications avec l'intervention de gènes codant pour l'extensine, des celluloses synthases, cellulases, glucan endo-1,3-beta-D-glucosidase et MERI-5 (xyloglucosyl transferase) dont l'expression varie lors du stress. Nous trouvons aussi des gènes codant pour la PAL, la caffeate o-methyltransferase et la caffeoyl-CoA-o-methyltransferase qui suggèrent également une modification de la teneur en lignines. L'expression et la régulation de ces gènes intervenant dans la formation de la paroi diffèrent entre les deux écotypes.

Au niveau moléculaire, nous avons mis en évidence des gènes différentiellement exprimés depuis la perception du signal avec l'intervention du Ca^{2+} jusqu'à la synthèse et la modification des protéines. En ce qui concerne la perception et les étapes précoces de la transmission du signal, nous avons une régulation de récepteurs à protéine kinase, d'une protéine à sept hélices transmembranaires, d'un récepteur protéine kinase CLAVATA1, de phospholipases, d'IMP et de calmoduline. La détection de facteurs de transcriptions fait suite aux cascades de transduction du signal. Dans cette étude nous avons mis en évidence des protéines en doigt de zinc dont certaines sont induites par le stress et d'autres inhibées et pour lesquelles nous observons des régulations contrastées entre écotypes. Nous avons détecté sept autres facteurs de transcription dont un facteur EREBP suggérant un rôle de l'éthylène dans la réponse au stress. Cette constatation est renforcée par la régulation de la SAM synthase, la SAM décarboxylase et l'ACC oxydase qui sont également différentiellement exprimés entre écotypes. La régulation de l'expression des gènes semble également contrôlée par l'organisation même de l'ADN. En effet des gènes codant des histones (H2A, H2B et H4) ainsi qu'une protéine hétérochromatine-like sont régulés. La régulation semble aussi se faire au niveau de la traduction, ce que souligne la présence de gènes régulés différemment entre écotypes et codant un facteur EF1-gamma. Nous avons détecté aussi une série de gènes codant des sous unités ribosomales. L'effet du traitement et/ou de la provenance a également un impact sur les modification et dégradations des protéines avec la présence de gènes codant des sous-unités du complexe protéosomique, des peptidases (signal peptidase I, endopeptidase Clp, thimet oligopeptidase) et une alpha galactosyltransférase. L'abondance de transcrits de cathepsin B-like cystéine protéinase chez l'EM suggère une activité lysosomale plus élevée chez cet écotype.

Les mécanismes de détoxification sont mis en avant par la présence de gènes codant des métallothionéines, des catalases, une protéine chaperonne pour la SOD et des peroxydases. La réponse au stress hydrique est ici confirmée entre autre par les gènes codant des HSP, des LEA, Déhydrines, Défensine et protéines de réponse aux pathogènes. Des

régulations différentes entre écotypes sont, là encore, observées. Enfin, des gènes codant des protéines membranaires à rôle de transport sont détectés, notamment des transporteurs de peptides et d'acides aminés ainsi qu'une pompe à proton pouvant jouer un rôle dans les équilibres osmotiques.

D.3.2.3.5 Les gènes détectés sont-ils les même dans les deux expériences ?

Au total, 290 accessions n'ont été détectées que dans l'expérience SC (Tab. Ax-6), 444 accessions uniquement dans l'expérience SL (Tab. Ax-7) et 168 accessions ont été mises en évidence dans les deux expériences (Tab. Ax-8) *(Les gènes annotés « Putative protéine » ou « Unknown » ne sont pas inclus dans ces tableaux)*.

Le tableau D-10 indique le nombre et le pourcentage d'accessions sans fonction connue parmi les gènes spécifiques et communs aux deux expériences. Pour l'ensemble de la banque, nous avions environ 27 % de « putative » et 10% de « unknown » (Fig. D-14). Avec 22.7% de « putative » et 21.6% de « unknown » dans le SL, cela indique que la moitié des gènes de fonction inconnue détectés, présentent une homologie avec des séquences déjà identifiées. Les gènes différentiellement exprimés dans le SL semblent donc être moins « spécifiques » du pin maritime. Par contre, le fait d'obtenir plus de 30% de « putative » dans le SC et dans les gènes communs aux deux stress, ouvre la porte sur la découverte de gènes encore non étudiés, peut-être spécifiques à l'espèce et impliqués dans la réponse au stress hydrique des deux écotypes de pin maritime.

Concernant les fonctions des gènes rencontrés, la répartition des accessions en catégories fonctionnelles (Fig. D-36) indique que la proportion de gènes « non-classés » est plus faible dans les gènes « communs » aux deux expériences. Les gènes spécifiques du SL diffèrent de ceux du SC par une plus grande part de gènes du « métabolisme » et de l'« énergie ». La catégorie transcription occupe une place plus importante tandis que les catégories « réponse au stress » et « trafic intra cellulaire » sont moins représentée. La répartition des gènes communs se distingue par une faible proportion de gènes « non-classés » associée à une plus grande représentativité des classes « organisation cellulaire » et « réponse au stress ».

Les gènes détectés pour un effet simultané provenance et condition, ou présentant une interaction, sont des cas particulièrement intéressant car ils répondent au stress et différencient les deux écotypes dont la provenance marocaine est plus résistante à la sécheresse. Dans cet objectif, il conviendra d'étudier les gènes sur leur profil d'expression, ce que nous ferons dans le chapitre suivant en utilisant des méthodes de regroupement.

Tableau D-10 : Proportion d'accessions annotées « Putative » ou « Unknown » par rapport au nombre de gènes détectées spécifiquement dans l'expérience stress court, stress long ou dans les deux expériences. Type d'effet significatif détecté : P= Provenance, C= Condition et I=Interaction.

	Spécifiques Stress Court				Spécifiques Stress Long				Communs			
	Putative		Unknown		Putative		Unknown		Putative		Unknown	
effet P,PI	72	24,8%	18	6,2%	7	1,6%	3	0,7%	_	_	_	_
effet C,CI	8	2,8%	2	0,7%	82	18,5%	32	7,2%	_	_	_	_
effet PC, PCI	7	2,4%	1	0,3%	12	2,7%	1	0,2%	_	_	_	_
Total (avec effet I)	88	30,3%	21	7,2%	101	22,7%	96	21,6%	54	32,1%	5	3,0%

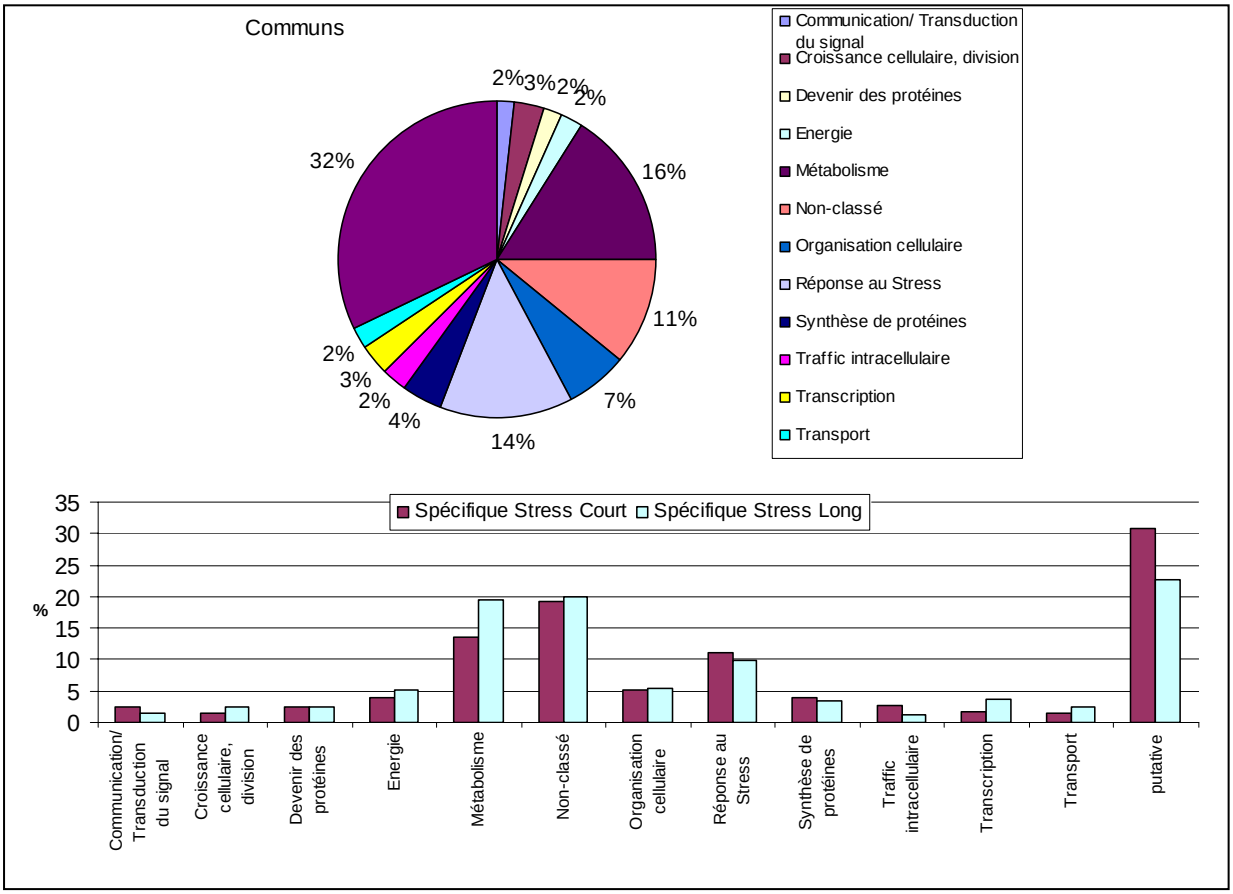


Figure D-36 : répartition dans les catégories fonctionnelles des accessions détectées spécifiquement dans l'expérience stress court, stress long et communs aux deux expériences.

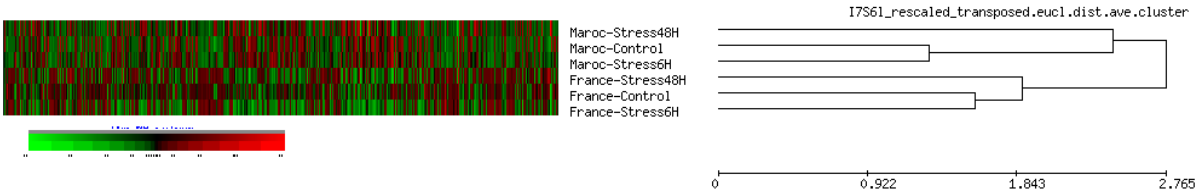


Figure D-37 : Regroupement des données centrées sur la moyenne des gènes différentiellement exprimés dans l'expérience stress court. Classification hiérarchique UPGMA, distances euclidiennes, regroupement par échantillon.

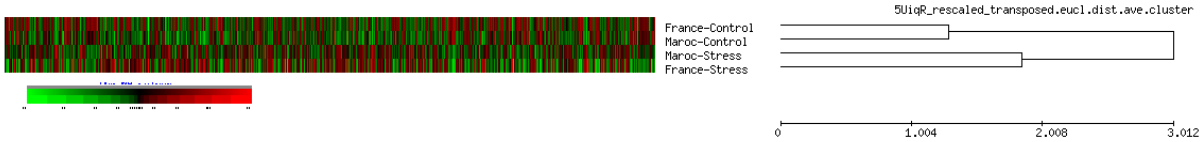


Figure D-38 : Regroupement des données centrées sur la moyenne des gènes différentiellement exprimés dans l'expérience stress long. Classification hiérarchique UPGMA, distances euclidiennes, regroupement par échantillon.

Nous avons par exemple le facteur d'élongation 1-gamma (BX252622) surexprimé lors du stress et plus abondant chez les marocains. Il y a également un gène codant une plantacyanine (BX678708) dont l'expression diminue avec le stress dans la provenance landaise mais se maintient dans les marocains. L'accession (BX681748) code pour une enzyme de la voie alternative sans mévalonate de la biosynthèse des terpènes. Son expression diminue lors du stress mais reste supérieure pour l'écotype marocain. C'est également le cas de la GPP synthase (BX254083) et de la pinène synthase (BX682869).
Nous retrouvons aussi les protéines de choc thermique (BX252065, BX252366, BX681237) ainsi que la protéine induite par l'ABA (BX679279), déjà signalées comme induite lors du stress et plus abondantes chez les marocains.
Il serait intéressant d'étudier plus avant le groupe de facteurs de transcriptions mis en évidence dans les deux expériences et caractérisé par des effets provenance et condition simultanés. Ces facteurs sont un facteur ERE (BX680424), un facteur G-box (BX682416), une protéine F-box (BX254733), un Rof1 (BX251374) et un gène présentant une homologie avec une protéine HIT-10 à doigt de zinc (BX253595).

D.3.2.4 Regroupement et classification des données d'expression

L'étude statistique par analyse de variance a permis de détecter les gènes différentiellement exprimés pour les facteurs provenance, condition et interaction. Les résultats précédents ont permis d'identifier de nombreux gènes impliqués dans la réponse au stress ou différentiellement exprimés entre les deux écotypes. Cependant, ce type d'analyse ne donne pas la possibilité d'identifier les gènes sur le sens de variation (surexpression vs sous expression), leur profil d'expression et l'amplitude de la régulation. Les méthodes utilisées dans ce chapitre vont permettre de répondre à ces questions et notamment de regrouper les gènes co-régulés.

D.3.2.4.1 Organisation et profil d'expression des groupes générés

Les données de chaque gène ont été centrées sur la moyenne du gène pour les raisons évoquées dans la partie matériel et méthodes. Nous nous intéresserons ici uniquement à la classification des gènes différentiellement exprimés.
Le regroupement des données par type d'échantillon donne les arbres des figures D-37 et D-38, respectivement pour l'expérience SC et SL. Dans l'expérience « stress court », les échantillons sont regroupés par provenance puis les échantillons les plus proches au sein des provenances sont ceux des conditions contrôle et 6 heures de stress, l'échantillon 48H se

détachant des 2 précédents. Concernant l'expérience « stress long », les échantillons sont cette fois ci regroupés par condition. Cette approche confirme les résultats obtenus précédemment qui montraient pour l'expérience SC un effet provenance majoritaire et pour l'expérience SL un effet condition beaucoup plus marqué.

Le regroupement par gène aboutit aux figures D-39 et D-40. Les couleurs permettent de distinguer des groupes de gènes dont les profils de régulation sont proches. Pour le SC, nous décelons trois profils de régulation principaux (Fig. D-39). Le groupe supérieur A correspond à des gènes dont l'expression est plus forte dans l'écotype français (EF). Le groupe intermédiaire B rassemble les gènes plus exprimés dans l'EM quelque soit la condition ainsi que des gènes dont l'expression diminue lors du stress. Le groupe inférieur C correspond globalement aux gènes surexprimés au cours du stress (surtout à 48H). Pour ce qui est de la réponse au SL, nous distinguons sur la figure D-40 un bloc supérieur A de gènes surexprimés lors du stress, suivi d'un groupe B de gènes dont le niveau d'expression diffère entre provenances et avec le stress. La moitié inférieure de l'arbre est composée de deux groupes, le plus important en taille, C, englobe les gènes régulés négativement lors du stress et l'extrémité inférieure D de l'arbre est occupée par des gènes fortement régulés positivement au cours du stress. Le fort niveau de régulation des gènes du D est sans doute à l'origine de la création d'un groupe distinct du A. Ces représentations sont finalement assez peu informatives car elles rendent difficile la détermination des groupes de profil de régulation et la synthèse des informations. Par contre, il est possible de couper l'arbre à un niveau seuil pour obtenir un nombre de groupes restreint et donner un profil type des gènes pour chaque sous groupe. C'est ce que permet le module SOTArray (cf. matériel et méthodes). Les gènes sont regroupés à la fois en fonction de leur profil et de l'amplitude de la régulation.

L'arbre correspondant à l'expérience SC est représenté par la figure D-41. Le seuil de variabilité intra-nœud proposé par défaut dans la suite EPCLUST est de 0.405. Les gènes sont divisés en 7 groupes organisés en 3 branches principales (comme à la figure D-39). La branche supérieure (61+85=146 gènes) correspond à des gènes régulés négativement dont l'expression est plus élevée chez l'EM, une branche intermédiaire (47+39+16=102 gènes) de gènes plus exprimés chez l'EM et dont l'expression tend à augmenter avec le stress, et la branche inférieure (25+185=210 gènes) qui regroupe les gènes plus exprimés chez l'EF. Dans cette branche, le groupe 1 rassemble à lui seul 185 gènes soit 40 % des accessions détectées. Si l'on regarde les groupes 3 à 7 (Fig. D-41), nous avons 248 gènes (16+39+47+85+61) gènes globalement plus exprimés dans l'EM et 210 gènes (groupes 1 et 2 (185+25)) plus exprimés dans l'EF quelque soit la condition. Le nombre de gènes plus exprimés dans chaque écotype

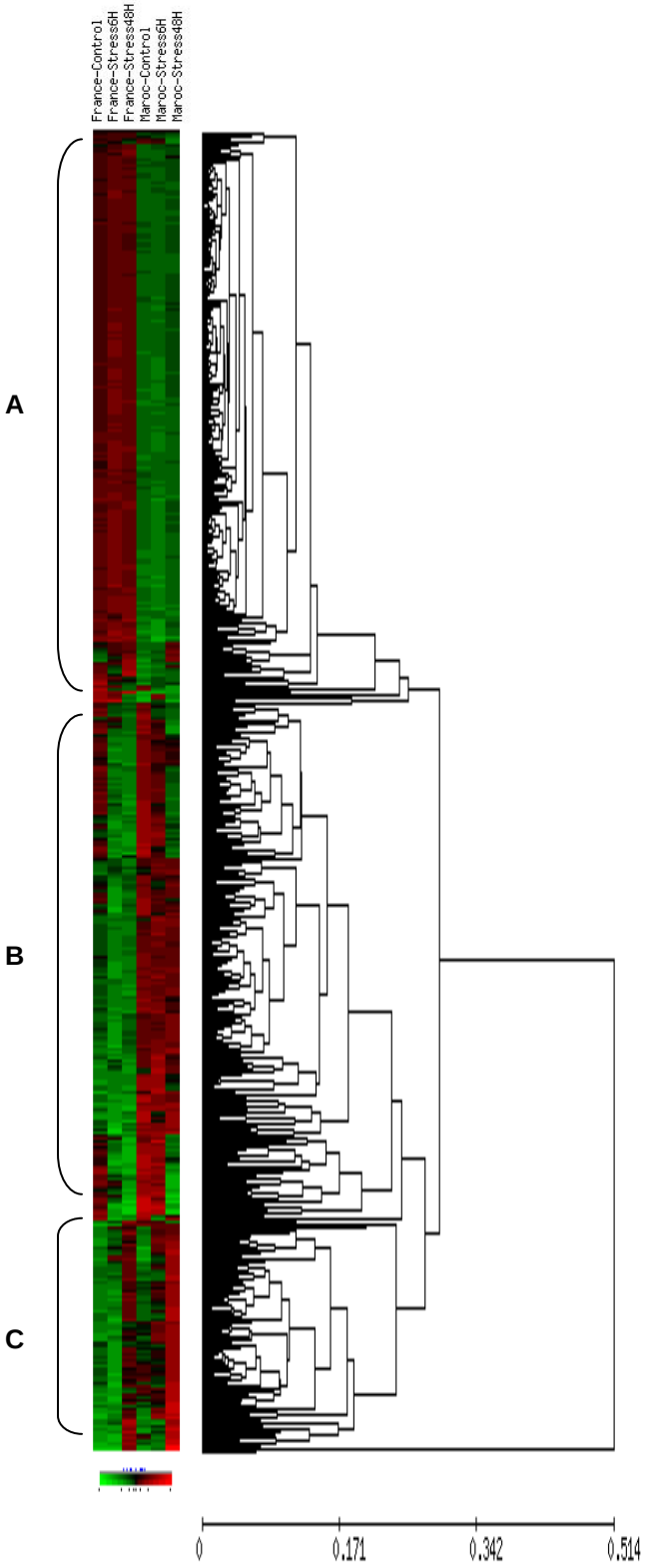


Figure D-39 : Regroupement des données des 458 gènes différentiellement exprimés dans l'expérience stress court (SC). Classification hiérarchique UPGMA, distances euclidiennes, regroupement par gènes.

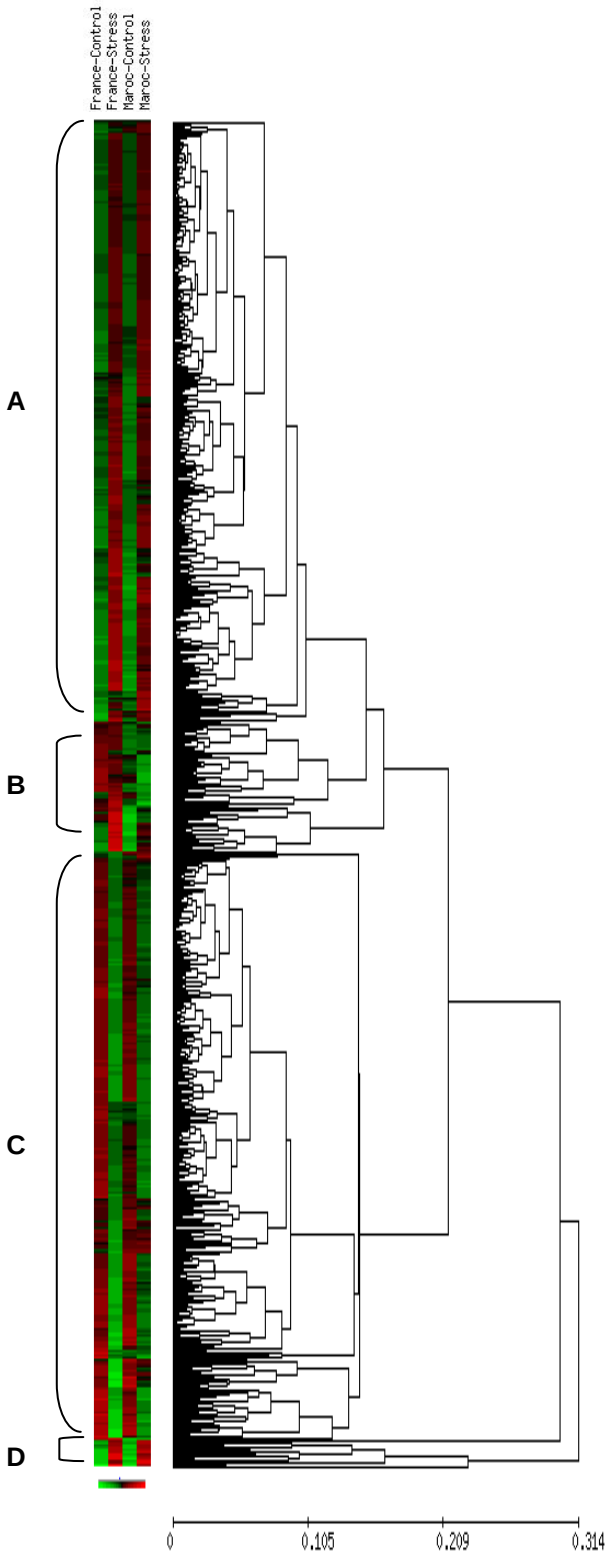


Figure D-40 : Regroupement des données des 612 gènes différentiellement exprimés dans l'expérience stress long (SL). Classification hiérarchique UPGMA, distances euclidiennes, regroupement par gènes.

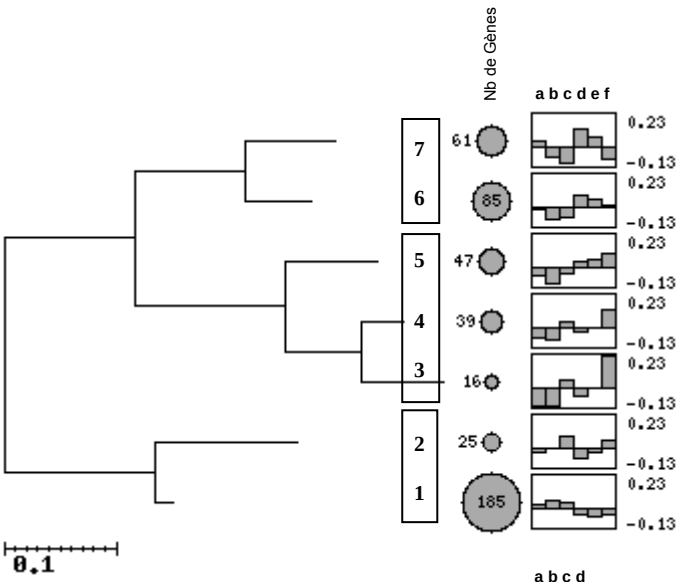


Figure D-41 : Regroupement des gènes différentiellement exprimés dans l'expérience stress court par SOTArray (Classification hiérarchique UPGMA, distances euclidiennes).
De gauche à droite, a) France Control (FrCc), b) Fr6H, c) Fr48H, d) Maroc Control (McCc), e) Mc6H, f) Mc48H.

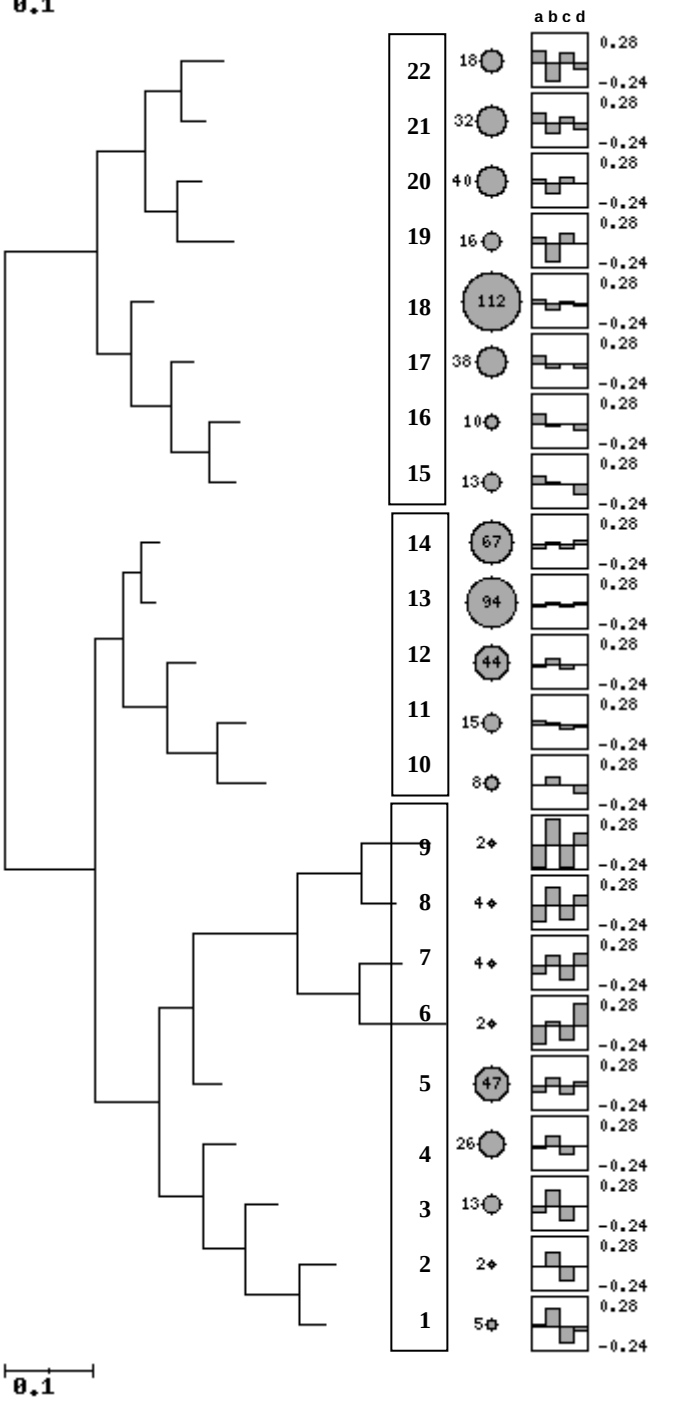


Figure D-42: Regroupement des gènes différentiellement exprimés dans l'expérience stress long par SOTArray (Classification hiérarchique UPGMA, distances euclidiennes).
De gauche à droite, a) France Control (FrCL), b) France Stress Long (FrSL), c) Maroc Control (McCL), d) Maroc Stress Long (McSL).

est donc similaire mais 146 des gènes surexprimés chez l'EM répondent également au stress par une diminution de l'expression. La quantité de transcrit de ces gènes dont l'expression diminue, restera donc plus élevée chez l'EM. Un maintien d'une quantité plus élevée de transcrits quelque soit la condition peut être une stratégie de réponse au stress.

L'arbre généré pour l'expérience SL est plus éclaté et comporte 22 groupes organisés également en 3 branches principales (Fig. D-42). Le seuil de variabilité intra-nœud proposé par défaut est égal à 0.175. La branche supérieure (279 gènes) englobe les gènes sous exprimés lors du stress, la branche inférieure (105 gènes) les gènes surexprimés et la branche centrale les gènes faiblement régulés (228 gènes). Nous détectons donc en SL près de trois fois plus de gènes sous exprimés lors du stress que de gènes surexprimés.

La seconde méthode de classification fournie par le logiciel Expander permet de synthétiser encore plus les données d'expression en modulant l'homogénéité minimum attendue dans chaque cluster. Dans le cas du SC, la figure D-43 illustre les profils moyens et les écart-types des 6 clusters obtenus par défaut avec une homogénéité moyenne de 0.949. Sur les 458 accessions, seuls 22 singletons n'ont pas pu être affecté à un cluster. Pour l'expérience SL, le nombre de clusters générés par défaut étant trop important, la valeur d'homogénéité moyenne intra cluster visée a été rehaussée à 0.9 et l'homogénéité moyenne obtenue est de 0.963. Expander génère alors 7 clusters différents et 37 singletons (Fig. D-44). Une représentation sous la forme d'une matrice des données clustérisées est également accessible (Fig. D-43 et D-44). Les détails de tous les profils d'expression d'un cluster donné sont présentés aux figures Ax-1 et Ax-2.

Le but de ces deux méthodes de regroupement est de produire deux niveaux d'information. Le premier, synthétique est obtenu par Expander (Fig. D-43 et D-44) et le second, généré par SOTArray (Fig. D-41 et D-42) permet un regroupement plus détaillé par classification hiérarchique.

La liste des gènes détectés dans les deux séries d'expériences et leurs affectations aux groupes générés par SOTArray et Expander sont accessible au tableau Ax-9.

Les groupes des deux méthodes montrent des similarités qui valident les analyses. Si l'on s'intéresse par exemple aux regroupements générés sur les données du SC (Tab. Ax-9), nous pouvons voir que les groupes sont relativement conservés (Cluster 1 Expander = groupe 1 SOTArray +qq gènes groupe 2 ; Cluster 2 Expander = groupes 6 et 7 SOTArray ; Cluster 3 Expander = groupes 3 + 4 + 5 SOTArray ; Cluster 4 Expander = groupe 5 + qq gènes groupe 6 ; Cluster 5 Expander = groupe 7 + qq gènes 1 et 6 ; Cluster 6 = groupe 2 + qq gènes groupes 3 et 4). Au bilan, mis à part quelques gènes clustérisés avec d'autres groupes, le groupe 7 de

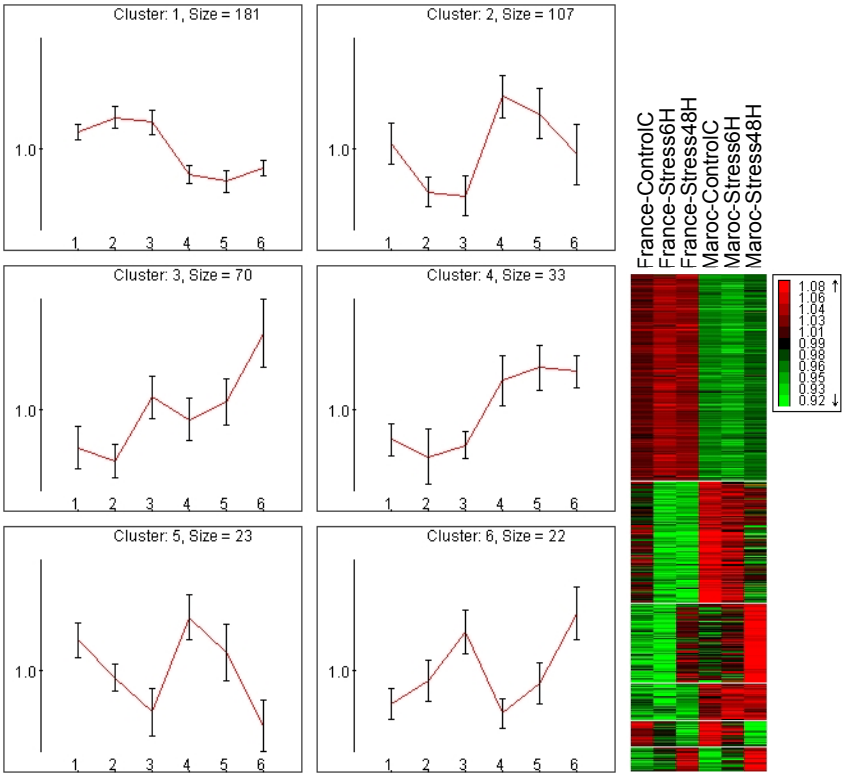


Figure D-43 : Clusters et matrice des données clustérisées, générés par la méthode « click » de Expander sur les gènes détectés dans l'expérience stress court.

1) France-ContrôleC, 2) France Stress6H, 3) France Stress48H, 4) Maroc-ContrôleC, 5) Maroc-Stress6H, 6) Maroc-Stress48H.

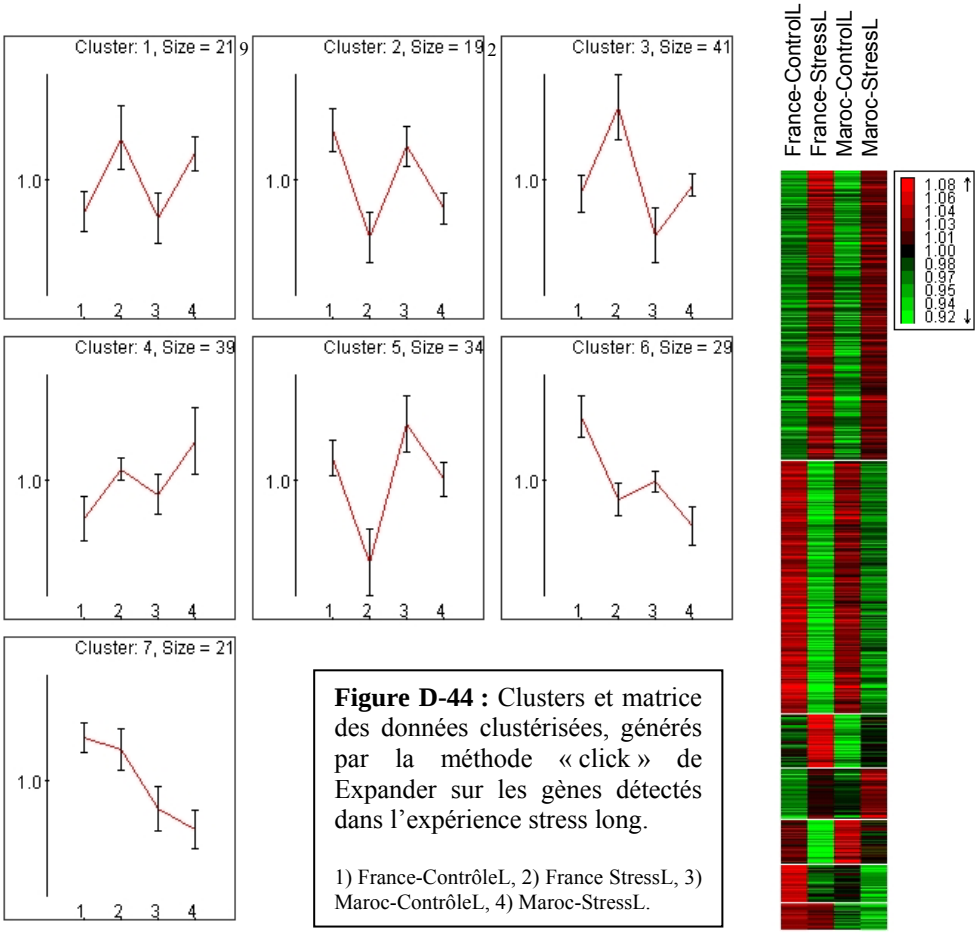


Figure D-44 : Clusters et matrice des données clustérisées, générés par la méthode « click » de Expander sur les gènes détectés dans l'expérience stress long.

1) France-ContrôleL, 2) France StressL, 3) Maroc-ContrôleL, 4) Maroc-StressL.

SOTArray a été réparti dans les clusters 2 et 5 de Expander et une partie des gènes du groupe 5 SOTArray sont réaffectés au Cluster 3. Les principaux regroupements sont donc conservés entre les deux méthodes.

D.3.2.4.2 Plasticité moléculaire et gènes d'intérêt

Une des questions essentielles à laquelle nous voulons répondre est d'identifier les gènes qui sont responsables de la différence de tolérance à la sécheresse entre les deux écotypes de pin maritime. Cette question en amène d'autre ; on peut se demander si ces gènes seront détectés parmi les gènes régulés en stress court, en stress long, ou encore différentiellement exprimés entre écotypes.

Avant toute chose, il convient d'identifier ce qui va permettre à une plante de répondre à son environnement. Un même génotype va pouvoir présenter différents phénotypes selon le milieu. L'existence de cette plasticité phénotypique est indispensable à la survie de la plante dans des conditions qui ne sont jamais optimales à sa croissance. Il est possible de considérer la plante comme un système dynamique en équilibre avec son milieu. L'environnement n'étant jamais stable, le système plante va devoir s'acclimater en permanence pour conserver une certaine homéostasie. Même si la régulation des gènes, leur expression et la traduction en protéines peut être rapide, il est peu probable que leur régulation soit l'unique source du maintien de cet équilibre dynamique. L'ensemble de l'organisation de la plante, sa structure et les métabolites cellulaires offrent une plasticité propre pour répondre aux fluctuations des conditions du milieu. Dans un milieu relativement stable, l'expression de l'ensemble des transcrits assurera le renouvellement des protéines nécessaires au maintien de cet état. L'extraction et l'analyse des transcrits dans cette phase permet d'accéder à une représentation ponctuelle du transcriptome. Deux génotypes (ou deux écotypes) n'adopteront pas forcément la même structure, la même composition en métabolites et donc le même transcriptome pour atteindre leur état d'équilibre et n'auront pas forcément la même plasticité. Les gènes différentiellement exprimés entre écotypes, quelque soit le niveau de stress, pourront donc indiquer une adaptation à des milieux différents due à la sélection naturelle. Dans notre cas, une des composantes différentiant les biotopes des deux écotypes est le niveau d'aridité. Une partie de ces gènes pourra donc être impliquée dans l'adaptation constitutive à une sécheresse chronique.

Un changement brutal de l'environnement (ex : Stress Court) va nécessiter un changement rapide de l'état initial (contrôle) vers le nouvel état (stressé) pour rétablir un nouvel équilibre. Cela va nécessiter un ajustement du métabolome et par conséquent du

transcriptome. Les gènes régulés doivent permettre d'atteindre rapidement cet objectif. Une plante rencontre très fréquemment des variations brutales et transitoires de son environnement. Pour ce type de changement d'état très limité dans le temps, il est peu concevable que la plante réponde par des modifications de fond touchant sa structure et son organisation (ex : croissance racinaire, chute des feuilles…). Le changement d'état d'équilibre doit faire appel à une plasticité du système faisant intervenir des gènes de réponse « d'urgence » (ex : flux ioniques et équilibre osmotique). Les gènes différentiellement exprimés au cours de ce type de traitement pourront refléter la dynamique de réponse au stress. Si ces gènes peuvent participer à l'aptitude à résister au stress hydrique, ils ne sont pas obligatoirement impliqués dans l'acclimatation ou l'adaptation des écotypes à des déficits en eau plus long.

Le maintien d'un état de stress prolongé (Stress long) correspond à un nouvel état « stable ». Par rapport à un état « contrôle », la plante doit mettre en place un système lui permettant de conserver ce nouvel équilibre. Dans ce cas, des modifications de fond, de structure sont certainement mises en œuvre. En analysant le transcriptome après 3 semaines de stress, nous ne pouvons pas mettre en évidence les gènes intervenant dans ce changement d'état mais nous avons accès aux transcrits des gènes nécessaires à cette nouvelle situation d'équilibre. Là encore, les deux écotypes n'adoptent pas forcément les mêmes réponses pour atteindre cet état d'équilibre. L'expérience SL, devrait donc nous permettre de détecter les gènes mis en œuvre pour maintenir le nouvel état « stressé » et de comparer les différences entre écotypes.

La plasticité moléculaire des deux écotypes sera donc multiple. Nous aurons une composante dynamique avec la réponse à court terme et une composante d'acclimatation avec la réponse à long terme. A cette plasticité des écotypes vis-à-vis de la réponse au stress osmotique, se combine leur aptitude propre d'origine adaptative. Au niveau du transcriptome, les deux provenances apporteront des réponses communes face au stress en régulant l'expression de certains gènes. L'identification de ces gènes permettra d'étudier les gènes de réponse générale au stress (effet condition quelque soit l'écotype). Par ailleurs, des gènes avec des niveaux d'expression différents quelque soit la condition vont permettre de différencier les deux écotypes (effet provenance). Nous retrouvons aussi des gènes régulés au cours du stress et dont l'expression est différente entre les deux écotypes. C'est dans cette catégorie que nous pensons avoir le plus de chance d'identifier des gènes pouvant expliquer les différentes stratégies de réponse moléculaire voire la différence de résistance au stress hydrique entre les deux écotypes. Certains gènes seront détectés à la fois en SC et en SL et

pourront être considérés comme des gènes du stress, d'autres seront plus spécifiques de la réponse d'urgence ou de l'acclimatation.

Dans l'expérience SC, les clusters 1 et 4 (Fig. D-43) regroupent les gènes différentiellement exprimés entre provenances et les clusters 5 et 6 rassemblent les gènes de réponse au stress. Les clusters qui vont nous intéresser plus particulièrement sont les 2 et 3 où les gènes sont régulés par le stress et diffèrent entre provenances. Pour l'expérience SL (Fig. D-44), les gènes qui nous intéresseront le plus seront ceux des clusters 3, 4, 5 et 6 qui montrent une nette régulation lors du stress et une différence d'expression marquée entre écotypes.

Afin de nous focaliser sur une liste de gènes la plus ciblée possible en terme de profil d'expression, un croisement des données des clusters d'intérêt des expériences SC et SL a été réalisé (Tab. D-11). Les gènes appartenant aux clusters d'intérêt et communs aux deux expériences ainsi que les gènes de ces clusters qui sont « spécifiques » du SC ou du SL ont été retenus. La liste des **215 accessions** ainsi retenues est indiquée au tableau Ax-10. C'est essentiellement de cette liste que des gènes ont été choisis pour vérifier leur expression par PCR quantitative. Au total, 41 gènes n'ont pas été affectés à un cluster Expander, que ce soit dans l'une ou l'autre des expériences. Ces accessions ont probablement un profil d'expression atypique. La liste de ces accessions est accessible au tableau Ax-11.

D.3.2.4.3 Profils d'expression et fonctions associées

L'étude des fonctions et catégories fonctionnelles des gènes constituant un même groupe (Tab. Ax-9) n'a pas fait apparaître de corrélation nette entre un profil de régulation particulier (un groupe) et une catégorie de fonctions régulées. Que ce soit dans le cas des clusters d'Expander ou des groupes de SOTArray, il existe une grande variabilité de fonctions et de catégories fonctionnelles au sein de chaque regroupement. Nous allons toutefois passer en revue les fonctions des gènes des différents clusters du tableau Ax-9.

D.3.2.4.3.1 Gènes dont l'expression diffère entre écotypes

Dans le SC, cela concerne les clusters 1 et 4 (Fig. D-43). Sur les 181 gènes du cluster 1, surexprimés dans l'EF, seulement 23 sont également présents et affectés à un groupe dans le SL. Près de la moitié (47%) des accessions sont de fonction inconnue (putative + unknown). L'EF est notamment caractérisé par une expression plus importante que l'EM pour des gènes liés à la réponse au stress comme des LEA (BX666039, BX681471), des peroxydases (BX254690, AL751229), des protéines de résistance à la maladie (BX255778,

Tableau D-11 : Répartition des gènes différentiellement exprimés dans les expériences SC et SL (nombre de gènes) en fonction de leur affectation dans les clusters générés par Expander. En vert, sélection des gènes présentant un profil d'intérêt. « na » : gènes non affectés à un cluster Expander de l'expérience concernée (gène non détecté dans cette expérience ou faisant partie des singletons du clustering)

Stress Court \ Stress Long		Numéro du cluster "Expander"							
		1	2	**3**	**4**	**5**	**6**	7	na
Numéro du cluster "Expander"	1	16	0	0	5	0	0	2	158
	2	0	39	0	1	11	4	3	49
	3	23	2	1	5	0	0	0	39
	4	3	2	0	2	3	0	0	23
	5	0	10	0	0	2	2	3	6
	6	9	0	1	1	0	0	0	11
	na	168	139	39	25	18	23	13	41

BX253926), des endochitinases (CR392715, BX249660) et une protéine de réponse à la déshydratation RD22 (CR392961). Il semble également que la synthèse des parois distingue les deux écotypes. En effet, nous retrouvons surexprimées dans l'EF : la PAL (BX682063), des pectates lyases (BX680185, 250283), une Glucan endo-1,3-beta-glucosidase (AL751122) et la 4CL. Des sous-unités ribosomales (25S, 30S(S16), 40S, 60S(L41)) et la protéine ribosomale L10 sont aussi plus exprimées chez l'EF. Il est intéressant de noter que l'EF accumule les transcrits de Bétaïne-aldéhyde déshydrogénase (EC 1.2.1.8) (BX682419) qui permet la synthèse de bétaïne qui est un osmoprotectant. Ces résultats semblent indiquer que l'EF aurait une voie de synthèse de la lignine plus active que l'EM, accumulerait des dérivés de la bétaïne et exprimerait plus certains gènes de réponse au stress (LEA, peroxydases, disease resistance) et ce, de façon constitutive quelque soit le traitement. Ces caractéristiques peuvent avoir une forte répercussion sur les aptitudes des 2 écotypes à répondre au stress. Cette hypothèse ne peut pas être directement confirmée par les résultats du SL où 15 des 21 gènes du groupe 7 sont de fonctions inconnues (Fig. D-44). Nous retrouvons par contre deux gènes codant des Beta-glucosidase qui vont pouvoir agir sur les parois et une Défensine, qui sont plus exprimés dans l'EF.

A l'inverse, nous avons beaucoup moins de gènes caractéristiques de l'EM. Dans le SC, le cluster 4 (Fig. D-43) comprend 33 gènes dont seulement 9 n'ont pas de fonction connue. Parmi les gènes surexprimés de façon constitutive chez l'EM nous trouvons des gènes en rapport avec la détoxification et l'oxydation comme la Glutathione transferase (EC 2.5.1.18) (BX681235), un précurseur de la peroxydase (EC 1.11.1.7) (AL750365), la Thioredoxin type H (BX255712) et l'Alcohol dehydrogenase class III (BX252271) -oxidation des longues chaînes aliphatiques-. Des gènes vont intervenir sur les protéines (Endopeptidase Clp subunit (EC 3.4.21.92) (AL749972)) ou vont faciliter leur repliement (Peptidylprolyl isomerase (EC 5.2.1.8)(AL750449)). Il est intéressant de remarquer que nous trouvons des gènes impliqués dans le transport de protons avec une Carbonate dehydratase (EC 4.2.1.1)(AL750008) qui catalyse la réaction d'hydratation du carbone ($H^+ + HCO_3^- <-> CO_2 + H_2O$) et la présence d'une Proton-exporting ATPase (EC 3.6.3.6)(CR393662) qui nécessite de l'énergie. Nous trouvons l'Uridylate kinase (EC 3.6.3.6)(BX682721) qui catalyse la réaction de conversion de l'ATP et l'AMP en ADP qui pourrait renforcer cette idée à moins que sa surexpression ne soit associée aussi à la modification des protéines soulignée par l'ADP-ribosylation factor (BX254611) et un gène codant la sous-unité alpha du coatomère (BX252434). Enfin, l'EM surexprime une protéine de choc thermique 17.9kDa de classe II (BX252065). Il semble donc que l'EM se distingue de l'EF en maintenant de façon constitutive des mécanismes de

détoxification, de stabilisation et modification des protéines et le maintien d'un gradient de protons plus actif. La présence d'un gradient de protons pourrait servir au maintien d'un gradient d'ions ou d'autres composés (co-transport/antiport) pouvant jouer un rôle dans l'osmorégulation. En SL, aucun profil ne correspond à des gènes caractéristiques de l'EM dont l'expression serait plus forte que celle de l'EF quelque soit la condition.

Une partie des gènes exprimés différemment entre les deux écotypes et dont l'expression ne dépend pas de la condition ont des fonctions pouvant avoir un impact sur les capacités constitutives des 2 écotypes à répondre au stress hydrique. Il s'agit peut être de stratégies différentes des écotypes pour atteindre l'état d'équilibre avec leur milieu en maintenant à un niveau basal l'expression de ces gènes et ainsi permettre aux deux écotypes de faire face aux contraintes de leur environnement avec une plasticité optimale. **Pour résumer, l'EF diffèrerait de l'EM par l'activation des mécanismes touchant la composition des parois et la production de bétaïne, tandis que l'EM diffèrerait de l'EF par des mécanismes de détoxification et probablement un maintien de gradients ioniques plus actifs.** A savoir si ces mécanismes favoriseraient l'adaptation de chaque écotype dans leur milieu naturel, est encore trop prématuré, mais une investigation des patrons de diversité nucléotidique de ces gènes devrait permettre d'apporter des éléments de réponse à cette hypothèse.

D.3.2.4.3.2 Gènes dont l'expression dépend de la condition

En SC, peu de gènes (cluster 5 = 23 ; cluster 6 = 22) répondant au stress ont une expression similaire dans les 2 écotypes. Le cluster 5 (Fig. D-43) regroupe 23 gènes sous exprimés dans les deux écotypes. Il comporte des gènes codant une protéine riche en proline (BX679458) et une arabinogalactane protéine PtaAGP5 (BX679871) que l'on aurait plutôt attendue comme induite par le stress. Dans ce cluster nous trouvons aussi des précurseurs de peroxidases (CR392585, CR392453, BX679939) qui sont sous exprimés lors du stress et des enzymes agissant sur les oligosaccharides comme une xyloglucosyl transferase (EC 2.4.1.207)(CR393553) et un précurseur de la beta-glucosidase (EC 3.2.1.21)(CR392379).

Dans le cluster 6 du SC, nous retrouvons des gènes classiquement surexprimés lors du stress comme une Déhydrine (BX249564), une protéine de réponse à l'éthylène (CR394461), une protéine de choc thermique de 70kDa (BX251816), une protéine hydrophobe RCI2 (induite par les basses températures et le sel) et une protéine reliée aux pathogènes (BX680166). Nous trouvons aussi une protéine à 7 domaines transmembranaires (CR382434), un récepteur à la protéine kinase (BX682453) qui interviennent dans la transduction du signal.

Les gènes dont le profil dépend de la condition sont beaucoup plus nombreux dans le SL. Le cluster 1 (Fig. D-44) qui correspond aux gènes induits par le stress regroupe 219 gènes tandis que les gènes réprimés sont au nombre de 192 dans le cluster 2. Nous ne détaillerons pas les fonctions des gènes qui sont présentés dans le tableau Ax-9. Cependant, nous retrouvons dans les gènes surexprimés les gènes classiquement décrits comme induits par le stress comme les protéines LEA, Déhydrine, HSP, ASR ainsi que des gènes associés à la transduction du signal, des facteurs de transcriptions et à la détoxification. Les gènes sous exprimés couvrent l'ensemble des catégories fonctionnelles.

D.3.2.4.3.3 Gènes répondant différemment entre écotypes lors du stress

Il s'agit des gènes dont la plasticité moléculaire dépend de l'écotype. En SC, la différence de niveau d'expression entre les deux écotypes est plus marquée pour les **gènes surexprimés lors du stress qui sont également plus exprimés dans l'EM** et regroupés dans le cluster 3 (Fig. D-43). Pour ce cluster, l'augmentation du niveau d'expression chez l'EF après 48H de stress correspond au niveau d'expression de base de l'EM. L'induction de ces gènes lors du stress, associée à une expression plus élevée chez l'EM résistant à la sécheresse, en font des candidats privilégiés pour la dynamique de la réponse adaptative au stress osmotique. En d'autres termes, ces gènes seront les marqueurs de la capacité de l'EM, résistant, à retrouver rapidement un nouvel état d'équilibre compatible avec l'apparition d'un stress violent. Sur ces 70 gènes nous trouvons 23 gènes de fonction inconnue. Parmi les gènes connus, nous avons des gènes de réponse au stress avec des protéines de choc thermique (classe I, small, 101)(BX252366, BX250781, BX681237) dont notre gène de contrôle positif HSP, des LEA (BX253089, BX250781), des Métallothionéines (BX255766, BX252844) et des protéines de réponse aux pathogènes avec la Chitinase (EC 3.2.1.14)(AL750218, BX677767, CR393567). Une protéine de la membrane plasmique induite par l'ABA est également présente. Dans la réponse à court terme, l'EM induit plus l'expression de gènes impliqués dans la perception et la régulation de la transcription : Phospholipase D alpha 1 (BX254449), RING finger protein 26 (BX252953), Zinc finger protein HIT-10 (BX253595), G-box binding factor 4 (BX682416), histone H4 (AL750508). Des gènes favorisant la traduction des ARNm : Elongation factor 1-gamma (BX252622) et Gb|AAD26879.1 (BX680013) et pouvant intervenir dans les modifications post-traductionnelles : Ubiquitin-conjugating enzyme (EC 6.3.2.19)(BX682459, BX255062, AL750092, BX250267), sont également détectés. La surexpression par l'EM de la Cellulose synthase (BE123499) et d'un précurseur de l'Elastine (BX255377) pourrait également indiquer un renforcement des parois

des cellules à moins que ce ne soit pour une élongation cellulaire mais la diminution de l'expression des gènes du métabolisme (cluster 2) accompagnant la réponse au SC ne va pas dans ce sens.

L'EM semble donc se distinguer de l'EF dans la réponse à court terme par 1) une plus forte activation de gènes intervenant dans la régulation de l'expression et la synthèse des protéines 2) la surexpression de protéines de réponse au stress 3) une synthèse de cellulose et d'élastine.

Dans le SL, les gènes présentant le même type de profil se répartissent dans les clusters 4 et 3 (Fig. D-44). Le cluster 4 regroupe les gènes induits dont l'expression est plus forte chez l'EM pour lequel le niveau d'expression de base correspond à celui de l'EF après 3 semaines de stress. Sur les 39 gènes, 21 n'ont pas de fonction connue. Comme en SC, l'EM induit plus que l'EF l'expression de gènes en rapport avec la transcription : Zinc finger protein HIT-10 (BX253595), F-box protein (BX254733), Homeobox-leucine zipper protein (BX678438). La présence d'une protéine heterochromatine-like suggère la possibilité d'un autre mode de régulation de l'expression en modulant l'accessibilité à l'ADN. En SL, il semble que l'EM privilégie la synthèse de flavonoïdes -Flavonol 3-O-glucosyltransferase (EC 2.4.1.91) (AL750486)- et de la SOD –« Copper chaperone for Superoxide dismutase » (BX677784)- pour la détoxification. Nous retrouvons également en SL chez l'EM des gènes codant des protéines de choc thermique (small, STI1)(BX250781, CR393209) dont notre gène de contrôle positif HSP. Nous relevons enfin la présence d'une fonction de protéine chaperonne dnaJ qui participe activement à la réponse aux stress osmotiques et thermiques en prévenant l'agrégation des protéines dénaturées par le stress en interagissant avec le dnaK. La surexpression de cette protéine est donc un avantage certain pour l'EM en situation de stress osmotique. A l'inverse, le cluster 3 regroupe les gènes induits dont l'expression est plus forte chez l'EF dont le niveau d'expression de base correspond à celui de l'EM après 3 semaines de stress. Sur les 41 gènes, 12 n'ont pas de fonction connue. L'EF exprime plus en SL des gènes de réponse au stress comme des peptides antimicrobiens (Al7506636, BX681928, CR394499), des protéines « Aluminium-induced protein-like » (BX683057, BX253756), un précurseur de peroxydase (BXBX680225) une protéine homologue au cytochrome P450 (BX679683) et particulièrement des protéines chaperonnes dnaJ (BX679383, AL751212, BX681402). La plupart des autres gènes de fonction connue interviennent dans le métabolisme et l'énergie. Il semble donc que l'EF surexprime des gènes lui permettant de maintenir son métabolisme tandis que l'EM met en œuvre des mécanismes de réponse peut

être plus spécifiques, la majorité des gènes (cluster 4) étant de fonction inconnue, les autres correspondant à des fonctions de régulation, stabilisation des protéines et de réponse au stress. **Nous trouvons 5 gènes communs aux clusters 3 du SC et 4 du SL qui constituent donc des gènes candidats potentiels. Les accessions CR394334 et BX678079 n'ont pas de fonction connue. Les accessions BX250781 et BX000656 correspondent respectivement à une HSP de faible poids moléculaire et à notre gène de contrôle positif HSP. La dernière accession, peut-être la plus intéressante, présente une homologie faible avec une protéine en doigt de Zinc HIT-10 mais également des homologies avec des séquences de fonction inconnue d'*Arabidopsis* (35%) et du riz (22%).**

Les **gènes sous-exprimés lors du stress et différemment exprimés entre écotypes** sont rassemblés dans le cluster 2 en SC (Fig. D-43). Le niveau d'expression de ces gènes reste plus élevé chez l'EM que chez l'EF, atteignant le niveau d'expression du contrôle de l'EF après 48H de stress. Les gènes répondant au stress de ce groupe seront donc relativement plus exprimé dans l'EM. Il peut s'agir d'un mécanisme actif de la part de l'EM. Sur les 107 gènes, 22 n'ont pas de fonction connue. Nous retrouvons un facteur de transcription Rof1 (BX251374), un facteur EREBP (BX680424) et les histones H2A (BX252660, BX252614) et H4 (BX679616). Nous trouvons des gènes de réponse au stress pour lesquels une surexpression est généralement rencontrée : précurseur de peroxydase (BX675142, CR392767, BX678783, BX678596, CR393156), LEA (BX678224, BX677964), Abscissic stress ripening protein (BX251094, BX682324), protéine induite par la déshydratation (BX681136) et Défensine (BX681602). Ce cluster se distingue par la présence d'une série de gènes en rapport avec la biosynthèse de composés terpéniques. Nous retrouvons une Isopentenyl-diphosphate delta-isomerase II (EC 5.3.3.2) (CR392581), une 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase (EC 4.6.1.12) (BX681748) (voie sans mévalonate), la Geranyl diphosphate synthase (BX254083) et la Pinène synthase (EC 4.2.3.14) (BX682869, CR393796). Ces résultats concorderaient avec l'hypothèse d'une teneur également plus élevée en pinène au niveau racinaire chez l'EM que le dosage sur les parties aériennes avait souligné. Nous citerons enfin la Cholinephosphate cytidylyltransferase B (EC 2.7.7.15) (BX678890) qui catalyse la réaction de CDP-choline en choline phosphate qui est le précurseur de la choline et permet la synthèse de bétaïne. Nous avions vu que la Bétaïne-aldéhyde déshydrogénase (EC 1.2.1.8) était plus exprimée dans l'EF, les deux écotypes pourraient donc avoir des stratégies de régulation différentes pour la synthèse de bétaïne. L'EF pourrait adopter une expression constitutive plus élevée de l'enzyme (EC 1.2.1.8) qui n'est pas régulé

par le stress tandis que l'EM maintiendrait une expression plus forte de l'enzyme (EC 2.7.7.15) qui diminue lors du stress.

En ce qui concerne le SL, les gènes régulés négativement lors du stress et présentant une expression la plus contrastée entre les deux écotypes appartiennent aux clusters 6 et 5 (Fig. D-44). Dans le cluster 6, sur les 29 gènes, 7 n'ont pas de fonction connue. Les gènes régulés négativement ont ici un niveau d'expression supérieur dans l'EF, atteignant le niveau basal du contrôle de l'EM après 3 semaines de stress. Nous retiendrons la présence de quelques gènes de réponse au stress codant une Metallothionein-like protein EMB30 (AL749965), une Disease resistance response protein 206 (CR394442) et une Dehydration stress-induced protein (BX681136). La voie de biosynthèse de la lignine est représentée par la C4H (Trans-cinnamate 4-monooxygenase (EC 1.14.13.11)) (BX681070) et la Caffeate O-methyltransferase (EC 2.1.1.68) qui donne du Ferulate précurseur du feruloyl-CoA. La présence de ces deux enzymes renforce l'hypothèse d'une voie de biosynthèse plus active de la lignine chez l'EF que nous avions évoquée en retrouvant les transcrits de la PAL et de la 4CL plus accumulés chez l'EF. Le cluster 5 correspond à des gènes plus exprimés dans l'EM et dont le niveau d'expression diminue faiblement après 3 semaines de stress, pour atteindre celui du contrôle de l'EF. Le profil évoque un maintien de l'expression chez l'EM tandis que l'expression chute chez l'EF. Parmi les fonctions des gènes concernés nous remarquerons la présence de protéines « Metallothionein-like (EMB30) » (AL750160, CR392866, BX249127) différentes de celles citées précédemment.

D.3.2.4.3.4 Gènes dont le profil d'expression est atypique

Les 41 gènes qui n'ont pas pu être affectés à un cluster du SC et du SL par Expander sont présentés dans le tableau Ax-11. Aucune fonction connue n'a pu être trouvée pour 13 d'entre eux. Les 11 premiers gènes de la liste ont cependant été assignés aux clusters 10 à 14 de SOTArray (Fig. D-42) qui correspondent à la branche des gènes très faiblement régulés. Cette faible régulation est certainement à l'origine de leur non affectation à un cluster d'Expander.

Pour les autres gènes, nous remarquons la présence de 12 accessions correspondant à des clones de *Pinus taeda* (Identifiant commençant par NX). Nous en avions en effet sélectionné 48 pour le dépôt parmi lesquels 14 accessions ont été détectées différentiellement exprimées. La quasi-totalité des accessions se trouvent donc parmi les gènes non regroupés. Cette constatation soulève la question de la qualité d'hybridation entre séquences d'espèces différentes ou une possible contamination de la copie de la banque américaine d'où sont tirés

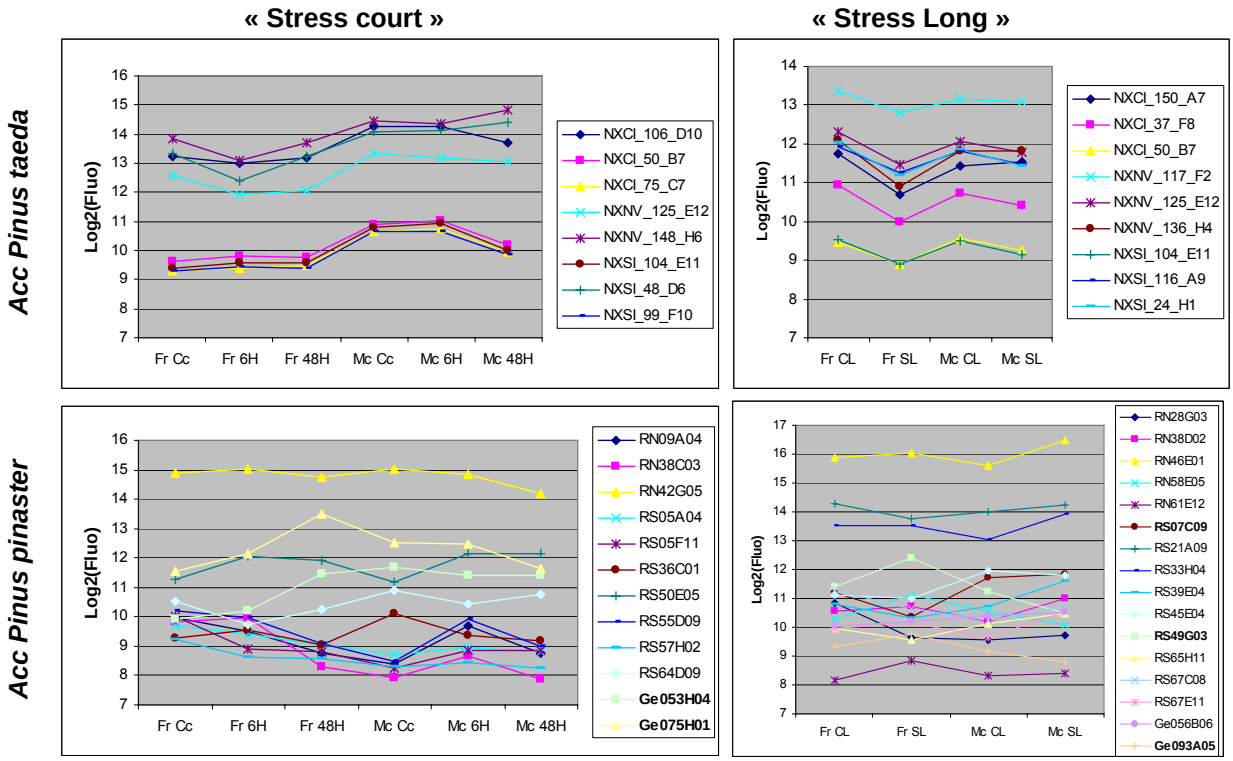


Figure D-45 : Profils d'expression (SC et SL) des accessions présentant des profils atypiques et qui n'ont pas été affecté à un cluster par Expander. Les profils des accessions de pinus teada et du pin maritime sont représentés séparément.

ces 48 clones. La figure D-45 présente les profils d'expression des accessions non affectées à un cluster par Expander. Nous pouvons remarquer que l'ensemble des accessions de *Pinus taeda* ont des profils très similaires ce qui est surprenant compte tenu du fait que les fonctions de ces accessions sont très différentes. Il y a vraisemblablement un problème avec ces clones ou l'hybridation. Les accessions de pin maritime non groupées sont dans l'ensemble faiblement régulées en SC comme en SL. Les accessions (BX680855) RS49G03 et (BX254079) Ge093A05 en SL, correspondant respectivement à une Alcool déshydrogénase (ADH) (EC 1.1.1.1) et une Inositol oxygénase, font exception et présentent un profil d'interaction avec une induction lors du stress pour l'EF et une diminution chez l'EM. La détection de l'ADH qui est un enzyme marqueur de l'hypoxie (Fukao and Bailey-Serres 2004), pourrait indiquer une différence d'oxygénation de certains bacs ou une sensibilité différente des écotypes. A l'inverse, une protéine Metallothionein-like (RS07C09) diminue lors du stress chez l'EF alors que son expression ne bouge pas chez l'EM. En SC, l'interaction concerne Ge075H01 de fonction inconnue et une protéine Germin (oxalate oxidase)-like (Ge053H04) qui augmente lors du stress chez l'EF pour atteindre à 48H le niveau d'expression de l'EM qui reste constant. **Les gènes que nous venons de citer ont des profils de régulation particulièrement intéressants car ils répondent de façon contrastée entre les deux écotypes et jouent potentiellement un rôle dans la réponse contrastée des écotypes au stress hydrique.**

Mis à part ces quelques gènes dont le profil d'interaction atypique est à l'origine de l'absence d'affectation à un cluster, le non regroupement vient donc essentiellement du fait que les gènes présentent une très faible régulation.

D.3.3 PCR quantitative en temps réel

Cette méthode permet d'accéder de façon précise et spécifique au niveau d'expression d'un gène dans un échantillon d'ARN total donné. Il s'agissait d'utiliser une technique non basée sur l'hybridation pour quantifier de façon précise le niveau d'expression des gènes sélectionnés à partir des résultats des microarrays, afin de confirmer leur implication potentielle dans la résistance au stress hydrique.

D.3.3.1 Choix des gènes

La sélection des gènes s'est faite préférentiellement parmi les 215 gènes d'intérêt évoqués au paragraphe D.3.2.4.2.

Une première liste d'environ 75 gènes a été sélectionnée arbitrairement en fonction de leur annotation et du cluster auquel ils appartenaient. Nous avons réparti le choix entre les différents groupes du Tableau Ax-10 formé par les clusters Expander et les sous-clusters SOTArray. Afin d'être le plus spécifique possible lors de l'amplification PCR, nous avons vérifié la position de la séquence disponible pour une accession donnée par rapport à la séquence consensus du TC correspondant du TIGR. Si la séquence n'était pas dans la partie 3' du gène, une séquence de pin maritime du même TC était choisie lorsqu'elle était disponible. La détermination des amorces a été conduite comme décrite dans la partie Matériel et Méthode. Dans 40% des cas (30 gènes), le dessin des amorces n'a pas pu être réalisé à cause de la position trop en 5' des EST de pin maritime disponibles, d'une longueur trop courte ou de l'absence d'amorce répondant à nos exigences. Sur les gènes restants, environ 30% (13 gènes) donnaient une amplification PCR faible, multi-bandes ou l'amplification en PCR quantitative avait une efficacité faible (<94%). Au final, 31 gènes ont passé l'ensemble des critères de sélection. Ils sont indiqués au tableau D-12. Leur expression a été mesurée sur l'ensemble des échantillons des expériences stress court et stress long.

Nous avons également sélectionné quatre gènes supplémentaires n'appartenant pas à la liste des 215 gènes et présentant un effet provenance, condition ou dont la fonction semble impliquée dans la réponse au stress hydrique. Notre choix s'est porté sur : 1) le gène d'accession (BX250268) de fonction non connue qui est présent dans les clusters Expander 1 du stress court et 7 du stress long (Fig. D-43 et D-44) et semble indépendant du stress mais différentiellement exprimé entre provenances, 2) un gène homologue à une épimérase / déshydratase (CR392825), situé dans les clusters 5 du SC et du SL (choix effet condition et fonction), 3) les gènes d'accession (BX680424 et BX682869) qui sont respectivement homologues à un facteur de liaison à un élément de réponse à l'éthylène *(EREBP)* et à la Pinène synthase, qui appartiennent aux clusters Expander 2 du SC et du SL (choix fonction).

L'expression des gènes de la voie de biosynthèse de l'éthylène a été mesurée pour les enzymes indiqués à la figure C-12. Aucune séquence n'a été trouvée pour l'ACC N-malonyl-transférase. Les amorces pour l'Adenosylhomocysteine nucléosidase (MTA nucléosidase) (EC 3.2.2.9) ont été dessinées sur BX254748 (TC6693), celles de l'ACC oxydase (EC 1.14.17.4) à partir de la séquence consensus du TC58524 (BX252194, BX252232), celles de l'ACC synthase (EC 4.4.1.14) à partir de la séquence consensus du TC69587, celles de la SAM synthase (EC 2.5.1.6) sur l'EST BX251224 du TC58204 et celles de l'ACC désaminase sur la séquence complète de la partie codante du pin maritime que nous avons séquencée (cf. D.4.5.)

Tableau D-12 : Liste des gènes étudiés en PCR quantitative et les accessions des ESTs sur lesquelles ont été dessinées les amorces. (31 gènes choisis d'après leur profil et 4 gènes complémentaires)

Gène étudié		Séquences utilisées		
Accession	Identifiant	Accession	Identifiant	Annotation
AL749954	AS03A12	AL749954	AS03A12	Thimet oligopeptidase (EC 3.4.24.15)
AL749965	AS03B12	AL749965	AS03B12	Metallothionein-like protein EMB30
AL750591	RN05B11	AL750591	RN05B11	Xyloglucan:xyloglucosyl transferase (EC 2.4.1.207)
AL751204	RS06C03	AL751204	RS06C03	T29M8.12 protein
BX250140	Ge032H07	BX250140	Ge032H07	Defender against cell death
BX252860	Ge071H04	BX252860	Ge071H04	5-methyltetrahydropteroyltriglutamate--homocysteine methyltra
BX253583	Ge085E07	BX253583	Ge085E07	Endoglucanase
BX253595	Ge085G01	BX680293 BX679631	RS41F03 RS30B09	Zinc finger protein HIT-10, partial (3%)
BX254093	Ge093C03	BX682524 BX682995	Ge117H08 Ge126F10	caffeate O-methyltransferase [Pinus taeda]
BX255062	Ge110B02	BX255062	Ge110B02	Ubiquitin-conjugating enzyme E2 (EC 6.3.2.19)
BX255377	Ge005B04	BX255377	Ge005B04	Elastin precursor - human (fragment) {Homo sapiens;} , partial
BX255395	Ge005C11	BX255395	Ge005C11	putative protein
BX677134	RN18B10	BX677134	RN18B10	putative protein
BX677784	RN44F10	BX677784	RN44F10	Copper chaperone for Superoxide dismutase
BX678596	RS11G02	BX678596	RS11G02	Peroxidase precursor (EC 1.11.1.7)
BX678613	RS12A03	BX678613	RS12A03	Gamma-thionin homolog precursor
BX678783	RS15A09	BX678783	RS15A09	Peroxidase precursor (EC 1.11.1.7)
BX679434	RS25D07	BX679434	RS25D07	putative protein
BX679792	RS34A05	BX679054 BX680429	RS19A06 RS43C03	PrMC3
BX680204	RS40E02	BX680204	RS40E02	4-hydroxyphenylpyruvate dioxygenase (EC 1.13.11.27)
BX680432	RS43C07	BX680432	RS43C07	Pinus taeda clone PtaAGP5
BX681070	RS52G08	CR354679 BX251017	Ge124G07 Ge044C04	Trans-cinnamate 4-monooxygenase (EC 1.14.13.11)
BX681080	RS53A02	BX681080	RS53A02	Acetylornithine deacetylase (EC 3.5.1.16)
BX681136	RS54B08	BX681136	RS54B08	Dehydration stress-induced protein
BX681189	RS55B01	BX681189	RS55B01	putative protein
BX681928	RS66H05	BX681928	RS66H05	Antimicrobial peptide
CR392910	RN49H03	BX677194	RN19A11	Germin-like protein
CR393156	RN53C07	CR393156	RN53C07	Peroxidase precursor (EC 1.11.1.7)
CR393209	RN53H11	CR393209	RN53H11	Heat shock protein STI1
CR393223	RN54B06	CR393223	RN54B06	Heterochromatin-like protein
CR393241	RN54D04	CR393241	RN54D04	Plasma membrane intrinsic polypeptide
BX250268	Ge034E09	BX250268	Ge034E09	putative protein
BX680424	RS43C10	CR393430	RN60A09	Ethylene responsive element binding factor
BX682869	Ge122H09	BX682869	Ge122H09	Pinene synthase (EC 4.2.3.14)
CR392825	RN48H03	BX250552	Ge038D08	Epimerase/dehydratase

En ce qui concerne le **gène de contrôle**, après avoir testé trois gènes potentiels issus de l'analyse des données des puces et au vu de la stabilité d'expression sur les parties racinaires de pin maritime en PCR quantitative, notre choix s'est porté sur le clone **Ge066D02**. L'EST porte l'accession **BX252550** et la séquence est homologue à une **protéine ribosomale 40S S27**.

D.3.3.2 Niveau d'expression des gènes

Les données d'expression obtenues par PCR quantitative ont été mesurées pour l'ensemble des échantillons du SC et du stress long SL. Par contre, les niveaux d'expression calculés à partir des données d'hybridation sont représentés sur deux graphiques séparés, vu que les données ont été normalisées, ajustées et traitées indépendamment pour les deux types d'expérience. Les figures de l'annexe Ax-3 correspondent aux gènes détectés en SC et en SL, tandis que les gènes détectés lors de l'expérience SL ou SC uniquement sont présentés respectivement dans les figures des annexes Ax-4 et Ax-5. La valeur d'expression indiquée en ordonnée correspond au ratio d'expression par rapport à l'échantillon le plus faiblement exprimé pour lequel le ratio est égal à 1. Une expression de 4 signifie que l'échantillon accumule 4 fois plus de transcrits que l'échantillon dont l'expression est la plus faible.

D.3.3.2.1 Gènes vérifiés conjointement en stress court et stress long

Les résultats de PCR quantitative confirment les données d'expression des puces pour les gènes RN49H03, RS06C03, RS15A09, RN54D04 et RS34A05 (Fig. Ax-3). Les sens de variation et les profils d'expression sont conservés mais l'amplitude de la régulation ne l'est pas. Les niveaux de régulation de ces gènes obtenus par analyse des données d'hybridation atteignaient un maximum de 6 fois pour le stress court et de 3.5 fois pour le SL. Les résultats de PCR quantitative sur ces mêmes gènes indiquent des expressions régulées par un facteur pouvant atteindre 20 fois.

A titre d'exemple, nous avons pu vérifier que l'accession CR392910 (RN49H03), homologue à une « Germin-like protein », est plus exprimée dans l'EM où l'expression diminue lors du stress court et augmente légèrement lors du SL. Dans l'EF, l'expression est plus faible et pratiquement non régulée. La fonction exacte de ce type de protéines n'est pas connue mais il a été rapporté des activités superoxyde dismutase (Segarra *et al.* 2003) et une régulation de leur expression au cours du stress notamment chez l'orge où certains gènes sont surexprimés dans les racines de jeunes plants élevés en hydroponie et soumis à un stress salin (Hurkman *et al.* 1991).

Le gène RS06C03 (AL751204) (Fig. Ax-3B) annoté protéine T29M8.12 -à activité mono-oxygénase- a été assemblé par le TIGR avec une autre séquence et le consensus présente une homologie avec une Dimethylaniline mono-oxygenase (N-oxide-forming) (EC 1.14.13.8) aussi appelée FMO (Flavin containing mono-oxygenase). Ce gène aurait donc un rôle possible dans le maintien de l'équilibre redox de la cellule. Dans notre cas, le niveau de transcrits diminue légèrement avec le stress chez les 2 écotypes mais est supérieur chez l'EM.

Le gène RS15A09 (BX678783) (Fig. Ax-3C) qui code une peroxydase est également régulé négativement au cours du stress et plus exprimé chez l'EM. Ce type de régulation n'est pas forcément en contradiction avec la surexpression des peroxydases souvent rapportée lors de stress. Les peroxydases permettent d'éviter les oxydations non contrôlées par les AOS (Active oxygen species) qui sont normalement produites par de nombreux processus métaboliques (Shigeoka *et al.* 2002). Dans la cellule, différentes formes de peroxydases contrôlées par des gènes distincts sont présentes. L'expression de certains membres de cette famille pourrait être co-régulée avec les voies métaboliques génératrices d'AOS tandis que d'autres seraient liés à la réponse au stress. Dans le cas présent, ce gène indique peut être une diminution de l'activité d'une voie métabolique associée.

Le gène RN54D04 (CR393241) (Fig. Ax-3D) est homologue à des protéines de membrane et à un polypeptide de 22 kDa (DREPP4) (O24124) identifié chez le tabac et régulé lors du développement. L'expression diminue lors du stress et les données de PCR quantitative indiquent un niveau plus élevé chez l'EM, nettement marqué dans l'expérience SL.

L'expression du gène RS34A05 (BX679792) (Fig. Ax-3E) en PCR quantitative confirme pratiquement les données des puces. En effet, le sens de variation est conservé dans le SL avec une sous expression en stress et un niveau plus élevé chez l'EM. En revanche, en SC, le niveau d'expression de la provenance marocaine à 48H est plus bas que prévu. Compte tenu des écart-types des données d'hybridation, nous pouvons considérer que les résultats des deux méthodes ne sont pas très éloignés. Ce gène homologue au gène PrMC3 mis en évidence au niveau des cônes males de *Pinus radiata* (Walden *et al.* 1999) est donc régulé négativement par le stress et plus exprimé chez le marocain en condition de stress.

Les résultats de PCR quantitative des gènes Ge005C11 (Fig. Ax-3F), RS25D07 (Fig. Ax-3G), Ge110B02 (Fig. Ax-3H), RS54B08 (Fig. Ax-3I) et Ge085G01 (Fig. Ax-3J) ne sont par contre pas similaires aux résultats obtenus par microarrays. Nous remarquons que ces gènes, excepté RS25D07 (Fig. Ax-3G), ont une amplitude de régulation entre l'échantillon le

moins exprimé et le plus exprimé comprise entre 3 et 7 fois. Ces gènes sont donc plus faiblement régulés que les gènes cités précédemment dont l'amplitude était au minimum de 10 fois et au maximum de 20 fois.

Si on considère les données du SL, ces gènes (sauf RS54B08) appartiennent aux clusters 10, 11 et 13 générés par SOTArray qui correspondent aux très faibles variations d'expression des données d'hybridation. Cela indiquerait que nos données et l'analyse statistique ne permettraient pas d'estimer correctement les régulations de faible amplitude.

En SC, ils appartiennent tous (sauf RS54B08) au cluster 3 d'Expander et au groupe 4 de SOTArray. Considérant uniquement les données de SC, compte tenu des faibles régulations et des écart-types, les effets condition et provenance sont confirmés pour Ge005C11 (BX255395) (Fig. Ax-3F) pour lequel, en SC, l'EM exprime plus le transcrit et l'échantillon 48H est effectivement le plus exprimé. Le gène RS25D07 (BX679434) (Fig. Ax-3G) détecté pour un effet provenance présente bien un niveau plus élevé dans l'EM, même si l'expression de l'échantillon « Mc48H » apparaît comme plus exprimé que ce qu'indiquaient les donnés des puces. Les résultats de qPCR des trois autres gènes (Fig. Ax-3 H, I, J) ne correspondent par contre absolument pas à ceux des puces. Pour Ge085G01 (BX253595) (Fig. Ax-3J) cela peut s'expliquer par l'amplification d'une séquence ne correspondant pas au gène attendu sachant que le dessin des amorces a été réalisé sur les EST RS41F03 (BX680293) et RS30B09 (BX679631) du même contig. Pour les gènes Ge110B02 et RS54B08, cela peut provenir d'une hybridation croisée d'ADNc d'autres gènes sur le spot ou de l'amplification en PCR quantitative d'un autre membre d'une famille multigénique tandis que le spot traduirait l'expression cumulée de l'ensemble des membres de la famille. Ge110B02 (Fig. Ax-3H) est homologue à une enzyme de conjugaison à l'Ubiquitine et RS54B08 à (Fig. Ax-3I) une protéine induite par la déshydratation.

Dans le cas du gène RS54B08 (BX681136) (Fig. Ax-3I), les profils d'expression obtenus en qPCR et indiquant une surexpression des transcrits lors du stress, correspondent mieux à ce à quoi nous pouvions nous attendre d'après l'annotation. La régulation de l'expression diffère complètement entre les deux écotypes, l'EF maintient une expression élevée quelque soit la condition tandis que l'EM augmente son expression uniquement lors du stress. La détermination des amorces ayant été réalisée sur l'EST lui-même et l'amplification étant spécifique, les données des puces pourraient donc résulter d'une hybridation de la sonde avec des ADNc d'un autre membre de la même famille.

L'expression des gènes supplémentaires confirme en partie, pour trois d'entre eux, les résultats d'hybridation.

Le gène Ge034E09 (BX250268) (Fig. Ax-3K), sans homologie connue, présente bien un effet provenance marqué en SL comme en SC avec une expression de 30 à plus de 60 fois supérieure pour l'EF. De plus les résultats de PCR quantitative indiquent nettement une surexpression lors du stress chez l'EF aussi bien à 48 heures qu'en SL. Ce gène semble bien être caractéristique de l'EF et impliqué dans la réponse au stress chez cet écotype.

Le gène RN48H03 (CR392825) (Fig. Ax-3L), homologue à une épimérase/déshydratase, présente bien un effet condition avec sous expression en SL. En SC, l'effet condition mis en évidence sur les données d'hybridation avec de forts écart-types est moins net et le profil d'expression en qPCR de l'EM avec une forte expression à 6H diffère de celui des puces.

Le gène Ge122H09 (BX682869) (Fig. Ax-3M), codant une pinène synthase, est bien sous exprimé en SL et en SC même si le profil d'expression de l'EM en SC est caractérisé par une surexpression à 6 heures. L'effet provenance qui avait été détecté en SC sur les données de puces n'est par contre plus observé.

Nous pouvons faire le même constat pour RS43C10 (BX680424) (Fig. Ax-3N) codant un facteur de liaison aux éléments de réponse à l'éthylène (EREBP). Les données d'expression du stress court ne concordent pas pour l'EM.

Il est intéressant de noter que les 3 gènes des figures Ax-3 L, M et K présentent tous une surexpression transitoire forte à 6H chez l'EM avant de diminuer leur expression tandis que l'EF, qui a une expression basale plus forte, n'a pas se comportement. Ces gènes pourraient donc intervenir dans la réponse « immédiate » au stress et refléter des stratégies différentes. **L'EM adopterait une régulation précoce plus efficace, ce que les résultats précédents avaient également suggéré.**

La différence de profil d'expression peut s'expliquer par un dessin des amorces sur les EST RN60A09 (CR393430) pour RS43C10 et Ge038D08 (BX250552) pour RN48H03. Cependant, la surexpression à 6 heures chez l'EM détectée en PCR quantitative et non présente dans les données d'hybridation se retrouve pour plusieurs gènes (Ge085G01, Ge110B02, Ge122H09, RN48H03, Ge110B2). Cela pourrait indiquer aussi un biais de normalisation des données d'hybridation.

Si l'on considère maintenant les profils d'expression des gènes obtenus en PCR quantitative, nous pouvons remarquer que pratiquement tous les gènes retenus répondent au stress et présentent des régulations différentes entre les deux écotypes. **L'objectif initial d'identifier**

des gènes répondant au stress court et au stress long mais distinguant les deux provenances est donc atteint. Par contre, la majorité des gènes présentent une diminution de l'expression lors du stress. Ce type de régulation peut résulter d'une réponse passive au stress par ralentissement des voies métaboliques touchées lors de la réaction de la plante face au stress. Mais la diminution de l'expression d'inhibiteur peut entrainer aussi une activation de certains enzymes et des produits associés. Les surexpressions à 6H suivies d'une sous-expression suggèrent plus une implication possible des gènes concernés dans la réponse immédiate de maintien de l'homéostasie plutôt qu'une réponse d'ajustement de l'expression pour atteindre le nouvel état d'équilibre en stress.

D.3.3.2.2 Gènes vérifiés en stress long

Sur la figure Ax-4, nous pouvons voir que le gène RS66H05 (BX681928) (Fig. Ax-4A) présente effectivement un effet provenance et condition. Ce gène homologue à un peptide antimicrobien est surexprimé lors du stress et son niveau de transcrit est plus élevé dans l'EF.
Le gène RS40E02 (BX680204) (Fig. Ax-4B), détecté pour l'accumulation de ses transcrits avec le stress, présente bien ce comportement mais l'amplitude de la régulation est plus faible que celle prédite par les données d'hybridation.
La régulation négative de l'expression a pu être vérifiée par PCR quantitative pour les gènes RS43C07 (BX680432) (Fig. Ax-4C) –no hit-, RN05B11 (AL750591) (Fig. Ax-4D) -Xyloglucan:xyloglucosyl transferase (EC 2.4.1.207)-, Ge071H04 (BX252860) (Fig. Ax-4E) -5-methyltetrahydropteroyl triglutamate--homocysteine methyltransferase-, Ge093C03 (BX254093) (Fig. Ax-4F) -caffeate O-methyltransferase- et RS52G08 (BX681070) (Fig. Ax-4G) -Trans-cinnamate 4-monooxygenase (EC 1.14.13.11)-.
Les gènes Ge085E07 (BX253583) (Fig. Ax-4H) –Endoglucanase-, AS03B12 (AL749965) (Fig. Ax-4I) -protéine Metallothionein-like EMB30- et AS03A12 (AL749954) (Fig. Ax-4J) -Thimet oligopeptidase (EC 3.4.24.15)- détectés pour leur régulation négative présentent également cette réponse en PCR quantitative mais les niveaux d'expression entre provenances et conditions ne sont pas totalement conservés. Le gène AS03B12 (AL749965) ne correspond plus à nos critères de sélection initiaux, l'expression du gène étant régulée de la même manière pour les deux provenances.
Les gènes RS55B01 (BX681189) (Fig. Ax-4K) –*protéine putative*- et RN44F10 (BX677784) (Fig. Ax-4L) –protéine chaperonne à cuivre pour la SOD- sont bien surexprimés lors du stress mais leurs niveaux d'expression ne correspondent pas exactement entre les 2 techniques. Par contre, pour ces deux gènes, la régulation diffère effectivement entre les deux provenances.

Deux gènes ont des expressions non vérifiées par PCR quantitative. Il s'agit d'une part du gène RN53H11 (CR393209) (Fig. Ax-4M) codant une protéine de choc thermique et d'autre part du gène RN54B06 (CR393223) (Fig. Ax-4N) homologue à une protéine « heterochromatine-like ». Les transcrits amplifiés par PCR présentent dans les deux cas une inhibition de l'expression par le stress. Comme évoqué dans le paragraphe précédent, nous retrouvons ces deux gènes dans les groupes 10 et 11 de SOTArray qui correspondent à des faibles régulations, ce que confirme la qPCR. La détection de la régulation des gènes classés dans ces groupes semble encore une fois poser problème. Si les résultats de qPCR ne confirment pas ceux des puces, les profils obtenus sont par contre très intéressants car l'EM surexprime ces gènes de réponse au stress en SL par rapport à l'EF. Une vérification de leur comportement dans différentes populations pourrait être envisagée.

D.3.3.2.3 Gènes vérifiés en stress court

A la figure Ax-5, les profils d'expression obtenus pour les gènes RN53C07 (Fig. Ax-5A), RS11G02 (Fig. Ax-5B), RN18B10 (Fig. Ax-5C) et RS53A02 (Fig. Ax-5D) confirment parfaitement les données d'hybridation. Nous avons donc des gènes présentant un effet condition et provenance simultané. Les profils d'expression des gènes RN53C07 (CR393156) et RS11G02 (BX678596) codant des précurseurs de peroxydases sont pratiquement identiques. Le gène RN18B10 (BX677134) sans homologie connue a une expression pratiquement constante, excepté pour l'EM à 48H qui surexprime ce gène d'un facteur 8, ce qui en fait un gène candidat potentiel. Ce type de régulation peut en effet annoncer une activation de mécanismes de réponse au stress de l'écotype résistant.

Le gène RS53A02 (BX681080) (Fig. Ax-5D) homologue à une Acetylornithine deacetylase (EC 3.5.1.16) est faiblement régulé avec une légère augmentation des transcrits à 48H chez l'EM et une diminution de l'expression à 6 heures chez l'EF. Bien que l'amplitude de la régulation ne soit pas très élevée, l'expression induite à 48H chez l'EM accompagnée d'un niveau plus élevé chez l'EM en SC et en SL en fait un gène potentiellement intéressant. De plus, cet enzyme permet, à partir de l'acethylornithine, de produire de l'ornithine qui par décarboxylation peut donner de la putrescine et par conséquent de la spermidine et de la spermine. L'implication de ces composés dans la réponse au stress renforce l'idée d'un rôle possible de ce gène dans la différence de résistance entre les 2 écotypes.

Le gène RS12A03 (BX678613) (Fig. Ax-5E) –homologue à un précurseur de gamma-thionin qui intervient dans la réponse au stress en tant que Défensine-, détecté pour un effet

provenance, a bien un niveau de transcrit supérieur dans l'EM mais l'amplitude de la régulation est très faible et n'atteint que 2 fois en qPCR.
Les gènes Ge005B04 (BX255377) (Fig. Ax-5F) et Ge032H07 (BX250140) (Fig. Ax-5G) n'ont pas été validés. Les profils diffèrent totalement. Pour le gène Ge005B04, homologue à un précurseur de l'élastine, l'amplification PCR indique une diminution de l'expression à 48H et une régulation identique entre les deux écotypes, excepté en SL où la diminution d'expression est moins forte dans l'EM. Une surexpression d'élastine en SL chez l'EM pourrait être une stratégie de réponse de l'EM au stress hydrique. Pour Ge032H07 –gène de défense contre la mort cellulaire-, la régulation est faible. L'expression ne varie pas en SC pour l'EM et une surexpression apparaît à 6H pour l'EF.

Pour plus de la moitié des gènes sélectionnés, nous avons donc vérifié les résultats obtenus avec les puces. Les gènes présentent une régulation différente de leur expression lors du stress selon l'écotype. Cependant, le contraste d'expression entre écotypes lors du stress (niveau de base d'un écotype correspondant au niveau en stress de l'autre) qui avait été observé sur le clustering des données des puces et avait donné lieu à la sélection de ces gènes n'est pas aussi marqué en qPCR. Le choix de gènes candidats potentiels sur la base des données de qPCR devra donc être plus nuancé et tenir compte des profils où les 2 écotypes régulent différemment lors du stress.

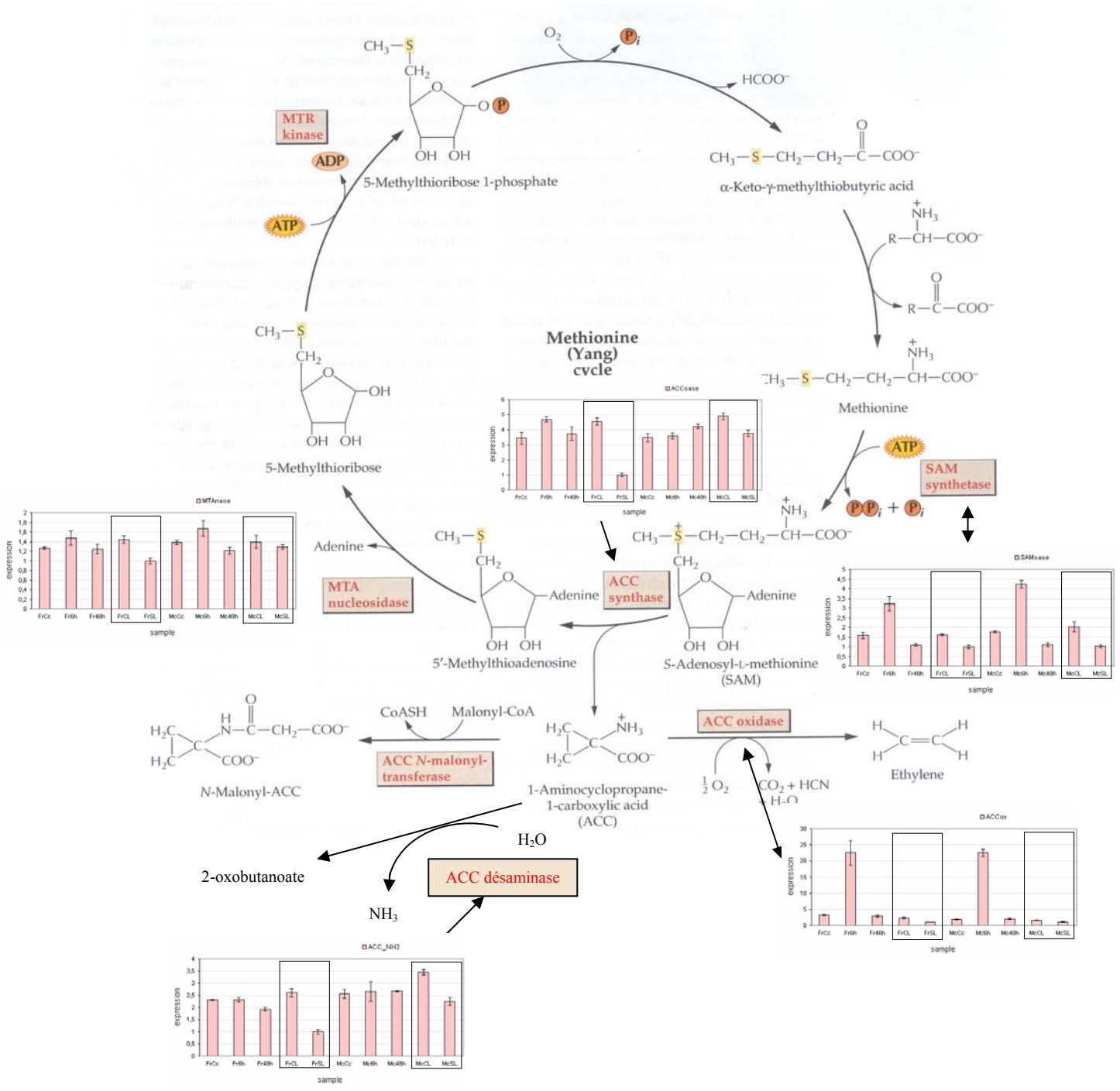
D.3.3.2.4 Gènes de la biosynthèse de l'éthylène

L'expression des gènes codant les enzymes de la voie de biosynthèse de l'éthylène est indiquée à la figure D-46. Nous allons nous intéresser tout d'abord aux enzymes du cycle de la méthionine.

L'expression de la SAM synthétase varie entre 1 et 2 fois entre les provenances et les conditions de stress, excepté pour le stress à 6H où l'EF augmente son expression de 1.5 pour le contrôle à 3 fois et l'EM passe de 1.75 à plus de 4 fois. L'expression du contrôle long est légèrement plus élevée pour l'EM (2 vs 1.5) mais en SL la quantité de transcrits est identique entre les deux provenances.

L'ACC synthase est plus fortement régulé, surtout pour l'EM. En SC, cet écotype surexprime légèrement le gène à 6H tandis qu'en SL la transcription est réduite de 4.5 fois. L'EM maintient une expression proche de 3.5 fois avec une légère augmentation de la quantité de transcrits pour le contrôle long atteignant 4.5 fois.

Figure D-46 : Voie de biosynthèse de l'éthylène et expression des gènes codant les enzymes associés.



L'expression de la MTA nucléosidase est relativement stable autour de 1.3 fois. Les valeurs extrêmes concernent l'EF en SL et l'EM à 6H de stress qui atteignent respectivement 1 fois et 1.7 fois.

Si l'on considère les expressions des transcrits de ces 3 enzymes, en supposant que la synthèse protéique ainsi que l'activité enzymatique est corrélée à l'accumulation des transcrits, nous pouvons nous attendre à une accumulation d'ACC à 6 heures de stress pour les deux provenances. Cette accumulation devrait être légèrement plus élevée chez l'EM. Nous devrions également avoir une forte diminution d'ACC chez l'EF en SL.

Cet ACC est le substrat des réactions faisant intervenir l'ACC oxydase et l'ACC désaminase. L'expression de l'ACC désaminase varie pour l'expérience SL. L'EF diminue l'expression du gène par un facteur 2.5. Les transcrits sont plus exprimés chez l'EM en contrôle long (3.5 fois) et leur quantité diminue avec le stress pour atteindre 2.25 fois.

La production de transcrits homologues à l'ACC oxydase augmente fortement chez les deux provenances à 6H de stress pour atteindre un niveau d'expression proche de 23 fois. Nous nous attendons donc à une surproduction massive d'éthylène à ce stade dans les racines. La régulation en stress long semble indiquer une quantité de transcrits supérieure pour l'EF et une diminution de l'expression lors du stress. Cela ne se répercute par contre pas forcément par une diminution proportionnelle de la production d'éthylène. En effet, l'expression de l'ACC désaminase est également réduite lors du stress ce qui signifie que la quantité d'ACC disponible pour l'ACC oxydase est augmentée ce qui peut maintenir partiellement l'efficacité de la réaction de production de l'éthylène.
En étudiant les données d'hybridation, nous avions cité une accession (BX251044) codant une ACC oxydase dont l'expression augmentait lors du stress long ce qui n'est pas le cas ici. Il faut rappeler que plusieurs consensus du TIGR sont homologues à l'ACC oxydase et que deux regroupent des séquences de pin maritime. L'accession (BX251044) appartient au TC73594 tandis que la séquence amplifiée en PCR quantitative est issue du TC58524. Il serait intéressant de dessiner des amorces sur l'autre consensus. Il existe donc probablement deux gènes distincts codant l'ACC oxydase qui peuvent être régulés par des voies distinctes.

D.3.4 Discussion

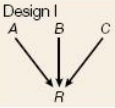
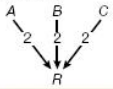
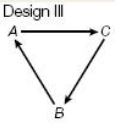
Pour cette étude d'expression, nous avons adopté un plan d'expérience complexe en « loop design ». Le choix de ce type d'expérience est dicté d'une part par le nombre de lames disponibles et d'autre part par l'avantage en terme de puissance de détection des variations

d'expression d'un tel dispositif. En effet, l'organisation des hybridations va influer sur la variance inhérente à l'expérience pour l'estimation des effets testés (Speed 2003). La figure D-47, tirée d'une publication de Yang et Speed (2002) montre la variance obtenue pour différents plans d'expérience. Il apparaît clairement que pour un même nombre de lames, un « loop design » permet d'estimer un effet avec une variance réduite. Dans le cas d'une expérience à un seul facteur, la variance moyenne des estimateurs passe de 2 à 0.67, tandis que dans le cas d'un plan factoriel 2x2 le gain en terme de variance est obtenu en maintenant une variance réduite quelque soit l'effet testé. Toutefois, ce type de dispositif présente un inconvénient majeur. La diminution de la variance découle de la possibilité de tester l'effet d'un facteur par comparaison directe et indirecte. Sur le design III à gauche de la figure D-47, cela se traduit par la possibilité de comparer A et B en tenant compte des données de l'hybridation de A et B sur la même lame ou de comparer A et B en comparant A et C puis C et B. Ce dispositif circulaire est très sensible à la perte de données car la perte des données d'une lame va rompre la boucle. Plus le nombre de lames va augmenter, plus le risque de perdre des données pour un gène va augmenter. Il est possible dans ce cas de réaliser des hybridations croisées comme dans le design III à droite de la figure D-47 afin de limiter le nombre d'étapes de la voie de comparaison indirecte.

La stratégie de normalisation des données n'a pas été fixée dès le début de l'analyse. Différentes approches ont été testées avant d'adopter celle décrite dans la thèse. C'est en étudiant la structuration des données que nous avons identifié successivement les facteurs correctifs à appliquer. Les deux étapes successives de correction des données en sont l'illustration. La première corrige les variations intra-lame et la seconde les variations inter-lames. Les données de microarrays comportent de nombreux biais qu'il faut absolument corriger avant toute interprétation. J'insiste particulièrement sur ce point qu'est l'analyse de la qualité des données sur lesquelles vont être réalisées les tests statistiques pour détecter des gènes différentiellement exprimés. La normalisation et la correction des données vont dépendre du plan d'expérience et des questions auxquelles nous voulons répondre. C'est pourquoi le modèle global de normalisation inclus, entre autres, les effets techniques liés à la provenance, condition et interaction provenance x condition pour fixer un effet global sur l'ensemble des gènes. Chaque gène est dans ce cas considéré comme une répétition. Il aurait pu être intéressant d'appliquer un modèle mixte avec certains facteurs comme l'effet lame comme aléatoire mais ce type de modèle est très mal pris en charge par le logiciel de statistiques R. Pour la même raison, il n'a pas été possible d'appliquer un modèle d'analyse de variance unique en incluant un effet « gène » ce qui aurait permis de fixer les variations du

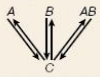
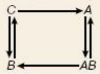
Figure D-47 : Impact du plan d'expérience sur la puissance de détection des variations d'expression.

Single-factor experiments

Design choices	Number of slides	Units of material (number of samples)	Average variance
Indirect designs			
Design I 	3	$A = B = C = 1$	2.00
Design II 	6	$A = B = C = 2$	1.00
Direct design			
Design III 	3	$A = B = C = 2$	0.67

Variance of estimated effects for three different designs of single-factor experiments. σ^2 was set to 1 throughout.

2 × 2 factorial experiments

Design choices	Main effect *A*	Main effect *B*	Interaction *A.B*
Indirect design			
Design I 	0.50	0.50	1.50
A balance of direct and indirect design			
Design II 	0.67	0.43	0.67
Design III	0.50	0.50	1.00
Design IV	N/A	0.30	0.67

Variance of estimated effects for four different designs of 2 × 2 factorial experiments. σ^2 was set to 1 throughout.

modèle global tout en effectuant la détection des gènes différentiellement exprimés. Nous avons donc opté pour une analyse en 2 temps.

L'étude de la réponse au stress à court terme et à long terme a permis d'accéder à plusieurs types d'information. Le premier constat se situe au niveau de l'effet significatif détecté. Dans l'expérience stress court, les gènes sont principalement détectés pour un effet provenance tandis que dans l'expérience stress long la détection concernait surtout un effet condition. Comme nous avons pu le voir, si un biais des données concernant les provenances n'est pas à exclure pour l'expérience stress court –3 séries d'hybridations distinctes pouvant être à l'origine d'un effet « série » partiellement confondu à l'effet provenance-, on peut toutefois s'attendre à ce que la part de variation d'expression des gènes due au facteur condition soit plus faible dans le stress court que dans le stress long. Seulement 48 heures séparent les plants stressés des plants contrôles dans un cas, contre 3 semaines dans l'autre. L'étude de la réponse précoce est cependant informative en termes d'effet condition et de dynamique de réponse du transcriptome aux variations brutales de l'environnement. Le pin maritime répond dès 6 heures de stress en régulant l'expression de certains gènes comme ceux impliqués dans la synthèse de l'éthylène. Nous avons aussi relevé des gènes dont le niveau d'expression diffère entre les deux provenances et peut refléter une adaptation d'un écotype à son environnement.

Parmi les gènes détectés, nous avons également retrouvé des gènes connus pour leur implication dans la réponse au stress ce qui conforte la démarche technique et l'analyse des données réalisée. Il reste toutefois délicat de sélectionner les gènes différentiellement exprimés sur la seule base des p-value et des coefficients de corrélation du modèle.

Un modèle d'analyse de variance va tester l'identité de deux groupes sur la base de leur moyenne et des variances intra et inter groupes. Les gènes dont les profils d'expressions sont opposés mais pour lesquels les moyennes des 2 provenances ainsi que celles des conditions sont identiques, risquent de ne pas être détectés si les variances sont importantes. L'analyse de variance seule, ne rend pas non plus compte dans quelle mesure un échantillon diffère des autres.

L'étude des profils d'expression est donc indispensable à l'interprétation correcte des résultats. Le regroupement des données par des méthodes de classification est un outil visuel supplémentaire pour sélectionner les gènes sur la base de leur profil d'expression et de leur norme de réaction.

Considérant les résultats de PCR quantitative, il serait peut être judicieux de ne pas prendre en compte les gènes détectés pour des faibles amplitudes de régulation qui sont regroupés comme nous l'avons vu dans les clusters 10 à 13 par SOTArray en stress long.

Les données de PCR quantitative ont également souligné la nécessité d'une vérification des données d'expression par deux techniques différentes. La technologie des puces à ADNc pose le problème de la spécificité de l'hybridation des sondes sur les cibles :
1) plusieurs membres de familles multigéniques sont susceptibles d'hybrider sur un même spot ; 2) des ARNm codant pour des enzymes de fonctions proches avec des séquences proches peuvent également être à l'origine d'hybridations non spécifiques ; 3) des écotypes différents dont la séquence d'ADN diverge -même au niveau du même membre- peuvent entrainer un différentiel d'hybridation dû à la variation allèlique, si bien que les puces d'Affymetrix sont également utilisées pour génotyper des individus (Butcher *et al.* 2004; Matsuzaki *et al.* 2004a; Matsuzaki *et al.* 2004b).
A l'opposé, la PCR quantitative est très spécifique ce qui peut expliquer que l'expression des gènes n'a pas toujours été confirmée. Pour de nouvelles études, il serait intéressant d'effectuer deux dépôts par gène, l'un couvrant la région 5' et l'autre la région 3' afin de pouvoir estimer la spécificité du signal de fluorescence. Les puces à oligomères (60mers) constituent une excellente alternative, quoique très coûteuses, car elles présentent les mêmes avantages de spécificité que la PCR quantitative.

Ce type d'analyse d'expression à haut débit sert de point de départ pour orienter les recherches sur un nombre plus réduit de gènes présentant des régulations pour le caractère étudié. Nous avons mis en évidence des gènes de fonctions connues et d'autres sans homologie avec des séquences connues. Dans les deux cas, c'est la régulation de leur expression différente entre les deux écotypes qui fait de ces gènes des candidats potentiels pour l'adaptation au stress osmotique au stade juvénile. Il restera à vérifier que ce sont toujours de bons candidats pour le stress hydrique en menant des études d'association de ces gènes à des caractères impliqués dans la réponse au stress hydrique sur des populations contrastées pour leur adaptation à la sécheresse.
A partir de cette masse de données d'expression, il est également possible d'adopter une démarche plus ciblée en partant d'un gène régulé codant un enzyme de fonction connue et en étudiant les enzymes des voies biochimiques dans lesquelles le gène d'intérêt intervient. C'est ce que nous avons tenté de faire avec les gènes codants pour les enzymes de la voie de synthèse de l'éthylène.

D.4 Analyse du protéome

D.4.1 Détection des protéines

Avant de comparer les résultats obtenus dans le cadre de cette thèse avec les données disponibles sur le protéome, il est important de préciser quelques points sur la technique d'électrophorèse bidimensionnelle utilisée pour détecter les protéines. L'intégralité du protéome n'est pas accessible. Seule une fraction, déterminée par la solubilité et le point isoélectrique des protéines, peut être étudiée. Une autre limite concerne l'abondance des protéines et la taille des spots visibles avec les techniques de coloration. Il faut également garder à l'esprit qu'un spot ne correspond pas forcément à une seule protéine. Plusieurs protéines peuvent migrer dans le gel au même emplacement. La taille du spot détecté reflètera donc la quantité de la protéine majoritaire et des éventuelles protéines en mélange. A l'inverse, le produit de la traduction des transcrits d'un gène donné de migrera pas uniquement au niveau d'un spot unique. Des allèles différents, des produits d'épissages alternatifs ou des membres de familles multigéniques pourront être à l'origine de spots distincts. De plus, les modifications post-traductionnelles comme des glycosylations vont influer sur la migration et la position du spot correspondant. Les variations observées devront donc être interprétées comme des modifications de l'expression d'une des formes possible de la protéine.

D.4.2 Identification des spots protéiques différentiellement exprimés

Les protéines des tissus racinaires des plants de l'expérience « stress long » ont été extraites et analysées par électrophorèse bidimensionnelle par Manon Moreau dans le cadre de son stage de Master 2. Après détection et nettoyage des gels, 1201 spots ont été retenus. L'analyse de variance a donné les résultats suivants : 696 spots ont au moins un effet significatif, 41% des spots présentent un effet provenance, 39% ont un effet traitement et 31% montrent un effet d'interaction (Fig. D-48)(Moreau 2005). Le choix des spots à séquencer, effectué essentiellement sur la base d'un effet d'interaction et de la taille du spot, a abouti à la sélection de 192 spots qui ont été prélevés et séquencés. L'annotation et le profil d'expression des spots ne présentant pas de mélange de peptides et pour lesquels le résultat de la recherche d'homologie sur les bases de données SwissProt et TIGR était sans ambiguïté, sont listés dans le Tableau Ax-12 (accessible également en interactif par l'URL : http://www.pierroton.inra.fr/biogeco/genetique/theses.html).

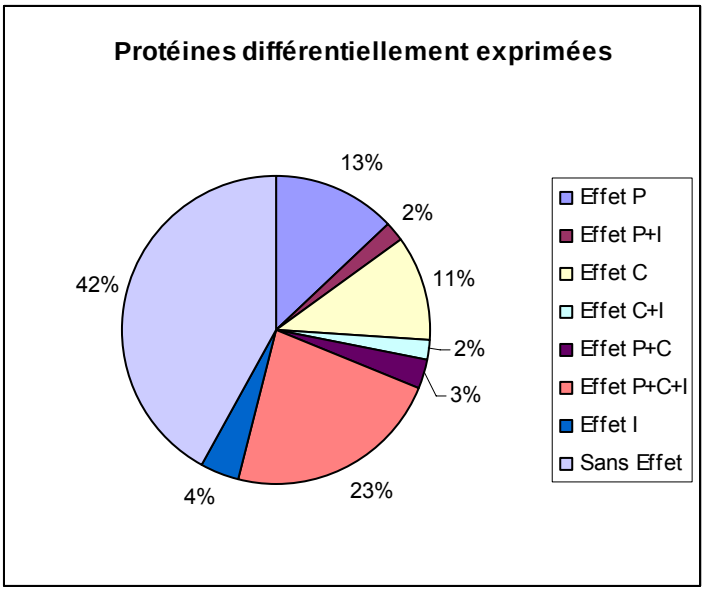


Figure D-48 : Résultats des analyses de variances pour la détection des protéines différentiellement exprimées. P = Provenance, C = Condition et I = Interaction. (d'après Moreau (2005))

Ces résultats diffèrent nettement de ceux que nous avons obtenus lors de l'analyse des microarrays. En effet, 23 % des spots présentent un effet triple provenance, condition et interaction tandis que l'effet d'interaction était peu détecté sur les données des transcrits. Même si la détection d'un effet d'interaction seul est faible (4%), cela reste nettement supérieur aux résultats des puces où l'effet d'interaction représente une part infime des effets détectés. Ces résultats peuvent provenir du traitement des données utilisées ou des gènes/protéines étudiés mais il est également probable que la déconnection entre la quantité de transcrits et la quantité de protéines accumulées se traduisent par un contraste plus marqué et donc détectable entre les stratégies de régulation des deux écotypes au niveau des protéines.

D.4.3 Annotation des spots protéiques

L'étape d'annotation fonctionnelle des spots séquencés est un exercice particulièrement délicat et diffère de celle des EST. L'attribution d'une fonction ne se fait pas par recherche d'homologie entre la séquence de la protéine et les bases de données publiques car la séquence complète de la protéine n'est pas connue. La LC-MS/MS donne une information sur la masse d'une partie des peptides obtenus après digestion et sur la masse des sous-peptides obtenus par fragmentation de certains d'entre eux. Ces informations permettent de rechercher dans les bases de données des protéines dont la digestion permettrait d'obtenir un spectre MS/MS similaire. En fonction du nombre de peptides permettant l'identification de la séquence (minimum de 2) et de leur annotation, lorsque c'est possible, une fonction est assignée à la protéine. De plus, contrairement aux EST où un clone contient un insert, plusieurs protéines peuvent être prélevées en mélange au sein d'un spot. Les peptides obtenus peuvent donc provenir de protéines différentes de fonctions différentes. L'attribution d'une fonction est donc beaucoup plus délicate que dans le cas des transcrits. Les recherches d'homologie étaient classiquement faites contre la base de données SwissProt. Cette base, comme nous l'avons déjà exposé est constituée de protéines très bien annotées mais en nombre limité. Considérant les bases d'EST utilisées pour les études de transcrits, j'ai suggéré de compléter les recherches d'homologie sur la base Pinus du TIGR dont l'annotation est déterminée par recherche d'homologie contre des banques protéiques et qui permet de couvrir également des protéines non représentées dans la base SwissProt ou dont la séquence serait plus spécifique du genre Pinus. Il est donc possible de déterminer la fonction au vu des résultats des recherches de concordance de spectres MS/MS avec des séquences protéiques de la base SwissProt et de séquences traduites du TIGR. De plus, l'obtention d'un TC homologue permet de rechercher si un des EST le composant a été sélectionné pour l'Unigène

de pin maritime défini lors de cette thèse, ce qui facilite la comparaison entre les données d'expression protéique et transcriptomique. Les accessions déposées sur la lame, appartenant aux TC homologues aux peptides, ont été ajoutées dans la dernière colonne du tableau Ax-12.

D.4.4 Expression du protéome et du transcriptome

En analysant les fonctions des protéines annotées et détectées comme différentiellement exprimées (Tab. Ax-12), nous remarquons que l'on retrouve souvent des fonctions également rencontrées dans la liste des gènes détectés par microarrays. Le tableau D-13, extrait de cette liste (Tab. Ax-4), reprend les accessions de fonction commune. Nous comparerons les données de protéomique avec nos données de transcriptomique en suivant les classes fonctionnelles proposées. Les profils d'expression des spots sont accessibles par l'URL http://www.pierroton.inra.fr/biogeco/genetique/theses.html sur la version interactive du tableau Ax-12. Pour certains spots, une homologie a été trouvée pour un TC du TIGR contenant une accession déposée sur la puce et différentiellement exprimée. Dans ces cas précis, les profils d'expression des spots et des accessions correspondantes sont indiqués dans la figure D-50.

D.4.4.1 « Métabolisme des acides aminés »

Dans cette classe, nous retrouvons 4 spots (149/150/154/155) correspondant à une 5-methyltetrahydropteroyltriglutamate-homocystéine methyltransférase (EC 2.1.1.14). La présence de spots proches correspondant à cette fonction peut s'expliquer, comme évoqué précédemment, par la présence d'allèles différents et/ou de membres de familles multigéniques et/ou de modifications post-traductionnelles. Nous avions détecté et vérifié en PCR quantitative l'expression de l'accession Ge071H04 (BX252860) (Fig. Ax-4E) qui baissait lors du stress et était légèrement supérieure chez l'EF. Au niveau protéique, nous avons observé des comportements opposés entre les écotypes avec une accumulation chez l'EF et une baisse chez l'EM. Dans les deux études, nous observons donc une régulation des produits des gènes de cette famille. Cet enzyme (EC 2.1.1.14) permet la conversion de la L-homocystéine à la L-methionine (Fig. D-49). Une régulation de sa synthèse suggère un autre rôle que celui de production d'acide aminé pour la synthèse de protéines. Or, sous l'action de la methionine adenosyl-transferase (EC 2.5.1.6), la L-methionine donne la S-adenosyl-L-methionine. Trois spots (521/533/549) différentiellement exprimés pour cet enzyme ont été détectés. Deux sont induits par le stress chez l'EF et inhibés chez l'EM, et un spot (533) présente une diminution d'accumulation pour les 2 écotypes lors du stress comme l'accession (BE643924) que nous avons analysée en PCR quantitative. Une des utilisations de la S-

adenosyl-L-methionine dans la cellule est de servir de donneur de méthyle dans la réaction de méthylation de l'ADN par l'ADN methyltransférase (EC 2.1.1.37) qui transfère le groupement méthyle de la S-adenosyl-L-methionine sur la cytosine pour donner la 5-methyl-cytosine. Une hypothèse serait donc une régulation de la production de S-adenosyl-L-methionine permettant la méthylation de l'ADN qui serait régulée lors du stress. Les résultats du protéome indiqueraient une accumulation de méthionine chez l'EM lors du stress et une diminution pour l'EM. Dans l'hypothèse d'une relation entre cette accumulation et l'activation de la méthylation, l'EF aurait donc une méthylation plus active (inhibition de gènes). Cependant, l'action d'une Adenosylmethionine décarboxylase (EC 2.5.1.16) sur la S-adenosyl-L-methionine et la putrescine donne la 5'-Methylthio-adenosine et la spermidine. Cet enzyme fait également partie des spots différentiellement exprimés (spot724) et n'est détecté que dans l'EM contrôle. Une autre hypothèse serait donc la régulation de la synthèse de spermidine qui intervient dans la réponse au stress et pour laquelle nous avons détecté des gènes régulés dans l'analyse du transcriptome.

D.4.4.2 « Biosynthèse des métabolites secondaires »

Les spots de cette catégorie indiquent une régulation des enzymes intervenant dans la voie de synthèse des composés terpéniques. Comme dans l'analyse du transcriptome, nous retrouvons une variation de la Geranyl diphosphate synthase (spot1344) et de la Pinène synthase (EC 4.2.3.14) pour laquelle 2 spots distincts ont été identifiés (spot1351/323). Les spots 1344 et 1351 ne sont visibles que dans l'échantillon stressé de l'EF, ce qui n'est absolument pas confirmé au niveau des transcrits homologues. L'accession BX682869 (pinène synthase) avait une expression plus forte chez l'EF en SL et diminuait en stress chez les 2 écotypes, et l'accession BX254083 (Fig. D-50O) indiquait une diminution d'expression en stress pour les deux écotypes.

Il faut également relever la présence d'un autre enzyme représenté par le spot 172 classé dans « métabolisme des lipides » homologue à une 4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase (EC 1.17.4.3) qui permet également de catalyser une réaction pouvant intervenir dans la voie alternative sans mévalonate de biosynthèse des terpènes dont nous avions parlé dans la synthèse bibliographique. Cette voie se pose comme la voie principale chez les plantes.

La détection de ces spots différentiellement exprimés, renforcée par la détection de transcrits de même fonction, suggère donc une régulation différente entre les deux écotypes

Tableau D-13 : Extrait du tableau Ax-4 des gènes dont la fonction correspond à celle des protéines différentiellement exprimées.
Pvalprov/cond/inter : *valeur de la probabilité associée au test pour les facteurs provenance, condition et interaction* ; r^2 prov/ cond/ inter/ tot/ PCI/ max : *part de la variabilité expliquée par les facteurs provenance (P), condition (C), intéraction (I) et pour la totalité du modèle, l'ensemble des facteurs P,C et I, et par le facteur expliquant le plus de variabilité*. Data > Bkg : nb de valeurs supérieures au bruit de fond.
Fact prov/cond : *niveau de régulation pour les facteurs P ou C.*

Gènes Différentiellement Exprimés dans l'expérience Stress Long

Acc Number	ID	Pvalprov	r2 prov	Pvalcond	r2 cond	Pvalinter	r2 inter	r2 tot	r2 PCI	r2 max	data > Bkg	Moy France	Moy Maroc	Moy Cont	Moy 3S	fact prov	fact cond	Annotation	catégorie fonctionelle
BX252228	Ge061F04	8,17E-01	0	4,01E-07	59	1,00E+00	0	67	60	59	20	9,413	9,306	8,959	9,782		1,77	(2R)-phospho-3-sulfolactate synthase (EC 4.4.-.-)	Non-classé
BX252860	Ge071H04	2,95E-01	6	1,62E-05	47	1,00E+00	2	59	55	47	35	10,824	10,516	11,078	10,324		1,69	5-methyltetrahydropteroyltriglutamate–homocysteine methyltra	Métabolisme
BX255342	Ge004F07	7,09E-01	2	9,75E-04	35	1,00E+00	0	49	37	35	29	10,055	10,113	10,412	9,777		1,55	Actin	Organisation cellulaire
BE607163	NXCI_37_F8	7,48E-01	1	1,71E-05	48	8,86E-01	7	58	56	48	34	10,463	10,565	10,854	10,153		1,63	adenosine kinase (AK)	Métabolisme
BG040863	NXSI_116_A9	9,36E-01	0	4,22E-04	40	1,00E+00	3	45	43	40	40	11,583	11,628	11,875	11,327		1,46	adenosine kinase (AK)	Métabolisme
BX249835	Ge029C07	9,80E-01	0	1,56E-04	44	1,00E+00	1	48	45	44	21	9,227	9,226	9,572	8,882		1,61	Adenosine kinase (EC 2.7.1.20)	Métabolisme
BX680855	RS49G03	8,94E-05	35	2,06E-01	3	1,54E-03	26	71	65	35	38	11,911	10,867	11,347	11,640	2,06		Alcohol dehydrogenase (EC 1.1.1.1)	Métabolisme
BX249652	Ge026F11	1,80E-01	5	2,16E-08	60	1,00E+00	3	81	68	60	24	9,946	9,624	9,259	10,347		2,13	ASR protein, partial (39%)	Réponse au Stress
BX251974	Ge058A01	9,35E-01	0	1,45E-04	41	1,00E+00	5	52	46	41	40	11,914	11,937	12,275	11,571		1,63	ATP synthase beta chain, mitochondrial precursor	Energie
CR392379	RN31A10	9,86E-06	55	5,53E-02	8	1,00E+00	1	66	63	55	27	10,009	9,277	9,847	9,586	1,66		Beta-glucosidase precursor (EC 3.2.1.21)	Métabolisme
BX681997	RS67H05	9,63E-06	58	1,23E-01	5	1,00E+00	0	65	64	58	23	9,547	8,650	9,324	9,053	1,86		Beta-glucosidase, chloroplast precursor (EC 3.2.1.21)	Métabolisme
BX681672	RS62G04	9,52E-01	0	1,53E-04	45	1,00E+00	0	61	45	45	20	8,226	8,212	8,032	8,388		1,28	Carboxypeptidase C (EC 3.4.16.5)	Devenir des protéines
BX678807	RS15D05	5,83E-01	3	9,96E-06	53	1,00E+00	1	57	56	53	26	9,404	9,242	9,666	9,013		1,57	Carboxypeptidase C precursor (EC 3.4.16.5)	Devenir des protéines
BX679422	RS25B09	5,41E-01	3	8,90E-07	58	1,00E+00	2	65	63	58	20	9,425	9,274	9,016	9,714		1,62	Cysteine synthase, chloroplast precursor (EC 2.5.1.47)	Métabolisme
BX254083	Ge093A12	8,69E-01	0	7,70E-11	73	7,58E-02	7	83	80	73	30	9,470	9,450	10,081	8,843		2,36	Geranyl diphosphate synthase, complete	Métabolisme
BX251816	Ge055H04	8,95E-01	0	1,06E-03	37	1,00E+00	0	39	38	37	20	8,852	8,902	8,562	9,181		1,54	Heat Shock 70 kd protein	Réponse au Stress
CR393209	RN53H11	2,78E-01	8	4,77E-04	35	1,00E+00	1	50	44	35	38	10,272	10,685	10,069	10,805		1,67	Heat shock protein STI1	Réponse au Stress
AL749654	AN05E01	4,50E-02	4	3,78E-14	82	1,00E+00	1	91	87	82	19	8,685	8,447	8,053	9,075		2,03	Oxygen-evolving enhancer protein 1, chloroplast precursor	Energie
BX682918	Ge123E04	8,86E-01	0	2,55E-04	42	1,00E+00	2	47	44	42	14	8,557	8,667	8,986	8,216		1,70	Phenylcoumaran benzylic ether reductase homolog TH1, comp	Transcription
BX682869	Ge122H09	1,54E-01	6	1,66E-08	61	3,46E-01	6	76	74	61	29	9,629	9,211	10,081	8,843		2,36	Pinene synthase (EC 4.2.3.1)	Métabolisme
BX253505	Ge084D05	8,37E-01	1	1,02E-03	34	1,00E+00	5	44	40	34	40	12,004	11,874	12,289	11,615		1,60	Pinus taeda water deficit inducible protein LP3	Réponse au Stress
BX682906	Ge123D02	7,82E-01	0	1,22E-08	67	6,75E-01	5	74	72	67	30	9,619	9,738	10,200	9,133		2,10	Probable ATP-dependent Clp protease proteolytic su	Non-classé
BX680864	RS49H06	7,45E-01	1	8,38E-06	38	1,00E+00	1	70	40	38	30	10,211	10,242	10,583	9,864		1,65	Pyruvate decarboxylase (EC 4.1.1.1)	Energie
CR394212	RN75C02	2,52E-03	27	3,00E-04	28	1,00E+00	5	63	60	28	36	10,810	10,318	10,857	10,369		1,40	Pyruvate decarboxylase isozyme 2 (EC 4.1.1.1) (PDC)	Energie
BE643924	NXCI_50_B7	2,69E-01	8	4,68E-04	35	1,00E+00	2	50	46	35	20	9,175	9,420	9,512	9,035		1,39	S-adenosyl-Met synthase	Métabolisme
CR394120	RN74B04	1,75E-01	10	1,03E-04	41	1,00E+00	4	59	54	41	20	9,012	8,618	8,489	9,187		1,62	Serine/threonine specific protein phosphatase 2C (EC 3.1.3.16	Métabolisme
AL749979	AS03D06	9,78E-01	0	1,97E-07	65	1,00E+00	0	67	66	65	38	10,031	10,056	9,654	10,427		1,71	Thioredoxin	Energie
BX252988	Ge076B04	1,00E-06	51	4,12E-04	18	1,00E+00	3	76	72	51	36	10,897	9,886	10,208	10,777	2,01	1,48	Thioredoxin H-type (TRX-H)	Energie
BX253001	Ge076C12	7,45E-01	0	3,77E-11	76	3,70E-01	4	83	81	76	34	11,007	11,123	12,026	10,082		3,85	Tubulin alpha chain	Organisation cellulaire
BX252968	Ge075G10	6,37E-01	1	1,61E-10	71	8,75E-02	7	82	79	71	40	12,190	12,327	12,798	11,691		2,15	Tubulin alpha chain	Organisation cellulaire
BX249241	Ge021G02	5,80E-01	1	3,19E-10	73	4,86E-01	4	80	78	73	32	10,656	10,880	11,561	9,930		3,10	Tubulin alpha chain	Organisation cellulaire
BX249627	Ge026D05	6,06E-01	1	2,61E-09	64	9,66E-02	8	81	73	64	27	9,490	9,620	9,946	9,177		1,70	Tubulin beta chain	Organisation cellulaire
BX252002	Ge058D06	6,88E-01	1	1,42E-06	60	1,00E+00	0	62	61	60	30	9,550	9,690	10,034	9,177		1,81	Tubulin beta chain	Organisation cellulaire
BX681614	RS62A02	9,85E-01	0	2,02E-06	56	1,00E+00	2	62	58	56	28	9,387	9,424	9,819	8,985		1,78	Tubulin beta chain	Organisation cellulaire
CR393797	RN65E06	4,46E-01	4	2,26E-04	31	4,99E-01	9	61	44	31	40	11,223	11,388	11,459	11,119		1,27	Tubulin beta chain	Organisation cellulaire
BX000608	Tubuline	7,26E-01	1	3,05E-08	65	6,63E-01	5	72	71	65	31	10,481	10,639	11,217	9,871		2,54	Tubuline	Organisation cellulaire

Figure D-49 : Métabolisme de la Méthionine et méthylation de l'ADN.

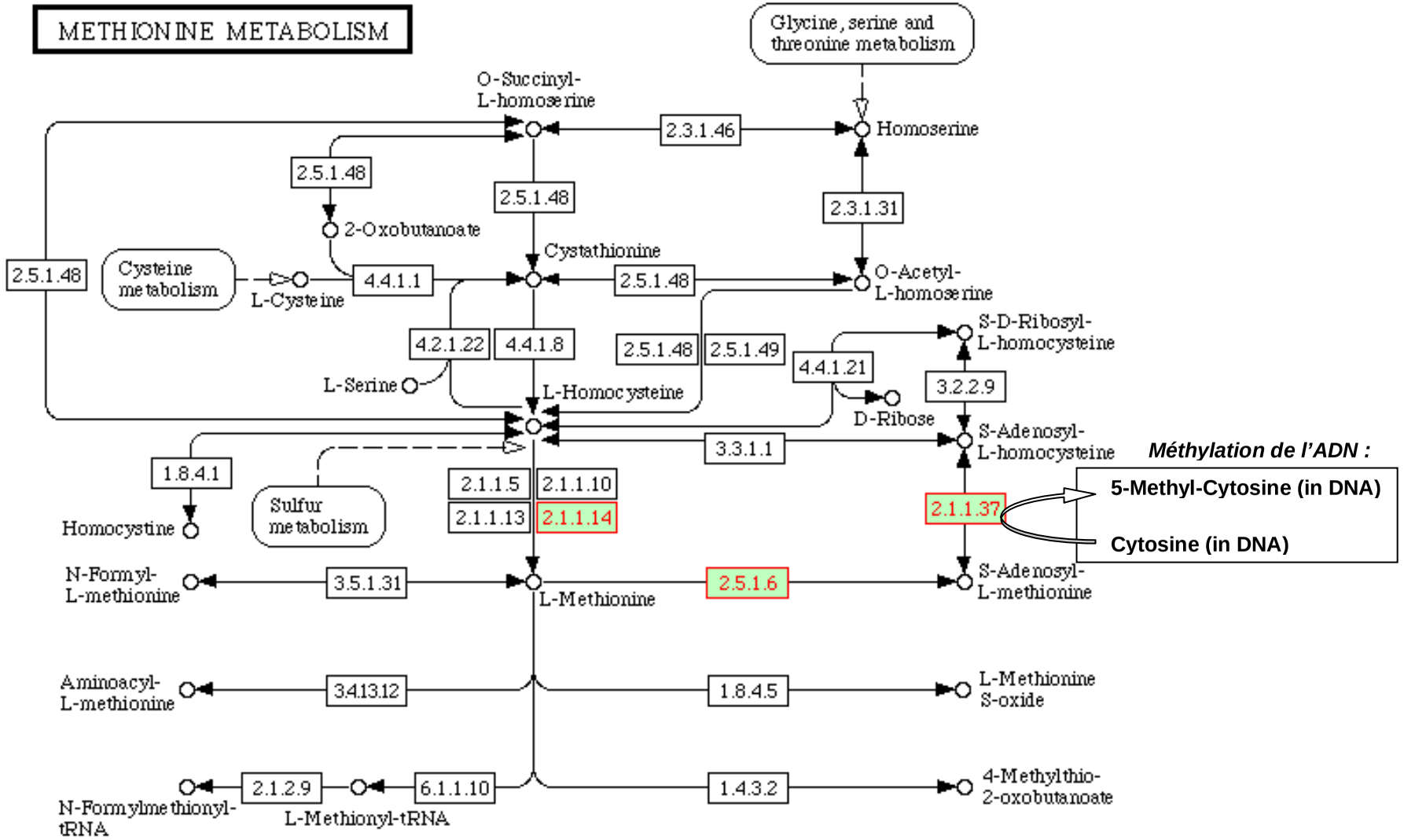
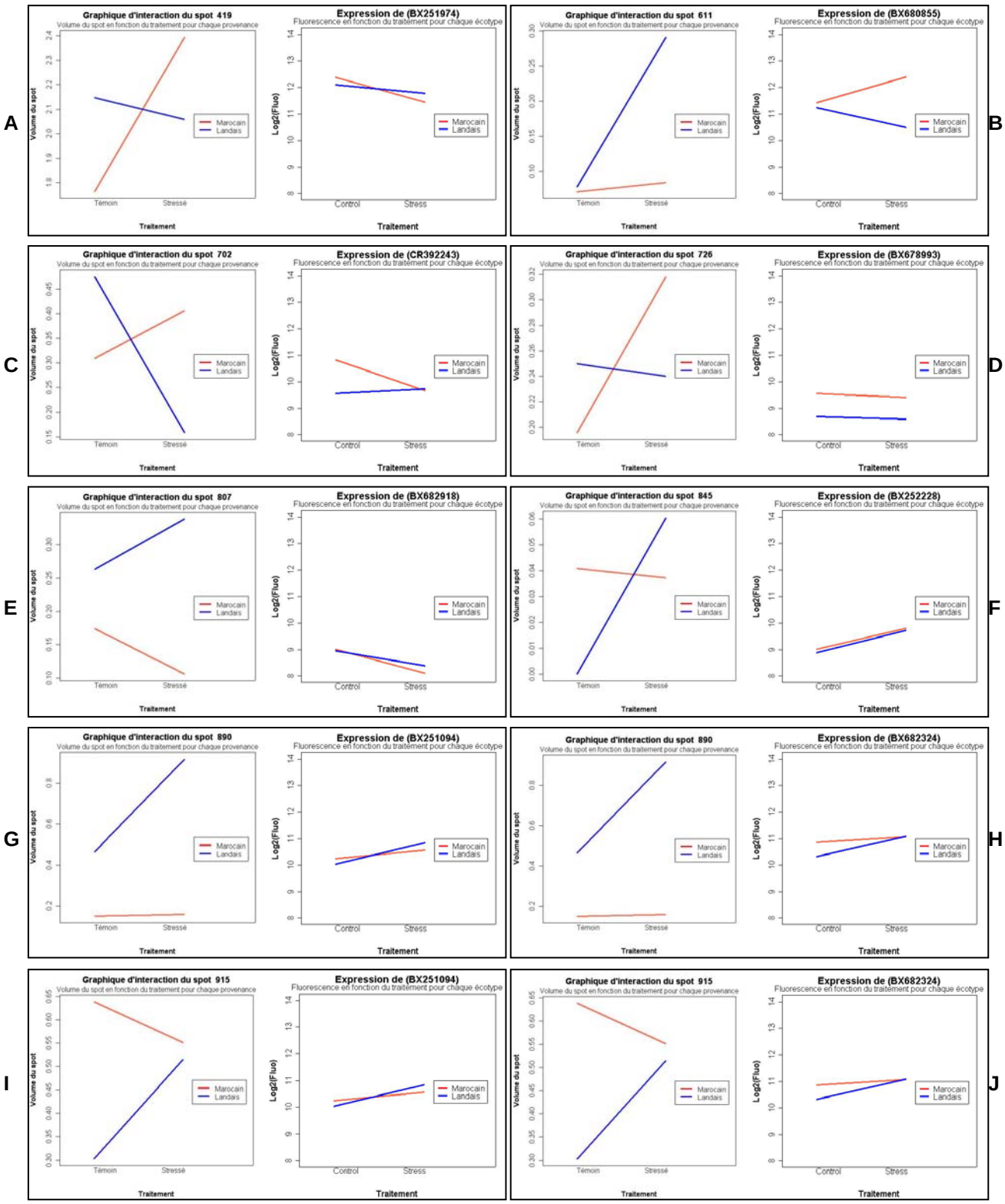
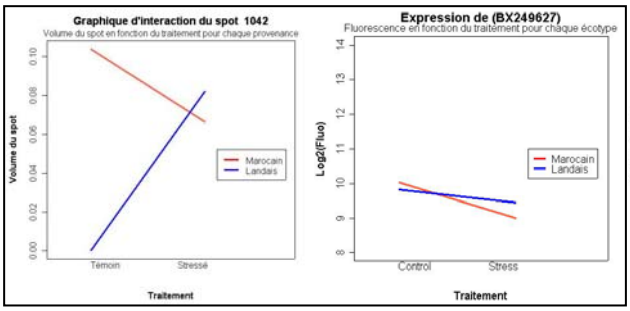


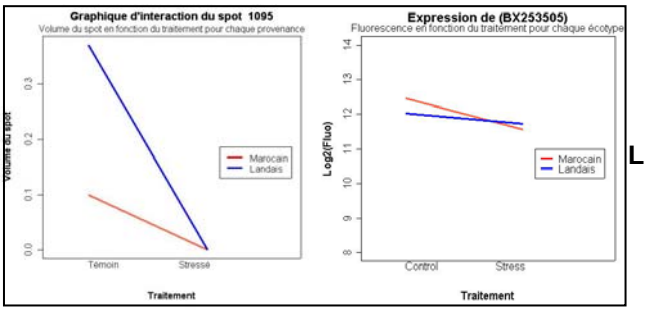
Figure D-50 : Comparatif des profils d'expression des protéines (à gauche) et des transcrits du TC homologue (à droite).



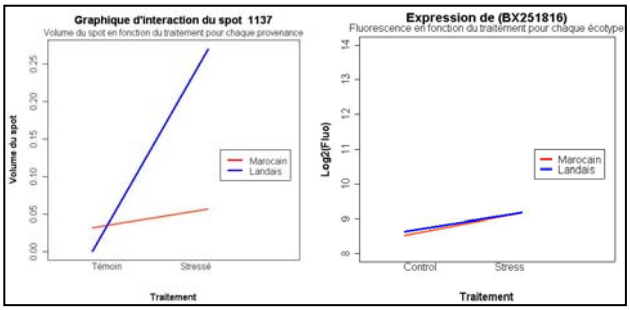
K


Graphique d'interaction du spot 1042
Volume du spot en fonction du traitement pour chaque provenance
Volume du spot
Témoin
Stressé
Traitement
Marocain
Landais
Expression de (BX249627)
Fluorescence en fonction du traitement pour chaque écotype
Log2(Fluo)
Control
Stress
Traitement
Marocain
Landais

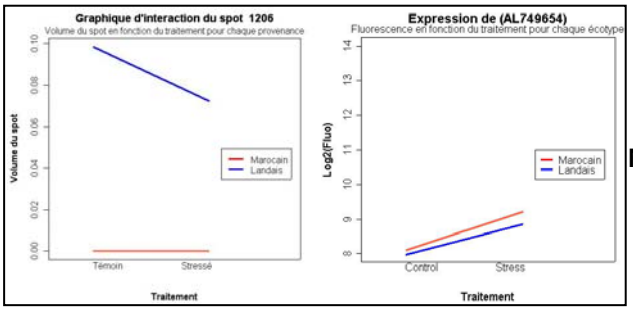
L


Graphique d'interaction du spot 1095
Volume du spot en fonction du traitement pour chaque provenance
Volume du spot
Témoin
Stressé
Traitement
Marocain
Landais
Expression de (BX253505)
Fluorescence en fonction du traitement pour chaque écotype
Log2(Fluo)
Control
Stress
Traitement
Marocain
Landais

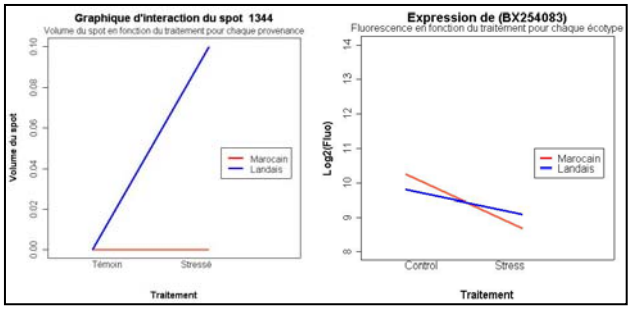
M


Graphique d'interaction du spot 1137
Volume du spot en fonction du traitement pour chaque provenance
Volume du spot
Témoin
Stressé
Traitement
Marocain
Landais
Expression de (BX251816)
Fluorescence en fonction du traitement pour chaque écotype
Log2(Fluo)
Control
Stress
Traitement
Marocain
Landais

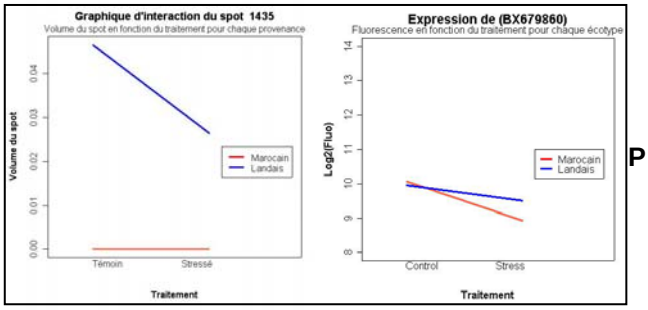
N


Graphique d'interaction du spot 1206
Volume du spot en fonction du traitement pour chaque provenance
Volume du spot
Témoin
Stressé
Traitement
Marocain
Landais
Expression de (AL749654)
Fluorescence en fonction du traitement pour chaque écotype
Log2(Fluo)
Control
Stress
Traitement
Marocain
Landais

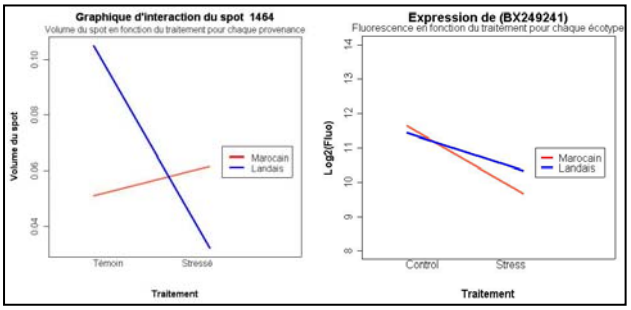
O


Graphique d'interaction du spot 1344
Volume du spot en fonction du traitement pour chaque provenance
Volume du spot
Témoin
Stressé
Traitement
Marocain
Landais
Expression de (BX254083)
Fluorescence en fonction du traitement pour chaque écotype
Log2(Fluo)
Control
Stress
Traitement
Marocain
Landais

P


Graphique d'interaction du spot 1435
Volume du spot en fonction du traitement pour chaque provenance
Volume du spot
Témoin
Stressé
Traitement
Marocain
Landais
Expression de (BX679860)
Fluorescence en fonction du traitement pour chaque écotype
Log2(Fluo)
Control
Stress
Traitement
Marocain
Landais

Q


Graphique d'interaction du spot 1464
Volume du spot en fonction du traitement pour chaque provenance
Volume du spot
Témoin
Stressé
Traitement
Marocain
Landais
Expression de (BX249241)
Fluorescence en fonction du traitement pour chaque écotype
Log2(Fluo)
Control
Stress
Traitement
Marocain
Landais

pour la synthèse des composés terpéniques au cours du stress. Nous rappellerons que les hormones comme l'ABA et les GA partagent des précurseurs communs avec les terpènes.

D.4.4.3 « Métabolisme »

Dans cette catégorie, nous retrouvons des fonctions que nous avons déjà rencontrées lors de l'analyse des transcrits différentiellement exprimés. Une Pyruvate décarboxylase isoenzyme 1 (EC 4.1.1.1) (spot288) augmente lors du stress et très fortement chez l'EF. Des transcrits de mêmes fonctions ont été détectés mais aucun de l'isoforme 1. Une chaine beta de l'ATP synthase (EC 3.6.3.14) (spot419) est accumulée lors du stress chez l'EM et diminue chez l'EF. Sur la figure D-50A, les transcrits ne concordent pas, nous avions une légère interaction non significative et de sens opposée à celle observée pour le spot 419. La Sérine carboxypeptidase (EC 3.4.16.5) (spot456) n'est pas régulée chez l'EM et diminue chez l'EF. Une Thioredoxine réductase (EC 1.8.1.9) (spot793) n'est pas régulée chez l'EM et diminue chez l'EF. Une Cyanogenic beta-glucosidase (EC 3.2.1.21) (spot1230) qui diminue chez l'EF et s'accumule en stress chez l'EM.

La régulation de la beta-glucosidase peut indiquer une régulation de la conversion de la cellulose en cellobiose sur lequel agit l'enzyme pour donner le D-glucose. Cela peut indiquer une régulation du métabolisme ou une modification des parois par régulation de la dégradation des composés cellulosiques.

Nous remarquerons également les spots 807 et 808 dont la fonction est une Phenylcoumaran benzylic ether réductase (PCBERs) qui a été décrite par Gang *et al.* (1999) sur *Pinus taeda*. La séquence de cet enzyme est très proche de celle de la pinoresinol-lariciresinol reductase (PLR) et de celle des isoflavones réductases (IFR). Il est impliqué dans la synthèse des composés de défense de la plante, dérivés des phenylpropanoïdes. L'expression du transcrit BX682918 du TC homologue (Fig. D-50E) indique une diminution d'expression lors du stress mais moins prononcée chez l'EF ce qui pourrait éventuellement être compatible avec la variation d'accumulation du spot 807.

Enfin le spot 611, homologue à l'Alcool déshydrogénase (EC 1.1.1.1) qui est un enzyme intervenant dans de nombreuses voies métaboliques, est également détecté. De nouveau, le profil du transcrit BX680855 (Fig. D-50B) n'explique absolument pas l'accumulation du spot 611 et présente même un profil d'expression totalement opposé.

D.4.4.4 « Cytosquelette »

L'actine et la tubuline sont de nouveau mises en évidence. Leur rôle dans le stress peut être structural dans le maintien de la conformation cellulaire et/ou intervenir dans le trafic intracellulaire de composés. Les profils d'expression de l'actine ne correspondent cependant pas à ce qui avait pu être détecté en étudiant les transcrits. Nous avions observé une diminution de l'expression lors du stress pour les 2 écotypes (BX255342). Le spot 600 montre un comportement opposé. Concernant la tubuline, les 2 spots détectés, respectivement homologues aux chaînes alpha (spot1464) et beta (1042), ont des sens de variation opposés (Fig. D-50Q et K). Les transcrits homologues aux tubulines ont des profils d'expression assez homogènes présentant une diminution de l'expression chez les 2 écotypes mais moins forte pour l'EM. Les profils des accessions BX249627 et BX249241 appartenant respectivement aux TC homologues aux spots 1042 et 1464 n'expliquent pas les accumulations observées.

D.4.4.5 « Réponse au stress »

Parmi les HSP détectées, l'homologie avec la protéine de choc thermique 70 kDa pour plusieurs spots (spot631/243/637/1137(Fig. D-50M)/1210/1211) confirme son implication dans la réponse au stress hydrique. Les données de transcrits indiquaient une surexpression lors du stress dans les 2 écotypes. Au niveau protéique, elle est représentée par plusieurs spots dont certains (spot243/1210/1211) ont une quantité diminuée chez l'EF voire même ne sont détectés que pour les échantillons contrôles de l'EF (spot1210/1211) ou de l'EM (spot637). Ces résultats confirment donc l'implication de cette HSP 70kDa dans la réponse au stress. La détection de spots distincts présents seulement dans l'échantillon témoin pour chaque écotype pourrait être expliquée par l'existence de deux formes allèliques différentes de la protéine et spécifiques de chaque écotype, ainsi que par la présence possible de modifications post-traductionnelles intervenant lors du stress qui se traduiraient par une migration du spot à un autre emplacement. Une autre HSP notée « STI » (spot231), n'est détectée sur les gels 2D que dans les échantillons témoin de l'EM. Au niveau des transcrits, l'expression était détectée dans les deux écotypes avec une augmentation lors du stress. Il est possible que nous soyons là encore, en présence de formes allèliques différentes ne permettant la détection du spot que dans l'EM. Les épissages alternatifs et les modifications post-traductionnelles sont également possibles, pouvant amener à des protéines migrant à des endroits très différents sur le gel.

Les autres fonctions détectées dans les deux approches sont des ASR (Abscisic stress ripening) (spot890/915) et des LP3 (spot897/1095). Pour la « water deficit inducible protein LP3 », les résultats des protéines comme des transcrits indiquent une diminution au cours du

stress (Fig. D-50L) et non une induction comme on aurait pu s'y attendre. En effet, Chang *et al.* (1996) avaient montré lors d'une expérience de stress hydrique sur des plants en pots de *Pinus taeda* que les transcrits homologues à cette protéines étaient induits dans les racines lors du stress. Dans le cas de l'ASR (Fig. D-50G/H/I/J), les résultats des transcrits et des protéines pourraient globalement expliquer le sens de régulation, avec une induction forte lors du stress chez l'EF tandis que l'EM n'induit pratiquement pas ce gène. Par contre, les transcrits ne permettent pas d'expliquer l'accumulation du spot 890 chez l'EF et du spot 915 chez l'EM.

D.4.4.6 Autres catégories

Les autres fonctions que l'on a pu détecter et que nous ne ferons que citer, correspondent à une Phosphosulfolactate synthase (EC 4.4.1.19) (spot845) (Fig. D-50F) qui n'est pas régulée par le stress chez l'EF et exprimée fortement uniquement en stress chez l'EM, une protéine Sérine/thréonine phosphatase isoenzyme 2 (spot1341) détectée uniquement en stress dans l'EF et une (Probable) ATP-dependent Clp protéase (spot116) plus exprimée et non régulée dans l'EF et exprimée uniquement en stress chez l'EM.

D.4.5 De la protéine au gène : cas de l'ACC désaminase

Les résultats de l'analyse du protéome mettent en évidence le spot 1385 qui n'est détecté que dans l'échantillon stressé de l'EF. La fonction attribuée à ce spot est une putative 1-aminocyclopropane-1-carboxylate désaminase (EC 3.5.99.7). Le rôle de cet enzyme dans la régulation de l'éthylène est bien connu et documenté dans la littérature. Il est exprimé par les bactéries du sol et les microorganismes symbiotiques. Hontzeas *et al.* (2004) ont ainsi isolé, exprimé et caractérisé la 1-aminocyclopropane-1-carboxylate désaminase de la rhizobactérie (*Pseudomonas putida* UW4). Les études de Ma *et al.* (2003; 2004) semblent montrer que la bactérie symbiotique (*Rhizobium leguminosarum* Biovar viciae, *Sinorhizobium meliloti*) utiliserait cet enzyme pour moduler la teneur en éthylène de la plante hôte (Pois, Trèfle) et favoriser la nodulation. De nombreuses études rapportent l'utilisation de promoteurs de croissance à activité ACC désaminase pour la construction de plantes transgéniques (Grichko and Glick 2001b; Penrose and Glick 2003). Ce type de construction confère une résistance à différents stress chez les plants transformés (Grichko and Glick 2001a; Robison *et al.* 2001; Mayak *et al.* 2004; Stearns *et al.* 2005). Dans le cas de la tomate, Mayak *et al.* (2004) ont observé une résistance au stress salin des plants transformés. Cependant, à notre connaissance, aucune étude n'a encore décrit une régulation de la teneur en éthylène par une ACC désaminase endogène de la plante. Nous avons donc recherché également l'existence de

transcrits homologues dans les bases d'EST. Chez les conifères, nous disposions pour l'ACC désaminase du TC44655 composé de l'EST de pin maritime Ge071A04 (BX252805) et de l'EST de pin *taeda* (BQ107915). Nous avions également l'accession BE662323. Ces EST ne couvraient pas l'intégralité de la partie codante. Par ailleurs, une recherche sur les bases de données a permis d'identifier le gène OJ1611_C08.26 homologue à l'ACC désaminase chez le riz dont l'ARNm complet correspondant est le XM_468034. Une recherche d'homologie de cette séquence a permis d'identifier : les accessions gi|20259954 et gi|30694324 d'*Arabidopsis* ; le singleton CO484002 (homologues en 5') et le TC8605 (homologue en 3') (séparés d'une 50^{aine} de bases manquantes) du projet épicéa du TIGR ; une séquence de bouleau (AY154652). La séquence codante complète du gène n'étant pas déterminée chez les conifères, nous avons décidé de la séquencer et de déterminer s'il pouvait s'agir d'une séquence contaminante d'origine bactérienne. L'alignement des séquences du riz, de l'épicéa et de pins a permis d'identifier des régions conservées en 5' au niveau nucléotidique sur lesquelles nous avons déterminé différentes amorces. Le dessin des amorces de la région 3' s'est fait sur l'EST disponible. Ces amorces nous ont permis d'amplifier sur ADNc l'intégralité de la partie codante et de cloner ce gène. Depuis l'obtention de cette séquence de pin maritime, un nouvel assemblage d'EST du TIGR (version 6.0) a été généré en incluant les nouveaux EST produits sur *Pinus taeda*. Le TC44655 sur lequel nous nous étions basé, a été éclaté en deux nouveaux consensus : le TC60605 qui est constitué de 6 séquences dont celle de pin maritime (BX252805) et l'accession BE662323 et le TC60606 regroupant deux EST. Le TC60605 couvre maintenant l'intégralité de la partie codante et l'alignement confirme la séquence que nous avions obtenue (Fig. Ax-6). Après séquençage, une analyse phylogénétique des séquences traduites de l'ACC désaminase de plantes supérieures et de séquences de bactéries a été réalisée en utilisant le logiciel MEGA version 3.0 (Kumar *et al.* 2004) avec la méthode UPGMA. La figure D-51 montre que la phylogénie des séquences permet de bien séparer les différents groupes taxonomiques. Les séquences des plantes supérieures sont bien rassemblées et se distinguent des séquences des bactéries. Les séquences de *Pyrococcus* et de *T. maritima* sont phylogénétiquement entre les autres bactéries et les plantes supérieures. Deux séquences provenant d'*A. thaliana* se détachent et ne font vraisemblablement pas partie de la même famille. Ces résultats semblent soutenir l'hypothèse d'une ACC désaminase endogène chez les plantes supérieures. L'amplification sur ADN génomique a été testée et permet l'amplification de fragments correspondant à la taille attendue. Pour vérifier que la séquence identifiée code effectivement pour une protéine à activité ACC désaminase, nous avons réalisé une construction plasmidique clonée dans un

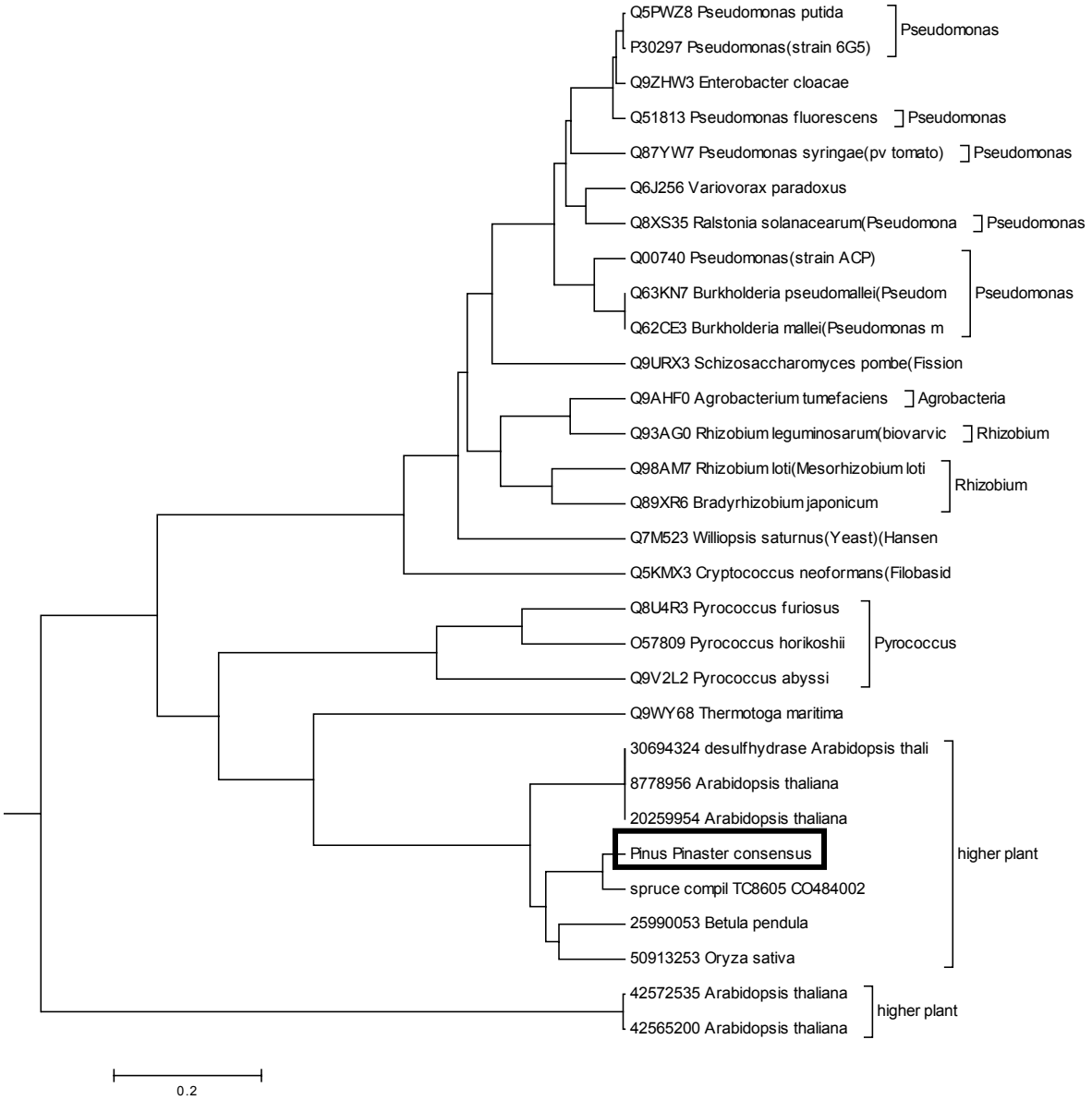


Figure D-51 : Arbre de similarité des séquences homologues à l'ACC désaminase chez les plantes supérieures et les bactéries. Les numéros d'accessions indiqués correspondent au « gi » ou au « gb »

vecteur d'expression. Il restera à l'issue de cette thèse à continuer les étapes de production, de purification et le test d'activité de la protéine de fusion pour confirmer la présence d'une ACC désaminase endogène chez les plantes supérieures. Si l'activité de désamination de l'ACC est confirmée, il sera intéressant de poursuivre plus avant cette étude même si les résultats d'expression de transcrits et de protéines sont opposés. En effet, les résultats de PCR quantitative sur ce transcrit ont montré une différence dans l'expression et la régulation entre les deux écotypes avec une accumulation plus élevée chez l'EM et une diminution lors du stress. Au niveau protéique, le spot n'était détecté que dans l'échantillon landais stressé. Cette déconnection entre régulation transcriptionnelle et traductionnelle montre combien il faut rester prudent quant à l'interprétation des données du transcriptome.

D.4.6 Discussion

L'étude des transcrits nous a permis d'analyser les régulations d'expression des gènes individuellement. Elle ne nous permet cependant pas de prédire la quantité de protéines produite ou de prendre en compte l'ensemble des modifications depuis la traduction de l'ARNm, en passant par les modifications post-traductionnelles, jusqu'à la forme finale de la protéine. L'électrophorèse bidimensionnelle nous donne accès à une fraction du protéome au moment du prélèvement. Les variations de quantité de protéines au niveau d'un spot sont le résultat de l'ensemble des réactions qui se sont produites depuis la traduction jusqu'à la protéine détectée. En se plaçant au niveau protéique, nous diminuons le nombre d'étapes séparant les mesures des répercussions physiologiques observées. Nous avons vu que cette approche ne permet cependant pas de s'affranchir de la présence de protéines en mélange, de formes allèliques et de familles multigéniques modifiant la migration des protéines et l'attribution de la fonction des protéines détectées reste délicate. Nous avons aussi constaté qu'un certain nombre de spots correspondaient à des protéines de fonction également détectée dans les transcrits différentiellement exprimés. Cependant, les profils d'expression des spots restent rarement confirmés par les transcrits (Fig. D-50). Il apparaît ainsi clairement qu'il n'est pas possible de relier directement l'expression des transcrits avec les variations de quantité d'un spot. Outre le fait qu'il y ait une déconnection entre le transcrit et le spot, il n'est pas certain que le TC homologue détecté soit effectivement le gène concerné, l'annotation des spots se situant plus au niveau de la fonction que du gène.

Pour conclure, ces deux approches sont donc complémentaires et permettent d'identifier les fonctions, éventuellement les voies métaboliques, régulées lors d'une expérience. En ce sens, l'approche protéomique peut orienter l'étude vers des gènes candidats

dont l'expression est régulée et pour lesquels des spots de fonctions similaires sont différentiellement exprimés. L'étude des transcrits permet toutefois d'accéder à des gènes tels que les facteurs de transcriptions agissant en amont lors du stress.

Conclusions et Perspectives

E CONCLUSIONS ET PERSPECTIVES

E.1 Développement et utilisation d'une approche à haut débit pour l'étude de la réponse au stress hydrique.

L'étude de la régulation du transcriptome par une approche de type puce à ADNc nécessite de disposer d'une quantité de séquences suffisante. Sur une espèce où cette source de données n'est pas disponible, cela demande un travail de préparation considérable. Pour le pin maritime, la coopération des biologistes et des bio-informaticiens a permis de mener à bien un programme de séquençage de près de 10000 EST dans le cadre de cette thèse. Nous disposons maintenant d'une chaîne d'assemblage et d'annotation d'EST qui permet de traiter les 26000 EST actuellement disponibles. L'assemblage des séquences et la génération de séquences consensus se sont révélés un moyen efficace pour reconstituer la séquence des gènes. Nous avons pu déterminer qu'avec le choix d'un séquençage en 5', la couverture de l'extrémité 3' des gènes était faible. Un séquençage en 3' d'une liste de clones choisis augmenterait considérablement la quantité d'information disponible. Dans un programme de séquençage il est donc intéressant de ne pas tenir compte uniquement de la redondance mais également de la couverture des gènes par les EST.

L'étape d'annotation permet d'assigner une fonction aux EST. Une annotation fine comme celle réalisée pour les séquences de racines a demandé beaucoup de temps (estimé à 5 mois). Si une assignation précise de la fonction est intéressante, dans le cas de la détection de gènes différentiellement exprimés, il aurait été possible de n'annoter que les gènes statistiquement significatifs. Cependant, ce travail va pouvoir maintenant servir de base pour faciliter l'annotation des nouveaux projets d'assemblage d'EST qui vont ainsi pouvoir hériter de ces annotations.

Les outils bioinformatiques élargissent le champ d'analyse des données. La détection de gènes différentiellement exprimés par Northern électronique a permis d'identifier des transcrits dont certains présentent des fonctions décrites dans la littérature comme étant régulées au cours du stress hydrique. Cependant, nous avons vu que ce type d'approche est sensible à la qualité des données analysées. Les incertitudes sur l'identification des transcrits issus d'un gène et leur représentativité dans les banques limitent l'interprétation possible des résultats. La recherche de différences d'expression dans les bases de données publiques selon cette méthode ne rend compte que de la représentativité des EST générés à partir des banques construites qui ne reflètent pas forcément la réalité biologique.

Nous avons choisi de générer un Unigène de près de 7500 accessions en sélectionnant un représentant par contig et en limitant au maximum la redondance dans les singletons. Dans l'hypothèse ou la taille de l'Unigène ne serait pas une limite, la sélection d'une séquence en 3' et d'une séquence en 5' permettrait de comparer l'hybridation des deux dépôts et éventuellement de différentier l'hybridation spécifique des transcrits d'un membre, de celui de la famille du gène.

Le choix d'un plan d'expérience de type « loop design » s'est révélé une bonne option en offrant la possibilité de comparer l'expression de deux écotypes et 5 traitements avec un nombre minimum de lame. L'inconvénient est cependant une faible robustesse du système en cas de données manquantes. Il convient donc de combiner de multiples boucles plutôt que de réaliser une boucle unique. La normalisation et le traitement statistique des données de fluorescence sont une étape cruciale pour éliminer les biais provenant de la technique et de l'expérimentation utilisée.

L'analyse des données d'expression des puces nous a permis d'identifier 458 « gènes » régulés lors de la réponse à court terme, 612 « gènes » régulés dans la réponse à long terme et 168 accessions sont communes aux deux expériences. Les gènes détectés par le dispositif expérimental mis en place, diffèrent essentiellement pour le facteur « provenance » prépondérant dans la réponse à court terme et pour le facteur « traitement » prépondérant dans la réponse à long terme. Les fonctions des gènes régulés couvrent l'ensemble des catégories fonctionnelles. Nous détectons des gènes impliqués dans la perception, la transduction du signal et la régulation de l'expression avec des facteurs de transcription, des histones et des gènes pouvant intervenir dans l'accessibilité de l'ADN. Nous trouvons également des gènes impliqués dans la traduction, la modification et la dégradation des protéines. Nous avons également mis en évidence des gènes connus pour leur implication dans la réponse au stress avec des LEA, Déhydrines, HSP, AGP et des gènes de détoxification. Des gènes pouvant intervenir dans la régulation de l'osmolarité ont également été mis en évidence. De plus, près de 30% des gènes détectés sont de fonction encore inconnue.

Le choix d'effectuer un clustering uniquement sur les gènes différentiellement exprimés s'est révélé une méthode efficace pour se focaliser sur l'étude des 215 gènes dont le profil d'expression indiquait une régulation contrastée entre les deux écotypes soumis à la contrainte osmotique.

Dans ces gènes d'intérêt, nous avons identifié des transcrits surexprimés, correspondant à des gènes de réponse au stress (HSP, Métallothionéines, Chitinase, synthèse des flavonoïdes, SOD, protéines chaperonnes dnaJ, « Aluminium-induced », Peroxydases, Cytochrome P450).

D'autres transcrits sont impliqués dans la régulation de l'expression (Protéine membranaire induite par l'ABA, « RING finger », « Zinc finger », « G-box binding factor », «F-box », « Homéobox-leucine zipper », Histone H4, « Heterochromatine-like », Phospholipase D, Elongation factor 1-gamma …). Mais nous avons aussi trouvé des transcrits agissant sur la composition des parois (Cellulose synthase, Elastine) et des gènes du métabolisme et de l'énergie.

Nous avons également détecté des gènes sous-exprimés différemment entre écotypes lors du stress. Les fonctions associées touchent la régulation de la transcription (FT Rof1, EREBP, Histones H2A, H4), Péroxydases, LEA, protéines « Abscissic stress ripening», « induites par la déshydratation », Métallothionéines, « Disease resistance », Défensine. Des transcrits sont impliqués dans la biosynthèse des composés terpéniques (IPP isomérase, GPP synthase, Pinène synthase, MECDP-synthase), d'autres peuvent être impliqués dans la synthèse de la lignine (C4H, Caffeate O-methyltransferase) et nous avons détecté des enzymes pouvant intervenir dans la production de Bétaïne (EC 2.7.7.15, EC 1.2.1.8).

La régulation de certains de ces gènes semble indiquer que les écotypes adoptent des stratégies différentes pour répondre au stress osmotique. L'écotype français (EF) adopterait une expression constitutive plus élevée de gènes non régulés lors du stress sur des mécanismes touchant la composition des parois –en particulier la synthèse de la lignine- et la production de bétaïne, tandis que l'écotype marocain (EM) diffèrerait de l'EF en surexprimant des gènes régulés lors du stress. Les mécanismes concernés touchent la détoxification, la stabilisation et modification des protéines et probablement le maintien de gradients ioniques plus actifs. De plus, l'EM semble se distinguer de l'EF dans la réponse à court terme par 1) une plus forte activation de gènes intervenant dans la régulation de l'expression et la synthèse des protéines 2) la surexpression de protéines de réponse au stress et 3) une synthèse de cellulose et d'élastine.

L'analyse de l'expression de 35 gènes par PCR quantitative a souligné la difficulté de confirmer les données des microarrays. Ce constat peut trouver son explication dans la spécificité d'analyse de la PCR et l'étude de l'expression de membres spécifiques de familles multigéniques. Nous avons pu montrer entre autres la régulation des gènes impliqués dans la voie de biosynthèse de l'éthylène. Plus particulièrement, nous avons mis en évidence l'existence d'une ACC désaminase endogène chez les plantes supérieures qui pourrait jouer un rôle dans la régulation de la production de cette hormone.

La comparaison des données d'expression avec les données de protéomique a souligné la complémentarité de ces deux approches pour la prise en compte des modifications post-

traductionnelles et du découplage existant entre l'expression d'un transcrit et l'accumulation de la protéine.

E.2 Perspectives de cette étude

L'utilisation des microarrays n'est pas une finalité en soit. Elle constitue le point de départ d'études plus ciblées. Ce type d'approche permet d'avoir rapidement une vue globale des gènes dont l'expression est potentiellement intéressante pour des études plus précises d'expression, de transformation génétique, de cartographie génétique ou d'étude de diversité.

E.2.1 Exploitation des ressources mises en place

Un des objectifs initiaux de cette thèse était de créer une ressource de séquences de pin maritime par la génération d'EST. La mise à disposition de cette base de connaissance et des outils bioinformatiques associés ouvre maintenant de nombreuses perspectives. Nous avons vu que de telles ressources améliorent considérablement l'identification des gènes, des transcrits et des protéines étudiés. L'accès à la séquence partielle ou totale de plusieurs milliers de gène permet d'envisager leur étude tant au niveau de leur expression que de leur diversité nucléotidique (Le Dantec *et al.* 2004; Pot *et al.* 2005). Lorsqu'elles sont disponibles, il est également possible d'identifier et d'étudier les différents membres de familles multigéniques.

Nous avons développé l'utilisation de la technologie des puces à ADNc au laboratoire. Avec le séquençage supplémentaire de 10000 EST issus d'une banque « bourgeons » nous venons de créer un nouvel Unigène plus complet (12K éléments) pour des études d'expression sur le stress hydrique ainsi que pour la formation du bois. Le nouveau réseau d'excellence Evoltree qui porte sur la biodiversité des écosystèmes terrestres nous permettra également d'enrichir cette collection.

Comme nous l'avons réalisé pour la voie de synthèse de l'éthylène, il serait maintenant intéressant d'étudier par PCR quantitative des voies biochimiques supposées impliquées dans la réponse au stress hydrique ou pour lesquelles des gènes différentiellement exprimés entre écotypes ont été identifiés. Cela permettrait de mieux cerner le rôle des enzymes et leur implication dans des processus biochimiques liés à l'adaptation.

Nous avons émis plusieurs hypothèses sur des stratégies de régulation différentes de certaines voies biochimiques des deux provenances. Il serait maintenant intéressant d'effectuer des mesures complémentaires sur les métabolites. Le couplage des données du métabolome avec

les données de transcriptome s'est révélé une aide précieuse pour l'interprétation des gènes impliqués dans la formation du bois dans le cadre de la thèse de Jorge Paiva (2006).

E.2.2 Potentiel adaptatif des gènes candidats

La capacité à répondre rapidement à un changement de contrainte environnementale va dépendre du niveau de diversité génétique non neutre et de la plasticité des arbres dans les populations naturelles et d'amélioration. La plasticité phénotypique va permettre à un génotype de produire différents phénotypes en réponse à des environnements distincts. Une part de cette plasticité est contrôlée génétiquement et reflètera la valeur adaptative du génotype. Dans cette thèse, nous avons exploré la plasticité moléculaire (propriété d'un génotype à modifier l'expression de son génome en réponse à son environnement) de deux écotypes en condition de stress osmotique. Nous avons identifié des gènes dont la norme de réaction en faisait des gènes candidats expressionnels potentiels. Les variations d'expression ne sont cependant pas forcément liées à un caractère adaptatif mais peuvent découler d'une plasticité « passive ». Il est donc indispensable de vérifier si la régulation d'un gène permet la sélection de génotypes qui auront une résistance accrue à la sécheresse. Pour constituer un bon gène candidat, un gène doit affecter l'expression de caractères adaptatifs et apporter un avantage sélectif au génotype sélectionné. En populations naturelles, nous devrions donc pouvoir observer des traces (signatures) de l'effet de la sélection naturelle sur les bons gènes candidats. La sélection naturelle, en favorisant les phénotypes les plus adaptés et en éliminant les moins adaptés, va modifier les distributions allèliques. Des tests de neutralité permettront alors de déterminer si les profils de diversité nucléotidique sont compatibles avec l'hypothèse de neutralité (Nielsen 2005). Si l'idéal est d'étudier directement l'effet de la sélection comme c'est le cas dans les populations naturelles, les temps de générations (plusieurs dizaines d'années pour les arbres) rendent impossible une telle approche sur des dispositifs expérimentaux. Par contre, les études d'association (Cardon and Bell 2001) constituent des méthodes indirectes, compatibles avec l'expérimentation. Dans cette optique des études d'association portant sur des populations naturelles contrastées sont en développement chez le pin maritime (thèse d'Emmanuelle Eveno en cours). Il s'agit de détecter et de valider des polymorphismes nucléotidiques en les associant à la variation phénotypique de caractères adaptatifs tels que la croissance, la biomasse ou l'efficience d'utilisation de l'eau (delta ^{13}C). Ces approches devraient permettre de valider les gènes candidats voire les polymorphismes d'intérêt. Cet outil moléculaire pourrait alors servir de critère diagnostique pour contrôler le niveau de diversité adaptative dans les futures variétés améliorées.

E.2.3 De l'organe à la cellule

Les résultats de cette étude renvoient une image de l'expression globale des transcrits des parties racinaires. Toutefois, les racines sont des organes complexes. Les transcrits exprimés dans les pointes racinaires, les tissus conducteurs lignifiés ou les zones d'absorption des éléments nutritifs et de l'eau, ne sont pas les mêmes. Une étude à l'échelle d'un tissu voire d'un type cellulaire fournirait certainement des informations plus précises sur les stratégies de régulation des gènes concernés. Depuis quelques années, il existe une technologie nommée « Laser-capture Microdissection » permettant de prélever des échantillons à l'échelle de la cellule (Day *et al.* 2005). Cette méthode associée à l'amplification linéaire de transcrits permet de réaliser des études d'expression. Ces outils ont déjà été utilisés chez *Arabidopsis* pour étudier le développement de l'embryon (Casson *et al.* 2005), ainsi que chez le maïs pour identifier les gènes différentiellement exprimés dans les cellules de l'épiderme ou des tissus vasculaires (Nakazono *et al.* 2003). Il est également possible de construire des banques d'ADNc à partir des cellules prélevées comme décrit par Asano et al (2002) sur le phloème du riz (Asano *et al.* 2002). Cette technique est actuellement en cours de développement au laboratoire et pourrait nous aider à mieux comprendre les régulations des transcrits intervenant lors de l'application d'un stress osmotique en distinguant les régulations des zones de croissance, d'absorption et les tissus vasculaires.

E.2.4 Une autre voie de régulation : les micros ARN

Les micros ARNs (miRNAs), sont des molécules d'ARN simple brins d'environ 22 nucléotides qui sont associées au complexe RISC (RNA-induced silencing complex). Cette catégorie d'ARN semble jouer un rôle important dans la régulation de l'expression des gènes, particulièrement au cours du développement (Kidner and Martienssen 2005). En s'associant à l'ARNm cible le complexe inhibe la traduction et provoque le clivage de l'ARNm. Ainsi, il apparaît clairement que la régulation de l'activité des gènes n'est pas limitée à une voie métabolique. Il semble donc indispensable de mener des études sur ce mécanisme qui s'annonce aussi important que la régulation de l'expression des gènes par des facteurs de transcription.

E.2.5 Contrôle épigénétique

Le dernier point que nous évoquerons concerne l'ensemble des modifications de l'expression des gènes héritables au cours des divisions cellulaires, réversibles (selon le type cellulaire) et qui ne s'accompagnent pas de changements de séquences nucléotidiques. Les

deux catégories principales de modifications épigénétiques connues, concernent les mécanismes de méthylation de l'ADN par différentes familles d'ADN Méthyltransférases (DNMT) (Finnegan *et al.* 1998; Laird 2003) et les modifications biochimiques des Histones (Fransz and de Jong 2002) qui peuvent être des Acétylation (Histone Acétyl Transférases (HAT)/ Histone Déacétylases (HDAC)), des Méthylations (Histone MéthylTransférases (HMT)) (Kinases et Phosphatases) (Martin and Zhang 2005) ou des Phosphorylations (He and Amasino 2005). L'acétylation favorise l'expression des gènes tandis que la méthylation la réprime.

Nous avons mis en évidence des gènes régulés codant des Histones et des gènes suggérant une implication des mécanismes de méthylation au cours du stress. Il est tout à fait envisageable que de tels contrôles épigénétiques interviennent dans l'expression différentielle des gènes entre les deux écotypes. Des études sur les niveaux de méthylation pourraient donc se révéler intéressantes.

Bibliographie

Abe, H., Yamaguchi-Shinozaki, K., Urao, T., Iwasaki, T., Hosokawa, D. and Shinozaki, K. (1997). "Role of Arabidopsis MYC and MYB homologs in drought- and abscisic acid-regulated gene expression." Plant Cell **9**(10): 1859-1868.

Akiyama, T. and Pillai, M.A. (2003). "Isolation and characterization of a gene for a repetitive proline rich protein (OsPRP) down-regulated during submergence in rice (Oryza sativa)." Physiol. Plantarum **118**(4): 507-513.

Alazard, P. and Raffin, A. (2002). Les gains génétiques des première et deuxième générations de vergers. ONF : Le progrès génétique en forêt. P. Edition**:** 27-40.

Alexandersson, E., Fraysse, L., Sjövall-Larsen, S., Gustavsson, S., Fellert, M., Karlsson, M., Johanson, U. and Kjellbom, P. (2005). "Whole Gene Family Expression and Drought Stress Regulation of Aquaporins." Plant Mol. Biol. **59**(3): 469-484.

Alia, R., Gil, L. and Pardos, J.A. (1993). "Performance of 43 *Pinus pinaster* on 5 Locations in Central Spain." Silvae Genet. **44**(2-3): 75-80.

Aloni, R., Langhans, M., Aloni, E. and Ullrich, C. (2004). "Role of cytokinin in the regulation of root gravitropism." Planta **220**(1): 177-182.

Alonso, J.M., Stepanova, A.N., Leisse, T.J.*, et al.* (2003). "Genome-Wide Insertional Mutagenesis of Arabidopsis thaliana." Science **301**(5633): 653-657.

Alscher, R.G., Erturk, N. and Heath, L.S. (2002). "Role of superoxide dismutases (SODs) in controlling oxidative stress in plants." J. Exp. Bot. **53**(372): 1331-1341.

Amasino, R. (2005). "1955: Kinetin Arrives. The 50th Anniversary of a New Plant Hormone." Plant Physiol. **138**(3): 1177-1184.

Aparicio, F., Thomas, C.L., Lederer, C., Niu, Y., Wang, D. and Maule, A.J. (2005). "Virus induction of heat shock protein 70 reflects a general response to protein accumulation in the plant cytosol." Plant Physiol. **138**(1).

Asai, T., Tena, G., Plotnikova, J., Willmann, M.R., Chiu, W.L., Gomez-Gomez, L., Boller, T., Ausubel, F.M. and Sheen, J. (2002). "MAP kinase signalling cascade in Arabidopsis innate immunity." Nature **415**(6875): 977-983.

Asano, T., Masumura, T., Kusano, H., Kikuchi, S., Kurita, A., Shimada, H. and Kadowaki, K.-i. (2002). "Construction of a specialized cDNA library from plant cells isolated by laser capture microdissection: toward comprehensive analysis of the genes expressed in the rice phloem." Plant J. **32**(3): 401-408.

Ascenzi, R. and Gantt, J.S. (1997). "A drought-stress-inducible histone gene in Arabidopsis thaliana is a member of a distinct class of plant linker histone variants." Plant Mol. Biol. **34**(4): 629-41.

Bacon, M.A., Ed. (2004). Water use efficiency in plant biology. Biological Sciences Series. Oxford, Blackwell Publishing.

Baradat, P. and Pastuszka, P. (1992). Le pin maritime. Amélioration des espèces végétales cultivées. A. Gallais and H. Bannerot. Paris, INRA Editions**:** 695-709.

Barkla, B.J., Vera-Estrella, R., Maldonado-Gama, M. and Pantoja, O. (1999). "Abscisic acid induction of vacuolar H+-ATPase activity in Mesembryanthemum crystallinum is developmentally regulated." Plant Physiol. **120**(3): 811-819.

Bartel, B. (1997). "Auxin Biosynthesis." Annu. Rev. Plant Phys. Plant Mol. Biol. **48**(1): 51-66.

Bartels, D. and Salamini, F. (2001). "Desiccation tolerance in the resurrection plant Craterostigma plantagineum. A contribution to the study of drought tolerance at the molecular level." Plant Physiol. **127**(4): 1346-1353.

Benjamini, Y. and Hochberg, Y. (1995). "Controlling the false discovering rate : a practical and powerful approach to multiple testing." Journal of the Royal Statistical Society Series B **57**: 289-300.

Bewley, J.D. (1997). "Seed Germination and Dormancy." Plant Cell **9**(7): 1055-1066.

Bianchi, M.W., Damerval, C. and Vartanian, N. (2002). "Identification of proteins regulated by cross-talk between drought and hormone pathways in Arabidopsis wild-type and auxin-insensitive mutants, axr1 and axr2." Funct. Plant Biol. **29**(1): 55-61.

Borevitz, J.O. and Chory, J. (2004). "Genomics tools for QTL analysis and gene discovery." Curr. Opin. Plant Biol. **7**(2): 132-136.

Bostock, R.M. and Quatrano, R.S. (1992). "Regulation of Em Gene Expression in Rice : Interaction between Osmotic Stress and Abscisic Acid." Plant Physiol. **98**(4): 1356-1363.

Bouche, N., Yellin, A., Snedden, W.A. and Fromm, H. (2005). "Plant-Specific Calmodulin-Binding Proteins." Annu. Rev. Plant. Biol. **56**(1): 435-466.

Boudsocq, M. and Lauriere, C. (2005). "Osmotic Signaling in Plants. Multiple Pathways Mediated by Emerging Kinase Families." Plant Physiol. **138**(3): 1185-1194.

Bray, E.A. (2002). "Abscisic acid regulation of gene expression during water-deficit stress in the era of the Arabidopsis thaliana genome." Plant, Cell Environ. **25**(2): 153-161.

Bray, E.A. (2004). "Genes commonly regulated by water-deficit stress in Arabidopsis thaliana." J. Exp. Bot. **55**(407): 2331-2341.

Brendel, O., Pot, D., Plomion, C., Rozenberg, P. and Guehl, J.M. (2002). "Genetic parameters and QTL analysis of delta 13C and ring width in maritime pine." Plant, Cell Environ. **25**(8): 945-953.

Broadmeadow, M.S.J., Ray, D. and Samuel, C.J.A. (2005). "Climate change and the future for broadleaved tree species in Britain." Forestry **78**(2): 145-161.

Brunner, A.M., Busov, V.B. and Strauss, S.H. (2004). "Poplar genome sequence: functional genomics in an ecologically dominant plant species." Trends Plant Sci. **9**(1): 49-56.

Buchanan, C.D., Lim, S., Salzman, R.A.*, et al.* (2005). "Sorghum bicolor's Transcriptome Response to Dehydration, High Salinity and ABA." Plant Mol. Biol. **58**(5): 699-720.

Buckley, T.N. (2005). "The control of stomata by water balance." New Phytol. **168**(2): 275-292.

Burban, C. and Petit, R.J. (2003). "Phylogeography of maritime pine inferred with organelle markers having contrasted inheritance." Mol. Ecol. **12**(6): 1487-1495.

Burghardt, M. and Riederer, M. (2003). "Ecophysiological relevance of cuticular transpiration of deciduous and evergreen plants in relation to stomatal closure and leaf water potential." J. Exp. Bot. **54**(389): 1941-1949.

Burke, J., Davison, D. and Hide, W. (1999). "d2_cluster: A Validated Method for Clustering EST and Full-Length cDNA Sequences." Genome Res. **9**(11): 1135-1142.

Burke, J., Wang, H., Hide, W. and Davison, D.B. (1998). "Alternative Gene Form Discovery and Candidate Gene Selection from Gene Indexing Projects." Genome Res. **8**(3): 276-290.

Butcher, L.M., Meaburn, E., Liu, L., Fernandes, C., Hill, L., Al-Chalabi, A., Plomin, R., Schalkwyk, L. and Craig, I.W. (2004). "Genotyping Pooled DNA on Microarrays: A Systematic Genome Screen of Thousands of SNPs in Large Samples to Detect QTLs for Complex Traits." Behavior Genetics **34**(5): 549-555.

Calamassi, R., Rocca, G.D., Falusi, M., Paolettiinstb, E. and Strati, S. (2001). "Resistance to water stress in seedlings of eight European provenances of Pinus halepensis Mill." Ann.Sci.For. **58**(6): 663-672.

Campalans, A., Messeguer, R., Goday, A. and Pages, M. (1999). "Plant responses to drought, from ABA signal transduction events to the action of the induced proteins." Plant Physiol. Bioch. **37**(5): 327-340.

Cardon, L.R. and Bell, J.I. (2001). "Association Study Designs for Complex Diseases." Nat. Rev. Genet. **2**(2): 91-99.

Caruso, A., Morabito, D., Delmotte, F., Kahlem, G. and Carpin, S. (2002). "Dehydrin induction during drought and osmotic stress in Populus." Plant Physiol. Bioch. **40**(12): 1033-1042.

Cassab, G.I. (1998). "Plant Cell Wall Proteins." Annu. Rev. Plant Phys. Plant Mol. Biol. **49**(1): 281-309.

Casson, S., Spencer, M., Walker, K. and Lindsey, K. (2005). "Laser capture microdissection for the analysis of gene expression during embryogenesis of Arabidopsis." Plant J. **42**(1): 111-123.

Chagne, D. (2004). Développement de marqueurs moléculaires chez le pin maritime (Pinus pinaster Ait.) et cartographie génétique comparée des conifères. Nnacy, Université Henri Poincaré, Nancy I, Thèse de Doctorat, 166.

Chang, C. and Bleecker, A.B. (2004). "Ethylene Biology. More Than a Gas." Plant Physiol. **136**(2): 2895-2899.

Chang, C. and Stadler, R. (2001). "Ethylene hormone receptor action in Arabidopsis." BioEssays **23**(7): 619-627.

Chang, S., Puryear, J. and Cairney, J. (1993). "A Simple and Efficient Method for Isolating RNA from Pine Trees." Plant Mol. Biol. Rep. **11**(2): 113-116.

Chang, S., Puryear, J.D., Dias, M.A.D.L., Funkhouser, E.A., Newton, R.J. and Cairney, J. (1996). "Gene expression under water deficit in loblolly pine (Pinus taeda): isolation and characterization of cDNA clones." Physiol. Plantarum **97**(1): 139-148.

Chaves, M.M., Maroco, J.P. and Pereira, J.S. (2003). "Understanding plant responses to drought — from genes to the whole plant." Funct. Plant Biol. **30**: 239-264.

Chen, L. and Bush, D.R. (1997). "LHT1, a lysine- and histidine-specific amino acid transporter in Arabidopsis." Plant Physiol. **115**(3): 1127-1134.

Chen, S., Wang, S., Huttermann, A. and Altman, A. (2002). "Xylem abscisic acid accelerates leaf abscission by modulating polyamine and ethylene synthesis in water-stressed intact poplar." Trees-Struct. Funct. **16**(1): 16-22.

Chen, Z., Hong, X., Zhang, H., Wang, Y., Li, X., Zhu, J.-K. and Gong, Z. (2005). "Disruption of the cellulose synthase gene, AtCesA8/IRX1, enhances drought and osmotic stress tolerance in Arabidopsis." Plant J. **43**(2): 273-283.

Choi, H., Hong, J., Ha, J., Kang, J. and Kim, S. (2000). "ABFs, a family of ABA-responsive element binding factors." J. Biol.Chem. **275**(3): 1723-1730.

Cochard, H., Froux, F., Mayr, S. and Coutand, C. (2004). "Xylem Wall Collapse in Water-Stressed Pine Needles." Plant Physiol. **134**(1): 401-408.

Correa-Aragunde, N., Graziano, M. and Lamattina, L. (2004). "Nitric oxide plays a central role in determining lateral root development in tomato." Planta **218**(6): 900-905.

Costa, P. (1999). Réponse moléculaire, physiologique et génétique du pin maritime à une contrainte hydrique. Nancy, Université Henri Poincaré, Nancy I, Thèse de Doctorat, 116.

Costa, P., Bahrman, N., Frigerio, J.-M., Kremer, A. and Plomion, C. (1998). "Water-deficit-responsive proteins in maritime pine." Plant Mol. Biol. **38**(4): 587-596.

Couée, I., Hummel, I., Sulmon, C., Gouesbet, G. and El Amrani, A. (2004). "Involvement of polyamines in root development." Plant Cell, Tissue and Organ Culture **76**(1): 1-10.

Crookshanks, M., Emmersen, J., Welinder, K.G. and Lehmann Nielsen, K. (2001). "The potato tuber transcriptome: analysis of 6077 expressed sequence tags." FEBS Lett. **506**(2): 123-126.

Crozier, A., Kamiya, Y., Bishop, G. and Yokota, T. (2000). Biosynthesis of Hormones and Elicitor Molecules. Biochemistry & Molecular Biology of Plants. B. B. Buchanan, W. Gruissem and L. J. Russell. Berkeley, American Society of Plant Physiologists**:** 850-929.

Cruiziat, P. and Tyree, M.T. (1990). "La montée de la sève dans les arbres." La Recherche **21**(220): 406-414.

Cushman, J.C. and Bohnert, H.J. (2000). "Genomic approaches to plant stress tolerance." Curr. Opin. Plant Biol. **3**(2): 117-124.

Damerval, C., De Vienne, D., Zivy, M. and Thiellement, H. (1986). "Technical improvements in two-dimensional electrophoresis increase the level of genetic variation detected in wheat-seedling proteins." Electrophoresis **7**(1): 52-54.

Danjon, F., Bert, D., Godin, C. and Trichet, P. (1999). "Structural root architecture of 5-year-old *Pinus pinaster* measured by 3D digitising and analysed with AMAPmod." Plant Soil **217**(1 - 2): 49-63.

Danjon, F., Fourcaud, T. and Bert, D. (2005). "Root architecture and wind-firmness of mature Pinus pinaster." New Phytol. **168**(2): 387-400.

Day, R.C., Grossniklaus, U. and Macknight, R.C. (2005). "Be more specific! Laser-assisted microdissection of plant cells." Trends Plant Sci. **10**(8): 397-406.

Deak, K.I. and Malamy, J. (2005). "Osmotic regulation of root system architecture." Plant J. **43**(1): 17-28.

Delzon, S., Sartore, M., Burlett, R., Dewar, R. and Loustau, D. (2004). "Hydraulic responses to height growth in maritime pine trees." Plant, Cell Environ. **27**(9): 1077-1087.

Diatchenko, L., Lau, Y.-F.C., Campbell, A.P.*, et al.* (1996). "Suppression subtractive hybridization: A method for generating differentially regulated or tissue-specific cDNA probes and libraries." PNAS **93**(12): 6025-6030.

Dong, H.-P., Yu, H., Bao, Z., Guo, X., Peng, J., Yao, Z., Chen, G., Qu, S. and Dong, H. (2005). "The ABI2-dependent abscisic acid signalling controls HrpN-induced drought tolerance in Arabidopsis." Planta **221**(3): 313-327.

Dong, J., Chen, C. and Chen, Z. (2003). "Expression profiles of the *Arabidopsis* WRKY gene superfamily during plant defense response." Plant Mol. Biol. **51**(1): 21-37.

Draghici, S. (2003). Data analysis tools for DNA microarrays, Chapman & Hall / CRC.477.

Dubos, C. (2001). Réponse moléculaire de jeunes plants de pin maritime soumis à un stress hydrique en milieu hydroponique, Henri Poincaré, Nancy 1.

Dubos, C. and Plomion, C. (2001). "Drought differentially affects expression of a PR-10 protein, in needles of maritime pine (Pinus pinaster Ait.) seedlings." J. Exp. Bot. **52**(358): 1143-1144.

Dubos, C. and Plomion, C. (2003). "Identification of water-deficit responsive genes in maritime pine (Pinus pinaster Ait.) roots." Plant Mol. Biol. **51**(2): 249-262.

Dubos, C., Provost, G.l., Pot, D., Salin, F., Lalane, C., Madur, D., Frigerio, J.M. and Plomion, C. (2003). "Identification and characterization of water-stress-responsive genes in hydroponically grown maritime pine (Pinus pinaster) seedlings." Tree Physiol. **23**(3): 169-179.

Dugas, D.V. and Bartel, B. (2004). "MicroRNA regulation of gene expression in plants." Curr. Opin. Plant Biol. **7**(5): 512-520.

Eapen, D., Barroso, M.L., Ponce, G., Campos, M.E. and Cassab, G.I. (2005). "Hydrotropism: root growth responses to water." Trends Plant Sci. **10**(1): 44-50.

El Hadj Moussa, F. (1989). Circulation de l'eau dans un ecosysteme de pin maritime en landes de gascogne, Thèse de Doctorat.

el-Khatib, R.T., Hamerlynck, E.P., Gallardo, F. and Kirby, E.G. (2004). "Transgenic poplar characterized by ectopic expression of a pine cytosolic glutamine synthetase gene exhibits enhanced tolerance to water stress." Tree Physiol. **24**(7): 729-736.

Elsik, C.G. and Williams, C.G. (2000). "Retroelements contribute to the excess low-copy-number DNA in pine." Mol. Genet. Genomics **264**(1 - 2): 47-55.

Enckevort, L., Droc, G.t., Piffanelli, P., *et al.* (2005). "EU-OSTID: A Collection of Transposon Insertional Mutants for Functional Genomics in Rice." Plant Mol. Biol. **59**(1): 99-110.

Faldt, J., Martin, D., Miller, B., Suman, R. and Bohlmann, J. (2003). "Traumatic resin defense in Norway spruce (Picea abies): methyl jasmonate-induced terpene synthase gene expression, and cDNA cloning and functional characterization of (+)-3-carene synthase." Plant Mol. Biol. **51**(1): 119-133.

Ferry-Dumazet, H., Houel, G., Montalent, P., *et al.* (2005). "PROTICdb: A web-based application to store, track, query, and compare plant proteome data." Proteomics **5**(8): 2069-2081.

Finnegan, E.J., Genger, R.K., Kovac, K., Peacock, W.J. and Dennis, E.S. (1998). "DNA methylation and the promotion of flowering by vernalization." PNAS **95**(10): 5824-5829.

Finnegan, E.J. and Kovac, K.A. (2000). "Plant DNA methyltransferases." Plant Mol. Biol. **43**(2 - 3): 189-201.

Finnegan, E.J., Peacock, W.J. and Dennis, E.S. (1996). "Reduced DNA methylation in Arabidopsis thaliana results in abnormal plant development." PNAS **93**(16): 8449-8454.

Fransz, P.F. and de Jong, J.H. (2002). "Chromatin dynamics in plants." Curr. Opin. Plant Biol. **5**(6): 560-567.

Froux, F. (2002). Caractéristiques hydrauliques, régulation stomatique et efficience d'utilisation de l'eau de quatre espèces de conifères méditerranéens (*Cupressus sempervirens, Cedrus atlantica, Pinus halepensis et Pinus nigra*). Nancy, Université Henri Poincaré - Nancy I, Thèse de Doctorat, 140.

Frugoli, J.A., Zhong, H., Nuccio, M.L., McCourt, P., McPeek, M.A., Thomas, T.L. and McClung, C.R. (1996). "Catalase is encoded by a multigene family in Arabidopsis thaliana (L.) Heynh." Plant Physiol. **112**(1): 327-336.

Fukao, T. and Bailey-Serres, J. (2004). "Plant responses to hypoxia – is survival a balancing act?" Trends Plant Sci. **9**(9): 449-456.

Galau, G.A., Hughes, D.W. and Dure, L., III (1986). "Abscisic acid induction of cloned cotton late embryogenesis-abundant (Lea) mRNAs." Plant Mol. Biol. **7**(3).

Gallardo, F., Fu, J., Jing, Z.P., Kirby, E.G. and Canovas, F.M. (2003). "Genetic modification of amino acid metabolism in woody plants." Plant Physiol. Bioch. **41**(6-7): 587-594.

Gang, D.R., Kasahara, H., Xia, Z.-Q., Vander Mijnsbrugge, K., Bauw, G., Boerjan, W., Van Montagu, M., Davin, L.B. and Lewis, N.G. (1999). "Evolution of Plant Defense Mechanisms. Relationships of Phenyl Coumaran Benzylic Ether Reductases to Pinoresinol-Lariciresinol and Isoflavone Reductases." J. Biol.Chem. **274**(11): 7516-7527.

Garbaye, J. (2000). "The role of ectomycorrhizal symbiosis in the resistance of forests to water stress." Outlook Agr. **29**(1): 63-63.

Gaudillere, J.-P. and Chasles, M. (1986). "Critères de choix de lampes pour culture de plantes en conditions contrôlées." Cah. Techn. INRA **13**: 27-36.

Gaxiola, R.A., Li, J., Undurraga, S., Dang, L.M., Allen, G.J., Alper, S.L. and Fink, G.R. (2001). "Drought- and salt-tolerant plants result from overexpression of the AVP1 H+-pump." PNAS **98**(20): 11444-11449.

GIEC (2001). Changements climatiques 2001: Rapport de synthèse. Résumé à l'intention des décideurs. Wembley, GIEC, 37.

Gion, J.-M., Lalanne, C., Le Provost, G., *et al.* (2005). "The proteome of maritime pine wood forming tissue." Proteomics **5**(14): 3731-3751.

Goddijn, O.J.M. and Dun, K.v. (1999). "Trehalose metabolism in plants." Trends Plant Sci. **4**(8): 315-319.

Gong, Z., Koiwa, H., Cushman, M.A., *et al.* (2001). "Genes That Are Uniquely Stress Regulated in Salt Overly Sensitive (sos) Mutants." Plant Physiol. **126**(1): 363-375.

Gorham, J., Hughes, L. and Wyn Jones, R.G. (1981). "Low-molecular-weight carbohydrates in some salt-stressed plants." Physiol. PLANT. **53**: 27-33.

Grelet, J., Benamar, A., Teyssier, E., Avelange-Macherel, M.-H., Grunwald, D. and Macherel, D. (2005). "Identification in Pea Seed Mitochondria of a Late-Embryogenesis Abundant Protein Able to Protect Enzymes from Drying." Plant Physiol. **137**(1): 157-167.

Grichko, V.P. and Glick, B.R. (2001a). "Amelioration of flooding stress by ACC deaminase-containingplant growth-promoting bacteria." Plant Physiol. Bioch. **39**: 11-17.

Grichko, V.P. and Glick, B.R. (2001b). "Flooding tolerance of transgenic tomato plants expressing the bacterial enzyme ACC deaminase controlled by the 35S, rolD or PRB-1b promoter." Plant Physiol. Bioch. **39**: 19-25.

Guan, L. and Scandalios, J.G. (1998). "Two structurally similar maize cytosolic superoxide dismutase genes, Sod4 and Sod4A, respond differentially to abscisic acid and high osmoticum." Plant Physiol. **117**(1): 217-24.

Guehl, J.M., Picon, C., Aussenac, G. and Gross, P. (1994). "Interactive effects of elevated CO2 and soil drought on growth and transpiration efficiency and its determinants in two European forest tree species." Tree Physiol. **14**: 707-724.

Guyon, J.P. (1980). Variabilité géographique et ecophysiologique du pin maritime (Pinus pinaster AIT.), Ecole Nationale des Ingénieurs des Travaux des Eaux et Forêts, Mémoire de fin d'étude, 80.

Guyon, J.P. (1982). Croissance en hauteur et comportement hydrique de sept provenances de Pin maritime en plantations comparatives et en milieu contrôlé, Bordeaux II, DEA.

Hare, P.D., Cress, W.A. and Staden, J.V. (1998). "Dissecting the roles of osmolyte accumulation during stress." Plant, Cell Environ. **21**(6): 535-553.

Harrak, H., Chamberland, H., Plante, M., Bellemare, G., Lafontaine, J.G. and Tabaeizadeh, Z. (1999). "A proline-, threonine-, and glycine-rich protein down-regulated by drought is localized in the cell wall of xylem elements." Plant Physiol. **121**(2): 557-564.

He, Y. and Amasino, R.M. (2005). "Role of chromatin modification in flowering-time control." Trends Plant Sci. **10**(1): 30-35.

Heath, L.S., Ramakrishnan, N., Sederoff, R.R., *et al.* (2002). "Studying the functional genomics of stress responses in loblolly pine with the Expresso microarray experiment management system." Compar. Funct. Genom. **3**(3): 226-243.

Heller, R., Esnault, R. and Lance, C. (1993). Physiologie Végétale - 1 - Nutrition. Paris, MASSON.294.

Hilbricht, T., Salamini, F. and Bartels, D. (2002). "CpR18, a novel SAP-domain plant transcription factor, binds to a promoter region necessary for ABA mediated expression of the CDeT27-45 gene from the resurrection plant Craterostigma plantagineum Hochst." Plant J. **31**(3): 293-303.

Hirayama, T., Ohto, C., Mizoguchi, T. and Shinozaki, K. (1995). "A gene encoding a phosphatidylinositol-specific phospholipase C is induced by dehydration and salt stress in Arabidopsis thaliana." PNAS **92**(9): 3903-7.

Hirochika, H., Guiderdoni, E., An, G.*, et al.* (2004). "Rice Mutant Resources for Gene Discovery." Plant Mol. Biol. **54**(3): 325-334.

Hontzeas, N., Zoidakis, J., Glick, B.R. and Abu-Omar, M.M. (2004). "Expression and characterization of 1-aminocyclopropane-1-carboxylate deaminase from the rhizobacterium Pseudomonas putida UW4: a key enzyme in bacterial plant growth promotion." BBA-Proteins Proteom. **1703**(1): 11-19.

Huang, J., Lin, S. and Wang, C. (2000). "A pollen-specific and desiccation-associated transcript in Lilium longiflorum during development and stress." Plant Cell Physiol. **41**(4): 477-485.

Huber, S.C. and Huber, J.L. (1996). "Role and regulation of sucrose-phosphate synthase in higher plants." Annu. Rev. Plant Phys. Plant Mol. Biol. **47**: 431-444.

Huh, G.H., Matsuura, Y., Meshi, T. and Iwabuchi, M. (1995). "Differential expression of the two types of histone H2A genes in wheat." BBA-Gene Struct. Expr. **1261**(1): 155-160.

Hukin, D., Cochard, H., Dreyer, E., Thiec, D.L. and Bogeat-Triboulot, M.B. (2005). "Cavitation vulnerability in roots and shoots: does Populus euphratica Oliv., a poplar from arid areas of Central Asia, differ from other poplar species?" J. Exp. Bot. **56**(418): 2003-2010.

Hurkman, W.J., Tao, H.P. and Tanaka, K. (1991). "Germin-Like Polypeptides Increase in Barley Roots during Salt Stress." Plant Physiol. **97**: 366-374.

Iba, K. (2002). "Acclimative Response to Temperature Stress in Higher Plants: Approaches of Gene Engineering for Temperature Tolerance." Annu. Rev. Plant. Biol. **53**(1): 225-245.

Ichimura, K., Mizoguchi, T., Irie, K., Morris, P., Giraudat, J., Matsumoto, K. and Shinozaki, K. (1998). "Isolation of ATMEKK1 (a MAP kinase kinase kinase)-interacting proteins and analysis of a MAP kinase cascade in Arabidopsis." Biochemical & Biophysical Research Communications. **253**(2): 532-43.

Igarashi, Y., Yoshiba, Y., Takeshita, T., Nomura, S., Otomo, J., Yamaguchi-Shinozaki, K. and Shinozaki, K. (2000). "Molecular cloning and characterization of a cDNA encoding proline transporter in rice." Plant Cell Physiol. **41**(6).

Imoto, K., Yokoyama, R. and Nishitani, K. (2005). "Comprehensive approach to genes involved in cell wall modifications in Arabidopsis thaliana." Plant Mol. Biol. **58**(2): 177-192.

Ingram, J. and Bartels, D. (1996). "The molecular basis of dehydration tolerance in plants." Annu. Rev. Plant Phys. Plant Mol. Biol. **47**: 377-403.

Initiative, T.A.G. (2000). "Analysis of the genome sequence of the flowering plant Arabidopsis thaliana." Nature **408**(6814): 796-815.

Inoue, T., Higuchi, M., Hashimoto, Y., Seki, M., Kobayashi, M., Kato, T., Tabata, S., Shinozaki, K. and Kakimoto, T. (2001). "Identification of CRE1 as a cytokinin receptor from Arabidopsis." Nature **409**(6823): 1060-1063.

Ishitani, M., Majumder, A.L., Bornhouser, A., Michalowski, C.B., Jensen, R.G. and Bohnert, H.J. (1996). "Coordinate transcriptional induction of myo-inositol metabolism during environmental stress." Plant J. **9**(4): 537-548.

Jang, J.Y., Kim, D.G., Kim, Y.O., Kim, J.S. and Kang, H. (2004). "An Expression Analysis of a Gene Family Encoding Plasma Membrane Aquaporins in Response to Abiotic Stresses in *Arabidopsis Thaliana*." Plant Mol. Biol. **54**(5): 713-725.

Jespersen, H.M., Kjaersgard, I.V.H., Ostergaard, L. and Welinder, K.G. (1997). "From sequence analysis of three novel ascorbate peroxidases from Arabidopsis thaliana to structure, function and evolution of seven types of ascorbate peroxidase." Biochem. J. **326**(2): 305-310.

Johnson, J.D., Tognetti, R. and Paris, P. (2002). "Water relations and gas exchange in poplar and willow under water stress and elevated atmospheric CO2." Physiol. Plantarum **115**(1): 93-100.

Johnson, K.L., Jones, B.J., Bacic, A. and Schultz, C.J. (2003). "The Fasciclin-Like Arabinogalactan Proteins of Arabidopsis. A Multigene Family of Putative Cell Adhesion Molecules." Plant Physiol. **133**(4): 1911-1925.

Johnson, P.R. and Ecker, J.R. (1998). "The Ethylene Gas Signal Transduction Pathway : A Molecular Perspective." Annu. Rev. Genet. **32**(1): 227-254.

Jones, A.M. (2002). "G-protein-coupled signaling in Arabidopsis." Curr. Opin. Plant Biol. **5**(5): 402-407.

Kawasaki, S., Borchert, C., Deyholos, M., Wang, H., Brazille, S., Kawai, K., Galbraith, D. and Bohnert, H.J. (2001). "Gene expression profiles during the initial phase of salt stress in rice." Plant Cell **13**(4): 889-905.

Kidner, C.A. and Martienssen, R.A. (2005). "The developmental role of microRNA in plants." Curr. Opin. Plant Biol. **8**(1): 38-44.

Kieber, J.J. (1997). "The Ethylen Response Pathway in Arabidopsis." Annu. Rev. Plant Phys. Plant Mol. Biol. **48**(1): 277-296.

Kizis, D. and Pages, M. (2002). "Maize DRE-binding proteins DBF1 and DBF2 are involved in rab17 regulation through the drought-responsive element in an ABA-dependent pathway." Plant J. **30**(6): 679-689.

Knight, H. and Knight, M.R. (2001). "Abiotic stress signalling pathways: specificity and cross-talk." Trends Plant Sci. **6**(6): 262-267.

Koch, G.W., Sillett, S.C., Jennings, G.M. and Davis, S.D. (2004). "The limits to tree height." Nature **428**(6985): 851-854.

Komis, G., Apostolakos, P. and Galatis, B. (2002). "Hyperosmotic Stress Induces Formation of Tubulin Macrotubules in Root-Tip Cells of Triticum turgidum: Their Probable Involvement in Protoplast Volume Control." Plant Cell Physiol. **43**(8): 911-922.

Koster, K.L. and Leopold, A.C. (1988). "Sugars and Desiccation Tolerance in Seeds." Plant Physiol. **88**(3): 829-832.

Kramer, P.J. (1983). Water Relations of Plants. San Diego, Academic Press, Inc.489.

Kudla, J., Xu, Q., Harter, K., Gruissem, W. and Luan, S. (1999). "Genes for calcineurin B-like proteins in Arabidopsis are differentially regulated by stress signals." PNAS **96**(8): 4718-23.

Kumar, S., Tamura, K. and Nei, M. (2004). "MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment." Brief. Bioinform. **5**(14): 150-163.

Kwart, M., Hirner, B., Hummel, S. and Frommer, W.B. (1993). "Differential expression of two related amino acid transporters with differing substrate specificity in Arabidopsis thaliana." Plant J. **4**(6): 993-1002.

Laird, P.W. (2003). "The Power and the Promise of DNA Methylation Markers." Nature Reviews Cancer **3**(4): 253-266.

Lamhamedi, M.S., Bernier, P.Y. and Fortin, J.A. (1992). "Hydraulic conductance and soil water potential at the soil-root interface of Pinus pinaster seedlings inoculated with different dikaryons of Pisolithus sp." Tree Physiol. **10**: 231-244.

Laporte, M.M., Shen, B. and Tarczynski, M.C. (2002). "Engineering for drought avoidance: expression of maize NADP-malic enzyme in tobacco results in altered stomatal function." J. Exp. Bot. **53**(369): 699-705.

Le Dantec, L., Chagné, D., Pot, D., *et al.* (2004). "Automated SNP Detection in Expressed Sequence Tags: Statistical Considerations and Application to Maritime Pine Sequences." Plant Mol. Biol. **54**(3): 461-470.

Lee, U., Wie, C., Escobar, M., Williams, B., Hong, S. and Vierling, E. (2005). "Genetic analysis reveals domain interactions of Arabidopsis Hsp100/ClpB and cooperation with the small heat shock protein chaperone system." Plant Cell **17**(2).

Leung, J. and Giraudat, J. (1998). "Abscissic Acid Signal Transduction." Annu. Rev. Plant Phys. Plant Mol. Biol. **49**(1): 199-222.

Li, J., Wang, X., Watson, M.B. and Assmann, S.M. (2000). "Regulation of abscisic acid-induced stomatal closure and anion channels by guard cell AAPK kinase." Science **287**(5451): 300-303.

Li, Y., Jones, L. and McQueen-Mason, S. (2003). "Expansins and cell growth." Curr. Opin. Plant Biol. **6**(6): 603-610.

Light, G., Mahan, J., Roxas, V. and Allen, R. (2005). "Transgenic cotton (Gossypium hirsutum L.) seedlings expressing a tobacco glutathione S-transferase fail to provide improved stress tolerance." Planta **222**(2): 346-354.

Llusia, J. and Penuelas, J. (2000). "Seasonal patterns of terpene content and emission from seven Mediterranean woody species in field conditions." Am. J. Bot. **87**(1): 133-140.

Llusià, J. and Penuelas, J. (1998). "Changes in terpene content and emission in potted Mediterranean woody plants under severe drought." Can. J. Bot. **76**(8): 1366-1373.
Lo Bianco, R., Rieger, M. and Sung, S.-J.S. (2000). "Effect of drought on sorbitol and sucrose metabolism in sinks and sources of peach." Physiol. Plantarum **108**(1): 71-78.
Loewus, F.A. and Murthy, P.P.N. (2000). "myo-Inositol metabolism in plants." Plant Sci. **150**(1): 1-19.
Loidl, P. (2004). "A plant dialect of the histone language." Trends Plant Sci. **9**(2): 84-90.
Loustau, D., Bosc, A., Colin, A.*, et al.* (2005). "Modeling climate change effects on the potential production of French plains forests at the sub-regional level." Tree Physiol. **25**(7): 813-823.
Loustau, D., Crepeau, S., Guye, M.G., Sartore, M. and Saur, E. (1995). "Growth and water relations of three geographically separate origins of maritime pine (Pinus pinaster Ait) under saline condition." Tree Physiol. **15**(9): 569-576.
Luan, S. (2002). "Signalling drought in guard cells." Plant, Cell Environ. **25**(2): 229-237.
Ma, W., Charles, T.C. and Glick, B.R. (2004). "Expression of an Exogenous 1-Aminocyclopropane-1-Carboxylate Deaminase Gene in Sinorhizobium meliloti Increases Its Ability To Nodulate Alfalfa." Appl. Environ. Microbiol. **70**(10): 5891-5897.
Ma, W., Guinel, F.C. and Glick, B.R. (2003). "Rhizobium leguminosarum Biovar viciae 1-Aminocyclopropane-1-Carboxylate Deaminase Promotes Nodulation of Pea Plants." Appl. Environ. Microbiol. **69**(8): 4396-4402.
Maillet, M. (1995). Biologie cellulaire. Paris, MASSON.328.
Marcotte Jr, W.R., Russell, S.H. and Quatrano, R.S. (1989). "Abscisic Acid-Responsive Sequences from the Em Gene of Wheat." Plant Cell **1**(10): 969-976.
Marè, C., Mazzucotelli, E., Crosatti, C., Francia, E., Stanca, A.â.m. and Cattivelli, L. (2004). "Hv-WRKY38: a new transcription factor involved in cold- and drought-response in barley." Plant Mol. Biol. **55**(3): 399-416.
Marjanovic, Z., Uehlein, N., Kaldenhoff, R., Zwiazek, J., Weiss, M., Hampp, R. and Nehls, U. (2005). "Aquaporins in poplar: What a difference a symbiont makes!" Planta **222**(2): 258-268.
Marron, N., Villar, M., Dreyer, E., Delay, D., Boudouresque, E., Petit, J.-M., Delmotte, F., Guehl, J.-M. and Brignolas, F. (2005). "Diversity of leaf traits related to productivity in 31 Populus deltoides × Populus nigra clones." Tree Physiol. **25**(4): 425-435.
Martin, C. and Zhang, Y. (2005). "The Diverse Functions of Histone Lysine Methylation." Nat. Rev. Mol. Cell Biol. **6**(11): 838-849.
Martin, O. (2002). Approches statistiques pour l'analyse de données de puces à ADN. Grenoble, Université Joseph FOURIER, Thèse de Doctorat, 148.
Martin-Tanguy, J. (2001). "Metabolism and function of polyamines in plants: recent development (new approaches)." Plant Growth Regul. **34**(1): 135-148.
Matsuzaki, H., Dong, S., Loi, H.*, et al.* (2004a). "Genotyping over 100,000 SNPs on a pair of oligonucleotide arrays." Nature Methods **1**: 109-111.
Matsuzaki, H., Loi, H., Dong, S.*, et al.* (2004b). "Parallel Genotyping of Over 10,000 SNPs Using a One-Primer Assay on a High-Density Oligonucleotide Array." Genome Res. **14**(3): 414-425.
Maugé, J. (1987). Le Pin Maritime. Paris, Institut pour le développement.191.
Mayak, S., Tirosh, T. and Glick, B.R. (2004). "Plant growth-promoting bacteria confer resistance in tomato plants to salt stress." Plant Physiol. Bioch. **42**(6): 565-572.
McCourt, P. (1999). "Genetic Analysis of Hormone Signaling." Annu. Rev. Plant Phys. Plant Mol. Biol. **50**(1): 219-243.
Merkouropoulos, G. and Shirsat, A. (2003). "The unusual Arabidopsis extensin gene atExt1 is expressed throughout plant development and is induced by a variety of biotic and abiotic stresses." Planta **217**(3): 356-366.
Mikami, K., Katagiri, T., Iuchi, S., Yamaguchi-Shinozaki, K. and Shinozaki, K. (1998). "A gene encoding phosphatidylinositol-4-phosphate 5-kinase is induced by water stress and abscisic acid in Arabidopsis thaliana." Plant J. **15**(4): 563-568.
Milioni, D. and Hatzopoulos, P. (1997). "Genomic organization of hsp90 gene family in Arabidopsis." Plant Mol. Biol. **35**(6): 955-961.

Mizoguchi, T., Hayashida, N., Yamaguchi-Shinozaki, K., Kamada, H. and Shinozaki, K. (1993). "ATMPKs: a gene family of plant MAP kinases in Arabidopsis thaliana." FEBS Lett. **336**(3): 440-444.

Mizoguchi, T., Irie, K., Hirayama, T., Hayashida, N., Yamaguchi-Shinozaki, K., Matsumoto, K. and Shinozaki, K. (1996). "A gene encoding a mitogen-activated protein kinase kinase kinase is induced simultaneously with genes for a mitogen-activated protein kinase and an S6 ribosomal protein kinase by touch, cold, and water stress in Arabidopsis thaliana." PNAS **93**(2): 765-769.

Monclus, R., Dreyer, E., Delmotte, F.*, et al.* (2005). "Productivity, leaf traits and carbon isotope discrimination in 29 Populus deltoides x P. nigra clones." New Phytol. **167**(1): 53-62.

Montané, J. (1994). La Forêt Landaise : une aventure de l'homme et son milieu. Toulouse, Privats.

Moreau, M. (2005). Etude des stratégies de réponse moléculaire à un stress hydrique au niveau des racines chez deux écotypes de pin maritime à l'aide d'une approche protéomique. Bordeaux, Université de Bordeaux I, Rapport de 2eme année de Master, 16.

Mowla, S.B., Thomson, J.A., Farrant, J.M. and Mundree1, S.G. (2002). "A novel stress-inducible antioxidant enzyme identified from the resurrection plant Xerophyta viscosa Baker." Planta **215**: 716-726.

Murray, B.G. (1998). "Nuclear DNA amounts in gymnosperms." Ann. Bot. **82**(SupA): 3-15.

Nadeau, J.H. and Frank, W.N. (2000). "The roads from phenotypic variation to gene discovery: mutagenesis versus QTLs." Nat. Genet. **25**: 381.

Nakashima, K., Satoh, R., Kiyosue, T., Yamaguchi-Shinozaki, K. and Shinozaki, K. (1998). "A gene encoding proline dehydrogenase is not only induced by proline and hypoosmolarity, but is also developmentally regulated in the reproductive organs of Arabidopsis." Plant Physiol. **118**(4): 1233-1241.

Nakazono, M., Qiu, F., Borsuk, L.A. and Schnable, P.S. (2003). "Laser-Capture Microdissection, a Tool for the Global Analysis of Gene Expression in Specific Plant Cell Types: Identification of Genes Expressed Differentially in Epidermal Cells or Vascular Tissues of Maize." Plant Cell **15**(3): 583-596.

Nambara, E. and Marion-Poll, A. (2005). "Abscissic Acid Biosynthesis and Catabolism." Annu. Rev. Plant. Biol. **56**(1): 165-185.

Nanjo, T., Kobayashi, M., Yoshiba, Y., Kakubari, Y., Yamaguchi-Shinozaki, K. and Shinozaki, K. (1999). "Antisense suppression of proline degradation improves tolerance to freezing and salinity in Arabidopsis thaliana." FEBS Lett. **461**(3): 205-210.

Nayyar, H. and Chander, S. (2004). "Protective Effects of Polyamines against Oxidative Stress Induced by Water and Cold Stress in Chickpea." Journal of Agronomy and Crop Science **190**(5): 355-365.

Nelson, D.E., Koukoumanos, M. and Bohnert, H.J. (1999). "Myo-inositol-dependent sodium uptake in ice plant." Plant Physiol. **119**(1): 165-172.

Neta-Sharir, I., Isaacson, T., Lurie, S. and Weiss, D. (2005). "Dual role for tomato heat shock protein 21: protecting photosystem II from oxidative stress and promoting color changes during fruit maturation." Plant Cell **17**(6).

Nguyen, A. and Lamant, A. (1988). "Pinitol and Myo-inositol Accumulation in Water-stressed Seedlings of Maritime pine." Phytochemestry **27**(11): 3423-3427.

Nguyen, A. and Lamant, A. (1989a). "Effect of water stress on potassium distribution in young seedlings of maritime pine (*Pinus pinaster* Ait.)." Ann.Sci.For. **46**(suppl.): 379-383.

Nguyen, A. and Lamant, A. (1989b). "Variation in growth and osmotic regulation of water-stressed maritime pine (*Pinus pinaster* Ait.) provenances." Tree Physiol. **5**: 123-133.

Nguyen-Queyrens, A., Costa, P., Loustau, D. and Plomion, C. (2002). "Osmotic adjustment in Pinus pinaster cuttings in response to a soil drying cycle." Ann. For. Sci. **59**(7): 795-799.

Nguyen-Queyrens, A., Ferhi, A., Loustau, D. and Guehl, J.-M. (1998). "Within-ring $\delta^{13}C$ spatial variability and interannual variations in wood cellulose of two contrasting provenances of *Pinus pinaster*." Can. J. Forest. Res. **28**(5): 766-773.

Nielsen, R. (2005). "Molecular Signatures of Natural Selection." Annu. Rev. Genet. **39**(1): 197-218.

Nouaim, R. and Chaussod, R. (1996). Rôle des mycorhizes dans l'alimentation hydrique et minérale des plantes, notamment des ligneux de zones arides. La mycorhization des plantes forestières

en milieu aride et semi-aride et la lutte contre la désertification dans le bassin méditerranéen, CIHEAM-IAMZ. **20:** 21-26.

Nyporko, A.Y., Demchuk, O.N. and Blume, Y.B. (2003). "Cold adaptation of plant microtubules: structural interpretation of primary sequence changes in a highly conserved region of alpha - tubulin." Cell Biol. Int. **27**(3): 241-243.

Ozturk, Z.N., Talamé, V., Deyholos, M., Michalowski, C.B., Galbraith, D.W., Gozukirmizi, N., Tuberosa, R. and Bohnert, H.J. (2002). "Monitoring large-scale changes in transcript abundance in drought- and salt-stressed barley." Plant Mol. Biol. **48**(5 - 6): 551-573.

Pantuwan, G. (2000). Yield responses to rice (Oryza sative L.) genotypes to water deficit in rainfed lowlands. St. Lucia, Qld.179.

Parvanova, D., Ivanov, S., Konstantinova, T., Karanov, E., Atanassov, A., Tsvetkov, T., Alexieva, V. and Djilianov, D. (2004). "Transgenic tobacco plants accumulating osmolytes show reduced oxidative damage under freezing stress." Plant Physiol. Bioch. **42**(1): 57-63.

Penna, S. (2003). "Building stress tolerance through over-producing trehalose in transgenic plants." Trends Plant Sci. **8**(8): 355-357.

Penrose, D.M. and Glick, B.R. (2003). "Methods for isolating and characterizing ACC deaminase-containing plant growth-promoting rhizobacteria." Physiol. Plantarum **118**(1): 10-15.

Penuelas, J. and Llusia, J. (1999). "Short-term responses of terpene emission rates to experimental changes of PFD in Pinus halepensis and Quercus ilex in summer field conditions." Environmental and Experimental Botany **42**(1): 61-68.

Pfaffl, M.W. (2001). "A new mathematical model for relative quantification in real-time RT-PCR." Nucl. Acids Res. **29**(9): 2003-2007.

Phillips, J.R., Hilbricht, T., Salamini, F. and Bartels, D. (2002). "A novel abscisic acid- and dehydration-responsive gene family from the resurrection plant Craterostigma plantagineum encodes a plastid-targeted protein with DNA-binding activity." Planta **215**(2): 258-266.

Picon, C., Guehl, J.-M. and Ferhi, A. (1996). "Leaf gas exchange and carbon isotope composition respons to drought in a drought-avoiding (*Pinus pinaster*) and a drought-tolerant (*Quercus petraea*) species under present and elevated atmospheric CO_2 concentrations." Plant, Cell Environ. **19**: 182-190.

Pot, D., McMillan, L., Echt, C., Le Provost, G., Garnier-Gere, P., Cato, S. and Plomion, C. (2005). "Nucleotide variation in genes involved in wood formation in two pine species." New Phytol. **167**(1): 101-112.

Prandl, R., Hinderhofer, K., Eggers-Schumacher, G. and Schoffl, F. (1998). "HSF3, a new heat shock factor from Arabidopsis thaliana, derepresses the heat shock response and confers thermotolerance when overexpressed in transgenic plants." Mol. Gen. Genet. **258**(3): 269-278.

Qiang, L., Kasuga, M., Sakuma, Y., Abe, H., Miura, S., Yamaguchi-Shinozaki, K. and Shinozaki, K. (1998). "Two transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain separate two cellular signal transduction pathways in drought- and low-temperature-responsive gene expression, respectively, in Arabidopsis." Plant Cell **10**(8): 1391-1406.

Rabbani, M.A., Maruyama, K., Abe, H., *et al.* (2003). "Monitoring Expression Profiles of Rice Genes under Cold, Drought, and High-Salinity Stresses and Abscisic Acid Application Using cDNA Microarray and RNA Gel-Blot Analyses." Plant Physiol. **133**(4): 1755-1767.

Ramanjulu, S. and Bartels, D. (2002). "Drought- and desiccation-induced modulation of gene expression in plants." Plant, Cell Environ. **25**(2): 141-151.

Rech, P. (2002). Régulation transcriptionnelle du gène codant pour l'alcool cinnamyl déshydrogénase. Toulouse, Université Paul Sabatier (Toulouse III).

Rentsch, D., Hirner, B., Schmelzer, E. and Frommer, W.B. (1996). "Salt stress-induced proline transporters and salt stress-repressed broad specificity amino acid permeases identified by suppression of a yeast amino acid permease-targeting mutant." Plant Cell **8**(8): 1437-1446.

Rhoads, D.M., White, S.J., Zhou, Y., Muralidharan, M. and Elthon, T.E. (2005). "Altered gene expression in plants with constitutive expression of a mitochondrial small heat shock protein suggests the involvement of retrograde regulation in the heat stress response." Physiol. Plantarum **123**(4): 435-444.

Richards, D.E., King, K.E., Ait-ali, T. and Harberd, N.P. (2001). "How Gibberellin Regulates Plant Growth and Development : A Molecular Genetic Analysis of Gibberellin Signaling." Annu. Rev. Plant Phys. Plant Mol. Biol. **52**(1): 67-88.

Roberts, D.R., Dumbroff, E.B. and Thompson, J.E. (1986). "Exogenous polyamines alter membrane fluidity in bean leaves - a basis for their potential misinterpretation of their true physiological role." Planta **167**: 395-401.

Roberts, J.K., DeSimone, N.A., Lingle, W.L. and Dure III, L. (1993). "Cellular Concentrations and Uniformity of Cell-Type Accumulation of Two Lea Proteins in Cotton Embryos." Plant Cell **5**(7): 769-780.

Roberts, S.K. and Snowman, B.N. (2000). "The effects of ABA on channel-mediated K+ transport across higher plant roots." J. Exp. Bot. **51**(350): 1585-1594.

Robison, M.M., Shah, S., Tamot, B., Pauls, K.P., Moffatt, B.A. and Glick, B.R. (2001). "Reduced symptoms of Verticillium wilt in transgenic tomato expressing a bacterial ACC deaminase." Mol. Plant Pathol. **2**(3): 135-145.

Rodrigo, M.J., Bockel, C., Blervacq, A.-S. and Bartels, D. (2004). "The novel gene CpEdi-9 from the resurrection plant C. plantagineum encodes a hydrophilic protein and is expressed in mature seeds as well as in response to dehydration in leaf phloem tissues." Planta **219**(4): 579-589.

Romualdi, C., Bortoluzzi, S., d'Alessi, F. and Danieli, G.A. (2003). "IDEG6: a web tool for detection of differentially expressed genes in multiple tag sampling experiments." Physiol. Genomics **12**(2): 159-162.

Rudd, S. (2003). "Expressed sequence tags: alternative or complement to whole genome sequences?" Trends Plant Sci. **8**(7): 321-329.

Ryser, U. (2003). "Protoxylem: the deposition of a network containing glycine-rich cell wall proteins starts in the cell corners in close association with the pectins of the middle lamella." Planta **216**(5): 854-864.

Sabirzhanova, I.B., Sabirzhanov, B.E., Chemeris, A.V., Veselov, D.S. and Kudoyarova, G.R. (2005). "Fast changes in expression of expansin gene and leaf extensibility in osmotically stressed maize plants." Plant Physiol. Bioch. **43**(4): 419-422.

Saijo, Y., Kinoshita, N., Ishiyama, K.*, et al.* (2001). "A Ca2+-dependent protein kinase that endows rice plants with cold- and salt-stress tolerance functions in vascular bundles." Plant Cell Physiol. **42**(11): 1228-1233.

Sakuma, Y., Liu, Q., Dubouzet, J.G., Abe, H., Shinozaki, K. and Yamaguchi-Shinozaki, K. (2002). "DNA-binding specificity of the ERF/AP2 domain of Arabidopsis DREBs, transcription factors involved in dehydration- and cold-inducible gene expression." Biochem. Bioph. Res. Co. **290**(3): 998-1009.

Sarda, X., Tousch, D., Ferrare, K., Cellier, F., Alcon, C., Dupuis, J.M., Casse, F. and Lamaze, T. (1999). "Characterization of closely related delta -TIP genes encoding aquaporins which are differentially expressed in sunflower roots upon water deprivation through exposure to air." Plant Mol. Biol. **40**(1): 179-191.

Sarrauste, N. (1982). Photosynthèse, respiration et répartition de matière sèche de jeunes plants de pin maritime (*Pinus pinaster Ait.*) appartenant à sept provenances et conduits selon deux traitements hydriques, Paris VII, DEA.

Sathyan, P. (2004). Identification of Drought Responsive Genes in Aleppo Pine (Pinus halepensis) and Loblolly Pine (Pinus taeda.L). Texas, Texas A&M University, 74.

Sauter, A., Davies, W.J. and Hartung, W. (2001). "The long-distance abscisic acid signal in the droughted plant: the fate of the hormone on its way from root to shoot." J. Exp. Bot. **52**(363): 1991-1997.

Savoure, A., Jaoua, S., Hua, X., Ardiles, W., Montagu, M.v. and Verbruggen, N. (1995). "Isolation, characterization, and chromosomal location of a gene encoding the DELTA 1-pyrroline-5-carboxylate synthetase in Arabidopsis thaliana." FEBS Lett. **372**(1): 13-19.

Schar, C. and Jendritzky, G. (2004). "Climate change: Hot news from summer 2003." Nature **432**(7017): 559-560.

Schar, C., Vidale, P.L., Luthi, D., Frei, C., Haberli, C., Liniger, M.A. and Appenzeller, C. (2004). "The role of increasing temperature variability in European summer heatwaves." Nature **427**(6972): 332-336.

Schroeder, J.I., Kwak, J.M. and Allen, G.J. (2001). "Guard cell abscisic acid signalling and engineering drought hardiness in plants." Nature **410**(6826): 327-330.

Scippa, G.S., Griffiths, A., Chiatante, D. and Bray, E.A. (2000). "The H1 histone variant of tomato, H1-S, is targeted to the nucleus and accumulates in chromatin in response to water-deficit stress." Planta **211**(2): 173-181.

Scott, P. (2000). "Resurrection Plants and the Secrets of Eternal Leaf." Ann. Bot. **85**(2): 159-166.

Segarra, C.I., Casalongue, C.A., Pinedo, M.L., Ronchi, V.P. and Conde, R.D. (2003). "A germin-like protein of wheat leaf apoplast inhibits serine proteases." J. Exp. Bot. **54**(386): 1335-1341.

Segura, A., Moreno, M., Madueno, F., Molina, A. and Garcia-Olmedo, F. (1999). "Snakin-1, a Peptide from Potato That Is Active Against Plant Pathogens." MPMI **12**(1): 16-23.

Seki, M., Narusaka, M., Ishida, J.*, et al.* (2002). "Monitoring the expression profiles of 7000 Arabidopsis genes under drought, cold and high-salinity stresses using a full-length cDNA microarray." Plant J. **31**(3): 279-292.

Sharan, R., Maron-Katz, A. and Shamir, R. (2003). "CLICK and EXPANDER: a system for clustering and visualizing gene expression data." Bioinformatics **19**(14): 1787-1799.

Sharp, R.E. and LeNoble, M.E. (2002). "ABA, ethylene and the control of shoot and root growth under water stress." J. Exp. Bot. **53**(366): 33-37.

Sheveleva, E., Chmara, W., Bohnert, H.J. and Jensen, R.G. (1997). "Increased salt and drought tolerance by D-ononitol production in transgenic Nicotiana tabacum L." Plant Physiol. **115**(3): 1211-1219.

Sheveleva, E.V., Marquez, S., Chmara, W., Zegeer, A., Jensen, R.G. and Bohnert, H.J. (1998). "Sorbitol-6-Phosphate Dehydrogenase Expression in Transgenic Tobacco . High Amounts of Sorbitol Lead to Necrotic Lesions." Plant Physiol. **117**(3): 831-839.

Shi, H., Ishitani, M., Kim, C. and Zhu, J. (2000). "The Arabidopsis thaliana salt tolerance gene SOS1 encodes a putative Na+/H+ antiporter." PNAS **97**(12): 6896-6901.

Shi, H., Lee, B.-h., Wu, S.-J. and Zhu, J.-K. (2003). "Overexpression of a plasma membrane Na+/H+ antiporter gene improves salt tolerance in Arabidopsis thaliana." Nat. Biotechnol. **21**(1): 81-85.

Shigeoka, S., Ishikawa, T., Tamoi, M., Miyagawa, Y., Takeda, T., Yabuta, Y. and Yoshimura, K. (2002). "Regulation and function of ascorbate peroxidase isoenzymes." J. Exp. Bot. **53**(372): 1305-1319.

Showalter, A.M. (1993). "Structure and Function of Plant Cell Wall Proteins." Plant Cell **5**: 9-13.

Siefritz, F., Tyree, M.T., Lovisolo, C., Schubert, A. and Kaldenhoff, R. (2002). "PIP1 plasma membrane aquaporins in tobacco: from cellular effects to function in plants." Plant Cell **14**(4): 869-876.

Simon, V., Clement, B., Riba, M.-L. and Torres, L. (1994). "The Landes experiment: Monoterpenes emitted from the maritime pine." J. Geophys. Res. **99**(D8): 16501-16510.

Smart, L.B., Moskal, W.A., Cameron, K.D. and Bennett, A.B. (2001). "MIP genes are down-regulated under drought stress in Nicotiana glauca." Plant Cell Physiol. **42**(7): 686-693.

Smirnoff, N. and Bryant, J.A. (1999). "DREB takes the stress out of growing up." Nat. Biotechnol. **17**(3): 229-30.

Snedden, W.A. and Fromm, H. (1997). "Characterization of the plant homologue of prohibitin, a gene associated with antiproliferative activity in mammalian cells." Plant Mol. Biol. **33**(4): 753-756.

Speed, T., Ed. (2003). Statistical analysis of gene expression microarray data, Chapman & Hall / CRC.

Stasolla, C., Zyl, L.v., Egertsdotter, U., Craig, D., Liu, W.B. and Sederoff, R.R. (2003). "The effects of polyethylene glycol on gene expression of developing white spruce somatic embryos." Plant Physiol. **131**(1): 49-60.

Stearns, J.C., Shah, S., Greenberg, B.M., Dixon, D.G. and Glick, B.R. (2005). "Tolerance of transgenic canola expressing 1-aminocyclopropane-1-carboxylic acid deaminase to growth inhibition by nickel." Plant Physiol. Bioch. **43**(7): 701-708.

Steudle, E. (2000). "Water uptake by roots: effects of water deficit." J. Exp. Bot. **51**(350): 1531-1542.

Steudle, E. (2001). "The Cohesion-tension Mechanism and the Acquisition of Water by Plant Roots." Annu. Rev. Plant Phys. Plant Mol. Biol. **52**(1): 847-875.

Strauss, S.H., Brunner, A.M., Busov, V.B., Ma, C. and Meilan, R. (2004). "Ten lessons from 15 years of transgenic Populus research." Forestry **77**(5): 455-465.

Streeter, J.G., Lohnes, D.G. and Fioritto, R.J. (2001). "Patterns of pinitol accumulation in soybean plants and relationships to drought tolerance." Plant, Cell Environ. **24**(4): 429-438.

Taji, T., Ohsumi, C., Iuchi, S., Seki, M., Kasuga, M., Kobayashi, M., Yamaguchi-Shinozaki, K. and Shinozaki, K. (2002). "Important roles of drought- and cold-inducible genes for galactinol synthase in stress tolerance in Arabidopsis thaliana." Plant J. **29**(4): 417-426.

Tang, W., Peng, X. and Newton, R.J. (2005). "Enhanced tolerance to salt stress in transgenic loblolly pine simultaneously expressing two genes encoding mannitol-1-phosphate dehydrogenase and glucitol-6-phosphate dehydrogenase." Plant Physiol. Bioch. **43**(2): 139-146.

Taylor, I.B., Burbidge, A. and Thompson, A.J. (2000). "Control of abscisic acid synthesis." J. Exp. Bot. **51**(350): 1563-1574.

"A. thaliana genome." (2000). Nature **408**(6814): 791.

Tiberi, R., Niccoli, A., Curini, M., Epifano, F., Marcotullio, M.C. and Rosati, O. (1999). "The Role of the Monoterpene Composition in Pinus spp. Needles, in Host Selection by the Pine Processionary Caterpillar, Thaumetopoea pityocampa." Phytoparasitica **27**(4): 263-272.

Tiburcio, A.F., Besford, R.T., Capell, T., Borell, A., Testillano, P.S. and Risueno, M.C. (1994). "Mechanism of polyamine action during senescence responses induced by osmotic stress." J. Exp. Bot. **45**: 1789-1800.

Tognetti, R., Michelozzi, M., Lauteri, M., Brugnoli, E. and Giannini, R. (2000). "Geographic variation in growth, carbon isotope discrimination, and monoterpene composition in Pinus pinaster Ait. provenances." Can. J. Forest. Res. **30**(11): 1682-1690.

Triboulot, M. (1996). Effets du stress osmotique et de la carence en potassium sur la croissance racinaire du pin maritime. Etude biomécanique du grandissement cellulaire. Nancy, Université Henri Poincaré - Nancy I, Thèse de Doctorat, 85.

Truernit, E., Schmid, J., Epple, P., Illig, J. and Sauer, N. (1996). "The sink-specific and stress-regulated Arabidopsis STP4 gene: enhanced expression of a gene encoding a monosaccharide transporter by wounding, elicitors, and pathogen challenge." Plant Cell **8**(12): 2169-2182.

Turtola, S., Manninen, A.-M., Rikala, R. and Kainulainen, P. (2003). "Drought Stress Alters the Concentration of Wood Terpenoids in Scots Pine and Norway Spruce Seedlings." J. Chem. Ecol. **29**(9): 1981-1995.

Tyerman, S.D., Niemietz, C.M. and Bramley, H. (2002). "Plant aquaporins: multifunctional water and solute channels with expanding roles." Plant, Cell Environ. **25**(2): 173-194.

Ueda, A., Shi, W., Sanmiya, K., Shono, M. and Takabe, T. (2001). "Functional analysis of salt-inducible proline transporter of barley roots." Plant Cell Physiol. **42**(11): 1282-1289.

Ueda, M., Koshino-Kimura, Y. and Okada, K. (2005). "Stepwise understanding of root development." Curr. Opin. Plant Biol. **8**(1): 71-76.

Urao, T., Katagiri, T., Mizoguchi, T., Yamaguchi-Shinozaki, K., Hayashida, N. and Shinozaki, K. (1994). "An Arabidopsis thaliana cDNA encoding Ca2+-dependent protein kinase." Plant Physiol. **105**(4): 1461-1462.

Urao, T., Yakubov, B., Satoh, R., Yamaguchi-Shinozaki, K., Seki, M., Hirayama, T. and Shinozaki, K. (1999). "A transmembrane hybrid-type histidine kinase in Arabidopsis functions as an osmosensor." Plant Cell **11**(9): 1743-1754.

Urao, T., Yamaguchi-Shinozaki, K., Urao, S. and Shinozaki, K. (1993). "An Arabidopsis myb homolog is induced by dehydration stress and its gene product binds to the conserved MYB recognition sequence." Plant Cell **5**(11): 1529-1539.

van der Weele, C.M., Spollen, W.G., Sharp, R.E. and Baskin, T.I. (2000). "Growth of Arabidopsis thaliana seedlings under water deficit studied by control of water potential in nutrient-agar media." J. Exp. Bot. **51**(350): 1555-1562.

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A. and Speleman, F. (2002). "Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes." Genome Biol. **3**(7): 1-12.

Verbruggen, N., Hua, X., May, M. and Montagu, M.v. (1996). "Environmental and developmental signals modulate proline homeostasis: evidence for a negative transcriptional regulator." PNAS **93**(16): 8787-8791.

Vicient, C., Roscoe, T. and Delseny, M. (1998). "Gene note. Characterization of an Em-like gene of Brassica napus." J. Exp. Bot. **49**(323): 1061-1062.

Vincent, D., Lapierre, C., Pollet, B., Cornic, G., Negroni, L. and Zivy, M. (2005). "Water Deficits Affect Caffeate O-Methyltransferase, Lignification, and Related Enzymes in Maize Leaves. A Proteomic Investigation." Plant Physiol. **137**(3): 949-960.

Wakamiya, I., Newton, R.J., Johnston, J.S. and Price, H.J. (1993). "Genome size and environmental factors in the genus Pinus." Am. J. Bot. **80**(11): 1235-1241.

Wakamiya, I., Price, H.J., Messina, M.G. and Newton, R.J. (1996). "Pine genome size diversity and water relations." Physiol. Plantarum **96**(1): 13-20.

Walden, A.R., Walter, C. and Gardner, R.C. (1999). "Genes Expressed in Pinus radiata Male Cones Include Homologs to Anther-Specific and Pathogenesis Response Genes." Plant Physiol. **121**(4): 1103-1116.

Walden, R., Cordeiro, A. and Tiburcio, A.F. (1997). "Polyamines: Small Molecules Triggering Pathways in Plant Growth and Development." Plant Physiol. **113**(4): 1009-1013.

Wang, W., Vinocur, B. and Altman, A. (2003). "Plant responses to drought, salinity and extreme temperatures: towards genetic engineering for stress tolerance." Planta **218**(1): 1-14.

Wang, W., Vinocur, B., Shoseyov, O. and Altman, A. (2004). "Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response." Trends Plant Sci. **9**(5): 244-252.

Wang, X.-Q., Tank, D.C. and Sang, T. (2000). "Phylogeny and Divergence Times in Pinaceae: Evidence from Three Genomes." Mol. Biol. Evol. **17**(5): 773-781.

Warren, C.R. and Adams, M.A. (2000). "Water availability and branch length determine Δ13C in foliage of Pinus pinaster." Tree Physiol. **20**: 637-643.

Watanabe, S., Hamano, M., Kakeshita, H., Bunai, K., Tojo, S., Yamaguchi, H., Fujita, Y., Wong, S.-L. and Yamane, K. (2003). "Mannitol-1-Phosphate Dehydrogenase (MtlD) Is Required for Mannitol and Glucitol Assimilation in Bacillus subtilis: Possible Cooperation of mtl and gut Operons." J. Bacteriol. **185**(16): 4816-4824.

Weber, A., Servaites, J.C., Geiger, D.R., Kofler, H., Hille, D., Groner, F., Hebbeker, U. and Flugge, U.I. (2000). "Identification, purification, and molecular cloning of a putative plastidic glucose translocator." Plant Cell **12**(5): 787-801.

Wilkinson, S. and Davies, W.J. (2002). "ABA-based chemical signalling: the co-ordination of responses to stress in plants." Plant, Cell Environ. **25**(2): 195-210.

Wise, M. (2003). "LEAping to conclusions: A computational reanalysis of late embryogenesis abundant proteins and their possible roles." BMC Bioinformatics **4**(1): 52.

Wu, H. (2003). MAANOVA: a software package for the analysis of spotted cDNA microarray experiments. The analysis of gene expression data: methods and software, Springer, NY**:** 455.

Wu, Y., Jeong, B.-R., Fry, S.C. and Boyer, J.S. (2005). "Change in XET activities, cell wall extensibility and hypocotyl elongation of soybean seedlings at low water potential." Planta **220**(4): 593-601.

Wu, Y., Thorne, E.T., Sharp, R.E. and Cosgrove, D.J. (2001). "Modification of Expansin Transcript Levels in the Maize Primary Root at Low Water Potentials." Plant Physiol. **126**(4): 1471-1479.

Wullschleger, S.D., Meinzer, F.C. and Vertessy, R.A. (1998). "A review of whole-plant water use studies in trees." Tree Physiol. **18**(8/9).

Xiong, L., Schumaker, K.S. and Zhu, J.-K. (2002). "Cell Signaling during Cold, Drought, and Salt Stress." Plant Cell **14**(90001): S165-183.

Yamaguchi-Shinozaki, K. and Shinozaki, K. (1994). "A novel cis-acting element in an Arabidopsis gene is involved in responsiveness to drought, low-temperature, or high-salt stress." Plant Cell **6**(2): 251-264.

Yang, S.-H., Wang, H., Sathyan, P., Stasolla, C. and Loopstra, C.A. (2005). "Real-time RT-PCR analysis of loblolly pine (Pinus taeda) arabinogalactan-protein and arabinogalactan-protein-like genes." Physiol. Plantarum **124**(1): 91-106.

Yang, Y., Qi, M. and Mei, C. (2004). "Endogenous salicylic acid protects rice plants from oxidative damage caused by aging as well as biotic and abiotic stress." Plant J. **40**(6): 909-919.

Yang, Y.H. and Speed, T. (2002). "Design issues for cDNA microarray experiments." Nat. Rev. Genet. **3**(8): 579-588.

Yokoyama, R. and Nishitani, K. (2001). "A Comprehensive Expression Analysis of all Members of a Gene Family Encoding Cell-Wall Enzymes Allowed us to Predict cis-Regulatory Regions Involved in Cell-Wall Construction in Specific Organs of Arabidopsis." Plant Cell Physiol. **42**(10): 1025-1033.

Yoshiba, Y., Kiyosue, T., Katagiri, T., Ueda, H., Mizoguchi, T., Yamaguchi-Shinozaki, K., Wada, K., Harada, Y. and Shinozaki, K. (1995). "Correlation between the induction of a gene for DELTA 1-pyrroline-5-carboxylate synthetase and the accumulation of proline in Arabidopsis thaliana under osmotic stress." Plant J. **7**(5): 751-760.

Yu, J., Hu, S., Wang, J.*, et al.* (2002). "A Draft Sequence of the Rice Genome (Oryza sativa L. ssp. indica)." Science **296**(5565): 79-92.

Annexes

Tableau Ax-1 : Résultat du northern *in silico* sur les données de l'assemblage sans regroupements. Chaque contig est considéré comme un gène différent (CT du projet « réponse au stress »). 279 gènes différentiellement exprimés, statistiquement significatifs pour au moins un des tests statistiques avec un risque α de 0.05. (Lib1 = banque RN (root control) ; Lib2 = banque RS (root stress))

Contig	Description	Lib1	Lib2	Lib1(norm)	Lib2(norm)	AC 1 2	Fisher 1 2	Chi2x2 1 2	R	Chi	Ratio d'expressions normalisées
CT469	Inositol 1-alpha-galactosyltransferase (EC 2.4.1.123)	9	0	25,7	0	0,002529	0,00533	0,006607	0,001034	0,006607	25,80
CT124	Phosphate carrier protein	8	0	22,9	0	0,004598	0,010007	0,010452	0,001978	0,010452	22,90
CT406	Ubiquitin / 40S ribosomal protein S27a	8	0	22,9	0	0,004598	0,010007	0,010452	0,001978	0,010452	22,90
CT131	Protein Farnesyl-diphosphate farnesyltransferase alpha (EC 2.5.1.21)	7	0	20	0	0,008362	0,018892	0,016623	0,003806	0,016623	20,00
CT456	Phytocyanin	7	0	20	0	0,008362	0,018892	0,016623	0,003806	0,016623	20,00
CT631	40S ribosomal protein S16	7	0	20	0	0,008362	0,018892	0,016623	0,003806	0,016623	20,00
CT635	At2g15910 protein	7	0	20	0	0,008362	0,018892	0,016623	0,003806	0,016623	20,00
CT693	Beta-galactosidase precursor (EC 3.2.1.23)	7	0	20	0	0,008362	0,018892	0,016623	0,003806	0,016623	20,00
CT1008	putative protein	6	0	17,2	0	0,015206	0,035889	0,026615	0,00738	0,026615	17,20
CT1066	Adenine phosphoribosyltransferase (EC 2.4.2.7)	6	0	17,2	0	0,015206	0,035889	0,026615	0,00738	0,026615	17,20
CT168	Peroxidase precursor (EC 1.11.1.7)	6	0	17,2	0	0,015206	0,035889	0,026615	0,00738	0,026615	17,20
CT170	Mitochondrial import receptor	6	0	17,2	0	0,015206	0,035889	0,026615	0,00738	0,026615	17,20
CT3	Enoyl-[acyl-carrier protein] reductase [NADH] (EC 1.3.1.9)	6	0	17,2	0	0,015206	0,035889	0,026615	0,00738	0,026615	17,20
CT388	Germin-like protein	6	0	17,2	0	0,015206	0,035889	0,026615	0,00738	0,026615	17,20
CT494	Minor extracellular protease vpr precursor	6	0	17,2	0	0,015206	0,035889	0,026615	0,00738	0,026615	17,20
CT538	Membrane-associated salt-inducible-like protein	6	0	17,2	0	0,015206	0,035889	0,026615	0,00738	0,026615	17,20
CT700	Plasma membrane intrinsic protein	6	0	17,2	0	0,015206	0,035889	0,026615	0,00738	0,026615	17,20
CT714	putative protein	6	0	17,2	0	0,015206	0,035889	0,026615	0,00738	0,026615	17,20
CT786	40S ribosomal protein S14	6	0	17,2	0	0,015206	0,035889	0,026615	0,00738	0,026615	17,20
CT881	Cytochrome P450 (EC 1.14.-.-)	6	0	17,2	0	0,015206	0,035889	0,026615	0,00738	0,026615	17,20
CT1234	H(+)-transporting two-sector ATPase epsilon chain, mitochondrial precursor (EC 3.6.3.14)	5	0	14,3	0	0,027652	0,068657	0,042993	0,014452	0,042993	14,30
CT1237	Polygalacturonase isoenzyme 1 beta subunit precursor	5	0	14,3	0	0,027652	0,068657	0,042993	0,014452	0,042993	14,30
CT1353	putative protein	5	0	14,3	0	0,027652	0,068657	0,042993	0,014452	0,042993	14,30
CT149	Acid phosphatase precursor 1 (EC 3.1.3.2)	5	0	14,3	0	0,027652	0,068657	0,042993	0,014452	0,042993	14,30
CT413	Peroxiredoxin, mitochondrial precursor	5	0	14,3	0	0,027652	0,068657	0,042993	0,014452	0,042993	14,30
CT477	Ethylene responsive element binding factor	5	0	14,3	0	0,027652	0,068657	0,042993	0,014452	0,042993	14,30
CT513	Methyl binding domain protein	5	0	14,3	0	0,027652	0,068657	0,042993	0,014452	0,042993	14,30
CT517	5S ribosomal RNA	5	0	14,3	0	0,027652	0,068657	0,042993	0,014452	0,042993	14,30
CT56	Polcalcin	5	0	14,3	0	0,027652	0,068657	0,042993	0,014452	0,042993	14,30
CT618	26S proteasome non-ATPase regulatory subunit	5	0	14,3	0	0,027652	0,068657	0,042993	0,014452	0,042993	14,30
CT645	Auxin-induced protein	5	0	14,3	0	0,027652	0,068657	0,042993	0,014452	0,042993	14,30
CT679	Lipase	5	0	14,3	0	0,027652	0,068657	0,042993	0,014452	0,042993	14,30
CT738	putative protein	5	0	14,3	0	0,027652	0,068657	0,042993	0,014452	0,042993	14,30
CT761	Phenylalanine ammonia-lyase (EC 4.3.1.5)	5	0	14,3	0	0,027652	0,068657	0,042993	0,014452	0,042993	14,30
CT85	putative protein	5	0	14,3	0	0,027652	0,068657	0,042993	0,014452	0,042993	14,30
CT860	putative protein	5	0	14,3	0	0,027652	0,068657	0,042993	0,014452	0,042993	14,30
CT923	Aspartate aminotransferase (EC 2.6.1.1)	5	0	14,3	0	0,027652	0,068657	0,042993	0,014452	0,042993	14,30
CT946	putative protein	5	0	14,3	0	0,027652	0,068657	0,042993	0,014452	0,042993	14,30
CT1	putative protein	4	0	11,5	0	0,050285	0,132366	0,070298	0,028691	0,070298	11,50
CT102	putative protein	4	0	11,5	0	0,050285	0,132366	0,070298	0,028691	0,070298	11,50
CT1039	Flavonoid 3',5'-hydroxylase (EC 1.14.-.-)	4	0	11,5	0	0,050285	0,132366	0,070298	0,028691	0,070298	11,50
CT1076	Unknown protein	4	0	11,5	0	0,050285	0,132366	0,070298	0,028691	0,070298	11,50
CT1141	putative protein	4	0	11,5	0	0,050285	0,132366	0,070298	0,028691	0,070298	11,50
CT1176	Peroxidase precursor (EC 1.11.1.7)	4	0	11,5	0	0,050285	0,132366	0,070298	0,028691	0,070298	11,50
CT1214	putative protein	4	0	11,5	0	0,050285	0,132366	0,070298	0,028691	0,070298	11,50
CT127	putative protein	4	0	11,5	0	0,050285	0,132366	0,070298	0,028691	0,070298	11,50
CT1271	putative protein	4	0	11,5	0	0,050285	0,132366	0,070298	0,028691	0,070298	11,50
CT1290	putative protein	4	0	11,5	0	0,050285	0,132366	0,070298	0,028691	0,070298	11,50
CT177	putative protein	4	0	11,5	0	0,050285	0,132366	0,070298	0,028691	0,070298	11,50
CT192	Adenosine kinase (EC 2.7.1.20)	4	0	11,5	0	0,050285	0,132366	0,070298	0,028691	0,070298	11,50
CT193	Syntaxin	4	0	11,5	0	0,050285	0,132366	0,070298	0,028691	0,070298	11,50
CT235	Peroxidase precursor (EC 1.11.1.7)	4	0	11,5	0	0,050285	0,132366	0,070298	0,028691	0,070298	11,50
CT24	putative protein	4	0	11,5	0	0,050285	0,132366	0,070298	0,028691	0,070298	11,50
CT277	putative protein	4	0	11,5	0	0,050285	0,132366	0,070298	0,028691	0,070298	11,50
CT296	putative protein	4	0	11,5	0	0,050285	0,132366	0,070298	0,028691	0,070298	11,50
CT299	putative protein	4	0	11,5	0	0,050285	0,132366	0,070298	0,028691	0,070298	11,50
CT309	Histone H4	4	0	11,5	0	0,050285	0,132366	0,070298	0,028691	0,070298	11,50
CT330	putative protein	4	0	11,5	0	0,050285	0,132366	0,070298	0,028691	0,070298	11,50
CT334	Germin-like protein subfamily 1 member 7 precursor	4	0	11,5	0	0,050285	0,132366	0,070298	0,028691	0,070298	11,50
CT364	Serine-rich protein	4	0	11,5	0	0,050285	0,132366	0,070298	0,028691	0,070298	11,50
CT37	Poly(rC)-binding protein	4	0	11,5	0	0,050285	0,132366	0,070298	0,028691	0,070298	11,50
CT383	putative protein	4	0	11,5	0	0,050285	0,132366	0,070298	0,028691	0,070298	11,50
CT408	putative protein	4	0	11,5	0	0,050285	0,132366	0,070298	0,028691	0,070298	11,50
CT464	Microsomal signal peptidase 22 kDa subunit	4	0	11,5	0	0,050285	0,132366	0,070298	0,028691	0,070298	11,50
CT514	NifU-like protein	4	0	11,5	0	0,050285	0,132366	0,070298	0,028691	0,070298	11,50
CT515	putative protein	4	0	11,5	0	0,050285	0,132366	0,070298	0,028691	0,070298	11,50
CT529	L-ascorbate peroxidase (EC 1.11.1.11)	4	0	11,5	0	0,050285	0,132366	0,070298	0,028691	0,070298	11,50
CT539	Unknown protein	4	0	11,5	0	0,050285	0,132366	0,070298	0,028691	0,070298	11,50
CT555	putative protein	4	0	11,5	0	0,050285	0,132366	0,070298	0,028691	0,070298	11,50
CT557	40S ribosomal protein S3a	4	0	11,5	0	0,050285	0,132366	0,070298	0,028691	0,070298	11,50
CT577	Unknown protein	4	0	11,5	0	0,050285	0,132366	0,070298	0,028691	0,070298	11,50
CT602	RNA-binding protein with multiple splicing homolog	4	0	11,5	0	0,050285	0,132366	0,070298	0,028691	0,070298	11,50
CT692	putative protein	4	0	11,5	0	0,050285	0,132366	0,070298	0,028691	0,070298	11,50
CT735	Zinc-binding protein	4	0	11,5	0	0,050285	0,132366	0,070298	0,028691	0,070298	11,50
CT741	putative protein	4	0	11,5	0	0,050285	0,132366	0,070298	0,028691	0,070298	11,50
CT817	Protein-methionine-S-oxide reductase (EC 1.8.4.6)	4	0	11,5	0	0,050285	0,132366	0,070298	0,028691	0,070298	11,50
CT829	Adenylate kinase (EC 2.7.4.3)	4	0	11,5	0	0,050285	0,132366	0,070298	0,028691	0,070298	11,50
CT847	putative protein	4	0	11,5	0	0,050285	0,132366	0,070298	0,028691	0,070298	11,50
CT855	Hypothetical 28.8 kDa protein	4	0	11,5	0	0,050285	0,132366	0,070298	0,028691	0,070298	11,50
CT859	Alpha-expansin precursor	4	0	11,5	0	0,050285	0,132366	0,070298	0,028691	0,070298	11,50
CT886	5-methyltetrahydropteroyltriglutamate--homocysteine S-methyltransferase (EC 2.1.1.14)	4	0	11,5	0	0,050285	0,132366	0,070298	0,028691	0,070298	11,50
CT891	Lysozyme precursor (EC 3.2.1.17) (Peptidoglycan hydrolase) (Autolysin) (Lysosyme)	4	0	11,5	0	0,050285	0,132366	0,070298	0,028691	0,070298	11,50
CT932	P-protein [Includes: Chorismate mutase (EC 5.4.99.	4	0	11,5	0	0,050285	0,132366	0,070298	0,028691	0,070298	11,50
CT953	putative protein	4	0	11,5	0	0,050285	0,132366	0,070298	0,028691	0,070298	11,50
CT219	Ethylene responsive element binding factor	9	1	25,7	3,5	0,011381	0,027628	0,025943	0,015462	0,025943	7,37
CT259	40S ribosomal protein S15	7	1	20	3,5	0,030109	0,081087	0,064399	0,047028	0,064399	5,71
CT145	60S ribosomal protein L10	6	1	17,2	3,5	0,047909	0,137871	0,102089	0,081483	0,102089	4,91
CT31	Protein-methionine-S-oxide reductase (EC 1.8.4.6)	6	1	17,2	3,5	0,047909	0,137871	0,102089	0,081483	0,102089	4,91
CT558	putative protein	6	1	17,2	3,5	0,047909	0,137871	0,102089	0,081483	0,102089	4,91
CT653	60S ribosomal protein L32	6	1	17,2	3,5	0,047909	0,137871	0,102089	0,081483	0,102089	4,91
CT667	Pinus taeda putative cell wall protein	6	1	17,2	3,5	0,047909	0,137871	0,102089	0,081483	0,102089	4,91
CT682	Osmotin-like protein precursor	6	1	17,2	3,5	0,047909	0,137871	0,102089	0,081483	0,102089	4,91
CT223	40S ribosomal protein S7	10	2	28,6	6,1	0,018592	0,077435	0,048227	0,037142	0,048227	4,69
CT166	Defensin	17	3	48,7	10,5	0,002198	0,00638	0,006895	0,004148	0,006895	4,64
CT389	Germin-like protein	8	2	22,9	6,1	0,041919	0,201202	0,111626	0,096813	0,111626	3,75
CT143	putative protein	16	4	45,8	13,1	0,007646	0,025059	0,024342	0,018902	0,024342	3,50
CT88	putative protein	12	3	34,4	10,5	0,017445	0,06755	0,051266	0,042032	0,051266	3,28
CT276	Late embryogenesis abundant protein	21	7	60,1	24,5	0,008567	0,036261	0,03293	0,028365	0,03293	2,45
CT431	Metallothionein-like protein EMB30	15	5	42,9	17,5	0,020025	0,076258	0,071621	0,063874	0,071621	2,45

Contig	Description	Lib1	Lib2	Lib1(norm)	Lib2(norm)	AC 1 2	Fisher 1 2	Chi2x2 1 2	R	Chi	Ratio d'expressions normalisées
CT108	Phosphopyruvate hydratase (EC 4.2.1.11)	43	15	123,1	52,5	0,000701	0,003285	0,003234	0,002591	0,003234	2,34
CT93	Naringenin-chalcone synthase (EC 2.3.1.74)	29	13	83	45,5	0,012837	0,085894	0,066176	0,062422	0,066176	1,82
CT20	Ubiquitin	20	10	57,3	34,1	0,036046	0,269515	0,197565	0,192359	0,197565	1,68
CT73	25S ribosomal subunit (AJ271027)	96	61	274,8	213,4	0,010638	0,123109	0,116429	0,118707	0,116429	1,29
CT155	Elongation factor 1-alpha	15	19	42,9	66,5	0,03357	0,227957	0,201328	0,203476	0,201328	1,55
CT224	putative protein	7	12	20	41,1	0,029126	0,164223	0,111279	0,111906	0,111279	2,06
CT286	Avr9/Cf-9 rapidly elicited protein 146	11	19	31,5	66,5	0,010805	0,06403	0,043162	0,043754	0,043162	2,11
CT390	Nucleoside diphosphate kinase I (EC 2.7.4.6)	5	9	14,3	31,5	0,041969	0,181668	0,146684	0,146889	0,146684	2,20
CT72	Polcalcin	7	14	20	48,1	0,013616	0,050176	0,045697	0,045639	0,045697	2,41
CT126	Tubulin alpha	3	7	8,6	24,5	0,041064	0,12507	0,111897	0,110229	0,111897	2,85
CT686	60S ribosomal protein L14	3	7	8,6	24,5	0,041064	0,12507	0,111897	0,110229	0,111897	2,85
CT717	Unknown protein	2	6	5,7	20,1	0,038712	0,151405	0,087987	0,084391	0,087987	3,53
CT86	putative protein	2	6	5,7	20,1	0,038712	0,151405	0,087987	0,084391	0,087987	3,53
CT70	putative protein	3	9	8,6	31,5	0,015251	0,043722	0,036599	0,034589	0,036599	3,66
CT67	putative protein	2	7	5,7	24,5	0,022402	0,088083	0,048	0,0445	0,048	4,30
CT994	Histone H1	2	7	5,7	24,5	0,022402	0,088083	0,048	0,0445	0,048	4,30
CT99	putative protein	3	11	8,6	38,5	0,005112	0,014235	0,011506	0,010161	0,011506	4,48
CT68	putative protein	2	8	5,7	27,4	0,012604	0,05048	0,026024	0,023103	0,026024	4,75
CT707	Ubiquitin	2	8	5,7	27,4	0,012604	0,05048	0,026024	0,023103	0,026024	4,75
CT997	Hypothetical 8.0 kDa protein	2	8	5,7	27,4	0,012604	0,05048	0,026024	0,023103	0,026024	4,75
CT531	Late embryogenesis abundant protein	3	12	8,6	41,1	0,002876	0,00796	0,006388	0,005404	0,006388	4,78
CT1105	Inosine-5'-monophosphate dehydrogenase (EC 1.1.1.2)	1	5	2,9	17,5	0,033516	0,096767	0,059054	0,051985	0,059054	6,03
CT1222	40S ribosomal protein S26 (S31)	1	5	2,9	17,5	0,033516	0,096767	0,059054	0,051985	0,059054	6,03
CT1315	Major pollen allergen Bet v 1-D/H / Intracellular pathogenesis-related protein	1	5	2,9	17,5	0,033516	0,096767	0,059054	0,051985	0,059054	6,03
CT699	Annexin	1	5	2,9	17,5	0,033516	0,096767	0,059054	0,051985	0,059054	6,03
CT934	Zinc finger protein 216	1	5	2,9	17,5	0,033516	0,096767	0,059054	0,051985	0,059054	6,03
CT71	Aucune Annotation	2	11	5,7	38,5	0,001992	0,004532	0,004062	0,003019	0,004062	6,75
CT668	Pinus taeda putative cell wall protein	1	6	2,9	20,1	0,017599	0,050834	0,030319	0,024798	0,030319	6,93
CT869	putative protein	1	6	2,9	20,1	0,017599	0,050834	0,030319	0,024798	0,030319	6,93
CT1166	60S ribosomal protein L37	2	13	5,7	45,5	0,000543	0,001222	0,001167	0,000746	0,001167	7,98
CT1269	40S ribosomal protein S27	1	7	2,9	24,5	0,009053	0,026421	0,01564	0,011752	0,01564	8,45
CT1031	CDH1-D	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT107	Phosphopyruvate hydratase (EC 4.2.1.11)	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1153	Coatomer zeta-1 subunit (Zeta-1 coat protein)	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT120	Thioredoxin H	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT132	Abscisic stress ripening protein	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1507	putative protein	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1509	L1332.6 protein	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1514	Hydroquinone glucosyltransferase (EC 2.4.1.218)	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1518	Harpin-induced protein (HIN1)	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1520	Phosphogluconate dehydrogenase (decarboxylating) (EC 1.1.1.44)	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1528	putative protein	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1531	T1G11.5 protein	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1536	Auxin-induced protein 15A	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1538	Unknown protein	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1541	Malate dehydrogenase (oxaloacetate decarboxylating) (NADP+) (EC 1.1.1.40)	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1542	Hypothetical 36.9 kDa protein	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1544	SUR2 protein (Syringomycin response protein 2)	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1557	Cyprosin precursor (EC 3.4.23.-) (Fragment)	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1572	Glutamate--ammonia ligase (EC 6.3.1.2)	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1577	putative protein	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1584	Ubiquinol-cytochrome C reductase iron-sulfur subunit	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1585	Alcohol dehydrogenase (EC 1.1.1.1)	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1587	putative protein	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1590	Rab GDP dissociation inhibitor alpha	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1591	putative protein	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1592	Alpha-glucosidase (EC 3.2.1.20)	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1596	Proteasome endopeptidase complex subunit alpha type 3 (EC 3.4.25.1)	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1598	WRKY transcription factor	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1609	Proteasome endopeptidase complex subunit alpha type 5 (EC 3.4.25.1)	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1614	Glutathione transferase (EC 2.5.1.18)	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1615	RNA-directed DNA polymerase (EC 2.7.7.49)	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1626	Unknown protein	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1634	Iron(III)-zinc(II) purple acid phosphatase precursor	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1639	F18K10.26 protein	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1644	Disease resistance response protein 206	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1648	Hypothetical 51.6 kDa protein	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1670	Flavonol 3-O-glucosyltransferase 5 (EC 2.4.1.91)	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1674	Tubulin beta chain	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1683	Unknown protein	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1684	putative protein	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1720	putative protein	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1725	60S ribosomal protein L41, mitochondrial precursor	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1731	putative protein	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1736	S-receptor kinase PK3	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1739	F12P19.3	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1767	putative protein	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1773	putative protein	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1774	putative protein	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1779	P0482D04.5 protein	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1780	NBS/LRR	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1782	putative protein	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1783	Ubiquitin--protein ligase-like protein (SUMO-1) (EC 6.3.2.19)	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1798	putative protein	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1802	40S ribosomal protein S25	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1807	Protein kinase CTR1 (EC 2.7.1.37)	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1811	40S ribosomal protein S20	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1812	NADH dehydrogenase (ubiquinone) 18 kDa subunit (EC 1.6.5.3), mitochondrial precursor	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1817	putative protein	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1820	Unknown protein	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1821	Phosphate-induced (phi-1) protein	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1822	Zinc finger protein	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1824	putative protein	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1830	putative protein	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1831	Signal peptidase I (EC 3.4.21.89)	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1833	Tubulin beta chain (Beta tubulin)	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1841	Cathepsin B-like cysteine proteinase precursor (EC 3.4.22.1)	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1850	ABA induced plasma membrane protein	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1870	5'-AMP-activated protein kinase, beta-2 subunit	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1877	Rac-GTP binding protein-like	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1878	Hypothetical 228.6 kDa protein	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1885	Acyl carrier protein, mitochondrial precursor (ACP)	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1893	putative protein	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1895	Hypothetical 68.3 kDa protein	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1907	putative protein	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1908	Hypothetical 12.2 kDa protein	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1927	putative protein	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1928	Arginine methyltransferase-like protein (EC 2.1.1.-)	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1938	Histone H1	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50

Contig	Description	Lib1	Lib2	Lib1(norm)	Lib2(norm)	AC 1 2	Fisher 1 2	Chi2x2 1 2	R	Chi	Ratio d'expressions normalisées
CT1939	H(+)-transporting two-sector ATPase delta chain (EC 3.6.3.14), mitochondrial precursor	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1940	putative protein	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1948	60S ribosomal protein L36	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1961	IAA-amino acid hydrolase 1 (EC 3.5.1.-)	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1962	Pantothenate kinase 4 (EC 2.7.1.33)	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1995	Allyl alcohol dehydrogenase-like protein	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT2000	putative protein	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT2107	T20M3.12 protein (Putative chitinase protein (EC 3.2.1.14))	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT2112	putative protein	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT2117	Dihydrokaempferol 4-reductase (EC 1.1.1.219)	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT312	Metallothionein-like protein	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT314	Metallothionein-like protein	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT342	Flavonoid 3',5'-hydroxylase (EC 1.14.-.-)	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT344	Flavonoid 3',5'-hydroxylase (EC 1.14.-.-)	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT600	Proline-rich protein / Cortical cell delineating protein precursor	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT664	Pinus taeda putative cell wall protein	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT884	Protein Z / serpin	0	3	0	10,5	0,050142	0,09113	0,0555	0,028596	0,0555	10,50
CT1120	Cysteine proteinase inhibitor (Cystatin)	0	4	0	13,1	0,022569	0,040993	0,027011	0,011487	0,027011	13,10
CT1431	Auxin-induced protein PCNT115	0	4	0	13,1	0,022569	0,040993	0,027011	0,011487	0,027011	13,10
CT1503	Unknown protein	0	4	0	13,1	0,022569	0,040993	0,027011	0,011487	0,027011	13,10
CT1522	Germin-like protein subfamily 2 member 1 precursor	0	4	0	13,1	0,022569	0,040993	0,027011	0,011487	0,027011	13,10
CT1556	Homeobox-leucine zipper protein	0	4	0	13,1	0,022569	0,040993	0,027011	0,011487	0,027011	13,10
CT1600	Xyloglucan:xyloglucosyl transferase (EC 2.4.1.207)	0	4	0	13,1	0,022569	0,040993	0,027011	0,011487	0,027011	13,10
CT1602	Peptidylprolyl isomerase (EC 5.2.1.8)	0	4	0	13,1	0,022569	0,040993	0,027011	0,011487	0,027011	13,10
CT1611	Unknown protein	0	4	0	13,1	0,022569	0,040993	0,027011	0,011487	0,027011	13,10
CT1627	Unknown protein	0	4	0	13,1	0,022569	0,040993	0,027011	0,011487	0,027011	13,10
CT1663	ATP-dependent RNA helicase	0	4	0	13,1	0,022569	0,040993	0,027011	0,011487	0,027011	13,10
CT1704	Glucan endo-1,3-beta-glucosidase precursor	0	4	0	13,1	0,022569	0,040993	0,027011	0,011487	0,027011	13,10
CT1723	F13M7.16 protein (Hypothetical protein)	0	4	0	13,1	0,022569	0,040993	0,027011	0,011487	0,027011	13,10
CT1744	Unknown protein	0	4	0	13,1	0,022569	0,040993	0,027011	0,011487	0,027011	13,10
CT1749	putative protein	0	4	0	13,1	0,022569	0,040993	0,027011	0,011487	0,027011	13,10
CT1751	putative protein	0	4	0	13,1	0,022569	0,040993	0,027011	0,011487	0,027011	13,10
CT1752	Ubiquitin--protein ligase E2-17 kDa (EC 6.3.2.19)	0	4	0	13,1	0,022569	0,040993	0,027011	0,011487	0,027011	13,10
CT1769	Nucleotide sugar epimerase-like protein	0	4	0	13,1	0,022569	0,040993	0,027011	0,011487	0,027011	13,10
CT1791	Unknown protein	0	4	0	13,1	0,022569	0,040993	0,027011	0,011487	0,027011	13,10
CT1805	MAP kinase	0	4	0	13,1	0,022569	0,040993	0,027011	0,011487	0,027011	13,10
CT1845	Hypothetical 11.2 kDa protein	0	4	0	13,1	0,022569	0,040993	0,027011	0,011487	0,027011	13,10
CT1855	Sex-determining region Y protein	0	4	0	13,1	0,022569	0,040993	0,027011	0,011487	0,027011	13,10
CT1937	Cytochrome P450 (EC 1.14.-.-)	0	4	0	13,1	0,022569	0,040993	0,027011	0,011487	0,027011	13,10
CT1968	Guanine deaminase (EC 3.5.4.3) (Guanase)	0	4	0	13,1	0,022569	0,040993	0,027011	0,011487	0,027011	13,10
CT2105	H(+)-transporting two-sector ATPase delta chain (EC 3.6.3.14), mitochondrial precursor	0	4	0	13,1	0,022569	0,040993	0,027011	0,011487	0,027011	13,10
CT2109	Geranylgeranyl pyrophosphate synthetase, chloroplast precursor (EC 2.5.1.30)	0	4	0	13,1	0,022569	0,040993	0,027011	0,011487	0,027011	13,10
CT387	Germin-like protein	0	4	0	13,1	0,022569	0,040993	0,027011	0,011487	0,027011	13,10
CT941	putative protein	0	4	0	13,1	0,022569	0,040993	0,027011	0,011487	0,027011	13,10
CT1482	Hypothetical 19.9 kDa protein	0	5	0	17,5	0,010158	0,018437	0,013415	0,004717	0,013415	17,50
CT1508	Isocitrate dehydrogenase (NADP+) (EC 1.1.1.42)	0	5	0	17,5	0,010158	0,018437	0,013415	0,004717	0,013415	17,50
CT1512	3-oxoacyl-[acyl-carrier protein] synthase III B (EC 2.3.1.41)	0	5	0	17,5	0,010158	0,018437	0,013415	0,004717	0,013415	17,50
CT1524	DNA-directed RNA polymerase II 14.4 kDa polypeptide	0	5	0	17,5	0,010158	0,018437	0,013415	0,004717	0,013415	17,50
CT1551	putative protein	0	5	0	17,5	0,010158	0,018437	0,013415	0,004717	0,013415	17,50
CT1610	Histone H3	0	5	0	17,5	0,010158	0,018437	0,013415	0,004717	0,013415	17,50
CT1622	60S ribosomal protein L23	0	5	0	17,5	0,010158	0,018437	0,013415	0,004717	0,013415	17,50
CT1640	Pollen specific protein SF3	0	5	0	17,5	0,010158	0,018437	0,013415	0,004717	0,013415	17,50
CT1654	Plasma membrane intrinsic protein	0	5	0	17,5	0,010158	0,018437	0,013415	0,004717	0,013415	17,50
CT1664	dTDP-glucose 4,6-dehydratase (EC 4.2.1.46)	0	5	0	17,5	0,010158	0,018437	0,013415	0,004717	0,013415	17,50
CT1721	60S ribosomal protein L38	0	5	0	17,5	0,010158	0,018437	0,013415	0,004717	0,013415	17,50
CT1746	Phosphoglucomutase-like protein	0	5	0	17,5	0,010158	0,018437	0,013415	0,004717	0,013415	17,50
CT1806	Inorganic polyphosphate/ATP-NAD kinase	0	5	0	17,5	0,010158	0,018437	0,013415	0,004717	0,013415	17,50
CT346	Flavonoid 3',5'-hydroxylase (EC 1.14.-.-)	0	5	0	17,5	0,010158	0,018437	0,013415	0,004717	0,013415	17,50
CT440	putative protein	0	5	0	17,5	0,010158	0,018437	0,013415	0,004717	0,013415	17,50
CT59	Polcalcin Calcium-binding protein	0	5	0	17,5	0,010158	0,018437	0,013415	0,004717	0,013415	17,50
CT60	Calcium binding protein	0	5	0	17,5	0,010158	0,018437	0,013415	0,004717	0,013415	17,50
CT98	putative protein	0	5	0	17,5	0,010158	0,018437	0,013415	0,004717	0,013415	17,50
CT1246	Alpha-1,4-glucan-protein synthase (UDP-forming) (EC 2.4.1.112)	0	6	0	20,1	0,004572	0,00829	0,006753	0,001965	0,006753	20,10
CT1564	60S ribosomal protein L44	0	6	0	20,1	0,004572	0,00829	0,006753	0,001965	0,006753	20,10
CT1666	putative protein	0	6	0	20,1	0,004572	0,00829	0,006753	0,001965	0,006753	20,10
CT1680	Unknown protein	0	6	0	20,1	0,004572	0,00829	0,006753	0,001965	0,006753	20,10
CT1529	Carboxypeptidase C (EC 3.4.16.5)	0	7	0	24,5	0,002058	0,003727	0,003433	0,000828	0,003433	24,50
CT1560	Eukaryotic translation initiation factor 5A	0	7	0	24,5	0,002058	0,003727	0,003433	0,000828	0,003433	24,50
CT275	Late embryogenesis abundant protein	0	9	0	31,5	0,000417	0,000753	0,000906	0,00015	0,000906	31,50
CT336	Germin-like protein	0	12	0	41,1	0,000038	0,000068	0,000127	0,000012	0,000127	41,10
CT1981	Nodulin-like protein (Putative nodulin protein)	0	15	0	52,5	0,000003	0,000006	0,000018	0,000001	0,000018	52,50

Tableau Ax-2 : Résultat du northern *in silico* sur les données de l'assemblage avec regroupements des ESTs en tenant compte des annotations. 247 gènes différentiellement exprimés, statistiquement significatifs pour au moins un des tests statistiques avec un risque α de 0.05. (Lib1 = banque RN (root control) ; Lib2 = banque RS (root stress))

Description	RN	RS	RN (norm)	RS (norm)	AC 1 2	Fisher 1 2	Chi2x2 1 2	R	Chi	Ratio d'expressions normalisées
Phosphate carrier protein	9	0	23,4	0	0,001807	0,00446	0,004848	0,000746	0,004848	23,40
Aspartate aminotransferase (EC 2.6.1.1)	8	0	20,8	0	0,003399	0,008671	0,007915	0,001477	0,007915	20,80
Ubiquitin / 40S ribosomal protein S27a	8	0	20,8	0	0,003399	0,008671	0,007915	0,001477	0,007915	20,80
Phytocyanin	16	1	41,6	2,1	0,000173	0,000393	0,000706	0,000171	0,000706	19,81
At2g15910 protein	7	0	18,2	0	0,006392	0,01691	0,012993	0,00294	0,012993	18,20
Minor extracellular protease vpr precursor (EC 3.4.21.-)	7	0	18,2	0	0,006392	0,01691	0,012993	0,00294	0,012993	18,20
Protein Farnesyl-diphosphate farnesyltransferase alpha (EC 2.5.1.21)	7	0	18,2	0	0,006392	0,01691	0,012993	0,00294	0,012993	18,20
Adenine phosphoribosyltransferase (EC 2.4.2.7)	6	0	15,6	0	0,012021	0,033081	0,021474	0,005899	0,021474	15,60
Enoyl-[acyl-carrier protein] reductase [NADH] (EC 1.3.1.9)	6	0	15,6	0	0,012021	0,033081	0,021474	0,005899	0,021474	15,60
Leucine-rich repeat protein SHOC-2 (Ras-binding protein)	6	0	15,6	0	0,012021	0,033081	0,021474	0,005899	0,021474	15,60
Membrane-associated salt-inducible-like protein	6	0	15,6	0	0,012021	0,033081	0,021474	0,005899	0,021474	15,60
Mitochondrial import receptor	6	0	15,6	0	0,012021	0,033081	0,021474	0,005899	0,021474	15,60
Polygalacturonase inhibitor precursor	6	0	15,6	0	0,012021	0,033081	0,021474	0,005899	0,021474	15,60
Unknown_820	6	0	15,6	0	0,012021	0,033081	0,021474	0,005899	0,021474	15,60
40S ribosomal protein S3a	5	0	13	0	0,022606	0,064938	0,035812	0,011953	0,035812	13,00
5S ribosomal RNA	5	0	13	0	0,022606	0,064938	0,035812	0,011953	0,035812	13,00
Ankyrin protein	5	0	13	0	0,022606	0,064938	0,035812	0,011953	0,035812	13,00
DNA-binding protein ESCAROLA	5	0	13	0	0,022606	0,064938	0,035812	0,011953	0,035812	13,00
Expansin-like precursor	5	0	13	0	0,022606	0,064938	0,035812	0,011953	0,035812	13,00
Methyl binding domain protein	5	0	13	0	0,022606	0,064938	0,035812	0,011953	0,035812	13,00
Peptide transporter PTR2-B	5	0	13	0	0,022606	0,064938	0,035812	0,011953	0,035812	13,00
P-protein [Includes: Chorismate mutase (EC 5.4.99.5)	5	0	13	0	0,022606	0,064938	0,035812	0,011953	0,035812	13,00
Thiazole biosynthetic enzyme, chloroplast precursor	5	0	13	0	0,022606	0,064938	0,035812	0,011953	0,035812	13,00
Unknown_1058	5	0	13	0	0,022606	0,064938	0,035812	0,011953	0,035812	13,00
Unknown_1477	5	0	13	0	0,022606	0,064938	0,035812	0,011953	0,035812	13,00
Unknown_175	5	0	13	0	0,022606	0,064938	0,035812	0,011953	0,035812	13,00
Unknown_845	5	0	13	0	0,022606	0,064938	0,035812	0,011953	0,035812	13,00
Unknown_971	5	0	13	0	0,022606	0,064938	0,035812	0,011953	0,035812	13,00
40S ribosomal protein S15	10	1	26	2,1	0,00495	0,013409	0,012074	0,006339	0,012074	12,38
Serine-rich protein	10	1	26	2,1	0,00495	0,013409	0,012074	0,006339	0,012074	12,38
Inositol 1-alpha-galactosyltransferase (EC 2.4.1.123)	9	1	23,4	2,1	0,008463	0,023896	0,019525	0,011487	0,019525	11,14
L-ascorbate oxidase precursor (EC 1.10.3.3)	9	1	23,4	2,1	0,008463	0,023896	0,019525	0,011487	0,019525	11,14
Adenosine kinase (EC 2.7.1.20)	4	0	10,4	0	0,042513	0,127922	0,060475	0,024559	0,060475	10,40
Alpha-expansin precursor	4	0	10,4	0	0,042513	0,127922	0,060475	0,024559	0,060475	10,40
Carbonate dehydratase (EC 4.2.1.1)	4	0	10,4	0	0,042513	0,127922	0,060475	0,024559	0,060475	10,40
Cyclin-dependent kinases regulatory subunit	4	0	10,4	0	0,042513	0,127922	0,060475	0,024559	0,060475	10,40
Dihydroorotase (EC 3.5.2.3) (Hypothetical protein)	4	0	10,4	0	0,042513	0,127922	0,060475	0,024559	0,060475	10,40
Microsomal signal peptidase 22 kDa subunit	4	0	10,4	0	0,042513	0,127922	0,060475	0,024559	0,060475	10,40
Pyruvate dehydrogenase (EC 1.2.4.1)	4	0	10,4	0	0,042513	0,127922	0,060475	0,024559	0,060475	10,40
Ribulose bisphosphate carboxylase small chain (EC 4.1.1.39)	4	0	10,4	0	0,042513	0,127922	0,060475	0,024559	0,060475	10,40
RNA-binding protein with multiple splicing homolog	4	0	10,4	0	0,042513	0,127922	0,060475	0,024559	0,060475	10,40
Unknown_1	4	0	10,4	0	0,042513	0,127922	0,060475	0,024559	0,060475	10,40
Unknown_1065	4	0	10,4	0	0,042513	0,127922	0,060475	0,024559	0,060475	10,40
Unknown_1260	4	0	10,4	0	0,042513	0,127922	0,060475	0,024559	0,060475	10,40
Unknown_1333	4	0	10,4	0	0,042513	0,127922	0,060475	0,024559	0,060475	10,40
Unknown_1394	4	0	10,4	0	0,042513	0,127922	0,060475	0,024559	0,060475	10,40
Unknown_1413	4	0	10,4	0	0,042513	0,127922	0,060475	0,024559	0,060475	10,40
Unknown_192	4	0	10,4	0	0,042513	0,127922	0,060475	0,024559	0,060475	10,40
Unknown_217	4	0	10,4	0	0,042513	0,127922	0,060475	0,024559	0,060475	10,40
Unknown_24	4	0	10,4	0	0,042513	0,127922	0,060475	0,024559	0,060475	10,40
Unknown_267	4	0	10,4	0	0,042513	0,127922	0,060475	0,024559	0,060475	10,40
Unknown_372	4	0	10,4	0	0,042513	0,127922	0,060475	0,024559	0,060475	10,40
Unknown_391	4	0	10,4	0	0,042513	0,127922	0,060475	0,024559	0,060475	10,40
Unknown_394	4	0	10,4	0	0,042513	0,127922	0,060475	0,024559	0,060475	10,40
Unknown_427	4	0	10,4	0	0,042513	0,127922	0,060475	0,024559	0,060475	10,40
Unknown_481	4	0	10,4	0	0,042513	0,127922	0,060475	0,024559	0,060475	10,40
Unknown_506	4	0	10,4	0	0,042513	0,127922	0,060475	0,024559	0,060475	10,40
Unknown_613	4	0	10,4	0	0,042513	0,127922	0,060475	0,024559	0,060475	10,40
Unknown_653	4	0	10,4	0	0,042513	0,127922	0,060475	0,024559	0,060475	10,40
Unknown_798	4	0	10,4	0	0,042513	0,127922	0,060475	0,024559	0,060475	10,40
Unknown_848	4	0	10,4	0	0,042513	0,127922	0,060475	0,024559	0,060475	10,40
Unknown_957	4	0	10,4	0	0,042513	0,127922	0,060475	0,024559	0,060475	10,40
Acid phosphatase (EC 3.1.3.2)	8	1	20,8	2,1	0,014324	0,042361	0,031672	0,020747	0,031672	9,90
40S ribosomal protein S14	7	1	18,2	2,1	0,023945	0,074578	0,051567	0,037328	0,051567	8,67
40S ribosomal protein S16	7	1	18,2	2,1	0,023945	0,074578	0,051567	0,037328	0,051567	8,67
60S ribosomal protein L10	7	1	18,2	2,1	0,023945	0,074578	0,051567	0,037328	0,051567	8,67
60S ribosomal protein L18	7	1	18,2	2,1	0,023945	0,074578	0,051567	0,037328	0,051567	8,67
Beta-fructofuranosidase (EC 3.2.1.26)	7	1	18,2	2,1	0,023945	0,074578	0,051567	0,037328	0,051567	8,67
60S acidic ribosomal protein	6	1	15,6	2,1	0,039402	0,130111	0,084321	0,066842	0,084321	7,43
60S ribosomal protein L32	6	1	15,6	2,1	0,039402	0,130111	0,084321	0,066842	0,084321	7,43
ADP,ATP carrier protein	6	1	15,6	2,1	0,039402	0,130111	0,084321	0,066842	0,084321	7,43
Calcineurin B	6	1	15,6	2,1	0,039402	0,130111	0,084321	0,066842	0,084321	7,43
Caltractin	6	1	15,6	2,1	0,039402	0,130111	0,084321	0,066842	0,084321	7,43
Hypothetical 28.8 kDa protein	6	1	15,6	2,1	0,039402	0,130111	0,084321	0,066842	0,084321	7,43
Polygalacturonase (EC 3.2.1.15)	6	1	15,6	2,1	0,039402	0,130111	0,084321	0,066842	0,084321	7,43
Syntaxin	6	1	15,6	2,1	0,039402	0,130111	0,084321	0,066842	0,084321	7,43
Unknown_656	6	1	15,6	2,1	0,039402	0,130111	0,084321	0,066842	0,084321	7,43
Ethylene responsive element binding factor	29	4	75,4	11,8	0,000012	0,000031	0,000062	0,000019	0,000062	6,39
Phenylalanine ammonia-lyase (EC 4.3.1.5)	11	2	28,6	5,9	0,00874	0,025751	0,022975	0,016288	0,022975	4,85
60S ribosomal protein L23A	9	2	23,4	5,9	0,021796	0,071089	0,056743	0,04614	0,056743	3,97
26S proteasome non-ATPase regulatory subunit	13	3	33,8	8,9	0,008308	0,025204	0,024255	0,018637	0,024255	3,80
Beta-glucanase	8	2	20,8	5,9	0,033537	0,116219	0,088908	0,076566	0,088908	3,53
Unknown_233	16	4	41,6	11,8	0,005061	0,022506	0,016062	0,012289	0,016062	3,53
Unknown_178	12	3	31,2	8,9	0,012695	0,040381	0,037138	0,030119	0,037138	3,51
Defensin	23	6	59,8	17,7	0,001308	0,004683	0,00471	0,003375	0,00471	3,38
Protein-methionine-S-oxide reductase (EC 1.8.4.6)	11	3	28,6	8,9	0,019099	0,064056	0,056624	0,048195	0,056624	3,21
Pectinesterase (EC 3.1.1.11)	14	4	36,4	11,8	0,011304	0,055528	0,036218	0,030356	0,036218	3,08
Hypothetical 19.3 kDa protein	10	3	26	8,9	0,02822	0,100446	0,085883	0,076257	0,085883	2,92
14-3-3 protein	9	3	23,4	8,9	0,040824	0,155391	0,129414	0,119108	0,129414	2,63
5-methyltetrahydropteroyltriglutamate--homocysteine S-methyltransferase (EC 2.1.1.14)	9	3	23,4	8,9	0,040824	0,155391	0,129414	0,119108	0,129414	2,63
60S ribosomal protein L5	9	3	23,4	8,9	0,040824	0,155391	0,129414	0,119108	0,129414	2,63
Auxin-induced protein	9	3	23,4	8,9	0,040824	0,155391	0,129414	0,119108	0,129414	2,63
Extensin	11	4	28,6	11,8	0,033536	0,129162	0,117287	0,108809	0,117287	2,42

Description	RN	RS	RN (norm)	RS (norm)	AC 1 2	Fisher 1 2	Chi2x2 1 2	R	Chi	Ratio d'expressions normalisées
Cyclin A/CDK2-associated protein	10	4	26	11,8	0,04625	0,19142	0,170661	0,162138	0,170661	2,20
Peroxiredoxin	10	4	26	11,8	0,04625	0,19142	0,170661	0,162138	0,170661	2,20
Catalase (EC 1.11.1.6)	16	7	41,6	20,7	0,02629	0,14334	0,114609	0,108901	0,114609	2,01
60S ribosomal protein L41	19	9	49,4	26,6	0,024418	0,13195	0,118758	0,114061	0,118758	1,86
Phosphopyruvate hydratase (EC 4.2.1.11)	44	22	114,5	64,1	0,004654	0,03441	0,027334	0,026111	0,027334	1,79
40S ribosomal protein S7	16	8	41,6	23,6	0,036931	0,221197	0,18457	0,179464	0,18457	1,76
Peroxidase precursor (EC 1.11.1.7)	76	38	197,7	112,3	0,000588	0,004341	0,003615	0,003465	0,003615	1,76
Metallothionein-like protein EMB30	18	9	46,8	26,6	0,03116	0,179603	0,159252	0,154544	0,159252	1,76
ADP-ribosylation factor	20	11	52	32,5	0,034901	0,213431	0,204726	0,200917	0,204726	1,60
Naringenin-chalcone synthase (EC 2.3.1.74)	43	24	111,9	70,9	0,010248	0,084495	0,069745	0,068386	0,069745	1,58
Translationally controlled tumor protein	25	17	65	50,2	0,04704	0,441322	0,408233	0,406318	0,408233	1,29
Phosphate-induced protein	30	21	78	62	0,043211	0,482136	0,417187	0,415795	0,417187	1,26
Ubiquitin	59	47	153,5	138,8	0,036265	0,625197	0,605389	0,604881	0,605389	1,11
Late embryogenesis abundant protein	51	52	132,7	153,6	0,031632	0,486898	0,45338	0,45549	0,45338	1,16
Ubiquitin--protein ligase (EC 6.3.2.19)	34	40	88,4	118,2	0,022732	0,2417	0,210342	0,212706	0,210342	1,34
Elongation factor 1-alpha	15	19	39	56,1	0,041589	0,305861	0,288754	0,28956	0,288754	1,44
Polcalcin	33	42	85,8	124,1	0,01388	0,130115	0,109456	0,111416	0,109456	1,45
Metallothionein-like protein	25	32	65	94,5	0,020924	0,182702	0,157075	0,158739	0,157075	1,45
Peptidylprolyl isomerase (EC 5.2.1.8)	15	20	39	59,1	0,03408	0,237918	0,220111	0,220991	0,220111	1,52
Germin-like protein	21	29	54,6	85,7	0,017284	0,119356	0,112051	0,113214	0,112051	1,57
Cell wall protein	18	26	46,8	76,8	0,017121	0,128936	0,101946	0,102792	0,101946	1,64
DnaJ protein homolog	6	10	15,6	29,5	0,048783	0,221562	0,208426	0,207618	0,208426	1,89
Tubulin beta chain (Beta tubulin)	6	10	15,6	29,5	0,048783	0,221562	0,208426	0,207618	0,208426	1,89
Unknown_319	7	12	18,2	35,5	0,035788	0,172293	0,153119	0,15241	0,153119	1,95
Avr9/Cf-9 rapidly elicited protein 146	12	21	31,2	62	0,011595	0,055966	0,052406	0,052253	0,052406	1,99
Flavonoid 3',5'-hydroxylase (EC 1.14.-.-)	13	23	33,8	67,9	0,008804	0,044655	0,039696	0,039583	0,039696	2,01
Histone H1	11	21	28,6	62	0,007929	0,049115	0,032681	0,032294	0,032681	2,17
Zinc finger protein	5	10	13	29,5	0,034403	0,193771	0,123167	0,121383	0,123167	2,27
Glutathione transferase (EC 2.5.1.18)	9	18	23,4	53,2	0,009922	0,051713	0,038451	0,037762	0,038451	2,27
Proteasome endopeptidase complex subunit (EC 3.4.25.1)	6	13	15,6	38,4	0,016969	0,067367	0,058896	0,057379	0,058896	2,46
Cellulase precursor (EC 3.2.1.4) (Endo-1,4-beta glucanase)	5	11	13	32,5	0,023431	0,085266	0,078497	0,076509	0,078497	2,50
Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.-)	3	7	7,8	20,7	0,047355	0,205207	0,141649	0,138128	0,141649	2,65
Naringenin 3-dioxygenase (EC 1.14.11.9)	3	7	7,8	20,7	0,047355	0,205207	0,141649	0,138128	0,141649	2,65
Eukaryotic translation initiation factor 5A	3	8	7,8	23,6	0,030489	0,128344	0,084901	0,08111	0,084901	3,03
Tubulin alpha chain	3	8	7,8	23,6	0,030489	0,128344	0,084901	0,08111	0,084901	3,03
Beta-glucosidase (EC 3.2.1.21)	2	6	5,2	17,7	0,044377	0,157965	0,110063	0,104338	0,110063	3,40
Geranylgeranyl pyrophosphate synthetase (EC2.5.1.10)	2	6	5,2	17,7	0,044377	0,157965	0,110063	0,104338	0,110063	3,40
MAP kinase homolog (EC 2.7.1.-)	2	6	5,2	17,7	0,044377	0,157965	0,110063	0,104338	0,110063	3,40
Unknown_176	2	6	5,2	17,7	0,044377	0,157965	0,110063	0,104338	0,110063	3,40
Unknown_74	3	9	7,8	26,6	0,019036	0,078704	0,05028	0,046734	0,05028	3,41
60S ribosomal protein L37	4	13	10,4	38,4	0,005264	0,015409	0,014189	0,012533	0,014189	3,69
Pathogenesis-related protein	4	14	10,4	41,4	0,003169	0,009103	0,008414	0,007161	0,008414	3,98
Unknown_70	2	7	5,2	20,7	0,026717	0,092529	0,062594	0,057154	0,062594	3,98
Unknown_189	3	11	7,8	32,5	0,006906	0,028233	0,017183	0,01482	0,017183	4,17
60S ribosomal protein L44	2	8	5,2	23,6	0,015638	0,053236	0,035391	0,030833	0,035391	4,54
Carboxypeptidase C (EC 3.4.16.5)	2	8	5,2	23,6	0,015638	0,053236	0,035391	0,030833	0,035391	4,54
Hypothetical 8.0 kDa protein	2	8	5,2	23,6	0,015638	0,053236	0,035391	0,030833	0,035391	4,54
Unknown_71	2	8	5,2	23,6	0,015638	0,053236	0,035391	0,030833	0,035391	4,54
Glucan endo-1,3-beta-D-glucosidase (EC 3.2.1.39)	2	10	5,2	29,5	0,005029	0,016908	0,011194	0,008659	0,011194	5,67
Unknown_75	2	10	5,2	29,5	0,005029	0,016908	0,011194	0,008659	0,011194	5,67
40S ribosomal protein S26 (S31)	1	5	2,6	14,8	0,038191	0,10484	0,072992	0,063332	0,072992	5,69
40S ribosomal protein S9 (S7)	1	5	2,6	14,8	0,038191	0,10484	0,072992	0,063332	0,072992	5,69
Aldehyde dehydrogenase (NAD+) 1A (EC 1.2.1.3)	1	5	2,6	14,8	0,038191	0,10484	0,072992	0,063332	0,072992	5,69
Inorganic polyphosphate/ATP-NAD kinase	1	5	2,6	14,8	0,038191	0,10484	0,072992	0,063332	0,072992	5,69
Inosine-5'-monophosphate dehydrogenase (EC 1.1.1.2)	1	5	2,6	14,8	0,038191	0,10484	0,072992	0,063332	0,072992	5,69
Major pollen allergen Bet v 1-D/H / Intracellular pathogenesis-related protein	1	5	2,6	14,8	0,038191	0,10484	0,072992	0,063332	0,072992	5,69
Pectate lyase (EC 4.2.2.2)	1	5	2,6	14,8	0,038191	0,10484	0,072992	0,063332	0,072992	5,69
40S ribosomal protein S27	2	11	5,2	32,5	0,002783	0,009377	0,006275	0,004524	0,006275	6,25
Harpin-induced protein	2	11	5,2	32,5	0,002783	0,009377	0,006275	0,004524	0,006275	6,25
Annexin	1	6	2,6	17,7	0,020864	0,056086	0,039115	0,031408	0,039115	6,81
Cathepsin B precursor (EC 3.4.22.1)	1	6	2,6	17,7	0,020864	0,056086	0,039115	0,031408	0,039115	6,81
Hypothetical 19.9 kDa protein	1	6	2,6	17,7	0,020864	0,056086	0,039115	0,031408	0,039115	6,81
LATERAL ORGAN BOUNDARIES protein (LOB)	1	6	2,6	17,7	0,020864	0,056086	0,039115	0,031408	0,039115	6,81
Myb-related protein	1	6	2,6	17,7	0,020864	0,056086	0,039115	0,031408	0,039115	6,81
RNA helicase	1	6	2,6	17,7	0,020864	0,056086	0,039115	0,031408	0,039115	6,81
Unknown_981	1	6	2,6	17,7	0,020864	0,056086	0,039115	0,031408	0,039115	6,81
Hydroquinone glucosyltransferase (EC 2.4.1.218)	1	7	2,6	20,7	0,011165	0,029619	0,021064	0,015477	0,021064	7,96
UDP-glucose : protein transglucosylase (EC 2.4.1.112)	1	7	2,6	20,7	0,011165	0,029619	0,021064	0,015477	0,021064	7,96
40S ribosomal protein S20	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
40S ribosomal protein S25	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
60S ribosomal protein L41, mitochondrial precursor	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
ABA induced plasma membrane protein	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
Allyl alcohol dehydrogenase-like protein	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
Alpha-galactosidase precursor (EC 3.2.1.22)	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
Arginine methyltransferase-like protein (EC 2.1.1.-)	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
ATP-binding cassette	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
Auxin-induced protein 15A	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
Cyprosin precursor (EC 3.4.23.-) (Fragment)	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
Disease resistance protein	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
F12P19.3	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
F18K10.26 protein	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
Gb|AAD48513.1	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
Glycine-rich RNA-binding protein 2, mitochondrial	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
Hypothetical 12.2 kDa protein	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
Hypothetical 228.6 kDa protein	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
Hypothetical 36.9 kDa protein	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
Hypothetical 51.6 kDa protein	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
IAA-amino acid hydrolase 1 (EC 3.5.1.-)	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
Iron(III)-zinc(II) purple acid phosphatase precursor	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
L1332.6 protein	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
Mitochondrial import receptor subunit TOM20	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
P0482D04.5 protein	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
Phosphogluconate dehydrogenase (decarboxylating) (EC 1.1.1.44)	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
Polyubiquitin	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
Protein kinase CTR1 (EC 2.7.1.37)	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
Rab GDP dissociation inhibitor alpha	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
RNA-directed DNA polymerase (EC 2.7.7.49)	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
Signal peptidase I (EC 3.4.21.89)	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
S-receptor kinase PK3	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
SUR2 protein (Syringomycin response protein 2)	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
T1G11.5 protein	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
Ubiquinol-cytochrome C reductase iron-sulfur subunit	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
Unknown_1635	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
Unknown_1657	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
Unknown_1707	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
Unknown_1717	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
Unknown_1721	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
Unknown_1816	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
Unknown_1853	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90

Description	RN	RS	RN (norm)	RS (norm)	AC 1 2	Fisher 1 2	Chi2x2 1 2	R	Chi	Ratio d'expressions normalisées
Unknown_1864	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
Unknown_1900	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
Unknown_1906	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
Unknown_1907	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
Unknown_1916	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
Unknown_1933	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
Unknown_1953	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
Unknown_1960	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
Unknown_1966	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
Unknown_2029	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
Unknown_2043	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
Unknown_2064	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
Unknown_2077	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
Unknown_2140	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
Unknown_2255	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
Unknown_980	0	3	0	8,9	0,054594	0,102621	0,064873	0,032833	0,064873	8,90
Zinc finger protein 216	1	8	2,6	23,6	0,005882	0,015485	0,01139	0,007588	0,01139	9,08
Caffeate O-methyltransferase (EC 2.1.1.68)	0	4	0	11,8	0,025564	0,04803	0,033017	0,013734	0,033017	11,80
F13M7.16 protein (Hypothetical protein)	0	4	0	11,8	0,025564	0,04803	0,033017	0,013734	0,033017	11,80
Glutamate--ammonia ligase (EC 6.3.1.2)	0	4	0	11,8	0,025564	0,04803	0,033017	0,013734	0,033017	11,80
Guanine deaminase (EC 3.5.4.3) (Guanase)	0	4	0	11,8	0,025564	0,04803	0,033017	0,013734	0,033017	11,80
Hypothetical 11.2 kDa protein	0	4	0	11,8	0,025564	0,04803	0,033017	0,013734	0,033017	11,80
Hypothetical 20.6 kDa protein	0	4	0	11,8	0,025564	0,04803	0,033017	0,013734	0,033017	11,80
Nucleotide sugar epimerase-like protein	0	4	0	11,8	0,025564	0,04803	0,033017	0,013734	0,033017	11,80
Secretory carrier membrane protein.	0	4	0	11,8	0,025564	0,04803	0,033017	0,013734	0,033017	11,80
Unknown_1053	0	4	0	11,8	0,025564	0,04803	0,033017	0,013734	0,033017	11,80
Unknown_1882	0	4	0	11,8	0,025564	0,04803	0,033017	0,013734	0,033017	11,80
Unknown_1884	0	4	0	11,8	0,025564	0,04803	0,033017	0,013734	0,033017	11,80
3-oxoacyl-[acyl-carrier protein] synthase III B (EC 2.3.1.41)	0	5	0	14,8	0,01197	0,022476	0,017141	0,005871	0,017141	14,80
60S ribosomal protein L36	0	5	0	14,8	0,01197	0,022476	0,017141	0,005871	0,017141	14,80
Alpha-glucosidase (EC 3.2.1.20)	0	5	0	14,8	0,01197	0,022476	0,017141	0,005871	0,017141	14,80
Homeobox-leucine zipper protein	0	5	0	14,8	0,01197	0,022476	0,017141	0,005871	0,017141	14,80
Isocitrate dehydrogenase (NADP+) (EC 1.1.1.42)	0	5	0	14,8	0,01197	0,022476	0,017141	0,005871	0,017141	14,80
NBS/LRR protein	0	5	0	14,8	0,01197	0,022476	0,017141	0,005871	0,017141	14,80
Pantothenate kinase (EC 2.7.1.33)	0	5	0	14,8	0,01197	0,022476	0,017141	0,005871	0,017141	14,80
Phosphoglucomutase-like protein	0	5	0	14,8	0,01197	0,022476	0,017141	0,005871	0,017141	14,80
Protein kinase (EC 2.7.1.-)	0	5	0	14,8	0,01197	0,022476	0,017141	0,005871	0,017141	14,80
Sex-determining region Y protein	0	5	0	14,8	0,01197	0,022476	0,017141	0,005871	0,017141	14,80
Unknown_1680	0	5	0	14,8	0,01197	0,022476	0,017141	0,005871	0,017141	14,80
Unknown_1796	0	5	0	14,8	0,01197	0,022476	0,017141	0,005871	0,017141	14,80
Unknown_188	0	5	0	14,8	0,01197	0,022476	0,017141	0,005871	0,017141	14,80
Unknown_538	0	5	0	14,8	0,01197	0,022476	0,017141	0,005871	0,017141	14,80
60S ribosomal protein L23	0	6	0	17,7	0,005605	0,010516	0,009018	0,002546	0,009018	17,70
Pollen specific protein	0	9	0	26,6	0,000575	0,001076	0,001379	0,000219	0,001379	26,60
Nodulin-like protein	0	19	0	56,1	0	0,000001	0,000003	0	0,000003	56,10

Tableau Ax-3 : Gènes différentiellement exprimés dans le **stress court** et paramètres statistiques associés

Pvalprov/cond/inter : *valeur de la probabilité associée au test pour les facteurs provenance, condition et interaction* ; r^2 prov/ cond/ inter/ tot/ PCI/ max : *part de la variabilité expliquée par les facteurs provenance (P), condition (C), intéraction (I) et pour la totalité du modèle, l'ensemble des facteurs P,C et I, et par le facteur expliquant le plus de variabilité*. Fact prov/cond : *niveau de régulation pour les facteurs P ou C.*

Acc Number	ID	Pvalprov	r2 prov	Pvalcond	r2 cond	Pvalinter	r2 inter	r2 tot	r2 PCI	r2 max	data > Bkg	Moy France	Moy Maroc	Moy Cont	Moy S6H	Moy S48H	fact prov	fact cond	Annotation	Catégories fonctionnelles
BX681136	RS54B08	7,44E-07	12	6,36E-19	66	1,00E+00	0	79	78	66	74	11,235	11,917	12,627	11,722	10,405	1,60	4,66	Dehydration stress-induced protein	Réponse au Stress
CR393156	RN53C07	2,50E-08	27	1,57E-11	47	1,00E+00	2	76	76	47	50	9,572	10,789	11,099	10,380	9,121	2,33	3,94	Peroxidase precursor (EC 1.11.1.7)	Réponse au Stress
BX000656	HSP	2,00E-06	10	6,36E-19	59	1,91E-02	5	81	74	59	58	9,341	10,573	8,934	9,148	11,661	2,35	6,62	Heat shock protein	Réponse au Stress
BX678783	RS15A09	2,11E-09	20	4,87E-17	51	2,09E-01	3	79	74	51	72	10,479	11,683	12,038	11,306	9,997	2,30	4,12	Peroxidase precursor (EC 1.11.1.7)	Réponse au Stress
BX253089	Ge077G07	8,33E-08	18	3,69E-15	50	6,22E-02	5	75	73	50	75	10,645	11,515	10,296	11,176	11,920	1,83	3,08	LEA, complete	Réponse au Stress
BX252366	Ge063E01	2,49E-06	10	5,67E-18	56	2,81E-02	5	80	71	56	54	9,088	9,840	8,803	9,358	10,330	1,68	2,88	Class I Heat Shock Protein	Réponse au Stress
CR393398	RN59F04	5,77E-09	25	6,46E-13	44	1,00E+00	0	71	70	44	76	11,143	12,104	12,303	11,983	10,733	1,95	2,97	putative protein	_
BX680845	RS49D06	2,04E-12	38	1,23E-11	32	1,00E+00	0	76	70	38	74	10,470	11,816	11,904	10,992	10,521	2,54	2,61	Polyphosphoinositide binding protein	Organisation cellulaire
CR394335	RN76H08	1,74E-12	36	1,15E-12	32	1,00E+00	1	78	69	36	76	10,799	12,205	12,266	11,371	10,868	2,65	2,64	Glutathione transferase (EC 2.5.1.18)	Métabolisme
BX254083	Ge093A12	1,70E-08	26	5,53E-11	39	1,00E+00	0	68	65	39	66	9,442	10,334	10,482	10,076	9,194	1,86	2,44	Geranyl diphosphate synthase, complete	Métabolisme
BX675142	RN10A03	3,62E-06	20	1,14E-09	42	1,00E+00	3	67	65	42	74	10,763	11,422	11,879	11,126	10,200	1,58	3,20	Peroxidase precursor (EC 1.11.1.7)	Réponse au Stress
CR392220	RN28D08	8,01E-10	37	4,00E-08	27	1,00E+00	0	65	64	37	76	12,713	13,601	13,704	13,083	12,681	1,85	2,03	putative protein	_
CR393359	RN59B03	4,92E-06	14	2,10E-12	45	8,64E-01	3	69	62	45	76	12,143	13,144	13,335	13,238	11,592	2,00	3,35	putative protein	_
BX681080	RS53A02	2,61E-07	22	2,39E-10	37	9,42E-01	3	66	62	37	76	11,702	12,570	11,736	11,669	12,926	1,83	2,39	Acetylornithine deacetylase (EC 3.5.1.16)	Devenir des protéines
non soumis	RN28G12	5,58E-07	19	1,91E-11	42	1,00E+00	0	67	61	42	75	10,967	11,411	11,941	11,059	10,499	1,36	2,72	putative protein	_
CR392399	RN31D03	8,39E-06	17	2,10E-10	43	1,00E+00	1	61	60	43	72	10,132	10,789	11,071	10,658	9,726	1,58	2,54	putative protein	_
CR392510	RN33D10	7,00E-07	17	6,54E-11	35	8,82E-02	6	70	58	35	76	11,776	13,005	13,175	12,481	11,580	2,34	3,02	putative protein	_
BX252660	Ge068G01	1,11E-06	22	7,58E-09	36	1,00E+00	0	59	58	36	58	9,348	9,757	9,967	9,561	9,133	1,33	1,78	Histone H2A	Organisation cellulaire
non soumis	RN32A07	1,00E-07	29	4,71E-06	22	2,41E-01	6	58	57	29	76	10,649	11,280	11,391	10,891	10,602	1,55	1,73	putative protein	_
BX679298	RS22F06	2,85E-07	21	7,71E-10	35	1,00E+00	0	66	56	35	76	11,192	12,154	11,267	11,309	12,411	1,95	2,21	putative protein	_
AL750051	AS05C03	1,03E-07	28	2,18E-06	23	1,00E+00	2	60	53	28	76	11,419	12,282	12,282	11,877	11,433	1,82	1,80	Hypothetical 49.0 kDa protein	Non-classé
BX680660	RS46H02	1,66E-07	26	5,03E-07	26	1,00E+00	0	60	52	26	76	11,037	11,979	11,121	11,430	12,040	1,92	1,89	putative protein	_
BX681748	RS63G11	1,53E-06	23	4,15E-07	29	1,00E+00	0	56	52	29	60	9,137	9,948	9,966	9,876	8,932	1,75	2,05	2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase (EC 4.6	Métabolisme
BX252163	Ge060F11	9,93E-08	27	7,44E-07	24	1,00E+00	1	61	52	27	76	10,737	11,535	10,830	11,069	11,565	1,74	1,66	Unknown	Non-classé
BX679222	RS21G07	1,29E-07	26	2,26E-06	22	7,64E-01	4	62	51	26	50	8,571	9,346	9,280	9,096	8,580	1,71	1,62	putative protein	_
CR393796	RN65E04	1,15E-06	25	7,44E-06	23	1,00E+00	3	54	50	25	28	7,993	8,604	8,629	8,348	7,955	1,53	1,60	Pinene synthase (EC 4.2.3.14)	Métabolisme
BX682869	Ge122H09	3,64E-06	20	2,26E-07	30	1,00E+00	0	57	50	30	61	9,310	10,371	10,472	9,945	9,172	2,09	2,46	Pinene synthase (EC 4.2.3.14)	Métabolisme
BX680940	RS51B12	6,92E-06	20	6,29E-07	29	1,00E+00	1	54	50	29	76	10,538	11,061	11,271	10,456	10,553	1,44	1,76	putative protein	_
BX680053	RS38F01	9,36E-09	23	2,83E-09	25	1,00E+00	0	73	49	25	76	11,808	10,990	11,111	11,034	11,882	1,76	1,80	putative protein	_
BX679791	RS34A04	1,64E-06	24	5,20E-06	24	1,00E+00	1	54	49	24	76	11,445	12,170	12,256	11,676	11,463	1,65	1,73	60S ribosomal protein	Synthèse de protéines
BX678491	RS10A02	8,35E-06	21	5,64E-06	26	1,00E+00	0	50	47	26	49	8,723	9,522	9,645	9,137	8,610	1,74	2,05	Pathogenesis-related protein 5 precursor (PR-5)	Réponse au Stress
BX250435	Ge036H06	2,14E-06	18	5,13E-08	27	1,00E+00	2	65	46	27	76	13,210	12,667	12,716	12,730	13,272	1,46	1,47	CTV.15, partial (41%)	Non-classé
BX252622	Ge068A05	5,02E-06	20	8,17E-06	22	1,00E+00	3	56	45	22	73	9,935	10,879	10,142	9,965	11,043	1,92	2,11	Elongation factor 1-gamma (EF-1-gamma) (eEF-1B gamma), p	Synthèse de protéines
CR392614	RN35E03	2,74E-07	28	1,11E-05	21	1,00E+00	1	56	50	28	64	9,499	10,334	10,252	10,208	9,429	1,78	1,77	Acid phosphatase	Métabolisme
CR392259	RN30A11	3,06E-07	28	5,82E-05	18	1,00E+00	1	55	47	28	71	10,238	11,398	11,229	11,164	10,237	2,23	1,99	Acid phosphatase precursor 1 (EC 3.1.3.2)	Métabolisme
BX255226	Ge113B09	2,60E-12	47	4,68E-04	10	5,07E-01	3	70	60	47	74	11,008	12,576	12,200	11,922	11,383	2,96	1,76	TrPRP2, partial (67%)	Non-classé
BX680425	RS43C11	6,20E-08	36	4,55E-04	15	1,00E+00	1	53	51	36	70	10,150	11,120	11,092	10,609	10,197	1,96	1,86	putative protein	_
BX253047	Ge077B10	7,32E-06	23	9,71E-04	15	2,09E-01	8	47	46	23	72	10,175	10,653	10,307	10,003	10,827	1,39	1,77	Aminotransferase	Métabolisme
BX682634	Ge119A01	3,97E-06	25	1,51E-04	19	1,00E+00	1	51	45	25	70	9,812	10,347	10,357	10,123	9,795	1,45	1,48	Elongation factor 1-gamma (EF-1-gamma)	Synthèse de protéines
CR394276	RN76B02	5,36E-06	23	5,10E-04	16	1,00E+00	3	50	42	23	66	9,518	10,227	9,667	9,793	10,226	1,64	1,47	putative protein	_
AL751204	RS06C03	7,73E-10	53	6,01E-02	5	1,00E+00	1	64	59	53	57	8,790	10,087	9,972	9,021	9,380	2,46		T29M8.12 protein	Métabolisme
BX251681	Ge053H04	5,52E-06	24	5,93E-03	11	2,98E-03	21	61	56	24	58	10,567	11,505	10,726	10,726	11,424	1,92		Germin(Oxalate oxidase)-like protein, partial (24%)	Croissance cellulaire, division
BX678613	RS12A03	2,30E-08	34	2,01E-03	11	2,81E-02	11	57	56	34	76	12,730	13,467	13,000	12,755	13,478	1,67		Gamma-thionin homolog precursor	Réponse au Stress
BX681156	RS54E10	2,86E-09	56	9,30E-01	0	1,00E+00	0	60	56	56	60	13,358	12,482	12,929	13,020	12,925	1,83		putative protein	_
BX681782	RS64F10	1,65E-09	54	8,62E-01	0	1,00E+00	1	64	55	54	60	10,758	12,343	11,418	11,412	11,621	3,00		putative protein	_
BX681161	RS54E06	3,49E-09	53	5,43E-01	1	1,00E+00	0	61	54	53	60	11,878	10,945	11,372	11,527	11,457	1,91		putative protein	_
CR394036	RN72F06	2,04E-12	53	6,36E-01	0	1,00E+00	1	66	54	53	60	8,693	10,106	9,523	9,477	9,319	2,66		Agglutinin alpha chain (MPA)	Organisation cellulaire
AL749979	AS03D06	1,03E-09	43	5,93E-03	8	1,00E+00	1	58	53	43	76	10,451	11,460	10,743	11,087	11,171	2,01		Thioredoxin	Energie
BX679792	RS34A05	1,99E-09	48	1,33E-01	4	1,00E+00	1	53	52	48	76	10,521	11,420	11,217	10,970	10,812	1,86		PrMC3	Non-classé
BX681067	RS52G04	2,08E-08	47	9,66E-02	5	1,00E+00	0	59	52	47	52	9,595	10,461	10,109	10,033	9,878	1,82		Embryogenesis-associated protein EMB7	Non-classé
BX677537	RN40H04	3,00E-09	51	7,88E-01	0	1,00E+00	0	63	52	51	60	12,729	12,030	12,402	12,427	12,391	1,62		Receptor protein kinase precursor	Communication/Transduction du signal
BX250268	Ge034E09	1,74E-12	46	4,05E-02	4	1,00E+00	1	71	51	46	50	9,907	8,220	8,852	8,785	9,349	3,22		putative protein	_
BX249833	Ge029C02	7,79E-08	48	4,00E-01	2	1,00E+00	1	52	51	48	60	10,665	11,529	11,149	11,125	10,961	1,82		Ras-like small monomeric GTP-binding protein, partial (65%)	Communication/Transduction du signal
CR393785	RN65C08	1,91E-09	48	1,75E-01	3	1,00E+00	0	55	51	48	75	12,000	11,080	11,368	11,602	11,682	1,89		Hypothetical 8.7 kDa protein	Non-classé
CR392543	RN33H08	4,81E-09	40	8,95E-03	8	1,00E+00	2	55	51	40	76	14,658	13,870	14,028	14,243	14,475	1,73		25S rRNA	Synthèse de protéines
BX679434	RS25D07	1,05E-10	42	1,54E-03	9	1,00E+00	0	68	51	42	76	11,581	13,177	12,798	12,278	12,035	3,02		putative protein	_
BG039488	NXSI_99_F10	1,65E-09	39	4,94E-02	5	1,89E-01	6	61	50	39	63	9,356	10,368	10,013	10,100	9,620	2,02		S-adenosyl-Met decarboxylase	Métabolisme
BX254984	Ge108G07	1,76E-09	47	2,25E-01	3	1,00E+00	0	53	50	47	76	14,201	13,189	13,536	13,550	13,885	2,02		putative protein	_
non soumis	RN28G08	3,53E-10	47	8,63E-01	0	1,00E+00	3	59	50	47	70	9,829	11,244	10,795	10,072	10,659	2,67		putative protein	_
BX681370	RS57H02	2,12E-07	28	1,49E-03	11	4,56E-02	10	56	49	28	40	8,836	8,315	8,734	8,503	8,414	1,44		Unknown	Non-classé
BX249490	Ge024F08	1,78E-09	48	6,01E-01	1	1,00E+00	0	51	49	48	76	14,140	13,441	13,680	13,815	13,842	1,62		putative protein	_
BX251839	Ge056B11	9,23E-12	48	8,78E-01	0	1,00E+00	0	66	49	48	64	9,550	10,898	10,395	9,962	10,306	2,55		Leucine-rich repeat-containing protein 2	Réponse au Stress
CR392910	RN49H03	5,22E-10	42	1,87E-02	6	1,00E+00	1	62	49	42	76	10,941	12,383	12,048	11,487	11,465	2,72		Germin-like protein	Croissance cellulaire, division
BG039825	NXSI_104_E11	9,14E-10	36	4,05E-02	4	4,14E-02	8	66	48	36	64	9,515	10,558	10,158	10,332	9,790	2,06		Ornitine transaminase	Métabolisme
BX249875	Ge029G04	2,50E-08	38	4,79E-03	10	1,00E+00	0	56	48	38	73	10,602	11,495	11,263	11,249	10,705	1,86		Glucan endo-1,3-beta-D-glucosidase (EC 3.2.1.39)	Organisation cellulaire
BX681204	RS55C05	1,85E-09	45	1,65E-01	3	1,00E+00	0	55	48	45	76	13,755	13,080	13,284	13,423	13,508	1,60		Cysteine synthase (EC 2.5.1.47)	Métabolisme
BE643924	NXCI_50_B7	5,72E-10	37	4,56E-02	4	1,17E-01	6	66	48	37	70	9,723	10,687	10,312	10,475	9,986	1,95		S-adenosyl-Met synthase	Métabolisme
CR392874	RN49D09	4,54E-09	45	2,73E-01	2	1,00E+00	0	50	48	45	76	12,628	11,750	12,019	12,185	12,310	1,84		Zinc finger protein 364	Transcription
BE761959	NXCI_75_C7	5,22E-10	37	9,68E-02	3	7,96E-02	7	66	47	37	62	9,391	10,432	10,019	10,157	9,713	2,06		Glutamate synthase	Métabolisme
BX677392	RN38C03	3,19E-06	23	1,89E-03	12	2,12E-02	13	58	47	23	39	9,175	8,098	8,754	9,137	8,102	2,11		putative protein	_
AL750008	AS03G06	3,53E-10	43	7,02E-01	0	9,11E-01	3	62	47	43	62	9,472	11,005	10,484	9,920	10,290	2,89		Carbonate dehydratase (EC 4.2.1.1)	Métabolisme
BX249968	Ge030H03	2,60E-09	46	5,69E-01	1	1,00E+00	0	52	47	46	76	12,597	11,563	11,947	12,053	12,167	2,05		Latex-abundant protein	Non-classé
BX252607	Ge067G04	1,53E-07	32	1,42E-03	13	1,00E+00	1	51	46	32	60	9,241	9,935	9,864	9,569	9,354	1,62		NIP3, partial (75%)	Réponse au Stress
BX254875	Ge106E11	1,87E-09	43	1,38E-01	3	1,00E+00	0	56	46	43	76	10,895	11,730	11,510	11,183	11,243	1,78		Probable aquaporin, complete	Transport
BX254006	Ge092B02	2,55E-10	45	2,92E-01	2	1,00E+00	0	62	46	45	76	14,420	13,674	13,979	13,917	14,149	1,68		Outer membrane protein (Fragment), partial (20%)	Non-classé
BX249881	Ge029G11	3,35E-07	45	5,40E-01	1	1,00E+00	0	48	46	45	60	11,253	12,038	11,739	11,480	11,591	1,72		ATP synthase gamma chain	Energie
BX678464	RS09E08	9,14E-10	44	3,51E-01	2	1,00E+00	0	58	46	44	76	14,528	13,736	14,029	14,086	14,215	1,73		60S ribosomal protein L41, mitochondrial precursor	Synthèse de protéines
BX255662	Ge008C06	2,61E-07	33	1,18E-01	4	2,09E-01	8	46	46	33	76	10,506	11,059	10,914	10,883	10,617	1,47		Glycine-rich protein homolog	Non-classé
CR392503	RN33D02	8,75E-09	43	1,95E-01	3	1,00E+00	0	50	46	43	76	12,344	11,547	11,747	12,057	12,030	1,74		putative protein	_
BX681362	RS57F06	3,32E-09	45	8,24E-01	0	1,00E+00	0	51	46	45	76	13,445	12,666	12,961	13,115	13,063	1,72		DnaJ protein homolog	Réponse au Stress
AL749902	AS02D12	5,97E-09	43	9,17E-01	0	1,00E+00	3	52	46	43	63	9,635	10,892	10,510	9,789	10,393	2,39		putative protein	_
BX249546	Ge025C11	2,15E-08	41	5,16E-01	1	1,00E+00	4	49	46	41	69	9,955	10,811	10,545	10,300	10,325	1,81		Calmodulin	Communication/Transduction du signal
BX249900	Ge030A10	3,46E-09	44	4,25E-01	1	1,00E+00	0	53	46	44	76	15,169	14,575	14,779	14,859	14,937	1,51		putative protein	_
BX681222	RS55E07	1,45E-08	38	4,38E-02	6	1,00E+00	1	53	45	38	76	10,877	11,830	11,580	11,392	11,154	1,94		putative protein	_
BX252065	Ge059C02	1,01E-07	37	3,54E-02	7	1,00E+00	1	46	45	37	76	14,390	15,050	14,510	14,997	14,814	1,58		17.9 kDa class II heat shock protein	Réponse au Stress
BX251535	Ge050H05	3,51E-07	32	4,85E-03	11	1,00E+00	1	47	45	32	76	11,319	12,141	12,148	11,453	11,521	1,77		60S ribosomal protein L39	Synthèse de protéines
AL751090	RS05A04	7,16E-07	27	4,94E-02	6	2,81E-02	12	57	45	27	48	9,365	8,806	9,187	9,180	8,938	1,47		Protein kinase (EC 2.7.1.-)	Non-classé
BX250876	Ge042D06	1,45E-08	44	9,59E-01	0	1,00E+00	1	46	45	44	76	13,976	12,991	13,434	13,448	13,490	1,98		putative protein	_
BX252614	Ge067H06	2,87E-08	42	2,57E-01	3	1,00E+00	0	48	45	42	64	9,564	10,264	10,140	9,697	9,887	1,62		Histone H2A variant	Organisation cellulaire
BX253820	Ge089A07	8,95E-12	41	2,91E-02	4	1,00E+00	0	72	45	41	34	8,809	7,961	8,295	8,179	8,552	1,80		Hypothetical 10.4 kDa protein (Expressed protein)	Non-classé
BX680441	RS43D06	4,81E-09	44	7,95E-01	0	1,00E+00	0	55	44	44	26	8,477	7,833	8,091	8,151	8,172	1,56		putative protein	_
BX678411	RS08G05	1,26E-08	43	5,47E-01	1	1,00E+00	0	48	44	43	76	12,435	11,685	11,946	12,062	12,127	1,68		RNA-binding protein KIAA0682	Non-classé
BX682765	Ge121D06	1,34E-07	30	1,36E-03	12	1,00E+00	2	55	44	30	61	9,293	10,369	10,095	10,205	9,384	2,11		Granule-bound glycogen [starch] synthase	Métabolisme
BX681296	RS56F09	1,05E-10	43	3,72E-01	1	1,00E+00	0	65	44	43	76	14,480	13,422	13,847	13,869	14,041	2,08		[Pyruvate dehydrogenase (Lipoamide)]-phosphatase (EC 3.1.3	Energie
AW870208	NXNV_125_E12	2,84E-08	39	7,48E-02	5	1,00E+00	0	50	44	39	76	12,244	13,201	12,983	12,644	12,562	1,94		Aspartate transaminase	Métabolisme
AL750449	RN02F06	2,02E-07	35	6,12E-02	6	1,00E+00	4	46	44	35	76	12,170	12,999	12,361	12,888	12,689	1,78		Peptidylprolyl isomerase (EC 5.2.1.8)	Devenir des protéines
AL750105	AS05G10	4,59E-08	35	1,73E-02	7	1,00E+00	2	53	44	35	53	9,630	8,897	9,130	8,981	9,536	1,66		Unknown	Non-classé
BX679331	RS23D01	3,35E-07	32	5,37E-03	11	1,00E+00	0	49	44	32	46	9,305	8,680	8,782	8,898	9,262	1,54		putative protein	_
BX681223	RS55F11	1,50E-08	42	5,82E-01	1	1,00E+00	1	48	44	42	76	14,324	13,450	13,768	13,856	13,971	1,83		putative protein	_
AL750365	RN01F10	3,72E-08	41	7,89E-01	0	1,00E+00	2	46	44	41	76	10,505	11,193	10,845	11,147	10,712	1,61		Peroxidase precursor (EC 1.11.1.7)	Réponse au Stress
CR393564	RN61G10	4,13E-09	41	1,97E-01	3	1,00E+00	0	55	44	41	76	14,751	13,939	14,228	14,252	14,471	1,76		Sorcin	Communication/Transduction du signal
BX681617	RS62A06	6,06E-09	42	3,52E-01	2	1,00E+00	0	54	44	42	76	15,127	14,356	14,631	14,683	14,816	1,71		putative protein	_
CR392961	RN50F01	1,22E-08	42	5,02E-01	1	1,00E+00	1	58	44	42	68	11,026	10,276	10,778	10,601	10,611	1,68		Dehydration-responsive protein RD22 precursor	Réponse au Stress
BX681552	RS61B05	6,11E-08	39	6,57E-01	1	1,00E+00	4	46	44	39	74	10,252	11,189	10,516	11,381	10,582	1,91		putative protein	_
AL750713	RN06E03	1,65E-09	43	9,00E-01	0	1,00E+00	0	57	43	43	21	8,202	7,623	7,908	7,869	7,904	1,49		L-ascorbate oxidase precursor (EC 1.10.3.3)	Métabolisme
BX682324	Ge115A12	5,26E-10	38	2,78E-02	5	1,00E+00	0	65	43	38	76	10,668	11,931	11,546	11,322	11,111	2,40		Abscisic stress ripening protein	Réponse au Stress
BX249494	Ge024F11	6,55E-09	41	2,31E-01	2	1,00E+00	0	53	43	41	76	10,718	11,756	11,433	11,192	11,130	2,05		Profilin	Traffic intracellulaire
BX682063	RS69E03	1,74E-12	42	2,75E-01	1	1,00E+00	0	74	43	42	76	13,808	12,526	13,073	13,003	13,281	2,43		Phenylalanine ammonia-lyase (EC 4.3.1.5)	Métabolisme
BX680952	RS51C03	4,48E-06	30	1,05E-02	12	1,00E+00	1	44	43	30	69	10,042	11,067	10,827	11,369	9,864	2,04		putative protein	_
BX255613	Ge007G03	1,15E-06	40	8,10E-01	0	1,00E+00	2	46	43	40	60	10,992	11,807	11,449	11,227	11,387	1,76		putative protein	_
non soumis	RN28G10	1,87E-09	43	7,51E-01	0	1,00E+00	0	57	43	43	76	13,033	12,201	12,566	12,527	12,670	1,78		putative protein	_

BC065529	Ha_PP758	1,15E-07	34	2,69E-02	7	1,00E+00	1	49	43	34	_	8,879	8,131	8,384	8,186	8,786	1,68	_	_
BX681871	RS66A11	2,95E-07	36	1,26E-01	5	1,00E+00	2	46	43	36	76	12,957	13,537	13,503	13,026	13,214	1,50	putative protein	_
BX252796	Ge070H02	5,83E-09	40	2,25E-01	2	1,00E+00	0	54	43	40	76	13,774	13,087	13,331	13,368	13,529	1,61	SOUL-like protein, partial (10%)	Non-classé
BX677642	RN42E06	6,06E-09	42	8,25E-01	0	1,00E+00	0	61	43	42	76	11,716	10,730	11,168	11,250	11,221	1,98	putative protein	_
BX252809	Ge071A10	1,09E-07	39	3,86E-01	2	1,00E+00	2	45	43	39	76	10,688	11,631	11,358	11,184	11,039	1,92	putative protein	_
BF049732	NXCl_106_D10	1,08E-07	38	3,28E-01	2	1,00E+00	2	48	43	38	76	13,160	14,066	13,741	13,703	13,413	1,87	Fumarase	Métabolisme
CR392181	RN27G03	1,69E-07	35	8,79E-02	5	1,00E+00	2	46	43	35	76	12,465	13,007	12,940	12,598	12,645	1,46	Annexin	Organisation cellulaire
BX249996	Ge031B08	5,54E-09	41	5,51E-01	1	1,00E+00	0	53	43	41	76	12,273	11,493	11,806	11,805	11,959	1,72	At3g03773 {Arabidopsis thaliana;} , partial (71%)	Non-classé
CR393619	RN62F01	1,87E-09	42	5,98E-01	1	1,00E+00	0	57	43	42	76	13,903	13,129	13,428	13,523	13,551	1,71	Peptide transporter PTR2-B	Transport
CR392504	RN33D03	4,96E-08	39	2,03E-01	3	1,00E+00	1	48	42	39	68	9,867	10,690	10,489	10,150	10,195	1,77	putative protein	_
non soumis	RN32C07	1,70E-08	42	9,00E-01	0	1,00E+00	1	49	42	42	76	13,659	12,577	13,141	12,914	13,148	2,12	putative protein	_
BX682787	Ge121G09	1,48E-07	37	1,75E-01	4	1,00E+00	1	44	42	37	76	11,033	11,745	11,630	11,190	11,310	1,64	60S ribosomal protein L31	Synthèse de protéines
BC065529	Ha_PP758_b	6,11E-08	34	2,53E-02	7	1,00E+00	1	52	42	34	_	10,362	9,118	9,516	9,370	10,128	2,37	_	_
CR392490	RN33B05	3,97E-09	36	2,60E-02	6	1,00E+00	0	61	42	36	76	15,149	14,366	14,607	14,669	14,919	1,72	putative protein	_
BX681612	RS61H11	3,38E-08	42	9,23E-01	0	1,00E+00	0	45	42	42	76	14,527	13,788	14,090	14,156	14,177	1,67	dUTP pyrophosphatase (EC 3.6.1.23)	Métabolisme
BX251778	Ge055C09	1,37E-07	38	1,46E-01	4	1,00E+00	0	47	42	38	32	8,487	8,011	8,148	8,218	8,360	1,39	Unknown	Non-classé
BX680100	RS39B01	5,09E-08	40	3,41E-01	2	1,00E+00	0	46	42	40	76	13,758	13,073	13,307	13,343	13,530	1,61	Nodulin-like protein (Putative nodulin protein)	Organisation cellulaire
BX255434	Ge005G04	7,33E-08	38	6,88E-01	1	1,00E+00	3	46	42	38	76	11,162	12,030	11,584	11,627	11,651	1,83	UDP-glucose 4-epimerase (EC 5.1.3.2)	Métabolisme
BX252271	Ge062B09	2,04E-08	41	5,96E-01	1	1,00E+00	0	48	42	41	76	10,808	11,682	11,253	11,170	11,347	1,83	Alcohol dehydrogenase class III (Glutathione-dependent forma	Métabolisme
BX681431	RS59B06	6,47E-08	38	1,41E-01	4	1,00E+00	1	48	42	38	76	13,173	12,257	12,580	12,474	12,949	1,89	Pathogenesis-related protein PR-4B precursor	Réponse au Stress
CR392730	RN37C05	1,60E-06	41	6,76E-01	1	1,00E+00	0	43	42	41	61	14,414	13,839	14,061	14,247	14,164	1,49	putative protein	_
BX250861	Ge042C04	2,07E-06	37	1,32E-01	5	1,00E+00	0	48	42	37	48	9,405	10,210	10,050	9,344	9,788	1,75	OVCA2=CANDIDATE tumor suppressor, partial (9%)	Non-classé
BX250267	Ge034E07	3,51E-08	39	2,79E-01	2	1,00E+00	0	49	42	39	76	10,627	11,509	11,076	10,841	11,268	1,84	Ubiquitin-conjugating enzyme E2-17 kDa (EC 6.3.2.19)	Organisation cellulaire
BX252735	Ge069H08	2,86E-06	28	2,20E-03	14	1,00E+00	0	43	42	28	76	11,000	11,510	11,484	11,312	11,009	1,42	putative protein	_
BX254794	Ge104G10	1,45E-06	30	1,53E-02	9	1,00E+00	3	45	42	30	69	9,881	10,896	10,298	9,964	10,850	2,02	putative protein	_
AL751154	RS05F10	1,94E-07	29	1,92E-03	11	1,00E+00	2	57	42	29	76	14,209	13,507	13,660	13,798	14,085	1,63	Serine/threonine specific protein phosphatase 2C (EC 3.1.3.16	Métabolisme
BX678807	RS15D05	3,82E-07	33	2,29E-02	8	1,00E+00	0	45	42	33	68	9,805	10,669	10,565	10,118	10,025	1,82	Carboxypeptidase C precursor (EC 3.4.16.5)	Devenir des protéines
AL750839	RN07H06	4,81E-09	41	7,37E-01	0	1,00E+00	0	56	42	41	65	10,663	9,415	9,967	9,808	10,132	2,37	putative protein	_
BX249627	Ge026D05	4,18E-06	29	5,97E-03	13	1,00E+00	0	43	42	29	64	9,819	10,462	10,528	9,855	9,920	1,56	Tubulin beta chain	Organisation cellulaire
BX666039	RN08H11	3,94E-09	40	5,11E-01	1	1,00E+00	0	56	41	40	76	13,086	12,286	12,595	12,658	12,744	1,74	Late embryogenesis abundant protein	Réponse au Stress
AL750301	AS07H12	2,61E-08	38	1,68E-01	3	1,00E+00	0	51	41	38	76	12,254	11,416	11,673	11,823	11,955	1,79	Unknown	Non-classé
CR394037	RN72F07	7,98E-09	40	6,71E-01	1	1,00E+00	1	53	41	40	76	13,298	12,418	12,761	12,856	12,900	1,84	Coatomer beta' subunit (Beta'-coat protein)	Traffic intracellulaire
BX248870	Ge002A10	8,49E-06	25	1,15E-03	16	1,00E+00	1	42	41	25	74	10,578	11,154	11,229	10,783	10,572	1,49	60S acidic ribosomal protein P1	Synthèse de protéines
BE123499	NXNV_148_H6	9,54E-08	38	6,60E-01	1	1,00E+00	2	45	41	38	76	13,624	14,547	14,159	13,807	14,251	1,90	Cellulose synthase	Métabolisme
BX681114	RS53G02	2,74E-06	28	4,64E-03	12	1,00E+00	1	43	41	28	76	12,447	13,161	12,737	12,304	13,248	1,64	26S proteasome non-ATPase regulatory subunit 2	Devenir des protéines
BX681087	RS53B06	1,19E-06	29	1,81E-02	9	1,00E+00	3	45	41	29	76	14,545	14,061	14,414	14,302	14,150	1,40	SUR2 protein (Syringomycin response protein 2)	Métabolisme
BX681972	RS67E05	2,58E-10	41	7,56E-01	0	1,00E+00	0	65	41	41	76	13,728	12,658	13,153	13,048	13,253	2,10	putative protein	_
AL749580	AN01H05	2,37E-06	26	6,31E-03	11	1,00E+00	4	51	41	26	43	8,577	9,373	8,827	8,844	9,350	1,74	putative protein	_
BX679419	RS25A08	8,28E-08	40	5,73E-01	1	1,00E+00	0	45	41	40	76	13,879	13,116	13,380	13,466	13,560	1,70	putative protein	_
AL749972	AS03C11	6,90E-09	41	9,63E-01	0	1,00E+00	0	53	41	41	76	11,215	12,233	11,832	11,569	11,781	2,03	Endopeptidase Clp subunit (EC 3.4.21.92)	Devenir des protéines
BX678995	RS18B05	3,32E-06	28	7,21E-03	11	1,00E+00	2	42	41	28	76	11,188	11,803	11,440	11,059	11,879	1,53	putative protein	_
CR393841	RN65G12	1,77E-06	40	7,95E-01	0	1,00E+00	0	44	41	40	27	8,313	7,914	8,089	8,180	8,131	1,32	Unknown	Non-classé
CR393672	RN63C06	6,20E-08	40	9,55E-01	0	1,00E+00	0	44	41	40	76	14,824	13,985	14,324	14,447	14,403	1,79	putative protein	_
BX678890	RS16F01	1,75E-06	29	6,06E-03	12	1,00E+00	0	43	41	29	69	9,826	10,558	10,528	10,079	9,958	1,66	Cholinephosphate cytidylyltransferase B (EC 2.7.7.15)	Métabolisme
BX255565	Ge007C03	1,51E-06	28	3,00E-02	7	6,95E-01	5	46	41	28	76	12,449	13,368	13,327	12,502	12,788	1,89	Dephospho-CoA kinase (Dephosphocoenzyme A kinase) , par	Métabolisme
BX681588	RS61F06	4,54E-09	40	9,02E-01	0	1,00E+00	0	56	41	40	76	13,320	12,504	12,851	12,922	12,911	1,76	putative protein	_
CR394252	RN75G02	1,09E-07	40	9,09E-01	0	1,00E+00	0	41	41	40	76	14,606	13,896	14,214	14,229	14,254	1,64	putative protein	_
BX254293	Ge096C10	1,73E-06	28	1,01E-02	10	1,00E+00	3	46	41	28	50	8,967	9,584	9,080	9,413	9,440	1,53	Peptide methionine sulfoxide reductase (EC 1.8.4.6)	Réponse au Stress
BX255712	Ge008H07	9,83E-07	31	2,03E-01	3	5,46E-01	6	43	41	31	76	11,005	11,645	11,156	11,625	11,359	1,56	Thioredoxin H-type	Energie
CR393627	RN62F09	3,03E-07	39	9,87E-01	0	1,00E+00	2	41	41	39	55	9,688	9,005	9,337	9,279	9,403	1,60	putative protein	_
BX255626	Ge007H04	5,13E-09	39	7,60E-01	0	1,00E+00	1	56	41	39	76	10,679	11,611	11,242	11,100	11,136	1,91	Small zinc finger-like protein, complete	Transcription
CR392786	RN48A05	5,77E-09	40	8,10E-01	0	1,00E+00	0	55	41	40	76	12,119	11,112	11,557	11,522	11,665	2,01	Receptor protein kinase CLAVATA1 precursor (EC 2.7.1.-)	Communication/Transduction du signal
BX682978	Ge126D02	3,88E-07	35	7,97E-02	6	1,00E+00	0	46	40	35	74	8,353	9,041	8,856	8,694	8,528	1,61	putative protein	_
CR393730	RN64C07	1,90E-08	38	3,23E-01	2	1,00E+00	0	52	40	38	58	10,025	9,043	9,374	9,549	9,627	1,98	putative protein	_
CR392715	RN37A08	8,58E-07	32	2,66E-02	8	1,00E+00	0	42	40	32	76	14,446	13,972	13,990	14,245	14,388	1,39	Chitinase precursor (EC 3.2.1.14)	Réponse au Stress
BX252565	Ge066E10	3,06E-07	37	2,61E-01	3	1,00E+00	1	53	40	37	48	8,914	9,619	9,380	9,303	9,191	1,63	putative protein	_
BX677507	RN40E01	1,09E-07	38	9,28E-01	0	1,00E+00	3	45	40	38	72	10,016	10,792	10,410	10,599	10,327	1,71	Nuclear scaffold-like protein p76	Croissance cellulaire, division
BX681068	RS52G06	3,23E-06	26	1,93E-03	14	1,00E+00	0	45	40	26	74	10,561	11,285	11,245	10,897	10,646	1,65	Cysteine proteinase RD21A precursor (EC 3.4.22.-)	Devenir des protéines
CR392486	RN33A12	3,46E-09	40	6,91E-01	0	1,00E+00	0	57	40	40	76	13,026	12,020	12,459	12,410	12,592	2,01	ATP-dependent Clp protease proteolytic subunit (EC 3.4.21.92	Non-classé
BX250726	Ge040E06	1,62E-08	33	1,31E-02	7	1,00E+00	0	59	40	33	65	9,514	10,696	9,769	10,575	10,249	2,27	Unknown	Non-classé
BX253877	Ge089H04	2,54E-07	36	3,19E-01	2	1,00E+00	2	43	40	36	17	8,013	7,605	7,747	7,685	7,924	1,33	DNA repair protein RAD16, partial (6%)	Réponse au Stress
BX677733	RN44A03	3,00E-09	39	4,22E-01	1	1,00E+00	0	59	40	39	76	13,399	12,361	12,811	12,677	13,011	2,05	Unknown	Non-classé
CR393500	RN61A02	3,59E-06	28	5,86E-03	12	1,00E+00	1	42	40	28	76	11,028	11,878	11,354	10,927	11,958	1,80	putative protein	_
BX252371	Ge063E06	2,84E-07	37	3,39E-01	2	1,00E+00	1	41	40	37	75	10,739	11,520	11,435	10,803	11,067	1,72	60S ribosomal protein L31	Synthèse de protéines
BX677894	RN46B11	1,89E-07	38	8,57E-01	0	1,00E+00	2	42	40	38	53	9,388	8,733	8,965	9,082	9,102	1,57	Alcohol dehydrogenase (EC 1.1.1.1)	Métabolisme
BX680013	RS38B06	5,92E-06	26	9,22E-03	11	1,00E+00	3	41	40	26	76	10,589	11,241	10,844	10,465	11,328	1,57	GbJAAD26879.1	Synthèse de protéines
BX249189	Ge021A08	1,65E-07	39	8,52E-01	0	1,00E+00	1	41	40	39	76	11,204	11,970	11,622	11,783	11,478	1,70	putative protein	_
BX254449	Ge098G01	6,47E-07	33	5,64E-02	6	1,00E+00	1	42	40	33	66	9,683	10,370	9,931	9,887	10,267	1,61	Phospholipase D alpha 1 (PLD alpha 1) (Choline phosphatase	Communication/Transduction du signal
BX677476	RN40A05	1,69E-08	39	6,48E-01	1	1,00E+00	0	52	40	39	76	13,651	12,599	13,031	13,038	13,207	2,07	putative protein	_
BX679990	RS37G05	3,93E-07	39	7,18E-01	1	1,00E+00	0	44	40	39	70	13,348	12,475	12,847	12,776	12,993	1,83	Hypothetical protein (AT5g01750/T20L15_20)	Non-classé
BX680021	RS38C03	2,32E-06	29	1,32E-02	10	1,00E+00	1	42	40	29	76	10,975	11,755	11,815	10,889	11,245	1,72	Dihydroflavonol-4-reductase (EC 1.1.1.219) (DFR)	Métabolisme
BX251886	Ge056H01	2,20E-07	37	5,94E-01	1	1,00E+00	2	42	40	37	71	10,157	11,074	10,767	10,608	10,524	1,89	Anther specific protein / Pollen specific & Anther specific LAT5	Organisation cellulaire
BX249794	Ge027G10	6,02E-07	39	7,68E-01	0	1,00E+00	0	42	40	39	76	11,519	10,799	11,097	11,090	11,296	1,65	Beta-ketoacyl-ACP synthetase I-2, partial (91%)	Métabolisme
BX680909	RS50F11	3,68E-07	37	9,17E-01	0	1,00E+00	2	40	40	37	74	11,941	12,637	12,297	12,512	12,179	1,62	Pinus sylvestris class III peroxidase	Réponse au Stress
BX249359	Ge023B03	1,23E-07	39	9,28E-01	0	1,00E+00	1	43	40	39	76	14,154	12,909	13,456	13,428	13,591	2,37	putative protein	_
BX252657	Ge068F09	4,03E-07	35	2,10E-01	3	1,00E+00	1	43	40	35	72	10,093	10,836	10,691	10,286	10,384	1,67	60S ribosomal protein L35A	Synthèse de protéines
BX680327	RS42A07	5,72E-10	38	2,78E-01	2	1,00E+00	0	65	40	38	76	12,624	11,586	12,008	11,982	12,213	2,05	Hypothetical 85.6 kDa protein	Non-classé
BX252434	Ge064E04	2,03E-07	38	9,57E-01	0	1,00E+00	1	40	40	38	76	12,841	13,836	13,530	12,999	13,422	1,99	Coatomer alpha subunit	Traffic intracellulaire
AL750551	RN03G06	1,35E-07	37	2,24E-01	3	1,00E+00	0	45	40	37	76	12,403	11,765	11,938	12,129	12,166	1,56	Hypothetical 168.1 kDa protein	Non-classé
BX677769	RN44E06	3,72E-08	39	6,21E-01	1	1,00E+00	0	49	40	39	76	13,253	12,396	12,737	12,758	12,900	1,81	Enoyl-CoA hydratase (EC 4.2.1.17)	Métabolisme
AL750816	RN07F07	3,68E-08	38	4,03E-01	2	1,00E+00	0	50	40	38	76	10,834	10,140	10,435	10,282	10,625	1,62	putative protein	_
BX682459	Ge116H08	1,96E-06	26	1,14E-03	14	1,00E+00	0	51	40	26	70	9,786	10,554	9,936	10,120	10,470	1,70	Ubiquitin-conjugating enzyme (EC 6.3.2.19)	Organisation cellulaire
BX249297	Ge022D04	2,94E-07	39	9,88E-01	0	1,00E+00	0	41	39	39	76	11,617	10,873	11,198	11,217	11,259	1,67	Unknown	Non-classé
BX681158	RS54E01	9,00E-09	39	9,57E-01	0	1,00E+00	0	54	39	39	76	13,428	12,538	12,961	12,884	13,007	1,85	Portal protein (GPB) (Minor capsid protein B)	Traffic intracellulaire
BX252055	Ge059B02	1,35E-07	39	9,17E-01	0	1,00E+00	1	42	39	39	76	11,530	12,434	12,097	11,732	12,085	1,87	subtilisin-like serine proteinase - Arabidopsis thaliana {Arabido	Devenir des protéines
BX676834	RN12D04	1,08E-07	39	7,26E-01	1	1,00E+00	0	48	39	39	27	8,258	7,758	8,012	8,007	7,970	1,41	T17B22.3 protein	Non-classé
BX678309	RN69D12	8,93E-08	39	9,03E-01	0	1,00E+00	0	45	39	39	76	14,461	13,857	14,093	14,200	14,160	1,52	putative protein	_
CR394145	RN74D07	1,60E-10	38	5,04E-01	1	1,00E+00	0	68	39	38	76	14,358	13,397	13,812	13,777	13,945	1,95	GTP-binding protein SAR	Traffic intracellulaire
AL751076	RS03G09	5,26E-08	38	6,78E-01	1	1,00E+00	0	48	39	38	76	13,769	13,082	13,364	13,357	13,487	1,61	Unknown	Non-classé
CR393994	RN71F04	2,68E-07	38	9,19E-01	0	1,00E+00	0	39	39	38	76	13,413	12,635	12,950	12,980	13,076	1,72	Unknown	Non-classé
CR394236	RN75E09	8,27E-08	38	8,14E-01	0	1,00E+00	0	46	38	38	76	12,875	12,186	12,509	12,373	12,606	1,61	Kinesin light chain (KLC)	Traffic intracellulaire
CR394188	RN74H08	3,47E-08	38	6,28E-01	1	1,00E+00	0	50	39	38	76	12,807	11,979	12,314	12,330	12,460	1,78	Protein kinase-like (EC 2.7.1.37)	Non-classé
BX249660	Ge026G07	2,36E-07	38	6,82E-01	1	1,00E+00	0	40	39	38	69	10,262	9,666	9,824	10,074	10,000	1,51	ENDOCHITINASE (EC 3.2.1.14)	Réponse au Stress
BX255216	Ge112H12	9,93E-08	38	7,93E-01	0	1,00E+00	0	45	39	38	76	14,089	13,334	13,644	13,668	13,759	1,69	putative protein	_
BX679128	RS20C09	2,04E-07	38	8,70E-01	0	1,00E+00	0	41	38	38	76	14,088	13,355	13,654	13,693	13,759	1,66	putative protein	_
BX255778	Ge009F07	3,35E-08	38	4,30E-01	1	1,00E+00	0	50	39	38	76	13,746	12,975	13,238	13,387	13,420	1,71	Disease resistance protein-like	Réponse au Stress
BX677882	RN46A04	3,34E-09	38	6,04E-01	1	1,00E+00	0	60	39	38	76	13,777	12,592	13,133	12,955	13,303	2,27	putative protein	_
BX254883	Ge106G06	9,55E-08	38	5,49E-01	1	1,00E+00	0	46	39	38	76	12,713	11,964	12,236	12,326	12,404	1,68	Histidine ammonia-lyase (Histidase) , partial (4%)	Métabolisme
BX250119	Ge032F08	5,53E-07	38	7,89E-01	0	1,00E+00	1	40	39	38	37	8,850	8,338	8,535	8,566	8,676	1,43	Auxin response factor 1, partial (25%)	Transcription
AL751149	RS05F05	2,46E-07	37	6,69E-01	1	1,00E+00	0	41	38	37	76	13,312	12,630	12,855	13,022	13,013	1,60	Unknown	Non-classé
AL750459	RN02G04	4,45E-08	37	3,75E-01	2	1,00E+00	0	50	39	37	76	13,995	13,356	13,582	13,652	13,744	1,56	Ubiquinol-cytochrome C reductase	Energie
BX679198	RS21C03	7,89E-08	37	8,42E-01	0	1,00E+00	0	47	37	37	76	14,286	13,459	13,802	13,856	13,899	1,77	putative protein	_
BX252966	Ge075G09	1,11E-07	37	6,46E-01	1	1,00E+00	0	45	38	37	76	13,605	12,892	13,178	13,178	13,319	1,64	At5g20110, partial (46%)	Organisation cellulaire
AL751229	RS06E04	8,14E-08	37	8,05E-01	0	1,00E+00	1	47	38	37	76	14,124	13,210	13,712	13,482	13,672	1,88	Peroxidase precursor (EC 1.11.1.7)	Réponse au Stress
AL751348	RS07G08	6,83E-07	37	7,32E-01	1	1,00E+00	0	39	38	37	29	8,436	7,981	8,225	8,183	8,177	1,37	putative protein	_
BX254145	Ge094A09	1,17E-07	37	4,72E-01	1	1,00E+00	0	45	39	37	76	11,599	12,218	12,188	11,575	11,867	1,54	PR34, partial (4%)	Non-classé
CR394501	RN50D10	1,99E-07	37	9,91E-01	0	1,00E+00	0	47	37	37	76	12,619	11,745	12,137	12,011	12,199	1,83	putative protein	_
BX251533	Ge050H03	2,00E-07	37	5,84E-01	1	1,00E+00	0	42	38	37	76	13,548	12,807	13,059	13,198	13,239	1,67	At1g11170/T28P6_16, partial (58%)	Non-classé
AL750386	RN01H09	1,22E-08	37	2,25E-01	2	1,00E+00	0	56	39	37	72	10,627	9,867	10,178	10,062	10,386	1,69	30S ribosomal protein S16	Synthèse de protéines
BX254489	Ge099C04	5,35E-07	37	9,00E-01	0	1,00E+00	1	39	38	37	74	12,464	11,026	11,765	11,495	11,827	2,71	ATP-dependent DNA helicase (EC 3.6.1.-), partial (2%)	Réponse au Stress
BX254690	Ge102G05	7,58E-07	37	5,02E-01	2	1,00E+00	0	39	39	37	48	9,441	8,953	9,103	9,093	9,280	1,40	peroxidase	Réponse au Stress
CR392669	RN36C12	3,85E-07	37	6,94E-01	1	1,00E+00	0	39	37	37	76	13,880	13,188	13,422	13,565	13,585	1,62	putative protein	_
BX678298	RN69C12	3,42E-07	37	7,14E-01	1	1,00E+00	0	40	37	37	76	13,738	13,177	13,381	13,456	13,501	1,48	Unknown	Non-classé
BX681531	RS60G10	2,04E-07	37	8,02E-01	0	1,00E+00	0	43	37	37	76	14,728	14,187	14,457	14,433	14,435	1,46	Chalcone-flavonone isomerase	Métabolisme
CR393770	RN65A10	2,32E-08	37	2,41E-01	2	1,00E+00	0	54	39	37	76	14,148	13,225	13,541	13,656	13,797	1,90	Unknown	Non-classé
CR393484	RN60G02	3,68E-07	37	5,53E-01	1	1,00E+00	0	39	38	37	76	14,013	13,402	13,607	13,686	13,786	1,53	Aldose 1-epimerase precursor (EC 5.1.3.3)	Métabolisme

AL749609	AN02H06	4,88E-09	36	9,29E-01	0	1,00E+00	0	59	36	36	63	9,438	10,460	9,810	10,508	9,811	2,03	putative protein	_
BF609706	NXSI_48_D6	2,61E-07	36	6,90E-01	1	1,00E+00	1	42	38	36	76	13,088	14,190	13,714	13,356	13,820	2,15	malate dehydrogenase	Métabolisme
CR393875	RN66E08	3,24E-07	36	9,10E-01	0	1,00E+00	0	41	36	36	76	14,425	13,593	13,954	13,948	14,049	1,78	Calcium-transporting ATPase, plasma membrane-type	Transport
BX681127	RS54B10	4,47E-07	36	8,18E-01	0	1,00E+00	0	39	36	36	76	13,192	12,586	12,783	13,007	12,884	1,52	Glutathione transferase (EC 2.5.1.18)	Métabolisme
BX680185	RS40C03	1,31E-07	36	8,81E-01	0	1,00E+00	0	46	37	36	76	14,053	13,357	13,660	13,661	13,732	1,62	Pectate lyase (EC 4.2.2.2)	Organisation cellulaire
BX679594	RS29B09	1,87E-09	36	4,57E-01	1	1,00E+00	0	64	37	36	76	12,507	11,565	11,974	11,902	12,122	1,92	putative protein	_
BX681520	RS60F10	2,76E-07	36	7,80E-01	0	1,00E+00	3	42	39	36	76	10,371	10,851	10,734	10,235	10,756	1,39	Aldehyde dehydrogenase (NAD+) (EC 1.2.1.3)	Métabolisme
CR394186	RN74H06	1,06E-08	36	1,96E-01	3	1,00E+00	0	57	38	36	76	13,038	11,719	12,231	12,154	12,587	2,49	Phytocyanin	Energie
BX681382	RS58B04	4,79E-07	36	7,79E-01	0	1,00E+00	1	39	38	36	76	14,627	13,787	14,208	14,168	14,172	1,79	Ubiquitin-protein ligase 1	Organisation cellulaire
BX250510	Ge037H06	1,73E-06	36	8,96E-01	0	1,00E+00	0	38	36	36	57	10,322	12,209	11,468	11,148	11,484	3,70	putative protein	_
BX679475	RS26E10	5,54E-09	36	1,85E-01	2	1,00E+00	0	60	38	36	76	12,578	11,563	11,924	12,027	12,184	2,02	Ribosomal protein L10	Synthèse de protéines
BX681926	RS66H02	9,14E-10	36	5,33E-01	1	1,00E+00	0	66	36	36	76	14,748	13,831	14,231	14,185	14,354	1,89	putative protein	_
BX679085	RS19F02	1,15E-07	36	6,67E-01	1	1,00E+00	1	47	37	36	76	13,018	12,012	12,452	12,333	12,628	2,01	Hypothetical 20.6 kDa protein	Non-classé
BX677297	RN20B06	7,45E-06	36	5,03E-01	2	1,00E+00	0	41	38	36	58	10,160	11,017	10,528	10,942	10,367	1,81	putative protein	_
BX255103	Ge110H06	1,99E-08	36	6,65E-01	1	1,00E+00	0	55	36	36	76	14,356	13,566	13,916	13,847	14,026	1,73	putative protein	_
BX680179	RS40B08	5,84E-07	36	7,67E-01	0	1,00E+00	1	38	37	36	76	13,059	12,105	12,496	12,420	12,711	1,94	At2g44310 Hypothetical protein	Non-classé
BX681255	RS56B07	1,34E-07	35	6,96E-01	1	1,00E+00	0	47	36	35	76	14,232	13,577	13,863	13,795	13,973	1,57	putative protein	_
BX681245	RS56A07	1,09E-07	35	2,72E-01	2	1,00E+00	0	48	38	35	76	13,310	12,590	12,817	12,958	13,035	1,65	Dynamin-like protein	Traffic intracellulaire
BX678445	RS09C02	5,53E-07	35	6,06E-01	1	1,00E+00	0	39	36	35	76	12,208	11,560	11,780	11,883	11,949	1,57	Unknown	Non-classé
CR393361	RN59B05	1,99E-09	35	1,95E-01	2	1,00E+00	0	64	38	35	76	12,793	11,853	12,232	12,176	12,449	1,92	putative protein	_
AL751122	RS05C12	5,09E-08	35	3,30E-01	2	1,00E+00	0	56	37	35	76	12,461	11,624	11,980	11,943	12,179	1,79	Glucan endo-1,3-beta-glucosidase precursor	Organisation cellulaire
BX251918	Ge057C01	2,43E-07	35	9,08E-01	0	1,00E+00	1	44	36	35	73	10,094	11,006	10,671	10,390	10,588	1,88	Ubiquitin-conjugating enzyme (EC 6.3.2.19)	Organisation cellulaire
AL751056	4CL	6,20E-08	35	7,47E-01	0	1,00E+00	0	51	36	35	76	14,604	13,937	14,201	14,279	14,292	1,59	_	_
CR392472	RN32G05	4,47E-07	35	7,80E-01	0	1,00E+00	1	40	37	35	76	12,998	12,063	12,434	12,477	12,602	1,91	Calmodulin-related protein 2	Communication/Transduction du signal
BX254888	Ge106H02	4,44E-07	35	5,02E-01	1	1,00E+00	0	40	37	35	72	9,938	10,577	10,360	10,288	10,173	1,56	Ras-related protein Rab11B	Traffic intracellulaire
BX681821	RS65D03	2,84E-08	35	1,98E-01	3	1,00E+00	0	58	38	35	76	12,544	11,751	12,077	12,041	12,297	1,73	Aldehyde dehydrogenase (NAD+) (EC 1.2.1.3)	Métabolisme
BX680275	RS41D02	7,43E-08	35	4,91E-01	1	1,00E+00	0	51	37	35	76	11,571	10,810	11,083	11,210	11,239	1,69	putative protein	_
BX248924	Ge002G04	1,18E-07	35	3,21E-01	2	1,00E+00	0	48	37	35	76	12,762	11,942	12,243	12,252	12,474	1,76	At3g23090, partial (15%)	Non-classé
BX248830	Ge001E10	7,66E-07	35	6,47E-01	1	1,00E+00	0	37	36	35	76	12,367	11,726	11,964	11,965	12,142	1,56	putative protein	_
BX680818	RS49A01	1,71E-08	35	2,96E-01	2	1,00E+00	0	57	37	35	76	12,078	11,153	11,478	11,621	11,692	1,90	putative protein	_
BX677288	RN20A02	6,79E-07	35	6,25E-01	1	1,00E+00	0	38	36	35	76	13,532	12,957	13,144	13,265	13,297	1,49	putative protein	_
BX252329	Ge062H12	1,08E-07	35	8,82E-01	0	1,00E+00	0	49	35	35	76	14,722	14,064	14,356	14,346	14,416	1,58	Proteasome subunit beta type 5 precursor (20S proteasome s	Devenir des protéines
BX677223	RN19E11	7,71E-07	35	9,16E-01	0	1,00E+00	1	37	36	35	76	14,235	13,520	13,796	13,904	13,897	1,64	putative protein	_
BX679075	RS19D08	7,55E-07	35	7,63E-01	0	1,00E+00	0	37	36	35	76	12,745	11,911	12,228	12,281	12,406	1,78	putative protein	_
_	Ha_PP756_b	5,36E-06	35	9,23E-01	0	1,00E+00	0	44	35	35	_	8,244	7,862	8,025	8,135	8,063	1,30	_	_
BX678629	RS12C06	7,73E-06	35	4,74E-01	2	1,00E+00	0	42	37	35	60	12,546	11,986	12,208	12,334	12,326	1,47	Ubiquinol-cytochrome C reductase complex 7.8 kDa protein	Energie
BX252563	Ge066E08	2,76E-07	35	4,36E-01	2	1,00E+00	1	44	37	35	74	10,111	11,010	10,600	10,935	10,342	1,86	Aspartate aminotransferase, chloroplast precursor	Métabolisme
BX681544	RS61A06	3,82E-07	35	4,66E-01	2	1,00E+00	0	42	37	35	76	12,689	11,969	12,198	12,357	12,400	1,65	Lateral organ boundaries (LOB) domain protein	Non-classé
BX680085	RS38H10	1,99E-07	35	9,90E-01	0	1,00E+00	1	48	36	35	35	8,691	8,088	8,353	8,399	8,356	1,52	putative protein	_
BX679622	RS29H07	5,64E-07	35	9,18E-01	0	1,00E+00	0	39	35	35	76	14,437	13,730	14,032	14,051	14,110	1,63	putative protein	_
BX677557	RN41B08	3,67E-07	35	9,73E-01	0	1,00E+00	1	42	35	35	76	13,842	12,900	13,362	13,244	13,394	1,92	GASA5-like protein (Sb35)	Non-classé
BX682915	Ge123D11	3,41E-07	35	3,13E-01	2	1,00E+00	0	43	37	35	76	11,172	11,950	11,687	11,651	11,425	1,71	14-3-3 protein	Communication/Transduction du signal
BX254611	Ge101C06	6,47E-07	35	8,26E-01	0	1,00E+00	0	39	35	35	76	11,867	13,196	12,814	12,084	12,611	2,51	ADP-ribosylation factor	Traffic intracellulaire
BX681841	RS65F03	1,55E-07	35	5,06E-01	1	1,00E+00	0	47	36	35	76	13,048	12,320	12,688	12,765	12,576	1,66	NADPH dehydrogenase (EC 1.6.99.1)	Energie
BX681235	RS55H11	7,86E-08	35	8,66E-01	0	1,00E+00	0	51	35	35	76	10,775	11,734	11,207	11,657	11,116	1,94	Glutathione transferase (EC 2.5.1.18)	Métabolisme
CR394150	RN74E02	6,85E-07	35	7,26E-01	1	1,00E+00	0	38	36	35	76	14,161	13,540	13,765	13,842	13,902	1,54	putative protein	_
BX680710	RS47E08	5,94E-07	35	3,72E-01	2	1,00E+00	0	39	37	35	63	9,762	9,252	9,376	9,573	9,570	1,42	putative protein	_
CR393516	RN61B10	5,16E-07	35	8,85E-01	0	1,00E+00	1	40	36	35	76	13,991	13,133	13,497	13,507	13,605	1,81	Glutamate-1-semialdehyde 2,1-aminomutase (EC 5.4.3.8)	Métabolisme
BX253454	Ge083F12	1,16E-06	35	8,89E-01	0	1,00E+00	0	49	35	35	9	7,893	7,385	7,647	7,617	7,645	1,42	Patatin	Métabolisme
BX682440	Ge116F08	3,88E-07	35	4,91E-01	1	1,00E+00	0	42	36	35	76	10,438	11,171	10,929	10,792	10,732	1,66	Nuclear transport factor 2 (NTF-2).	Traffic intracellulaire
CR393030	RN51F07	1,11E-07	35	6,90E-01	1	1,00E+00	0	49	36	35	76	13,867	13,054	13,390	13,399	13,517	1,76	putative protein	_
BX253926	Ge090G07	1,09E-06	35	9,19E-01	0	1,00E+00	0	45	35	35	14	8,019	7,544	7,775	7,782	7,770	1,39	Disease resistance-like protein (Fragment), partial (16%)	Réponse au Stress
BX680229	RS40G08	7,16E-07	35	9,04E-01	0	1,00E+00	0	43	35	35	76	14,045	12,975	13,439	13,450	13,463	2,10	Hypothetical protein (At1g76160/T23E18_10)	Non-classé
BX252278	Ge062C05	6,65E-08	35	1,63E-01	3	1,00E+00	0	52	38	35	76	12,643	11,803	12,083	12,155	12,357	1,79	putative protein	_
AL749773	AN07G08	1,37E-07	35	7,66E-01	0	1,00E+00	0	48	35	35	75	10,038	10,833	10,459	10,359	10,516	1,73	Esterase D	Métabolisme
BX251766	Ge055B05	2,75E-07	35	7,42E-01	0	1,00E+00	0	44	35	35	76	12,407	11,749	12,005	12,078	12,109	1,58	putative protein	_
BX681771	RS64D09	7,70E-07	35	3,66E-01	2	1,00E+00	0	38	37	35	76	10,250	10,733	10,728	10,127	10,508	1,40	60S ribosomal protein L34	Synthèse de protéines
BX251165	Ge046B06	7,89E-08	35	5,62E-01	1	1,00E+00	0	51	36	35	76	14,677	13,832	14,177	14,160	14,338	1,80	putative protein	_
BX250283	Ge034G10	6,93E-07	35	9,12E-01	0	1,00E+00	0	39	35	35	76	12,424	11,736	12,013	12,097	12,092	1,61	Pectate lyase, partial (95%)	Organisation cellulaire
BX251094	Ge045C02	3,55E-06	35	2,57E-01	3	1,00E+00	0	40	38	35	65	9,963	10,904	10,708	9,967	10,272	1,92	Abscisic stress ripening protein	Réponse au Stress
CR354614	Ge124A01	3,95E-07	35	8,67E-01	0	1,00E+00	1	42	36	35	76	10,764	11,552	11,259	11,073	11,162	1,73	Blr1429 protein, partial (18%)	Non-classé
CR393481	RN60F10	1,87E-09	35	7,63E-01	0	1,00E+00	0	65	35	35	76	13,947	13,271	13,624	13,547	13,585	1,60	IN2-2 protein	Non-classé
AL750508	RN03C09	6,02E-07	35	1,56E-01	4	1,00E+00	1	40	39	35	76	10,652	11,308	10,897	10,955	11,132	1,58	Histone H4	Organisation cellulaire
BX680035	RS38D06	7,22E-07	35	8,44E-01	0	1,00E+00	0	38	35	35	76	12,204	11,586	11,806	11,970	11,898	1,53	putative protein	_
BX254656	Ge102B08	8,02E-08	35	9,78E-01	0	1,00E+00	0	51	35	35	76	14,113	13,018	13,562	13,432	13,576	2,14	CG33184-PA, isoform A, partial (9%)	Non-classé
CR393742	RN64E07	3,29E-07	35	5,85E-01	1	1,00E+00	0	43	36	35	76	13,463	12,491	12,858	12,919	13,072	1,96	S-adenosyl-L-methionine:salicylic acid carboxyl methyltransfer	Métabolisme
CR393662	RN63B06	1,37E-07	34	3,44E-01	2	1,00E+00	0	48	36	34	76	12,175	13,231	12,619	12,745	12,842	2,08	Proton-exporting ATPase (EC 3.6.3.6)	Transport
AL750092	AS05F09	1,99E-07	34	3,87E-01	2	1,00E+00	1	46	38	34	76	11,685	12,602	12,156	11,922	12,338	1,89	Ubiquitin-conjugating enzyme E2-17 kDa (EC 6.3.2.19)	Organisation cellulaire
BX681360	RS57F04	3,68E-07	34	1,41E-01	4	1,00E+00	0	47	39	34	51	9,916	8,954	9,223	9,576	9,566	1,95	MERI-5 protein precursor (Xyloglucan:xyloglucosyl transferase	Croissance cellulaire, division
BX681024	RS52B04	6,03E-09	34	3,56E-01	1	1,00E+00	0	61	36	34	76	14,954	14,146	14,465	14,491	14,622	1,75	Hypothetical 60.3 kDa protein	Non-classé
CR393402	RN59F08	1,09E-07	34	7,58E-01	0	1,00E+00	0	50	35	34	76	12,624	11,800	12,151	12,152	12,256	1,77	Unknown	Non-classé
BX253993	Ge091H07	3,56E-07	34	8,20E-01	0	1,00E+00	1	43	36	34	75	10,924	12,100	11,614	11,231	11,667	2,26	Ubiquitin-conjugating enzyme E2-17 kDa (EC 6.3.2.19)	Organisation cellulaire
BX677095	RN17D11	4,44E-07	34	5,76E-01	1	1,00E+00	0	42	35	34	76	13,979	13,268	13,515	13,640	13,679	1,64	putative protein	_
BX676912	RN13G11	2,74E-06	34	6,58E-01	1	1,00E+00	4	40	39	34	75	10,280	11,077	10,644	10,743	10,670	1,74	Acyl carrier protein, chloroplast precursor	Traffic intracellulaire
BX682036	RS68G05	1,13E-08	34	2,43E-01	2	1,00E+00	0	59	36	34	76	14,943	14,206	14,492	14,486	14,668	1,67	putative protein	_
AL750794	RN07D07	5,01E-07	34	5,34E-01	1	1,00E+00	1	46	37	34	38	8,716	8,284	8,452	8,502	8,550	1,35	Aldehyde dehydrogenase (NAD+) (EC 1.2.1.3)	Métabolisme
BX681698	RS63B07	3,29E-07	34	6,71E-01	1	1,00E+00	0	44	35	34	76	14,860	14,223	14,468	14,521	14,587	1,56	putative protein	_
BX680092	RS39A02	1,09E-07	34	3,62E-01	2	1,00E+00	0	50	36	34	76	13,625	12,776	13,069	13,201	13,280	1,80	T14P8.4 protein	Métabolisme
BX253462	Ge083G09	1,09E-06	34	9,87E-01	0	1,00E+00	0	39	34	34	43	9,105	8,464	8,750	8,696	8,805	1,56	putative protein	_
CR393571	RN61H07	6,77E-07	34	9,72E-02	5	1,00E+00	0	40	39	34	76	14,124	13,515	13,615	13,824	13,993	1,53	Aldose 1-epimerase precursor (EC 5.1.3.3)	Métabolisme
BX249896	Ge030A04	3,77E-07	34	5,90E-01	1	1,00E+00	0	44	35	34	76	11,536	10,920	11,157	11,183	11,289	1,53	Unknown	Non-classé
CR392159	RN27C08	6,65E-07	34	5,66E-01	1	1,00E+00	0	40	35	34	76	13,607	12,945	13,195	13,204	13,362	1,58	40S ribosomal protein	Synthèse de protéines
BX681118	RS53G08	1,15E-07	34	5,63E-01	1	1,00E+00	0	50	35	34	76	11,503	10,703	11,027	11,029	11,174	1,74	putative protein	_
BX678083	RN56A11	1,87E-07	34	5,04E-01	1	1,00E+00	0	48	35	34	76	13,566	12,771	13,062	13,155	13,234	1,73	putative protein	_
BX252185	Ge061A06	9,49E-07	34	5,07E-01	1	1,00E+00	2	38	37	34	68	9,675	10,699	9,966	10,764	10,125	2,03	Mitochondrial oxaloacetate transport protein	Transport
BX681437	RS59C04	4,56E-08	34	2,47E-01	2	1,00E+00	0	58	36	34	57	9,808	8,973	9,244	9,354	9,469	1,78	Electron transfer flavoprotein-ubiquinone oxidoreductase	Energie
CR393769	RN65A09	4,88E-07	34	6,21E-01	1	1,00E+00	0	42	35	34	76	13,708	12,820	13,125	13,327	13,309	1,85	Unknown	Non-classé
BX254258	Ge095G04	6,88E-07	34	6,18E-01	1	1,00E+00	0	40	35	34	67	9,929	10,725	10,394	10,603	10,135	1,74	LUMINAL BINDING PROTEIN	Réponse au Stress
BX681471	RS60A05	7,09E-07	34	8,32E-01	0	1,00E+00	0	40	34	34	76	13,125	12,429	12,720	12,731	12,818	1,62	Late embryogenesis abundant protein (Fragment)	Réponse au Stress
BX681554	RS61B07	4,10E-07	34	8,06E-01	0	1,00E+00	0	43	34	34	76	13,495	12,874	13,125	13,176	13,209	1,54	NADPH:quinone reductase (EC 1.6.5.5)	Energie
BX667580	RN09G05	3,33E-07	34	7,63E-01	0	1,00E+00	0	45	34	34	76	12,712	11,896	12,238	12,238	12,358	1,76	Unknown	Non-classé
BX254586	Ge100G06	5,94E-08	34	2,61E-01	2	1,00E+00	1	54	37	34	29	8,039	8,894	8,406	8,460	8,598	1,81	putative protein	_
BX682721	Ge120F04	8,85E-08	34	5,35E-01	1	1,00E+00	2	52	36	34	74	10,084	10,869	10,399	10,639	10,511	1,72	Uridylate kinase (EC 2.7.4.3)	Métabolisme
BX678390	RS08E04	1,65E-07	34	2,48E-01	3	1,00E+00	0	49	37	34	76	11,627	10,929	11,152	11,277	11,363	1,62	putative protein	_
BX682007	RS68A10	7,77E-07	34	2,53E-01	3	1,00E+00	0	39	37	34	76	12,448	11,847	11,996	12,186	12,242	1,52	Unknown	Non-classé
non soumis	Ge024B12	3,87E-07	34	7,89E-01	0	1,00E+00	2	44	36	34	76	15,427	15,042	15,202	15,331	15,180	1,31	putative protein	_
CR394502	RN50F08	1,09E-06	34	7,84E-01	0	1,00E+00	0	37	34	34	76	11,827	11,164	11,418	11,485	11,538	1,58	putative protein	_
BX681084	RS53B10	3,42E-07	34	9,21E-01	0	1,00E+00	0	45	34	34	76	15,117	14,474	14,766	14,737	14,819	1,56	putative protein	_
AL751163	RS05G07	6,83E-07	34	7,89E-01	0	1,00E+00	0	41	34	34	76	13,720	13,076	13,338	13,363	13,439	1,56	Hypothetical 78.8 kDa protein	Non-classé
BX682419	Ge116C12	1,00E-06	34	7,88E-01	0	1,00E+00	0	38	34	34	76	12,548	11,871	12,155	12,125	12,274	1,60	Betaine-aldehyde dehydrogenase (EC 1.2.1.8)	Métabolisme
BX252717	Ge069F09	4,44E-07	34	3,87E-01	2	1,00E+00	2	45	38	34	74	10,281	11,197	10,902	10,773	10,642	1,89	B2 protein	Non-classé
BX679939	RS36F06	7,85E-05	8	1,44E-16	60	1,08E-01	5	74	74	60	71	10,284	10,864	11,345	11,174	9,403	1,49	3,84	Peroxidase precursor (EC 1.11.1.7)	Réponse au Stress
BX250781	Ge041B11	7,69E-05	8	8,59E-17	57	5,54E-03	9	76	73	57	44	8,776	9,734	8,344	8,481	10,798	1,94	5,48	Low molecular weight heat-shock protein, partial (98%)	Réponse au Stress
BX680424	RS43C10	1,92E-05	9	1,31E-17	61	9,11E-01	2	77	72	61	67	10,166	11,119	11,690	11,033	9,337	1,94	5,11	Ethylene responsive element binding factor	Transcription
CR394334	RN76H07	9,86E-05	6	9,19E-19	59	1,76E-01	3	80	69	59	55	9,056	9,822	8,550	9,645	10,313	1,70	3,39	putative protein	_
BX679759	RS33B10	2,72E-05	13	3,76E-12	50	1,00E+00	0	64	64	50	76	12,083	12,582	13,022	12,237	11,691	1,41	2,52	Adenylate kinase (EC 2.7.4.3)	Métabolisme
BX254971	Ge108E04	2,46E-05	12	6,58E-13	48	1,00E+00	0	69	60	48	70	9,598	10,470	9,234	10,346	10,754	1,83	2,87	Alkaline alpha galactosidase II, partial (84%)	Métabolisme
CR393241	RN54D04	1,48E-05	22	1,04E-06	34	1,00E+00	0	56	56	34	58	9,259	10,003	10,357	9,565	9,047	1,67	2,48	Plasma membrane intrinsic polypeptide	Transport
BX252844	Ge071F04	4,24E-05	18	2,99E-07	35	1,00E+00	0	55	54	35	76	13,249	13,893	13,164	13,434	14,169	1,56	2,01	metallothionein-like protein	Métabolisme
BX681237	RS55H04	6,09E-05	14	2,17E-08	34	5,46E-01	4	58	52	34	72	9,888	10,470	9,836	10,030	10,685	1,50	1,80	Heat shock protein 101	Réponse au Stress
BX254251	Ge095F04	6,76E-05	16	1,30E-07	35	1,00E+00	1	51	51	35	76	13,502	14,036	13,448	13,306	14,448	1,45	2,21	cytosolic aldehyde dehydrogenase {Oryza sativa (japonica cult	Métabolisme
BX681985	RS67F12	1,19E-05	20	1,08E-06	29	1,00E+00	0	51	49	29	70	10,107	10,894	11,040	10,583	9,926	1,72	2,16	Isoflavone reductase homolog (EC 1.3.1.-)	Métabolisme
BX667535	RN09C02	5,70E-05	15	3,21E-07	30	1,00E+00	1	55	46	30	75	10,221	10,957	10,150	10,608	11,100	1,67	1,93	Carbon catabolite derepressing protein kinase	Métabolisme
BX255395	Ge005C11	2,38E-05	22	5,16E-05	24	1,00E+00	0	49	47	24	76	12,556	13,305	12,692	12,410	13,525	1,68	2,17	putative protein	_
BX681290	RS56F02	7,98E-05	19	2,76E-05	26	1,00E+00	0	44	44	26	76	10,896	11,418	11,638	10,987	10,773	1,44	1,82	40S ribosomal protein S11	Synthèse de protéines
CR394206	RN75B08	2,00E-05	18	5,26E-05	20	3,47E-01	6	54	44	20	74	11,138	11,952	11,947	11,545	11,159	1,76	1,73	L-ascorbate oxidase precursor (EC 1.10.3.3)	Métabolisme

BX255377	Ge005B04	2,21E-05	20	2,95E-05	23	1,00E+00	0	48	43	23	76	11,373	12,137	11,430	11,622	12,243	1,70	1,76	Elastin precursor - human (fragment) {Homo sapiens;} , partial	Réponse au Stress
BX249687	Ge027B05	7,84E-05	18	4,27E-05	24	1,00E+00	0	43	43	24	76	13,127	13,691	13,171	12,946	14,002	1,48	2,08	FKBP-like protein (Arabidopsis thaliana genomic DNA, chromo	Transcription
BX679837	RS34F05	4,81E-05	17	1,82E-05	23	1,00E+00	0	50	41	23	76	11,941	12,723	12,852	12,165	11,939	1,72	1,88	L3 cytoplasmic ribosomal protein	Synthèse de protéines
BX677134	RN18B10	2,30E-05	17	1,49E-04	16	1,06E-02	14	55	47	17	53	8,789	9,624	8,968	8,895	9,722	1,78	1,77	putative protein	_
BX252727	Ge069G09	1,03E-05	24	9,65E-04	16	1,00E+00	1	42	41	24	76	11,454	12,260	12,337	11,724	11,484	1,75	1,81	Esterase precursor (Early nodule-specific protein homolog) (L	Métabolisme
CR392325	RN30F07	2,20E-05	22	3,04E-04	19	1,00E+00	0	43	41	22	76	11,549	12,090	12,324	11,618	11,447	1,46	1,84	40S ribosomal protein S26 (S31)	Synthèse de protéines
BX252228	Ge061F04	2,07E-05	22	3,26E-04	19	1,00E+00	0	44	41	22	62	9,717	10,329	9,875	9,563	10,520	1,53	1,94	(2R)-phospho-3-sulfolactate synthase (EC 4.4.-.-)	Non-classé
BX253595	Ge085G01	6,94E-05	19	1,60E-04	21	1,00E+00	0	42	41	21	76	13,102	13,628	13,196	12,762	13,971	1,44	2,31	Zinc finger protein HIT-10, partial (3%)	Transcription
AL751155	RS05F11	3,72E-05	21	1,21E-01	5	8,87E-04	28	59	53	28	45	9,312	8,629	9,169	8,884	8,844	1,60		putative protein	_
CR394213	RN75C03	7,54E-05	24	2,64E-03	17	1,00E+00	1	47	42	24	60	12,697	13,724	13,021	12,077	13,959	2,04		Auxin-repressed 12.5 kDa protein	Non-classé
BX682894	Ge123B11	1,41E-05	24	1,06E-03	16	1,00E+00	1	42	41	24	71	9,988	10,447	10,578	10,038	9,980	1,37		60S ribosomal protein L38	Synthèse de protéines
BX682906	Ge123D02	1,27E-05	24	1,63E-03	15	1,00E+00	1	41	40	24	74	10,252	10,780	10,897	10,258	10,311	1,44		ATP-dependent Clp protease proteolytic su	Non-classé
CR392407	RN31D12	1,10E-05	34	7,42E-01	1	1,00E+00	0	40	35	34	60	12,223	11,429	11,777	11,987	11,844	1,73		putative protein	_
AL750400	RN02B01	1,91E-05	34	7,64E-01	1	1,00E+00	0	36	35	34	60	11,394	10,839	11,016	11,290	11,153	1,47		Copper chaperone for Superoxide dismutase	Réponse au Stress
BX667562	RN09E10	9,93E-04	5	9,19E-19	70	1,66E-01	4	80	79	70	55	9,542	9,845	10,676	9,796	8,681	1,23	3,99	putative protein	_
BX678596	RS11G02	8,87E-04	8	3,61E-13	57	1,00E+00	1	68	66	57	64	10,563	11,118	12,129	11,091	9,408	1,47	6,59	Peroxidase precursor (EC 1.11.1.7)	Réponse au Stress
AL750163	AS06E02	1,72E-04	11	4,11E-11	47	1,80E-01	6	66	63	47	51	9,205	9,707	8,898	9,727	9,908	1,42	2,01	putative protein	_
BX677767	RN44E03	2,72E-04	11	5,53E-11	49	1,00E+00	1	63	60	49	72	10,501	11,125	10,370	10,088	11,759	1,54	3,18	Chitinase precursor (EC 3.2.1.14)	Réponse au Stress
BX681602	RS61G11	5,89E-04	9	2,05E-11	50	1,00E+00	1	61	60	50	74	10,953	11,529	12,110	11,323	10,299	1,49	3,51	Defensin	Réponse au Stress
BX679279	RS22D07	4,47E-04	10	1,40E-10	46	1,00E+00	2	59	58	46	74	10,543	11,269	10,332	10,113	12,084	1,65	3,92	ABA induced plasma membrane protein	Réponse au Stress
AL749547	AN01B08	3,86E-04	8	3,76E-12	45	3,85E-01	4	67	57	45	48	8,849	9,469	8,711	8,847	9,883	1,54	2,25	Outer membrane lipoprotein blc precursor	Transport
AL751045	RS03D12	6,22E-04	10	4,88E-10	43	5,56E-01	4	59	56	43	76	10,951	11,374	10,765	10,760	11,876	1,34	2,17	putative protein	_
BX678079	RN56A06	1,50E-04	15	4,63E-07	33	1,00E+00	2	54	50	33	48	8,768	9,335	8,636	9,176	9,445	1,48	1,75	putative protein	_
CR392767	RN37G07	4,72E-04	13	1,04E-07	37	1,00E+00	0	51	50	37	61	9,560	9,969	10,353	9,687	9,245	1,33	2,15	Peroxidase precursor (EC 1.11.1.7)	Réponse au Stress
AL751295	RS07C01	1,49E-04	15	5,76E-07	32	1,00E+00	2	50	48	32	76	10,571	10,954	11,187	10,812	10,302	1,30	1,85	Proline-rich protein / Cortical cell delineating protein precursor	Non-classé
BX252953	Ge075E07	1,61E-04	15	2,09E-06	30	1,00E+00	2	48	47	30	76	13,410	13,880	13,353	13,159	14,304	1,38	2,21	RING finger protein 26	Transcription
CR394461	RN79E12	4,55E-04	13	1,54E-06	30	1,00E+00	2	48	45	30	76	13,688	13,266	13,173	13,308	13,881	1,34	1,63	Ethylene-responsive protein, putative	Réponse au Stress
BX681938	RS67A12	9,08E-04	11	3,40E-07	33	1,00E+00	0	49	45	33	76	13,514	13,783	14,244	13,393	13,193	1,21	2,07	Protein kinase (EC 2.7.1.-)	Métabolisme
BX676790	RN11A09	6,46E-04	8	2,82E-09	34	1,00E+00	2	64	44	34	76	11,947	11,995	12,668	11,814	11,328	1,03	2,53	putative protein	_
BX255766	Ge009E07	2,39E-04	15	7,58E-06	28	1,00E+00	1	45	44	28	76	13,267	13,900	13,290	13,129	14,235	1,55	2,15	metallothionein	Métabolisme
BX679616	RS29G03	3,25E-04	14	3,00E-06	30	1,00E+00	0	46	44	30	52	9,149	9,547	9,667	9,366	9,023	1,32	1,56	Histone H4	Organisation cellulaire
BX679916	RS36C01	5,24E-04	13	1,37E-05	26	1,16E-01	10	50	49	26	55	9,256	9,600	9,751	9,432	9,130	1,27	1,54	Metallophosphatase	Devenir des protéines
BX681234	RS55G07	1,94E-04	15	2,17E-05	24	1,39E-01	9	49	48	24	76	10,637	11,180	10,696	10,383	11,506	1,46	2,18	putative protein	_
AL750621	RN05E06	1,02E-04	17	2,83E-05	25	1,00E+00	1	44	43	25	74	10,198	10,712	10,762	10,870	9,896	1,43	1,96	Pectate lyase (EC 4.2.2.2)	Organisation cellulaire
BX252908	Ge074G03	1,61E-04	17	7,30E-05	23	1,00E+00	2	43	42	23	76	10,462	10,975	10,510	10,309	11,239	1,43	1,91	Cathepsin B-like cysteine proteinase, partial (68%)	Devenir des protéines
BX252225	Ge061E10	4,33E-04	14	1,98E-05	26	1,00E+00	1	44	42	26	58	9,316	9,826	9,287	9,335	10,052	1,42	1,70	putative protein	_
BX251374	Ge048G11	2,78E-04	16	3,13E-05	25	1,00E+00	0	41	41	25	76	12,414	13,005	13,205	12,885	12,108	1,51	2,14	Rof1, partial (88%)	Transcription
BX677964	RN47C10	3,38E-04	15	3,98E-05	25	1,00E+00	0	42	40	25	76	12,876	13,355	13,694	12,979	12,619	1,39	2,11	Late embryogenesis abundant protein	Réponse au Stress
BX678708	RS14B03	2,31E-04	15	7,74E-04	16	9,74E-02	11	44	43	16	76	11,927	12,675	13,223	11,172	12,091	1,68	4,14	Basic blue protein (Plantacyanin)	Energie
BX666044	RN09A04	7,48E-04	12	5,18E-02	6	8,87E-04	25	48	43	25	55	9,411	8,843	9,140	9,620	8,755	1,48		putative protein	_
BX681214	RS55D09	1,87E-04	16	2,65E-01	3	2,25E-03	22	46	41	22	58	9,720	9,020	9,296	9,959	9,021	1,62		putative protein	_
BX680492	RS44C05	2,65E-01	1	1,36E-18	73	1,00E+00	0	74	74	73	76	11,331	11,516	10,465	11,051	12,703		4,72	putative protein	_
BX254733	Ge103E05	1,93E-03	6	6,34E-15	58	1,17E-01	6	70	70	58	58	9,465	9,804	8,976	9,195	10,647		3,19	F-box protein (Fragment), partial (38%)	Transcription
BX680166	RS40A07	5,90E-01	0	6,58E-13	53	5,54E-03	12	66	65	53	34	8,346	8,376	7,840	8,512	8,826		1,98	Pathogenesis-related protein	Réponse au Stress
CR393563	RN61G09	1,66E-01	1	5,74E-13	54	1,91E-02	10	66	65	54	70	10,166	10,259	10,730	10,280	9,622		2,16	MERI-5 protein precursor (Xyloglucan:xyloglucosyl transferase	Croissance cellulaire, division
BX678010	RN55B05	8,27E-02	2	5,74E-13	56	2,46E-01	5	64	63	56	46	8,812	9,081	8,361	8,878	9,637		2,42	putative protein	_
BX251816	Ge055H04	2,59E-01	1	2,68E-13	54	2,81E-02	8	69	63	54	42	8,677	8,921	8,258	8,911	9,365		2,15	Heat shock 70kd protein	Réponse au Stress
BX678105	RN56D01	6,79E-01	0	1,15E-12	51	6,35E-03	12	66	63	51	74	11,027	11,155	11,786	11,152	10,317		2,77	putative protein	_
BX249564	Ge025E05	6,52E-01	0	2,49E-14	60	1,00E+00	0	67	60	60	71	10,583	10,389	9,624	10,626	11,305		3,21	Dehydrin	Réponse au Stress
CR392453	RN32D03	5,27E-02	3	2,56E-12	56	1,00E+00	2	61	60	56	70	10,209	10,400	11,016	10,209	9,618		2,63	Peroxidase precursor (EC 1.11.1.7)	Réponse au Stress
BX676888	RN13E06	1,53E-01	2	1,20E-12	55	1,00E+00	2	63	59	55	39	8,762	8,782	9,357	9,064	7,959		2,64	Carbonate dehydratase (EC 4.2.1.1)	Métabolisme
CR392825	RN48H03	6,88E-02	3	2,83E-09	54	1,00E+00	2	61	59	54	56	9,421	9,979	10,687	9,939	8,489		4,59	Epimerase/dehydratase	Métabolisme
CR393320	RN58E01	3,42E-03	7	5,42E-11	50	1,00E+00	0	58	58	50	51	9,078	9,337	9,711	9,101	8,754		1,94	Cytochrome P450 (EC 1.14.-.-)	Métabolisme
CR392585	RN35A10	6,37E-01	0	6,54E-11	49	8,82E-02	8	58	58	49	76	13,990	13,927	14,307	14,079	13,504		1,75	Peroxidase precursor (EC 1.11.1.7)	Réponse au Stress
CR393329	RN58F01	8,58E-03	5	1,91E-11	51	1,00E+00	0	60	57	51	66	9,882	10,520	9,582	9,453	11,392		3,83	Major pollen allergen Bet v 1-D/H / Intracellular pathogenesis-r	Non-classé
BX679297	RS22F03	6,54E-03	6	3,04E-11	48	1,00E+00	3	61	57	48	70	10,409	10,945	11,209	11,250	9,777		2,78	putative protein	_
AL750525	RN03E03	2,71E-03	7	2,17E-10	46	1,00E+00	2	58	56	46	62	9,556	10,051	10,582	9,386	9,273		2,48	MtN21 nodulin protein-like	Organisation cellulaire
BX249906	Ge030B06	8,26E-01	0	7,18E-11	52	1,00E+00	3	56	55	52	68	9,921	9,898	9,304	10,099	10,434		2,19	Hydrophobic protein RCI2 (Low temperature and salt induced)	Non-classé
AL750218	AS07A11	2,12E-02	5	2,28E-08	44	5,14E-01	6	58	55	44	56	9,118	9,496	8,945	8,722	9,984		2,40	Chitinase (EC 3.2.1.14)	Réponse au Stress
BX678088	RN56B05	7,99E-01	0	2,06E-10	52	1,00E+00	2	55	54	52	41	8,550	8,624	8,181	8,303	9,219		2,05	putative protein	_
BX253632	Ge086B12	2,24E-01	1	3,76E-12	52	1,00E+00	0	62	54	52	39	8,796	8,461	9,282	8,301	8,116		2,24	Snakin-1, partial (67%)	Réponse au Stress
BX680428	RS43C02	5,88E-02	3	2,06E-10	50	1,00E+00	0	58	53	50	42	8,729	9,156	8,387	8,584	9,718		2,51	putative protein	_
CR392434	RN31H05	3,85E-02	4	5,52E-10	47	1,00E+00	3	55	53	47	71	10,021	10,187	9,688	10,018	10,617		1,90	Seven transmembrane protein 1	Communication/Transduction du signal
BX679749	RS32G05	1,65E-01	2	1,96E-10	51	1,00E+00	1	54	53	51	54	9,538	9,288	8,929	8,810	10,302		2,81	putative protein	_
CR393567	RN61H01	2,70E-03	8	4,48E-09	42	1,00E+00	3	55	53	42	75	11,947	12,687	11,828	11,331	13,563		4,70	Chitinase precursor (EC 3.2.1.14)	Réponse au Stress
AL750498	RN03B11	8,43E-01	0	7,09E-09	52	1,00E+00	1	55	53	52	63	10,683	10,716	11,235	11,179	9,843		2,62	NADPH dehydrogenase (EC 1.6.99.1) / 12-OXOPHYTODIENC	Energie
BX681387	RS58C02	1,21E-01	3	5,11E-09	50	1,00E+00	1	53	53	50	52	9,429	9,224	8,753	8,918	10,238		2,80	Licheninase precursor (EC 3.2.1.73)	Réponse au Stress
CR393855	RN66C01	5,67E-02	3	5,09E-10	48	1,00E+00	1	54	53	48	76	13,107	13,299	13,755	13,016	12,745		2,01	putative protein	_
BX253650	Ge086D09	2,92E-02	5	7,58E-09	48	1,00E+00	0	56	53	48	29	8,493	8,263	7,906	8,304	8,861		1,94	Unknown	Non-classé
AL750706	RN06D08	2,09E-01	2	3,79E-10	49	1,00E+00	0	54	51	49	58	9,460	9,745	8,979	9,752	10,196		2,33	Nodulin	Organisation cellulaire
BX255769	Ge009E10	2,09E-02	5	4,57E-09	45	1,00E+00	1	51	50	45	74	10,647	10,911	11,246	10,834	10,262		1,98	T19F11.6 protein (NDR1/HIN1-like protein 1) (Harpin-induced f	Réponse au Stress
BX680242	RS41A02	3,01E-02	5	3,45E-09	45	1,00E+00	0	51	50	45	51	8,980	9,262	9,566	9,166	8,636		1,90	Disease resistance response protein 206	Réponse au Stress
AL751292	RS07B10	1,77E-02	5	8,00E-09	44	1,00E+00	0	54	49	44	64	9,692	10,094	9,522	9,649	10,523		2,00	Gamma-lyase	Métabolisme
BX681893	RS66D06	7,84E-02	3	2,55E-09	46	1,00E+00	1	51	49	46	58	9,289	9,417	9,880	9,111	8,953		1,90	putative protein	_
BX679871	RS35C05	3,61E-01	1	1,71E-08	41	2,09E-01	8	50	49	41	76	11,759	11,903	12,488	12,142	10,938		2,93	Pinus taeda clone PtaAGP5	Non-classé
BX680502	RS44D07	9,36E-01	0	9,39E-09	43	5,46E-01	5	54	48	43	74	10,747	10,771	10,298	10,708	11,252		1,94	putative protein	_
BX681731	RS63F03	9,21E-03	7	1,68E-08	41	1,00E+00	0	50	48	41	76	10,921	11,290	10,719	10,833	11,721		2,00	Anther specific LAT52 protein precursor	Organisation cellulaire
BX682416	Ge116C09	4,59E-02	3	1,14E-09	44	1,00E+00	1	57	48	44	60	9,398	9,786	9,071	9,795	10,030		1,94	G-box binding factor 4.	Transcription
AL750116	AS05H10	1,26E-01	2	1,00E-08	42	1,00E+00	3	54	48	42	49	9,068	9,455	8,857	8,864	9,989		2,19	EF-HAND containing protein-like	Non-classé
CR393111	RN52G02	9,23E-03	6	5,18E-09	41	1,00E+00	1	54	48	41	74	9,971	10,419	9,806	9,982	10,781		1,97	Unknown	Non-classé
BX255062	Ge110B02	6,37E-03	8	4,73E-08	39	1,00E+00	1	49	48	39	76	10,452	10,898	10,248	10,342	11,378		2,19	Ubiquitin-conjugating enzyme E2 (EC 6.3.2.19)	Organisation cellulaire
BX251544	Ge051A06	2,71E-01	1	7,82E-13	46	1,00E+00	0	70	47	46	76	11,513	11,071	11,810	11,347	10,670		2,20	putative protein	_
non soumis	RN59A03	1,01E-02	6	3,72E-08	39	1,00E+00	2	50	47	39	76	13,788	13,961	14,196	13,759	13,617		1,49	putative protein	_
CR392586	RN35A11	3,39E-02	4	2,26E-06	27	7,59E-03	15	52	47	27	65	9,724	10,058	10,312	9,651	9,618		1,62	putative protein	_
BX682453	Ge116H01	8,32E-01	0	3,72E-08	43	1,00E+00	4	52	47	43	49	9,092	9,179	8,730	9,092	9,651		1,89	Receptor-like protein kinase, partial (70%)	Communication/Transduction du signal
BX251314	Ge048A11	5,08E-01	1	4,74E-08	39	2,83E-01	7	49	47	39	42	8,995	8,881	8,652	8,895	9,264		1,53	putative protein	_
BX678993	RS18B02	4,86E-02	3	2,52E-09	42	1,00E+00	2	56	47	42	59	9,509	10,065	10,327	10,210	8,976		2,55	IN2-2 protein	Non-classé
BX680890	RS50C06	8,64E-03	7	1,41E-07	38	1,00E+00	1	49	46	38	43	8,776	9,239	8,575	8,838	9,603		2,04	NADPH:quinone reductase (EC 1.6.5.5)	Energie
BX681545	RS61A07	8,54E-03	7	2,06E-07	37	1,00E+00	1	46	45	37	56	9,355	9,688	10,132	9,208	9,092		2,06	Cell elongation protein DIMINUTO	Croissance cellulaire, division
BX677148	RN18D05	7,03E-02	4	8,14E-08	39	1,00E+00	2	46	45	39	68	9,653	9,977	9,365	9,569	10,479		2,16	Unknown	Non-classé
BX679458	RS26B10	1,11E-03	12	3,09E-06	30	1,00E+00	2	45	44	30	76	12,047	12,564	12,766	12,839	11,507		2,52	Proline-rich protein	Non-classé
BX253244	Ge080C01	8,39E-02	3	1,78E-07	38	1,00E+00	1	45	43	38	46	8,913	9,193	8,664	8,965	9,546		1,84	Lin1294 protein, partial (9%)	Non-classé
BX677447	RN39C09	1,12E-03	12	3,61E-06	30	1,00E+00	0	44	42	30	62	9,409	9,886	10,375	8,973	9,336		2,64	putative protein	_
BX253795	Ge088F01	1,31E-01	2	3,02E-07	36	1,00E+00	4	46	42	36	42	8,549	8,790	8,284	8,775	9,032		1,68	Unknown	Non-classé
BX680324	RS42A12	4,65E-02	4	2,32E-07	37	1,00E+00	1	45	42	37	72	10,092	10,493	9,883	9,999	10,949		2,09	Unknown	Non-classé
BX678294	RN69C07	3,05E-02	6	1,14E-06	36	1,00E+00	0	44	42	36	57	9,232	9,606	9,965	9,272	8,990		1,96	Nuclear transport factor 2 (NTF-2)	Traffic intracellulaire
BX250140	Ge032H07	1,89E-03	11	2,60E-06	31	1,00E+00	0	46	42	31	76	13,876	14,402	13,855	13,830	14,707		1,84	Defender against cell death	Réponse au Stress
BX255545	Ge007A07	1,61E-02	6	8,08E-07	34	1,00E+00	0	44	41	34	73	10,123	10,386	10,582	10,141	9,995		1,50	Prohibitin	Croissance cellulaire, division
AL750511	RN03D01	1,81E-01	2	8,08E-07	34	1,00E+00	4	44	41	34	49	8,804	8,928	9,262	8,871	8,446		1,76	PtaAGP5 (from Pinus taeda)	Non-classé
BX681992	RS67G12	5,45E-01	1	2,74E-06	34	6,47E-01	6	43	40	34	37	8,172	8,298	7,939	7,946	8,704		1,70	Fluconazole resistance protein 1	Transport
BX678224	RN68A05	2,67E-03	10	3,97E-06	30	1,00E+00	0	44	40	30	76	13,124	13,418	13,900	12,977	12,805		2,14	Late embryogenesis abundant protein	Réponse au Stress
BX682091	RS70D01	1,86E-01	2	8,31E-07	37	1,00E+00	1	40	40	37	76	11,404	11,566	11,074	11,164	12,143		2,10	putative protein	_
BX678858	RS16B02	3,58E-02	5	5,06E-06	31	1,00E+00	4	41	40	31	76	12,364	12,081	11,924	12,012	12,657		1,66	putative protein	_
BX249919	Ge030C08	7,10E-02	4	1,62E-06	35	1,00E+00	1	40	40	35	76	10,574	10,765	10,405	10,189	11,276		2,12	Unknown	Non-classé
BX250172	Ge033C10	5,93E-01	0	9,72E-08	38	1,00E+00	0	48	38	38	54	9,391	9,471	9,076	9,440	9,812		1,67	putative protein	_
BX680905	RS50E05	9,37E-01	0	2,24E-06	35	1,00E+00	2	38	37	35	76	11,694	11,758	11,223	12,102	12,028		1,84	Fumarate hydratase, mitochondrial precursor (EC 4.2.1.2)	Energie
BX681270	RS56D01	3,85E-02	5	1,80E-06	34	1,00E+00	0	40	39	34	44	8,715	9,008	9,270	8,924	8,405		1,82	putative protein	_
BX680817	RS49A12	5,41E-01	1	4,74E-06	34	1,00E+00	0	35	35	34	65	9,736	9,620	9,270	9,474	10,238		1,96	putative protein	_
CR392379	RN31A10	9,57E-01	0	4,73E-06	34	1,00E+00	0	35	34	34	70	10,253	10,190	10,635	9,986	9,925		1,64	Beta-glucosidase precursor (EC 3.2.1.21)	Métabolisme
BX680966	RS51E10	6,35E-01	0	2,98E-06	34	1,00E+00	5	41	39	34	66	9,643	9,783	10,032	9,933	9,240		1,73	putative protein	_
BX250343	Ge035G01	5,41E-02	6	3,45E-05	34	1,00E+00	1	43	41	34	28	8,345	8,188	7,983	8,182	8,590		1,52	putative protein	_
BX677655	RN42G05	1,57E-02	7	5,42E-05	24	1,76E-01	9	42	40	24	76	14,883	14,675	14,958	14,940	14,468		1,40	Flavonoid 3',5'-hydroxylase (EC 1.14.-.-)	Métabolisme
AL750567	RN03H10	1,15E-03	16	2,29E-04	26	1,00E+00	1	44	43	26	50	9,732	10,204	10,347	9,725	9,722		1,54	60S ribosomal protein L14	Synthèse de protéines
CR392581	RN35A05	3,38E-03	13	2,60E-04	27	1,00E+00	0	41	41	27	60	10,834	11,310	11,523	10,976	10,675		1,80	Isopentenyl-diphosphate delta-isomerase II (EC 5.3.3.2)	Métabolisme
BX252971	Ge075H01	5,09E-01	0	2,11E-02	6	8,49E-09	46	58	53	46	76	12,427	12,199	12,052	12,336	12,561			putative protein	_

Tableau Ax-4 : Gènes différentiellement exprimés dans le **stress long** et paramètres statistiques associés
Légende : cf. Tab.Ax-3.

Acc Number	ID	Pvalprov	r2 prov	Pvalcond	r2 cond	Pvalinter	r2 inter	r2 tot	r2 PCI	r2 max	data > Bkg	Moy France	Moy Maroc	Moy Cont	Moy 3S	fact prov	fact cond	Annotation	catégorie fonctionelle
BX252065	Ge059C02	5,61E-08	15	1,13E-15	53	2,29E-09	21	95	89	53	40	13,147	14,220	14,504	12,648	2,10	3,62	17.9 kDa class II heat shock protein	Réponse au Stress
BX000656	HSP	5,01E-07	9	3,82E-19	79	6,99E-05	6	96	94	79	30	9,818	10,552	9,067	11,156	1,66	4,25	putative protein	_
BX682091	RS70D01	3,75E-07	18	1,40E-15	74	1,00E+00	0	93	92	74	40	11,323	10,561	10,253	11,783	1,70	2,89	putative protein	_
CR394499	RN50C09	5,61E-08	39	2,92E-10	46	4,99E-01	3	88	87	46	40	11,464	10,337	10,404	11,622	2,18	2,33	Antimicrobial peptide	Réponse au Stress
CR393398	RN59F04	5,01E-07	21	1,77E-13	61	4,15E-01	2	91	85	61	33	10,193	11,271	11,453	9,796	2,11	3,15	putative protein	_
AL750636	RN05F08	8,79E-08	50	6,69E-07	30	1,00E+00	1	83	80	50	40	12,587	10,975	11,335	12,550	3,06	2,32	Antimicrobial peptide	Réponse au Stress
non soumis	RN28G12	6,58E-06	14	8,94E-14	62	4,05E-02	4	92	80	62	30	9,593	10,468	10,705	9,181	1,83	2,88	putative protein	_
AL750498	RN03B11	9,86E-06	28	1,49E-08	44	3,98E-02	9	83	80	44	28	10,187	9,451	10,348	9,438	1,67	1,88	NADPH dehydrogenase (EC 1.6.99.1) / 12-OXOPHYTODIENC	Energie
BX679343	RS23E08	5,01E-07	42	5,66E-07	32	1,00E+00	2	81	76	42	31	10,590	9,227	9,484	10,607	2,57	2,18	Chaperone protein dnaJ	Réponse au Stress
BX681246	RS56A08	5,01E-07	46	3,19E-06	29	1,00E+00	1	80	75	46	40	12,817	11,298	11,647	12,771	2,87	2,18	putative protein	_
BX681928	RS66H05	5,01E-07	41	3,41E-07	33	1,00E+00	0	82	74	41	40	13,137	11,510	11,820	13,152	3,09	2,52	Antimicrobial peptide	Réponse au Stress
CR394036	RN72F06	9,86E-06	39	5,89E-06	32	1,00E+00	0	76	72	39	16	8,417	9,323	9,165	8,394	1,87	1,71	Agglutinin alpha chain (MPA)	Organisation cellulaire
BX680492	RS44C05	5,01E-07	44	7,35E-06	25	1,00E+00	0	83	68	44	30	10,090	9,200	9,412	10,112	1,85	1,63	putative protein	_
BX252988	Ge076B04	1,00E-06	51	4,12E-04	18	1,00E+00	3	76	72	51	36	10,897	9,886	10,208	10,777	2,01	1,48	Thioredoxin H-type (TRX-H)	Energie
BX678105	RN56D01	5,01E-07	42	3,31E-01	1	3,56E-05	28	82	71	42	36	10,727	9,569	10,171	10,358	2,23		putative protein	_
BX254079	Ge093A05	2,07E-06	47	2,62E-01	2	1,21E-03	23	76	72	47	20	9,545	8,972	9,255	9,377	1,49		Inositol oxygenase (Myo-inositol oxygenase) (Aldehyde reduct	Métabolisme
BX250268	Ge034E09	2,14E-09	76	3,24E-01	1	1,00E+00	0	82	78	76	25	9,910	8,150	9,095	9,317	3,39		putative protein	_
BX682328	Ge115B05	9,86E-06	48	1,20E-03	18	1,00E+00	1	70	68	48	29	10,228	9,428	10,147	9,669	1,74		putative protein	_
CR392220	RN28D08	5,01E-07	66	8,37E-01	0	1,00E+00	1	71	67	66	40	13,461	12,574	13,082	13,131	1,85		putative protein	_
BX681997	RS67H05	9,63E-06	58	1,23E-01	5	1,00E+00	0	65	64	58	23	9,547	8,650	9,324	9,053	1,86		Beta-glucosidase, chloroplast precursor (EC 3.2.1.21)	Métabolisme
CR392379	RN31A10	9,86E-06	55	5,53E-02	8	1,00E+00	1	66	63	55	27	10,009	9,277	9,847	9,586	1,66		Beta-glucosidase precursor (EC 3.2.1.21)	Métabolisme
BX679792	RS34A05	6,63E-06	41	5,47E-03	10	1,00E+00	3	76	54	41	28	9,529	10,341	10,032	9,676	1,76		PrMC3	Non-classé
CR392105	RN12D09	5,31E-05	13	1,45E-11	49	3,56E-05	16	90	77	49	40	12,258	13,215	13,421	11,861	1,94	2,95	Pinus taeda calcium binding protein	Non-classé
BX678810	RS15D08	1,75E-05	15	2,65E-13	67	1,78E-01	3	90	86	67	26	9,969	9,073	8,757	10,464	1,86	3,26	putative protein	_
BX679939	RS36F06	1,81E-05	23	1,27E-10	61	1,00E+00	0	85	84	61	28	8,977	9,657	9,790	8,709	1,60	2,11	Peroxidase precursor (EC 1.11.1.7)	Réponse au Stress
BX249884	Ge029H02	3,01E-05	27	8,53E-08	41	3,02E-02	10	81	77	41	26	9,815	9,183	9,200	9,925	1,55	1,65	Phosphoenolpyruvate carboxykinase (EC 4.1.4.49)	Energie
CR394078	RN73D08	8,59E-05	22	1,54E-08	46	4,30E-02	9	82	77	46	22	9,329	8,572	8,533	9,520	1,69	1,98	Licheninase precursor (EC 3.2.1.73)	Réponse au Stress
BX681871	RS66A11	5,75E-05	36	4,30E-04	20	9,04E-01	4	72	60	36	40	11,571	12,066	11,936	11,602	1,41	1,26	putative protein	_
BX681954	RS67C08	3,01E-05	31	8,21E-02	4	5,82E-05	31	78	66	31	40	10,973	10,462	10,861	10,676	1,42		37 kDa inner envelope membrane protein	Non-classé
BX680855	RS49G03	8,94E-05	35	2,06E-01	3	1,54E-03	26	71	65	35	38	11,911	10,867	11,347	11,640	2,06		Alcohol dehydrogenase (EC 1.1.1.1)	Métabolisme
AL751302	RS07C09	2,28E-05	33	7,44E-03	9	2,09E-01	7	77	49	33	38	10,784	11,784	11,413	10,955	2,00		Metallothionein-like protein	Métabolisme
CR392325	RN30F07	1,68E-05	47	4,60E-01	2	1,00E+00	0	69	49	47	40	11,525	11,006	11,369	11,267	1,43		40S ribosomal protein S26 (S31)	Synthèse de protéines
BX251675	Ge053G10	1,54E-05	57	5,81E-01	1	1,00E+00	0	63	58	57	32	10,393	9,349	10,054	9,896	2,06		Poly(A)-binding protein, partial (84%)	Synthèse de protéines
BX682057	RS69D04	6,41E-04	5	4,72E-17	82	3,24E-04	5	96	92	82	26	9,523	9,110	8,386	10,384	1,33	4,00	putative protein	_
BX250781	Ge041B11	1,63E-04	11	5,12E-14	74	4,79E-02	5	90	89	74	23	9,004	9,682	8,431	10,119	1,60	3,22	Low molecular weight heat-shock protein, partial (98%)	Réponse au Stress
AL750468	RN02H01	7,88E-04	11	3,21E-12	69	1,41E-01	5	87	85	69	30	10,535	9,814	9,433	11,060	1,65	3,09	F21O3.6 Hypothetical protein	Non-classé
CR393329	RN58F01	6,74E-04	14	8,53E-11	63	8,60E-02	6	85	83	63	26	10,228	9,507	9,226	10,654	1,65	2,69	Major pollen allergen Bet v 1-D/H / Intracellular pathogenesis-r	Non-classé
BX667551	RN09D10	1,14E-04	18	8,85E-11	62	1,00E+00	2	85	81	62	29	10,321	9,588	9,400	10,656	1,66	2,39	4-hydroxyphenylpyruvate dioxygenase (EC 1.13.11.27)	Métabolisme
BX681136	RS54B08	1,07E-04	22	4,49E-09	52	1,81E-01	6	82	80	52	36	11,371	10,429	11,711	10,277	1,92	2,70	Dehydration stress-induced protein	Réponse au Stress
BX254733	Ge103E05	8,98E-04	18	7,98E-08	45	2,23E-01	6	81	69	45	20	8,626	9,216	8,476	9,273	1,51	1,74	Putative F-box protein (Fragment), partial (38%)	Transcription
BX253595	Ge085G01	8,95E-04	13	3,80E-08	33	1,65E-02	9	86	54	33	40	12,111	12,555	12,009	12,569	1,36	1,47	Zinc finger protein HIT-10, partial (3%)	Transcription
non soumis	RN32A07	8,93E-04	26	1,64E-05	34	1,00E+00	1	70	61	34	24	9,263	9,824	9,773	9,202	1,48	1,48	putative protein	_
BX680191	RS40C09	7,81E-04	32	1,80E-04	30	1,00E+00	2	64	63	32	30	10,264	9,409	10,318	9,527	1,81	1,73	putative protein	_
BX252185	Ge061A06	8,98E-04	30	7,64E-04	22	2,83E-01	10	66	61	30	20	8,745	9,255	9,165	8,733	1,42	1,35	Mitochondrial oxaloacetate transport protein	Transport
BX682032	RS68F08	8,95E-04	33	8,21E-04	24	1,00E+00	0	69	57	33	40	13,923	14,459	13,899	14,318	1,45	1,34	putative protein	_
BX680909	RS50F11	8,93E-04	21	5,65E-02	5	3,56E-05	37	76	64	37	40	10,755	10,459	10,715	10,557	1,23		Pinus sylvestris class III peroxidase	Réponse au Stress
BX681552	RS61B05	3,45E-04	36	2,54E-03	18	2,55E-01	10	71	64	36	28	9,235	9,827	9,625	9,240	1,51		putative protein	_
BX679759	RS33B10	1,03E-04	49	1,77E-01	5	1,00E+00	0	59	54	49	40	13,159	11,959	12,860	12,497	2,30		Adenylate kinase (EC 2.7.4.3)	Métabolisme
BX679434	RS25D07	1,14E-04	48	1,54E-01	6	1,00E+00	0	59	54	48	40	12,644	12,018	12,289	12,497	1,54		putative protein	_
BX681865	RS65H11	1,76E-04	31	4,44E-01	1	1,65E-02	17	72	50	31	31	9,765	10,309	10,033	9,932	1,46		Dihydropinosylvin synthase (EC 2.3.1.-)	Métabolisme
CR393627	RN62F09	3,50E-04	48	8,26E-01	0	1,00E+00	0	53	49	48	19	9,284	8,516	8,945	9,009	1,70		putative protein	_
CR392200	RN28B03	9,90E-04	41	3,83E-01	3	1,00E+00	5	53	48	41	28	9,869	9,179	9,681	9,505	1,61		Defensin	Réponse au Stress
CR394232	RN75E03	8,95E-04	35	2,00E-02	12	1,00E+00	1	61	48	35	24	9,882	9,164	9,788	9,401	1,64		putative protein	_
non soumis	RN59A03	8,95E-04	41	2,20E-01	5	1,00E+00	0	54	46	41	40	13,845	13,312	13,544	13,720	1,45		putative protein	_
CR393855	RN66C01	3,82E-04	43	6,38E-01	1	1,00E+00	0	57	44	43	40	13,703	12,854	13,424	13,303	1,80		putative protein	_
AL751204	RS06C03	1,70E-03	33	1,61E-03	23	1,00E+00	0	72	56	33	20	8,496	9,006	8,894	8,526	1,42	1,29	T29M8.12 protein	Métabolisme
CR394054	RN73A01	1,05E-03	34	4,75E-03	18	1,00E+00	2	60	55	34	40	14,901	14,543	14,618	14,897	1,28	1,21	putative protein	_
BX679276	RS22D04	3,20E-03	31	4,62E-03	20	1,00E+00	2	56	54	31	30	10,199	9,668	10,191	9,782	1,44	1,33	putative protein	_
AL750133	AS06B05	4,48E-03	28	2,24E-03	23	1,00E+00	2	57	52	28	28	10,015	9,265	10,027	9,403	1,68	1,54	putative protein	_
BX679259	RS22B08	1,36E-03	32	6,04E-03	17	1,00E+00	1	61	49	32	40	11,909	10,938	11,197	11,844	1,96	1,57	putative protein	_
CR393324	RN58E05	9,09E-03	16	1,35E-02	10	4,02E-05	43	72	69	43	36	10,696	10,310	10,394	10,690	1,31		Epoxide hydrolase (EC 3.3.2.3)	Réponse au Stress
BX681938	RS67A12	4,15E-03	20	3,21E-01	2	3,56E-05	46	70	68	46	40	12,729	12,236	12,449	12,615	1,41		Protein kinase (EC 2.7.1.-)	Métabolisme
BX678224	RN68A05	7,56E-03	19	3,11E-01	3	1,10E-04	42	68	64	42	40	12,711	12,155	12,393	12,584	1,47		Late embryogenesis abundant protein	Réponse au Stress
BX680131	RS39E04	3,57E-03	22	7,28E-01	1	1,38E-03	30	68	53	30	39	10,567	11,158	10,764	10,843	1,51		putative protein	_
BX255062	Ge110B02	2,22E-03	37	9,61E-02	8	1,00E+00	0	51	45	37	20	9,385	9,033	9,157	9,332	1,28		Ubiquitin-conjugating enzyme E2 (EC 6.3.2.19)	Organisation cellulaire
BX251837	Ge056B06	2,22E-03	39	2,08E-01	6	1,00E+00	0	48	45	39	36	10,108	10,453	10,175	10,317	1,27		Protein transport protein SEC22	Traffic intracellulaire
BX680940	RS51B12	4,04E-03	33	1,01E-01	8	1,00E+00	4	51	45	33	36	10,698	10,332	10,648	10,455	1,29		putative protein	_
AL750626	RN05E11	4,09E-03	34	1,22E-01	8	1,00E+00	1	49	42	34	24	9,819	9,016	9,674	9,322	1,75		putative protein	_
BX252363	Ge063D09	8,83E-03	33	3,57E-01	4	1,00E+00	5	47	42	33	26	9,554	9,000	9,432	9,249	1,47		Unknown	Non-classé
BX679160	RS20G09	3,20E-03	38	8,45E-01	0	1,00E+00	2	46	40	38	26	9,851	9,268	9,644	9,592	1,50		putative protein	_
BX680587	RS45E04	2,46E-03	39	6,87E-01	1	1,00E+00	0	47	40	39	40	11,059	11,880	11,448	11,327	1,77		Unknown	Non-classé
BX255048	Ge109H02	4,35E-03	38	8,94E-01	0	1,00E+00	0	42	38	38	18	8,868	8,371	8,652	8,686	1,41		Unknown	Non-classé
BX682761	Ge121C11	5,70E-03	36	8,92E-01	0	1,00E+00	1	42	37	36	36	10,614	10,013	10,352	10,396	1,52		putative protein	_
BX252848	Ge071F08	6,24E-03	36	8,91E-01	0	1,00E+00	0	42	36	36	20	9,404	8,947	9,204	9,237	1,37		putative protein	_
BX681153	RS54D06	2,88E-03	35	3,20E-01	4	1,00E+00	1	56	40	35	32	10,441	9,770	10,257	10,087	1,59		putative protein	_
BX678708	RS14B03	4,58E-01	1	4,07E-15	64	1,10E-08	24	93	89	64	40	12,361	12,190	13,349	11,237		4,32	Basic blue protein (Plantacyanin)	Energie
BX677917	RN46E01	5,34E-01	0	3,21E-12	30	6,38E-07	13	95	43	30	40	15,964	16,048	15,771	16,225		1,37	putative protein	_
BX678354	RS08A06	6,30E-01	1	1,30E-09	43	3,56E-05	21	87	65	43	40	11,101	11,239	11,433	10,879		1,47	Pinus taeda Trans-cinnamate 4-monooxygenase (EC 1.14.13.1	Métabolisme
BF220737	NXCI_150_A7	4,98E-01	1	5,31E-08	35	1,30E-05	28	89	64	35	39	11,260	11,467	11,625	10,963		1,58	Arginine decarboxylase	Métabolisme
CR392796	RN48B11	1,61E-02	7	1,61E-10	55	1,10E-04	18	86	80	55	40	12,669	12,505	12,323	12,883		1,47	putative protein	_
BX679749	RS32G05	8,83E-03	3	5,20E-17	86	3,03E-03	5	95	94	86	26	9,302	8,906	8,100	10,087		3,97	putative protein	_
BX251374	Ge048G11	5,98E-02	3	9,43E-15	77	5,81E-03	6	91	87	77	37	11,096	11,501	12,113	10,403		3,27	Rof1, partial (88%)	Transcription
BX680863	RS49H05	5,26E-01	1	3,87E-12	73	3,96E-03	11	86	85	73	30	9,974	9,783	10,715	9,080		3,11	putative protein	_
BX681964	RS67D07	3,29E-01	2	8,54E-12	72	9,36E-03	10	86	84	72	30	10,203	10,438	10,932	9,662		2,41	putative protein	_
BX680021	RS38C03	4,14E-01	1	4,40E-12	70	1,41E-03	12	87	84	70	30	10,219	10,446	10,950	9,669		2,43	Dihydroflavonol-4-reductase (EC 1.1.1.219) (DFR)	Métabolisme
BX676949	RN14F10	4,04E-03	10	1,43E-10	57	1,51E-03	13	86	80	57	34	10,494	11,082	11,361	10,096		2,40	putative protein	_
BX677658	RN42H03	3,02E-02	11	1,20E-06	38	3,88E-03	19	76	68	38	27	10,063	9,531	10,410	9,291		2,17	Nonspecific lipid-transfer protein precursor (LTP)	Transport
AW985459	NXNV_136_H4	1,53E-01	6	6,02E-07	40	1,54E-03	21	77	67	40	40	11,510	11,822	11,992	11,278		1,64	Aspartate transaminase	Métabolisme
BX681976	RS67E11	8,63E-01	0	3,38E-06	43	8,64E-03	21	70	64	43	36	10,058	10,105	9,838	10,315		1,39	Histidine decarboxylase (EC 4.1.1.22)	Métabolisme
BX252366	Ge063E01	6,63E-01	0	7,18E-25	98	1,00E+00	0	98	98	98	30	9,551	9,600	8,542	10,600		4,16	Class I Heat Shock Protein	Réponse au Stress
BX682159	RS72E06	7,65E-02	2	3,18E-16	88	1,00E+00	1	92	92	88	30	10,870	10,366	9,294	12,042		6,72	Flavonoid 3',5'-hydroxylase (EC 1.14.-.-)	Métabolisme
BX676897	RN13F04	5,20E-01	1	1,69E-14	88	1,00E+00	0	90	89	88	40	11,156	11,243	10,710	11,671		1,95	Syringomycin biosynthesis enzyme-like protein	Métabolisme
BX679598	RS29C05	5,68E-02	4	2,20E-13	81	6,82E-02	4	91	89	81	25	9,683	8,979	7,787	10,706		7,57	putative protein	_
CR393567	RN61H01	1,43E-01	3	1,10E-13	85	1,00E+00	0	88	88	85	40	11,619	11,371	10,841	12,198		2,56	Chitinase precursor (EC 3.2.1.14)	Réponse au Stress
AL749654	AN05E01	4,50E-02	4	3,78E-14	82	1,00E+00	1	91	87	82	19	8,685	8,447	8,053	9,075		2,03	Oxygen-evolving enhancer protein 1, chloroplast precursor	Energie
BX680601	RS45G02	7,31E-01	0	2,79E-13	84	8,45E-01	2	87	86	84	20	8,858	8,925	8,502	9,267		1,70	Hypothetical 25.9 kDa protein	Non-classé
BX679298	RS22F06	4,20E-01	1	3,63E-13	81	3,39E-01	3	87	86	81	40	11,018	11,228	10,374	11,829		2,74	putative protein	_
BX249564	Ge025E05	8,04E-01	0	1,08E-17	86	1,00E+00	0	94	86	86	28	9,681	9,814	9,011	10,457		2,72	dehydrin	Réponse au Stress
CR392510	RN33D10	1,34E-01	4	2,52E-12	79	4,99E-01	3	86	85	79	35	10,919	11,447	12,326	9,935		5,24	putative protein	_
BX679458	RS26B10	3,18E-01	1	3,10E-14	81	3,83E-01	3	90	85	81	32	10,933	11,267	12,084	10,049		4,10	Proline-rich protein	Non-classé
BX255575	Ge007D01	5,25E-01	1	1,10E-13	82	5,08E-01	2	92	85	82	40	11,360	11,213	10,803	11,928		2,18	putative protein	_
BX679871	RS35C05	8,44E-02	4	2,92E-12	78	1,00E+00	2	86	84	78	30	10,193	10,689	11,362	9,421		3,84	Pinus taeda clone PtaAGP5	Non-classé
CR393359	RN59B03	4,15E-02	6	1,27E-11	76	1,00E+00	1	84	84	76	36	11,152	11,640	12,204	10,491		3,28	putative protein	_
AL749568	AN01E10	1,38E-02	9	1,27E-10	72	1,00E+00	2	84	83	72	20	8,927	8,512	8,185	9,284		2,14	putative protein	_
BX252727	Ge069G09	4,56E-02	7	8,44E-11	72	8,29E-01	3	82	82	72	29	10,011	10,537	11,077	9,365		3,28	Esterase precursor (Early nodule-specific protein homolog) (L	Métabolisme
CR392453	RN32D03	9,78E-01	0	2,77E-11	80	1,00E+00	2	82	82	80	28	9,439	9,473	10,325	8,580		3,35	Peroxidase precursor (EC 1.11.1.7)	Réponse au Stress
CR392767	RN37G07	7,98E-01	0	1,90E-11	78	4,31E-01	4	84	82	78	22	8,929	9,039	9,616	8,330		2,44	Peroxidase precursor (EC 1.11.1.7)	Réponse au Stress
CR392528	RN33F11	3,22E-02	8	8,64E-11	74	1,00E+00	0	82	82	74	28	9,969	9,712	9,469	10,263		1,73	putative protein	_

BX678400	RS08F03	1,01E-01	5	6,27E-11	75	1,00E+00	1	82	82	75	26	10,042	9,723	9,343	10,486	2,21	Glutathione transferase (EC 2.5.1.18)	Métabolisme
AL751210	RS06C09	9,53E-01	0	7,96E-11	76	2,61E-01	5	82	81	76	30	9,683	9,711	10,521	8,868	3,14	putative protein	_
BX253001	Ge076C12	7,45E-01	0	3,77E-11	76	3,70E-01	4	83	81	76	34	11,007	11,123	12,026	10,082	3,85	Tubulin alpha chain	Organisation cellulaire
CR392399	RN31D03	3,18E-01	2	2,30E-11	73	1,21E-01	6	84	81	73	30	9,764	9,528	10,399	8,940	2,75	putative protein	_
BX681602	RS61G11	4,41E-02	7	1,71E-10	71	1,00E+00	2	82	80	71	30	10,484	9,917	11,084	9,431	3,15	Defensin	Réponse au Stress
BX254083	Ge093A12	8,69E-01	0	7,70E-11	73	7,58E-02	7	83	80	73	30	9,470	9,450	10,081	8,843	2,36	Geranyl diphosphate synthase, complete	Métabolisme
BX681235	RS55H11	2,57E-01	3	1,36E-10	77	1,00E+00	0	81	80	77	39	10,568	10,318	9,881	11,055	2,26	Glutathione transferase (EC 2.5.1.18)	Métabolisme
BX677733	RN44A03	6,91E-01	1	8,00E-11	78	1,00E+00	0	81	79	78	40	11,697	11,783	11,376	12,087	1,64	Unknown	Non-classé
BX678416	RS08G10	7,69E-01	0	1,57E-14	79	1,00E+00	0	91	79	79	25	9,722	9,815	9,351	10,167	1,76	Chlorophyll A-B binding protein, chloroplastic	Energie
BX252968	Ge075G10	6,37E-01	1	1,61E-10	71	8,75E-02	7	82	79	71	40	12,190	12,327	12,798	11,691	2,15	Tubulin alpha chain	Organisation cellulaire
AL749666	AN05G02	6,31E-01	1	4,38E-09	75	1,00E+00	3	80	79	75	13	8,216	8,097	7,720	8,529	1,75	Photosystem I reaction center subunit N, chloropla	Energie
BX255662	Ge008C06	7,79E-01	0	1,74E-10	72	1,23E-01	7	81	79	72	30	10,001	9,934	10,722	9,227	2,82	Glycine-rich protein homolog	Non-classé
BX681400	RS58E03	1,65E-01	5	7,54E-10	69	4,86E-01	5	80	78	69	20	9,111	8,872	8,549	9,482	1,91	Unknown	Non-classé
BX249241	Ge021G02	5,80E-01	1	3,19E-10	73	4,86E-01	4	80	78	73	32	10,656	10,880	11,561	9,930	3,10	Tubulin alpha chain	Organisation cellulaire
BX677767	RN44E03	7,69E-01	0	7,28E-11	78	1,00E+00	0	82	78	78	33	10,072	10,188	9,604	10,632	2,04	Chitinase precursor (EC 3.2.1.14)	Réponse au Stress
BX253756	Ge088A07	4,15E-03	14	7,60E-09	57	2,23E-01	7	79	78	57	38	11,006	10,323	10,069	11,397	2,51	Aluminum-induced protein-like protein, partial (67%)	Réponse au Stress
BX679165	RS20G05	4,02E-02	8	1,12E-09	61	4,19E-02	9	81	78	61	34	10,761	11,333	11,716	10,264	2,74	putative protein	_
BX678362	RS08B03	8,19E-01	0	3,17E-10	77	1,00E+00	0	79	77	77	40	11,470	11,488	11,188	11,766	1,49	Protein-methionine-S-oxide reductase (EC 1.8.4.6)	Réponse au Stress
BX255369	Ge005A06	5,95E-01	1	8,44E-10	75	1,00E+00	0	77	77	75	40	10,752	10,873	10,385	11,217	1,78	Glycine-rich RNA-binding protein, complete	Communication/Transduction du signal
CR394310	RN76F04	5,34E-01	2	6,59E-10	74	1,00E+00	1	81	76	74	20	8,849	8,715	8,427	9,158	1,66	Transcription factor HBP-1b homolog	Transcription
BX682993	Ge126F07	5,74E-01	2	5,42E-09	71	9,36E-01	4	78	76	71	35	9,432	9,692	10,340	8,742	3,03	UDP-glucose 4-epimerase (EC 5.1.3.2)	Métabolisme
CR393620	RN62F02	9,99E-01	0	2,03E-08	74	1,00E+00	2	77	76	74	15	8,274	8,257	7,769	8,667	1,86	putative protein	_
BX680486	RS44B07	9,52E-01	0	1,51E-09	72	7,84E-01	4	77	76	72	30	9,938	9,892	10,573	9,266	2,47	Poly(A)-binding protein	Transcription
AL751292	RS07B10	5,56E-02	8	4,11E-09	65	1,00E+00	2	78	75	65	22	9,616	9,196	8,913	9,983	2,10	Gamma-lyase	Métabolisme
non soumis	RN28G08	7,91E-02	7	9,88E-09	56	2,09E-02	12	79	75	56	32	9,757	10,096	9,455	10,330	1,83	putative protein	_
AL750677	RN06B02	2,37E-01	4	4,38E-09	69	1,00E+00	1	76	75	69	31	10,194	9,750	9,183	10,849	3,17	putative protein	_
BX681985	RS67F12	3,07E-01	3	3,99E-09	67	7,53E-01	4	77	75	67	30	9,749	9,438	10,304	8,945	2,57	Isoflavone reductase homolog (EC 1.3.1.-)	Métabolisme
BX249398	Ge023F01	4,83E-01	2	3,04E-09	68	6,06E-01	5	77	75	68	20	8,884	8,662	8,269	9,321	2,07	Auxin-repressed 12.5 kd protein	Non-classé
BX253036	Ge077A02	1,96E-01	4	5,35E-10	62	4,95E-02	8	82	74	62	40	12,004	12,257	12,512	11,699	1,76	Methionine synthase, partial (63%)	Synthèse de protéines
AL751212	RS06C11	2,19E-02	11	1,81E-08	56	1,62E-01	8	78	74	56	26	9,851	9,205	8,929	10,257	2,51	Chaperone protein dnaJ	Réponse au Stress
BX682453	Ge116H01	4,53E-01	2	4,38E-09	71	1,00E+00	0	75	74	71	20	8,886	8,723	8,409	9,232	1,77	Receptor-like protein kinase, partial (70%)	Communication/Transduction du signal
BX681290	RS56F02	2,73E-01	4	4,07E-09	66	6,75E-01	4	77	74	66	30	9,710	10,027	10,410	9,263	2,21	40S ribosomal protein S11	Synthèse de protéines
CR394220	RN75C12	3,61E-03	16	2,51E-08	55	1,00E+00	3	77	74	55	22	9,660	9,199	9,076	9,875	1,74	Peptide transporter PTR2-B	Transport
BX677636	RN42D06	1,32E-01	6	7,00E-09	67	1,00E+00	0	76	74	67	28	9,925	10,078	9,760	10,213	1,37	Unknown	Non-classé
CR393563	RN61G09	7,16E-02	9	3,07E-08	61	7,76E-01	4	74	74	61	30	9,954	9,619	10,257	9,383	1,83	MERI-5 protein precursor (Xyloglucan:xyloglucosyl transferase	Croissance cellulaire, division
BX682869	Ge122H09	1,54E-01	6	1,66E-08	61	3,46E-01	6	76	74	61	29	9,629	9,211	10,081	8,843	2,36	Pinene synthase (EC 4.2.3.1)	Métabolisme
CR393691	RN63E10	1,95E-03	19	4,33E-08	55	1,00E+00	0	76	73	55	32	10,377	9,636	9,486	10,675	2,28	putative protein	_
BX254971	Ge108E04	8,04E-01	0	8,29E-09	70	1,00E+00	3	74	73	70	20	9,259	9,322	8,776	9,792	2,02	Alkaline alpha galactosidase II, partial (84%)	Métabolisme
BX249627	Ge026D05	6,06E-01	1	2,61E-09	64	9,66E-02	8	81	73	64	27	9,490	9,620	9,946	9,177	1,70	Tubulin beta chain	Organisation cellulaire
BX253089	Ge077G07	8,55E-01	0	1,30E-08	68	6,34E-01	5	74	73	68	30	9,919	9,877	9,460	10,344	1,85	LEA, complete	Réponse au Stress
BX681166	RS54F02	1,40E-02	5	1,11E-13	64	8,60E-02	4	91	73	64	40	10,733	10,527	10,090	11,211	2,18	NBS/LRR	Réponse au Stress
BX679999	RS35D10	9,70E-01	0	6,04E-09	71	1,00E+00	2	74	73	71	40	12,337	12,444	11,170	13,590	5,35	Unknown	Non-classé
BX249127	Ge020C11	1,23E-03	20	1,23E-07	49	8,50E-01	4	76	72	49	32	9,863	10,729	10,842	9,577	2,40	Metallothionein-like protein EMB30	Métabolisme
BX249875	Ge029G04	5,55E-01	2	2,82E-09	70	1,00E+00	1	77	72	70	30	9,716	9,906	10,350	9,235	2,17	Glucan endo-1,3-beta-glucosidase precursor	Organisation cellulaire
AL749902	AS02D12	4,04E-03	13	1,53E-08	48	1,86E-02	11	81	72	48	24	9,448	9,909	9,234	10,030	1,74	putative protein	_
BX682906	Ge123D02	7,82E-01	0	1,22E-08	67	6,75E-01	5	74	72	67	30	9,619	9,738	10,200	9,133	2,10	Probable ATP-dependent Clp protease proteolytic su	Non-classé
CR392722	RN37B06	1,69E-01	5	7,53E-09	64	1,00E+00	3	77	72	64	29	9,147	9,570	9,932	8,701	2,35	Peroxidase precursor (EC 1.11.1.7)	Réponse au Stress
AL749783	AS01A02	6,29E-01	1	3,05E-08	65	6,34E-01	5	73	71	65	21	9,385	9,263	8,953	9,719	1,70	H(+)-transporting two-sector ATPase delta subunit (EC 3.6.3.1	Energie
BX000608	Tubuline	7,26E-01	1	3,05E-08	65	6,63E-01	5	72	71	65	31	10,481	10,639	11,217	9,871	2,54	Tubuline	Organisation cellulaire
BX679279	RS22D07	5,81E-01	2	6,88E-09	69	1,00E+00	0	75	71	69	38	10,462	10,274	9,827	10,947	2,17	ABA induced plasma membrane protein	Réponse au Stress
BX681039	RS52D12	1,42E-01	6	2,28E-08	60	4,70E-01	6	76	71	60	32	10,223	10,616	10,953	9,808	2,21	putative protein	_
BX251976	Ge058A03	1,39E-03	23	8,43E-07	47	1,00E+00	2	72	71	47	20	9,337	9,025	8,987	9,436	1,36	Unknown	Non-classé
AL751104	RS05B06	2,09E-01	3	8,73E-11	67	1,00E+00	1	84	71	67	40	10,522	10,691	10,314	10,865	1,46	DNA polymerase III	Croissance cellulaire, division
BX679225	RS21G10	4,43E-01	3	3,05E-08	67	1,00E+00	1	72	71	67	20	8,728	8,584	8,306	9,034	1,66	putative protein	_
BX681172	RS54G01	2,97E-01	4	2,42E-08	59	1,41E-01	9	76	71	59	32	10,479	10,747	11,149	10,024	2,18	Endosperm-specific protein-like protein	Non-classé
CR392544	RN33H09	4,36E-01	3	2,42E-08	67	1,00E+00	0	72	70	67	22	9,510	9,303	8,990	9,864	1,83	putative protein	_
BX252031	Ge058G08	8,15E-01	0	3,07E-09	69	1,00E+00	1	77	70	69	24	8,874	9,015	9,484	8,376	2,16	arabinogalactan protein [Pinus taeda]	Réponse au Stress
BX679860	RS35A06	3,40E-01	4	5,21E-07	56	2,30E-01	10	71	70	56	30	9,494	9,735	10,023	9,168	1,81	2-alpha-hydroxytaxane 2-O-benzoyltransferase (EC 2.3.1.166)	Métabolisme
CR392434	RN31H05	4,98E-01	2	1,26E-08	61	2,93E-01	7	76	70	61	24	9,639	9,504	9,003	10,167	2,24	Seven transmembrane protein 1	Communication/Transduction du signal
CR394335	RN76H08	3,03E-02	13	2,14E-07	57	1,00E+00	0	71	70	57	34	10,448	10,846	11,015	10,199	1,76	Glutathione transferase (EC 2.5.1.18)	Métabolisme
BX681832	RS65E04	7,75E-02	8	4,72E-08	55	3,10E-01	7	76	70	55	26	9,727	9,369	9,082	10,086	2,01	Pathogenesis-related protein PRB1-2 precursor	Réponse au Stress
BX679599	RS29C06	4,15E-01	3	4,15E-09	67	1,00E+00	0	77	70	67	40	11,518	11,650	11,321	11,820	1,41	putative protein	_
AL751122	RS05C12	1,54E-01	3	8,54E-12	66	1,00E+00	0	87	70	66	40	11,230	11,406	11,029	11,571	1,46	Glucan endo-1,3-beta-D-glucosidase (EC 3.2.1.39)	Organisation cellulaire
BX252074	Ge059D03	1,70E-01	7	8,33E-08	63	1,00E+00	0	70	70	63	22	8,832	9,228	9,579	8,401	2,26	arabinogalactan protein	Réponse au Stress
BX680204	RS40E02	6,37E-03	16	1,79E-07	51	1,00E+00	3	74	69	51	30	10,335	9,560	9,399	10,651	2,38	4-hydroxyphenylpyruvate dioxygenase (EC 1.13.11.27)	Métabolisme
BX252380	Ge063F05	7,76E-01	1	5,78E-08	67	1,00E+00	2	70	69	67	20	8,512	8,438	8,122	8,843	1,65	putative protein	_
BX675142	RN10A03	4,92E-01	3	6,25E-07	64	1,00E+00	2	70	69	64	27	9,512	9,945	10,636	8,773	3,64	Peroxidase precursor (EC 1.11.1.7)	Réponse au Stress
BX677882	RN46A04	8,43E-01	0	1,84E-08	69	1,00E+00	0	72	69	69	40	11,877	11,968	11,507	12,320	1,76	putative protein	_
CR393412	RN59G08	8,86E-01	0	5,50E-09	68	1,00E+00	0	76	69	68	40	10,843	10,926	10,522	11,231	1,63	putative protein	_
BX677604	RN41H04	9,25E-01	0	5,86E-08	69	1,00E+00	0	69	69	69	40	12,790	12,808	12,589	13,005	1,33	Hypothetical 20.8 kDa protein	Non-classé
AL750545	RN03F11	1,85E-01	6	3,92E-08	63	1,00E+00	0	73	69	63	36	10,524	10,130	9,811	10,921	2,16	Catalase isozyme (EC 1.11.1.6).	Réponse au Stress
BX249499	Ge024G05	1,16E-01	5	6,27E-10	63	1,00E+00	0	82	69	63	40	11,277	11,439	11,055	11,628	1,49	Pherophorin-dz1 protein precursor, partial (7%)	Non-classé
BX677876	RN45H07	5,35E-01	2	3,24E-07	66	1,00E+00	1	72	68	66	20	8,627	8,397	8,020	8,952	1,91	putative protein	_
BX681199	RS55C12	4,72E-01	3	2,89E-08	66	1,00E+00	0	72	68	66	22	9,481	9,265	8,964	9,826	1,82	Unknown	Non-classé
BX678152	RN57A03	5,48E-01	2	1,18E-07	65	1,00E+00	1	68	68	65	19	8,829	8,990	9,322	8,465	1,81	Peroxidase precursor (EC 1.11.1.7)	Réponse au Stress
CR394223	RN75D03	7,61E-02	10	2,76E-07	58	1,00E+00	0	69	68	58	39	10,932	11,350	11,574	10,625	1,93	putative protein	_
BX249652	Ge026F11	1,80E-01	5	2,16E-08	60	1,00E+00	3	81	68	60	24	9,946	9,624	9,259	10,347	2,13	ASR protein, partial (39%)	Réponse au Stress
AL749609	AN02H06	4,44E-01	3	9,78E-08	64	1,00E+00	0	69	68	64	20	9,231	9,112	8,891	9,477	1,50	putative protein	_
BX251416	Ge049C07	4,72E-01	3	3,40E-07	61	1,00E+00	4	74	68	61	21	8,879	8,613	8,307	9,208	1,87	Oxygen-evolving enhancer protein 3-1, chloroplast precursor	Energie
BX254281	Ge096B03	5,56E-02	11	1,38E-06	49	4,13E-01	8	69	68	49	28	9,962	9,489	9,293	10,253	1,95	Ripening-associated protein (Fragment)	Réponse au Stress
CR392825	RN48H03	2,12E-03	23	2,85E-06	44	1,00E+00	1	75	68	44	21	8,740	9,314	9,426	8,564	1,82	Epimerase/dehydratase	Métabolisme
BX677447	RN39C09	3,32E-01	3	7,08E-09	62	1,00E+00	2	77	67	62	17	8,665	8,936	9,218	8,329	1,85	putative protein	_
CR393662	RN63B06	2,35E-01	5	3,05E-08	60	1,00E+00	3	75	67	60	40	11,661	11,429	11,196	11,940	1,67	Proton-exporting ATPase (EC 3.6.3.6)	Transport
BX249919	Ge030C08	1,75E-01	7	7,78E-07	53	5,01E-01	7	68	67	53	31	10,027	9,861	9,732	10,190	1,37	Unknown	Non-classé
BX680451	RS43E08	7,89E-01	0	5,33E-08	66	1,00E+00	0	73	67	66	40	11,682	11,582	11,073	12,207	2,19	putative protein	_
BX250683	Ge040A04	7,51E-01	1	1,97E-07	61	1,00E+00	4	69	67	61	30	9,442	9,572	10,031	8,957	2,10	Fasciclin-like arabinogalactan protein FLA8	Croissance cellulaire, division
BX255769	Ge009E10	9,09E-01	0	6,63E-08	65	1,00E+00	2	70	66	65	30	9,855	9,948	10,371	9,413	1,94	T19F11.6 protein (NDR1/HIN1-like protein 1) (Harpin-induced	Réponse au Stress
CR392836	RN48F07	4,91E-01	3	2,57E-07	62	1,00E+00	2	68	66	62	25	9,673	9,772	9,488	9,937	1,36	Unknown	Non-classé
BX252277	Ge062C04	4,72E-01	3	4,01E-07	59	1,00E+00	4	68	66	59	30	10,342	10,107	10,761	9,735	2,04	Harpin-induced protein	Réponse au Stress
AL749979	AS03D06	9,78E-01	0	1,97E-07	65	1,00E+00	0	67	66	65	38	10,031	10,056	9,654	10,427	1,71	Thioredoxin	Energie
BX679879	RS35D06	9,67E-01	0	2,30E-07	63	1,00E+00	2	67	66	63	40	12,006	12,016	11,782	12,237	1,37	putative protein	_
CR392980	RN51A02	4,88E-01	3	1,02E-07	63	1,00E+00	0	70	66	63	20	9,195	9,295	9,037	9,432	1,32	40S ribosomal protein S8	Synthèse de protéines
BX681925	RS66H01	9,80E-01	0	2,73E-08	65	1,00E+00	0	73	65	65	40	11,510	11,554	11,264	11,791	1,44	Hydroxymethylglutaryl-CoA synthase (EC 2.3.3.10)	Métabolisme
BX250481	Ge037E07	6,05E-01	2	3,39E-07	63	1,00E+00	0	66	65	63	16	8,748	8,931	9,318	8,324	1,99	putative protein	_
BX253114	Ge078C04	5,69E-01	2	4,45E-07	60	1,00E+00	3	70	65	60	21	9,330	9,050	8,615	9,764	2,22	Alkaline alpha galactosidase II, partial (84%)	Métabolisme
CR392217	RN28D05	2,85E-01	5	5,51E-07	53	4,31E-01	7	70	65	53	29	9,336	9,746	10,081	8,918	2,24	Basic blue protein (Cusacyanin) (Plantacyanin)	Energie
BX251932	Ge057D07	7,86E-01	1	4,72E-08	64	1,00E+00	1	72	65	64	24	8,880	9,048	9,503	8,392	2,16	arabinogalactan protein	Non-classé
BX249385	Ge023D08	6,59E-01	1	1,07E-07	63	1,00E+00	1	70	65	63	40	11,459	11,390	11,831	11,032	1,74	60S ribosomal protein	Synthèse de protéines
BX677573	RN41D05	6,04E-01	1	2,75E-08	64	1,00E+00	0	73	65	64	40	12,058	12,168	11,861	12,343	1,40	Yippee-like protein	Non-classé
BX255417	Ge005E10	5,52E-01	2	8,43E-07	61	1,00E+00	1	67	65	61	29	9,955	9,712	10,460	9,272	2,28	P20-Arc, partial (92%)	Non-classé
BX678372	RS08C03	9,35E-01	0	4,97E-09	65	1,00E+00	0	77	65	65	40	12,499	12,563	12,268	12,782	1,43	Cationic peroxidase precursor (EC 1.11.1.7)	Réponse au Stress
BX682416	Ge116C09	7,47E-01	1	8,68E-07	63	1,00E+00	0	65	64	63	20	8,867	8,772	8,474	9,180	1,63	G-box binding factor 4.	Transcription
BX681748	RS63G11	9,92E-01	0	3,44E-07	61	1,00E+00	3	67	64	61	21	8,928	8,962	9,448	8,435	2,02	2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase (EC 4.6	Métabolisme
BX681234	RS55G07	6,09E-01	2	1,11E-06	53	2,83E-01	9	67	64	53	30	9,740	9,621	10,059	9,326	1,66	putative protein	_
CR393320	RN58E01	2,65E-01	6	2,20E-06	53	8,86E-01	6	65	64	53	19	8,991	8,741	9,269	8,514	1,69	Cytochrome P450 (EC 1.14.-.-)	Métabolisme
BX251676	Ge053G11	8,04E-01	0	2,20E-06	49	6,17E-02	15	67	64	49	18	8,776	8,725	9,082	8,428	1,57	Xyloglucan:xyloglucosyl transferase (EC 2.4.1.207)	Croissance cellulaire, division
BX677769	RN44E06	7,60E-01	1	2,28E-08	63	1,00E+00	0	74	64	63	40	11,906	11,994	11,697	12,185	1,40	Enoyl-CoA hydratase (EC 4.2.1.17)	Métabolisme
BX681149	RS54D01	6,55E-01	1	3,80E-07	62	1,00E+00	0	66	64	62	29	9,826	9,774	9,602	10,009	1,33	Unknown	Non-classé
BX681263	RS56C03	1,53E-01	6	7,25E-08	55	1,00E+00	3	74	64	55	40	10,465	10,295	10,097	10,697	1,52	EDS1-like protein	Réponse au Stress
BX252809	Ge071A10	2,17E-01	6	8,28E-07	55	1,00E+00	3	67	64	55	32	10,144	10,460	10,708	9,833	1,83	putative protein	_
BX251276	Ge047F03	9,07E-01	0	5,75E-07	62	1,00E+00	2	65	64	62	20	9,048	9,011	8,747	9,319	1,49	putative protein	_

RN03G12	8,76E-01	0	3,49E-09	61	1,00E+00	2	79	63	61	40	11,662	11,689	11,367	11,979	1,53	Plasma membrane H+ ATPase (EC 3.6.1.3)	Transport
RN50F06	5,81E-01	2	1,32E-07	62	1,00E+00	0	76	63	62	20	8,538	8,635	8,374	8,777	1,32	Unknown	Non-classé
RS25B09	5,41E-01	3	8,90E-07	58	1,00E+00	2	65	63	58	20	9,425	9,274	9,016	9,714	1,62	Cysteine synthase, chloroplast precursor (EC 2.5.1.47)	Métabolisme
RN54H02	9,04E-01	0	2,59E-06	51	1,91E-01	12	65	63	51	35	10,285	10,316	9,735	10,860	2,18	Germin-like protein subfamily 1 member 7 precursor	Croissance cellulaire, division
RN28G10	7,99E-01	0	1,41E-07	62	1,00E+00	1	69	63	62	40	12,402	12,388	12,127	12,666	1,45	putative protein	_
Ge112H01	4,52E-01	3	1,28E-06	53	5,41E-01	7	67	63	53	22	9,788	9,535	9,273	10,101	1,77	putative protein	_
RS54E01	7,92E-01	0	5,42E-09	63	1,00E+00	0	78	63	63	40	12,270	12,343	12,097	12,501	1,32	Portal protein (GPB) (Minor capsid protein B)	Traffic intracellulaire
RN32G07	4,68E-01	3	7,78E-07	60	1,00E+00	0	64	63	60	22	9,492	9,265	8,968	9,835	1,82	Glutamate synthase (NADH) (Sb14) (EC 1.4.1.14)(Picea maria	Métabolisme
RS41A02	8,47E-01	0	6,14E-07	62	1,00E+00	1	64	63	62	17	8,754	8,832	9,194	8,377	1,76	Disease resistance response protein 206	Réponse au Stress
RS38F06	6,09E-01	2	8,45E-07	61	1,00E+00	0	63	63	61	23	8,899	9,086	9,463	8,485	1,97	F14L17.20 protein	Non-classé
RS05C10	9,28E-01	0	6,69E-07	61	1,00E+00	1	64	63	61	20	9,542	9,528	9,287	9,787	1,41	S-adenosylmethionine-dependent methyltransferase T12H1.6	Métabolisme
AS05C03	4,06E-02	14	4,02E-06	48	1,00E+00	0	65	63	48	33	10,330	10,666	10,758	10,171	1,50	Hypothetical 49.0 kDa protein	Non-classé
Ge040H03	3,94E-01	4	7,47E-07	57	1,00E+00	1	66	63	57	40	11,777	11,480	11,172	12,144	1,96	F11M15.7 protein	Transcription
RS42A12	5,13E-01	3	4,64E-07	60	1,00E+00	0	66	62	60	30	9,879	9,665	9,354	10,233	1,84	Unknown	Non-classé
Ge025A06	2,59E-01	5	7,13E-07	52	1,00E+00	5	69	62	52	20	8,933	8,705	8,544	9,140	1,51	putative protein	_
Ge056H01	3,74E-01	4	1,11E-06	55	1,00E+00	3	66	62	55	30	9,702	9,986	10,249	9,382	1,82	Anther specific protein / Pollen specific & Anther specific LAT5	Organisation cellulaire
RN49G02	6,55E-01	2	1,44E-06	60	1,00E+00	1	65	62	60	38	10,800	10,915	10,540	11,156	1,53	putative protein	_
RN19E10	7,16E-02	13	7,82E-06	49	1,00E+00	1	65	62	49	20	9,130	8,798	8,679	9,301	1,54	Hypothetical 40.2 kDa protein	Non-classé
Ge039B07	8,61E-01	0	1,17E-06	57	1,00E+00	5	65	62	57	21	9,067	9,134	9,629	8,558	2,10	phosphate-induced protein	Non-classé
Ge116E03	9,98E-01	0	7,13E-07	62	1,00E+00	0	64	62	62	26	9,277	9,302	9,756	8,818	1,91	Adenosylhomocysteinase (EC 3.3.1.1)	Métabolisme
Ge005G04	7,62E-01	1	1,47E-06	58	1,00E+00	3	63	61	58	30	10,278	10,388	10,749	9,895	1,81	UDP-glucose 4-epimerase (EC 5.1.3.2)	Métabolisme
Ge102A03	1,59E-01	8	1,03E-06	54	1,00E+00	0	67	61	54	39	10,653	10,298	10,108	10,914	1,75	At1g64110/F22C12_22, partial (47%)	Non-classé
Ge040B08	3,20E-01	4	1,97E-07	54	1,00E+00	3	72	61	54	20	9,010	8,750	8,484	9,328	1,79	putative protein	_
RS15G11	1,13E-02	13	2,20E-07	46	1,00E+00	2	76	61	46	28	9,917	9,176	9,051	10,190	2,20	4-hydroxyphenylpyruvate dioxygenase (EC 1.13.11.27)	Métabolisme
Ge058D06	6,88E-01	1	1,42E-06	60	1,00E+00	0	62	61	60	30	9,550	9,690	10,034	9,177	1,81	Tubulin beta chain	Organisation cellulaire
Ge067G04	3,30E-01	5	3,01E-06	53	1,00E+00	2	63	61	53	20	8,975	9,161	9,342	8,757	1,50	NIP3, partial (75%)	Réponse au Stress
RS21G07	7,82E-01	1	2,11E-06	56	1,00E+00	4	62	61	56	17	8,744	8,838	9,174	8,389	1,72	putative protein	_
Ge054B09	9,25E-01	0	7,95E-07	60	1,00E+00	1	64	60	60	30	9,275	9,367	9,799	8,824	1,97	Zinc-finger protein NOLZ1, partial (3%)	Transcription
RN54D04	7,96E-02	10	4,98E-06	49	1,00E+00	1	77	60	49	21	8,825	9,521	9,632	8,483	2,22	Plasma membrane intrinsic polypeptide	Transport
Ge088E12	8,90E-02	11	3,82E-06	48	1,00E+00	1	65	60	48	18	8,782	9,188	9,325	8,564	1,69	Unknown	Non-classé
AN01B08	5,96E-01	2	1,01E-06	58	1,00E+00	0	64	60	58	22	9,370	9,166	8,860	9,716	1,81	Outer membrane lipoprotein blc precursor	Transport
Ge025D03	9,35E-01	0	3,16E-07	60	1,00E+00	0	68	60	60	22	9,825	9,830	9,596	10,058	1,38	Citrate synthase	Métabolisme
Ge037D04	8,09E-01	1	2,54E-06	53	8,07E-01	6	63	60	53	15	8,655	8,588	8,968	8,289	1,60	(1-4)-beta-mannan endohydrolase precursor (EC 3.2.1.78).	Métabolisme
RS40H04	2,84E-02	13	7,95E-07	47	1,00E+00	0	72	60	47	20	8,670	8,910	8,567	8,965	1,32	Alcohol oxidase 2	Métabolisme
RS40C01	1,54E-01	9	9,87E-06	49	1,00E+00	2	61	60	49	20	9,413	9,067	8,894	9,655	1,69	Membrane transporter protein	Transport
RS46F06	1,44E-01	6	8,05E-08	53	1,00E+00	1	75	60	53	23	9,667	9,045	8,731	10,106	2,59	hsr203J	Non-classé
RN09A05	6,05E-01	2	1,76E-06	57	1,00E+00	0	62	60	57	40	12,656	12,541	12,355	12,866	1,43	putative protein	_
Ge061F04	8,17E-01	0	4,01E-07	59	1,00E+00	0	67	60	59	20	9,413	9,306	8,959	9,782	1,77	(2R)-phospho-3-sulfolactate synthase (EC 4.4.-.-)	Non-classé
RS07C01	6,42E-02	13	8,87E-06	46	1,00E+00	0	63	59	46	30	9,702	10,186	10,311	9,480	1,78	Proline-rich protein / Cortical cell delineating protein precursor	Non-classé
RN03D01	3,89E-01	4	5,42E-07	56	1,00E+00	0	68	59	56	11	8,326	8,594	8,858	8,009	1,80	PtaAGP5 (from Pinus taeda)	Non-classé
RS58H06	5,13E-01	2	5,31E-08	57	1,00E+00	0	74	59	57	20	9,024	9,266	8,716	9,525	1,75	Hydroquinone glucosyltransferase (EC 2.4.1.218)	Energie
RN05H02	6,68E-01	2	4,18E-06	57	1,00E+00	0	62	59	57	33	10,743	10,619	11,159	10,177	1,98	60S ribosomal protein L27a	Synthèse de protéines
Ge060D11	9,90E-01	0	2,57E-06	57	1,00E+00	2	61	59	57	18	8,988	8,979	9,415	8,553	1,82	Hypothetical protein yfbG	Métabolisme
GP	9,34E-01	0	2,57E-06	52	6,12E-01	7	65	59	52	14	8,665	8,665	9,086	8,245	1,79	GP	Métabolisme
RS41A09	5,23E-01	2	6,56E-07	56	1,00E+00	0	70	59	56	22	8,958	9,258	9,649	8,520	2,19	Glucan endo-1,3-beta-glucosidase precursor	Organisation cellulaire
Ge124G05	8,53E-01	0	5,57E-06	58	1,00E+00	0	60	59	58	14	8,576	8,673	9,073	8,160	1,88	Unknown	Non-classé
RN41B01	6,75E-01	1	2,84E-06	57	1,00E+00	0	60	59	57	40	13,089	13,032	12,869	13,263	1,31	Glycosyl transferase-like protein	Organisation cellulaire
RN51H01	9,99E-01	0	2,52E-06	58	1,00E+00	0	60	59	58	40	11,566	11,574	11,361	11,777	1,33	26S protease regulatory subunit 4 homolog	Devenir des protéines
RS62A02	9,85E-01	0	2,02E-06	56	1,00E+00	2	62	58	56	28	9,387	9,424	9,819	8,985	1,78	Tubulin beta chain	Organisation cellulaire
RN48A05	7,51E-01	1	2,60E-06	57	1,00E+00	0	61	58	57	40	10,832	10,901	10,651	11,068	1,34	Receptor protein kinase CLAVATA1 precursor (EC 2.7.1.-)	Communication/Transduction du signal
Ge062E09	2,83E-01	6	6,63E-06	52	1,00E+00	0	60	58	52	25	9,829	9,600	9,443	10,032	1,50	Nuc-1 negative regulatory protein preg	Non-classé
RN49F07	9,80E-01	0	2,53E-06	58	1,00E+00	0	61	58	58	24	9,067	9,046	9,688	8,430	2,39	Fasciclin-like arabinogalactan-protein 2	Croissance cellulaire, division
RN46B05	9,17E-01	0	1,69E-06	58	1,00E+00	0	62	58	58	40	12,138	12,168	11,980	12,320	1,27	Pyruvate kinase (EC 2.7.1.40)	Energie
Ge063C05	5,81E-01	2	6,63E-06	52	1,00E+00	3	60	58	52	24	9,187	9,367	9,619	8,899	1,65	putative protein	_
RS34A04	6,81E-01	1	6,27E-06	55	1,00E+00	1	58	57	55	36	10,838	10,710	11,148	10,426	1,65	60S ribosomal protein	Synthèse de protéines
RN42A04	7,09E-01	1	5,09E-06	55	1,00E+00	1	59	57	55	30	10,388	10,269	10,831	9,850	1,97	putative protein	_
RS62F06	4,98E-01	2	1,23E-07	55	1,00E+00	0	76	57	55	39	12,387	12,580	12,156	12,781	1,54	Unknown	Non-classé
RS61A07	6,93E-01	1	6,56E-06	55	1,00E+00	1	58	57	55	18	8,809	8,914	9,191	8,511	1,60	Cell elongation protein DIMINUTO	Croissance cellulaire, division
RS10D05	4,11E-02	12	3,01E-06	44	1,00E+00	1	69	57	44	30	9,720	10,562	10,780	9,334	2,72	putative protein	_
Ge066B06	5,11E-01	3	6,33E-06	53	1,00E+00	0	59	57	53	20	8,929	8,829	8,707	9,072	1,29	Potyvirus VPg interacting protein, partial (55%)	Non-classé
Ge079F01	8,47E-01	0	2,89E-06	56	1,00E+00	0	61	57	56	14	8,545	8,655	8,984	8,193	1,73	putative protein	_
RS14D11	4,33E-01	3	2,54E-07	54	1,00E+00	0	72	57	54	40	12,851	12,981	12,717	13,089	1,29	Myosin V	Traffic intracellulaire
RN30E11	9,35E-01	0	9,90E-06	52	1,00E+00	4	58	57	52	30	9,320	9,337	9,714	8,940	1,71	Omega-6 fatty acid desaturase, endoplasmic reticulum	Métabolisme
RS21C03	7,12E-01	1	3,70E-06	55	1,00E+00	0	60	57	55	40	13,249	13,215	13,075	13,397	1,25	putative protein	_
RS09B05	4,90E-01	3	1,16E-06	52	1,00E+00	1	67	57	52	40	12,332	12,512	12,130	12,678	1,46	Homeobox-leucine zipper protein	Transcription
RS21D07	5,63E-01	3	5,21E-06	54	1,00E+00	0	60	57	54	40	11,290	11,402	11,131	11,539	1,33	SelT-like protein precursor	Non-classé
RS15D05	5,83E-01	3	9,96E-06	53	1,00E+00	1	57	56	53	26	9,404	9,242	9,666	9,013	1,57	Carboxypeptidase C precursor (EC 3.4.16.5)	Devenir des protéines
Ge029D04	5,25E-01	3	1,93E-06	53	1,00E+00	0	65	56	53	20	9,442	9,317	9,145	9,639	1,41	At2g43970/F6E13.10 {Arabidopsis thaliana;} , partial (41%)	Non-classé
Ge086B12	8,98E-01	0	1,20E-06	54	1,00E+00	1	66	55	54	12	8,417	8,571	8,978	7,979	2,00	Snakin-1, partial (67%)	Réponse au Stress
AS03A10	4,60E-01	2	1,84E-07	42	2,09E-02	11	87	55	42	20	8,843	8,813	9,235	8,430	1,75	Hypothetical 38.0 kDa protein (Putative dehydrogenase)	Non-classé
Ge007C03	9,83E-01	0	6,63E-06	45	2,94E-01	10	65	55	45	40	11,765	11,785	12,025	11,521	1,42	Dephospho-CoA kinase (Dephosphocoenzyme A kinase) , par	Métabolisme
RS08H04	5,63E-01	3	5,37E-06	52	1,00E+00	0	61	54	52	40	12,636	12,751	12,492	12,872	1,30	putative protein	_
RN03A11	2,68E-01	5	7,37E-07	49	1,00E+00	0	71	54	49	40	11,554	11,771	11,389	11,892	1,42	Flavonol 3-O-glucosyltransferase (EC 2.4.1.91)	Métabolisme
Ge123H04	9,77E-01	0	5,57E-06	53	1,00E+00	0	60	53	53	40	10,364	10,382	9,966	10,777	1,75	Zinc-finger protein, partial (53%)	Transcription
AS02H08	4,48E-02	9	2,23E-07	44	1,00E+00	0	77	53	44	40	10,751	10,985	10,653	11,037	1,30	putative protein	_
RN28C02	6,11E-01	2	1,96E-06	51	1,00E+00	0	66	52	51	40	11,559	11,745	11,324	11,942	1,54	Phenylalanine ammonia-lyase (EC 4.3.1.5)	Métabolisme
RN18D05	9,04E-01	0	9,87E-06	51	1,00E+00	1	59	52	51	20	9,302	9,361	8,967	9,685	1,64	Unknown	Non-classé
Ge094B08	4,54E-01	3	2,86E-06	47	1,00E+00	2	68	52	47	40	11,490	11,228	10,982	11,789	1,75	Hypothetical protein	Non-classé
Ge098C03	5,80E-01	2	6,14E-07	45	6,12E-01	5	74	52	45	35	10,636	10,988	11,308	10,245	2,09	Ubiquitin	Organisation cellulaire
Ge038A02	6,36E-01	2	7,56E-06	47	1,00E+00	1	63	50	47	11	8,347	8,538	8,773	8,073	1,62	Beta-galactosidase , partial (37%)	Métabolisme
RS66F12	5,84E-01	2	6,27E-06	46	1,00E+00	1	65	49	46	40	11,979	12,141	11,806	12,282	1,39	Unknown	Non-classé
RS29B09	1,42E-01	7	2,38E-06	42	1,00E+00	0	72	49	42	40	11,376	11,646	11,244	11,723	1,39	putative protein	_
RN60F01	7,51E-01	1	2,38E-06	46	1,00E+00	0	69	46	46	40	12,777	12,885	12,625	13,015	1,31	Phospholipase D alpha 1 precursor (EC 3.1.4.4)	Communication/Transduction du signal
RN45G02	5,47E-01	2	1,37E-07	44	1,00E+00	1	78	46	44	40	11,520	11,636	11,351	11,782	1,35	putative protein	_
RS03E11	9,44E-01	0	5,21E-06	43	1,00E+00	1	68	45	43	29	9,780	9,881	10,370	9,271	2,14	putative protein	_
RN73D02	7,81E-01	1	4,78E-06	44	1,00E+00	0	68	45	44	40	13,104	13,229	12,926	13,382	1,37	Cytochrome P450 77A3 (EC 1.14.-.-)	Métabolisme
RN31G10	1,17E-01	7	2,64E-06	37	1,00E+00	0	74	44	37	40	11,764	11,920	11,690	11,962	1,21	putative protein	_
RS07C12	5,78E-01	1	3,06E-06	39	1,00E+00	0	83	41	39	40	11,808	11,826	11,557	12,072	1,43	Hypothetical 58.9 kDa protein	Non-classé
RS49H06	7,45E-01	1	8,38E-06	38	1,00E+00	1	70	40	38	30	10,211	10,242	10,583	9,864	1,65	Pyruvate decarboxylase (EC 4.1.1.1)	Energie
Ge064E04	4,04E-03	12	4,95E-05	18	2,72E-04	21	82	52	21	40	11,784	12,158	12,114	11,754	1,28	Coatomer alpha subunit	Traffic intracellulaire
RN59H07	2,94E-02	12	1,52E-05	31	1,65E-03	23	74	65	31	40	12,446	12,926	12,986	12,290	1,62	Naringenin-chalcone synthase (EC 2.3.1.74)	Métabolisme
Ge106G10	2,48E-03	22	2,07E-05	34	2,57E-01	9	73	64	34	18	9,006	8,341	8,364	9,085	1,65	Phosphoenolpyruvate carboxykinase (EC 4.1.4.49)	Energie
RS50C06	4,36E-01	4	1,01E-05	52	6,63E-01	7	70	63	52	17	8,641	8,326	8,062	8,899	1,79	NADPH:quinone reductase (EC 1.6.5.5)	Energie
Ge058E02	1,16E-01	11	1,26E-05	48	1,00E+00	3	63	62	48	20	9,285	8,912	8,782	9,532	1,68	putative protein	_
RS30A09	1,60E-01	9	2,79E-05	41	2,55E-01	11	65	61	41	28	9,531	9,844	10,092	9,260	1,78	Root cap protein	Transport
Ge127F05_b	3,49E-01	5	2,68E-05	46	5,40E-01	9	61	60	46	#N/A	11,671	11,384	11,176	11,954	1,72	Aluminum-induced protein-like protein, partial (65%)	Réponse au Stress
RN47F03	1,79E-01	8	3,83E-05	39	1,95E-01	12	63	59	39	18	8,809	8,530	8,428	8,967	1,45	Beta-fructofuranosidase (EC 3.2.1.26), cell wall isozyme precu	Métabolisme
Ge048D03	8,13E-01	1	1,38E-05	49	5,01E-01	9	59	58	49	20	8,779	8,724	8,491	9,023	1,45	putative protein	_
RN76H07	1,16E-01	10	1,08E-05	44	1,00E+00	4	64	58	44	20	8,804	9,058	8,589	9,223	1,55	putative protein	_
RS27D07	7,70E-01	1	4,35E-05	50	9,11E-01	7	64	58	50	26	9,813	9,849	9,592	10,036	1,36	Inosine-5'-monophosphate dehydrogenase (EC 1.1.1.2)	Métabolisme
RN40A01	9,06E-01	0	1,16E-05	54	1,00E+00	3	59	58	54	15	8,704	8,762	9,151	8,256	1,86	Glutamate decarboxylase (EC 4.1.1.15)	Métabolisme
RN09E10	6,21E-01	2	1,65E-05	52	1,00E+00	3	59	57	52	17	8,912	8,798	9,190	8,538	1,57	putative protein	_
Ge007G08	2,19E-01	8	1,26E-05	48	1,00E+00	2	60	57	48	31	10,397	10,158	10,635	9,967	1,59	putative protein	_
RN38F06	3,07E-01	6	1,08E-05	50	1,00E+00	1	59	57	50	20	9,500	9,310	9,160	9,688	1,44	Short-chain type dehydrogenase/reductase (EC 1.-.-.-)	Métabolisme
RS22F03	4,64E-01	4	3,17E-05	48	1,00E+00	4	59	57	48	29	9,798	9,608	10,054	9,355	1,62	putative protein	_
Ge108C11	8,60E-01	0	2,21E-05	43	1,81E-01	13	61	56	43	20	9,108	9,162	9,325	8,934	1,31	transitional endoplasmic reticulum ATPase - Arabidopsis thalia	Métabolisme
RN37H08	8,16E-01	1	1,00E-05	54	1,00E+00	2	57	56	54	18	8,917	8,847	9,226	8,552	1,60	Unknown	Non-classé
NXCl_37_F8	7,48E-01	1	1,71E-05	48	8,86E-01	7	58	56	48	34	10,463	10,565	10,854	10,153	1,63	adenosine kinase (AK)	Métabolisme
Ge118H04	2,53E-01	7	7,33E-05	38	2,83E-01	11	60	56	38	25	9,220	9,016	9,416	8,862	1,47	Shikimate 5-dehydrogenase (EC 1.1.1.25)	Métabolisme
RS31C02	1,96E-01	8	3,72E-05	42	1,00E+00	6	60	56	42	25	9,948	9,396	9,182	10,272	2,13	Cytochrome P450 (EC 1.14.-.-), putative	Métabolisme
RN35F03	7,73E-01	1	2,27E-05	39	6,17E-02	16	65	55	39	37	11,565	11,611	12,241	10,926	2,49	Caffeoyl-CoA O-methyltransferase (EC 2.1.1.104)	Métabolisme

Ge071H04	2,95E-01	6	1,62E-05	47	1,00E+00	2	59	55	47	35	10,824	10,516	11,078	10,324	1,69	5-methyltetrahydropteroyltriglutamate--homocysteine methyltra	Métabolisme
RS53C12	8,46E-01	0	1,10E-05	53	1,00E+00	2	57	55	53	20	8,955	8,905	9,301	8,568	1,66	Phosphogluconate dehydrogenase (decarboxylating) (EC 1.1.1	Energie
RS65E10	6,42E-01	2	1,26E-05	51	1,00E+00	2	57	55	51	40	10,454	10,362	10,202	10,633	1,35	putative protein	_
RN09F07	8,51E-01	0	2,72E-05	51	1,00E+00	4	57	55	51	40	11,009	11,037	10,877	11,150	1,21	Presenilin	Croissance cellulaire, division
RN41G01	6,29E-01	2	1,50E-05	52	1,00E+00	1	56	54	52	20	9,052	9,119	8,903	9,255	1,28	putative protein	_
Ge056D01	7,18E-01	1	1,13E-05	53	1,00E+00	0	57	54	53	20	9,040	8,929	8,714	9,278	1,48	Hypothetical protein	Non-classé
AN05A11	1,46E-01	10	2,50E-05	44	1,00E+00	1	60	54	44	30	10,173	9,812	10,445	9,611	1,78	putative protein	_
RS47B12	1,32E-01	12	9,53E-05	42	1,00E+00	0	55	54	42	30	9,978	9,649	9,550	10,142	1,51	Hypothetical protein (DNAJ domain-containing unknown protei	Réponse au Stress
Ge069H08	8,86E-01	0	1,82E-05	52	1,00E+00	2	55	54	52	30	10,171	10,202	10,439	9,928	1,42	putative protein	_
RN47C06	1,32E-01	10	1,65E-05	44	1,00E+00	0	62	54	44	40	11,401	11,673	11,269	11,750	1,40	putative protein	_
RS57E07	1,03E-01	12	5,20E-05	41	1,00E+00	0	58	54	41	27	10,057	9,720	9,623	10,222	1,52	Late embryogenesis abundant protein	Réponse au Stress
RN06D08	9,14E-01	0	1,15E-05	52	1,00E+00	1	65	54	52	20	9,110	8,978	8,635	9,397	1,70	Nodulin	Organisation cellulaire
RN55F05	9,28E-01	0	1,95E-05	51	1,00E+00	2	55	53	51	29	9,480	9,415	9,842	9,065	1,71	Phenylalanine ammonia-lyase (EC 4.3.1.5)	Métabolisme
RN55B08	5,00E-01	4	4,86E-05	44	1,00E+00	5	56	53	44	28	9,576	9,259	9,912	8,987	1,90	Trans-cinnamate 4-monooxygenase (EC 1.14.13.11)	Métabolisme
RS43C07	1,09E-02	19	7,13E-05	33	1,00E+00	0	65	53	33	22	9,015	10,038	10,050	8,798	2,38	putative protein	Non-classé
Ge036E11	5,29E-01	3	5,30E-05	45	1,00E+00	4	54	53	45	18	8,831	9,012	9,265	8,541	1,65	putative protein	_
RS12H05	9,85E-01	0	1,26E-05	52	1,00E+00	1	57	53	52	30	9,941	9,899	10,375	9,473	1,87	Alpha-1,4-glucan-protein synthase (UDP-forming) (EC 2.4.1.11	Métabolisme
Ge061E10	7,70E-01	1	1,15E-05	51	1,00E+00	1	58	53	51	20	9,223	9,277	8,854	9,635	1,72	putative protein	_
RN45B11	5,12E-01	4	4,69E-05	46	1,00E+00	3	55	53	46	16	8,669	8,489	8,934	8,260	1,59	Chloroplast nucleoid DNA binding protein	Devenir des protéines
RS32A02	4,66E-01	4	2,31E-05	48	1,00E+00	0	56	53	48	40	13,220	13,060	12,926	13,385	1,38	Glutathione transferase (EC 2.5.1.18)	Métabolisme
Ge098C07	7,45E-01	1	2,29E-05	49	1,00E+00	3	56	52	49	38	10,915	11,044	11,374	10,559	1,76	Glycine-rich protein	Non-classé
RS40F01	9,51E-01	0	3,64E-05	50	1,00E+00	2	53	52	50	18	8,708	8,693	9,020	8,384	1,55	Unknown	Non-classé
Ge057A05	3,01E-01	6	1,95E-05	46	1,00E+00	0	60	52	46	21	8,849	9,184	9,353	8,613	1,67	Clt1, partial (98%)	Communication/Transduction du signal
RS64B10	5,90E-01	3	6,08E-05	44	1,00E+00	4	54	52	44	20	9,231	9,000	8,737	9,540	1,75	putative protein	_
RS25H03	9,36E-01	0	4,44E-05	50	1,00E+00	1	55	51	50	18	8,527	8,444	8,058	8,887	1,78	putative protein	_
RN43C10	8,74E-01	0	3,06E-05	50	1,00E+00	1	53	51	50	40	11,124	11,095	10,910	11,315	1,32	RuBisCO subunit binding-protein beta subunit	Synthèse de protéines
Ge121D06	7,03E-01	1	2,35E-05	49	1,00E+00	0	55	51	49	23	9,363	9,193	9,664	8,926	1,67	Granule-bound glycogen [starch] synthase	Métabolisme
RS57G02	6,78E-01	1	1,93E-05	41	4,54E-01	8	64	51	41	28	9,390	9,545	9,724	9,180	1,46	Tonoplast intrinsic protein	Transport
RS55H04	6,09E-01	3	4,69E-05	48	1,00E+00	0	52	51	48	20	9,583	9,433	9,231	9,815	1,50	Heat shock protein 101	Réponse au Stress
RN52G02	9,28E-01	0	1,61E-05	50	1,00E+00	1	57	51	50	20	9,350	9,291	9,087	9,565	1,39	Unknown	Non-classé
RS49A07	2,97E-01	7	5,05E-05	44	1,00E+00	0	56	51	44	38	10,597	10,225	10,042	10,854	1,75	Arginine methyltransferase-like protein (EC 2.1.1.-)	Métabolisme
RN62F03	4,98E-01	4	1,75E-05	47	1,00E+00	1	59	51	47	40	12,397	12,259	12,148	12,535	1,31	putative protein	_
Ge067D04	7,73E-01	1	2,03E-05	40	2,71E-01	10	65	51	40	40	12,277	12,214	12,531	11,972	1,47	Abscisic acid insensitive 5 (ABI5) (BZIP protein), partial (17%)	Transcription
Ge035G11	9,98E-01	0	2,33E-05	49	1,00E+00	1	55	51	49	30	9,509	9,564	9,937	9,126	1,75	Wound-induced basic protein	Réponse au Stress
Ge094A09	2,08E-01	5	1,99E-05	29	1,18E-02	17	74	51	29	40	11,203	11,495	11,065	11,574	1,42	PR34, partial (4%)	Non-classé
RS28F02	9,09E-01	0	1,07E-05	50	1,00E+00	1	60	51	50	30	9,575	9,550	10,021	9,108	1,88	Alpha-1,4-glucan-protein synthase (UDP-forming) (EC 2.4.1.11	Métabolisme
NXNV_125_E12	9,63E-01	0	3,84E-05	44	1,00E+00	6	57	51	44	40	11,886	11,927	12,211	11,595	1,53	Aspartate transaminase	Métabolisme
Ge125H03	9,09E-01	0	5,77E-05	39	2,55E-01	12	61	51	39	48	10,388	10,423	10,929	9,875	2,08	Unknown	Non-classé
Ge029H11	9,78E-01	0	9,16E-05	37	1,63E-01	14	60	51	37	40	13,130	13,130	13,348	12,912	1,35	Plasma membrane intrinsic protein (Aquaporin)	Transport
RS57A01	1,51E-01	10	7,41E-05	40	1,00E+00	0	65	51	40	18	8,391	8,600	8,323	8,659	1,26	putative protein	_
Ge043A10	6,69E-01	2	2,69E-05	47	1,00E+00	1	56	50	47	11	8,650	8,542	9,054	8,160	1,86	MERI-5 protein precursor (Endo-xyloglucan transferase) / Bras	Organisation cellulaire
AS03B12	1,86E-01	8	4,79E-05	40	1,00E+00	3	61	50	40	26	9,790	9,364	10,019	9,220	1,74	Metallothionein-like protein EMB30	Métabolisme
RS52G08	4,33E-02	11	6,94E-05	28	8,75E-02	11	76	50	28	31	10,479	10,299	10,614	10,153	1,38	Trans-cinnamate 4-monooxygenase (EC 1.14.13.11)	Métabolisme
Ge050H05	2,73E-01	7	7,91E-05	43	1,00E+00	0	54	50	43	38	10,953	10,704	11,149	10,558	1,51	60S ribosomal protein L39	Synthèse de protéines
RN44G03	6,48E-01	2	5,20E-05	48	1,00E+00	0	52	50	48	37	10,132	10,193	9,990	10,323	1,26	Phospholipase C	Communication/Transduction du signal
AS03A12	2,69E-01	6	1,74E-05	41	1,00E+00	3	64	50	41	32	10,287	9,907	10,541	9,729	1,75	Thimet oligopeptidase (EC 3.4.24.15)	Devenir des protéines
RS47B01	3,94E-01	5	9,16E-05	43	1,00E+00	2	54	50	43	24	9,689	9,344	9,139	9,964	1,77	Hypothetical 34.5 kDa protein	Non-classé
RS57A03	9,80E-01	0	1,93E-05	49	1,00E+00	0	57	50	49	20	8,964	8,968	8,656	9,275	1,54	Chloroplast nucleoid DNA binding protein, putative	Devenir des protéines
RS62C04	9,30E-01	0	3,76E-05	48	1,00E+00	1	54	50	48	18	8,319	8,282	8,062	8,546	1,40	Receptor protein kinase TMK1 precursor	Communication/Transduction du signal
Ge044C12	7,08E-01	1	3,86E-05	43	1,00E+00	5	59	49	43	34	10,529	10,598	11,021	10,092	1,90	Cyclophilin, partial (98%)	Réponse au Stress
RN44B02	9,58E-01	0	9,53E-05	45	1,00E+00	4	51	49	45	26	9,417	9,392	9,665	9,148	1,43	Cell division cycle protein 48 homolog (Valosin containing prot	Croissance cellulaire, division
RN69H07	4,44E-01	5	9,39E-05	43	1,00E+00	1	53	49	43	31	10,420	10,273	10,629	10,094	1,45	Vegetatible incompatibility protein HET-E-1	Non-classé
Ge056D04	7,32E-01	1	3,18E-05	46	1,00E+00	1	56	48	46	37	10,668	10,823	11,067	10,393	1,60	40S ribosomal protein S27	Synthèse de protéines
RN47H03	4,58E-01	5	9,41E-05	43	1,00E+00	0	53	48	43	29	9,164	9,423	9,630	8,905	1,65	Tomato extensin mRNA	Organisation cellulaire
Ge104C09	5,25E-01	4	8,20E-05	44	1,00E+00	0	53	48	44	20	8,899	9,113	9,301	8,668	1,55	Sterol-C-methyltransferase.	Métabolisme
RN03H07	4,60E-01	4	9,58E-05	42	1,00E+00	2	54	48	42	30	9,503	9,872	10,105	9,196	1,88	Alpha galactosyltransferase-like protein	Devenir des protéines
RN49D11	9,46E-01	0	6,94E-05	48	1,00E+00	0	50	48	48	30	9,735	9,794	10,106	9,411	1,62	40S ribosomal protein S16	Synthèse de protéines
RN20D07	9,52E-01	0	3,72E-05	47	1,00E+00	1	55	48	47	20	8,639	8,668	8,315	8,987	1,59	Glutathione transferase (EC 2.5.1.18)	Métabolisme
Ge120A02	7,99E-01	1	8,12E-05	46	1,00E+00	0	51	47	46	17	8,713	8,855	9,153	8,387	1,70	Anter-specific proline-rich protein	Non-classé
Ge036H06	5,88E-01	2	1,74E-05	44	1,00E+00	1	62	46	44	40	12,425	12,547	12,278	12,669	1,31	CTV.15, partial (41%)	Non-classé
RS38H04	9,25E-01	0	2,49E-05	46	1,00E+00	0	58	46	46	40	12,359	12,295	12,135	12,531	1,32	AAA-type ATPase	Energie
Ge091C11	7,46E-01	1	9,41E-05	45	1,00E+00	0	51	46	45	20	9,216	9,112	8,958	9,391	1,35	Putative receptor protein kinase TMK1 precursor (EC 2.7.1.-)	Métabolisme
RS52C04	9,15E-01	0	3,78E-05	43	1,00E+00	2	58	46	43	20	9,306	9,323	9,102	9,523	1,34	Unknown	Non-classé
Ge031B01	3,23E-03	14	2,70E-05	22	7,93E-02	9	79	45	22	28	9,205	9,740	9,673	9,165	1,42	Similarity to late embryogenesis abundant protein, partial (31%	Réponse au Stress
RN27G07	8,39E-01	0	6,72E-05	43	1,00E+00	2	55	45	43	30	9,705	9,829	10,060	9,449	1,53	putative protein	_
RS34F05	5,01E-01	4	9,47E-05	40	1,00E+00	1	56	45	40	37	11,073	10,811	11,303	10,633	1,59	L3 cytoplasmic ribosomal protein	Synthèse de protéines
RS61H03	7,41E-01	1	2,24E-05	41	1,00E+00	3	63	45	41	40	10,562	10,640	10,359	10,827	1,38	Unknown	Non-classé
RS25B12	8,10E-01	1	1,12E-05	43	1,00E+00	0	64	44	43	40	11,635	11,761	11,439	11,932	1,41	Unknown	Non-classé
Ge026E03	6,44E-01	1	1,74E-05	38	1,00E+00	0	73	40	38	18	8,426	8,405	8,277	8,594	1,25	Phosphoenolpyruvate carboxylase , partial (50%)	Energie
RN75C02	2,52E-03	27	3,00E-04	28	1,00E+00	5	63	60	28	36	10,810	10,318	10,857	10,369	1,40	Pyruvate decarboxylase isozyme 2 (EC 4.1.1.1) (PDC)	Energie
RN28G03	1,59E-02	17	6,31E-04	22	1,10E-02	21	67	60	22	26	10,237	9,648	10,317	9,686	1,55	putative protein	_
Ge093C03	3,43E-02	16	1,40E-04	33	2,91E-01	11	62	60	33	18	8,936	8,555	9,060	8,506	1,47	caffeate O-methyltransferase [Pinus taeda]	Métabolisme
Ge122G03	2,22E-03	26	1,53E-04	30	1,00E+00	3	65	59	30	23	9,282	8,925	9,349	8,929	1,34	F3O9.22 protein.	Non-classé
RN56A06	6,77E-03	26	1,00E-03	27	1,00E+00	4	63	57	27	20	8,445	8,906	8,423	8,877	1,37	putative protein	_
AS06G11	1,39E-03	30	9,84E-04	23	1,00E+00	2	63	55	30	40	15,078	14,539	14,646	15,079	1,35	Hypothetical protein DR0254	Non-classé
RN66G01	3,25E-02	19	3,50E-04	33	1,00E+00	2	56	55	33	20	8,696	9,110	9,136	8,587	1,46	MERI-5 protein precursor (Xyloglucan:xyloglucosyl transferase	Croissance cellulaire, division
Ge029B11	1,13E-02	23	8,58E-04	27	1,00E+00	5	58	55	27	32	10,292	9,670	10,360	9,726	1,55	At1g07510/F22G5_9, partial (32%)	Non-classé
Ge039D02	1,28E-01	14	4,52E-04	41	1,00E+00	0	66	55	41	16	8,336	8,220	8,148	8,431	1,22	putative protein	_
RS43C11	1,44E-01	12	1,72E-04	42	1,00E+00	0	62	54	42	28	9,232	9,894	10,066	8,992	2,11	putative protein	_
RN74B04	1,75E-01	10	1,03E-04	41	1,00E+00	4	59	54	41	20	9,012	8,618	8,489	9,187	1,62	Serine/threonine specific protein phosphatase 2C (EC 3.1.3.16	Métabolisme
RN27F05	2,15E-02	20	3,01E-04	32	1,00E+00	1	58	54	32	25	9,727	9,152	9,837	9,157	1,60	Hypothetical 68.3 kDa protein	Non-classé
Ge111D09	6,37E-03	25	5,67E-04	28	1,00E+00	1	59	54	28	33	10,224	9,797	9,831	10,275	1,36	eukaryotic initiation factor 3E subunit {Arabidopsis thaliana;} , p	Transcription
RN60D09	8,78E-01	0	1,04E-04	50	1,00E+00	3	55	53	50	23	9,027	9,103	9,489	8,711	1,72	putative protein	_
Ge133B11	1,77E-02	21	3,80E-04	32	1,00E+00	0	57	53	32	36	10,741	10,190	10,212	10,829	1,53	Wsv162 (WSSV218), partial (23%)	Non-classé
RN61E12	5,93E-01	3	3,60E-04	38	4,31E-01	12	57	52	38	17	8,516	8,357	8,210	8,650	1,36	putative protein	_
AS01F08	5,23E-01	4	4,37E-04	40	8,67E-01	9	54	52	40	30	9,959	10,265	10,648	9,622	2,04	Naringenin-chalcone synthase (EC 2.3.1.74)	Métabolisme
RN11G12	1,88E-01	7	1,80E-04	31	1,28E-01	13	72	52	31	40	12,661	13,309	13,339	12,343	1,99	Naringenin-chalcone synthase (EC 2.3.1.74)	Métabolisme
RN19D06	3,72E-01	6	1,40E-04	43	1,00E+00	3	55	52	43	30	10,091	9,885	10,287	9,702	1,50	putative protein	_
RN27G03	1,54E-01	11	1,12E-04	41	1,00E+00	0	54	52	41	40	11,627	11,990	12,128	11,416	1,64	Annexin	Organisation cellulaire
RN38D02	7,73E-01	1	1,40E-04	35	1,03E-01	16	59	52	35	40	10,647	10,602	10,427	10,831	1,32	putative protein	_
RS58E06	2,83E-02	18	1,91E-04	33	1,00E+00	0	60	51	33	23	9,653	9,158	9,160	9,751	1,51	DnaJ-like protein	Réponse au Stress
RN02G02	6,46E-02	16	8,17E-04	30	1,00E+00	4	54	50	30	28	10,277	9,717	9,704	10,402	1,62	Protein-methionine-S-oxide reductase (EC 1.8.4.6)	Réponse au Stress
RS12A09	2,23E-01	9	1,09E-04	40	1,00E+00	1	55	50	40	28	9,855	9,503	9,258	10,171	1,88	IN2-1 protein	Non-classé
RN49H03	4,11E-02	19	1,00E-03	30	1,00E+00	0	52	50	30	36	10,001	10,492	9,905	10,491	1,50	Germin-like protein	Croissance cellulaire, division
RS59E06	4,25E-01	5	1,86E-04	40	1,00E+00	4	52	49	40	25	8,981	9,300	9,494	8,723	1,71	Nuclear polyadenylated RNA-binding protein NAB4	Transcription
RN73D11	5,52E-02	18	8,78E-04	30	1,00E+00	1	52	49	30	20	9,285	8,856	8,852	9,376	1,44	Syntaxin 71	Traffic intracellulaire
Ge108G08	6,29E-01	2	1,71E-04	39	9,88E-01	7	53	49	39	40	11,965	12,169	12,425	11,668	1,69	Ubiquitin	Organisation cellulaire
Ge084A09	6,68E-01	2	6,97E-04	36	6,34E-01	10	51	48	36	33	10,509	10,684	10,895	10,226	1,59	Peptidyl-prolyl cis-trans isomerase (EC 5.2.1.8) (Cyclophilin)	Devenir des protéines
RN73B08	1,09E-01	13	4,52E-04	32	1,00E+00	3	55	48	32	30	10,252	9,738	10,428	9,666	1,70	putative protein	_
Ge115E03	3,05E-01	8	2,44E-04	40	1,00E+00	0	49	48	40	22	9,661	9,490	9,384	9,801	1,34	putative protein	_
RS33E01	6,75E-01	2	2,88E-04	46	1,00E+00	0	51	48	46	38	11,958	12,018	11,836	12,164	1,26	putative protein	_
Ge097C07	3,57E-01	6	1,53E-04	42	1,00E+00	0	51	48	42	18	8,492	8,315	8,221	8,621	1,32	CoA-thioester hydrolase CHY1 (3-hydroxyisobutyryl-coenzyme	Non-classé
Ge108H10	9,28E-01	0	2,13E-04	43	1,00E+00	5	51	48	43	17	8,647	8,637	8,951	8,366	1,50	Phosphate-induced protein	Non-classé
RS03A03	8,77E-01	0	6,50E-04	46	1,00E+00	1	48	48	46	13	8,649	8,712	8,999	8,358	1,56	Glucan endo-1,3-beta-glucosidase precursor	Organisation cellulaire
RS26C01	9,35E-02	15	6,23E-04	33	1,00E+00	0	51	48	33	28	9,919	9,643	9,604	10,014	1,33	PGPD14 protein (Pollen germination related protein)	Non-classé
AS07F11	4,72E-01	5	1,79E-04	41	1,00E+00	2	51	47	41	20	8,943	8,738	8,590	9,132	1,46	Unknown	Non-classé
RS43C10	6,44E-01	2	1,57E-04	44	1,00E+00	1	48	47	44	29	9,289	9,434	9,686	9,009	1,60	Ethylene responsive element binding factor	Transcription
RS15F01	6,94E-02	14	2,26E-04	33	1,00E+00	0	59	47	33	21	9,327	9,613	9,257	9,626	1,29	Acyltransferase homolog	Métabolisme
RN09F03	3,86E-01	6	1,71E-04	39	1,00E+00	3	54	47	39	26	9,762	9,880	9,663	9,955	1,22	putative protein	_
RS73A01	1,88E-01	11	5,54E-04	36	1,00E+00	0	48	47	36	23	9,692	9,402	9,311	9,841	1,44	Histone H2A	Organisation cellulaire
RS30C11	7,63E-02	16	9,00E-04	30	1,00E+00	1	52	47	30	40	11,630	11,955	11,990	11,530	1,38	Gamma tonoplast intrinsic protein	Transport

RN09B03	8,77E-01	0	1,11E-04	46	1,00E+00	1	49	47	46	40	13,281	13,199	12,944	13,554	1,53	Abscisic stress ripening protein	Réponse au Stress
Ge009E08	2,71E-01	8	5,85E-04	34	1,00E+00	4	50	47	34	37	10,871	10,564	11,056	10,440	1,53	60S ribosomal protein L11	Synthèse de protéines
Ge033C10	7,82E-01	1	1,21E-04	44	1,00E+00	1	50	47	44	20	9,057	8,958	8,735	9,299	1,48	putative protein	_
Ge058A01	9,35E-01	0	1,45E-04	41	1,00E+00	5	52	46	41	40	11,914	11,937	12,275	11,571	1,63	ATP synthase beta chain, mitochondrial precursor	Energie
RN18A04	7,60E-01	1	2,77E-04	41	1,00E+00	3	47	46	41	18	8,797	8,685	9,074	8,430	1,56	Histone H2B	Organisation cellulaire
Ge030G11	7,51E-01	1	1,40E-04	45	1,00E+00	0	48	46	45	40	11,249	11,165	11,013	11,417	1,32	putative protein	_
Ge114B12	9,52E-01	0	1,40E-04	45	1,00E+00	0	47	46	45	30	9,361	9,358	9,746	8,973	1,71	Histone H2A	Organisation cellulaire
RS39C05	9,78E-01	0	1,24E-04	45	1,00E+00	1	49	46	45	30	9,924	9,969	10,248	9,636	1,53	Nuclear polyadenylated RNA-binding protein	Transcription
RS18E12	9,74E-02	14	8,27E-04	30	1,00E+00	2	53	46	30	23	9,678	9,177	9,158	9,798	1,56	Signal peptidase I (EC 3.4.21.89)	Devenir des protéines
RS39G10	2,26E-01	10	6,51E-04	36	1,00E+00	0	51	46	36	20	8,716	8,530	8,454	8,813	1,28	RNA helicase	Synthèse de protéines
RS65D10	2,60E-01	9	7,64E-04	35	1,00E+00	2	47	46	35	18	8,667	8,433	8,360	8,786	1,34	Flavonol 3-O-glucosyltransferase 5 (EC 2.4.1.91)	Métabolisme
NXCI_50_B7	2,69E-01	8	4,68E-04	35	1,00E+00	2	50	46	35	20	9,175	9,420	9,512	9,035	1,39	S-adenosyl-Met synthase	Métabolisme
Ge023B09	9,00E-01	0	1,57E-04	45	1,00E+00	0	47	46	45	40	11,364	11,390	11,191	11,558	1,29	glutathione-dependent formaldehyde dehydrogenase {Arabidop	Métabolisme
RN27F04	8,28E-01	1	1,13E-04	45	1,00E+00	0	49	46	45	20	9,202	9,107	8,897	9,431	1,45	putative protein	_
Ge029C07	9,80E-01	0	1,56E-04	44	1,00E+00	1	48	45	44	21	9,227	9,226	9,572	8,882	1,61	Adenosine kinase (EC 2.7.1.20)	Métabolisme
RN75B08	9,99E-01	0	3,80E-04	43	1,00E+00	2	46	45	43	30	10,197	10,195	10,571	9,822	1,68	L-ascorbate oxidase precursor (EC 1.10.3.3)	Métabolisme
RN46D11	5,56E-02	16	7,45E-04	29	1,00E+00	0	56	45	29	36	10,113	10,300	10,078	10,297	1,16	putative protein	_
Ge020B11	9,66E-01	0	1,33E-04	43	1,00E+00	2	50	45	43	13	8,633	8,595	8,963	8,272	1,61	Dihydrolipoyllysine-residue acetyltransferase component of pyr	Energie
Ge125A05	2,70E-01	9	5,17E-04	36	1,00E+00	0	48	45	36	32	8,422	8,832	8,964	8,207	1,69	Unknown	Non-classé
RN59B05	9,35E-01	0	1,32E-04	45	1,00E+00	0	48	45	45	40	11,344	11,331	11,106	11,572	1,38	putative protein	_
RS62G04	9,52E-01	0	1,53E-04	45	1,00E+00	0	61	45	45	20	8,226	8,212	8,032	8,388	1,28	Carboxypeptidase C (EC 3.4.16.5)	Devenir des protéines
RN09D11	4,31E-01	5	4,70E-04	36	1,00E+00	3	49	45	36	16	8,702	8,455	8,897	8,309	1,50	Chloroplast nucleoid DNA binding protein	Devenir des protéines
Ge005C11	6,70E-01	2	6,64E-04	28	1,41E-01	15	58	45	28	40	12,586	12,700	12,420	12,843	1,34	putative protein	_
Ge075B02	6,11E-01	2	1,03E-04	42	1,00E+00	0	54	45	42	28	9,738	9,545	9,957	9,365	1,51	ATP citrate lyase a-subunit, partial (84%)	Métabolisme
RS36D05	6,74E-01	2	1,74E-04	43	1,00E+00	0	49	45	43	20	9,172	9,083	8,912	9,360	1,36	Iron(III)-zinc(II) purple acid phosphatase precursor	Métabolisme
Ge032B10	9,52E-01	0	1,51E-04	45	1,00E+00	0	48	45	45	36	10,543	10,507	10,877	10,181	1,62	Unknown	Non-classé
RN51F07	4,51E-01	5	2,69E-04	40	1,00E+00	0	49	45	40	40	12,745	12,901	12,624	12,991	1,29	putative protein	_
Ge007F02	4,37E-01	5	2,68E-04	39	1,00E+00	1	50	45	39	40	12,132	12,356	11,952	12,491	1,45	Unknown	Non-classé
RS51F01	9,96E-01	0	1,87E-04	42	1,00E+00	2	49	45	42	30	10,011	10,048	10,404	9,647	1,69	Hypothetical 19.9 kDa protein	Non-classé
RS19C01	7,91E-02	14	8,21E-04	29	1,00E+00	1	55	45	29	22	9,212	8,802	8,793	9,302	1,42	putative protein	_
Ge092D12	2,95E-01	8	6,13E-04	36	1,00E+00	0	51	44	36	20	9,298	9,008	8,904	9,448	1,46	Remorin 1, partial (57%)	Non-classé
NXNV_117_F2	9,48E-01	0	5,10E-04	33	4,13E-01	11	53	44	33	40	13,091	13,128	13,288	12,924	1,29	glyceraldehyde-3-P-dehydrogenase	Métabolisme
RS28B01	8,19E-01	1	2,44E-04	43	1,00E+00	0	50	44	43	28	9,768	9,691	9,549	9,947	1,32	putative protein	_
Ge026H10	5,25E-01	4	5,41E-04	34	1,00E+00	6	55	44	34	24	9,864	9,572	9,426	10,040	1,53	putative protein	_
Ge123E04	8,86E-01	0	2,55E-04	42	1,00E+00	2	47	44	42	14	8,557	8,667	8,986	8,216	1,70	Phenylcoumaran benzylic ether reductase homolog TH1, comp	Transcription
AS02H06	1,85E-01	10	5,98E-04	33	1,00E+00	1	52	44	33	20	8,751	8,557	8,459	8,887	1,35	Naringenin 3-dioxygenase (EC 1.14.11.9)	Métabolisme
RS53G08	3,39E-01	6	1,89E-04	37	1,00E+00	1	55	44	37	40	10,458	10,680	10,332	10,762	1,35	putative protein	_
RN53H11	2,78E-01	8	4,77E-04	35	1,00E+00	1	50	44	35	38	10,272	10,685	10,069	10,805	1,67	Heat shock protein STI1	Réponse au Stress
Ge068A05	8,77E-01	0	4,11E-04	40	1,00E+00	3	45	44	40	28	9,897	9,823	9,590	10,145	1,47	Elongation factor 1-gamma (EF-1-gamma) (eEF-1B gamma), p	Synthèse de protéines
Ge109C07	5,38E-01	4	7,41E-04	37	1,00E+00	2	44	44	37	25	9,298	9,133	9,487	8,976	1,42	Glucose-1-phosphate adenylyltransferase small subunit, chlor	Non-classé
RN65E06	4,46E-01	4	2,26E-04	31	4,99E-01	9	61	44	31	40	11,223	11,388	11,459	11,119	1,27	Tubulin beta chain	Organisation cellulaire
Ge068G01	2,53E-01	9	6,42E-04	34	1,00E+00	0	50	43	34	18	8,595	8,879	8,950	8,467	1,40	Histone H2A	Organisation cellulaire
Ge095A09	9,61E-01	0	5,15E-04	42	1,00E+00	1	44	43	42	19	8,619	8,576	8,285	8,887	1,52	Unknown	Non-classé
NXSI_116_A9	9,36E-01	0	4,22E-04	40	1,00E+00	3	45	43	40	40	11,583	11,628	11,875	11,327	1,46	adenosine kinase (AK)	Métabolisme
RS25A08	6,88E-01	2	2,92E-04	41	1,00E+00	0	47	43	41	40	13,013	12,956	12,823	13,158	1,26	putative protein	_
Ge111E05	3,89E-01	6	8,84E-04	33	1,00E+00	4	49	43	33	32	10,843	10,511	11,056	10,365	1,61	40S ribosomal protein S23	Synthèse de protéines
Ge064E07	6,15E-01	3	7,45E-04	34	1,00E+00	6	48	43	34	29	9,731	9,603	9,882	9,478	1,32	Diphosphate--fructose-6-phosphate 1-phosphotransferase beta	Energie
RN09F01	9,02E-01	0	1,05E-04	43	1,00E+00	0	53	43	43	40	12,138	12,201	11,979	12,347	1,29	Amino acid permease (Amino acid carrier, putative)	Transport
RN37A12	5,55E-01	3	1,79E-04	40	1,00E+00	0	52	43	40	20	9,496	9,418	9,276	9,654	1,30	NADH dehydrogenase (quinone) 23 kDa subunit (1.6.99.5)	Energie
Ge002D03	8,82E-01	0	1,51E-04	42	1,00E+00	1	51	43	42	23	9,270	9,268	9,545	8,994	1,46	Lanatoside 15'-O-acetylesterase precursor, partial (28%)	Métabolisme
Ge062G04	7,57E-01	1	4,37E-04	41	1,00E+00	1	44	43	41	19	8,926	9,041	9,276	8,668	1,52	33 kDa secretory protein-like, partial (13%)	Non-classé
RS43A10	8,85E-01	0	3,66E-04	42	1,00E+00	0	43	43	42	15	8,525	8,468	8,820	8,185	1,55	F12F1.23 protein (Hypothetical 36.7 kDa protein)	Non-classé
RN43A04	8,98E-01	0	1,90E-04	43	1,00E+00	0	48	43	43	20	8,725	8,658	8,457	8,940	1,40	9-cis-epoxycarotenoid dioxygenase	Réponse au Stress
Ge024C08	4,72E-01	4	1,11E-04	37	1,00E+00	2	58	43	37	14	8,056	8,152	7,865	8,324	1,37	Mitochondrial aldehyde dehydrogenase, partial (91%)	Métabolisme
RS45E06	2,97E-01	6	1,15E-04	36	1,00E+00	0	59	43	36	40	11,534	11,728	11,437	11,787	1,27	Unknown	Non-classé
RS60F10	7,74E-01	1	2,11E-04	39	1,00E+00	3	52	43	39	27	9,601	9,691	9,836	9,438	1,32	Aldehyde dehydrogenase (NAD+) (EC 1.2.1.3)	Métabolisme
RN64G04	9,81E-01	0	4,82E-04	37	1,00E+00	6	48	43	37	40	11,093	11,121	11,324	10,884	1,36	40S ribosomal protein S2	Synthèse de protéines
Ge120E04	6,98E-01	2	3,89E-04	38	1,00E+00	3	48	43	38	30	9,878	9,667	10,167	9,420	1,68	Cellulose synthase	Organisation cellulaire
RS20C02	2,86E-01	7	4,70E-04	30	9,36E-01	6	58	43	30	28	9,939	10,543	10,735	9,626	2,16	Leucoanthocyanidin dioxygenase (EC 1.-.-.-) (LDOX)	Métabolisme
RN44F10	5,81E-01	3	7,28E-04	37	1,00E+00	2	48	43	37	25	9,634	9,725	9,497	9,848	1,28	Copper chaperone for Superoxide dismutase	Réponse au Stress
RS65A06	6,77E-01	2	5,36E-04	40	1,00E+00	0	47	43	40	40	11,469	11,415	11,299	11,613	1,24	Oxysterol-binding protein homolog	Transport
Ge002B04	6,75E-01	2	3,20E-04	39	1,00E+00	1	48	42	39	40	10,606	10,709	10,443	10,852	1,33	putative protein	_
Ge097D04	7,22E-01	2	7,19E-04	40	1,00E+00	1	50	42	40	16	8,396	8,229	8,017	8,583	1,48	Unknown	Non-classé
Ge116E05	7,07E-01	2	5,99E-04	39	1,00E+00	2	43	42	39	19	8,701	8,804	8,951	8,534	1,33	Inositolphosphorylceramide-B C-26 hydroxylase	Métabolisme
Ge115B11	6,88E-01	2	5,17E-04	39	1,00E+00	1	44	42	39	30	9,800	9,660	10,074	9,414	1,58	Surfeit 1.	Non-classé
Ge004C06	7,05E-01	2	2,86E-04	40	1,00E+00	0	48	42	40	38	10,510	10,351	10,160	10,734	1,49	catalase 1 [Pinus pinea]	Réponse au Stress
RS41G02	7,73E-01	1	2,13E-04	40	1,00E+00	1	54	42	40	19	8,330	8,314	8,150	8,497	1,27	Structural maintenance of chromosomes (SMC)-like protein	Croissance cellulaire, division
Ge044E10	2,80E-01	8	8,93E-04	33	1,00E+00	0	51	42	33	21	9,307	8,842	8,721	9,485	1,70	1-aminocyclopropane-1-carboxylate oxidase (EC 1.-.-.-)	Réponse au Stress
RS20F02	9,68E-01	0	2,20E-04	42	1,00E+00	0	48	42	42	40	11,411	11,396	11,225	11,584	1,28	Proteasome endopeptidase complex subunit alpha type 6 (EC	Devenir des protéines
Ge051G02	7,62E-01	1	1,55E-04	40	1,00E+00	1	61	42	40	14	8,139	8,198	7,977	8,354	1,30	Eukaryotic peptide chain release factor subunit 1	Synthèse de protéines
RS51G11	8,63E-01	0	1,91E-04	41	1,00E+00	1	51	42	41	40	12,122	12,182	11,993	12,299	1,24	Hyphally regulated protein precursor	Non-classé
RS61G10	5,30E-01	4	7,20E-04	37	1,00E+00	0	43	42	37	29	9,820	9,949	9,695	10,048	1,28	putative protein	_
Ge123B11	9,80E-01	0	4,52E-04	42	1,00E+00	0	42	42	42	28	9,343	9,334	9,621	9,058	1,48	60S ribosomal protein L38	Synthèse de protéines
NXSI_24_H1	9,60E-01	0	5,36E-04	39	1,00E+00	3	44	42	39	40	11,612	11,637	11,963	11,281	1,60	Cellulose synthase	Métabolisme
RN50D10	6,77E-01	2	4,36E-04	39	1,00E+00	1	47	42	39	40	11,549	11,598	11,421	11,716	1,23	putative protein	_
RS63E04	9,35E-01	0	5,39E-04	40	1,00E+00	2	43	42	40	17	8,542	8,575	8,803	8,307	1,41	Inositol-1(or 4)-monophosphatase (EC 3.1.3.25)	Communication/Transduction du signal
RS66F09	4,90E-01	5	4,60E-04	37	1,00E+00	0	48	42	37	23	9,736	9,594	9,440	9,919	1,39	Embryogenesis-associated protein EMB3	Non-classé
Ge041H12	5,41E-01	4	3,60E-04	38	1,00E+00	0	49	42	38	20	8,673	8,588	8,482	8,796	1,24	Ubiquitin carboxyl-terminal hydrolase (EC 3.1.2.	Organisation cellulaire
RN42G03	4,43E-01	5	6,23E-04	34	1,00E+00	2	50	41	34	17	8,503	8,759	8,873	8,338	1,45	putative protein	_
RS52B05	1,86E-01	9	5,99E-04	31	1,00E+00	1	58	41	31	40	10,901	11,046	10,808	11,116	1,24	Unknown	Non-classé
Ge078D03	7,60E-01	1	5,84E-04	39	1,00E+00	0	42	41	39	30	9,707	9,638	9,910	9,449	1,38	40S ribosomal protein S9 (S7)	Synthèse de protéines
RS34G02	9,19E-01	0	3,74E-04	39	1,00E+00	2	48	41	39	30	9,822	9,827	9,648	10,000	1,28	Unknown	Non-classé
RS61F06	9,69E-01	0	2,11E-04	40	1,00E+00	0	50	41	40	40	12,011	12,039	11,839	12,206	1,29	putative protein	_
RS53G02	1,60E-01	9	6,38E-04	28	1,00E+00	3	59	41	28	40	12,238	12,037	11,938	12,377	1,36	26S proteasome non-ATPase regulatory subunit 2	Devenir des protéines
RN61D03	8,48E-01	1	4,70E-04	39	1,00E+00	1	45	41	39	30	9,673	9,595	9,915	9,368	1,46	Malate dehydrogenase, mitochondrial precursor	Métabolisme
RS29H05	9,58E-01	0	4,12E-04	40	1,00E+00	0	45	40	40	20	8,871	8,859	8,656	9,077	1,34	Serine acetyltransferase	Métabolisme
RN06C10	9,70E-01	0	8,81E-04	37	1,00E+00	3	42	40	37	30	10,024	9,993	10,377	9,646	1,66	Thiazole biosynthetic enzyme, chloroplast precursor	Energie
Ge074H10	9,93E-01	0	6,41E-04	39	1,00E+00	2	43	40	39	30	10,021	10,026	10,279	9,767	1,43	60S ribosomal protein	Synthèse de protéines
Ge085H08	7,70E-01	1	8,27E-04	37	1,00E+00	3	43	40	37	26	9,453	9,335	9,682	9,130	1,47	Adenosylhomocysteinase (EC 3.3.1.1)	Métabolisme
RN18A09	9,35E-01	0	8,98E-04	38	1,00E+00	3	41	40	38	19	9,010	9,037	9,293	8,747	1,46	GTP-binding nuclear protein RAN-A1	Traffic intracellulaire
RS37G05	6,29E-01	2	4,41E-04	37	1,00E+00	0	52	40	37	38	12,722	12,855	12,589	12,967	1,30	Hypothetical protein (AT5g01750/T20L15_20)	Non-classé
RS17D12	5,78E-01	3	7,64E-04	36	1,00E+00	0	44	40	36	40	14,340	14,611	13,997	14,901	1,87	small subunit ribosomal RNA gene	Synthèse de protéines
Ge002A10	8,92E-01	0	7,43E-04	39	1,00E+00	0	40	40	39	30	10,249	10,304	10,580	9,962	1,53	60S ribosomal protein	Synthèse de protéines
RN64E07	9,60E-01	0	3,75E-04	39	1,00E+00	0	47	39	39	40	12,522	12,540	12,362	12,696	1,26	S-adenosyl-L-methionine:salicylic acid carboxyl methyltransfer	Métabolisme
RN32B03	8,13E-01	1	7,66E-04	39	1,00E+00	0	40	40	39	20	8,639	8,576	8,404	8,823	1,34	ATAF1-like protein (Fragment)	Non-classé
Ge100E02	9,63E-01	0	6,66E-04	39	1,00E+00	1	61	39	39	18	8,528	8,492	8,324	8,659	1,26	Unknown	Non-classé
Ge110H06	9,85E-01	0	8,70E-04	38	1,00E+00	1	40	39	38	40	13,730	13,739	13,550	13,917	1,29	putative protein	_
Ge075G09	8,78E-01	0	3,61E-04	38	1,00E+00	0	49	39	38	40	12,796	12,872	12,643	13,010	1,29	At5g20110, partial (46%)	Organisation cellulaire
RN75E09	7,57E-01	1	3,82E-04	38	1,00E+00	0	53	39	38	38	11,882	11,958	11,771	12,056	1,22	Kinesin light chain (KLC)	Traffic intracellulaire
RN44C09	9,51E-01	0	7,76E-04	38	1,00E+00	0	46	38	38	15	8,604	8,716	9,090	8,157	1,91	putative protein	_
RS52H08	7,93E-01	1	8,45E-04	37	1,00E+00	1	42	39	37	40	11,202	11,084	10,879	11,431	1,47	Zinc finger protein 216	Transcription
Ge065A11	7,60E-01	1	8,98E-04	37	1,00E+00	0	41	39	37	32	10,190	10,337	10,557	9,940	1,53	ENDOCHITINASE (EC 3.2.1.14)	Réponse au Stress
Ge064F03	7,26E-01	2	8,70E-04	36	1,00E+00	0	43	38	36	30	9,590	9,425	9,806	9,242	1,48	Ubiquinol-cytochrome C reductase iron-sulfur subunit (EC 1.10	Energie
RN60D06	5,81E-01	3	9,99E-04	35	1,00E+00	0	43	39	35	20	9,027	8,866	8,738	9,187	1,37	Serine-rich protein	Non-classé
Ge047H03	8,60E-01	0	3,20E-04	35	1,00E+00	1	54	37	35	30	9,789	9,784	9,982	9,592	1,31	60S ribosomal protein L10	Synthèse de protéines
Ge040B09	7,31E-01	1	5,99E-04	35	1,00E+00	0	49	36	35	21	9,552	9,523	9,382	9,699	1,25	F16A14.7, partial (50%)	Non-classé
NXSI_104_E11	7,02E-01	2	8,68E-04	35	1,00E+00	2	45	39	35	20	9,214	9,307	9,509	8,994	1,43	Ornitine transaminase	Métabolisme
RN30E06	9,25E-01	0	7,57E-04	35	1,00E+00	3	47	38	35	29	9,360	9,315	9,591	9,094	1,41	DNA-binding protein ESCAROLA	Transcription
Ge004F07	7,09E-01	2	9,75E-04	35	1,00E+00	0	49	37	35	29	10,055	10,113	10,412	9,777	1,55	Actin	Organisation cellulaire
RS37H01	4,43E-01	4	1,57E-04	35	1,00E+00	0	59	39	35	40	11,482	11,687	11,364	11,764	1,32	Related to chitinase 3 precursor protein	Réponse au Stress
RN05B11	5,03E-01	4	3,31E-04	34	1,00E+00	0	58	38	34	14	8,464	8,822	8,953	8,219	1,66	Xyloglucan:xyloglucosyl transferase (EC 2.4.1.207)	Croissance cellulaire, division

RN76G05	6,88E-01	1	1,57E-04	34	1,00E+00	0	60	36	34	20	8,515	8,497	8,193	8,822	1,55	putative protein	_
RS56H05	8,95E-01	0	9,83E-04	34	1,00E+00	4	46	38	34	40	10,957	10,957	10,762	11,152	1,31	Lil3 protein	Transcription
RS21A09	7,10E-01	1	1,68E-03	14	5,26E-05	35	76	49	35	40	14,031	14,120	14,169	13,964	1,15	Zinc finger protein 216	Transcription
RN61A02	1,40E-02	16	1,21E-03	19	1,54E-03	28	69	64	28	40	10,696	11,161	10,648	11,117	1,38	putative protein	_
RS15A09	9,35E-02	19	3,76E-03	31	1,00E+00	3	55	53	31	28	9,378	9,812	9,816	9,248	1,48	Peroxidase precursor (EC 1.11.1.7)	Réponse au Stress
RN54B06	1,09E-02	27	4,93E-03	22	1,00E+00	0	51	49	27	40	12,390	12,704	12,379	12,653	1,21	Heterochromatin-like protein	Organisation cellulaire
RS34H11	4,11E-02	20	2,57E-03	25	1,00E+00	4	51	49	25	28	9,872	9,503	9,925	9,524	1,32	Calmodulin	Communication/Transduction du signal
RS65C05	2,27E-02	21	1,48E-03	26	1,00E+00	2	59	49	26	21	9,258	8,914	8,948	9,279	1,26	putative protein	_
RN49C12	1,01E-01	14	7,24E-03	20	2,57E-01	14	51	49	20	37	10,815	11,267	11,259	10,732	1,44	Metallothionein-like protein EMB30	Métabolisme
RS42B09	6,86E-01	2	1,01E-03	31	2,58E-01	15	50	47	31	40	11,672	11,855	12,186	11,305	1,84	Naringenin-chalcone synthase (EC 2.3.1.74)	Métabolisme
Ge085E07	1,34E-01	13	2,00E-03	28	1,00E+00	5	49	46	28	19	9,064	8,664	9,174	8,634	1,45	ENDOGLUCANASE	Organisation cellulaire
RN17G03	2,00E-01	11	2,14E-03	29	1,00E+00	6	46	46	29	36	10,313	9,940	9,860	10,467	1,52	Expansin-like precursor	Croissance cellulaire, division
AS06D09	4,90E-02	16	1,09E-03	25	1,00E+00	5	59	46	25	32	9,590	10,443	10,416	9,447	1,96	Metallothionein-like protein	Métabolisme
RS31C10	6,27E-01	3	1,84E-03	38	1,00E+00	5	53	46	38	19	8,242	8,304	8,084	8,422	1,26	putative protein	_
RN73B01	1,60E-01	12	3,51E-03	25	8,45E-01	8	48	45	25	18	8,633	8,349	8,317	8,722	1,32	Unknown	Non-classé
RN56B05	8,61E-01	1	1,06E-03	41	1,00E+00	3	49	45	41	16	8,179	8,207	7,958	8,375	1,34	putative protein	_
RS66D06	1,09E-01	15	2,17E-03	28	1,00E+00	1	47	45	28	22	9,191	8,852	9,297	8,814	1,40	putative protein	_
RS55B01	5,18E-02	17	1,59E-03	26	1,00E+00	1	55	44	26	21	9,429	9,061	9,088	9,476	1,31	putative protein	_
RS06C01	1,94E-02	21	2,99E-03	22	1,00E+00	0	56	44	22	25	9,772	9,100	9,797	9,209	1,50	P0022F10.3 protein	Non-classé
AS03C11	7,57E-01	1	1,23E-03	29	3,10E-01	13	51	44	29	38	10,774	10,691	10,938	10,544	1,31	Endopeptidase Clp subunit (EC 3.4.21.92)	Devenir des protéines
RS39F01	4,21E-01	6	1,89E-03	30	1,00E+00	7	45	43	30	18	8,559	8,771	8,860	8,427	1,35	GPI-anchored protein	Non-classé
RN63G04	4,31E-01	6	1,20E-03	36	1,00E+00	1	49	43	36	25	9,910	9,681	9,521	10,061	1,45	Ubiquitin conjugating enzyme spm2	Organisation cellulaire
AS03G06	5,41E-01	4	2,58E-03	27	4,53E-01	12	47	43	27	30	9,765	9,977	9,536	10,163	1,54	Carbonate dehydratase (EC 4.2.1.1)	Métabolisme
RN35A11	4,31E-02	19	2,85E-03	24	1,00E+00	1	53	43	24	22	9,505	8,990	9,551	9,048	1,42	putative protein	_
RN74F01	4,64E-01	6	1,10E-03	36	1,00E+00	1	45	43	36	20	8,788	8,559	8,433	8,936	1,42	ADP,ATP carrier protein (ADP/ATP translocase)	Transport
RN63F04	3,67E-01	7	3,43E-03	28	1,00E+00	7	43	42	28	26	9,698	9,437	9,853	9,334	1,43	Unknown	Non-classé
RN64C02	3,67E-01	7	1,56E-03	32	1,00E+00	2	45	41	32	30	9,942	10,084	9,839	10,159	1,25	Chaperone protein dnaJ	Réponse au Stress
RS16C11	2,41E-01	10	3,39E-03	27	1,00E+00	5	45	41	27	20	9,144	8,838	8,796	9,248	1,37	putative protein	_
RS08A05	9,99E-01	0	1,23E-03	33	1,00E+00	8	44	41	33	30	9,823	9,794	10,107	9,516	1,51	Hypothetical 12.0 kDa protein	Non-classé
Ge068A11	9,34E-01	0	1,09E-03	34	1,00E+00	6	44	41	34	15	8,706	8,726	9,009	8,419	1,51	putative protein	_
RS62D03	6,09E-01	3	1,73E-03	32	1,00E+00	5	42	41	32	32	10,014	9,839	9,693	10,195	1,42	Unknown	Non-classé
RS40G02	1,16E-01	13	1,69E-03	26	1,00E+00	2	54	41	26	27	9,754	9,240	9,245	9,853	1,52	Peroxidase precursor (EC 1.11.1.7)	Réponse au Stress
RN79C01	8,52E-02	16	3,34E-03	24	1,00E+00	1	50	41	24	19	9,157	8,683	9,225	8,710	1,43	Disease resistance response protein 206	Réponse au Stress
RS08E10	2,46E-01	10	7,97E-03	24	1,00E+00	7	41	41	24	36	10,294	10,029	9,993	10,383	1,31	Alternative oxidase 1, mitochondrial precursor	Energie
Ge084D05	8,37E-01	1	1,02E-03	34	1,00E+00	5	44	40	34	40	12,004	11,874	12,289	11,615	1,60	Pinus taeda water deficit inducible protein LP3	Réponse au Stress
RN78G11	9,15E-01	0	1,02E-03	35	1,00E+00	5	43	40	35	25	9,201	9,255	9,543	8,903	1,56	putative protein	_
RS33H04	8,47E-01	0	7,66E-03	18	4,23E-02	22	56	40	22	40	13,509	13,493	13,333	13,672	1,26	putative protein	_
RN55B05	8,41E-01	1	1,06E-03	38	1,00E+00	0	43	39	38	20	8,703	8,741	8,447	8,990	1,46	putative protein	_
Ge051A06	9,96E-01	0	1,02E-03	38	1,00E+00	1	39	38	38	39	10,897	10,908	11,183	10,620	1,48	putative protein	_
Ge055H04	8,95E-01	0	1,06E-03	37	1,00E+00	0	39	38	37	20	8,852	8,902	8,562	9,181	1,54	Heat Shock 70 kd protein	Réponse au Stress
Ge062E12	1,00E+00	0	1,06E-03	37	1,00E+00	0	40	37	37	20	8,777	8,771	8,578	8,972	1,31	RNA-binding protein RZ-1, partial (9%)	Transcription
RS47B09	9,25E-01	0	1,37E-03	36	1,00E+00	0	38	37	36	29	9,570	9,549	9,873	9,251	1,54	Anter-specific proline-rich protein / Lanatoside 15'-O-acetyleste	Métabolisme
RS49D06	9,02E-01	0	1,63E-03	36	1,00E+00	0	37	36	36	32	10,457	10,502	10,709	10,241	1,38	Polyphosphoinositide binding protein	Organisation cellulaire
Ge098F09	6,52E-01	3	1,40E-03	36	1,00E+00	0	39	38	36	15	8,410	8,520	8,659	8,249	1,33	high mobility group protein MNB1b - maize (fragment) {Zea ma	Non-classé
Ge007A07	9,28E-01	0	1,46E-03	36	1,00E+00	1	38	37	36	30	9,751	9,703	10,019	9,445	1,49	Prohibitin	Croissance cellulaire, division
Ge069A05	9,82E-01	0	1,17E-03	35	1,00E+00	2	41	37	35	24	9,244	9,265	9,526	8,978	1,46	Cellulose synthase	Organisation cellulaire
Ge024F11	9,28E-01	0	1,49E-03	35	1,00E+00	1	39	37	35	38	10,678	10,617	10,944	10,363	1,50	Profilin	Traffic intracellulaire
RS16A04	5,96E-01	3	1,34E-03	35	1,00E+00	0	40	39	35	16	8,480	8,642	8,784	8,306	1,39	Caffeate O-methyltransferase (EC 2.1.1.68)	Métabolisme
RS66E03	8,53E-01	1	1,48E-03	35	1,00E+00	1	39	36	35	19	8,621	8,550	8,399	8,786	1,31	Unknown	Non-classé
Ge074D02	6,92E-01	2	1,09E-03	35	1,00E+00	2	43	39	35	28	9,450	9,537	9,723	9,246	1,39	Similarity to ABA induced plasma membrane protein, partial (5	Réponse au Stress
RN54G03	7,73E-01	1	2,77E-03	34	1,00E+00	1	43	36	34	32	10,259	10,397	9,820	10,805	1,98	Transaldolase (EC 2.2.1.2)	Métabolisme
Ge092E09	7,32E-01	2	1,81E-03	34	1,00E+00	2	38	38	34	15	8,427	8,320	8,641	8,127	1,43	serine protease	Métabolisme
Ge029E02	7,76E-01	1	1,84E-03	34	1,00E+00	3	39	38	34	18	8,926	8,811	9,155	8,605	1,46	subtilisin-like proteinase - Arabidopsis thaliana {Arabidopsis th	Devenir des protéines
Ge006B03	8,54E-01	1	2,48E-03	34	1,00E+00	0	35	35	34	30	9,838	9,924	10,188	9,557	1,55	acetone-cyanohydrin lyase [imported] - Arabidopsis thaliana, p	Non-classé
Ge050B02	7,51E-01	2	3,67E-03	34	1,00E+00	1	51	36	34	2	7,746	7,860	7,648	7,929	1,22	Unknown	Non-classé
RS22E05	9,99E-01	0	1,11E-03	34	1,00E+00	2	44	36	34	40	10,686	10,657	10,515	10,833	1,25	putative protein	_
Ge063D06	6,56E-01	3	3,68E-03	34	1,00E+00	1	40	37	34	28	9,892	9,761	9,605	10,063	1,37	Sec13p, partial (90%)	Traffic intracellulaire
RN73A05	8,88E-01	0	7,24E-03	34	1,00E+00	0	42	34	34	17	8,181	8,242	8,101	8,317	1,16	putative protein	_
RS36C02	5,81E-01	4	2,26E-03	34	1,00E+00	0	42	37	34	18	8,419	8,250	8,121	8,560	1,36	Protein kinase (EC 2.7.1.-)	Métabolisme
RN70H03	9,66E-01	0	2,51E-03	34	1,00E+00	1	36	35	34	26	9,676	9,657	9,380	9,957	1,49	putative protein	_

Tableau Ax-5 : Classement des gènes dont l'expression est significativement régulée par catégories fonctionelles pour les expériences de Stress Court et de Stress Long.
P,p = effet provenance significatif, C,c = effet condition significatif, I,i = interaction PC significative. Lettre minuscule = p-value comprise entre 0.01 et 0.001, Lettre majuscule = p-value <0.001, M+ / F+ =surexpression Maroc / France ; S+ / S- = sur/sous-expression au cours du stress

Communication/Transduction du signal

Accession	Identifiant	Annotation	Effet	*sens*
Stress Court				
BX682915	Ge123D11	14-3-3 protein	P	*M+*
BX249546	Ge025C11	Calmodulin	P	*M+*
CR392472	RN32G05	Calmodulin-related protein 2	P	*F+*
BX254449	Ge098G01	Phospholipase D alpha 1 (PLD alpha 1) (Choline phosphatase	P	*M+*
BX249833	Ge029C02	Ras-like small monomeric GTP-binding protein, partial (65%)	P	*M+*
CR392786	RN48A05	Receptor protein kinase CLAVATA1 precursor (EC 2.7.1.-)	P	*F+*
BX677537	RN40H04	Receptor protein kinase precursor	P	*F+*
BX682453	Ge116H01	Receptor-like protein kinase, partial (70%)	C	*S+*
CR392434	RN31H05	Seven transmembrane protein 1	C	*S+*
CR393564	RN61G10	Sorcin	P	*F+*
Stress Long				
BX679850	RS34H11	Calmodulin	c	*S-*
BX251900	Ge057A05	Clt1, partial (98%)	C	*S-*
BX255369	Ge005A06	Glycine-rich RNA-binding protein, complete	C	*S+*
BX681723	RS63E04	Inositol-1(or 4)-monophosphatase (EC 3.1.3.25)	C	*S-*
BX677790	RN44G03	Phospholipase C	C	*S+*
CR393473	RN60F01	Phospholipase D alpha 1 precursor (EC 3.1.4.4)	C	*S+*
CR392786	RN48A05	Receptor protein kinase CLAVATA1 precursor (EC 2.7.1.-)	C	*S+*
BX681632	RS62C04	Receptor protein kinase TMK1 precursor	C	*S+*
BX682453	Ge116H01	Receptor-like protein kinase, partial (70%)	C	*S+*
CR392434	RN31H05	Seven transmembrane protein 1	C	*S+*

Croissance cellulaire, division

Accession	Identifiant	Annotation	Effet	*sens*
Stress Court				
BX681545	RS61A07	Cell elongation protein DIMINUTO	C	*S-*
BX251681	Ge053H04	Germin(Oxalate oxidase)-like protein, partial (24%)	Pi	
CR392910	RN49H03	Germin-like protein	P	*M+*
CR393563	RN61G09	MERI-5 protein precursor (Xyloglucan:xyloglucosyl transferase	C	*S-*
BX681360	RS57F04	MERI-5 protein precursor (Xyloglucan:xyloglucosyl transferase	P	*F+*
BX677507	RN40E01	Nuclear scaffold-like protein p76	P	*M+*
BX255545	Ge007A07	Prohibitin	C	*S-*
Stress Long				
BX677740	RN44B02	Cell division cycle protein 48 homolog (Valosin containing prote	C	*S-*
BX681545	RS61A07	Cell elongation protein DIMINUTO	C	*S-*
AL751104	RS05B06	DNA polymerase III	C	*S+*
BX677108	RN17G03	Expansin-like precursor	c	*S+*
BX250683	Ge040A04	Fasciclin-like arabinogalactan protein FLA8	C	*S-*
CR392893	RN49F07	Fasciclin-like arabinogalactan-protein 2	C	*S-*
CR392910	RN49H03	Germin-like protein	C	*S+*
CR393282	RN54H02	Germin-like protein subfamily 1 member 7 precursor	C	*S+*
CR393887	RN66G01	MERI-5 protein precursor (Xyloglucan:xyloglucosyl transferase	C	*S-*
CR393563	RN61G09	MERI-5 protein precursor (Xyloglucan:xyloglucosyl transferase	C	*S-*
BX667570	RN09F07	Presenilin	C	*S+*
BX255545	Ge007A07	Prohibitin	c	*S-*
BX680304	RS41G02	Structural maintenance of chromosomes (SMC)-like protein	C	*S+*
AL750591	RN05B11	Xyloglucan:xyloglucosyl transferase (EC 2.4.1.207)	C	*S-*
BX251676	Ge053G11	Xyloglucan:xyloglucosyl transferase (EC 2.4.1.207)	C	*S-*

Devenir des protéines

Accession	Identifiant	Annotation	Effet	*sens*
Stress Court				
BX681114	RS53G02	26S proteasome non-ATPase regulatory subunit 2	P	*M+*
BX681080	RS53A02	Acetylornithine deacetylase (EC 3.5.1.16)	PC	*S+M+*
BX678807	RS15D05	Carboxypeptidase C precursor (EC 3.4.16.5)	P	*M+*
BX252908	Ge074G03	Cathepsin B-like cysteine proteinase, partial (68%)	PC	*S+M+*
BX681068	RS52G06	Cysteine proteinase RD21A precursor (EC 3.4.22.-)	P	*M+*
AL749972	AS03C11	Endopeptidase Clp subunit (EC 3.4.21.92)	P	*M+*
BX679916	RS36C01	Metallophosphatase	PC	*S-M+*
AL750449	RN02F06	Peptidylprolyl isomerase (EC 5.2.1.8)	P	*M+*
BX252329	Ge062H12	Proteasome subunit beta type 5 precursor (20S proteasome su	P	*F+*
BX252055	Ge059B02	subtilisin-like serine proteinase - Arabidopsis thaliana {Arabido	P	*M+*
Stress Long				
CR393041	RN51H01	26S protease regulatory subunit 4 homolog	C	*S+*
BX681114	RS53G02	26S proteasome non-ATPase regulatory subunit 2	C	*S+*
AL750564	RN03H07	Alpha galactosyltransferase-like protein	C	*S-*
BX681672	RS62G04	Carboxypeptidase C (EC 3.4.16.5)	C	*S+*
BX678807	RS15D05	Carboxypeptidase C precursor (EC 3.4.16.5)	C	*S-*
BX677823	RN45B11	Chloroplast nucleoid DNA binding protein	C	*S-*
BX667552	RN09D11	Chloroplast nucleoid DNA binding protein	C	*S-*
BX681320	RS57A03	Chloroplast nucleoid DNA binding protein, putative	C	*S+*
AL749972	AS03C11	Endopeptidase Clp subunit (EC 3.4.21.92)	c	*S-*
BX253484	Ge084A09	Peptidyl-prolyl cis-trans isomerase (EC 5.2.1.8) (Cyclophilin)	C	*S-*
BX679155	RS20F02	Proteasome endopeptidase complex subunit alpha type 6 (EC :	C	*S+*
BX679023	RS18E12	Signal peptidase I (EC 3.4.21.89)	C	*S+*
BX249855	Ge029E02	subtilisin-like proteinase - Arabidopsis thaliana {Arabidopsis tha	c	*S-*
AL749954	AS03A12	Thimet oligopeptidase (EC 3.4.24.15)	C	*S-*

Energie

Accession	Identifiant	Annotation	Effet	*sens*
Stress Court				
BX681296	RS56F09	[Pyruvate dehydrogenase (Lipoamide)]-phosphatase (EC 3.1.3.	P	*F+*
BX249881	Ge029G11	ATP synthase gamma chain	P	*M+*
BX678708	RS14B03	Basic blue protein (Plantacyanin)	PC	*S-M+*
BX681437	RS59C04	Electron transfer flavoprotein-ubiquinone oxidoreductase	P	*F+*
BX680905	RS50E05	Fumarate hydratase, mitochondrial precursor (EC 4.2.1.2)	C	*S+*
BX681841	RS65F03	NADPH dehydrogenase (EC 1.6.99.1)	P	*F+*
AL750498	RN03B11	NADPH dehydrogenase (EC 1.6.99.1) / 12-OXOPHYTODIENO	C	*S-*
BX680890	RS50C06	NADPH:quinone reductase (EC 1.6.5.5)	C	*S+*
BX681554	RS61B07	NADPH:quinone reductase (EC 1.6.5.5)	P	*F+*
CR394186	RN74H06	Phytocyanin	P	*F+*
AL749979	AS03D06	Thioredoxin	P	*M+*
BX255712	Ge008H07	Thioredoxin H-type	P	*M+*
AL750459	RN02G04	Ubiquinol-cytochrome C reductase	P	*F+*
BX678629	RS12C06	Ubiquinol-cytochrome C reductase complex 7.8 kDa protein	P	*F+*
Stress Long				
BX680080	RS38H04	AAA-type ATPase	C	*S+*
BX678396	RS08E10	Alternative oxidase 1, mitochondrial precursor	c	*S+*
BX251974	Ge058A01	ATP synthase beta chain, mitochondrial precursor	C	*S-*
CR392217	RN28D05	Basic blue protein (Cusacyanin) (Plantacyanin)	C	*S-*
BX678708	RS14B03	Basic blue protein (Plantacyanin)	CI	
BX678416	RS08G10	Chlorophyll A-B binding protein, chloroplastic	C	*S+*
BX249115	Ge020B11	Dihydrolipoyllysine-residue acetyltransferase component of pyr	C	*S-*
BX252436	Ge064E07	Diphosphate--fructose-6-phosphate 1-phosphotransferase beta	C	*S-*
AL749783	AS01A02	H(+)-transporting two-sector ATPase delta subunit (EC 3.6.3.14	C	*S+*
BX681414	RS58H06	Hydroquinone glucosyltransferase (EC 2.4.1.218)	C	*S+*
CR392718	RN37A12	NADH dehydrogenase (quinone) 23 kDa subunit (1.6.99.5)	C	*S+*
AL750498	RN03B11	NADPH dehydrogenase (EC 1.6.99.1) / 12-OXOPHYTODIENO	PC	*S-F+*
BX680890	RS50C06	NADPH:quinone reductase (EC 1.6.5.5)	C	*S+*
AL749654	AN05E01	Oxygen-evolving enhancer protein 1, chloroplast precursor	C	*S+*
BX251416	Ge049C07	Oxygen-evolving enhancer protein 3-1, chloroplast precursor	C	*S+*
BX254884	Ge106G10	Phosphoenolpyruvate carboxykinase (EC 4.1.1.49)	pC	*S+F+*
BX249884	Ge029H02	Phosphoenolpyruvate carboxykinase (EC 4.1.1.49)	PC	*S+F+*
BX249635	Ge026E03	Phosphoenolpyruvate carboxylase , partial (50%)	C	*S+*
BX681089	RS53C12	Phosphogluconate dehydrogenase (decarboxylating) (EC 1.1.1	C	*S-*
AL749666	AN05G02	Photosystem I reaction center subunit N, chloropla	C	*S+*
BX680864	RS49H06	Pyruvate decarboxylase (EC 4.1.1.1)	C	*S-*
CR394212	RN75C02	Pyruvate decarboxylase isozyme 2 (EC 4.1.1.1) (PDC)	pC	*S-F+*
BX677890	RN46B05	Pyruvate kinase (EC 2.7.1.40)	C	*S+*
AL750697	RN06C10	Thiazole biosynthetic enzyme, chloroplast precursor	C	*S-*
AL749979	AS03D06	Thioredoxin	C	*S+*
BX252988	Ge076B04	Thioredoxin H-type (TRX-H)	PC	*S+F+*
BX252441	Ge064F03	Ubiquinol-cytochrome C reductase iron-sulfur subunit (EC 1.10.	C	*S-*

Métabolisme

Accession	Identifiant	Annotation	Effet	*sens*
Stress Court				
BX681748	RS63G11	2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase (EC 4.6.	PC	*S-M+*
CR392614	RN35E03	Acid phosphatase	PC	*S-M+*
CR392259	RN30A11	Acid phosphatase precursor 1 (EC 3.1.3.2)	PC	*S-M+*
BX679759	RS33B10	Adenylate kinase (EC 2.7.4.3)	PC	*S-M+*
BX677894	RN46B11	Alcohol dehydrogenase (EC 1.1.1.1)	P	*F+*
BX252271	Ge062B09	Alcohol dehydrogenase class III (Glutathione-dependent forma	P	*M+*
BX681520	RS60F10	Aldehyde dehydrogenase (NAD+) (EC 1.2.1.3)	P	*M+*
BX681821	RS65D03	Aldehyde dehydrogenase (NAD+) (EC 1.2.1.3)	P	*F+*
AL750794	RN07D07	Aldehyde dehydrogenase (NAD+) (EC 1.2.1.3)	P	*F+*
CR393484	RN60G02	Aldose 1-epimerase precursor (EC 5.1.3.3)	P	*F+*
CR393571	RN61H07	Aldose 1-epimerase precursor (EC 5.1.3.3)	P	*F+*
BX254971	Ge108E04	Alkaline alpha galactosidase II, partial (84%)	PC	*S+M+*
BX253047	Ge077B10	Aminotransferase	PC	*S+M+*
BX252563	Ge066E08	Aspartate aminotransferase, chloroplast precursor	P	*M+*
AW870208	NXNV_125_E12	Aspartate transaminase	P	*M+*
CR392379	RN31A10	Beta-glucosidase precursor (EC 3.2.1.21)	C	*S-*
BX682419	Ge116C12	Betaine-aldehyde dehydrogenase (EC 1.2.1.8)	P	*F+*
BX249794	Ge027G10	Beta-ketoacyl-ACP synthetase I-2, partial (91%)	P	*F+*
BX667535	RN09C02	Carbon catabolite derepressing protein kinase	PC	*S+M+*
BX676888	RN13E06	Carbonate dehydratase (EC 4.2.1.1)	C	*S-*
AL750008	AS03G06	Carbonate dehydratase (EC 4.2.1.1)	P	*M+*
BE123499	NXNV_148_H6	Cellulose synthase	P	*M+*
BX681531	RS60G10	Chalcone-flavonone isomerase	P	*F+*
BX678890	RS16F01	Cholinephosphate cytidylyltransferase B (EC 2.7.7.15)	P	*M+*
BX681204	RS55C05	Cysteine synthase (EC 2.5.1.47)	P	*F+*
CR393320	RN58E01	Cytochrome P450 (EC 1.14.-.-)	C	*S-*
BX254251	Ge095F04	cytosolic aldehyde dehydrogenase {Oryza sativa (japonica cultiv	PC	*S+M+*
BX255565	Ge007C03	Dephospho-CoA kinase (Dephosphocoenzyme A kinase) , par	P	*M+*
BX680021	RS38C03	Dihydroflavonol-4-reductase (EC 1.1.1.219) (DFR)	P	*M+*
BX681612	RS61H11	dUTP pyrophosphatase (EC 3.6.1.23)	P	*F+*
BX677769	RN44E06	Enoyl-CoA hydratase (EC 4.2.1.17)	P	*F+*
CR392825	RN48H03	Epimerase/dehydratase	C	*S-*
AL749773	AN07G08	Esterase D	P	*M+*
BX252727	Ge069G09	Esterase precursor (Early nodule-specific protein homolog) (La	PC	*S-M+*
BX677655	RN42G05	Flavonoid 3',5'-hydroxylase (EC 1.14.-.-)	C	*S-*
BF049732	NXCI_106_D10	Fumarase	P	*M+*
AL751292	RS07B10	Gamma-lyase	C	*S+*
BX254083	Ge093A12	Geranyl diphosphate synthase, complete	PC	*S-M+*
BE761959	NXCI_75_C7	Glutamate synthase	P	*M+*
CR393516	RN61B10	Glutamate-1-semialdehyde 2,1-aminomutase (EC 5.4.3.8)	P	*F+*
BX681127	RS54B10	Glutathione transferase (EC 2.5.1.18)	P	*F+*
BX681235	RS55H11	Glutathione transferase (EC 2.5.1.18)	P	*M+*
CR394335	RN76H08	Glutathione transferase (EC 2.5.1.18)	PC	*S-M+*
BX682765	Ge121D06	Granule-bound glycogen [starch] synthase	P	*M+*
BX254883	Ge106G06	Histidine ammonia-lyase (Histidase) , partial (4%)	P	*F+*
BX681985	RS67F12	Isoflavone reductase homolog (EC 1.3.1.-)	PC	*S-M+*
CR392581	RN35A05	Isopentenyl-diphosphate delta-isomerase II (EC 5.3.3.2)	C	*S-*
AL750713	RN06E03	L-ascorbate oxidase precursor (EC 1.10.3.3)	P	*F+*
CR394206	RN75B08	L-ascorbate oxidase precursor (EC 1.10.3.3)	PC	*S-M+*
BF609706	NXSI_48_D6	malate dehydrogenase	P	*M+*
BX255766	Ge009E07	metallothionein	PC	*S+M+*
BX252844	Ge071F04	metallothionein-like protein	PC	*S+M+*
BG039825	NXSI_104_E11	Ornitine transaminase	P	*M+*
BX253454	Ge083F12	Patatin	P	*F+*
BX682063	RS69E03	Phenylalanine ammonia-lyase (EC 4.3.1.5)	P	*F+*
BX682869	Ge122H09	Pinene synthase (EC 4.2.3.14)	PC	*S-M+*
CR393796	RN65E04	Pinene synthase (EC 4.2.3.14)	PC	*S-M+*
BX681938	RS67A12	Protein kinase (EC 2.7.1.-)	PC	*S-M+*
CR393742	RN64E07	S-adenosyl-L-methionine:salicylic acid carboxyl methyltransfera	P	*F+*
BG039488	NXSI_99_F10	S-adenosyl-Met decarboxylase	P	*M+*
BE643924	NXCI_50_B7	S-adenosyl-Met synthase	P	*M+*
AL751154	RS05F10	Serine/threonine specific protein phosphatase 2C (EC 3.1.3.16)	P	*F+*
BX681087	RS53B06	SUR2 protein (Syringomycin response protein 2)	P	*F+*
BX680092	RS39A02	T14P8.4 protein	P	*F+*
AL751204	RS06C03	T29M8.12 protein	P	*M+*
BX255434	Ge005G04	UDP-glucose 4-epimerase (EC 5.1.3.2)	P	*M+*
BX682721	Ge120F04	Uridylate kinase (EC 2.7.4.3)	P	*M+*

Accession	Identifiant	Annotation	Effet	*sens*
Stress Long				
BX250471	Ge037D04	(1-4)-beta-mannan endohydrolase precursor (EC 3.2.1.78).	C	*S-*
BX679860	RS35A06	2-alpha-hydroxytaxane 2-O-benzoyltransferase (EC 2.3.1.166)	C	*S-*
BX681748	RS63G11	2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase (EC 4.6.	C	*S-*
BX680204	RS40E02	4-hydroxyphenylpyruvate dioxygenase (EC 1.13.11.27)	pC	*S+F+*
BX678841	RS15G11	4-hydroxyphenylpyruvate dioxygenase (EC 1.13.11.27)	C	*S+*
BX667551	RN09D10	4-hydroxyphenylpyruvate dioxygenase (EC 1.13.11.27)	PC	*S+F+*
BX252860	Ge071H04	5-methyltetrahydropteroyltriglutamate--homocysteine methyltra	C	*S-*
BX678822	RS15F01	Acyltransferase homolog	C	*S+*
BG040863	NXSI_116_A9	adenosine kinase (AK)	C	*S-*
BE607163	NXCI_37_F8	adenosine kinase (AK)	C	*S-*
BX249835	Ge029C07	Adenosine kinase (EC 2.7.1.20)	C	*S-*
BX682426	Ge116E03	Adenosylhomocysteinase (EC 3.3.1.1)	C	*S-*
BX253613	Ge085H08	Adenosylhomocysteinase (EC 3.3.1.1)	C	*S-*
BX679759	RS33B10	Adenylate kinase (EC 2.7.4.3)	P	*F+*
BX680855	RS49G03	Alcohol dehydrogenase (EC 1.1.1.1)	Pi	
BX680234	RS40H04	Alcohol oxidase 2	C	*S+*
BX681520	RS60F10	Aldehyde dehydrogenase (NAD+) (EC 1.2.1.3)	C	*S-*
BX254971	Ge108E04	Alkaline alpha galactosidase II, partial (84%)	C	*S+*
BX253114	Ge078C04	Alkaline alpha galactosidase II, partial (84%)	C	*S+*
BX678666	RS12H05	Alpha-1,4-glucan-protein synthase (UDP-forming) (EC 2.4.1.11	C	*S-*
BX679575	RS28F02	Alpha-1,4-glucan-protein synthase (UDP-forming) (EC 2.4.1.11	C	*S-*
BX680681	RS47B09	Anter-specific proline-rich protein / Lanatoside 15'-O-acetyleste	c	*S-*
BF220737	NXCI_150_A7	Arginine decarboxylase	Cl	
BX680824	RS49A07	Arginine methyltransferase-like protein (EC 2.1.1.-)	C	*S+*
AW870208	NXNV_125_E12	Aspartate transaminase	C	*S-*
AW985459	NXNV_136_H4	Aspartate transaminase	Ci	
BX252930	Ge075B02	ATP citrate lyase a-subunit, partial (84%)	C	*S-*
BX677977	RN47F03	Beta-fructofuranosidase (EC 3.2.1.26), cell wall isozyme precur	C	*S+*
BX250516	Ge038A02	Beta-galactosidase , partial (37%)	C	*S-*
CR392379	RN31A10	Beta-glucosidase precursor (EC 3.2.1.21)	P	*F+*
BX681997	RS67H05	Beta-glucosidase, chloroplast precursor (EC 3.2.1.21)	P	*F+*
BX678852	RS16A04	Caffeate O-methyltransferase (EC 2.1.1.68)	c	*S-*
BX254093	Ge093C03	caffeate O-methyltransferase [Pinus taeda]	C	*S-*
CR392624	RN35F03	Caffeoyl-CoA O-methyltransferase (EC 2.1.1.104)	C	*S-*
AL750008	AS03G06	Carbonate dehydratase (EC 4.2.1.1)	c	*S+*
BG039035	NXSI_24_H1	Cellulose synthase	C	*S-*
BX249550	Ge025D03	Citrate synthase	C	*S+*
BX679422	RS25B09	Cysteine synthase, chloroplast precursor (EC 2.5.1.47)	C	*S+*
CR393320	RN58E01	Cytochrome P450 (EC 1.14.-.-)	C	*S-*
BX679683	RS31C02	Cytochrome P450 (EC 1.14.-.-), putative	C	*S+*
CR394072	RN73D02	Cytochrome P450 77A3 (EC 1.14.-.-)	C	*S+*
BX255565	Ge007C03	Dephospho-CoA kinase (Dephosphocoenzyme A kinase) , par	C	*S-*
BX680021	RS38C03	Dihydroflavonol-4-reductase (EC 1.1.1.219) (DFR)	Ci	
BX681865	RS65H11	Dihydropinosylvin synthase (EC 2.3.1.-)	P	*M+*
BX677769	RN44E06	Enoyl-CoA hydratase (EC 4.2.1.17)	C	*S+*
CR392825	RN48H03	Epimerase/dehydratase	pC	*S-M+*
BX252727	Ge069G09	Esterase precursor (Early nodule-specific protein homolog) (La	C	*S-*
BX682159	RS72E06	Flavonoid 3',5'-hydroxylase (EC 1.14.-.-)	C	*S+*
AL750486	RN03A11	Flavonol 3-O-glucosyltransferase (EC 2.4.1.91)	C	*S+*
BX681828	RS65D10	Flavonol 3-O-glucosyltransferase 5 (EC 2.4.1.91)	C	*S+*
AL751292	RS07B10	Gamma-lyase	C	*S+*
BX254083	Ge093A12	Geranyl diphosphate synthase, complete	C	*S-*
BX677473	RN40A01	Glutamate decarboxylase (EC 4.1.1.15)	C	*S-*
CR392473	RN32G07	Glutamate synthase (NADH) (Sb14) (EC 1.4.1.14)(Picea maria	C	*S+*
BX677309	RN20D07	Glutathione transferase (EC 2.5.1.18)	C	*S+*
BX679708	RS32A02	Glutathione transferase (EC 2.5.1.18)	C	*S+*
BX681235	RS55H11	Glutathione transferase (EC 2.5.1.18)	C	*S+*
BX678400	RS08F03	Glutathione transferase (EC 2.5.1.18)	C	*S+*
CR394335	RN76H08	Glutathione transferase (EC 2.5.1.18)	C	*S-*
BX249364	Ge023B09	glutathione-dependent formaldehyde dehydrogenase {Arabidop	C	*S+*
AW784085	NXNV_117_F2	glyceraldehyde-3-P-dehydrogenase	C	*S-*
BX000600	GP	GP	C	*S-*
BX682765	Ge121D06	Granule-bound glycogen [starch] synthase	C	*S-*
BX681976	RS67E11	Histidine decarboxylase (EC 4.1.1.22)	Cl	
BX681925	RS66H01	Hydroxymethylglutaryl-CoA synthase (EC 2.3.3.10)	C	*S+*
BX252145	Ge060D11	Hypothetical protein yfbG	C	*S-*
BX679517	RS27D07	Inosine-5'-monophosphate dehydrogenase (EC 1.1.1.2)	C	*S+*
BX254079	Ge093A05	Inositol oxygenase (Myo-inositol oxygenase) (Aldehyde reducta	Pi	
BX682428	Ge116E05	Inositolphosphorylceramide-B C-26 hydroxylase	C	*S-*
BX679924	RS36D05	Iron(III)-zinc(II) purple acid phosphatase precursor	C	*S+*
BX681985	RS67F12	Isoflavone reductase homolog (EC 1.3.1.-)	C	*S-*
BX248894	Ge002D03	Lanatoside 15'-O-acetylesterase precursor, partial (28%)	C	*S-*
CR394206	RN75B08	L-ascorbate oxidase precursor (EC 1.10.3.3)	C	*S-*
BX679131	RS20C02	Leucoanthocyanidin dioxygenase (EC 1.-.-.-) (LDOX)	C	*S-*
CR393529	RN61D03	Malate dehydrogenase, mitochondrial precursor	C	*S-*
AL750160	AS06D09	Metallothionein-like protein	c	*S-*
AL751302	RS07C09	Metallothionein-like protein	Pc	*S-M+*
AL749965	AS03B12	Metallothionein-like protein EMB30	C	*S-*
CR392866	RN49C12	Metallothionein-like protein EMB30	c	*S-*
BX249127	Ge020C11	Metallothionein-like protein EMB30	pC	*S-M+*
BX253036	Ge077A02	Methionine synthase, partial (63%)	C	*S-*
BX249460	Ge024C08	Mitochondrial aldehyde dehydrogenase, partial (91%)	C	*S+*
AL749938	AS02H06	Naringenin 3-dioxygenase (EC 1.14.11.9)	C	*S+*
AL749834	AS01F08	Naringenin-chalcone synthase (EC 2.3.1.74)	C	*S-*
BX680340	RS42B09	Naringenin-chalcone synthase (EC 2.3.1.74)	c	*S-*
CR392095	RN11G12	Naringenin-chalcone synthase (EC 2.3.1.74)	C	*S-*
CR393420	RN59H07	Naringenin-chalcone synthase (EC 2.3.1.74)	Ci	
CR392294	RN30E11	Omega-6 fatty acid desaturase, endoplasmic reticulum	C	*S-*
BG039825	NXSI_104_E11	Ornitine transaminase	C	*S-*
BX678049	RN55F05	Phenylalanine ammonia-lyase (EC 4.3.1.5)	C	*S-*
CR392209	RN28C02	Phenylalanine ammonia-lyase (EC 4.3.1.5)	C	*S+*
BX682869	Ge122H09	Pinene synthase (EC 4.2.3.14)	C	*S-*
BX678354	RS08A06	Pinus taeda Trans-cinnamate 4-monooxygenase (EC 1.14.13.1	Cl	
BX679917	RS36C02	Protein kinase (EC 2.7.1.-)	c	*S+*
BX681938	RS67A12	Protein kinase (EC 2.7.1.-)	pI	
BX253959	Ge091C11	Putative receptor protein kinase TMK1 precursor (EC 2.7.1.-)	C	*S+*
CR393742	RN64E07	S-adenosyl-L-methionine:salicylic acid carboxyl methyltransfera	C	*S+*
BE643924	NXCI_50_B7	S-adenosyl-Met synthase	C	*S-*
AL751120	RS05C10	S-adenosylmethionine-dependent methyltransferase T12H1.6	C	*S+*
BX679620	RS29H05	Serine acetyltransferase	C	*S+*
BX254041	Ge092E09	serine protease	c	*S-*
CR394120	RN74B04	Serine/threonine specific protein phosphatase 2C (EC 3.1.3.16)	C	*S+*
BX682627	Ge118H04	Shikimate 5-dehydrogenase (EC 1.1.1.25)	C	*S-*
BX677417	RN38F06	Short-chain type dehydrogenase/reductase (EC 1.-.-.-)	C	*S+*
BX254773	Ge104C09	Sterol-C-methyltransferase.	C	*S-*
BX676897	RN13F04	Syringomycin biosynthesis enzyme-like protein	C	*S+*
AL751204	RS06C03	T29M8.12 protein	pc	*S-M+*
CR393272	RN54G03	Transaldolase (EC 2.2.1.2)	c	*S+*
BX678013	RN55B08	Trans-cinnamate 4-monooxygenase (EC 1.14.13.11)	C	*S-*
BX681070	RS52G08	Trans-cinnamate 4-monooxygenase (EC 1.14.13.11)	C	*S-*
BX254966	Ge108C11	transitional endoplasmic reticulum ATPase - Arabidopsis thaliar	C	*S-*
BX255434	Ge005G04	UDP-glucose 4-epimerase (EC 5.1.3.2)	C	*S-*
BX682993	Ge126F07	UDP-glucose 4-epimerase (EC 5.1.3.2)	C	*S-*

Non-classé Accession	Identifiant	Annotation	Effet	sens
Stress Court				
BX252228	Ge061F04	(2R)-phospho-3-sulfolactate synthase (EC 4.4.-.-)	PC	*S+M+*
BX251533	Ge050H03	At1g11170/T28P6_16, partial (58%)	P	*F+*
BX680179	RS40B08	At2g44310 Hypothetical protein	P	*F+*
BX249996	Ge031B08	At3g03773 {Arabidopsis thaliana;} , partial (71%)	P	*F+*
BX248924	Ge002G04	At3g23090, partial (15%)	P	*F+*
BX682906	Ge123D02	ATP-dependent Clp protease proteolytic su	P	*M+*
CR392486	RN33A12	ATP-dependent Clp protease proteolytic subunit (EC 3.4.21.92)	P	*F+*
CR394213	RN75C03	Auxin-repressed 12.5 kDa protein	P	*M+*
BX252717	Ge069F09	B2 protein	P	*M+*
CR354614	Ge124A01	Blr1429 protein, partial (18%)	P	*M+*
BX254656	Ge102B08	CG33184-PA, isoform A, partial (9%)	P	*F+*
BX250435	Ge036H06	CTV.15, partial (41%)	PC	*S+F+*
AL750116	AS05H10	EF-HAND containing protein-like	C	*S+*
BX681067	RS52G04	Embryogenesis-associated protein EMB7	P	*M+*
BX677557	RN41B08	GASA5-like protein (Sb35)	P	*F+*
BX255662	Ge008C06	Glycine-rich protein homolog	P	*M+*
BX249906	Ge030B06	Hydrophobic protein RCI2 (Low temperature and salt induced)	C	*S+*
BX253820	Ge089A07	Hypothetical 10.4 kDa protein (Expressed protein)	P	*F+*
AL750551	RN03G06	Hypothetical 168.1 kDa protein	P	*F+*
BX679085	RS19F02	Hypothetical 20.6 kDa protein	P	*F+*
AL750051	AS05C03	Hypothetical 49.0 kDa protein	PC	*S-M+*
BX681024	RS52B04	Hypothetical 60.3 kDa protein	P	*F+*
AL751163	RS05G07	Hypothetical 78.8 kDa protein	P	*F+*
CR393785	RN65C08	Hypothetical 8.7 kDa protein	P	*F+*
BX680327	RS42A07	Hypothetical 85.6 kDa protein	P	*F+*
BX680229	RS40G08	Hypothetical protein (At1g76160/T23E18_10)	P	*F+*
BX679990	RS37G05	Hypothetical protein (AT5g01750/T20L15_20)	P	*F+*
BX678993	RS18B02	IN2-2 protein	C	*S-*
CR393481	RN60F10	IN2-2 protein	P	*F+*
BX681544	RS61A06	Lateral organ boundaries (LOB) domain protein	P	*F+*
BX249968	Ge030H03	Latex-abundant protein	P	*F+*
BX253244	Ge080C01	Lin1294 protein, partial (9%)	C	*S+*
CR393329	RN58F01	Major pollen allergen Bet v 1-D/H / Intracellular pathogenesis-re	C	*S+*
BX254006	Ge092B02	Outer membrane protein (Fragment), partial (20%)	P	*F+*
BX250861	Ge042C04	OVCA2=CANDIDATE tumor suppressor, partial (9%)	P	*M+*
BX679871	RS35C05	Pinus taeda clone PtaAGP5	C	*S-*
BX254145	Ge094A09	PR34, partial (4%)	P	*M+*
BX679792	RS34A05	PrMC3	P	*M+*
BX679458	RS26B10	Proline-rich protein	C	*S-*
AL751295	RS07C01	Proline-rich protein / Cortical cell delineating protein precursor	PC	*S-M+*
AL751090	RS05A04	Protein kinase (EC 2.7.1.-)	P	*F+*
CR394188	RN74H08	Protein kinase-like (EC 2.7.1.37)	P	*F+*
AL750511	RN03D01	PtaAGP5 (from Pinus taeda)	C	*S-*
BX678411	RS08G05	RNA-binding protein KIAA0682	P	*F+*
BX252796	Ge070H02	SOUL-like protein, partial (10%)	P	*F+*
BX676834	RN12D04	T17B22.3 protein	P	*F+*
BX255226	Ge113B09	TrPRP2, partial (67%)	PC	*S-M+*
BX253650	Ge086D09	Unknown	C	*S+*
CR393111	RN52G02	Unknown	C	*S+*
BX677148	RN18D05	Unknown	C	*S+*
BX253795	Ge088F01	Unknown	C	*S+*
BX680324	RS42A12	Unknown	C	*S+*
BX249919	Ge030C08	Unknown	C	*S+*
BX681370	RS57H02	Unknown	P	*F+*
AL750105	AS05G10	Unknown	P	*F+*
BX251778	Ge055C09	Unknown	P	*F+*
AL750301	AS07H12	Unknown	P	*F+*
CR393841	RN65G12	Unknown	P	*F+*
BX250726	Ge040E06	Unknown	P	*M+*
BX677733	RN44A03	Unknown	P	*F+*
BX249297	Ge022D04	Unknown	P	*F+*
AL751076	RS03G09	Unknown	P	*F+*
CR393994	RN71F04	Unknown	P	*F+*
AL751149	RS05F05	Unknown	P	*F+*
BX678298	RN69C12	Unknown	P	*F+*
CR393770	RN65A10	Unknown	P	*F+*
BX678445	RS09C02	Unknown	P	*F+*
CR393402	RN59F08	Unknown	P	*F+*
BX249896	Ge030A04	Unknown	P	*F+*
CR393769	RN65A09	Unknown	P	*F+*
BX667580	RN09G05	Unknown	P	*F+*
BX682007	RS68A10	Unknown	P	*F+*
BX252163	Ge060F11	Unknown	PC	*S+M+*

Accession	Identifiant	Annotation	Effet	sens
Stress Long				
BX252228	Ge061F04	(2R)-phospho-3-sulfolactate synthase (EC 4.4.-.-)	C	*S+*
BX252312	Ge062G04	33 kDa secretory protein-like, partial (13%)	C	*S-*
BX681954	RS67C08	37 kDa inner envelope membrane protein	PI	
BX255465	Ge006B03	acetone-cyanohydrin lyase [imported] - Arabidopsis thaliana, pa	c	*S-*
BX682681	Ge120A02	Anter-specific proline-rich protein	C	*S-*
BX249830	Ge029B11	At1g07510/F22G5_9, partial (32%)	C	*S-*
BX254648	Ge102A03	At1g64110/F22C12_22, partial (47%)	C	*S+*
BX249846	Ge029D04	At2g43970/F6E13.10 {Arabidopsis thaliana;} , partial (41%)	C	*S+*
CR392445	RN32B03	ATAF1-like protein (Fragment)	C	*S+*
BX249398	Ge023F01	Auxin-repressed 12.5 kd protein	C	*S+*
BX254347	Ge097C07	CoA-thioester hydrolase CHY1 (3-hydroxyisobutyryl-coenzyme	C	*S+*
BX250435	Ge036H06	CTV.15, partial (41%)	C	*S+*
BX681915	RS66F09	Embryogenesis-associated protein EMB3	C	*S+*
BX681172	RS54G01	Endosperm-specific protein-like protein	C	*S-*
BX680404	RS43A10	F12F1.23 protein (Hypothetical 36.7 kDa protein)	C	*S-*
BX680058	RS38F06	F14L17.20 protein	C	*S-*
BX250697	Ge040B09	F16A14.7, partial (50%)	C	*S+*
AL750468	RN02H01	F21O3.6 Hypothetical protein	PC	*S+F+*
BX682853	Ge122G03	F3O9.22 protein.	pC	*S-F+*
BX255014	Ge109C07	Glucose-1-phosphate adenylyltransferase small subunit, chloro	C	*S-*
BX254410	Ge098C07	Glycine-rich protein	C	*S-*
BX255662	Ge008C06	Glycine-rich protein homolog	C	*S-*
BX680139	RS39F01	GPI-anchored protein	c	*S-*
BX254445	Ge098F09	high mobility group protein MNB1b - maize (fragment) {Zea ma	c	*S-*
BX680652	RS46F06	hsr203J	C	*S+*
BX680987	RS51G11	Hyphally regulated protein precursor	C	*S+*
BX678353	RS08A05	Hypothetical 12.0 kDa protein	c	*S-*
BX680979	RS51F01	Hypothetical 19.9 kDa protein	C	*S-*
BX677604	RN41H04	Hypothetical 20.8 kDa protein	C	*S+*
BX680601	RS45G02	Hypothetical 25.9 kDa protein	C	*S+*
BX680674	RS47B01	Hypothetical 34.5 kDa protein	C	*S+*
AL749952	AS03A10	Hypothetical 38.0 kDa protein (Putative dehydrogenase)	C	*S-*
BX677222	RN19E10	Hypothetical 40.2 kDa protein	C	*S+*
AL750051	AS05C03	Hypothetical 49.0 kDa protein	C	*S-*
AL751305	RS07C12	Hypothetical 58.9 kDa protein	C	*S+*
CR392178	RN27F05	Hypothetical 68.3 kDa protein	C	*S-*
BX251850	Ge056D01	Hypothetical protein	C	*S+*
BX254155	Ge094B08	Hypothetical protein	C	*S+*
BX679990	RS37G05	Hypothetical protein (AT5g01750/T20L15_20)	C	*S+*
AL750195	AS06G11	Hypothetical protein DR0254	pC	*S+F+*
BX678617	RS12A09	IN2-1 protein	C	*S+*
CR393329	RN58F01	Major pollen allergen Bet v 1-D/H / Intracellular pathogenesis-re	PC	*S+F+*
BX252300	Ge062E09	Nuc-1 negative regulatory protein preg	C	*S+*
AL751202	RS06C01	P0022F10.3 protein	c	*S-*
BX255417	Ge005E10	P20-Arc, partial (92%)	C	*S-*
BX679459	RS26C01	PGPD14 protein (Pollen germination related protein)	C	*S+*
BX249499	Ge024G05	Pherophorin-dz1 protein precursor, partial (7%)	C	*S+*
BX254994	Ge108H10	Phosphate-induced protein	C	*S-*
BX250614	Ge039B07	phosphate-induced protein	C	*S-*
CR392105	RN12D09	Pinus taeda calcium binding protein	PCI	
BX679871	RS35C05	Pinus taeda clone PtaAGP5	C	*S-*
BX680432	RS43C07	Pinus taeda clone PtaAGP5	C	*S-*
BX252532	Ge066B06	Potyvirus VPg interacting protein, partial (55%)	C	*S+*
BX254145	Ge094A09	PR34, partial (4%)	C	*S+*
BX679792	RS34A05	PrMC3	Pc	*S-M+*
BX682906	Ge123D02	Probable ATP-dependent Clp protease proteolytic su	C	*S-*
BX679458	RS26B10	Proline-rich protein	C	*S-*
AL751295	RS07C01	Proline-rich protein / Cortical cell delineating protein precursor	C	*S-*
AL750511	RN03D01	PtaAGP5 (from Pinus taeda)	C	*S-*
BX254033	Ge092D12	Remorin 1, partial (57%)	C	*S+*
BX679201	RS21D07	SelT-like protein precursor	C	*S+*
CR393456	RN60D06	Serine-rich protein	C	*S+*
BX682334	Ge115B11	Surfeit 1.	C	*S-*
BX679999	RS35D10	Unknown	C	*S+*
BX254570	Ge100E02	Unknown	C	*S+*
BX254217	Ge095A09	Unknown	C	*S+*
BX250080	Ge032B10	Unknown	C	*S-*
BX680213	RS40F01	Unknown	C	*S-*
CR393111	RN52G02	Unknown	C	*S+*
BX679842	RS34G02	Unknown	C	*S+*
BX681032	RS52C04	Unknown	C	*S+*
BX677148	RN18D05	Unknown	C	*S+*
BX681900	RS66E03	Unknown	c	*S+*
CR354677	Ge124G05	Unknown	C	*S-*
CR392777	RN37H08	Unknown	C	*S-*
BX679424	RS25B12	Unknown	C	*S+*
BX251479	Ge050B02	Unknown	c	*S+*
BX681605	RS61H03	Unknown	C	*S+*
BX254354	Ge097D04	Unknown	C	*S+*
BX677733	RN44A03	Unknown	C	*S+*
BX681149	RS54D01	Unknown	C	*S+*
BX681640	RS62D03	Unknown	c	*S+*
BX681917	RS66F12	Unknown	C	*S+*
CR392966	RN50F06	Unknown	C	*S+*
BX680324	RS42A12	Unknown	C	*S+*
BX681664	RS62F06	Unknown	C	*S+*
CR392836	RN48F07	Unknown	C	*S+*
AL750277	AS07F11	Unknown	C	*S+*
BX681199	RS55C12	Unknown	C	*S+*
BX255599	Ge007F02	Unknown	C	*S+*
CR393695	RN63F04	Unknown	c	*S-*
BX680588	RS45E06	Unknown	C	*S+*
CR354695	Ge125A05	Unknown	C	*S-*
BX681025	RS52B05	Unknown	C	*S+*
BX677636	RN42D06	Unknown	C	*S+*
BX253794	Ge088E12	Unknown	C	*S-*
CR394058	RN73B01	Unknown	c	*S+*
BX249919	Ge030C08	Unknown	C	*S+*
BX681400	RS58E03	Unknown	C	*S+*
CR354742	Ge125H03	Unknown	C	*S-*
BX252363	Ge063D09	Unknown	p	*F+*
BX255048	Ge109H02	Unknown	p	*F+*
BX680587	RS45E04	Unknown	p	*M+*
BX251976	Ge058A03	Unknown	pC	*S+F+*
BX678345	RN69H07	Vegetatible incompatibility protein HET-E-1	C	*S-*
BX784344	Ge133B11	Wsv162 (WSSV218), partial (23%)	C	*S+*
BX677573	RN41D05	Yippee-like protein	C	*S+*

Organisation cellulaire

Accession	Identifiant	Annotation	Effet	*sens*
Stress Court				
CR394036	RN72F06	Agglutinin alpha chain (MPA)	P	*M+*
CR392181	RN27G03	Annexin	P	*M+*
BX681731	RS63F03	Anther specific LAT52 protein precursor	C	*S+*
BX251886	Ge056H01	Anther specific protein / Pollen specific & Anther specific LAT52	P	*M+*
BX252966	Ge075G09	At5g20110, partial (46%)	P	*F+*
BX249875	Ge029G04	Glucan endo-1,3-beta-D-glucosidase (EC 3.2.1.39)	P	*M+*
AL751122	RS05C12	Glucan endo-1,3-beta-glucosidase precursor	P	*F+*
BX252660	Ge068G01	Histone H2A	PC	*S-M+*
BX252614	Ge067H06	Histone H2A variant	P	*M+*
AL750508	RN03C09	Histone H4	P	*M+*
BX679616	RS29G03	Histone H4	PC	*S-M+*
AL750525	RN03E03	MtN21 nodulin protein-like	C	*S-*
AL750706	RN06D08	Nodulin	C	*S+*
BX680100	RS39B01	Nodulin-like protein (Putative nodulin protein)	P	*F+*
BX680185	RS40C03	Pectate lyase (EC 4.2.2.2)	P	*F+*
AL750621	RN05E06	Pectate lyase (EC 4.2.2.2)	PC	*S-M+*
BX250283	Ge034G10	Pectate lyase, partial (95%)	P	*F+*
BX680845	RS49D06	Polyphosphoinositide binding protein	PC	*S-M+*
BX249627	Ge026D05	Tubulin beta chain	P	*M+*
BX682459	Ge116H08	Ubiquitin-conjugating enzyme (EC 6.3.2.19)	P	*M+*
BX251918	Ge057C01	Ubiquitin-conjugating enzyme (EC 6.3.2.19)	P	*M+*
BX255062	Ge110B02	Ubiquitin-conjugating enzyme E2 (EC 6.3.2.19)	C	*S+*
BX250267	Ge034E07	Ubiquitin-conjugating enzyme E2-17 kDa (EC 6.3.2.19)	P	*M+*
AL750092	AS05F09	Ubiquitin-conjugating enzyme E2-17 kDa (EC 6.3.2.19)	P	*M+*
BX253993	Ge091H07	Ubiquitin-conjugating enzyme E2-17 kDa (EC 6.3.2.19)	P	*M+*
BX681382	RS58B04	Ubiquitin-protein ligase 1	P	*F+*

Accession	Identifiant	Annotation	Effet	*sens*
Stress Long				
BX255342	Ge004F07	Actin	C	*S-*
CR394036	RN72F06	Agglutinin alpha chain (MPA)	PC	*S-M+*
CR392181	RN27G03	Annexin	C	*S-*
BX251886	Ge056H01	Anther specific protein / Pollen specific & Anther specific LAT52	C	*S-*
BX252966	Ge075G09	At5g20110, partial (46%)	C	*S+*
BX252682	Ge069A05	Cellulose synthase	c	*S-*
BX682714	Ge120E04	Cellulose synthase	C	*S-*
BX253583	Ge085E07	Endoglucanase	c	*S-*
AL751122	RS05C12	Glucan endo-1,3-beta-D-glucosidase (EC 3.2.1.39)	C	*S+*
AL751004	RS03A03	Glucan endo-1,3-beta-glucosidase precursor	C	*S-*
BX249875	Ge029G04	Glucan endo-1,3-beta-glucosidase precursor	C	*S-*
BX680247	RS41A09	Glucan endo-1,3-beta-glucosidase precursor	C	*S-*
BX677552	RN41B01	Glycosyl transferase-like protein	C	*S+*
CR393223	RN54B06	Heterochromatin-like protein	c	*S+*
BX682259	Ge114B12	Histone H2A	C	*S-*
BX252660	Ge068G01	Histone H2A	C	*S-*
BX682172	RS73A01	Histone H2A	C	*S+*
BX677123	RN18A04	Histone H2B	C	*S-*
BX250930	Ge043A10	MERI-5 protein precursor (Endo-xyloglucan transferase) / Bras	C	*S-*
AL750706	RN06D08	Nodulin	C	*S+*
BX680845	RS49D06	Polyphosphoinositide binding protein	c	*S-*
BX677988	RN47H03	Tomato extensin mRNA	C	*S-*
BX249241	Ge021G02	Tubulin alpha chain	C	*S-*
BX253001	Ge076C12	Tubulin alpha chain	C	*S-*
BX252968	Ge075G10	Tubulin alpha chain	C	*S-*
BX681614	RS62A02	Tubulin beta chain	C	*S-*
BX252002	Ge058D06	Tubulin beta chain	C	*S-*
CR393797	RN65E06	Tubulin beta chain	C	*S-*
BX249627	Ge026D05	Tubulin beta chain	C	*S-*
BX000608	Tubuline	Tubuline	C	*S-*
BX254985	Ge108G08	Ubiquitin	C	*S-*
BX254406	Ge098C03	Ubiquitin	C	*S-*
BX250835	Ge041H12	Ubiquitin carboxyl-terminal hydrolase (EC 3.1.2.15)	C	*S+*
CR393703	RN63G04	Ubiquitin conjugating enzyme spm2	c	*S+*
BX255062	Ge110B02	Ubiquitin-conjugating enzyme E2 (EC 6.3.2.19)	p	*F+*

Réponse au Stress

Accession	Identifiant	Annotation	Effet	*sens*
Stress Court				
BX252065	Ge059C02	17.9 kDa class II heat shock protein	P	*M+*
BX679279	RS22D07	ABA induced plasma membrane protein	PC	*S+M+*
BX682324	Ge115A12	Abscisic stress ripening protein	P	*M+*
BX251094	Ge045C02	Abscisic stress ripening protein	P	*M+*
BX254489	Ge099C04	ATP-dependent DNA helicase (EC 3.6.1.-), partial (2%)	P	*F+*
AL750218	AS07A11	Chitinase (EC 3.2.1.14)	C	*S+*
CR393567	RN61H01	Chitinase precursor (EC 3.2.1.14)	C	*S+*
CR392715	RN37A08	Chitinase precursor (EC 3.2.1.14)	P	*F+*
BX677767	RN44E03	Chitinase precursor (EC 3.2.1.14)	PC	*S+M+*
BX252366	Ge063E01	Class I Heat Shock Protein	PC	*S+M+*
AL750400	RN02B01	Copper chaperone for Superoxide dismutase	P	*F+*
BX250140	Ge032H07	Defender against cell death	C	*S+*
BX681602	RS61G11	Defensin	PC	*S-M+*
BX681136	RS54B08	Dehydration stress-induced protein	PC	*S-M+*
CR392961	RN50F01	Dehydration-responsive protein RD22 precursor	P	*F+*
BX249564	Ge025E05	Dehydrin	C	*S+*
BX255778	Ge009F07	Disease resistance protein-like	P	*F+*
BX680242	RS41A02	Disease resistance response protein 206	C	*S-*
BX253926	Ge090G07	Disease resistance-like protein (Fragment), partial (16%)	P	*F+*
BX253877	Ge089H04	DNA repair protein RAD16, partial (6%)	P	*F+*
BX681362	RS57F06	DnaJ protein homolog	P	*F+*
BX255377	Ge005B04	Elastin precursor - human (fragment) {Homo sapiens;} , partial	PC	*S+M+*
BX249660	Ge026G07	ENDOCHITINASE (EC 3.2.1.14)	P	*F+*
CR394461	RN79E12	Ethylene-responsive protein, putative	PC	*S+F+*
BX678613	RS12A03	Gamma-thionin homolog precursor	P	*M+*
BX251816	Ge055H04	Heat shock 70kd protein	C	*S+*
BX000656	HSP	Heat shock protein	PC	*F+*
BX681237	RS55H04	Heat shock protein 101	PC	*S+M+*
BX678224	RN68A05	Late embryogenesis abundant protein	C	*S-*
BX666039	RN08H11	Late embryogenesis abundant protein	P	*F+*
BX677964	RN47C10	Late embryogenesis abundant protein	PC	*S-M+*
BX681471	RS60A05	Late embryogenesis abundant protein (Fragment)	P	*F+*
BX253089	Ge077G07	LEA, complete	PC	*S+M+*
BX251839	Ge056B11	Leucine-rich repeat-containing protein 2	P	*M+*
BX681387	RS58C02	Licheninase precursor (EC 3.2.1.73)	C	*S+*
BX250781	Ge041B11	Low molecular weight heat-shock protein, partial (98%)	PCi	
BX254258	Ge095G04	LUMINAL BINDING PROTEIN	P	*M+*
BX252607	Ge067G04	NIP3, partial (75%)	P	*M+*
BX680166	RS40A07	Pathogenesis-related protein	Ci	
BX678491	RS10A02	Pathogenesis-related protein 5 precursor (PR-5)	PC	*S-M+*
BX681431	RS59B06	Pathogenesis-related protein PR-4B precursor	P	*F+*
BX254293	Ge096C10	Peptide methionine sulfoxide reductase (EC 1.8.4.6)	P	*M+*
BX254690	Ge102G05	peroxidase	P	*F+*
CR392585	RN35A10	Peroxidase precursor (EC 1.11.1.7)	C	*S-*
CR392453	RN32D03	Peroxidase precursor (EC 1.11.1.7)	C	*S-*
AL750365	RN01F10	Peroxidase precursor (EC 1.11.1.7)	P	*M+*
AL751229	RS06E04	Peroxidase precursor (EC 1.11.1.7)	P	*F+*
BX679939	RS36F06	Peroxidase precursor (EC 1.11.1.7)	PC	*S-M+*
BX678783	RS15A09	Peroxidase precursor (EC 1.11.1.7)	PC	*S-M+*
CR393156	RN53C07	Peroxidase precursor (EC 1.11.1.7)	PC	*S-M+*
BX675142	RN10A03	Peroxidase precursor (EC 1.11.1.7)	PC	*S-M+*
BX678596	RS11G02	Peroxidase precursor (EC 1.11.1.7)	PC	*S-M+*
CR392767	RN37G07	Peroxidase precursor (EC 1.11.1.7)	PC	*S-M+*
BX680909	RS50F11	Pinus sylvestris class III peroxidase	P	*M+*
BX253632	Ge086B12	Snakin-1, partial (67%)	C	*S-*
BX255769	Ge009E10	T19F11.6 protein (NDR1/HIN1-like protein 1) (Harpin-induced p	C	*S-*

Accession	Identifiant	Annotation	Effet	*sens*
Stress Long				
BX252065	Ge059C02	17.9 kDa class II heat shock protein	PCI	
BX251044	Ge044E10	1-aminocyclopropane-1-carboxylate oxidase (EC 1.-.-.-)	C	*S+*
BX677677	RN43A04	9-cis-epoxycarotenoid dioxygenase	C	*S+*
BX679279	RS22D07	ABA induced plasma membrane protein	C	*S+*
BX666054	RN09B03	Abscisic stress ripening protein	C	*S+*
BX683057	Ge127F05_b	Aluminum-induced protein-like protein, partial (65%)	C	*S+*
BX253756	Ge088A07	Aluminum-induced protein-like protein, partial (67%)	pC	*S+F+*
BX681928	RS66H05	Antimicrobial peptide	PC	*S+F+*
AL750636	RN05F08	Antimicrobial peptide	PC	*S+F+*
CR394499	RN50C09	Antimicrobial peptide	PC	*S+F+*
BX252074	Ge059D03	arabinogalactan protein	C	*S-*
BX251932	Ge057D07	arabinogalactan protein	C	*S-*
BX252031	Ge058G08	arabinogalactan protein [Pinus taeda]	C	*S-*
BX249652	Ge026F11	ASR protein, partial (39%)	C	*S+*
BX255310	Ge004C06	catalase 1 [Pinus pinea]	C	*S+*
AL750545	RN03F11	Catalase isozyme (EC 1.11.1.6).	C	*S+*
BX678372	RS08C03	Cationic peroxidase precursor (EC 1.11.1.7)	C	*S+*
CR393726	RN64C02	Chaperone protein dnaJ	c	*S+*
AL751212	RS06C11	Chaperone protein dnaJ	C	*S+*
BX679343	RS23E08	Chaperone protein dnaJ	PC	*S+F+*
BX677767	RN44E03	Chitinase precursor (EC 3.2.1.14)	C	*S+*
CR393567	RN61H01	Chitinase precursor (EC 3.2.1.14)	C	*S+*
BX252366	Ge063E01	Class I Heat Shock Protein	C	*S+*
BX677784	RN44F10	Copper chaperone for Superoxide dismutase	C	*S+*
BX251025	Ge044C12	Cyclophilin, partial (98%)	C	*S-*
BX681602	RS61G11	Defensin	C	*S-*
CR392200	RN28B03	Defensin	P	*F+*
BX681136	RS54B08	Dehydration stress-induced protein	PC	*S-F+*
BX249564	Ge025E05	dehydrin	C	*S+*
BX680242	RS41A02	Disease resistance response protein 206	C	*S-*
CR394442	RN79C01	Disease resistance response protein 206	c	*S-*
BX681402	RS58E06	DnaJ-like protein	C	*S+*
BX681263	RS56C03	EDS1-like protein	C	*S+*
BX252471	Ge065A11	ENDOCHITINASE (EC 3.2.1.14)	C	*S-*
CR393324	RN58E05	Epoxide hydrolase (EC 3.3.2.3)	pI	
BX252277	Ge062C04	Harpin-induced protein	C	*S-*
BX251816	Ge055H04	Heat Shock 70 kd protein	c	*S+*
BX681237	RS55H04	Heat shock protein 101	C	*S+*
CR393209	RN53H11	Heat shock protein STI1	C	*S+*
BX680673	RS47B12	Hypothetical protein (DNAJ domain-containing unknown protein	C	*S+*
BX681356	RS57E07	Late embryogenesis abundant protein	C	*S+*
BX678224	RN68A05	Late embryogenesis abundant protein	pI	
BX253089	Ge077G07	LEA, complete	C	*S+*
CR394078	RN73D08	Licheninase precursor (EC 3.2.1.73)	PC	*S+F+*
BX250781	Ge041B11	Low molecular weight heat-shock protein, partial (98%)	PC	*S+M+*
BX681166	RS54F02	NBS/LRR	C	*S+*
BX252607	Ge067G04	NIP3, partial (75%)	C	*S-*
BX681832	RS65E04	Pathogenesis-related protein PRB1-2 precursor	C	*S+*
CR392453	RN32D03	Peroxidase precursor (EC 1.11.1.7)	C	*S-*
BX678152	RN57A03	Peroxidase precursor (EC 1.11.1.7)	C	*S-*
BX675142	RN10A03	Peroxidase precursor (EC 1.11.1.7)	C	*S-*
CR392722	RN37B06	Peroxidase precursor (EC 1.11.1.7)	C	*S-*
BX680225	RS40G02	Peroxidase precursor (EC 1.11.1.7)	c	*S+*
BX678783	RS15A09	Peroxidase precursor (EC 1.11.1.7)	c	*S-*
CR392767	RN37G07	Peroxidase precursor (EC 1.11.1.7)	C	*S-*
BX679939	RS36F06	Peroxidase precursor (EC 1.11.1.7)	PC	*S-M+*
BX680909	RS50F11	Pinus sylvestris class III peroxidase	PI	
BX253505	Ge084D05	Pinus taeda water deficit inducible protein LP3	c	*S-*
BX678362	RS08B03	Protein-methionine-S-oxide reductase (EC 1.8.4.6)	C	*S+*
AL750457	RN02G02	Protein-methionine-S-oxide reductase (EC 1.8.4.6)	C	*S+*
BX679992	RS37H01	Related to chitinase 3 precursor protein	C	*S+*
BX254281	Ge096B03	Ripening-associated protein (Fragment)	C	*S+*
BX252884	Ge074D02	Similarity to ABA induced plasma membrane protein, partial (52	c	*S-*
BX249989	Ge031B01	Similarity to late embryogenesis abundant protein, partial (31%)	pC	*S-M+*
BX253632	Ge086B12	Snakin-1, partial (67%)	C	*S-*
BX255769	Ge009E10	T19F11.6 protein (NDR1/HIN1-like protein 1) (Harpin-induced p	C	*S-*
BX250350	Ge035G11	Wound-induced basic protein	C	*S-*

Synthèse de protéines

Accession	Identifiant	Annotation	Effet	*sens*	Accession	Identifiant	Annotation	Effet	*sens*
Stress Court					***Stress Long***				
CR392543	RN33H08	25S rRNA	P	*F+*	BX681290	RS56F02	40S ribosomal protein S11	C	*S-*
AL750386	RN01H09	30S ribosomal protein S16	P	*F+*	CR392876	RN49D11	40S ribosomal protein S16	C	*S-*
CR392159	RN27C08	40S ribosomal protein	P	*F+*	CR393754	RN64G04	40S ribosomal protein S2	C	*S-*
BX681290	RS56F02	40S ribosomal protein S11	PC	*S-M+*	BX255132	Ge111E05	40S ribosomal protein S23	C	*S-*
CR392325	RN30F07	40S ribosomal protein S26 (S31)	PC	*S-M+*	CR392325	RN30F07	40S ribosomal protein S26 (S31)	P	*F+*
BX248870	Ge002A10	60S acidic ribosomal protein P1	P	*M+*	BX251853	Ge056D04	40S ribosomal protein S27	C	*S-*
BX679791	RS34A04	60S ribosomal protein	PC	*S-M+*	CR392980	RN51A02	40S ribosomal protein S8	C	*S+*
AL750567	RN03H10	60S ribosomal protein L14	C	*S-*	BX253123	Ge078D03	40S ribosomal protein S9 (S7)	C	*S-*
BX682787	Ge121G09	60S ribosomal protein L31	P	*M+*	BX252922	Ge074H10	60S ribosomal protein	C	*S-*
BX252371	Ge063E06	60S ribosomal protein L31	P	*M+*	BX248870	Ge002A10	60S ribosomal protein	C	*S-*
BX681771	RS64D09	60S ribosomal protein L34	P	*M+*	BX679791	RS34A04	60S ribosomal protein	C	*S-*
BX252657	Ge068F09	60S ribosomal protein L35A	P	*M+*	BX249385	Ge023D08	60S ribosomal protein	C	*S-*
BX682894	Ge123B11	60S ribosomal protein L38	P	*M+*	BX251297	Ge047H03	60S ribosomal protein L10	C	*S-*
BX251535	Ge050H05	60S ribosomal protein L39	P	*M+*	BX255767	Ge009E08	60S ribosomal protein L11	C	*S-*
BX678464	RS09E08	60S ribosomal protein L41, mitochondrial precursor	P	*F+*	AL750653	RN05H02	60S ribosomal protein L27a	C	*S-*
BX682634	Ge119A01	Elongation factor 1-gamma (EF-1-gamma)	PC	*S-M+*	BX682894	Ge123B11	60S ribosomal protein L38	C	*S-*
BX252622	Ge068A05	Elongation factor 1-gamma (EF-1-gamma) (eEF-1B gamma), p	PC	*S+M+*	BX251535	Ge050H05	60S ribosomal protein L39	C	*S-*
BX680013	RS38B06	Gb|AAD26879.1	P	*M+*	BX251593	Ge051G02	Eukaryotic peptide chain release factor subunit 1	C	*S+*
BX679837	RS34F05	L3 cytoplasmic ribosomal protein	PC	*S-M+*	BX679837	RS34F05	L3 cytoplasmic ribosomal protein	C	*S-*
BX679475	RS26E10	Ribosomal protein L10	P	*F+*	BX251675	Ge053G10	Poly(A)-binding protein, partial (84%)	P	*F+*
					BX680147	RS39G10	RNA helicase	C	*S+*
					BX678954	RS17D12	small subunit ribosomal RNA gene	C	*S+*
					BX252622	Ge068A05	Elongation factor 1-gamma (EF-1-gamma) (eEF-1B gamma), p	C	*S+*

Trafic intracellulaire

Accession	Identifiant	Annotation	Effet	*sens*	Accession	Identifiant	Annotation	Effet	*sens*
Stress Court					***Stress Long***				
BX676912	RN13G11	Acyl carrier protein, chloroplast precursor	P	*M+*	BX252434	Ge064E04	Coatomer alpha subunit	pCl	
BX254611	Ge101C06	ADP-ribosylation factor	P	*M+*	BX677127	RN18A09	GTP-binding nuclear protein RAN-A1	C	*S-*
BX252434	Ge064E04	Coatomer alpha subunit	P	*M+*	CR394236	RN75E09	Kinesin light chain (KLC)	C	*S+*
CR394037	RN72F07	Coatomer beta' subunit (Beta'-coat protein)	P	*F+*	BX678737	RS14D11	Myosin V	C	*S+*
BX681245	RS56A07	Dynamin-like protein	P	*F+*	BX681158	RS54E01	Portal protein (GPB) (Minor capsid protein B)	C	*S+*
CR394145	RN74D07	GTP-binding protein SAR	P	*F+*	BX249494	Ge024F11	Profilin	c	*S-*
CR394236	RN75E09	Kinesin light chain (KLC)	P	*F+*	BX251837	Ge056B06	Protein transport protein SEC22	p	*M+*
BX678294	RN69C07	Nuclear transport factor 2 (NTF-2)	C	*S-*	BX252360	Ge063D06	Sec13p, partial (90%)	c	*S+*
BX682440	Ge116F08	Nuclear transport factor 2 (NTF-2).	P	*M+*	CR394081	RN73D11	Syntaxin 71	C	*S+*
BX681158	RS54E01	Portal protein (GPB) (Minor capsid protein B)	P	*F+*	BX677699	RN43C10	Chaperonin-60 beta subunit precursor, partial (91%)	C	*S+*
BX249494	Ge024F11	Profilin	P	*M+*					
BX254888	Ge106H02	Ras-related protein Rab11B	P	*M+*					

Transcription

Accession	Identifiant	Annotation	Effet	*sens*	Accession	Identifiant	Annotation	Effet	*sens*
Stress Court					***Stress Long***				
BX250119	Ge032F08	Auxin response factor 1, partial (25%)	P	*F+*	BX252601	Ge067D04	Abscisic acid insensitive 5 (ABI5) (BZIP protein), partial (17%)	C	*S-*
BX680424	RS43C10	Ethylene responsive element binding factor	PC	*S-M+*	CR392290	RN30E06	DNA-binding protein ESCAROLA	C	*S-*
BX249687	Ge027B05	FKBP-like protein (Arabidopsis thaliana genomic DNA, chromo	PC	*S+M+*	BX680424	RS43C10	Ethylene responsive element binding factor	C	*S-*
BX682416	Ge116C09	G-box binding factor 4.	C	*S+*	BX255127	Ge111D09	eukaryotic initiation factor 3E subunit {Arabidopsis thaliana;} , p	pC	*S+F+*
BX254733	Ge103E05	Putative F-box protein (Fragment), partial (38%)	C	*S+*	BX250756	Ge040H03	F11M15.7 protein	C	*S+*
BX252953	Ge075E07	RING finger protein 26	PC	*S+M+*	BX682416	Ge116C09	G-box binding factor 4.	C	*S+*
BX251374	Ge048G11	Rof1, partial (88%)	PC	*S-M+*	BX678438	RS09B05	Homeobox-leucine zipper protein	C	*S+*
BX255626	Ge007H04	Small zinc finger-like protein, complete	P	*M+*	BX681312	RS56H05	Lil3 protein	C	*S+*
CR392874	RN49D09	Zinc finger protein 364	P	*F+*	BX680112	RS39C05	Nuclear polyadenylated RNA-binding protein	C	*S-*
BX253595	Ge085G01	Zinc finger protein HIT-10, partial (3%)	PC	*S+M+*	BX681454	RS59E06	Nuclear polyadenylated RNA-binding protein NAB4	C	*S-*
					BX682918	Ge123E04	Phenylcoumaran benzylic ether reductase homolog TH1, comp	C	*S-*
					BX680486	RS44B07	Poly(A)-binding protein	C	*S-*
					BX254733	Ge103E05	Putative F-box protein (Fragment), partial (38%)	PC	*S+M+*
					BX252302	Ge062E12	RNA-binding protein RZ-1, partial (9%)	c	*S+*
					BX251374	Ge048G11	Rof1, partial (88%)	Ci	
					CR394310	RN76F04	Transcription factor HBP-1b homolog	C	*S+*
					BX681072	RS52H08	Zinc finger protein 216	C	*S+*
					BX679178	RS21A09	Zinc finger protein 216	cI	
					BX253595	Ge085G01	Zinc finger protein HIT-10, partial (3%)	PC	*S+M+*
					BX251702	Ge054B09	Zinc-finger protein NOLZ1, partial (3%)	C	*S-*
					BX682945	Ge123H04	Zinc-finger protein, partial (53%)	C	*S+*

Transport

Accession	Identifiant	Annotation	Effet	*sens*		Accession	Identifiant	Annotation	Effet	*sens*
Stress Court						***Stress Long***				
BX681992	RS67G12	Fluconazole resistance protein 1	C	*S+*		CR394157	RN74F01	ADP,ATP carrier protein (ADP/ATP translocase)	c	*S+*
CR393619	RN62F01	Peptide transporter PTR2-B	P	*F+*	i+	BX667565	RN09F01	Amino acid permease (Amino acid carrier, putative)	C	*S+*
CR393875	RN66E08	Calcium-transporting ATPase, plasma membrane-type	P	*F+*	r	BX679636	RS30C11	Gamma tonoplast intrinsic protein	C	*S-*
CR393662	RN63B06	Proton-exporting ATPase (EC 3.6.3.6)	P	*M+*	i+	BX680183	RS40C01	Membrane transporter protein	C	*S+*
BX252185	Ge061A06	Mitochondrial oxaloacetate transport protein	P	*M+*	r	BX252185	Ge061A06	Mitochondrial oxaloacetate transport protein	PC	*S-M+*
AL749547	AN01B08	Outer membrane lipoprotein blc precursor	PC	*S+M+*		BX677658	RN42H03	Nonspecific lipid-transfer protein precursor (LTP)	Ci	
CR393241	RN54D04	Plasma membrane intrinsic polypeptide	PC	*S-M+*	i	AL749547	AN01B08	Outer membrane lipoprotein blc precursor	C	*S+*
BX254875	Ge106E11	Probable aquaporin, complete	P	*M+*		BX681797	RS65A06	Oxysterol-binding protein homolog	C	*S+*
					i	CR394220	RN75C12	Peptide transporter PTR2-B	pC	*S+F+*
					i	AL750557	RN03G12	Plasma membrane H+ ATPase (EC 3.6.1.3)	C	*S+*
					r	CR393241	RN54D04	Plasma membrane intrinsic polypeptide	C	*S-*
					r	BX249892	Ge029H11	Plasma membrane intrinsic protein (Aquaporin)	C	*S-*
					i	CR393662	RN63B06	Proton-exporting ATPase (EC 3.6.3.6)	C	*S+*
					r	BX679628	RS30A09	Root cap protein	C	*S-*
					r	BX681365	RS57G02	Tonoplast intrinsic protein	C	*S-*

Tableau Ax-6 : Liste des gènes mis en évidence uniquement dans l'expérience Stress Court. Les gènes sans fonction connue ne sont pas indiqués.

Effet : *effet significatif à l'origine de la détection du gène (C/P/PC/Ci/Pi) = facteur condition (C), provenance (P); P et C, C et interaction (i), P et i)* ; sens : *sens de variation (S+/S- = induction/repression lors du stress, M+/F+ = surexpression Maroc/France)*

Acc Number	ID	Annotation	Catégorie fonctionelle	effet	sens
BX680905	RS50E05	Fumarate hydratase, mitochondrial precursor (EC 4.2.1.2)	Energie	C	S+
CR392581	RN35A05	Isopentenyl-diphosphate delta-isomerase II (EC 5.3.3.2)	Energie	C	S-
BX676888	RN13E06	Carbonate dehydratase (EC 4.2.1.1)	Métabolisme	C	S-
BX677655	RN42G05	Flavonoid 3',5'-hydroxylase (EC 1.14.-.-)	Métabolisme	C	S-
AL750116	AS05H10	EF-HAND containing protein-like	Non-classé	C	S+
BX249906	Ge030B06	Hydrophobic protein RCI2 (Low temperature and salt induced)	Non-classé	C	S+
BX678993	RS18B02	IN2-2 protein	Non-classé	C	S-
BX253244	Ge080C01	Lin1294 protein, partial (9%)	Non-classé	C	S+
BX681731	RS63F03	Anther specific LAT52 protein precursor	Organisation cellulaire	C	S+
AL750525	RN03E03	MtN21 nodulin protein-like	Organisation cellulaire	C	S-
AL750218	AS07A11	Chitinase (EC 3.2.1.14)	Réponse au Stress	C	S+
BX250140	Ge032H07	Defender against cell death	Réponse au Stress	C	S+
BX681387	RS58C02	Licheninase precursor (EC 3.2.1.73)	Réponse au Stress	C	S+
CR392585	RN35A10	Peroxidase precursor (EC 1.11.1.7)	Réponse au Stress	C	S-
AL750567	RN03H10	60S ribosomal protein L14	Synthèse de protéines	C	S-
BX678294	RN69C07	Nuclear transport factor 2 (NTF-2)	Traffic intracellulaire	C	S-
BX681992	RS67G12	Fluconazole resistance protein 1	Transport	C	S+
BX680166	RS40A07	Pathogenesis-related protein	Réponse au Stress	Ci	
BX682915	Ge123D11	14-3-3 protein	Communication/Transduction du signal	P	M+
BX249546	Ge025C11	Calmodulin	Communication/Transduction du signal	P	M+
CR392472	RN32G05	Calmodulin-related protein 2	Communication/Transduction du signal	P	F+
BX254449	Ge098G01	Phospholipase D alpha 1 (PLD alpha 1) (Choline phosphatase 1) (Phosphatidylcholine-l	Communication/Transduction du signal	P	M+
BX249833	Ge029C02	Ras-like small monomeric GTP-binding protein, partial (65%)	Communication/Transduction du signal	P	M+
BX677537	RN40H04	Receptor protein kinase precursor	Communication/Transduction du signal	P	F+
CR393564	RN61G10	Sorcin	Communication/Transduction du signal	P	F+
BX681360	RS57F04	MERI-5 protein precursor (Xyloglucan:xyloglucosyl transferase (EC 2.4.1.207)) / Brassin	Croissance cellulaire, division	P	F+
BX677507	RN40E01	Nuclear scaffold-like protein p76	Croissance cellulaire, division	P	M+
BX681068	RS52G06	Cysteine proteinase RD21A precursor (EC 3.4.22.-)	Devenir des protéines	P	M+
AL750449	RN02F06	Peptidylprolyl isomerase (EC 5.2.1.8)	Devenir des protéines	P	M+
BX252329	Ge062H12	Proteasome subunit beta type 5 precursor (20S proteasome subunit E) (Proteasome ep	Devenir des protéines	P	F+
BX252055	Ge059B02	subtilisin-like serine proteinase - Arabidopsis thaliana {Arabidopsis thaliana;} , partial (7	Devenir des protéines	P	M+
BX681296	RS56F09	[Pyruvate dehydrogenase (Lipoamide)]-phosphatase (EC 3.1.3.43)	Energie	P	F+
BX249881	Ge029G11	ATP synthase gamma chain	Energie	P	M+
BX681437	RS59C04	Electron transfer flavoprotein-ubiquinone oxidoreductase	Energie	P	F+
BX681841	RS65F03	NADPH dehydrogenase (EC 1.6.99.1)	Energie	P	F+
BX681554	RS61B07	NADPH:quinone reductase (EC 1.6.5.5)	Energie	P	F+
CR394186	RN74H06	Phytocyanin	Energie	P	F+
BX255712	Ge008H07	Thioredoxin H-type	Energie	P	M+
AL750459	RN02G04	Ubiquinol-cytochrome C reductase	Energie	P	F+
BX678629	RS12C06	Ubiquinol-cytochrome C reductase complex 7.8 kDa protein	Energie	P	F+
BX677894	RN46B11	Alcohol dehydrogenase (EC 1.1.1.1)	Métabolisme	P	F+
BX252271	Ge062B09	Alcohol dehydrogenase class III (Glutathione-dependent formaldehyde dehydrogenase)	Métabolisme	P	M+
AL750794	RN07D07	Aldehyde dehydrogenase (NAD+) (EC 1.2.1.3)	Métabolisme	P	F+
BX681821	RS65D03	Aldehyde dehydrogenase (NAD+) (EC 1.2.1.3)	Métabolisme	P	F+
CR393484	RN60G02	Aldose 1-epimerase precursor (EC 5.1.3.3)	Métabolisme	P	F+
CR393571	RN61H07	Aldose 1-epimerase precursor (EC 5.1.3.3)	Métabolisme	P	F+
BX252563	Ge066E08	Aspartate aminotransferase, chloroplast precursor	Métabolisme	P	M+
BX682419	Ge116C12	Betaine-aldehyde dehydrogenase (EC 1.2.1.8)	Métabolisme	P	F+
BX249794	Ge027G10	Beta-ketoacyl-ACP synthetase I-2, partial (91%)	Métabolisme	P	F+
BE123499	NXNV_148_H6	Cellulose synthase	Métabolisme	P	M+
BX681531	RS60G10	Chalcone-flavonone isomerase	Métabolisme	P	F+
BX678890	RS16F01	Cholinephosphate cytidylyltransferase B (EC 2.7.7.15)	Métabolisme	P	M+
BX681204	RS55C05	Cysteine synthase (EC 2.5.1.47)	Métabolisme	P	F+
BX681612	RS61H11	dUTP pyrophosphatase (EC 3.6.1.23)	Métabolisme	P	F+
AL749773	AN07G08	Esterase D	Métabolisme	P	M+
BF049732	NXCI_106_D10	Fumarase	Métabolisme	P	M+
BE761959	NXCI_75_C7	Glutamate synthase	Métabolisme	P	M+
CR393516	RN61B10	Glutamate-1-semialdehyde 2,1-aminomutase (EC 5.4.3.8)	Métabolisme	P	F+
BX681127	RS54B10	Glutathione transferase (EC 2.5.1.18)	Métabolisme	P	F+
BX254883	Ge106G06	Histidine ammonia-lyase (Histidase) , partial (4%)	Métabolisme	P	F+
AL750713	RN06E03	L-ascorbate oxidase precursor (EC 1.10.3.3)	Métabolisme	P	F+
BF609706	NXSI_48_D6	malate dehydrogenase	Métabolisme	P	M+
BX253454	Ge083F12	Patatin	Métabolisme	P	F+
BX682063	RS69E03	Phenylalanine ammonia-lyase (EC 4.3.1.5)	Métabolisme	P	F+
BG039488	NXSI_99_F10	S-adenosyl-Met decarboxylase	Métabolisme	P	M+
AL751154	RS05F10	Serine/threonine specific protein phosphatase 2C (EC 3.1.3.16)	Métabolisme	P	F+
BX681087	RS53B06	SUR2 protein (Syringomycin response protein 2)	Métabolisme	P	F+
BX680092	RS39A02	T14P8.4 protein	Métabolisme	P	F+
BX682721	Ge120F04	Uridylate kinase (EC 2.7.4.3)	Métabolisme	P	M+
BX251533	Ge050H03	At1g11170/T28P6_16, partial (58%)	Non-classé	P	F+
BX680179	RS40B08	At2g44310 Hypothetical protein	Non-classé	P	F+
BX249996	Ge031B08	At3g03773 {Arabidopsis thaliana;} , partial (71%)	Non-classé	P	F+
BX248924	Ge002G04	At3g23090, partial (15%)	Non-classé	P	F+
CR392486	RN33A12	ATP-dependent Clp protease proteolytic subunit (EC 3.4.21.92)	Non-classé	P	F+
CR394213	RN75C03	Auxin-repressed 12.5 kDa protein	Non-classé	P	M+
BX252717	Ge069F09	B2 protein	Non-classé	P	M+
CR354614	Ge124A01	Blr1429 protein, partial (18%)	Non-classé	P	M+
BX254656	Ge102B08	CG33184-PA, isoform A, partial (9%)	Non-classé	P	F+
BX681067	RS52G04	Embryogenesis-associated protein EMB7	Non-classé	P	M+
BX677557	RN41B08	GASA5-like protein (Sb35)	Non-classé	P	F+

BX253820	Ge089A07	Hypothetical 10.4 kDa protein (Expressed protein)	Non-classé	P	*F+*
AL750551	RN03G06	Hypothetical 168.1 kDa protein	Non-classé	P	*F+*
BX679085	RS19F02	Hypothetical 20.6 kDa protein	Non-classé	P	*F+*
BX681024	RS52B04	Hypothetical 60.3 kDa protein	Non-classé	P	*F+*
AL751163	RS05G07	Hypothetical 78.8 kDa protein	Non-classé	P	*F+*
CR393785	RN65C08	Hypothetical 8.7 kDa protein	Non-classé	P	*F+*
BX680327	RS42A07	Hypothetical 85.6 kDa protein	Non-classé	P	*F+*
BX680229	RS40G08	Hypothetical protein (At1g76160/T23E18_10)	Non-classé	P	*F+*
CR393481	RN60F10	IN2-2 protein	Non-classé	P	*F+*
BX681544	RS61A06	Lateral organ boundaries (LOB) domain protein	Non-classé	P	*F+*
BX249968	Ge030H03	Latex-abundant protein	Non-classé	P	*F+*
BX254006	Ge092B02	Outer membrane protein (Fragment), partial (20%)	Non-classé	P	*F+*
BX250861	Ge042C04	OVCA2=CANDIDATE tumor suppressor, partial (9%)	Non-classé	P	*M+*
AL751090	RS05A04	Protein kinase (EC 2.7.1.-)	Non-classé	P	*F+*
CR394188	RN74H08	Protein kinase-like (EC 2.7.1.37)	Non-classé	P	*F+*
BX678411	RS08G05	RNA-binding protein KIAA0682	Non-classé	P	*F+*
BX252796	Ge070H02	SOUL-like protein, partial (10%)	Non-classé	P	*F+*
BX676834	RN12D04	T17B22.3 protein	Non-classé	P	*F+*
BX252614	Ge067H06	Histone H2A variant	Organisation cellulaire	P	*M+*
AL750508	RN03C09	Histone H4	Organisation cellulaire	P	*M+*
BX680100	RS39B01	Nodulin-like protein (Putative nodulin protein)	Organisation cellulaire	P	*F+*
BX680185	RS40C03	Pectate lyase (EC 4.2.2.2)	Organisation cellulaire	P	*F+*
BX250283	Ge034G10	Pectate lyase, partial (95%)	Organisation cellulaire	P	*F+*
BX251918	Ge057C01	Ubiquitin-conjugating enzyme (EC 6.3.2.19)	Organisation cellulaire	P	*M+*
BX682459	Ge116H08	Ubiquitin-conjugating enzyme (EC 6.3.2.19)	Organisation cellulaire	P	*M+*
AL750092	AS05F09	Ubiquitin-conjugating enzyme E2-17 kDa (EC 6.3.2.19)	Organisation cellulaire	P	*M+*
BX250267	Ge034E07	Ubiquitin-conjugating enzyme E2-17 kDa (EC 6.3.2.19)	Organisation cellulaire	P	*M+*
BX253993	Ge091H07	Ubiquitin-conjugating enzyme E2-17 kDa (EC 6.3.2.19)	Organisation cellulaire	P	*M+*
BX681382	RS58B04	Ubiquitin-protein ligase 1	Organisation cellulaire	P	*F+*
BX251094	Ge045C02	Abscisic stress ripening protein	Réponse au Stress	P	*M+*
BX682324	Ge115A12	Abscisic stress ripening protein	Réponse au Stress	P	*M+*
BX254489	Ge099C04	ATP-dependent DNA helicase (EC 3.6.1.-), partial (2%)	Réponse au Stress	P	*F+*
CR392715	RN37A08	Chitinase precursor (EC 3.2.1.14)	Réponse au Stress	P	*F+*
AL750400	RN02B01	Copper chaperone for Superoxide dismutase	Réponse au Stress	P	*F+*
CR392961	RN50F01	Dehydration-responsive protein RD22 precursor	Réponse au Stress	P	*F+*
BX255778	Ge009F07	Disease resistance protein-like	Réponse au Stress	P	*F+*
BX253926	Ge090G07	Disease resistance-like protein (Fragment), partial (16%)	Réponse au Stress	P	*F+*
BX253877	Ge089H04	DNA repair protein RAD16, partial (6%)	Réponse au Stress	P	*F+*
BX681362	RS57F06	DnaJ protein homolog	Réponse au Stress	P	*F+*
BX249660	Ge026G07	ENDOCHITINASE (EC 3.2.1.14)	Réponse au Stress	P	*F+*
BX678613	RS12A03	Gamma-thionin homolog precursor	Réponse au Stress	P	*M+*
BX666039	RN08H11	Late embryogenesis abundant protein	Réponse au Stress	P	*F+*
BX681471	RS60A05	Late embryogenesis abundant protein (Fragment)	Réponse au Stress	P	*F+*
BX251839	Ge056B11	Leucine-rich repeat-containing protein 2	Réponse au Stress	P	*M+*
BX254258	Ge095G04	LUMINAL BINDING PROTEIN	Réponse au Stress	P	*M+*
BX681431	RS59B06	Pathogenesis-related protein PR-4B precursor	Réponse au Stress	P	*F+*
BX254293	Ge096C10	Peptide methionine sulfoxide reductase (EC 1.8.4.6)	Réponse au Stress	P	*M+*
BX254690	Ge102G05	peroxidase	Réponse au Stress	P	*F+*
AL750365	RN01F10	Peroxidase precursor (EC 1.11.1.7)	Réponse au Stress	P	*M+*
AL751229	RS06E04	Peroxidase precursor (EC 1.11.1.7)	Réponse au Stress	P	*F+*
BX254875	Ge106E11	Probable aquaporin, complete	Réponse au Stress	P	*M+*
CR392543	RN33H08	25S rRNA	Synthèse de protéines	P	*F+*
AL750386	RN01H09	30S ribosomal protein S16	Synthèse de protéines	P	*F+*
CR392159	RN27C08	40S ribosomal protein	Synthèse de protéines	P	*F+*
BX252371	Ge063E06	60S ribosomal protein L31	Synthèse de protéines	P	*M+*
BX682787	Ge121G09	60S ribosomal protein L31	Synthèse de protéines	P	*M+*
BX681771	RS64D09	60S ribosomal protein L34	Synthèse de protéines	P	*M+*
BX252657	Ge068F09	60S ribosomal protein L35A	Synthèse de protéines	P	*M+*
BX678464	RS09E08	60S ribosomal protein L41, mitochondrial precursor	Synthèse de protéines	P	*F+*
BX680013	RS38B06	Gb|AAD26879.1	Synthèse de protéines	P	*M+*
BX679475	RS26E10	Ribosomal protein L10	Synthèse de protéines	P	*F+*
BX676912	RN13G11	Acyl carrier protein, chloroplast precursor	Traffic intracellulaire	P	*M+*
BX254611	Ge101C06	ADP-ribosylation factor	Traffic intracellulaire	P	*M+*
CR394037	RN72F07	Coatomer beta' subunit (Beta'-coat protein)	Traffic intracellulaire	P	*F+*
BX681245	RS56A07	Dynamin-like protein	Traffic intracellulaire	P	*F+*
CR394145	RN74D07	GTP-binding protein SAR	Traffic intracellulaire	P	*F+*
BX682440	Ge116F08	Nuclear transport factor 2 (NTF-2).	Traffic intracellulaire	P	*M+*
BX254888	Ge106H02	Ras-related protein Rab11B	Traffic intracellulaire	P	*M+*
BX250119	Ge032F08	Auxin response factor 1, partial (25%)	Transcription	P	*F+*
BX255626	Ge007H04	Small zinc finger-like protein, complete	Transcription	P	*M+*
CR392874	RN49D09	Zinc finger protein 364	Transcription	P	*F+*
CR393875	RN66E08	Calcium-transporting ATPase, plasma membrane-type	Transport	P	*F+*
CR393619	RN62F01	Peptide transporter PTR2-B	Transport	P	*F+*
BX682634	Ge119A01	Elongation factor 1-gamma (EF-1-gamma)	Croissance cellulaire, division	PC	*S-M+*
BX681080	RS53A02	Acetylornithine deacetylase (EC 3.5.1.16)	Devenir des protéines	PC	*S+M+*
BX252908	Ge074G03	Cathepsin B-like cysteine proteinase, partial (68%)	Devenir des protéines	PC	*S+M+*
BX679916	RS36C01	Metallophosphatase	Devenir des protéines	PC	*S-M+*
CR392614	RN35E03	Acid phosphatase	Métabolisme	PC	*S-M+*
CR392259	RN30A11	Acid phosphatase precursor 1 (EC 3.1.3.2)	Métabolisme	PC	*S-M+*
BX253047	Ge077B10	Aminotransferase	Métabolisme	PC	*S+M+*
BX667535	RN09C02	Carbon catabolite derepressing protein kinase	Métabolisme	PC	*S+M+*
BX254251	Ge095F04	cytosolic aldehyde dehydrogenase {Oryza sativa (japonica cultivar-group);} , partial (96%	Métabolisme	PC	*S+M+*
BX255766	Ge009E07	metallothionein	Métabolisme	PC	*S+M+*
BX252844	Ge071F04	metallothionein-like protein	Métabolisme	PC	*S+M+*
CR393796	RN65E04	Pinene synthase (EC 4.2.3.14)	Métabolisme	PC	*S-M+*
BX255226	Ge113B09	TrPRP2, partial (67%)	Non-classé	PC	*S-M+*
BX679616	RS29G03	Histone H4	Organisation cellulaire	PC	*S-M+*
AL750621	RN05E06	Pectate lyase (EC 4.2.2.2)	Organisation cellulaire	PC	*S-M+*
BX255377	Ge005B04	Elastin precursor - human (fragment) {Homo sapiens;} , partial (50%)	Réponse au Stress	PC	*S+M+*
CR394461	RN79E12	Ethylene-responsive protein, putative	Réponse au Stress	PC	*S+F+*
BX677964	RN47C10	Late embryogenesis abundant protein	Réponse au Stress	PC	*S-M+*
BX678491	RS10A02	Pathogenesis-related protein 5 precursor (PR-5)	Réponse au Stress	PC	*S-M+*
CR393156	RN53C07	Peroxidase precursor (EC 1.11.1.7)	Réponse au Stress	PC	*S-M+*
BX678596	RS11G02	Peroxidase precursor (EC 1.11.1.7)	Réponse au Stress	PC	*S-M+*
BX249687	Ge027B05	FKBP-like protein (Arabidopsis thaliana genomic DNA, chromosome 3, P1 clone: MIL23)	Transcription	PC	*S+M+*
BX252953	Ge075E07	RING finger protein 26	Transcription	PC	*S+M+*
BX251681	Ge053H04	Germin(Oxalate oxidase)-like protein, partial (24%)	Croissance cellulaire, division	Pi	

Tableau Ax-7 : Liste des gènes mis en évidence uniquement dans l'expérience Stress Long. Les gènes sans fonction connue ne sont pas indiqués. Légende cf. Tab. Ax-6

Acc Number	ID	Annotation	Catégorie fonctionelle	effet	sens
BX255369	Ge005A06	Glycine-rich RNA-binding protein, complete	Communication/Transduction du signal	C	S+
CR393473	RN60F01	Phospholipase D alpha 1 precursor (EC 3.1.4.4)	Communication/Transduction du signal	C	S+
BX251900	Ge057A05	Clt1, partial (98%)	Communication/Transduction du signal	C	S-
BX677790	RN44G03	Phospholipase C	Communication/Transduction du signal	C	S+
BX681632	RS62C04	Receptor protein kinase TMK1 precursor	Communication/Transduction du signal	C	S+
BX681723	RS63E04	Inositol-1(or 4)-monophosphatase (EC 3.1.3.25)	Communication/Transduction du signal	C	S-
BX679850	RS34H11	Calmodulin	Communication/Transduction du signal	c	S-
AL751104	RS05B06	DNA polymerase III	Croissance cellulaire, division	C	S+
BX250683	Ge040A04	Fasciclin-like arabinogalactan protein FLA8	Croissance cellulaire, division	C	S-
BX251676	Ge053G11	Xyloglucan:xyloglucosyl transferase (EC 2.4.1.207)	Croissance cellulaire, division	C	S-
CR393282	RN54H02	Germin-like protein subfamily 1 member 7 precursor	Croissance cellulaire, division	C	S+
CR392893	RN49F07	Fasciclin-like arabinogalactan-protein 2	Croissance cellulaire, division	C	S-
BX667570	RN09F07	Presenilin	Croissance cellulaire, division	C	S+
BX677740	RN44B02	Cell division cycle protein 48 homolog (Valosin containing protein homolog) (vcp)	Croissance cellulaire, division	C	S-
CR393887	RN66G01	MERI-5 protein precursor (Xyloglucan:xyloglucosyl transferase (EC 2.4.1.207)) / Brassin	Croissance cellulaire, division	C	S-
BX680304	RS41G02	Structural maintenance of chromosomes (SMC)-like protein	Croissance cellulaire, division	C	S+
AL750591	RN05B11	Xyloglucan:xyloglucosyl transferase (EC 2.4.1.207)	Croissance cellulaire, division	C	S-
BX677108	RN17G03	Expansin-like precursor	Croissance cellulaire, division	c	S+
CR393041	RN51H01	26S protease regulatory subunit 4 homolog	Devenir des protéines	C	S+
BX677823	RN45B11	Chloroplast nucleoid DNA binding protein	Devenir des protéines	C	S-
AL749954	AS03A12	Thimet oligopeptidase (EC 3.4.24.15)	Devenir des protéines	C	S-
BX681320	RS57A03	Chloroplast nucleoid DNA binding protein, putative	Devenir des protéines	C	S+
AL750564	RN03H07	Alpha galactosyltransferase-like protein	Devenir des protéines	C	S-
BX253484	Ge084A09	Peptidyl-prolyl cis-trans isomerase (EC 5.2.1.8) (Cyclophilin)	Devenir des protéines	C	S-
BX679023	RS18E12	Signal peptidase I (EC 3.4.21.89)	Devenir des protéines	C	S+
BX681672	RS62G04	Carboxypeptidase C (EC 3.4.16.5)	Devenir des protéines	C	S+
BX667552	RN09D11	Chloroplast nucleoid DNA binding protein	Devenir des protéines	C	S-
BX679155	RS20F02	Proteasome endopeptidase complex subunit alpha type 6 (EC 3.4.25.1)	Devenir des protéines	C	S+
BX249855	Ge029E02	subtilisin-like proteinase - Arabidopsis thaliana {Arabidopsis thaliana;} , partial (12%)	Devenir des protéines	c	S-
AL749654	AN05E01	Oxygen-evolving enhancer protein 1, chloroplast precursor	Energie	C	S+
BX678416	RS08G10	Chlorophyll A-B binding protein, chloroplastic	Energie	C	S+
AL749666	AN05G02	Photosystem I reaction center subunit N, chloropla	Energie	C	S+
AL749783	AS01A02	H(+)-transporting two-sector ATPase delta subunit (EC 3.6.3.14)	Energie	C	S+
BX251416	Ge049C07	Oxygen-evolving enhancer protein 3-1, chloroplast precursor	Energie	C	S+
CR392217	RN28D05	Basic blue protein (Cusacyanin) (Plantacyanin)	Energie	C	S-
BX681414	RS58H06	Hydroquinone glucosyltransferase (EC 2.4.1.218)	Energie	C	S+
BX677890	RN46B05	Pyruvate kinase (EC 2.7.1.40)	Energie	C	S+
BX680864	RS49H06	Pyruvate decarboxylase (EC 4.1.1.1)	Energie	C	S-
BX681089	RS53C12	Phosphogluconate dehydrogenase (decarboxylating) (EC 1.1.1.44)	Energie	C	S-
BX680080	RS38H04	AAA-type ATPase	Energie	C	S+
BX249635	Ge026E03	Phosphoenolpyruvate carboxylase , partial (50%)	Energie	C	S+
BX251974	Ge058A01	ATP synthase beta chain, mitochondrial precursor	Energie	C	S-
BX249115	Ge020B11	Dihydrolipoyllysine-residue acetyltransferase component of pyruvate dehydrogenase con	Energie	C	S-
BX252436	Ge064E07	Diphosphate--fructose-6-phosphate 1-phosphotransferase beta subunit (EC 2.7.1.90) (P	Energie	C	S-
CR392718	RN37A12	NADH dehydrogenase (quinone) 23 kDa subunit (1.6.99.5)	Energie	C	S+
AL750697	RN06C10	Thiazole biosynthetic enzyme, chloroplast precursor	Energie	C	S-
BX252441	Ge064F03	Ubiquinol-cytochrome C reductase iron-sulfur subunit (EC 1.10.2.2)	Energie	C	S-
BX678396	RS08E10	Alternative oxidase 1, mitochondrial precursor	Energie	c	S+
BX682159	RS72E06	Flavonoid 3',5'-hydroxylase (EC 1.14.-.-)	Métabolisme	C	S+
BX676897	RN13F04	Syringomycin biosynthesis enzyme-like protein	Métabolisme	C	S+
BX678400	RS08F03	Glutathione transferase (EC 2.5.1.18)	Métabolisme	C	S+
BX682993	Ge126F07	UDP-glucose 4-epimerase (EC 5.1.3.2)	Métabolisme	C	S-
BX253036	Ge077A02	Methionine synthase, partial (63%)	Métabolisme	C	S-
BX679860	RS35A06	2-alpha-hydroxytaxane 2-O-benzoyltransferase (EC 2.3.1.166)	Métabolisme	C	S-
BX681925	RS66H01	Hydroxymethylglutaryl-CoA synthase (EC 2.3.3.10)	Métabolisme	C	S+
BX253114	Ge078C04	Alkaline alpha galactosidase II, partial (84%)	Métabolisme	C	S+
BX679422	RS25B09	Cysteine synthase, chloroplast precursor (EC 2.5.1.47)	Métabolisme	C	S+
CR392473	RN32G07	Glutamate synthase (NADH) (Sb14) (EC 1.4.1.14)(Picea mariana)	Métabolisme	C	S+
AL751120	RS05C10	S-adenosylmethionine-dependent methyltransferase T12H1.6 protein	Métabolisme	C	S+
BX682426	Ge116E03	Adenosylhomocysteinase (EC 3.3.1.1)	Métabolisme	C	S-
BX678841	RS15G11	4-hydroxyphenylpyruvate dioxygenase (EC 1.13.11.27)	Métabolisme	C	S+
BX249550	Ge025D03	Citrate synthase	Métabolisme	C	S+
BX250471	Ge037D04	(1-4)-beta-mannan endohydrolase precursor (EC 3.2.1.78).	Métabolisme	C	S-
BX680234	RS40H04	Alcohol oxidase 2	Métabolisme	C	S+
BX252145	Ge060D11	Hypothetical protein yfbG	Métabolisme	C	S-
BX000600	GP	GP	Métabolisme	C	S-
CR392294	RN30E11	Omega-6 fatty acid desaturase, endoplasmic reticulum	Métabolisme	C	S-
AL750486	RN03A11	Flavonol 3-O-glucosyltransferase (EC 2.4.1.91)	Métabolisme	C	S+
CR392209	RN28C02	Phenylalanine ammonia-lyase (EC 4.3.1.5)	Métabolisme	C	S+
BX250516	Ge038A02	Beta-galactosidase , partial (37%)	Métabolisme	C	S-
CR394072	RN73D02	Cytochrome P450 77A3 (EC 1.14.-.-)	Métabolisme	C	S+
BX677977	RN47F03	Beta-fructofuranosidase (EC 3.2.1.26), cell wall isozyme precursor	Métabolisme	C	S+
BX679517	RS27D07	Inosine-5'-monophosphate dehydrogenase (EC 1.1.1.2)	Métabolisme	C	S+
BX677473	RN40A01	Glutamate decarboxylase (EC 4.1.1.15)	Métabolisme	C	S-
BX677417	RN38F06	Short-chain type dehydrogenase/reductase (EC 1.-.-.-)	Métabolisme	C	S+
BX254966	Ge108C11	transitional endoplasmic reticulum ATPase - Arabidopsis thaliana {Arabidopsis thaliana;}	Métabolisme	C	S-
BE607163	NXCI_37_F8	adenosine kinase (AK)	Métabolisme	C	S-
BX682627	Ge118H04	Shikimate 5-dehydrogenase (EC 1.1.1.25)	Métabolisme	C	S-
BX679683	RS31C02	Cytochrome P450 (EC 1.14.-.-), putative	Métabolisme	C	S+
CR392624	RN35F03	Caffeoyl-CoA O-methyltransferase (EC 2.1.1.104)	Métabolisme	C	S-
BX252860	Ge071H04	5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase	Métabolisme	C	S-
BX678049	RN55F05	Phenylalanine ammonia-lyase (EC 4.3.1.5)	Métabolisme	C	S-
BX678013	RN55B08	Trans-cinnamate 4-monooxygenase (EC 1.14.13.11)	Métabolisme	C	S-
BX678666	RS12H05	Alpha-1,4-glucan-protein synthase (UDP-forming) (EC 2.4.1.112)	Métabolisme	C	S-
BX679708	RS32A02	Glutathione transferase (EC 2.5.1.18)	Métabolisme	C	S+
BX677699	RN43C10	Chaperonin-60 beta subunit precursor, partial (91%)	Métabolisme	C	S+
BX680824	RS49A07	Arginine methyltransferase-like protein (EC 2.1.1.-)	Métabolisme	C	S+
BX679575	RS28F02	Alpha-1,4-glucan-protein synthase (UDP-forming) (EC 2.4.1.112)	Métabolisme	C	S-
AL749965	AS03B12	Metallothionein-like protein EMB30	Métabolisme	C	S-
BX681070	RS52G08	Trans-cinnamate 4-monooxygenase (EC 1.14.13.11)	Métabolisme	C	S-
BX254773	Ge104C09	Sterol-C-methyltransferase.	Métabolisme	C	S-
BX677309	RN20D07	Glutathione transferase (EC 2.5.1.18)	Métabolisme	C	S+
BX253959	Ge091C11	Putative receptor protein kinase TMK1 precursor (EC 2.7.1.-)	Métabolisme	C	S+
BX254093	Ge093C03	caffeate O-methyltransferase [Pinus taeda]	Métabolisme	C	S-
CR394120	RN74B04	Serine/threonine specific protein phosphatase 2C (EC 3.1.3.16)	Métabolisme	C	S+
AL749834	AS01F08	Naringenin-chalcone synthase (EC 2.3.1.74)	Métabolisme	C	S-
CR392095	RN11G12	Naringenin-chalcone synthase (EC 2.3.1.74)	Métabolisme	C	S-
BX678822	RS15F01	Acyltransferase homolog	Métabolisme	C	S+
BX681828	RS65D10	Flavonol 3-O-glucosyltransferase 5 (EC 2.4.1.91)	Métabolisme	C	S+

BX249364	Ge023B09	glutathione-dependent formaldehyde dehydrogenase {Arabidopsis thaliana;} , partial (43	Métabolisme	C	S+
BX249835	Ge029C07	Adenosine kinase (EC 2.7.1.20)	Métabolisme	C	S-
BX252930	Ge075B02	ATP citrate lyase a-subunit, partial (84%)	Métabolisme	C	S-
BX679924	RS36D05	Iron(III)-zinc(II) purple acid phosphatase precursor	Métabolisme	C	S+
AW784085	NXNV_117_F2	glyceraldehyde-3-P-dehydrogenase	Métabolisme	C	S-
AL749938	AS02H06	Naringenin 3-dioxygenase (EC 1.14.11.9)	Métabolisme	C	S+
BG040863	NXSI_116_A9	adenosine kinase (AK)	Métabolisme	C	S-
BX248894	Ge002D03	Lanatoside 15'-O-acetylesterase precursor, partial (28%)	Métabolisme	C	S-
BX249460	Ge024C08	Mitochondrial aldehyde dehydrogenase, partial (91%)	Métabolisme	C	S+
BX679131	RS20C02	Leucoanthocyanidin dioxygenase (EC 1.-.-.-) (LDOX)	Métabolisme	C	S-
BX682428	Ge116E05	Inositolphosphorylceramide-B C-26 hydroxylase	Métabolisme	C	S-
BG039035	NXSI_24_H1	Cellulose synthase	Métabolisme	C	S-
CR393529	RN61D03	Malate dehydrogenase, mitochondrial precursor	Métabolisme	C	S-
BX679620	RS29H05	Serine acetyltransferase	Métabolisme	C	S+
BX253613	Ge085H08	Adenosylhomocysteinase (EC 3.3.1.1)	Métabolisme	C	S-
CR392866	RN49C12	Metallothionein-like protein EMB30	Métabolisme	c	S-
BX680340	RS42B09	Naringenin-chalcone synthase (EC 2.3.1.74)	Métabolisme	c	S-
AL750160	AS06D09	Metallothionein-like protein	Métabolisme	c	S-
BX680681	RS47B09	Anter-specific proline-rich protein / Lanatoside 15'-O-acetylesterase precursor	Métabolisme	c	S-
BX678852	RS16A04	Caffeate O-methyltransferase (EC 2.1.1.68)	Métabolisme	c	S-
CR393272	RN54G03	Transaldolase (EC 2.2.1.2)	Métabolisme	c	S+
BX254041	Ge092E09	serine protease	Métabolisme	c	S-
BX679917	RS36C02	Protein kinase (EC 2.7.1.-)	Métabolisme	c	S+
BX680601	RS45G02	Hypothetical 25.9 kDa protein	Non-classé	C	S+
BX249398	Ge023F01	Auxin-repressed 12.5 kd protein	Non-classé	C	S+
BX681172	RS54G01	Endosperm-specific protein-like protein	Non-classé	C	S-
BX677604	RN41H04	Hypothetical 20.8 kDa protein	Non-classé	C	S+
BX249499	Ge024G05	Pherophorin-dz1 protein precursor, partial (7%)	Non-classé	C	S+
BX677573	RN41D05	Yippee-like protein	Non-classé	C	S+
BX255417	Ge005E10	P20-Arc, partial (92%)	Non-classé	C	S-
BX680058	RS38F06	F14L17.20 protein	Non-classé	C	S-
BX677222	RN19E10	Hypothetical 40.2 kDa protein	Non-classé	C	S+
BX250614	Ge039B07	phosphate-induced protein	Non-classé	C	S-
BX254648	Ge102A03	At1g64110/F22C12_22, partial (47%)	Non-classé	C	S+
BX680652	RS46F06	hsr203J	Non-classé	C	S+
BX252300	Ge062E09	Nuc-1 negative regulatory protein preg	Non-classé	C	S+
BX252532	Ge066B06	Potyvirus VPg interacting protein, partial (55%)	Non-classé	C	S+
BX679201	RS21D07	SelT-like protein precursor	Non-classé	C	S+
BX249846	Ge029D04	At2g43970/F6E13.10 {Arabidopsis thaliana;} , partial (41%)	Non-classé	C	S+
AL749952	AS03A10	Hypothetical 38.0 kDa protein (Putative dehydrogenase)	Non-classé	C	S-
BX254155	Ge094B08	Hypothetical protein	Non-classé	C	S+
AL751305	RS07C12	Hypothetical 58.9 kDa protein	Non-classé	C	S+
BX251850	Ge056D01	Hypothetical protein	Non-classé	C	S+
BX680432	RS43C07	Pinus taeda clone PtaAGP5	Non-classé	C	S-
BX254410	Ge098C07	Glycine-rich protein	Non-classé	C	S-
BX680674	RS47B01	Hypothetical 34.5 kDa protein	Non-classé	C	S+
BX678345	RN69H07	Vegetatible incompatibility protein HET-E-1	Non-classé	C	S-
BX682681	Ge120A02	Anter-specific proline-rich protein	Non-classé	C	S-
BX249830	Ge029B11	At1g07510/F22G5_9, partial (32%)	Non-classé	C	S-
CR392178	RN27F05	Hypothetical 68.3 kDa protein	Non-classé	C	S-
BX784344	Ge133B11	Wsv162 (WSSV218), partial (23%)	Non-classé	C	S+
BX678617	RS12A09	IN2-1 protein	Non-classé	C	S+
BX254347	Ge097C07	CoA-thioester hydrolase CHY1 (3-hydroxyisobutyryl-coenzyme A hydrolase), partial (79%	Non-classé	C	S+
BX254994	Ge108H10	Phosphate-induced protein	Non-classé	C	S-
BX679459	RS26C01	PGPD14 protein (Pollen germination related protein)	Non-classé	C	S+
BX680979	RS51F01	Hypothetical 19.9 kDa protein	Non-classé	C	S-
BX254033	Ge092D12	Remorin 1, partial (57%)	Non-classé	C	S+
BX255014	Ge109C07	Glucose-1-phosphate adenylyltransferase small subunit, chloroplast precursor (ADP-glu	Non-classé	C	S-
BX252312	Ge062G04	33 kDa secretory protein-like, partial (13%)	Non-classé	C	S-
BX680404	RS43A10	F12F1.23 protein (Hypothetical 36.7 kDa protein)	Non-classé	C	S-
BX682334	Ge115B11	Surfeit 1.	Non-classé	C	S-
BX680987	RS51G11	Hyphally regulated protein precursor	Non-classé	C	S+
BX681915	RS66F09	Embryogenesis-associated protein EMB3	Non-classé	C	S+
CR392445	RN32B03	ATAF1-like protein (Fragment)	Non-classé	C	S+
CR393456	RN60D06	Serine-rich protein	Non-classé	C	S+
BX250697	Ge040B09	F16A14.7, partial (50%)	Non-classé	C	S+
AL751202	RS06C01	P0022F10.3 protein	Non-classé	c	S-
BX680139	RS39F01	GPI-anchored protein	Non-classé	c	S-
BX678353	RS08A05	Hypothetical 12.0 kDa protein	Non-classé	c	S-
BX254445	Ge098F09	high mobility group protein MNB1b - maize (fragment) {Zea mays;} , partial (52%)	Non-classé	c	S-
BX255465	Ge006B03	acetone-cyanohydrin lyase [imported] - Arabidopsis thaliana, partial (37%)	Non-classé	c	S-
BX253001	Ge076C12	Tubulin alpha chain	Organisation cellulaire	C	S-
BX252968	Ge075G10	Tubulin alpha chain	Organisation cellulaire	C	S-
BX249241	Ge021G02	Tubulin alpha chain	Organisation cellulaire	C	S-
BX000608	Tubuline	Tubuline	Organisation cellulaire	C	S-
BX252002	Ge058D06	Tubulin beta chain	Organisation cellulaire	C	S-
BX680247	RS41A09	Glucan endo-1,3-beta-glucosidase precursor	Organisation cellulaire	C	S-
BX677552	RN41B01	Glycosyl transferase-like protein	Organisation cellulaire	C	S+
BX681614	RS62A02	Tubulin beta chain	Organisation cellulaire	C	S-
BX254406	Ge098C03	Ubiquitin	Organisation cellulaire	C	S-
BX250930	Ge043A10	MERI-5 protein precursor (Endo-xyloglucan transferase) / Brassinosteroid-regulated prot	Organisation cellulaire	C	S-
BX677988	RN47H03	Tomato extensin mRNA	Organisation cellulaire	C	S-
BX254985	Ge108G08	Ubiquitin	Organisation cellulaire	C	S-
AL751004	RS03A03	Glucan endo-1,3-beta-glucosidase precursor	Organisation cellulaire	C	S-
BX682172	RS73A01	Histone H2A	Organisation cellulaire	C	S+
BX677123	RN18A04	Histone H2B	Organisation cellulaire	C	S-
BX682259	Ge114B12	Histone H2A	Organisation cellulaire	C	S-
CR393797	RN65E06	Tubulin beta chain	Organisation cellulaire	C	S-
BX682714	Ge120E04	Cellulose synthase	Organisation cellulaire	C	S-
BX250835	Ge041H12	Ubiquitin carboxyl-terminal hydrolase (EC 3.1.2.	Organisation cellulaire	C	S+
BX255342	Ge004F07	Actin	Organisation cellulaire	C	S-
CR393223	RN54B06	Heterochromatin-like protein	Organisation cellulaire	c	S+
BX253583	Ge085E07	ENDOGLUCANASE	Organisation cellulaire	c	S-
CR393703	RN63G04	Ubiquitin conjugating enzyme spm2	Organisation cellulaire	c	S+
BX252682	Ge069A05	Cellulose synthase	Organisation cellulaire	c	S-
BX678362	RS08B03	Protein-methionine-S-oxide reductase (EC 1.8.4.6)	Réponse au Stress	C	S+
AL751212	RS06C11	Chaperone protein dnaJ	Réponse au Stress	C	S+
BX681166	RS54F02	NBS/LRR	Réponse au Stress	C	S+
CR392722	RN37B06	Peroxidase precursor (EC 1.11.1.7)	Réponse au Stress	C	S-
BX252031	Ge058G08	arabinogalactan protein [Pinus taeda]	Réponse au Stress	C	S-
BX681832	RS65E04	Pathogenesis-related protein PRB1-2 precursor	Réponse au Stress	C	S+
BX252074	Ge059D03	arabinogalactan protein	Réponse au Stress	C	S-
AL750545	RN03F11	Catalase isozyme (EC 1.11.1.6).	Réponse au Stress	C	S+
BX678152	RN57A03	Peroxidase precursor (EC 1.11.1.7)	Réponse au Stress	C	S-

BX249652	Ge026F11	ASR protein, partial (39%)	Réponse au Stress	C	*S+*
BX254281	Ge096B03	Ripening-associated protein (Fragment)	Réponse au Stress	C	*S+*
BX252277	Ge062C04	Harpin-induced protein	Réponse au Stress	C	*S-*
BX251932	Ge057D07	arabinogalactan protein	Réponse au Stress	C	*S-*
BX678372	RS08C03	Cationic peroxidase precursor (EC 1.11.1.7)	Réponse au Stress	C	*S+*
BX681263	RS56C03	EDS1-like protein	Réponse au Stress	C	*S+*
BX683057	Ge127F05_b	Aluminum-induced protein-like protein, partial (65%)	Réponse au Stress	C	*S+*
BX680673	RS47B12	Hypothetical protein (DNAJ domain-containing unknown protein)	Réponse au Stress	C	*S+*
BX681356	RS57E07	Late embryogenesis abundant protein	Réponse au Stress	C	*S+*
BX250350	Ge035G11	Wound-induced basic protein	Réponse au Stress	C	*S-*
BX251025	Ge044C12	Cyclophilin, partial (98%)	Réponse au Stress	C	*S-*
BX681402	RS58E06	DnaJ-like protein	Réponse au Stress	C	*S+*
AL750457	RN02G02	Protein-methionine-S-oxide reductase (EC 1.8.4.6)	Réponse au Stress	C	*S+*
BX666054	RN09B03	Abscisic stress ripening protein	Réponse au Stress	C	*S+*
CR393209	RN53H11	Heat shock protein STI1	Réponse au Stress	C	*S+*
BX677677	RN43A04	9-cis-epoxycarotenoid dioxygenase	Réponse au Stress	C	*S+*
BX677784	RN44F10	Copper chaperone for Superoxide dismutase	Réponse au Stress	C	*S+*
BX255310	Ge004C06	catalase 1 [Pinus pinea]	Réponse au Stress	C	*S+*
BX251044	Ge044E10	1-aminocyclopropane-1-carboxylate oxidase (EC 1.-.-.-)	Réponse au Stress	C	*S+*
BX252471	Ge065A11	ENDOCHITINASE (EC 3.2.1.14)	Réponse au Stress	C	*S-*
BX679992	RS37H01	Related to chitinase 3 precursor protein	Réponse au Stress	C	*S+*
CR393726	RN64C02	Chaperone protein dnaJ	Réponse au Stress	c	*S+*
BX680225	RS40G02	Peroxidase precursor (EC 1.11.1.7)	Réponse au Stress	c	*S+*
CR394442	RN79C01	Disease resistance response protein 206	Réponse au Stress	c	*S-*
BX253505	Ge084D05	Pinus taeda water deficit inducible protein LP3	Réponse au Stress	c	*S-*
BX252884	Ge074D02	Similarity to ABA induced plasma membrane protein, partial (52%)	Réponse au Stress	c	*S-*
CR392980	RN51A02	40S ribosomal protein S8	Synthèse de protéines	C	*S+*
BX249385	Ge023D08	60S ribosomal protein	Synthèse de protéines	C	*S-*
AL750653	RN05H02	60S ribosomal protein L27a	Synthèse de protéines	C	*S-*
BX251853	Ge056D04	40S ribosomal protein S27	Synthèse de protéines	C	*S-*
CR392876	RN49D11	40S ribosomal protein S16	Synthèse de protéines	C	*S-*
BX255767	Ge009E08	60S ribosomal protein L11	Synthèse de protéines	C	*S-*
BX680147	RS39G10	RNA helicase	Synthèse de protéines	C	*S+*
BX255132	Ge111E05	40S ribosomal protein S23	Synthèse de protéines	C	*S-*
CR393754	RN64G04	40S ribosomal protein S2	Synthèse de protéines	C	*S-*
BX251593	Ge051G02	Eukaryotic peptide chain release factor subunit 1	Synthèse de protéines	C	*S+*
BX253123	Ge078D03	40S ribosomal protein S9 (S7)	Synthèse de protéines	C	*S-*
BX252922	Ge074H10	60S ribosomal protein	Synthèse de protéines	C	*S-*
BX678954	RS17D12	small subunit ribosomal RNA gene	Synthèse de protéines	C	*S+*
BX251297	Ge047H03	60S ribosomal protein L10	Synthèse de protéines	C	*S-*
BX678737	RS14D11	Myosin V	Traffic intracellulaire	C	*S+*
CR394081	RN73D11	Syntaxin 71	Traffic intracellulaire	C	*S+*
BX677127	RN18A09	GTP-binding nuclear protein RAN-A1	Traffic intracellulaire	C	*S-*
BX252360	Ge063D06	Sec13p, partial (90%)	Traffic intracellulaire	c	*S+*
CR394310	RN76F04	Transcription factor HBP-1b homolog	Transcription	C	*S+*
BX680486	RS44B07	Poly(A)-binding protein	Transcription	C	*S-*
BX250756	Ge040H03	F11M15.7 protein	Transcription	C	*S+*
BX251702	Ge054B09	Zinc-finger protein NOLZ1, partial (3%)	Transcription	C	*S-*
BX678438	RS09B05	Homeobox-leucine zipper protein	Transcription	C	*S+*
BX682945	Ge123H04	Zinc-finger protein, partial (53%)	Transcription	C	*S+*
BX252601	Ge067D04	Abscisic acid insensitive 5 (ABI5) (BZIP protein), partial (17%)	Transcription	C	*S-*
BX681454	RS59E06	Nuclear polyadenylated RNA-binding protein NAB4	Transcription	C	*S-*
BX680112	RS39C05	Nuclear polyadenylated RNA-binding protein	Transcription	C	*S-*
BX682918	Ge123E04	Phenylcoumaran benzylic ether reductase homolog TH1, complete	Transcription	C	*S-*
BX681072	RS52H08	Zinc finger protein 216	Transcription	C	*S+*
CR392290	RN30E06	DNA-binding protein ESCAROLA	Transcription	C	*S-*
BX681312	RS56H05	Lil3 protein	Transcription	C	*S+*
BX252302	Ge062E12	RNA-binding protein RZ-1, partial (9%)	Transcription	c	*S+*
AL750557	RN03G12	Plasma membrane H+ ATPase (EC 3.6.1.3)	Transport	C	*S+*
BX680183	RS40C01	Membrane transporter protein	Transport	C	*S+*
BX679628	RS30A09	Root cap protein	Transport	C	*S-*
BX681365	RS57G02	Tonoplast intrinsic protein	Transport	C	*S-*
BX249892	Ge029H11	Plasma membrane intrinsic protein (Aquaporin)	Transport	C	*S-*
BX679636	RS30C11	Gamma tonoplast intrinsic protein	Transport	C	*S-*
BX667565	RN09F01	Amino acid permease (Amino acid carrier, putative)	Transport	C	*S+*
BX681797	RS65A06	Oxysterol-binding protein homolog	Transport	C	*S+*
CR394157	RN74F01	ADP,ATP carrier protein (ADP/ATP translocase)	Transport	c	*S+*
BX678354	RS08A06	Pinus taeda Trans-cinnamate 4-monooxygenase (EC 1.14.13.11)	Métabolisme	CI	
BF220737	NXCI_150_A7	Arginine decarboxylase	Métabolisme	CI	
AW985459	NXNV_136_H4	Aspartate transaminase	Métabolisme	Ci	
BX681976	RS67E11	Histidine decarboxylase (EC 4.1.1.22)	Métabolisme	Ci	
CR393420	RN59H07	Naringenin-chalcone synthase (EC 2.3.1.74)	Métabolisme	Ci	
BX679178	RS21A09	Zinc finger protein 216	Transcription	cI	
BX677658	RN42H03	Nonspecific lipid-transfer protein precursor (LTP)	Transport	Ci	
BX681997	RS67H05	Beta-glucosidase, chloroplast precursor (EC 3.2.1.21)	Métabolisme	P	*F+*
BX681865	RS65H11	Dihydropinosylvin synthase (EC 2.3.1.-)	Métabolisme	P	*M+*
CR392200	RN28B03	Defensin	Réponse au Stress	P	*F+*
BX251675	Ge053G10	Poly(A)-binding protein, partial (84%)	Synthèse de protéines	P	*F+*
BX251837	Ge056B06	Protein transport protein SEC22	Traffic intracellulaire	p	*M+*
BX252988	Ge076B04	Thioredoxin H-type (TRX-H)	Energie	PC	*S+F+*
BX249884	Ge029H02	Phosphoenolpyruvate carboxykinase (EC 4.1.4.49)	Energie	PC	*S+F+*
BX254884	Ge106G10	Phosphoenolpyruvate carboxykinase (EC 4.1.4.49)	Energie	pC	*S+F+*
CR394212	RN75C02	Pyruvate decarboxylase isozyme 2 (EC 4.1.1.1) (PDC)	Energie	pC	*S-F+*
AL751302	RS07C09	Metallothionein-like protein	Métabolisme	Pc	*S-M+*
BX667551	RN09D10	4-hydroxyphenylpyruvate dioxygenase (EC 1.13.11.27)	Métabolisme	PC	*S+F+*
BX249127	Ge020C11	Metallothionein-like protein EMB30	Métabolisme	pC	*S-M+*
BX680204	RS40E02	4-hydroxyphenylpyruvate dioxygenase (EC 1.13.11.27)	Métabolisme	pC	*S+F+*
AL750468	RN02H01	F21O3.6 Hypothetical protein	Non-classé	PC	*S+F+*
BX682853	Ge122G03	F3O9.22 protein.	Non-classé	pC	*S-F+*
AL750195	AS06G11	Hypothetical protein DR0254	Non-classé	pC	*S+F+*
CR394499	RN50C09	Antimicrobial peptide	Réponse au Stress	PC	*S+F+*
AL750636	RN05F08	Antimicrobial peptide	Réponse au Stress	PC	*S+F+*
BX679343	RS23E08	Chaperone protein dnaJ	Réponse au Stress	PC	*S+F+*
BX681928	RS66H05	Antimicrobial peptide	Réponse au Stress	PC	*S+F+*
CR394078	RN73D08	Licheninase precursor (EC 3.2.1.73)	Réponse au Stress	PC	*S+F+*
BX253756	Ge088A07	Aluminum-induced protein-like protein, partial (67%)	Réponse au Stress	pC	*S+F+*
BX249989	Ge031B01	Similarity to late embryogenesis abundant protein, partial (31%)	Réponse au Stress	pC	*S-M+*
BX255127	Ge111D09	eukaryotic initiation factor 3E subunit {Arabidopsis thaliana;} , partial (96%)	Transcription	pC	*S+F+*
CR394220	RN75C12	Peptide transporter PTR2-B	Transport	pC	*S+F+*
CR392105	RN12D09	Pinus taeda calcium binding protein	Non-classé	PCI	
BX254079	Ge093A05	Inositol oxygenase (Myo-inositol oxygenase) (Aldehyde reductase-like 6) , partial (51%)	Métabolisme	Pi	
BX680855	RS49G03	Alcohol dehydrogenase (EC 1.1.1.1)	Métabolisme	Pi	
BX681954	RS67C08	37 kDa inner envelope membrane protein	Non-classé	PI	
CR393324	RN58E05	Epoxide hydrolase (EC 3.3.2.3)	Réponse au Stress	pI	

Tableau Ax-8 : Liste des gènes mis en évidence dans les deux expériences. Les gènes sans fonction connue ne sont pas indiqués. Légende cf. Tab. Ax-6

Acc Number	ID	Annotaion	Catégorie fonctionelle	effet SC	effet SL	*sens SC*	*sens SL*
CR392786	RN48A05	Receptor protein kinase CLAVATA1 precursor (EC 2.7.1.-)	Communication/Transduction du signal	P	C	*F+*	*S+*
BX682453	Ge116H01	Receptor-like protein kinase, partial (70%)	Communication/Transduction du signal	C	C	*S+*	*S+*
CR392434	RN31H05	Seven transmembrane protein 1	Communication/Transduction du signal	C	C	*S+*	*S+*
BX681545	RS61A07	Cell elongation protein DIMINUTO	Croissance cellulaire, division	C	C	*S-*	*S-*
BX252622	Ge068A05	Elongation factor 1-gamma (EF-1-gamma) (eEF-1B gamma), partial (95%)	Croissance cellulaire, division	PC	C	*S+M+*	*S+*
CR392910	RN49H03	Germin-like protein	Croissance cellulaire, division	P	C	*M+*	*S+*
CR393563	RN61G09	MERI-5 protein precursor (Xyloglucan:xyloglucosyl transferase (EC 2.4.1.207)) / Brassin	Croissance cellulaire, division	C	C	*S-*	*S-*
BX255545	Ge007A07	Prohibitin	Croissance cellulaire, division	C	c	*S-*	*S-*
BX681114	RS53G02	26S proteasome non-ATPase regulatory subunit 2	Devenir des protéines	P	C	*M+*	*S+*
BX678807	RS15D05	Carboxypeptidase C precursor (EC 3.4.16.5)	Devenir des protéines	P	C	*M+*	*S-*
AL749972	AS03C11	Endopeptidase Clp subunit (EC 3.4.21.92)	Devenir des protéines	P	c	*M+*	*S-*
BX678708	RS14B03	Basic blue protein (Plantacyanin)	Energie	PC	CI	*S-M+*	
AL750498	RN03B11	NADPH dehydrogenase (EC 1.6.99.1) / 12-OXOPHYTODIENOATE reductase	Energie	C	PC	*S-*	*S-F+*
BX680890	RS50C06	NADPH:quinone reductase (EC 1.6.5.5)	Energie	C	C	*S+*	*S+*
AL749979	AS03D06	Thioredoxin	Energie	P	C	*M+*	*S+*
BX681748	RS63G11	2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase (EC 4.6.1.12)	Métabolisme	PC	C	*S-M+*	*S-*
BX679759	RS33B10	Adenylate kinase (EC 2.7.4.3)	Métabolisme	PC	P	*S-M+*	*F+*
BX681520	RS60F10	Aldehyde dehydrogenase (NAD+) (EC 1.2.1.3)	Métabolisme	P	C	*M+*	*S-*
BX254971	Ge108E04	Alkaline alpha galactosidase II, partial (84%)	Métabolisme	PC	C	*S+M+*	*S+*
AW870208	NXNV_125_E12	Aspartate transaminase	Métabolisme	P	C	*M+*	*S-*
CR392379	RN31A10	Beta-glucosidase precursor (EC 3.2.1.21)	Métabolisme	C	P	*S-*	*F+*
AL750008	AS03G06	Carbonate dehydratase (EC 4.2.1.1)	Métabolisme	P	c	*M+*	*S+*
CR393320	RN58E01	Cytochrome P450 (EC 1.14.-.-)	Métabolisme	C	C	*S-*	*S-*
BX255565	Ge007C03	Dephospho-CoA kinase (Dephosphocoenzyme A kinase) , partial (17%)	Métabolisme	P	C	*M+*	*S-*
BX680021	RS38C03	Dihydroflavonol-4-reductase (EC 1.1.1.219) (DFR)	Métabolisme	P	Ci	*M+*	
BX677769	RN44E06	Enoyl-CoA hydratase (EC 4.2.1.17)	Métabolisme	P	C	*F+*	*S+*
CR392825	RN48H03	Epimerase/dehydratase	Métabolisme	C	pC	*S-*	*S-M+*
BX252727	Ge069G09	Esterase precursor (Early nodule-specific protein homolog) (Latex allergen Hev b 13) , p	Métabolisme	PC	C	*S-M+*	*S-*
AL751292	RS07B10	Gamma-lyase	Métabolisme	C	C	*S+*	*S+*
BX254083	Ge093A12	Geranyl diphosphate synthase, complete	Métabolisme	PC	C	*S-M+*	*S-*
BX681235	RS55H11	Glutathione transferase (EC 2.5.1.18)	Métabolisme	P	C	*M+*	*S+*
CR394335	RN76H08	Glutathione transferase (EC 2.5.1.18)	Métabolisme	PC	C	*S-M+*	*S-*
BX682765	Ge121D06	Granule-bound glycogen [starch] synthase	Métabolisme	P	C	*M+*	*S-*
BX681985	RS67F12	Isoflavone reductase homolog (EC 1.3.1.-)	Métabolisme	PC	C	*S-M+*	*S-*
CR394206	RN75B08	L-ascorbate oxidase precursor (EC 1.10.3.3)	Métabolisme	PC	C	*S-M+*	*S-*
BG039825	NXSI_104_E11	Ornitine transaminase	Métabolisme	P	C	*M+*	*S-*
BX682869	Ge122H09	Pinene synthase (EC 4.2.3.14)	Métabolisme	PC	C	*S-M+*	*S-*
BX681938	RS67A12	Protein kinase (EC 2.7.1.-)	Métabolisme	PC	pI	*S-M+*	
CR393742	RN64E07	S-adenosyl-L-methionine:salicylic acid carboxyl methyltransferase	Métabolisme	P	C	*F+*	*S+*
BE643924	NXCI_50_B7	S-adenosyl-Met synthase	Métabolisme	P	C	*M+*	*S-*
AL751204	RS06C03	T29M8.12 protein	Métabolisme	P	pc	*M+*	*S-M+*
BX255434	Ge005G04	UDP-glucose 4-epimerase (EC 5.1.3.2)	Métabolisme	P	C	*M+*	*S-*
BX252228	Ge061F04	(2R)-phospho-3-sulfolactate synthase (EC 4.4.-.-)	Non-classé	PC	C	*S+M+*	*S+*
BX250435	Ge036H06	CTV.15, partial (41%)	Non-classé	PC	C	*S+F+*	*S+*
BX255662	Ge008C06	Glycine-rich protein homolog	Non-classé	P	C	*M+*	*S-*
AL750051	AS05C03	Hypothetical 49.0 kDa protein	Non-classé	PC	C	*S-M+*	*S-*
BX679990	RS37G05	Hypothetical protein (AT5g01750/T20L15_20)	Non-classé	P	C	*F+*	*S+*
CR393329	RN58F01	Major pollen allergen Bet v 1-D/H / Intracellular pathogenesis-related protein	Non-classé	C	PC	*S+*	*S+F+*
BX679871	RS35C05	Pinus taeda clone PtaAGP5	Non-classé	C	C	*S-*	*S-*
BX254145	Ge094A09	PR34, partial (4%)	Non-classé	P	C	*M+*	*S+*
BX679792	RS34A05	PrMC3	Non-classé	P	Pc	*M+*	*S-M+*
BX682906	Ge123D02	Probable ATP-dependent Clp protease proteolytic su	Non-classé	P	C	*M+*	*S-*
BX679458	RS26B10	Proline-rich protein	Non-classé	C	C	*S-*	*S-*
AL751295	RS07C01	Proline-rich protein / Cortical cell delineating protein precursor	Non-classé	PC	C	*S-M+*	*S-*
AL750511	RN03D01	PtaAGP5 (from Pinus taeda)	Non-classé	C	C	*S-*	*S-*
CR394036	RN72F06	Agglutinin alpha chain (MPA)	Organisation cellulaire	P	PC	*M+*	*S-M+*
CR392181	RN27G03	Annexin	Organisation cellulaire	P	C	*M+*	*S-*
BX251886	Ge056H01	Anther specific protein / Pollen specific & Anther specific LAT52 protein precursor	Organisation cellulaire	P	C	*M+*	*S-*
BX252966	Ge075G09	At5g20110, partial (46%)	Organisation cellulaire	P	C	*F+*	*S+*
AL751122	RS05C12	Glucan endo-1,3-beta-D-glucosidase (EC 3.2.1.39)	Organisation cellulaire	P	C	*F+*	*S+*
BX249875	Ge029G04	Glucan endo-1,3-beta-glucosidase precursor	Organisation cellulaire	P	C	*M+*	*S-*
BX252660	Ge068G01	Histone H2A	Organisation cellulaire	PC	C	*S-M+*	*S-*
AL750706	RN06D08	Nodulin	Organisation cellulaire	C	C	*S+*	*S+*
BX680845	RS49D06	Polyphosphoinositide binding protein	Organisation cellulaire	PC	c	*S-M+*	*S-*
BX249627	Ge026D05	Tubulin beta chain	Organisation cellulaire	P	C	*M+*	*S-*
BX255062	Ge110B02	Ubiquitin-conjugating enzyme E2 (EC 6.3.2.19)	Organisation cellulaire	C	p	*S+*	*F+*
BX252065	Ge059C02	17.9 kDa class II heat shock protein	Réponse au Stress	P	PCI	*M+*	
BX679279	RS22D07	ABA induced plasma membrane protein	Réponse au Stress	PC	C	*S+M+*	*S+*
CR393567	RN61H01	Chitinase precursor (EC 3.2.1.14)	Réponse au Stress	C	C	*S+*	*S+*
BX677767	RN44E03	Chitinase precursor (EC 3.2.1.14)	Réponse au Stress	PC	C	*S+M+*	*S+*
BX252366	Ge063E01	Class I Heat Shock Protein	Réponse au Stress	PC	C	*S+M+*	*S+*
BX681602	RS61G11	Defensin	Réponse au Stress	PC	C	*S-M+*	*S-*
BX681136	RS54B08	Dehydration stress-induced protein	Réponse au Stress	PC	PC	*S-M+*	*S-F+*
BX249564	Ge025E05	dehydrin	Réponse au Stress	C	C	*S+*	*S+*
BX680242	RS41A02	Disease resistance response protein 206	Réponse au Stress	C	C	*S-*	*S-*
BX251816	Ge055H04	Heat Shock 70 kd protein	Réponse au Stress	C	c	*S+*	*S+*
BX681237	RS55H04	Heat shock protein 101	Réponse au Stress	PC	C	*S+M+*	*S+*
BX678224	RN68A05	Late embryogenesis abundant protein	Réponse au Stress	C	pI	*S-*	
BX253089	Ge077G07	LEA, complete	Réponse au Stress	PC	C	*S+M+*	*S+*
BX250781	Ge041B11	Low molecular weight heat-shock protein, partial (98%)	Réponse au Stress	PCi	PC		*S+M+*
BX252607	Ge067G04	NIP3, partial (75%)	Réponse au Stress	P	C	*M+*	*S-*
CR392453	RN32D03	Peroxidase precursor (EC 1.11.1.7)	Réponse au Stress	C	C	*S-*	*S-*
BX675142	RN10A03	Peroxidase precursor (EC 1.11.1.7)	Réponse au Stress	PC	C	*S-M+*	*S-*
CR392767	RN37G07	Peroxidase precursor (EC 1.11.1.7)	Réponse au Stress	PC	C	*S-M+*	*S-*
BX678783	RS15A09	Peroxidase precursor (EC 1.11.1.7)	Réponse au Stress	PC	c	*S-M+*	*S-*
BX679939	RS36F06	Peroxidase precursor (EC 1.11.1.7)	Réponse au Stress	PC	PC	*S-M+*	*S-M+*
BX680909	RS50F11	Pinus sylvestris class III peroxidase	Réponse au Stress	P	PI	*M+*	
BX253632	Ge086B12	Snakin-1, partial (67%)	Réponse au Stress	C	C	*S-*	*S-*
BX255769	Ge009E10	T19F11.6 protein (NDR1/HIN1-like protein 1) (Harpin-induced protein)	Réponse au Stress	C	C	*S-*	*S-*
BX681290	RS56F02	40S ribosomal protein S11	Synthèse de protéines	PC	C	*S-M+*	*S-*
CR392325	RN30F07	40S ribosomal protein S26 (S31)	Synthèse de protéines	PC	P	*S-M+*	*F+*
BX248870	Ge002A10	60S ribosomal protein	Synthèse de protéines	P	C	*M+*	*S-*
BX679791	RS34A04	60S ribosomal protein	Synthèse de protéines	PC	C	*S-M+*	*S-*
BX682894	Ge123B11	60S ribosomal protein L38	Synthèse de protéines	P	C	*M+*	*S-*
BX251535	Ge050H05	60S ribosomal protein L39	Synthèse de protéines	P	C	*M+*	*S-*
BX679837	RS34F05	L3 cytoplasmic ribosomal protein	Synthèse de protéines	PC	C	*S-M+*	*S-*
BX252434	Ge064E04	Coatomer alpha subunit	Traffic intracellulaire	P	pCI	*M+*	
CR394236	RN75E09	Kinesin light chain (KLC)	Traffic intracellulaire	P	C	*F+*	*S+*
BX681158	RS54E01	Portal protein (GPB) (Minor capsid protein B)	Traffic intracellulaire	P	C	*F+*	*S+*
BX249494	Ge024F11	Profilin	Traffic intracellulaire	P	c	*M+*	*S-*
BX680424	RS43C10	Ethylene responsive element binding factor	Transcription	PC	C	*S-M+*	*S-*
BX682416	Ge116C09	G-box binding factor 4.	Transcription	C	C	*S+*	*S+*
BX254733	Ge103E05	Putative F-box protein (Fragment), partial (38%)	Transcription	C	PC	*S+*	*S+M+*
BX251374	Ge048G11	Rof1, partial (88%)	Transcription	PC	Ci	*S-M+*	
BX253595	Ge085G01	Zinc finger protein HIT-10, partial (3%)	Transcription	PC	PC	*S+M+*	*S+M+*
BX252185	Ge061A06	Mitochondrial oxaloacetate transport protein	Transport	P	PC	*M+*	*S-M+*
AL749547	AN01B08	Outer membrane lipoprotein blc precursor	Transport	PC	C	*S+M+*	*S+*
CR393241	RN54D04	Plasma membrane intrinsic polypeptide	Transport	PC	C	*S-M+*	*S-*
CR393662	RN63B06	Proton-exporting ATPase (EC 3.6.3.6)	Transport	P	C	*M+*	*S+*

Tableau Ax-9 : Affectation des accessions aux groupes générés par les logiciels SOTArrays (classification hiérarchique UPGMA en distance euclidienne) et Expander pour les gènes différentiellement exprimés des expériences stress court et stress long. (na = non détecté en SC ou SL / non affecté à un groupe (singletons du clustering))

Accession	Identifiant	Groupe Expander Stress Court	Groupe SOTArray Stress Court	Groupe Expander Stress Long	Groupe SOTArray Stress Long	Annotation	Catégorie fonctionnelle
BX677882	RN46A04	1_XpanSC	1_euSC	1_XpanSL	10_euSL	putative protein	_
BX677733	RN44A03	1_XpanSC	1_euSC	1_XpanSL	10_euSL	Unknown	Non-classé
AL751122	RS05C12	1_XpanSC	1_euSC	1_XpanSL	10_euSL	Glucan endo-1,3-beta-glucosidase precursor	Organisation cellulaire
BX255103	Ge110H06	1_XpanSC	1_euSC	1_XpanSL	11_euSL	putative protein	_
BX679198	RS21C03	1_XpanSC	1_euSC	1_XpanSL	11_euSL	putative protein	_
BX679419	RS25A08	1_XpanSC	1_euSC	1_XpanSL	11_euSL	putative protein	_
BX681588	RS61F06	1_XpanSC	1_euSC	1_XpanSL	11_euSL	putative protein	_
CR393361	RN59B05	1_XpanSC	1_euSC	1_XpanSL	11_euSL	putative protein	_
CR394501	RN60D10	1_XpanSC	1_euSC	1_XpanSL	11_euSL	putative protein	_
non soumis	RN28G10	1_XpanSC	1_euSC	1_XpanSL	11_euSL	putative protein	_
CR392786	RN48A05	1_XpanSC	1_euSC	1_XpanSL	11_euSL	Receptor protein kinase CLAVATA1 precursor (EC 2.7.1.-)	Communication/Transduction du signal
BX677769	RN44E06	1_XpanSC	1_euSC	1_XpanSL	11_euSL	Enoyl-CoA hydratase (EC 4.2.1.17)	Métabolisme
CR393742	RN64E07	1_XpanSC	1_euSC	1_XpanSL	11_euSL	S-adenosyl-L-methionine:salicylic acid carboxyl methyltransfera	Métabolisme
BX252966	Ge075G09	1_XpanSC	1_euSC	1_XpanSL	11_euSL	At5g20110, partial (46%)	Organisation cellulaire
BX681158	RS54E01	1_XpanSC	1_euSC	1_XpanSL	11_euSL	Portal protein (GPB) (Minor capsid protein B)	Traffic intracellulaire
CR394236	RN75E09	1_XpanSC	1_euSC	1_XpanSL	11_euSL	Kinesin light chain (KLC)	Traffic intracellulaire
BX679594	RS29B09	1_XpanSC	1_euSC	4_XpanSL	11_euSL	putative protein	_
BX681118	RS53G08	1_XpanSC	1_euSC	4_XpanSL	11_euSL	putative protein	_
CR393030	RN51F07	1_XpanSC	1_euSC	4_XpanSL	11_euSL	putative protein	_
BX679990	RS37G05	1_XpanSC	1_euSC	4_XpanSL	11_euSL	Hypothetical protein (AT5g01750/T20L15_20)	Non-classé
BX250268	Ge034E09	1_XpanSC	1_euSC	7_XpanSL	1_euSL	putative protein	_
CR393627	RN62F09	1_XpanSC	1_euSC	7_XpanSL	13_euSL	putative protein	_
AL751056	4CL	1_XpanSC	1_euSC	na	na	_	_
non soumis	Ha_PP756_b	1_XpanSC	1_euSC	na	na	_	_
BC065529	Ha_PP758_b	1_XpanSC	1_euSC	na	na	_	_
AL750816	RN07F07	1_XpanSC	1_euSC	na	na	putative protein	_
AL750839	RN07H06	1_XpanSC	1_euSC	na	na	putative protein	_
AL751348	RS07G08	1_XpanSC	1_euSC	na	na	putative protein	_
BX248830	Ge001E10	1_XpanSC	1_euSC	na	na	putative protein	_
BX249359	Ge023B03	1_XpanSC	1_euSC	na	na	putative protein	_
BX249490	Ge024F08	1_XpanSC	1_euSC	na	na	putative protein	_
BX249900	Ge030A10	1_XpanSC	1_euSC	na	na	putative protein	_
BX250876	Ge042D06	1_XpanSC	1_euSC	na	na	putative protein	_
BX251165	Ge046B06	1_XpanSC	1_euSC	na	na	putative protein	_
BX251766	Ge055B05	1_XpanSC	1_euSC	na	na	putative protein	_
BX252278	Ge062C05	1_XpanSC	1_euSC	na	na	putative protein	_
BX253462	Ge083G09	1_XpanSC	1_euSC	na	na	putative protein	_
BX254984	Ge108G07	1_XpanSC	1_euSC	na	na	putative protein	_
BX255216	Ge112H12	1_XpanSC	1_euSC	na	na	putative protein	_
BX677095	RN17D11	1_XpanSC	1_euSC	na	na	putative protein	_
BX677223	RN19E11	1_XpanSC	1_euSC	na	na	putative protein	_
BX677288	RN20A02	1_XpanSC	1_euSC	na	na	putative protein	_
BX677476	RN40A05	1_XpanSC	1_euSC	na	na	putative protein	_
BX677642	RN42E06	1_XpanSC	1_euSC	na	na	putative protein	_
BX678083	RN56A11	1_XpanSC	1_euSC	na	na	putative protein	_
BX678309	RN69D12	1_XpanSC	1_euSC	na	na	putative protein	_
BX678390	RS08E04	1_XpanSC	1_euSC	na	na	putative protein	_
BX679075	RS19D08	1_XpanSC	1_euSC	na	na	putative protein	_
BX679128	RS20C09	1_XpanSC	1_euSC	na	na	putative protein	_
BX679622	RS29H07	1_XpanSC	1_euSC	na	na	putative protein	_
BX680035	RS38D06	1_XpanSC	1_euSC	na	na	putative protein	_
BX680085	RS38H10	1_XpanSC	1_euSC	na	na	putative protein	_
BX680275	RS41D02	1_XpanSC	1_euSC	na	na	putative protein	_
BX680441	RS43D06	1_XpanSC	1_euSC	na	na	putative protein	_
BX680710	RS47E08	1_XpanSC	1_euSC	na	na	putative protein	_
BX680818	RS49A01	1_XpanSC	1_euSC	na	na	putative protein	_
BX681084	RS53B10	1_XpanSC	1_euSC	na	na	putative protein	_
BX681156	RS54E10	1_XpanSC	1_euSC	na	na	putative protein	_
BX681161	RS54E06	1_XpanSC	1_euSC	na	na	putative protein	_
BX681223	RS55F11	1_XpanSC	1_euSC	na	na	putative protein	_
BX681255	RS56B07	1_XpanSC	1_euSC	na	na	putative protein	_
BX681617	RS62A06	1_XpanSC	1_euSC	na	na	putative protein	_
BX681698	RS63B07	1_XpanSC	1_euSC	na	na	putative protein	_
BX681926	RS66H02	1_XpanSC	1_euSC	na	na	putative protein	_
BX681972	RS67E05	1_XpanSC	1_euSC	na	na	putative protein	_
BX682036	RS68G05	1_XpanSC	1_euSC	na	na	putative protein	_
CR392407	RN31D12	1_XpanSC	1_euSC	na	na	putative protein	_
CR392490	RN33B05	1_XpanSC	1_euSC	na	na	putative protein	_
CR392503	RN33D02	1_XpanSC	1_euSC	na	na	putative protein	_
CR392669	RN36C12	1_XpanSC	1_euSC	na	na	putative protein	_
CR392730	RN37C05	1_XpanSC	1_euSC	na	na	putative protein	_
CR393672	RN63C06	1_XpanSC	1_euSC	na	na	putative protein	_
CR393730	RN64C07	1_XpanSC	1_euSC	na	na	putative protein	_
CR394150	RN74E02	1_XpanSC	1_euSC	na	na	putative protein	_
CR394252	RN75G02	1_XpanSC	1_euSC	na	na	putative protein	_
CR394502	RN50F08	1_XpanSC	1_euSC	na	na	putative protein	_
non soumis	Ge024B12	1_XpanSC	1_euSC	na	na	putative protein	_
non soumis	RN32C07	1_XpanSC	1_euSC	na	na	putative protein	_
BX677537	RN40H04	1_XpanSC	1_euSC	na	na	Receptor protein kinase precursor	Communication/Transduction du signal
CR392472	RN32G05	1_XpanSC	1_euSC	na	na	Calmodulin-related protein 2	Communication/Transduction du signal
CR393564	RN61G10	1_XpanSC	1_euSC	na	na	Sorcin	Communication/Transduction du signal
BX681360	RS57F04	1_XpanSC	1_euSC	na	na	MERI-5 protein precursor (Xyloglucan:xyloglucosyl transferase	Croissance cellulaire, division
BX252329	Ge062H12	1_XpanSC	1_euSC	na	na	Proteasome subunit beta type 5 precursor (20S proteasome su	Devenir des protéines
AL750459	RN02G04	1_XpanSC	1_euSC	na	na	Ubiquinol-cytochrome C reductase	Energie
BX678629	RS12C06	1_XpanSC	1_euSC	na	na	Ubiquinol-cytochrome C reductase complex 7.8 kDa protein	Energie
BX681296	RS56F09	1_XpanSC	1_euSC	na	na	[Pyruvate dehydrogenase (Lipoamide)]-phosphatase (EC 3.1.3.	Energie
BX681437	RS59C04	1_XpanSC	1_euSC	na	na	Electron transfer flavoprotein-ubiquinone oxidoreductase	Energie
BX681554	RS61B07	1_XpanSC	1_euSC	na	na	NADPH:quinone reductase (EC 1.6.5.5)	Energie
BX681841	RS65F03	1_XpanSC	1_euSC	na	na	NADPH dehydrogenase (EC 1.6.99.1)	Energie
CR394186	RN74H06	1_XpanSC	1_euSC	na	na	Phytocyanin	Energie
AL750713	RN06E03	1_XpanSC	1_euSC	na	na	L-ascorbate oxidase precursor (EC 1.10.3.3)	Métabolisme
AL750794	RN07D07	1_XpanSC	1_euSC	na	na	Aldehyde dehydrogenase (NAD+) (EC 1.2.1.3)	Métabolisme
AL751154	RS05F10	1_XpanSC	1_euSC	na	na	Serine/threonine specific protein phosphatase 2C (EC 3.1.3.16)	Métabolisme
BX249794	Ge027G10	1_XpanSC	1_euSC	na	na	Beta-ketoacyl-ACP synthetase I-2, partial (91%)	Métabolisme
BX253454	Ge083F12	1_XpanSC	1_euSC	na	na	Patatin	Métabolisme
BX254883	Ge106G06	1_XpanSC	1_euSC	na	na	Histidine ammonia-lyase (Histidase) , partial (4%)	Métabolisme
BX677894	RN46B11	1_XpanSC	1_euSC	na	na	Alcohol dehydrogenase (EC 1.1.1.1)	Métabolisme
BX680092	RS39A02	1_XpanSC	1_euSC	na	na	T14P8.4 protein	Métabolisme
BX681087	RS53B06	1_XpanSC	1_euSC	na	na	SUR2 protein (Syringomycin response protein 2)	Métabolisme
BX681127	RS54B10	1_XpanSC	1_euSC	na	na	Glutathione transferase (EC 2.5.1.18)	Métabolisme
BX681204	RS55C05	1_XpanSC	1_euSC	na	na	Cysteine synthase (EC 2.5.1.47)	Métabolisme
BX681531	RS60G10	1_XpanSC	1_euSC	na	na	Chalcone-flavonone isomerase	Métabolisme
BX681612	RS61H11	1_XpanSC	1_euSC	na	na	dUTP pyrophosphatase (EC 3.6.1.23)	Métabolisme
BX681821	RS65D03	1_XpanSC	1_euSC	na	na	Aldehyde dehydrogenase (NAD+) (EC 1.2.1.3)	Métabolisme
BX682063	RS69E03	1_XpanSC	1_euSC	na	na	Phenylalanine ammonia-lyase (EC 4.3.1.5)	Métabolisme
BX682419	Ge116C12	1_XpanSC	1_euSC	na	na	Betaine-aldehyde dehydrogenase (EC 1.2.1.8)	Métabolisme
CR393484	RN60G02	1_XpanSC	1_euSC	na	na	Aldose 1-epimerase precursor (EC 5.1.3.3)	Métabolisme
CR393516	RN61B10	1_XpanSC	1_euSC	na	na	Glutamate-1-semialdehyde 2,1-aminomutase (EC 5.4.3.8)	Métabolisme
AL750301	AS07H12	1_XpanSC	1_euSC	na	na	Unknown	Non-classé
AL750551	RN03G06	1_XpanSC	1_euSC	na	na	Hypothetical 168.1 kDa protein	Non-classé
AL751076	RS03G09	1_XpanSC	1_euSC	na	na	Unknown	Non-classé
AL751149	RS05F05	1_XpanSC	1_euSC	na	na	Unknown	Non-classé
AL751163	RS05G07	1_XpanSC	1_euSC	na	na	Hypothetical 78.8 kDa protein	Non-classé
BX248924	Ge002G04	1_XpanSC	1_euSC	na	na	At3g23090, partial (15%)	Non-classé
BX249297	Ge022D04	1_XpanSC	1_euSC	na	na	Unknown	Non-classé
BX249896	Ge030A04	1_XpanSC	1_euSC	na	na	Unknown	Non-classé
BX249968	Ge030H03	1_XpanSC	1_euSC	na	na	Latex-abundant protein	Non-classé
BX249996	Ge031B08	1_XpanSC	1_euSC	na	na	At3g03773 {Arabidopsis thaliana;} , partial (71%)	Non-classé
BX251533	Ge050H03	1_XpanSC	1_euSC	na	na	At1g11170/T28P6_16, partial (58%)	Non-classé
BX251778	Ge055C09	1_XpanSC	1_euSC	na	na	Unknown	Non-classé
BX252796	Ge070H02	1_XpanSC	1_euSC	na	na	SOUL-like protein, partial (10%)	Non-classé
BX253820	Ge089A07	1_XpanSC	1_euSC	na	na	Hypothetical 10.4 kDa protein (Expressed protein)	Non-classé
BX254006	Ge092B02	1_XpanSC	1_euSC	na	na	Outer membrane protein (Fragment), partial (20%)	Non-classé
BX254656	Ge102B08	1_XpanSC	1_euSC	na	na	CG33184-PA, isoform A, partial (9%)	Non-classé
BX667580	RN09G05	1_XpanSC	1_euSC	na	na	Unknown	Non-classé
BX676834	RN12D04	1_XpanSC	1_euSC	na	na	T17B22.3 protein	Non-classé
BX677557	RN41B08	1_XpanSC	1_euSC	na	na	GASA5-like protein (Sb35)	Non-classé
BX678298	RN69C12	1_XpanSC	1_euSC	na	na	Unknown	Non-classé
BX678411	RS08G05	1_XpanSC	1_euSC	na	na	RNA-binding protein KIAA0682	Non-classé
BX678445	RS09C02	1_XpanSC	1_euSC	na	na	Unknown	Non-classé
BX679085	RS19F02	1_XpanSC	1_euSC	na	na	Hypothetical 20.6 kDa protein	Non-classé
BX680179	RS40B08	1_XpanSC	1_euSC	na	na	At2g44310 Hypothetical protein	Non-classé
BX680229	RS40G08	1_XpanSC	1_euSC	na	na	Hypothetical protein (At1g76160/T23E18_10)	Non-classé
BX680327	RS42A07	1_XpanSC	1_euSC	na	na	Hypothetical 85.6 kDa protein	Non-classé

BX681024	RS52B04	1_XpanSC	1_euSC	na	na	Hypothetical 60.3 kDa protein	Non-classé
BX681544	RS61A06	1_XpanSC	1_euSC	na	na	Lateral organ boundaries (LOB) domain protein	Non-classé
BX682007	RS68A10	1_XpanSC	1_euSC	na	na	Unknown	Non-classé
CR392486	RN33A12	1_XpanSC	1_euSC	na	na	ATP-dependent Clp protease proteolytic subunit (EC 3.4.21.92)	Non-classé
CR393402	RN59F08	1_XpanSC	1_euSC	na	na	Unknown	Non-classé
CR393481	RN60F10	1_XpanSC	1_euSC	na	na	IN2-2 protein	Non-classé
CR393769	RN65A09	1_XpanSC	1_euSC	na	na	Unknown	Non-classé
CR393770	RN65A10	1_XpanSC	1_euSC	na	na	Unknown	Non-classé
CR393785	RN65C08	1_XpanSC	1_euSC	na	na	Hypothetical 8.7 kDa protein	Non-classé
CR393841	RN65G12	1_XpanSC	1_euSC	na	na	Unknown	Non-classé
CR393994	RN71F04	1_XpanSC	1_euSC	na	na	Unknown	Non-classé
CR394188	RN74H08	1_XpanSC	1_euSC	na	na	Protein kinase-like (EC 2.7.1.37)	Non-classé
BX250283	Ge034G10	1_XpanSC	1_euSC	na	na	Pectate lyase, partial (95%)	Organisation cellulaire
BX680100	RS39B01	1_XpanSC	1_euSC	na	na	Nodulin-like protein (Putative nodulin protein)	Organisation cellulaire
BX680185	RS40C03	1_XpanSC	1_euSC	na	na	Pectate lyase (EC 4.2.2.2)	Organisation cellulaire
BX681382	RS58B04	1_XpanSC	1_euSC	na	na	Ubiquitin-protein ligase 1	Organisation cellulaire
AL750400	RN02B01	1_XpanSC	1_euSC	na	na	Copper chaperone for Superoxide dismutase	Réponse au Stress
AL751229	RS06E04	1_XpanSC	1_euSC	na	na	Peroxidase precursor (EC 1.11.1.7)	Réponse au Stress
BX249660	Ge026G07	1_XpanSC	1_euSC	na	na	ENDOCHITINASE (EC 3.2.1.14)	Réponse au Stress
BX253877	Ge089H04	1_XpanSC	1_euSC	na	na	DNA repair protein RAD16, partial (6%)	Réponse au Stress
BX253926	Ge090G07	1_XpanSC	1_euSC	na	na	Disease resistance-like protein (Fragment), partial (16%)	Réponse au Stress
BX254489	Ge099C04	1_XpanSC	1_euSC	na	na	ATP-dependent DNA helicase (EC 3.6.1.-), partial (2%)	Réponse au Stress
BX254690	Ge102G05	1_XpanSC	1_euSC	na	na	peroxidase	Réponse au Stress
BX255778	Ge009F07	1_XpanSC	1_euSC	na	na	Disease resistance protein-like	Réponse au Stress
BX666039	RN08H11	1_XpanSC	1_euSC	na	na	Late embryogenesis abundant protein	Réponse au Stress
BX681362	RS57F06	1_XpanSC	1_euSC	na	na	DnaJ protein homolog	Réponse au Stress
BX681431	RS59B06	1_XpanSC	1_euSC	na	na	Pathogenesis-related protein PR-4B precursor	Réponse au Stress
BX681471	RS60A05	1_XpanSC	1_euSC	na	na	Late embryogenesis abundant protein (Fragment)	Réponse au Stress
CR392961	RN50F01	1_XpanSC	1_euSC	na	na	Dehydration-responsive protein RD22 precursor	Réponse au Stress
AL750386	RN01H09	1_XpanSC	1_euSC	na	na	30S ribosomal protein S16	Synthèse de protéines
BX678464	RS09E08	1_XpanSC	1_euSC	na	na	60S ribosomal protein L41, mitochondrial precursor	Synthèse de protéines
BX679475	RS26E10	1_XpanSC	1_euSC	na	na	Ribosomal protein L10	Synthèse de protéines
CR392159	RN27C08	1_XpanSC	1_euSC	na	na	40S ribosomal protein	Synthèse de protéines
CR392543	RN33H08	1_XpanSC	1_euSC	na	na	25S rRNA	Synthèse de protéines
BX681245	RS56A07	1_XpanSC	1_euSC	na	na	Dynamin-like protein	Traffic intracellulaire
CR394037	RN72F07	1_XpanSC	1_euSC	na	na	Coatomer beta' subunit (Beta'-coat protein)	Traffic intracellulaire
CR394145	RN74D07	1_XpanSC	1_euSC	na	na	GTP-binding protein SAR	Traffic intracellulaire
BX250119	Ge032F08	1_XpanSC	1_euSC	na	na	Auxin response factor 1, partial (25%)	Transcription
CR392874	RN49D09	1_XpanSC	1_euSC	na	na	Zinc finger protein 364	Transcription
CR393619	RN62F01	1_XpanSC	1_euSC	na	na	Peptide transporter PTR2-B	Transport
CR393875	RN66E08	1_XpanSC	1_euSC	na	na	Calcium-transporting ATPase, plasma membrane-type	Transport
BX250435	Ge036H06	1_XpanSC	2_euSC	4_XpanSL	11_euSL	CTV.15, partial (41%)	Non-classé
BX679331	RS23D01	1_XpanSC	2_euSC	na	na	putative protein	_
BX680053	RS38F01	1_XpanSC	2_euSC	na	na	putative protein	_
CR393571	RN61H07	1_XpanSC	2_euSC	na	na	Aldose 1-epimerase precursor (EC 5.1.3.3)	Métabolisme
AL750105	AS05G10	1_XpanSC	2_euSC	na	na	Unknown	Non-classé
CR392715	RN37A08	1_XpanSC	2_euSC	na	na	Chitinase precursor (EC 3.2.1.14)	Réponse au Stress
BX678807	RS15D05	2_XpanSC	6_euSC	2_XpanSL	17_euSL	Carboxypeptidase C precursor (EC 3.4.16.5)	Devenir des protéines
BX682765	Ge121D06	2_XpanSC	6_euSC	2_XpanSL	17_euSL	Granule-bound glycogen [starch] synthase	Métabolisme
BX679791	RS34A04	2_XpanSC	6_euSC	2_XpanSL	17_euSL	60S ribosomal protein	Synthèse de protéines
BX252735	Ge069H08	2_XpanSC	6_euSC	2_XpanSL	18_euSL	putative protein	_
BX255545	Ge007A07	2_XpanSC	6_euSC	2_XpanSL	18_euSL	Prohibitin	Croissance cellulaire, division
BX255565	Ge007C03	2_XpanSC	6_euSC	2_XpanSL	18_euSL	Dephospho-CoA kinase (Dephosphocoenzyme A kinase) , part	Métabolisme
CR394206	RN75B08	2_XpanSC	6_euSC	2_XpanSL	18_euSL	L-ascorbate oxidase precursor (EC 1.10.3.3)	Métabolisme
BX249627	Ge026D05	2_XpanSC	6_euSC	2_XpanSL	18_euSL	Tubulin beta chain	Organisation cellulaire
CR392181	RN27G03	2_XpanSC	6_euSC	2_XpanSL	18_euSL	Annexin	Organisation cellulaire
BX252607	Ge067G04	2_XpanSC	6_euSC	2_XpanSL	18_euSL	NIP3, partial (75%)	Réponse au Stress
BX248870	Ge002A10	2_XpanSC	6_euSC	2_XpanSL	18_euSL	60S acidic ribosomal protein P1	Synthèse de protéines
BX682894	Ge123B11	2_XpanSC	6_euSC	2_XpanSL	18_euSL	60S ribosomal protein L38	Synthèse de protéines
BX249494	Ge024F11	2_XpanSC	6_euSC	2_XpanSL	18_euSL	Profilin	Traffic intracellulaire
BX252809	Ge071A10	2_XpanSC	6_euSC	2_XpanSL	20_euSL	putative protein	_
BX679222	RS21G07	2_XpanSC	6_euSC	2_XpanSL	20_euSL	putative protein	_
BX251886	Ge056H01	2_XpanSC	6_euSC	2_XpanSL	20_euSL	Anther specific protein / Pollen specific & Anther specific LAT52	Organisation cellulaire
BX681290	RS56F02	2_XpanSC	6_euSC	2_XpanSL	20_euSL	40S ribosomal protein S11	Synthèse de protéines
BX680021	RS38C03	2_XpanSC	6_euSC	2_XpanSL	21_euSL	Dihydroflavonol-4-reductase (EC 1.1.1.219) (DFR)	Métabolisme
BX682906	Ge123D02	2_XpanSC	6_euSC	2_XpanSL	21_euSL	ATP-dependent Clp protease proteolytic su	Non-classé
BX249875	Ge029G04	2_XpanSC	6_euSC	2_XpanSL	21_euSL	Glucan endo-1,3-beta-D-glucosidase (EC 3.2.1.39)	Organisation cellulaire
BX252727	Ge069G09	2_XpanSC	6_euSC	2_XpanSL	22_euSL	Esterase precursor (Early nodule-specific protein homolog) (La	Métabolisme
BX255662	Ge008C06	2_XpanSC	6_euSC	2_XpanSL	22_euSL	Glycine-rich protein homolog	Non-classé
CR392910	RN49H03	2_XpanSC	6_euSC	4_XpanSL	10_euSL	Germin-like protein	Croissance cellulaire, division
BX681871	RS66A11	2_XpanSC	6_euSC	5_XpanSL	18_euSL	putative protein	_
AL751204	RS06C03	2_XpanSC	6_euSC	5_XpanSL	18_euSL	T29M8.12 protein	Métabolisme
AL750051	AS05C03	2_XpanSC	6_euSC	5_XpanSL	18_euSL	Hypothetical 49.0 kDa protein	Non-classé
BX680425	RS43C11	2_XpanSC	6_euSC	5_XpanSL	19_euSL	putative protein	_
non soumis	RN32A07	2_XpanSC	6_euSC	5_XpanSL	20_euSL	putative protein	_
BX251535	Ge050H05	2_XpanSC	6_euSC	6_XpanSL	17_euSL	60S ribosomal protein L39	Synthèse de protéines
BX679837	RS34F05	2_XpanSC	6_euSC	6_XpanSL	17_euSL	L3 cytoplasmic ribosomal protein	Synthèse de protéines
CR392220	RN28D08	2_XpanSC	6_euSC	7_XpanSL	13_euSL	putative protein	_
CR392325	RN30F07	2_XpanSC	6_euSC	7_XpanSL	13_euSL	40S ribosomal protein S26 (S31)	Synthèse de protéines
BX680940	RS51B12	2_XpanSC	6_euSC	7_XpanSL	15_euSL	putative protein	_
BX254145	Ge094A09	2_XpanSC	6_euSC	na	10_euSL	PR34, partial (4%)	Non-classé
BX679434	RS25D07	2_XpanSC	6_euSC	na	13_euSL	putative protein	_
BX679792	RS34A05	2_XpanSC	6_euSC	na	20_euSL	PrMC3	Non-classé
BX249189	Ge021A08	2_XpanSC	6_euSC	na	na	putative protein	_
BX252565	Ge066E10	2_XpanSC	6_euSC	na	na	putative protein	_
BX681222	RS55E07	2_XpanSC	6_euSC	na	na	putative protein	_
BX682978	Ge126D02	2_XpanSC	6_euSC	na	na	putative protein	_
CR392504	RN33D03	2_XpanSC	6_euSC	na	na	putative protein	_
BX249546	Ge025C11	2_XpanSC	6_euSC	na	na	Calmodulin	Communication/Transduction du signal
BX249833	Ge029C02	2_XpanSC	6_euSC	na	na	Ras-like small monomeric GTP-binding protein, partial (65%)	Communication/Transduction du signal
BX682915	Ge123D11	2_XpanSC	6_euSC	na	na	14-3-3 protein	Communication/Transduction du signal
BX677507	RN40E01	2_XpanSC	6_euSC	na	na	Nuclear scaffold-like protein p76	Croissance cellulaire, division
BX681068	RS52G06	2_XpanSC	6_euSC	na	na	Cysteine proteinase RD21A precursor (EC 3.4.22.-)	Devenir des protéines
BX249881	Ge029G11	2_XpanSC	6_euSC	na	na	ATP synthase gamma chain	Energie
BX252563	Ge066E08	2_XpanSC	6_euSC	na	na	Aspartate aminotransfesrase, chloroplast precursor	Métabolisme
BX678890	RS16F01	2_XpanSC	6_euSC	na	na	Cholinephosphate cytidylyltransferase B (EC 2.7.7.15)	Métabolisme
BX250861	Ge042C04	2_XpanSC	6_euSC	na	na	OVCA2=CANDIDATE tumor suppressor, partial (9%)	Non-classé
BX252717	Ge069F09	2_XpanSC	6_euSC	na	na	B2 protein	Non-classé
BX255226	Ge113B09	2_XpanSC	6_euSC	na	na	TrPRP2, partial (67%)	Non-classé
BX681067	RS52G04	2_XpanSC	6_euSC	na	na	Embryogenesis-associated protein EMB7	Non-classé
BX252614	Ge067H06	2_XpanSC	6_euSC	na	na	Histone H2A variant	Organisation cellulaire
BX679616	RS29G03	2_XpanSC	6_euSC	na	na	Histone H4	Organisation cellulaire
BX251094	Ge045C02	2_XpanSC	6_euSC	na	na	Abscisic stress ripening protein	Réponse au Stress
BX254258	Ge095G04	2_XpanSC	6_euSC	na	na	LUMINAL BINDING PROTEIN	Réponse au Stress
BX254875	Ge106E11	2_XpanSC	6_euSC	na	na	Probable aquaporin, complete	Transport
BX682324	Ge115A12	2_XpanSC	6_euSC	na	na	Abscisic stress ripening protein	Réponse au Stress
AL750567	RN03H10	2_XpanSC	6_euSC	na	na	60S ribosomal protein L14	Synthèse de protéines
BX252371	Ge063E06	2_XpanSC	6_euSC	na	na	60S ribosomal protein L31	Synthèse de protéines
BX252657	Ge068F09	2_XpanSC	6_euSC	na	na	60S ribosomal protein L35A	Synthèse de protéines
BX682787	Ge121G09	2_XpanSC	6_euSC	na	na	60S ribosomal protein L31	Synthèse de protéines
BX682634	Ge119A01	2_XpanSC	6_euSC	na	na	Elongation factor 1-gamma (EF-1-gamma)	Synthèse de protéines
BX254888	Ge106H02	2_XpanSC	6_euSC	na	na	Ras-related protein Rab11B	Traffic intracellulaire
BX682440	Ge116F08	2_XpanSC	6_euSC	na	na	Nuclear transport factor 2 (NTF-2).	Traffic intracellulaire
BX681545	RS61A07	2_XpanSC	7_euSC	2_XpanSL	18_euSL	Cell elongation protein DIMINUTO	Croissance cellulaire, division
BX680845	RS49D06	2_XpanSC	7_euSC	2_XpanSL	18_euSL	Polyphosphoinositide binding protein	Organisation cellulaire
BX680424	RS43C10	2_XpanSC	7_euSC	2_XpanSL	18_euSL	Ethylene responsive element binding factor	Transcription
CR394335	RN76H08	2_XpanSC	7_euSC	2_XpanSL	20_euSL	Glutathione transferase (EC 2.5.1.18)	Métabolisme
BX680242	RS41A02	2_XpanSC	7_euSC	2_XpanSL	20_euSL	Disease resistance response protein 206	Réponse au Stress
CR392399	RN31D03	2_XpanSC	7_euSC	2_XpanSL	21_euSL	putative protein	_
BX254083	Ge093A12	2_XpanSC	7_euSC	2_XpanSL	21_euSL	Geranyl diphosphate synthase, complete	Métabolisme
BX681748	RS63G11	2_XpanSC	7_euSC	2_XpanSL	21_euSL	2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase (EC 4.6.	Métabolisme
BX681985	RS67F12	2_XpanSC	7_euSC	2_XpanSL	21_euSL	Isoflavone reductase homolog (EC 1.3.1.-)	Métabolisme
BX682869	Ge122H09	2_XpanSC	7_euSC	2_XpanSL	21_euSL	Pinene synthase (EC 4.2.3.14)	Métabolisme
BX255769	Ge009E10	2_XpanSC	7_euSC	2_XpanSL	21_euSL	T19F11.6 protein (NDR1/HIN1-like protein 1) (Harpin-induced p	Réponse au Stress
BX681602	RS61G11	2_XpanSC	7_euSC	2_XpanSL	21_euSL	Defensin	Réponse au Stress
CR392510	RN33D10	2_XpanSC	7_euSC	2_XpanSL	22_euSL	putative protein	_
CR393359	RN59B03	2_XpanSC	7_euSC	2_XpanSL	22_euSL	putative protein	_
BX675142	RN10A03	2_XpanSC	7_euSC	2_XpanSL	22_euSL	Peroxidase precursor (EC 1.11.1.7)	Réponse au Stress
CR392767	RN37G07	2_XpanSC	7_euSC	2_XpanSL	22_euSL	Peroxidase precursor (EC 1.11.1.7)	Réponse au Stress
BX251374	Ge048G11	2_XpanSC	7_euSC	2_XpanSL	22_euSL	Rof1, partial (88%)	Transcription
BX252660	Ge068G01	2_XpanSC	7_euSC	5_XpanSL	18_euSL	Histone H2A	Organisation cellulaire
CR393398	RN59F04	2_XpanSC	7_euSC	5_XpanSL	19_euSL	putative protein	_
non soumis	RN28G12	2_XpanSC	7_euSC	5_XpanSL	19_euSL	putative protein	_
CR393241	RN54D04	2_XpanSC	7_euSC	5_XpanSL	19_euSL	Plasma membrane intrinsic polypeptide	Transport
AL751295	RS07C01	2_XpanSC	7_euSC	5_XpanSL	20_euSL	Proline-rich protein / Cortical cell delineating protein precursor	Non-classé
BX678783	RS15A09	2_XpanSC	7_euSC	5_XpanSL	20_euSL	Peroxidase precursor (EC 1.11.1.7)	Réponse au Stress
CR392586	RN35A11	2_XpanSC	7_euSC	6_XpanSL	15_euSL	putative protein	_
BX681136	RS54B08	2_XpanSC	7_euSC	6_XpanSL	16_euSL	Dehydration stress-induced protein	Réponse au Stress
BX681938	RS67A12	2_XpanSC	7_euSC	na	14_euSL	Protein kinase (EC 2.7.1.-)	Métabolisme
BX678224	RN68A05	2_XpanSC	7_euSC	na	14_euSL	Late embryogenesis abundant protein	Réponse au Stress
BX680962	RS51C03	2_XpanSC	7_euSC	na	na	putative protein	_
BX681270	RS56D01	2_XpanSC	7_euSC	na	na	putative protein	_
CR392581	RN35A05	2_XpanSC	7_euSC	na	na	Isopentenyl-diphosphate delta-isomerase II (EC 5.3.3.2)	Métabolisme
CR392259	RN30A11	2_XpanSC	7_euSC	na	na	Acid phosphatase precursor 1 (EC 3.1.3.2)	Métabolisme
CR392614	RN35E03	2_XpanSC	7_euSC	na	na	Acid phosphatase	Métabolisme
CR393796	RN65E04	2_XpanSC	7_euSC	na	na	Pinene synthase (EC 4.2.3.14)	Métabolisme
BX678993	RS18B02	2_XpanSC	7_euSC	na	na	IN2-2 protein	Non-classé
AL750525	RN03E03	2_XpanSC	7_euSC	na	na	MtN21 nodulin protein-like	Organisation cellulaire
AL750621	RN05E06	2_XpanSC	7_euSC	na	na	Pectate lyase (EC 4.2.2.2)	Organisation cellulaire
BX677964	RN47C10	2_XpanSC	7_euSC	na	na	Late embryogenesis abundant protein	Réponse au Stress
BX678491	RS10A02	2_XpanSC	7_euSC	na	na	Pathogenesis-related protein 5 precursor (PR-5)	Réponse au Stress
BX678596	RS11G02	2_XpanSC	7_euSC	na	na	Peroxidase precursor (EC 1.11.1.7)	Réponse au Stress
CR393156	RN53C07	2_XpanSC	7_euSC	na	na	Peroxidase precursor (EC 1.11.1.7)	Réponse au Stress

BX678294	RN69C07	2_XpanSC	7_euSC	na	na	Nuclear transport factor 2 (NTF-2)	Traffic intracellulaire
BX252622	Ge068A05	3_XpanSC	3_euSC	1_XpanSL	10_euSL	Elongation factor 1-gamma (EF-1-gamma) (eEF-1B gamma), p	Synthèse de protéines
CR393329	RN58F01	3_XpanSC	3_euSC	1_XpanSL	3_euSL	Major pollen allergen Bet v 1-D/H / Intracellular pathogenesis-re	Non-classé
BX254971	Ge108E04	3_XpanSC	3_euSC	1_XpanSL	5_euSL	Alkaline alpha galactosidase II, partial (84%)	Métabolisme
BX253089	Ge077G07	3_XpanSC	3_euSC	1_XpanSL	5_euSL	LEA, complete	Réponse au Stress
BX677767	RN44E03	3_XpanSC	3_euSC	1_XpanSL	5_euSL	Chitinase precursor (EC 3.2.1.14)	Réponse au Stress
BX679279	RS22D07	3_XpanSC	3_euSC	1_XpanSL	5_euSL	ABA induced plasma membrane protein	Réponse au Stress
CR393567	RN61H01	3_XpanSC	3_euSC	1_XpanSL	5_euSL	Chitinase precursor (EC 3.2.1.14)	Réponse au Stress
AL749547	AN01B08	3_XpanSC	3_euSC	1_XpanSL	5_euSL	Outer membrane lipoprotein blc precursor	Transport
BX252366	Ge063E01	3_XpanSC	3_euSC	1_XpanSL	8_euSL	Class I Heat Shock Protein	Réponse au Stress
CR394334	RN76H07	3_XpanSC	3_euSC	4_XpanSL	10_euSL	putative protein	_
BX250781	Ge041B11	3_XpanSC	3_euSC	4_XpanSL	6_euSL	Low molecular weight heat-shock protein, partial (98%)	Réponse au Stress
BX000656	HSP	3_XpanSC	3_euSC	4_XpanSL	6_euSL	Heat shock protein	Réponse au Stress
BX677134	RN18B10	3_XpanSC	3_euSC	na	na	putative protein	_
BX680428	RS43C02	3_XpanSC	3_euSC	na	na	putative protein	_
AL750116	AS05H10	3_XpanSC	3_euSC	na	na	EF-HAND containing protein-like	Non-classé
BX249919	Ge030C08	3_XpanSC	4_euSC	1_XpanSL	10_euSL	Unknown	Non-classé
BX677148	RN18D05	3_XpanSC	4_euSC	1_XpanSL	10_euSL	Unknown	Non-classé
CR393111	RN52G02	3_XpanSC	4_euSC	1_XpanSL	10_euSL	Unknown	Non-classé
BX681237	RS55H04	3_XpanSC	4_euSC	1_XpanSL	10_euSL	Heat shock protein 101	Réponse au Stress
BX680890	RS50C06	3_XpanSC	4_euSC	1_XpanSL	4_euSL	NADPH:quinone reductase (EC 1.6.5.5)	Energie
AL751292	RS07B10	3_XpanSC	4_euSC	1_XpanSL	4_euSL	Gamma-lyase	Métabolisme
BX252225	Ge061E10	3_XpanSC	4_euSC	1_XpanSL	5_euSL	putative protein	_
BX252228	Ge061F04	3_XpanSC	4_euSC	1_XpanSL	5_euSL	(2R)-phospho-3-sulfolactate synthase (EC 4.4.-.-)	Non-classé
BX680324	RS42A12	3_XpanSC	4_euSC	1_XpanSL	5_euSL	Unknown	Non-classé
AL750706	RN06D08	3_XpanSC	4_euSC	1_XpanSL	5_euSL	Nodulin	Organisation cellulaire
BX682416	Ge116C09	3_XpanSC	4_euSC	1_XpanSL	5_euSL	G-box binding factor 4.	Transcription
BX679298	RS22F06	3_XpanSC	4_euSC	1_XpanSL	7_euSL	putative protein	_
BX681234	RS55G07	3_XpanSC	4_euSC	2_XpanSL	18_euSL	putative protein	_
BX255062	Ge110B02	3_XpanSC	4_euSC	3_XpanSL	13_euSL	Ubiquitin-conjugating enzyme E2 (EC 6.3.2.19)	Organisation cellulaire
BX678079	RN56A06	3_XpanSC	4_euSC	4_XpanSL	10_euSL	putative protein	_
BX253595	Ge085G01	3_XpanSC	4_euSC	4_XpanSL	10_euSL	Zinc finger protein HIT-10, partial (3%)	Transcription
BX255395	Ge005C11	3_XpanSC	4_euSC	na	11_euSL	putative protein	_
AL750163	AS06E02	3_XpanSC	4_euSC	na	na	putative protein	_
AL751045	RS03D12	3_XpanSC	4_euSC	na	na	putative protein	_
BX252908	Ge074G03	3_XpanSC	4_euSC	na	na	Cathepsin B-like cysteine proteinase, partial (68%)	Devenir des protéines
BX681080	RS53A02	3_XpanSC	4_euSC	na	na	Acetylornithine deacetylase (EC 3.5.1.16)	Devenir des protéines
BX252844	Ge071F04	3_XpanSC	4_euSC	na	na	metallothionein-like protein	Métabolisme
BX253047	Ge077B10	3_XpanSC	4_euSC	na	na	Aminotransferase	Métabolisme
BX254251	Ge095F04	3_XpanSC	4_euSC	na	na	cytosolic aldehyde dehydrogenase {Oryza sativa (japonica culti	Métabolisme
BX255766	Ge009E07	3_XpanSC	4_euSC	na	na	metallothionein	Métabolisme
BX667535	RN09C02	3_XpanSC	4_euSC	na	na	Carbon catabolite derepressing protein kinase	Métabolisme
BX253244	Ge080C01	3_XpanSC	4_euSC	na	na	Lin1294 protein, partial (9%)	Non-classé
BX253795	Ge088F01	3_XpanSC	4_euSC	na	na	Unknown	Non-classé
CR394213	RN75C03	3_XpanSC	4_euSC	na	na	Auxin-repressed 12.5 kDa protein	Non-classé
BX681731	RS63F03	3_XpanSC	4_euSC	na	na	Anther specific LAT52 protein precursor	Organisation cellulaire
AL750218	AS07A11	3_XpanSC	4_euSC	na	na	Chitinase (EC 3.2.1.14)	Réponse au Stress
BX250140	Ge032H07	3_XpanSC	4_euSC	na	na	Defender against cell death	Réponse au Stress
BX249687	Ge027B05	3_XpanSC	4_euSC	na	na	FKBP-like protein (Arabidopsis thaliana genomic DNA, chromos	Transcription
BX252953	Ge075E07	3_XpanSC	4_euSC	na	na	RING finger protein 26	Transcription
AL749979	AS03D06	3_XpanSC	5_euSC	1_XpanSL	10_euSL	Thioredoxin	Energie
BX681114	RS53G02	3_XpanSC	5_euSC	1_XpanSL	12_euSL	26S proteasome non-ATPase regulatory subunit 2	Devenir des protéines
BX681520	RS60F10	3_XpanSC	5_euSC	2_XpanSL	18_euSL	Aldehyde dehydrogenase (NAD+) (EC 1.2.1.3)	Métabolisme
CR393500	RN61A02	3_XpanSC	5_euSC	na	10_euSL	putative protein	_
AL749580	AN01H05	3_XpanSC	5_euSC	na	na	putative protein	_
BX254586	Ge100G06	3_XpanSC	5_euSC	na	na	putative protein	_
BX254794	Ge104G10	3_XpanSC	5_euSC	na	na	putative protein	_
BX678995	RS18B05	3_XpanSC	5_euSC	na	na	putative protein	_
BX680660	RS46H02	3_XpanSC	5_euSC	na	na	putative protein	_
CR394276	RN76B02	3_XpanSC	5_euSC	na	na	putative protein	_
BX254449	Ge098G01	3_XpanSC	5_euSC	na	na	Phospholipase D alpha 1 (PLD alpha 1) (Choline phosphatase	Communication/Transduction du signal
BE123499	NXNV_148_H6	3_XpanSC	5_euSC	na	na	Cellulose synthase	Métabolisme
BX252163	Ge060F11	3_XpanSC	5_euSC	na	na	Unknown	Non-classé
AL750092	AS05F09	3_XpanSC	5_euSC	na	na	Ubiquitin-conjugating enzyme E2-17 kDa (EC 6.3.2.19)	Organisation cellulaire
AL750508	RN03C09	3_XpanSC	5_euSC	na	na	Histone H4	Organisation cellulaire
BX250267	Ge034E07	3_XpanSC	5_euSC	na	na	Ubiquitin-conjugating enzyme E2-17 kDa (EC 6.3.2.19)	Organisation cellulaire
BX682459	Ge116H08	3_XpanSC	5_euSC	na	na	Ubiquitin-conjugating enzyme (EC 6.3.2.19)	Organisation cellulaire
BX254293	Ge096C10	3_XpanSC	5_euSC	na	na	Peptide methionine sulfoxide reductase (EC 1.8.4.6)	Réponse au Stress
BX255377	Ge005B04	3_XpanSC	5_euSC	na	na	Elastin precursor - human (fragment) {Homo sapiens;} , partial	Réponse au Stress
BX678613	RS12A03	3_XpanSC	5_euSC	na	na	Gamma-thionin homolog precursor	Réponse au Stress
BX680013	RS38B06	3_XpanSC	5_euSC	na	na	GbJAAD26879.1	Synthèse de protéines
CR393662	RN63B06	4_XpanSC	5_euSC	1_XpanSL	10_euSL	Proton-exporting ATPase (EC 3.6.3.6)	Transport
AL749972	AS03C11	4_XpanSC	5_euSC	2_XpanSL	18_euSL	Endopeptidase Clp subunit (EC 3.4.21.92)	Devenir des protéines
BX255434	Ge005G04	4_XpanSC	5_euSC	2_XpanSL	18_euSL	UDP-glucose 4-epimerase (EC 5.1.3.2)	Métabolisme
AL749902	AS02D12	4_XpanSC	5_euSC	4_XpanSL	10_euSL	putative protein	_
non soumis	RN28G08	4_XpanSC	5_euSC	4_XpanSL	10_euSL	putative protein	_
BX252065	Ge059C02	4_XpanSC	5_euSC	5_XpanSL	19_euSL	17.9 kDa class II heat shock protein	Réponse au Stress
BX252185	Ge061A06	4_XpanSC	5_euSC	5_XpanSL	20_euSL	Mitochondrial oxaloacetate transport protein	Transport
AL750008	AS03G06	4_XpanSC	5_euSC	na	10_euSL	Carbonate dehydratase (EC 4.2.1.1)	Métabolisme
BX252434	Ge064E04	4_XpanSC	5_euSC	na	18_euSL	Coatomer alpha subunit	Traffic intracellulaire
BX250510	Ge037H06	4_XpanSC	5_euSC	na	na	putative protein	_
BX681782	RS64F10	4_XpanSC	5_euSC	na	na	putative protein	_
AL750449	RN02F06	4_XpanSC	5_euSC	na	na	Peptidylprolyl isomerase (EC 5.2.1.8)	Devenir des protéines
BX252055	Ge059B02	4_XpanSC	5_euSC	na	na	subtilisin-like serine proteinase - Arabidopsis thaliana {Arabido	Devenir des protéines
BX255712	Ge008H07	4_XpanSC	5_euSC	na	na	Thioredoxin H-type	Energie
AL749773	AN07G08	4_XpanSC	5_euSC	na	na	Esterase D	Métabolisme
BX252271	Ge062B09	4_XpanSC	5_euSC	na	na	Alcohol dehydrogenase class III (Glutathione-dependent forma	Métabolisme
BX682721	Ge120F04	4_XpanSC	5_euSC	na	na	Uridylate kinase (EC 2.7.4.3)	Métabolisme
BX250726	Ge040E06	4_XpanSC	5_euSC	na	na	Unknown	Non-classé
CR354614	Ge124A01	4_XpanSC	5_euSC	na	na	Blr1429 protein, partial (18%)	Non-classé
BX251918	Ge057C01	4_XpanSC	5_euSC	na	na	Ubiquitin-conjugating enzyme (EC 6.3.2.19)	Organisation cellulaire
BX253993	Ge091H07	4_XpanSC	5_euSC	na	na	Ubiquitin-conjugating enzyme E2-17 kDa (EC 6.3.2.19)	Organisation cellulaire
BX251839	Ge056B11	4_XpanSC	5_euSC	na	na	Leucine-rich repeat-containing protein 2	Réponse au Stress
BX254611	Ge101C06	4_XpanSC	5_euSC	na	na	ADP-ribosylation factor	Traffic intracellulaire
BX676912	RN13G11	4_XpanSC	5_euSC	na	na	Acyl carrier protein, chloroplast precursor	Traffic intracellulaire
AL749609	AN02H06	4_XpanSC	6_euSC	1_XpanSL	10_euSL	putative protein	_
BX681235	RS55H11	4_XpanSC	6_euSC	1_XpanSL	5_euSL	Glutathione transferase (EC 2.5.1.18)	Métabolisme
CR394036	RN72F06	4_XpanSC	6_euSC	5_XpanSL	19_euSL	Agglutinin alpha chain (MPA)	Organisation cellulaire
BX680909	RS50F11	4_XpanSC	6_euSC	na	13_euSL	Pinus sylvestris class III peroxidase	Réponse au Stress
BX681552	RS61B05	4_XpanSC	6_euSC	na	18_euSL	putative protein	_
BX255613	Ge007G03	4_XpanSC	6_euSC	na	na	putative protein	_
BX677297	RN20B06	4_XpanSC	6_euSC	na	na	putative protein	_
AL750365	RN01F10	4_XpanSC	6_euSC	na	na	Peroxidase precursor (EC 1.11.1.7)	Réponse au Stress
BX255626	Ge007H04	4_XpanSC	6_euSC	na	na	Small zinc finger-like protein, complete	Transcription
BX251544	Ge051A06	5_XpanSC	1_euSC	2_XpanSL	18_euSL	putative protein	_
BX253632	Ge086B12	5_XpanSC	1_euSC	2_XpanSL	21_euSL	Snakin-1, partial (67%)	Réponse au Stress
CR392585	RN35A10	5_XpanSC	1_euSC	na	na	Peroxidase precursor (EC 1.11.1.7)	Réponse au Stress
non soumis	RN59A03	5_XpanSC	6_euSC	na	13_euSL	putative protein	_
BX667562	RN09E10	5_XpanSC	7_euSC	2_XpanSL	17_euSL	putative protein	_
BX679297	RS22F03	5_XpanSC	7_euSC	2_XpanSL	17_euSL	putative protein	_
CR393563	RN61G09	5_XpanSC	7_euSC	2_XpanSL	17_euSL	MERI-5 protein precursor (Xyloglucan:xyloglucosyl transferase	Croissance cellulaire, division
CR393320	RN58E01	5_XpanSC	7_euSC	2_XpanSL	17_euSL	Cytochrome P450 (EC 1.14.-.-)	Métabolisme
AL750511	RN03D01	5_XpanSC	7_euSC	2_XpanSL	20_euSL	PtaAGP5 (from Pinus taeda)	Non-classé
BX679458	RS26B10	5_XpanSC	7_euSC	2_XpanSL	22_euSL	Proline-rich protein	Non-classé
BX679871	RS35C05	5_XpanSC	7_euSC	2_XpanSL	22_euSL	Pinus taeda clone PtaAGP5	Non-classé
CR392453	RN32D03	5_XpanSC	7_euSC	2_XpanSL	22_euSL	Peroxidase precursor (EC 1.11.1.7)	Réponse au Stress
BX679939	RS36F06	5_XpanSC	7_euSC	5_XpanSL	19_euSL	Peroxidase precursor (EC 1.11.1.7)	Réponse au Stress
CR392825	RN48H03	5_XpanSC	7_euSC	5_XpanSL	20_euSL	Epimerase/dehydratase	Métabolisme
AL750498	RN03B11	5_XpanSC	7_euSC	6_XpanSL	16_euSL	NADPH dehydrogenase (EC 1.6.99.1) / 12-OXOPHYTODIENO	Energie
BX681893	RS66D06	5_XpanSC	7_euSC	6_XpanSL	17_euSL	putative protein	_
CR393855	RN66C01	5_XpanSC	7_euSC	7_XpanSL	13_euSL	putative protein	_
BX679759	RS33B10	5_XpanSC	7_euSC	7_XpanSL	15_euSL	Adenylate kinase (EC 2.7.4.3)	Métabolisme
CR392379	RN31A10	5_XpanSC	7_euSC	7_XpanSL	15_euSL	Beta-glucosidase precursor (EC 3.2.1.21)	Métabolisme
BX678105	RN56D01	5_XpanSC	7_euSC	na	14_euSL	putative protein	_
BX676790	RN11A09	5_XpanSC	7_euSC	na	na	putative protein	_
BX680966	RS51E10	5_XpanSC	7_euSC	na	na	putative protein	_
BX676888	RN13E06	5_XpanSC	7_euSC	na	na	Carbonate dehydratase (EC 4.2.1.1)	Métabolisme
BX250172	Ge033C10	6_XpanSC	2_euSC	1_XpanSL	10_euSL	putative protein	_
BX678088	RN56B05	6_XpanSC	2_euSC	1_XpanSL	10_euSL	putative protein	_
CR392434	RN31H05	6_XpanSC	2_euSC	1_XpanSL	5_euSL	Seven transmembrane protein 1	Communication/Transduction du signal
BX249564	Ge025E05	6_XpanSC	2_euSC	1_XpanSL	7_euSL	Dehydrin	Réponse au Stress
BX679749	RS32G05	6_XpanSC	2_euSC	1_XpanSL	8_euSL	putative protein	_
BX680492	RS44C05	6_XpanSC	2_euSC	3_XpanSL	2_euSL	putative protein	_
BX250343	Ge035G01	6_XpanSC	2_euSC	na	na	putative protein	_
BX251314	Ge048A11	6_XpanSC	2_euSC	na	na	putative protein	_
BX678858	RS16B02	6_XpanSC	2_euSC	na	na	putative protein	_
BX680817	RS49A12	6_XpanSC	2_euSC	na	na	putative protein	_
BX249906	Ge030B06	6_XpanSC	2_euSC	na	na	Hydrophobic protein RCI2 (Low temperature and salt induced)	Non-classé
BX253650	Ge086D09	6_XpanSC	2_euSC	na	na	Unknown	Non-classé
BX680166	RS40A07	6_XpanSC	2_euSC	na	na	Pathogenesis-related protein	Réponse au Stress
BX681387	RS58C02	6_XpanSC	2_euSC	na	na	Licheninase precursor (EC 3.2.1.73)	Réponse au Stress
CR394461	RN79E12	6_XpanSC	2_euSC	na	na	Ethylene-responsive protein, putative	Réponse au Stress
BX681992	RS67G12	6_XpanSC	2_euSC	na	na	Fluconazole resistance protein 1	Transport
BX254733	Ge103E05	6_XpanSC	3_euSC	4_XpanSL	10_euSL	Putative F-box protein (Fragment), partial (38%)	Transcription
BX678010	RN55B05	6_XpanSC	4_euSC	1_XpanSL	10_euSL	putative protein	_
BX251816	Ge055H04	6_XpanSC	4_euSC	1_XpanSL	10_euSL	Heat shock 70kd protein	Réponse au Stress
BX682091	RS70D01	6_XpanSC	4_euSC	1_XpanSL	3_euSL	putative protein	_
BX682453	Ge116H01	6_XpanSC	4_euSC	1_XpanSL	5_euSL	Receptor-like protein kinase, partial (70%)	Communication/Transduction du signal
BX680502	RS44D07	6_XpanSC	4_euSC	na	na	putative protein	_
AL751155	RS05F11	na	1_euSC	na	na	putative protein	_
BX666044	RN09A04	na	1_euSC	na	na	putative protein	_
BX677392	RN38C03	na	1_euSC	na	na	putative protein	_
BX681214	RS55D09	na	1_euSC	na	na	putative protein	_

BX677655	RN42G05	na	1_euSC	na	na	Flavonoid 3',5'-hydroxylase (EC 1.14.-.-)	Métabolisme
AL751090	RS05A04	na	1_euSC	na	na	Protein kinase (EC 2.7.1.-)	Non-classé
BX681370	RS57H02	na	1_euSC	na	na	Unknown	Non-classé
BC065529	Ha_PP758	na	2_euSC	na	na	_	_
BX252971	Ge075H01	na	2_euSC	na	na	putative protein	_
BX680905	RS50E05	na	2_euSC	na	na	Fumarate hydratase, mitochondrial precursor (EC 4.2.1.2)	Energie
BF609706	NXSI_48_D6	na	5_euSC	na	na	malate dehydrogenase	Métabolisme
BX251681	Ge053H04	na	5_euSC	na	na	Germin(Oxalate oxidase)-like protein, partial (24%)	Croissance cellulaire, division
BX678708	RS14B03	na	6_euSC	2_XpanSL	22_euSL	Basic blue protein (Plantacyanin)	Energie
AW870208	NXNV_125_E12	na	6_euSC	na	18_euSL	Aspartate transaminase	Métabolisme
BE643924	NXCI_50_B7	na	6_euSC	na	18_euSL	S-adenosyl-Met synthase	Métabolisme
BG039825	NXSI_104_E11	na	6_euSC	na	18_euSL	Ornitine transaminase	Métabolisme
BX679916	RS36C01	na	6_euSC	na	na	Metallophosphatase	Devenir des protéines
BE761959	NXCI_75_C7	na	6_euSC	na	na	Glutamate synthase	Métabolisme
BF049732	NXCI_106_D10	na	6_euSC	na	na	Fumarase	Métabolisme
BG039488	NXSI_99_F10	na	6_euSC	na	na	S-adenosyl-Met decarboxylase	Métabolisme
BX681771	RS64D09	na	6_euSC	na	na	60S ribosomal protein L34	Synthèse de protéines
BX677447	RN39C09	na	7_euSC	2_XpanSL	20_euSL	putative protein	_
BX251276	Ge047F03	na	na	1_XpanSL	10_euSL	putative protein	_
BX251336	Ge048D03	na	na	1_XpanSL	10_euSL	putative protein	_
BX679686	RS31C10	na	na	1_XpanSL	10_euSL	putative protein	_
CR392177	RN27F04	na	na	1_XpanSL	10_euSL	putative protein	_
CR392898	RN49G02	na	na	1_XpanSL	10_euSL	putative protein	_
CR393412	RN59G08	na	na	1_XpanSL	10_euSL	putative protein	_
CR393972	RN70H03	na	na	1_XpanSL	10_euSL	putative protein	_
CR394320	RN76G05	na	na	1_XpanSL	10_euSL	putative protein	_
BX255369	Ge005A06	na	na	1_XpanSL	10_euSL	Glycine-rich RNA-binding protein, complete	Communication/Transduction du signal
BX681632	RS62C04	na	na	1_XpanSL	10_euSL	Receptor protein kinase TMK1 precursor	Communication/Transduction du signal
BX681320	RS57A03	na	na	1_XpanSL	10_euSL	Chloroplast nucleoid DNA binding protein, putative	Devenir des protéines
BX681414	RS58H06	na	na	1_XpanSL	10_euSL	Hydroquinone glucosyltransferase (EC 2.4.1.218)	Energie
AL751120	RS05C10	na	na	1_XpanSL	10_euSL	S-adenosylmethionine-dependent methyltransferase T12H1.6	Métabolisme
BX249460	Ge024C08	na	na	1_XpanSL	10_euSL	Mitochondrial aldehyde dehydrogenase, partial (91%)	Métabolisme
BX249550	Ge025D03	na	na	1_XpanSL	10_euSL	Citrate synthase	Métabolisme
BX253959	Ge091C11	na	na	1_XpanSL	10_euSL	Putative receptor protein kinase TMK1 precursor (EC 2.7.1.-)	Métabolisme
BX677309	RN20D07	na	na	1_XpanSL	10_euSL	Glutathione transferase (EC 2.5.1.18)	Métabolisme
BX679517	RS27D07	na	na	1_XpanSL	10_euSL	Inosine-5'-monophosphate dehydrogenase (EC 1.1.1.2)	Métabolisme
BX679620	RS29H05	na	na	1_XpanSL	10_euSL	Serine acetyltransferase	Métabolisme
BX679924	RS36D05	na	na	1_XpanSL	10_euSL	Iron(III)-zinc(II) purple acid phosphatase precursor	Métabolisme
BX681925	RS66H01	na	na	1_XpanSL	10_euSL	Hydroxymethylglutaryl-CoA synthase (EC 2.3.3.10)	Métabolisme
CR392209	RN28C02	na	na	1_XpanSL	10_euSL	Phenylalanine ammonia-lyase (EC 4.3.1.5)	Métabolisme
BX249499	Ge024G05	na	na	1_XpanSL	10_euSL	Pheorphorin-dz1 protein precursor, partial (7%)	Non-classé
BX249846	Ge029D04	na	na	1_XpanSL	10_euSL	At2g43970/F6E13.10 {Arabidopsis thaliana;} , partial (41%)	Non-classé
BX251850	Ge056D01	na	na	1_XpanSL	10_euSL	Hypothetical protein	Non-classé
BX254217	Ge095A09	na	na	1_XpanSL	10_euSL	Unknown	Non-classé
BX681032	RS52C04	na	na	1_XpanSL	10_euSL	Unknown	Non-classé
BX681664	RS62F06	na	na	1_XpanSL	10_euSL	Unknown	Non-classé
CR392445	RN32B03	na	na	1_XpanSL	10_euSL	ATAF1-like protein (Fragment)	Non-classé
CR392966	RN50F06	na	na	1_XpanSL	10_euSL	Unknown	Non-classé
BX677677	RN43A04	na	na	1_XpanSL	10_euSL	9-cis-epoxycarotenoid dioxygenase	Réponse au Stress
BX678362	RS08B03	na	na	1_XpanSL	10_euSL	Protein-methionine-S-oxide reductase (EC 1.8.4.6)	Réponse au Stress
BX681263	RS56C03	na	na	1_XpanSL	10_euSL	EDS1-like protein	Réponse au Stress
BX678954	RS17D12	na	na	1_XpanSL	10_euSL	small subunit ribosomal RNA gene	Synthèse de protéines
BX681072	RS52H08	na	na	1_XpanSL	10_euSL	Zinc finger protein 216	Transcription
BX682945	Ge123H04	na	na	1_XpanSL	10_euSL	Zinc-finger protein, partial (53%)	Transcription
AL750557	RN03G12	na	na	1_XpanSL	10_euSL	Plasma membrane H+ ATPase (EC 3.6.1.3)	Transport
BX248876	Ge002B04	na	na	1_XpanSL	11_euSL	putative protein	_
BX249964	Ge030G11	na	na	1_XpanSL	11_euSL	putative protein	_
BX666045	RN09A05	na	na	1_XpanSL	11_euSL	putative protein	_
BX677594	RN41G01	na	na	1_XpanSL	11_euSL	putative protein	_
BX677863	RN45G02	na	na	1_XpanSL	11_euSL	putative protein	_
BX678422	RS08H04	na	na	1_XpanSL	11_euSL	putative protein	_
BX679288	RS22E05	na	na	1_XpanSL	11_euSL	putative protein	_
BX679547	RS28B01	na	na	1_XpanSL	11_euSL	putative protein	_
BX679599	RS29C06	na	na	1_XpanSL	11_euSL	putative protein	_
BX679769	RS33E01	na	na	1_XpanSL	11_euSL	putative protein	_
BX679879	RS35D06	na	na	1_XpanSL	11_euSL	putative protein	_
CR394055	RN73A05	na	na	1_XpanSL	11_euSL	putative protein	_
BX677790	RN44G03	na	na	1_XpanSL	11_euSL	Phospholipase C	Communication/Transduction du signal
CR393473	RN60F01	na	na	1_XpanSL	11_euSL	Phospholipase D alpha 1 precursor (EC 3.1.4.4)	Communication/Transduction du signal
BX667570	RN09F07	na	na	1_XpanSL	11_euSL	Presenilin	Croissance cellulaire, division
BX680304	RS41G02	na	na	1_XpanSL	11_euSL	Structural maintenance of chromosomes (SMC)-like protein	Croissance cellulaire, division
BX679155	RS20F02	na	na	1_XpanSL	11_euSL	Proteasome endopeptidase complex subunit alpha type 6 (EC 3	Devenir des protéines
BX681672	RS62G04	na	na	1_XpanSL	11_euSL	Carboxypeptidase C (EC 3.4.16.5)	Devenir des protéines
CR393041	RN51H01	na	na	1_XpanSL	11_euSL	26S protease regulatory subunit 4 homolog	Devenir des protéines
BX249635	Ge026E03	na	na	1_XpanSL	11_euSL	Phosphoenolpyruvate carboxylase , partial (50%)	Energie
BX677890	RN46B05	na	na	1_XpanSL	11_euSL	Pyruvate kinase (EC 2.7.1.40)	Energie
BX680080	RS38H04	na	na	1_XpanSL	11_euSL	AAA-type ATPase	Energie
CR392718	RN37A12	na	na	1_XpanSL	11_euSL	NADH dehydrogenase (quinone) 23 kDa subunit (1.6.99.5)	Energie
BX249364	Ge023B09	na	na	1_XpanSL	11_euSL	glutathione-dependent formaldehyde dehydrogenase {Arabidop	Métabolisme
BX677699	RN43C10	na	na	1_XpanSL	11_euSL	RuBisCO subunit binding-protein beta subunit	Synthèse de protéines
CR394072	RN73D02	na	na	1_XpanSL	11_euSL	Cytochrome P450 77A3 (EC 1.14.-.-)	Métabolisme
AL751305	RS07C12	na	na	1_XpanSL	11_euSL	Hypothetical 58.9 kDa protein	Non-classé
BX250697	Ge040B09	na	na	1_XpanSL	11_euSL	F16A14.7, partial (50%)	Non-classé
BX254570	Ge100E02	na	na	1_XpanSL	11_euSL	Unknown	Non-classé
BX677573	RN41D05	na	na	1_XpanSL	11_euSL	Yippee-like protein	Non-classé
BX677604	RN41H04	na	na	1_XpanSL	11_euSL	Hypothetical 20.8 kDa protein	Non-classé
BX679201	RS21D07	na	na	1_XpanSL	11_euSL	SelT-like protein precursor	Non-classé
BX679424	RS25B12	na	na	1_XpanSL	11_euSL	Unknown	Non-classé
BX679842	RS34G02	na	na	1_XpanSL	11_euSL	Unknown	Non-classé
BX680987	RS51G11	na	na	1_XpanSL	11_euSL	Hyphally regulated protein precursor	Non-classé
BX681149	RS54D01	na	na	1_XpanSL	11_euSL	Unknown	Non-classé
BX681605	RS61H03	na	na	1_XpanSL	11_euSL	Unknown	Non-classé
BX681900	RS66E03	na	na	1_XpanSL	11_euSL	Unknown	Non-classé
CR392836	RN48F07	na	na	1_XpanSL	11_euSL	Unknown	Non-classé
BX250835	Ge041H12	na	na	1_XpanSL	11_euSL	Ubiquitin carboxyl-terminal hydrolase (EC 3.1.2.	Organisation cellulaire
BX677552	RN41B01	na	na	1_XpanSL	11_euSL	Glycosyl transferase-like protein	Organisation cellulaire
BX666054	RN09B03	na	na	1_XpanSL	11_euSL	Abscisic stress ripening protein	Réponse au Stress
BX678372	RS08C03	na	na	1_XpanSL	11_euSL	Cationic peroxidase precursor (EC 1.11.1.7)	Réponse au Stress
BX251593	Ge051G02	na	na	1_XpanSL	11_euSL	Eukaryotic peptide chain release factor subunit 1	Synthèse de protéines
CR392980	RN51A02	na	na	1_XpanSL	11_euSL	40S ribosomal protein S8	Synthèse de protéines
BX252302	Ge062E12	na	na	1_XpanSL	11_euSL	RNA-binding protein RZ-1, partial (9%)	Transcription
BX681312	RS56H05	na	na	1_XpanSL	11_euSL	Lil3 protein	Transcription
BX667565	RN09F01	na	na	1_XpanSL	11_euSL	Amino acid permease (Amino acid carrier, putative)	Transport
BX681797	RS65A06	na	na	1_XpanSL	11_euSL	Oxysterol-binding protein homolog	Transport
BX250630	Ge039D02	na	na	1_XpanSL	12_euSL	putative protein	_
BX681837	RS65E10	na	na	1_XpanSL	12_euSL	putative protein	_
BX682356	Ge115E03	na	na	1_XpanSL	12_euSL	putative protein	_
CR393621	RN62F03	na	na	1_XpanSL	12_euSL	putative protein	_
AL749938	AS02H06	na	na	1_XpanSL	12_euSL	Naringenin 3-dioxygenase (EC 1.14.11.9)	Métabolisme
BX677417	RN38F06	na	na	1_XpanSL	12_euSL	Short-chain type dehydrogenase/reductase (EC 1.-.-.-)	Métabolisme
BX679708	RS32A02	na	na	1_XpanSL	12_euSL	Glutathione transferase (EC 2.5.1.18)	Métabolisme
BX679917	RS36C02	na	na	1_XpanSL	12_euSL	Protein kinase (EC 2.7.1.-)	Métabolisme
BX681828	RS65D10	na	na	1_XpanSL	12_euSL	Flavonol 3-O-glucosyltransferase 5 (EC 2.4.1.91)	Métabolisme
AL750277	AS07F11	na	na	1_XpanSL	12_euSL	Unknown	Non-classé
BX252300	Ge062E09	na	na	1_XpanSL	12_euSL	Nuc-1 negative regulatory protein preg	Non-classé
BX252532	Ge066B06	na	na	1_XpanSL	12_euSL	Potyvirus VPg interacting protein, partial (55%)	Non-classé
BX254033	Ge092D12	na	na	1_XpanSL	12_euSL	Remorin 1, partial (57%)	Non-classé
BX254347	Ge097C07	na	na	1_XpanSL	12_euSL	CoA-thioester hydrolase CHY1 (3-hydroxyisobutyryl-coenzyme	Non-classé
BX254354	Ge097D04	na	na	1_XpanSL	12_euSL	Unknown	Non-classé
BX681640	RS62D03	na	na	1_XpanSL	12_euSL	Unknown	Non-classé
BX681915	RS66F09	na	na	1_XpanSL	12_euSL	Embryogenesis-associated protein EMB3	Non-classé
CR393456	RN60D06	na	na	1_XpanSL	12_euSL	Serine-rich protein	Non-classé
BX682172	RS73A01	na	na	1_XpanSL	12_euSL	Histone H2A	Organisation cellulaire
CR393703	RN63G04	na	na	1_XpanSL	12_euSL	Ubiquitin conjugating enzyme spm2	Organisation cellulaire
BX255310	Ge004C06	na	na	1_XpanSL	12_euSL	catalase 1 [Pinus pinea]	Réponse au Stress
BX680673	RS47B12	na	na	1_XpanSL	12_euSL	Hypothetical protein (DNAJ domain-containing unknown protein	Réponse au Stress
BX681356	RS57E07	na	na	1_XpanSL	12_euSL	Late embryogenesis abundant protein	Réponse au Stress
BX680147	RS39G10	na	na	1_XpanSL	12_euSL	RNA helicase	Synthèse de protéines
BX252360	Ge063D06	na	na	1_XpanSL	12_euSL	Sec13p, partial (90%)	Traffic intracellulaire
CR394157	RN74F01	na	na	1_XpanSL	12_euSL	ADP,ATP carrier protein (ADP/ATP translocase)	Transport
AL749568	AN01E10	na	na	1_XpanSL	3_euSL	putative protein	_
BX678810	RS15D08	na	na	1_XpanSL	3_euSL	putative protein	_
CR393691	RN63E10	na	na	1_XpanSL	3_euSL	putative protein	_
BX667551	RN09D10	na	na	1_XpanSL	3_euSL	4-hydroxyphenylpyruvate dioxygenase (EC 1.13.11.27)	Métabolisme
AL750468	RN02H01	na	na	1_XpanSL	3_euSL	F21O3.6 Hypothetical protein	Non-classé
BX680652	RS46F06	na	na	1_XpanSL	3_euSL	hsr203J	Non-classé
BX249522	Ge025A06	na	na	1_XpanSL	4_euSL	putative protein	_
BX252009	Ge058E02	na	na	1_XpanSL	4_euSL	putative protein	_
BX255208	Ge112H01	na	na	1_XpanSL	4_euSL	putative protein	_
BX680824	RS49A07	na	na	1_XpanSL	4_euSL	Arginine methyltransferase-like protein (EC 2.1.1.-)	Métabolisme
CR394120	RN74B04	na	na	1_XpanSL	4_euSL	Serine/threonine specific protein phosphatase 2C (EC 3.1.3.16)	Métabolisme
BX254648	Ge102A03	na	na	1_XpanSL	4_euSL	At1g64110/F22C12_22, partial (47%)	Non-classé
BX677222	RN19E10	na	na	1_XpanSL	4_euSL	Hypothetical 40.2 kDa protein	Non-classé
BX678617	RS12A09	na	na	1_XpanSL	4_euSL	IN2-1 protein	Non-classé
BX681832	RS65E04	na	na	1_XpanSL	4_euSL	Pathogenesis-related protein PRB1-2 precursor	Réponse au Stress
BX680183	RS40C01	na	na	1_XpanSL	4_euSL	Membrane transporter protein	Transport
CR394220	RN75C12	na	na	1_XpanSL	4_euSL	Peptide transporter PTR2-B	Transport
BX250696	Ge040B08	na	na	1_XpanSL	5_euSL	putative protein	_
BX252380	Ge063F05	na	na	1_XpanSL	5_euSL	putative protein	_
BX255575	Ge007D01	na	na	1_XpanSL	5_euSL	putative protein	_
BX677876	RN45H07	na	na	1_XpanSL	5_euSL	putative protein	_
BX679225	RS21G10	na	na	1_XpanSL	5_euSL	putative protein	_
BX679448	RS25H03	na	na	1_XpanSL	5_euSL	putative protein	_

BX680451	RS43E08	na	na	1_XpanSL	5_euSL	putative protein	_
BX681762	RS64B10	na	na	1_XpanSL	5_euSL	putative protein	_
CR392528	RN33F11	na	na	1_XpanSL	5_euSL	putative protein	_
CR392544	RN33H09	na	na	1_XpanSL	5_euSL	putative protein	_
CR393620	RN62F02	na	na	1_XpanSL	5_euSL	putative protein	_
CR393282	RN54H02	na	na	1_XpanSL	5_euSL	Germin-like protein subfamily 1 member 7 precursor	Croissance cellulaire, division
AL749654	AN05E01	na	na	1_XpanSL	5_euSL	Oxygen-evolving enhancer protein 1, chloroplast precursor	Energie
AL749666	AN05G02	na	na	1_XpanSL	5_euSL	Photosystem I reaction center subunit N, chloropla	Energie
AL749783	AS01A02	na	na	1_XpanSL	5_euSL	H(+)-transporting two-sector ATPase delta subunit (EC 3.6.3.14	Energie
BX251416	Ge049C07	na	na	1_XpanSL	5_euSL	Oxygen-evolving enhancer protein 3-1, chloroplast precursor	Energie
BX678416	RS08G10	na	na	1_XpanSL	5_euSL	Chlorophyll A-B binding protein, chloroplastic	Energie
BX676897	RN13F04	na	na	1_XpanSL	5_euSL	Syringomycin biosynthesis enzyme-like protein	Métabolisme
BX678400	RS08F03	na	na	1_XpanSL	5_euSL	Glutathione transferase (EC 2.5.1.18)	Métabolisme
BX679422	RS25B09	na	na	1_XpanSL	5_euSL	Cysteine synthase, chloroplast precursor (EC 2.5.1.47)	Métabolisme
CR392473	RN32G07	na	na	1_XpanSL	5_euSL	Glutamate synthase (NADH) (Sb14) (EC 1.4.1.14)(Picea maria	Métabolisme
CR393272	RN54G03	na	na	1_XpanSL	5_euSL	Transaldolase (EC 2.2.1.2)	Métabolisme
BX249398	Ge023F01	na	na	1_XpanSL	5_euSL	Auxin-repressed 12.5 kd protein	Non-classé
BX254155	Ge094B08	na	na	1_XpanSL	5_euSL	Hypothetical protein	Non-classé
BX680601	RS45G02	na	na	1_XpanSL	5_euSL	Hypothetical 25.9 kDa protein	Non-classé
BX680674	RS47B01	na	na	1_XpanSL	5_euSL	Hypothetical 34.5 kDa protein	Non-classé
BX681199	RS55C12	na	na	1_XpanSL	5_euSL	Unknown	Non-classé
BX681400	RS58E03	na	na	1_XpanSL	5_euSL	Unknown	Non-classé
AL750545	RN03F11	na	na	1_XpanSL	5_euSL	Catalase isozyme (EC 1.11.1.6).	Réponse au Stress
BX249652	Ge026F11	na	na	1_XpanSL	5_euSL	ASR protein, partial (39%)	Réponse au Stress
BX681166	RS54F02	na	na	1_XpanSL	5_euSL	NBS/LRR	Réponse au Stress
BX250756	Ge040H03	na	na	1_XpanSL	5_euSL	F11M15.7 protein	Transcription
CR394310	RN76F04	na	na	1_XpanSL	5_euSL	Transcription factor HBP-1b homolog	Transcription
AL750677	RN06B02	na	na	1_XpanSL	7_euSL	putative protein	_
BX253114	Ge078C04	na	na	1_XpanSL	7_euSL	Alkaline alpha galactosidase II, partial (84%)	Métabolisme
BX682057	RS69D04	na	na	1_XpanSL	8_euSL	putative protein	_
BX679999	RS35D10	na	na	1_XpanSL	8_euSL	Unknown	Non-classé
BX679598	RS29C05	na	na	1_XpanSL	9_euSL	putative protein	_
BX682159	RS72E06	na	na	1_XpanSL	9_euSL	Flavonoid 3',5'-hydroxylase (EC 1.14.-.-)	Métabolisme
AL750653	RN05H02	na	na	2_XpanSL	17_euSL	60S ribosomal protein L27a	Synthèse de protéines
BX255618	Ge007G08	na	na	2_XpanSL	17_euSL	putative protein	_
BX677213	RN19D06	na	na	2_XpanSL	17_euSL	putative protein	_
BX677614	RN42A04	na	na	2_XpanSL	17_euSL	putative protein	_
BX249855	Ge029E02	na	na	2_XpanSL	17_euSL	subtilisin-like proteinase - Arabidopsis thaliana {Arabidopsis tha	Devenir des protéines
BX677823	RN45B11	na	na	2_XpanSL	17_euSL	Chloroplast nucleoid DNA binding protein	Devenir des protéines
BX252441	Ge064F03	na	na	2_XpanSL	17_euSL	Ubiquinol-cytochrome C reductase iron-sulfur subunit (EC 1.10.	Energie
BX252930	Ge075B02	na	na	2_XpanSL	17_euSL	ATP citrate lyase a-subunit, partial (84%)	Métabolisme
BX254041	Ge092E09	na	na	2_XpanSL	17_euSL	serine protease	Métabolisme
BX678013	RN55B08	na	na	2_XpanSL	17_euSL	Trans-cinnamate 4-monooxygenase (EC 1.14.13.11)	Métabolisme
BX255014	Ge109C07	na	na	2_XpanSL	17_euSL	Glucose-1-phosphate adenylyltransferase small subunit, chloro	Non-classé
BX678345	RN69H07	na	na	2_XpanSL	17_euSL	Vegetatible incompatibility protein HET-E-1	Non-classé
BX682334	Ge115B11	na	na	2_XpanSL	17_euSL	Surfeit 1.	Non-classé
BX250930	Ge043A10	na	na	2_XpanSL	17_euSL	MERI-5 protein precursor (Endo-xyloglucan transferase) / Bras	Organisation cellulaire
BX677123	RN18A04	na	na	2_XpanSL	17_euSL	Histone H2B	Organisation cellulaire
BX682714	Ge120E04	na	na	2_XpanSL	17_euSL	Cellulose synthase	Organisation cellulaire
BX252277	Ge062C04	na	na	2_XpanSL	17_euSL	Harpin-induced protein	Réponse au Stress
BX252348	Ge063C05	na	na	2_XpanSL	18_euSL	putative protein	_
BX252626	Ge068A11	na	na	2_XpanSL	18_euSL	putative protein	_
CR393458	RN60D09	na	na	2_XpanSL	18_euSL	putative protein	_
CR394413	RN78G11	na	na	2_XpanSL	18_euSL	putative protein	_
non soumis	RN27G07	na	na	2_XpanSL	18_euSL	putative protein	_
BX681723	RS63E04	na	na	2_XpanSL	18_euSL	Inositol-1(or 4)-monophosphatase (EC 3.1.3.25)	Communication/Transduction du signal
BX251676	Ge053G11	na	na	2_XpanSL	18_euSL	Xyloglucan:xyloglucosyl transferase (EC 2.4.1.207)	Croissance cellulaire, division
BX677740	RN44B02	na	na	2_XpanSL	18_euSL	Cell division cycle protein 48 homolog (Valosin containing prote	Croissance cellulaire, division
BX253484	Ge084A09	na	na	2_XpanSL	18_euSL	Peptidyl-prolyl cis-trans isomerase (EC 5.2.1.8) (Cyclophilin)	Devenir des protéines
AL750697	RN06C10	na	na	2_XpanSL	18_euSL	Thiazole biosynthetic enzyme, chloroplast precursor	Energie
BX249115	Ge020B11	na	na	2_XpanSL	18_euSL	Dihydrolipoyllysine-residue acetyltransferase component of pyr	Energie
BX251974	Ge058A01	na	na	2_XpanSL	18_euSL	ATP synthase beta chain, mitochondrial precursor	Energie
BX252436	Ge064E07	na	na	2_XpanSL	18_euSL	Diphosphate--fructose-6-phosphate 1-phosphotransferase beta	Energie
BX680864	RS49H06	na	na	2_XpanSL	18_euSL	Pyruvate decarboxylase (EC 4.1.1.1)	Energie
BX681089	RS53C12	na	na	2_XpanSL	18_euSL	Phosphogluconate dehydrogenase (decarboxylating) (EC 1.1.1	Energie
BE607163	NXCI_37_F8	na	na	2_XpanSL	18_euSL	adenosine kinase (AK)	Métabolisme
BX248894	Ge002D03	na	na	2_XpanSL	18_euSL	Lanatoside 15'-O-acetylesterase precursor, partial (28%)	Métabolisme
BX249835	Ge029C07	na	na	2_XpanSL	18_euSL	Adenosine kinase (EC 2.7.1.20)	Métabolisme
BX250471	Ge037D04	na	na	2_XpanSL	18_euSL	(1-4)-beta-mannan endohydrolase precursor (EC 3.2.1.78).	Métabolisme
BX252145	Ge060D11	na	na	2_XpanSL	18_euSL	Hypothetical protein yfbG	Métabolisme
BX253613	Ge085H08	na	na	2_XpanSL	18_euSL	Adenosylhomocysteinase (EC 3.3.1.1)	Métabolisme
BX254773	Ge104C09	na	na	2_XpanSL	18_euSL	Sterol-C-methyltransferase.	Métabolisme
BX254966	Ge108C11	na	na	2_XpanSL	18_euSL	transitional endoplasmic reticulum ATPase - Arabidopsis thalian	Métabolisme
BX678049	RN55F05	na	na	2_XpanSL	18_euSL	Phenylalanine ammonia-lyase (EC 4.3.1.5)	Métabolisme
BX678354	RS08A06	na	na	2_XpanSL	18_euSL	Pinus taeda Trans-cinnamate 4-monooxygenase (EC 1.14.13.1	Métabolisme
BX678666	RS12H05	na	na	2_XpanSL	18_euSL	Alpha-1,4-glucan-protein synthase (UDP-forming) (EC 2.4.1.11	Métabolisme
BX678852	RS16A04	na	na	2_XpanSL	18_euSL	Caffeate O-methyltransferase (EC 2.1.1.68)	Métabolisme
BX679575	RS28F02	na	na	2_XpanSL	18_euSL	Alpha-1,4-glucan-protein synthase (UDP-forming) (EC 2.4.1.11	Métabolisme
BX680340	RS42B09	na	na	2_XpanSL	18_euSL	Naringenin-chalcone synthase (EC 2.3.1.74)	Métabolisme
BX680681	RS47B09	na	na	2_XpanSL	18_euSL	Anter-specific proline-rich protein / Lanatoside 15'-O-acetyleste	Métabolisme
BX682428	Ge116E05	na	na	2_XpanSL	18_euSL	Inositolphosphorylceramide-B C-26 hydroxylase	Métabolisme
CR392294	RN30E11	na	na	2_XpanSL	18_euSL	Omega-6 fatty acid desaturase, endoplasmic reticulum	Métabolisme
CR393529	RN61D03	na	na	2_XpanSL	18_euSL	Malate dehydrogenase, mitochondrial precursor	Métabolisme
AL749952	AS03A10	na	na	2_XpanSL	18_euSL	Hypothetical 38.0 kDa protein (Putative dehydrogenase)	Non-classé
BX250080	Ge032B10	na	na	2_XpanSL	18_euSL	Unknown	Non-classé
BX252312	Ge062G04	na	na	2_XpanSL	18_euSL	33 kDa secretory protein-like, partial (13%)	Non-classé
BX254410	Ge098C07	na	na	2_XpanSL	18_euSL	Glycine-rich protein	Non-classé
BX254445	Ge098F09	na	na	2_XpanSL	18_euSL	high mobility group protein MNB1b - maize (fragment) {Zea ma	Non-classé
BX254994	Ge108H10	na	na	2_XpanSL	18_euSL	Phosphate-induced protein	Non-classé
BX255465	Ge006B03	na	na	2_XpanSL	18_euSL	acetone-cyanohydrin lyase [imported] - Arabidopsis thaliana, pa	Non-classé
BX678353	RS08A05	na	na	2_XpanSL	18_euSL	Hypothetical 12.0 kDa protein	Non-classé
BX680213	RS40F01	na	na	2_XpanSL	18_euSL	Unknown	Non-classé
BX680404	RS43A10	na	na	2_XpanSL	18_euSL	F12F1.23 protein (Hypothetical 36.7 kDa protein)	Non-classé
BX680979	RS51F01	na	na	2_XpanSL	18_euSL	Hypothetical 19.9 kDa protein	Non-classé
CR392777	RN37H08	na	na	2_XpanSL	18_euSL	Unknown	Non-classé
AL751004	RS03A03	na	na	2_XpanSL	18_euSL	Glucan endo-1,3-beta-glucosidase precursor	Organisation cellulaire
BX252682	Ge069A05	na	na	2_XpanSL	18_euSL	Cellulose synthase	Organisation cellulaire
BX254985	Ge108G08	na	na	2_XpanSL	18_euSL	Ubiquitin	Organisation cellulaire
BX255342	Ge004F07	na	na	2_XpanSL	18_euSL	Actin	Organisation cellulaire
BX681614	RS62A02	na	na	2_XpanSL	18_euSL	Tubulin beta chain	Organisation cellulaire
BX682259	Ge114B12	na	na	2_XpanSL	18_euSL	Histone H2A	Organisation cellulaire
BX250350	Ge035G11	na	na	2_XpanSL	18_euSL	Wound-induced basic protein	Réponse au Stress
BX251025	Ge044C12	na	na	2_XpanSL	18_euSL	Cyclophilin, partial (98%)	Réponse au Stress
BX252471	Ge065A11	na	na	2_XpanSL	18_euSL	ENDOCHITINASE (EC 3.2.1.14)	Réponse au Stress
BX252884	Ge074D02	na	na	2_XpanSL	18_euSL	Similarity to ABA induced plasma membrane protein, partial (52	Réponse au Stress
BX253505	Ge084D05	na	na	2_XpanSL	18_euSL	Pinus taeda water deficit inducible protein LP3	Réponse au Stress
BX249385	Ge023D08	na	na	2_XpanSL	18_euSL	60S ribosomal protein	Synthèse de protéines
BX251297	Ge047H03	na	na	2_XpanSL	18_euSL	60S ribosomal protein L10	Synthèse de protéines
BX251853	Ge056D04	na	na	2_XpanSL	18_euSL	40S ribosomal protein S27	Synthèse de protéines
BX252922	Ge074H10	na	na	2_XpanSL	18_euSL	60S ribosomal protein	Synthèse de protéines
BX253036	Ge077A02	na	na	2_XpanSL	18_euSL	Methionine synthase, partial (63%)	Synthèse de protéines
BX253123	Ge078D03	na	na	2_XpanSL	18_euSL	40S ribosomal protein S9 (S7)	Synthèse de protéines
CR392876	RN49D11	na	na	2_XpanSL	18_euSL	40S ribosomal protein S16	Synthèse de protéines
CR393754	RN64G04	na	na	2_XpanSL	18_euSL	40S ribosomal protein S2	Synthèse de protéines
BX677127	RN18A09	na	na	2_XpanSL	18_euSL	GTP-binding nuclear protein RAN-A1	Traffic intracellulaire
BX252601	Ge067D04	na	na	2_XpanSL	18_euSL	Abscisic acid insensitive 5 (ABI5) (BZIP protein), partial (17%)	Transcription
BX680112	RS39C05	na	na	2_XpanSL	18_euSL	Nuclear polyadenylated RNA-binding protein	Transcription
CR392290	RN30E06	na	na	2_XpanSL	18_euSL	DNA-binding protein ESCAROLA	Transcription
BX249892	Ge029H11	na	na	2_XpanSL	18_euSL	Plasma membrane intrinsic protein (Aquaporin)	Transport
BX681365	RS57G02	na	na	2_XpanSL	18_euSL	Tonoplast intrinsic protein	Transport
BX679165	RS20G05	na	na	2_XpanSL	19_euSL	putative protein	_
CR392217	RN28D05	na	na	2_XpanSL	19_euSL	Basic blue protein (Cusacyanin) (Plantacyanin)	Energie
BX250409	Ge036E11	na	na	2_XpanSL	20_euSL	putative protein	_
BX253193	Ge079F01	na	na	2_XpanSL	20_euSL	putative protein	_
BX681039	RS52D12	na	na	2_XpanSL	20_euSL	putative protein	_
CR394223	RN75D03	na	na	2_XpanSL	20_euSL	putative protein	_
BX251900	Ge057A05	na	na	2_XpanSL	20_euSL	Clt1, partial (98%)	Communication/Transduction du signal
AL750564	RN03H07	na	na	2_XpanSL	20_euSL	Alpha galactosyltransferase-like protein	Devenir des protéines
AL749834	AS01F08	na	na	2_XpanSL	20_euSL	Naringenin-chalcone synthase (EC 2.3.1.74)	Métabolisme
BX250516	Ge038A02	na	na	2_XpanSL	20_euSL	Beta-galactosidase , partial (37%)	Métabolisme
BX679860	RS35A06	na	na	2_XpanSL	20_euSL	2-alpha-hydroxytaxane 2-O-benzoyltransferase (EC 2.3.1.166)	Métabolisme
BX253794	Ge088E12	na	na	2_XpanSL	20_euSL	Unknown	Non-classé
BX681172	RS54G01	na	na	2_XpanSL	20_euSL	Endosperm-specific protein-like protein	Non-classé
BX682681	Ge120A02	na	na	2_XpanSL	20_euSL	Anter-specific proline-rich protein	Non-classé
BX252002	Ge058D06	na	na	2_XpanSL	20_euSL	Tubulin beta chain	Organisation cellulaire
BX252968	Ge075G10	na	na	2_XpanSL	20_euSL	Tubulin alpha chain	Organisation cellulaire
BX254406	Ge098C03	na	na	2_XpanSL	20_euSL	Ubiquitin	Organisation cellulaire
BX677988	RN47H03	na	na	2_XpanSL	20_euSL	Tomato extensin mRNA	Organisation cellulaire
BX678152	RN57A03	na	na	2_XpanSL	20_euSL	Peroxidase precursor (EC 1.11.1.7)	Réponse au Stress
BX681454	RS59E06	na	na	2_XpanSL	20_euSL	Nuclear polyadenylated RNA-binding protein NAB4	Transcription
BX682918	Ge123E04	na	na	2_XpanSL	20_euSL	Phenylcoumaran benzylic ether reductase homolog TH1, comp	Transcription
BX679628	RS30A09	na	na	2_XpanSL	20_euSL	Root cap protein	Transport
BX000600	GP	na	na	2_XpanSL	21_euSL	GP	Métabolisme
BX000608	Tubuline	na	na	2_XpanSL	21_euSL	Tubuline	Organisation cellulaire
AL751055	RS03E11	na	na	2_XpanSL	21_euSL	putative protein	_
BX250481	Ge037E07	na	na	2_XpanSL	21_euSL	putative protein	_
BX677752	RN44C09	na	na	2_XpanSL	21_euSL	putative protein	_
BX681964	RS67D07	na	na	2_XpanSL	21_euSL	putative protein	_
BX250683	Ge040A04	na	na	2_XpanSL	21_euSL	Fasciclin-like arabinogalactan protein FLA8	Croissance cellulaire, division
CR392893	RN49F07	na	na	2_XpanSL	21_euSL	Fasciclin-like arabinogalactan-protein 2	Croissance cellulaire, division
BX677473	RN40A01	na	na	2_XpanSL	21_euSL	Glutamate decarboxylase (EC 4.1.1.15)	Métabolisme
BX682426	Ge116E03	na	na	2_XpanSL	21_euSL	Adenosylhomocysteinase (EC 3.3.1.1)	Métabolisme
CR392624	RN35F03	na	na	2_XpanSL	21_euSL	Caffeoyl-CoA O-methyltransferase (EC 2.1.1.104)	Métabolisme

BX250614	Ge039B07	na	na	2_XpanSL	21_euSL	phosphate-induced protein	Non-classé
BX251932	Ge057D07	na	na	2_XpanSL	21_euSL	arabinogalactan protein	Non-classé
BX255417	Ge005E10	na	na	2_XpanSL	21_euSL	P20-Arc, partial (92%)	Non-classé
BX680058	RS38F06	na	na	2_XpanSL	21_euSL	F14L17.20 protein	Non-classé
CR354677	Ge124G05	na	na	2_XpanSL	21_euSL	Unknown	Non-classé
CR354742	Ge125H03	na	na	2_XpanSL	21_euSL	Unknown	Non-classé
BX680247	RS41A09	na	na	2_XpanSL	21_euSL	Glucan endo-1,3-beta-glucosidase precursor	Organisation cellulaire
BX252031	Ge058G08	na	na	2_XpanSL	21_euSL	arabinogalactan protein [Pinus taeda]	Réponse au Stress
BX251702	Ge054B09	na	na	2_XpanSL	21_euSL	Zinc-finger protein NOLZ1, partial (3%)	Transcription
BX680486	RS44B07	na	na	2_XpanSL	21_euSL	Poly(A)-binding protein	Transcription
AL751210	RS06C09	na	na	2_XpanSL	22_euSL	putative protein	_
BX680863	RS49H05	na	na	2_XpanSL	22_euSL	putative protein	_
BX682993	Ge126F07	na	na	2_XpanSL	22_euSL	UDP-glucose 4-epimerase (EC 5.1.3.2)	Métabolisme
BX249241	Ge021G02	na	na	2_XpanSL	22_euSL	Tubulin alpha chain	Organisation cellulaire
BX253001	Ge076C12	na	na	2_XpanSL	22_euSL	Tubulin alpha chain	Organisation cellulaire
BX252074	Ge059D03	na	na	2_XpanSL	22_euSL	arabinogalactan protein	Réponse au Stress
CR392722	RN37B06	na	na	2_XpanSL	22_euSL	Peroxidase precursor (EC 1.11.1.7)	Réponse au Stress
BX681246	RS56A08	na	na	3_XpanSL	1_euSL	putative protein	_
AL750636	RN05F08	na	na	3_XpanSL	1_euSL	Antimicrobial peptide	Réponse au Stress
BX679343	RS23E08	na	na	3_XpanSL	1_euSL	Chaperone protein dnaJ	Réponse au Stress
BX681928	RS66H05	na	na	3_XpanSL	1_euSL	Antimicrobial peptide	Réponse au Stress
BX249673	Ge028H10	na	na	3_XpanSL	12_euSL	putative protein	_
BX678873	RS16C11	na	na	3_XpanSL	12_euSL	putative protein	_
BX679063	RS19C01	na	na	3_XpanSL	12_euSL	putative protein	_
BX681189	RS55B01	na	na	3_XpanSL	12_euSL	putative protein	_
BX681814	RS65C05	na	na	3_XpanSL	12_euSL	putative protein	_
CR392796	RN48B11	na	na	3_XpanSL	12_euSL	putative protein	_
CR394054	RN73A01	na	na	3_XpanSL	12_euSL	putative protein	_
BX677108	RN17G03	na	na	3_XpanSL	12_euSL	Expansin-like precursor	Croissance cellulaire, division
BX678396	RS08E10	na	na	3_XpanSL	12_euSL	Alternative oxidase 1, mitochondrial precursor	Energie
BX677977	RN47F03	na	na	3_XpanSL	12_euSL	Beta-fructofuranosidase (EC 3.2.1.26), cell wall isozyme precur	Métabolisme
AL750195	AS06G11	na	na	3_XpanSL	12_euSL	Hypothetical protein DR0254	Non-classé
BX251976	Ge058A03	na	na	3_XpanSL	12_euSL	Unknown	Non-classé
BX679459	RS26C01	na	na	3_XpanSL	12_euSL	PGPD14 protein (Pollen germination related protein)	Non-classé
CR394058	RN73B01	na	na	3_XpanSL	12_euSL	Unknown	Non-classé
BX683057	Ge127F05	na	na	3_XpanSL	12_euSL	#N/A	#N/A
BX255127	Ge111D09	na	na	3_XpanSL	12_euSL	eukaryotic initiation factor 3E subunit {Arabidopsis thaliana;} , p	Transcription
CR394499	RN50C09	na	na	3_XpanSL	2_euSL	Antimicrobial peptide	Réponse au Stress
BX678841	RS15G11	na	na	3_XpanSL	3_euSL	4-hydroxyphenylpyruvate dioxygenase (EC 1.13.11.27)	Métabolisme
BX680204	RS40E02	na	na	3_XpanSL	3_euSL	4-hydroxyphenylpyruvate dioxygenase (EC 1.13.11.27)	Métabolisme
AL751212	RS06C11	na	na	3_XpanSL	3_euSL	Chaperone protein dnaJ	Réponse au Stress
BX253756	Ge088A07	na	na	3_XpanSL	3_euSL	Aluminum-induced protein-like protein, partial (67%)	Réponse au Stress
CR394078	RN73D08	na	na	3_XpanSL	3_euSL	Licheninase precursor (EC 3.2.1.73)	Réponse au Stress
BX679259	RS22B08	na	na	3_XpanSL	4_euSL	putative protein	_
BX679023	RS18E12	na	na	3_XpanSL	4_euSL	Signal peptidase I (EC 3.4.21.89)	Devenir des protéines
BX249884	Ge029H02	na	na	3_XpanSL	4_euSL	Phosphoenolpyruvate carboxykinase (EC 4.1.4.49)	Energie
BX252988	Ge076B04	na	na	3_XpanSL	4_euSL	Thioredoxin H-type (TRX-H)	Energie
BX254884	Ge106G10	na	na	3_XpanSL	4_euSL	Phosphoenolpyruvate carboxykinase (EC 4.1.4.49)	Energie
BX679683	RS31C02	na	na	3_XpanSL	4_euSL	Cytochrome P450 (EC 1.14.-.-), putative	Métabolisme
BX784344	Ge133B11	na	na	3_XpanSL	4_euSL	Wsv162 (WSSV218), partial (23%)	Non-classé
AL750457	RN02G02	na	na	3_XpanSL	4_euSL	Protein-methionine-S-oxide reductase (EC 1.8.4.6)	Réponse au Stress
BX251044	Ge044E10	na	na	3_XpanSL	4_euSL	1-aminocyclopropane-1-carboxylate oxidase (EC 1.-.-.-)	Réponse au Stress
BX254281	Ge096B03	na	na	3_XpanSL	4_euSL	Ripening-associated protein (Fragment)	Réponse au Stress
BX680225	RS40G02	na	na	3_XpanSL	4_euSL	Peroxidase precursor (EC 1.11.1.7)	Réponse au Stress
BX681402	RS58E06	na	na	3_XpanSL	4_euSL	DnaJ-like protein	Réponse au Stress
CR394081	RN73D11	na	na	3_XpanSL	4_euSL	Syntaxin 71	Traffic intracellulaire
AL751104	RS05B06	na	na	4_XpanSL	10_euSL	DNA polymerase III	Croissance cellulaire, division
CR393209	RN53H11	na	na	4_XpanSL	10_euSL	Heat shock protein STI1	Réponse au Stress
BX678438	RS09B05	na	na	4_XpanSL	10_euSL	Homeobox-leucine zipper protein	Transcription
AL749940	AS02H08	na	na	4_XpanSL	11_euSL	putative protein	_
BX667566	RN09F03	na	na	4_XpanSL	11_euSL	putative protein	_
BX677915	RN46D11	na	na	4_XpanSL	11_euSL	putative protein	_
BX677962	RN47C06	na	na	4_XpanSL	11_euSL	putative protein	_
BX681318	RS57A01	na	na	4_XpanSL	11_euSL	putative protein	_
BX681601	RS61G10	na	na	4_XpanSL	11_euSL	putative protein	_
BX682032	RS68F08	na	na	4_XpanSL	11_euSL	putative protein	_
CR392429	RN31G10	na	na	4_XpanSL	11_euSL	putative protein	_
AL750486	RN03A11	na	na	4_XpanSL	11_euSL	Flavonol 3-O-glucosyltransferase (EC 2.4.1.91)	Métabolisme
BX678822	RS15F01	na	na	4_XpanSL	11_euSL	Acyltransferase homolog	Métabolisme
BX680234	RS40H04	na	na	4_XpanSL	11_euSL	Alcohol oxidase 2	Métabolisme
BX251479	Ge050B02	na	na	4_XpanSL	11_euSL	Unknown	Non-classé
BX255599	Ge007F02	na	na	4_XpanSL	11_euSL	Unknown	Non-classé
BX677636	RN42D06	na	na	4_XpanSL	11_euSL	Unknown	Non-classé
BX680588	RS45E06	na	na	4_XpanSL	11_euSL	Unknown	Non-classé
BX681025	RS52B05	na	na	4_XpanSL	11_euSL	Unknown	Non-classé
BX681917	RS66F12	na	na	4_XpanSL	11_euSL	Unknown	Non-classé
CR393223	RN54B06	na	na	4_XpanSL	11_euSL	Heterochromatin-like protein	Organisation cellulaire
BX677784	RN44F10	na	na	4_XpanSL	11_euSL	Copper chaperone for Superoxide dismutase	Réponse au Stress
BX679992	RS37H01	na	na	4_XpanSL	11_euSL	Related to chitinase 3 precursor protein	Réponse au Stress
CR393726	RN64C02	na	na	4_XpanSL	11_euSL	Chaperone protein dnaJ	Réponse au Stress
BX678737	RS14D11	na	na	4_XpanSL	11_euSL	Myosin V	Traffic intracellulaire
BX677653	RN42G03	na	na	5_XpanSL	18_euSL	putative protein	_
CR392866	RN49C12	na	na	5_XpanSL	18_euSL	Metallothionein-like protein EMB30	Métabolisme
CR393420	RN59H07	na	na	5_XpanSL	18_euSL	Naringenin-chalcone synthase (EC 2.3.1.74)	Métabolisme
BX680139	RS39F01	na	na	5_XpanSL	18_euSL	GPI-anchored protein	Non-classé
CR393797	RN65E06	na	na	5_XpanSL	18_euSL	Tubulin beta chain	Organisation cellulaire
BX679636	RS30C11	na	na	5_XpanSL	18_euSL	Gamma tonoplast intrinsic protein	Transport
BX676949	RN14F10	na	na	5_XpanSL	19_euSL	putative protein	_
BX678519	RS10D05	na	na	5_XpanSL	19_euSL	putative protein	_
AL750160	AS06D09	na	na	5_XpanSL	19_euSL	Metallothionein-like protein	Métabolisme
BX249127	Ge020C11	na	na	5_XpanSL	19_euSL	Metallothionein-like protein EMB30	Métabolisme
BX679131	RS20C02	na	na	5_XpanSL	19_euSL	Leucoanthocyanidin dioxygenase (EC 1.-.-.-) (LDOX)	Métabolisme
BX680432	RS43C07	na	na	5_XpanSL	19_euSL	putative protein	Non-classé
CR392105	RN12D09	na	na	5_XpanSL	19_euSL	Pinus taeda calcium binding protein	Non-classé
AL750591	RN05B11	na	na	5_XpanSL	20_euSL	Xyloglucan:xyloglucosyl transferase (EC 2.4.1.207)	Croissance cellulaire, division
CR393887	RN66G01	na	na	5_XpanSL	20_euSL	MERI-5 protein precursor (Xyloglucan:xyloglucosyl transferase	Croissance cellulaire, division
CR392095	RN11G12	na	na	5_XpanSL	20_euSL	Naringenin-chalcone synthase (EC 2.3.1.74)	Métabolisme
CR354695	Ge125A05	na	na	5_XpanSL	20_euSL	Unknown	Non-classé
BX249989	Ge031B01	na	na	5_XpanSL	20_euSL	Similarity to late embryogenesis abundant protein, partial (31%)	Réponse au Stress
AL750133	AS06B05	na	na	6_XpanSL	15_euSL	putative protein	_
BX680191	RS40C09	na	na	6_XpanSL	15_euSL	putative protein	_
AL751202	RS06C01	na	na	6_XpanSL	15_euSL	P0022F10.3 protein	Non-classé
CR392178	RN27F05	na	na	6_XpanSL	15_euSL	Hypothetical 68.3 kDa protein	Non-classé
CR394061	RN73B08	na	na	6_XpanSL	16_euSL	putative protein	_
BX679850	RS34H11	na	na	6_XpanSL	16_euSL	Calmodulin	Communication/Transduction du signal
CR394212	RN75C02	na	na	6_XpanSL	16_euSL	Pyruvate decarboxylase isozyme 2 (EC 4.1.1.1) (PDC)	Energie
BX249830	Ge029B11	na	na	6_XpanSL	16_euSL	At1g07510/F22G5_9, partial (32%)	Non-classé
BX682853	Ge122G03	na	na	6_XpanSL	16_euSL	F3O9.22 protein.	Non-classé
BX253583	Ge085E07	na	na	6_XpanSL	16_euSL	ENDOGLUCANASE	Organisation cellulaire
CR394442	RN79C01	na	na	6_XpanSL	16_euSL	Disease resistance response protein 206	Réponse au Stress
AL749637	AN05A11	na	na	6_XpanSL	17_euSL	putative protein	_
AL749954	AS03A12	na	na	6_XpanSL	17_euSL	Thimet oligopeptidase (EC 3.4.24.15)	Devenir des protéines
BX667552	RN09D11	na	na	6_XpanSL	17_euSL	Chloroplast nucleoid DNA binding protein	Devenir des protéines
AL749965	AS03B12	na	na	6_XpanSL	17_euSL	Metallothionein-like protein EMB30	Métabolisme
BX252860	Ge071H04	na	na	6_XpanSL	17_euSL	5-methyltetrahydropteroyltriglutamate--homocysteine methyltra	Métabolisme
BX254093	Ge093C03	na	na	6_XpanSL	17_euSL	caffeate O-methyltransferase [Pinus taeda]	Métabolisme
BX682627	Ge118H04	na	na	6_XpanSL	17_euSL	Shikimate 5-dehydrogenase (EC 1.1.1.25)	Métabolisme
CR393695	RN63F04	na	na	6_XpanSL	17_euSL	Unknown	Non-classé
BX255132	Ge111E05	na	na	6_XpanSL	17_euSL	40S ribosomal protein S23	Synthèse de protéines
BX255767	Ge009E08	na	na	6_XpanSL	17_euSL	60S ribosomal protein L11	Synthèse de protéines
BX677658	RN42H03	na	na	6_XpanSL	17_euSL	Nonspecific lipid-transfer protein precursor (LTP)	Transport
BX681070	RS52G08	na	na	6_XpanSL	18_euSL	Trans-cinnamate 4-monooxygenase (EC 1.14.13.11)	Métabolisme
BX252848	Ge071F08	na	na	7_XpanSL	13_euSL	putative protein	_
BX679160	RS20G09	na	na	7_XpanSL	13_euSL	putative protein	_
BX681153	RS54D06	na	na	7_XpanSL	13_euSL	putative protein	_
BX682761	Ge121C11	na	na	7_XpanSL	13_euSL	putative protein	_
BX255048	Ge109H02	na	na	7_XpanSL	13_euSL	Unknown	Non-classé
BX251675	Ge053G10	na	na	7_XpanSL	13_euSL	Poly(A)-binding protein, partial (84%)	Synthèse de protéines
BX252363	Ge063D09	na	na	7_XpanSL	14_euSL	Unknown	Non-classé
CR392200	RN28B03	na	na	7_XpanSL	14_euSL	Defensin	Réponse au Stress
AL750626	RN05E11	na	na	7_XpanSL	15_euSL	putative protein	_
BX679276	RS22D04	na	na	7_XpanSL	15_euSL	putative protein	_
BX682328	Ge115B05	na	na	7_XpanSL	15_euSL	putative protein	_
CR394232	RN75E03	na	na	7_XpanSL	15_euSL	putative protein	_
BX681997	RS67H05	na	na	7_XpanSL	15_euSL	Beta-glucosidase, chloroplast precursor (EC 3.2.1.21)	Métabolisme
BX677400	RN38D02	na	na	na	10_euSL	putative protein	_
BX681976	RS67E11	na	na	na	10_euSL	Histidine decarboxylase (EC 4.1.1.22)	Métabolisme
BX677917	RN46E01	na	na	na	11_euSL	putative protein	_
BX679783	RS33H04	na	na	na	11_euSL	putative protein	_
BX680131	RS39E04	na	na	na	11_euSL	putative protein	_
BX251837	Ge056B06	na	na	na	11_euSL	Protein transport protein SEC22	Traffic intracellulaire
CR393546	RN61E12	na	na	na	12_euSL	putative protein	_
BX681954	RS67C08	na	na	na	13_euSL	37 kDa inner envelope membrane protein	Non-classé
BX254079	Ge093A05	na	na	na	14_euSL	Inositol oxygenase (Myo-inositol oxygenase) (Aldehyde reducta	Métabolisme
BX680855	RS49G03	na	na	na	14_euSL	Alcohol dehydrogenase (EC 1.1.1.1)	Métabolisme
CR393324	RN58E05	na	na	na	14_euSL	Epoxide hydrolase (EC 3.3.2.3)	Réponse au Stress
CR392243	RN28G03	na	na	na	16_euSL	putative protein	_
AW784085	NXNV_117_F2	na	na	na	18_euSL	glyceraldehyde-3-P-dehydrogenase	Métabolisme
AW985459	NXNV_136_H4	na	na	na	18_euSL	Aspartate transaminase	Métabolisme
BF220737	NXCI_150_A7	na	na	na	18_euSL	Arginine decarboxylase	Métabolisme
BG039035	NXSI_24_H1	na	na	na	18_euSL	Cellulose synthase	Métabolisme
BG040863	NXSI_116_A9	na	na	na	18_euSL	adenosine kinase (AK)	Métabolisme
BX681865	RS65H11	na	na	na	18_euSL	Dihydropinosylvin synthase (EC 2.3.1.-)	Métabolisme
BX680587	RS45E04	na	na	na	18_euSL	Unknown	Non-classé
BX679178	RS21A09	na	na	na	18_euSL	Zinc finger protein 216	Transcription
AL751302	RS07C09	na	na	na	20_euSL	Metallothionein-like protein	Métabolisme

Figure Ax-1 : Détail des profils d'expression de toutes les accessions composant les clusters générés par Expander sur les données de l'expérience stress court. (1 = France-ControlC, 2=France-Stress6H, 3=France-Stress48H, 4 = Maroc-ControlC, 5= Maroc-Stress6H, 6=Maroc-Stress48H)

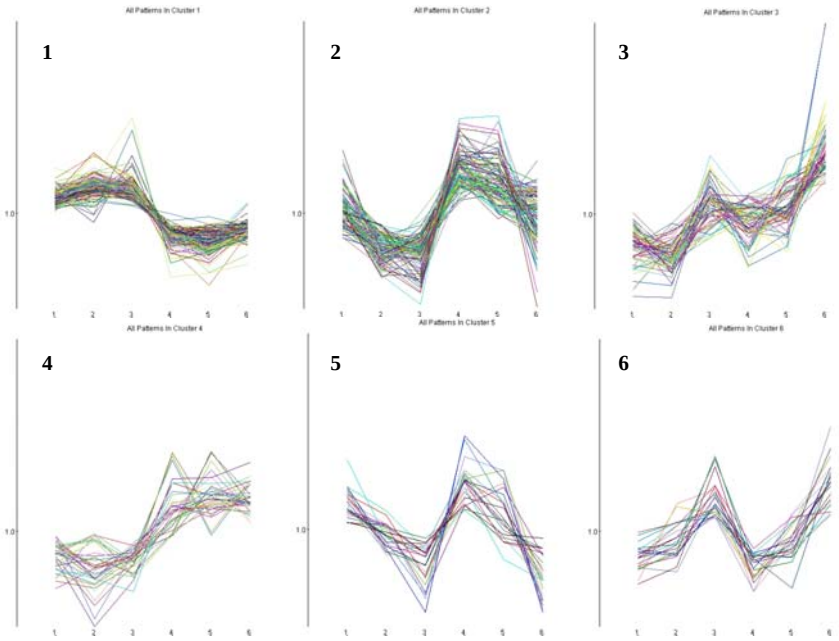


Figure Ax-2 : Détail des profils d'expression de toutes les accessions composant les clusters générés par Expander sur les données de l'expérience stress long. (1 = France-ControlL, 2=France-StressL, 3 = Maroc-ControlL, 4= Maroc-StressL)

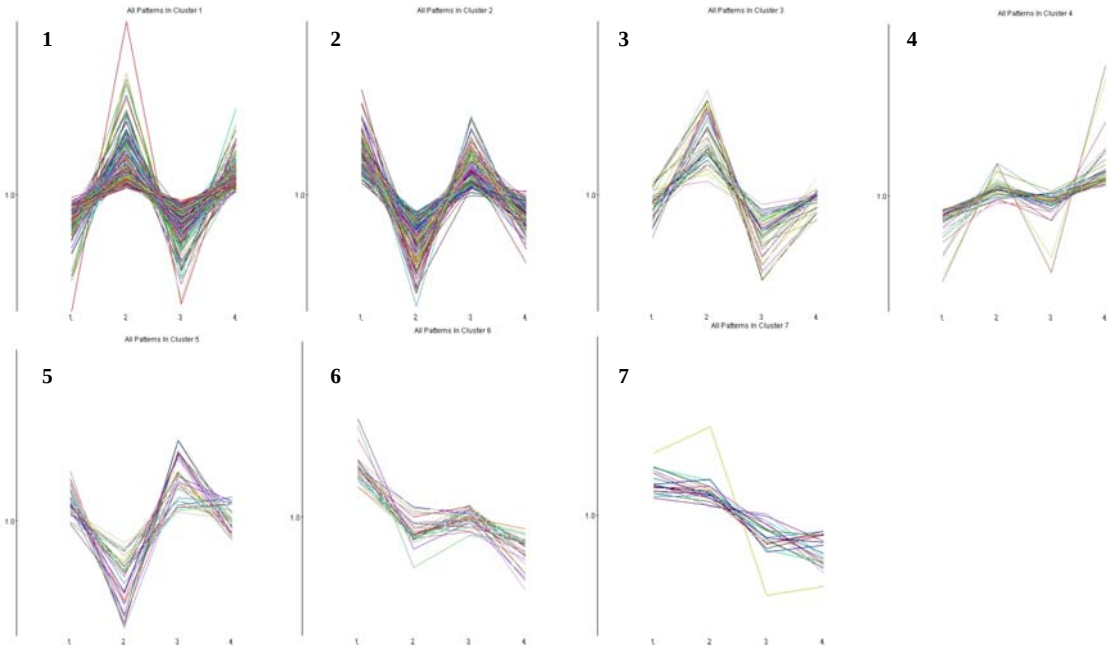


Tableau Ax-10 : Liste des 215 gènes des clusters présentant un profil d'intérêt. (en vert gènes testés en PCR quantitative.

Affectation groupe SL		Affectation groupe SC					
Expander	SOTArray	Expander	SOTArray	EST ID	nb EST	Accession	Annotation
3_XpanSL					**40**		
	1_sotaSL				4		
		na			**4**		
			na		4		
				RN05F08	1	AL750636	Antimicrobial peptide
				RS23E08	1	BX679343	Chaperone protein dnaJ
				RS56A08	1	BX681246	putative protein
				RS66H05	1	BX681928	Antimicrobial peptide
	12_sotaSL				16		
		na			**16**		
			na		16		
				AS06G11	1	AL750195	Hypothetical protein DR0254
				Ge026H10	1	BX249673	putative protein
				Ge058A03	1	BX251976	Unknown
				Ge111D09	1	BX255127	eukaryotic initiation factor 3E subunit {Arabidopsis thaliana;} , partial (96%)
				Ge127F05	1	BX683057	Aluminum-induced protein-like protein, partial (65%)
				RN17G03	1	BX677108	Expansin-like precursor
				RN47F03	1	BX677977	Beta-fructofuranosidase (EC 3.2.1.26), cell wall isozyme precursor
				RN48B11	1	CR392796	putative protein
				RN73A01	1	CR394054	putative protein
				RN73B01	1	CR394058	Unknown
				RS08E10	1	BX678396	Alternative oxidase 1, mitochondrial precursor
				RS16C11	1	BX678873	putative protein
				RS19C01	1	BX679063	putative protein
				RS26C01	1	BX679459	PGPD14 protein (Pollen germination related protein)
				RS55B01	1	BX681189	putative protein
				RS65C05	1	BX681814	putative protein
	13_sotaSL				1		
		3_XpanSC			**1**		
			4_sotaSC		1		
				Ge110B02	1	BX255062	Ubiquitin-conjugating enzyme E2 (EC 6.3.2.19)
	2_sotaSL				1		
		na			**1**		
			na		1		
				RN50C09	1	CR394499	Antimicrobial peptide
	3_sotaSL				5		
		na			**5**		
			na		5		
				Ge088A07	1	BX253756	Aluminum-induced protein-like protein, partial (67%)
				RN73D08	1	CR394078	Licheninase precursor (EC 3.2.1.73)
				RS06C11	1	AL751212	Chaperone protein dnaJ
				RS15G11	1	BX678841	4-hydroxyphenylpyruvate dioxygenase (EC 1.13.11.27)
				RS40E02	1	BX680204	4-hydroxyphenylpyruvate dioxygenase (EC 1.13.11.27)
	4_sotaSL				13		
		na			**13**		
			na		13		
				Ge029H02	1	BX249884	Phosphoenolpyruvate carboxykinase (EC 4.1.1.49)
				Ge044E10	1	BX251044	1-aminocyclopropane-1-carboxylate oxidase (EC 1.-.-.-)
				Ge076B04	1	BX252988	Thioredoxin H-type (TRX-H)
				Ge096B03	1	BX254281	Ripening-associated protein (Fragment)
				Ge106G10	1	BX254884	Phosphoenolpyruvate carboxykinase (EC 4.1.1.49)
				Ge133B11	1	BX784344	Wsv162 (WSSV218), partial (23%)
				RN02G02	1	AL750457	Protein-methionine-S-oxide reductase (EC 1.8.4.6)
				RN73D11	1	CR394081	Syntaxin 71
				RS18E12	1	BX679023	Signal peptidase I (EC 3.4.21.89)
				RS22B08	1	BX679259	putative protein
				RS31C02	1	BX679683	Cytochrome P450 (EC 1.14.-.-), putative
				RS40G02	1	BX680225	Peroxidase precursor (EC 1.11.1.7)
				RS58E06	1	BX681402	DnaJ-like protein
4_XpanSL					**31**		
	10_sotaSL				7		
		2_XpanSC			**1**		
			6_sotaSC		1		
				RN49H03	1	CR392910	Germin-like protein
		3_XpanSC			**3**		
			3_sotaSC		1		
				RN76H07	1	CR394334	putative protein
			4_sotaSC		2		
				Ge085G01	1	BX253595	Zinc finger protein HIT-10, partial (3%)
				RN56A06	1	BX678079	putative protein
		na			**3**		
			na		3		
				RN53H11	1	CR393209	Heat shock protein STI1
				RS05B06	1	AL751104	DNA polymerase III
				RS09B05	1	BX678438	Homeobox-leucine zipper protein
	11_sotaSL				22		
		na			**22**		
			na		22		
				AS02H08	1	AL749940	putative protein
				Ge007F02	1	BX255599	Unknown
				Ge050B02	1	BX251479	Unknown
				RN03A11	1	AL750486	Flavonol 3-O-glucosyltransferase (EC 2.4.1.91)
				RN09F03	1	BX667566	putative protein
				RN31G10	1	CR392429	putative protein
				RN42D06	1	BX677636	Unknown
				RN44F10	1	BX677784	Copper chaperone for Superoxide dismutase
				RN46D11	1	BX677915	putative protein
				RN47C06	1	BX677962	putative protein
				RN54B06	1	CR393223	Heterochromatin-like protein
				RN64C02	1	CR393726	Chaperone protein dnaJ
				RS14D11	1	BX678737	Myosin V
				RS15F01	1	BX678822	Acyltransferase homolog
				RS37H01	1	BX679992	Related to chitinase 3 precursor protein
				RS40H04	1	BX680234	Alcohol oxidase 2
				RS45E06	1	BX680588	Unknown
				RS52B05	1	BX681025	Unknown
				RS57A01	1	BX681318	putative protein
				RS61G10	1	BX681601	putative protein
				RS66F12	1	BX681917	Unknown
				RS68F08	1	BX682032	putative protein
	6_sotaSL				2		
		3_XpanSC			**2**		
			3_sotaSC		2		
				Ge041B11	1	BX250781	Low molecular weight heat-shock protein, partial (98%)
				HSP	1	BX000656	HSP
5_XpanSL					**29**		
	18_sotaSL				10		
		2_XpanSC			**4**		
			6_sotaSC		3		
				AS05C03	1	AL750051	Hypothetical 49.0 kDa protein
				RS06C03	1	AL751204	T29M8.12 protein
				RS66A11	1	BX681871	putative protein
			7_sotaSC		1		
				Ge068G01	1	BX252660	Histone H2A
		na			**6**		
			na		6		
				RN42G03	1	BX677653	putative protein
				RN49C12	1	CR392866	Metallothionein-like protein EMB30
				RN59H07	1	CR393420	Naringenin-chalcone synthase (EC 2.3.1.74)
				RN65E06	1	CR393797	Tubulin beta chain
				RS30C11	1	BX679636	Gamma tonoplast intrinsic protein
				RS39F01	1	BX680139	GPI-anchored protein
	19_sotaSL				11		
		2_XpanSC			**4**		
			6_sotaSC		1		
				RS43C11	1	BX680425	putative protein
			7_sotaSC		3		
				RN28G12	1	non soumis	putative protein
				RN54D04	1	CR393241	Plasma membrane intrinsic polypeptide
				RN59F04	1	CR393398	putative protein

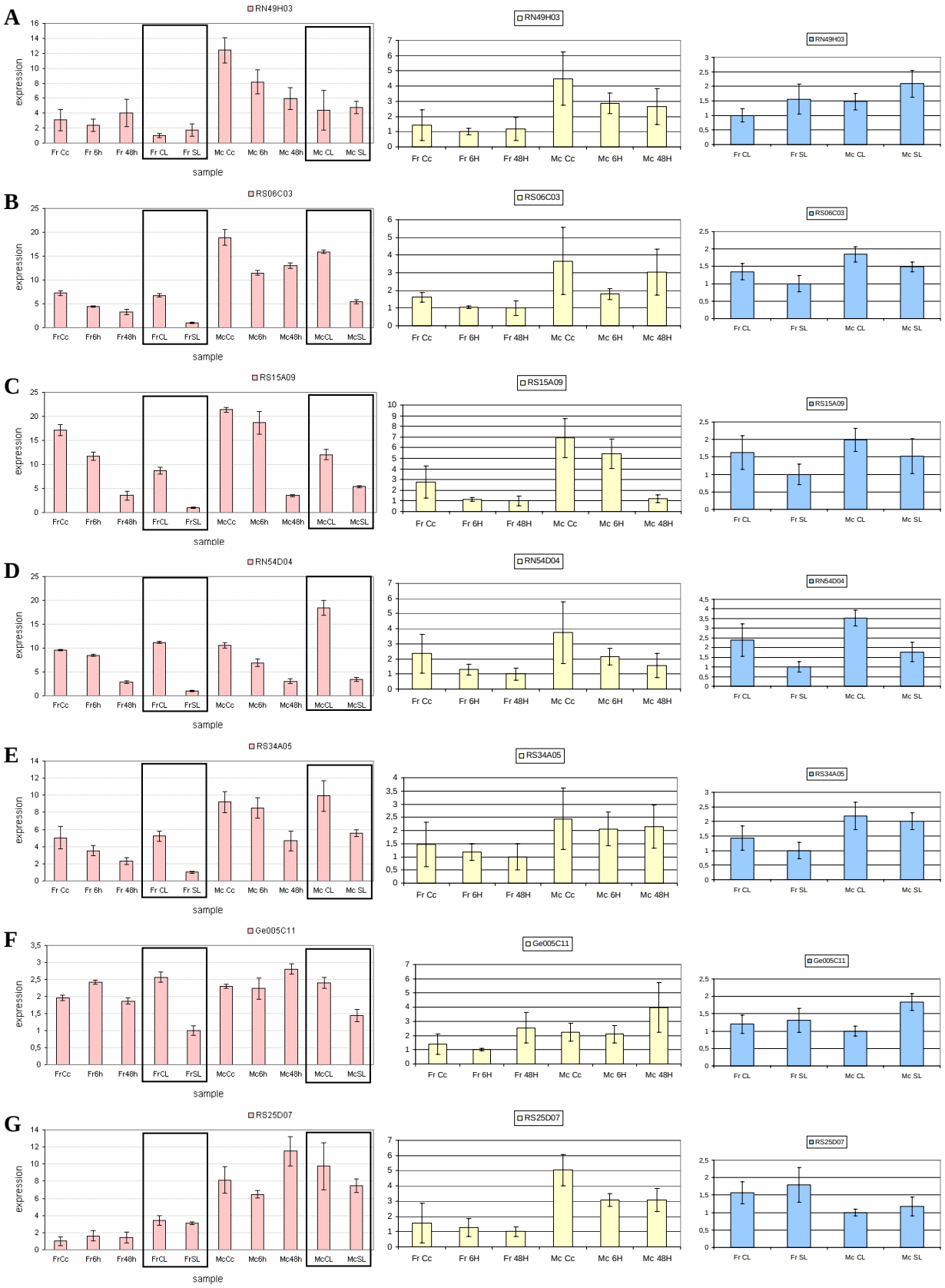
		na			**7**	
			na		7	
				AS06D09	1 AL750160	Metallothionein-like protein
				Ge020C11	1 BX249127	Metallothionein-like protein EMB30
				RN12D09	1 CR392105	Pinus taeda calcium binding protein
				RN14F10	1 BX676949	putative protein
				RS10D05	1 BX678519	putative protein
				RS20C02	1 BX679131	Leucoanthocyanidin dioxygenase (EC 1.-.-.-) (LDOX)
				RS43C07	1 BX680432	Pinus taeda clone PtaAGP5
	20_sotaSL				8	
		2_XpanSC			**3**	
			6_sotaSC		1	
				RN32A07	1 non soumis	putative protein
			7_sotaSC		2	
				RS07C01	1 AL751295	Proline-rich protein / Cortical cell delineating protein precursor
				RS15A09	1 BX678783	Peroxidase precursor (EC 1.11.1.7)
		na			**5**	
			na		5	
				Ge031B01	1 BX249989	Similarity to late embryogenesis abundant protein, partial (31%)
				Ge125A05	1 CR354695	Unknown
				RN05B11	1 AL750591	Xyloglucan:xyloglucosyl transferase (EC 2.4.1.207)
				RN11G12	1 CR392095	Naringenin-chalcone synthase (EC 2.3.1.74)
				RN66G01	1 CR393887	MERI-5 protein precursor (Xyloglucan:xyloglucosyl transferase (EC 2.4.1.207
6_XpanSL					**27**	
	15_sotaSL				5	
		2_XpanSC			**1**	
			7_sotaSC		1	
				RN35A11	1 CR392586	putative protein
		na			**4**	
			na		4	
				AS06B05	1 AL750133	putative protein
				RN27F05	1 CR392178	Hypothetical 68.3 kDa protein
				RS06C01	1 AL751202	P0022F10.3 protein
				RS40C09	1 BX680191	putative protein
	16_sotaSL				8	
		2_XpanSC			**1**	
			7_sotaSC		1	
				RS54B08	1 BX681136	Dehydration stress-induced protein
		na			**7**	
			na		7	
				Ge029B11	1 BX249830	At1g07510/F22G5_9, partial (32%)
				Ge085E07	1 BX253583	Endoglucanase
				Ge122G03	1 BX682853	F3O9.22 protein.
				RN73B08	1 CR394061	putative protein
				RN75C02	1 CR394212	Pyruvate decarboxylase isozyme 2 (EC 4.1.1.1) (PDC)
				RN79C01	1 CR394442	Disease resistance response protein 206
				RS34H11	1 BX679850	Calmodulin
	17_sotaSL				13	
		2_XpanSC			**2**	
			6_sotaSC		2	
				Ge050H05	1 BX251535	60S ribosomal protein L39
				RS34F05	1 BX679837	L3 cytoplasmic ribosomal protein
		na			**11**	
			na		11	
				AN05A11	1 AL749637	putative protein
				AS03A12	1 AL749954	Thimet oligopeptidase (EC 3.4.24.15)
				AS03B12	1 AL749965	Metallothionein-like protein EMB30
				Ge009E09	1 BX255767	60S ribosomal protein L11
				Ge071H04	1 BX252860	5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase
				Ge093C03	1 BX254093	caffeate O-methyltransferase [Pinus taeda]
				Ge111E05	1 BX255132	40S ribosomal protein S23
				Ge118H04	1 BX682627	Shikimate 5-dehydrogenase (EC 1.1.1.25)
				RN09D11	1 BX667552	Chloroplast nucleoid DNA binding protein
				RN42H03	1 BX677658	Nonspecific lipid-transfer protein precursor (LTP)
				RN63F04	1 CR393695	Unknown
	18_sotaSL				1	
		na			**1**	
			na		1	
				RS52G08	1 BX681070	Trans-cinnamate 4-monooxygenase (EC 1.14.13.11)
na					**88**	
	10_sotaSL				2	
		2_XpanSC			**1**	
			6_sotaSC		1	
				Ge094A09	1 BX254145	PR34, partial (4%)
		3_XpanSC			**1**	
			5_sotaSC		1	
				RN61A02	1 CR393500	putative protein
	11_sotaSL				1	
		3_XpanSC			**1**	
			4_sotaSC		1	
				Ge005C11	1 BX255395	putative protein
	13_sotaSL				1	
		2_XpanSC			**1**	
			6_sotaSC		1	
				RS25D07	1 BX679434	putative protein
	14_sotaSL				2	
		2_XpanSC			**2**	
			7_sotaSC		2	
				RN68A05	1 BX678224	Late embryogenesis abundant protein
				RS67A12	1 BX681938	Protein kinase (EC 2.7.1.-)
	20_sotaSL				1	
		2_XpanSC			**1**	
			6_sotaSC		1	
				RS34A05	1 BX679792	PrMC3
	na				81	
		2_XpanSC			**44**	
			6_sotaSC		30	
				Ge021A08	1 BX249189	putative protein
				Ge025C11	1 BX249546	Calmodulin
				Ge029C02	1 BX249833	Ras-like small monomeric GTP-binding protein, partial (65%)
				Ge029G11	1 BX249881	ATP synthase gamma chain
				Ge042C04	1 BX250861	OVCA2=CANDIDATE tumor suppressor, partial (9%)
				Ge045C02	1 BX251094	Abscisic stress ripening protein
				Ge063E06	1 BX252371	60S ribosomal protein L31
				Ge066E08	1 BX252563	Aspartate aminotransferase, chloroplast precursor
				Ge066E10	1 BX252565	putative protein
				Ge067H06	1 BX252614	Histone H2A variant
				Ge068F09	1 BX252657	60S ribosomal protein L35A
				Ge069F09	1 BX252717	B2 protein
				Ge095G04	1 BX254258	LUMINAL BINDING PROTEIN
				Ge106E11	1 BX254875	Probable aquaporin, complete
				Ge106H02	1 BX254888	Ras-related protein Rab11B
				Ge113B09	1 BX255226	TrPRP2, partial (67%)
				Ge115A12	1 BX682324	Abscisic stress ripening protein
				Ge116F08	1 BX682440	Nuclear transport factor 2 (NTF-2).
				Ge119A01	1 BX682634	Elongation factor 1-gamma (EF-1-gamma)
				Ge121G09	1 BX682787	60S ribosomal protein L31
				Ge123D11	1 BX682915	14-3-3 protein
				Ge126D02	1 BX682978	putative protein
				RN03H10	1 AL750567	60S ribosomal protein L14
				RN33D03	1 CR392504	putative protein
				RN40E01	1 BX677507	Nuclear scaffold-like protein p76
				RS16F01	1 BX678890	Cholinephosphate cytidylyltransferase B (EC 2.7.7.15)
				RS29G03	1 BX679616	Histone H4
				RS52G04	1 BX681067	Embryogenesis-associated protein EMB7
				RS52G06	1 BX681068	Cysteine proteinase RD21A precursor (EC 3.4.22.-)
				RS55E07	1 BX681222	putative protein
			7_sotaSC		14	
				RN03E03	1 AL750525	MtN21 nodulin protein-like
				RN05E06	1 AL750621	Pectate lyase (EC 4.2.2.2)
				RN30A11	1 CR392259	Acid phosphatase precursor 1 (EC 3.1.3.2)
				RN35A05	1 CR392581	Isopentenyl-diphosphate delta-isomerase II (EC 5.3.3.2)
				RN35E03	1 CR392614	Acid phosphatase
				RN47C10	1 BX677964	Late embryogenesis abundant protein

			RN53C07	1	CR393156	Peroxidase precursor (EC 1.11.1.7)
			RN65E04	1	CR393796	Pinene synthase (EC 4.2.3.14)
			RN69C07	1	BX678294	Nuclear transport factor 2 (NTF-2)
			RS10A02	1	BX678491	Pathogenesis-related protein 5 precursor (PR-5)
			RS11G02	1	BX678596	Peroxidase precursor (EC 1.11.1.7)
			RS18B02	1	BX678993	IN2-2 protein
			RS51C03	1	BX680952	putative protein
			RS56D01	1	BX681270	putative protein
	3_XpanSC			**37**		
		3_sotaSC		3		
			AS05H10	1	AL750116	EF-HAND containing protein-like
			RN18B10	1	BX677134	putative protein
			RS43C02	1	BX680428	putative protein
		4_sotaSC		17		
			AS06E02	1	AL750163	putative protein
			AS07A11	1	AL750218	Chitinase (EC 3.2.1.14)
			Ge009E07	1	BX255766	metallothionein
			Ge027B05	1	BX249687	FKBP-like protein (Arabidopsis thaliana genomic DNA, chromosome 3, P1 cl
			Ge032H07	1	BX250140	Defender against cell death
			Ge071F04	1	BX252844	metallothionein-like protein
			Ge074G03	1	BX252908	Cathepsin B-like cysteine proteinase, partial (68%)
			Ge075E07	1	BX252953	RING finger protein 26
			Ge077B10	1	BX253047	Aminotransferase
			Ge080C01	1	BX253244	Lin1294 protein, partial (9%)
			Ge088F01	1	BX253795	Unknown
			Ge095F04	1	BX254251	cytosolic aldehyde dehydrogenase {Oryza sativa (japonica cultivar-group);} ,
			RN09C02	1	BX667535	Carbon catabolite derepressing protein kinase
			RN75C03	1	CR394213	Auxin-repressed 12.5 kDa protein
			RS03D12	1	AL751045	putative protein
			RS53A02	1	BX681080	Acetylornithine deacetylase (EC 3.5.1.16)
			RS63F03	1	BX681731	Anther specific LAT52 protein precursor
		5_sotaSC		17		
			AN01H05	1	AL749580	putative protein
			AS05F09	1	AL750092	Ubiquitin-conjugating enzyme E2-17 kDa (EC 6.3.2.19)
			Ge005B04	1	BX255377	Elastin precursor - human (fragment) {Homo sapiens;} , partial (50%)
			Ge034E07	1	BX250267	Ubiquitin-conjugating enzyme E2-17 kDa (EC 6.3.2.19)
			Ge060F11	1	BX252163	Unknown
			Ge096C10	1	BX254293	Peptide methionine sulfoxide reductase (EC 1.8.4.6)
			Ge098G01	1	BX254449	Phospholipase D alpha 1 (PLD alpha 1) (Choline phosphatase 1) (Phosphat
			Ge100G06	1	BX254586	putative protein
			Ge104G10	1	BX254794	putative protein
			Ge116H08	1	BX682459	Ubiquitin-conjugating enzyme (EC 6.3.2.19)
			NXNV_148_H6	1	BE123499	Cellulose synthase
			RN03C09	1	AL750508	Histone H4
			RN76B02	1	CR394276	putative protein
			RS12A03	1	BX678613	Gamma-thionin homolog precursor
			RS18B05	1	BX678995	putative protein
			RS38B06	1	BX680013	Gb|AAD26879.1
			RS46H02	1	BX680660	putative protein
Total				**215**		

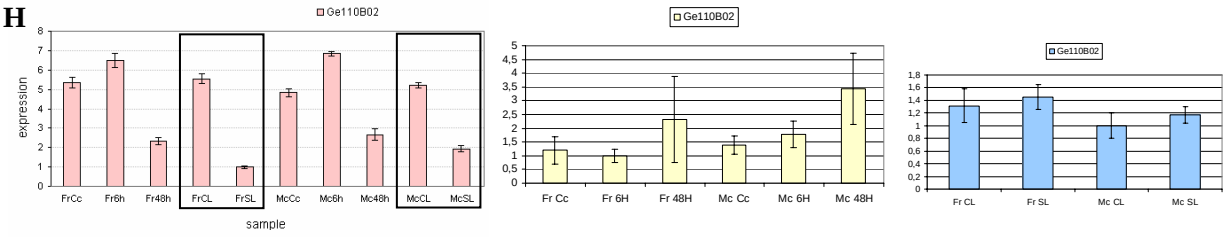
Tableau Ax-11 : Liste des singletons du clustering Expander communs aux deux expériences.

Affectation groupe SL		Affectation groupe SC					
Expander	SOTArray	Expander	SOTArray	EST ID	nb EST	Accession	Annotation
na							
	10_sotaSL				2		
		na			**2**		
			na		2		
				RN38D02	1	BX677400	putative protein
				RS67E11	1	BX681976	Histidine decarboxylase (EC 4.1.1.22)
	11_sotaSL				4		
		na			**4**		
			na		4		
				Ge056B06	1	BX251837	Protein transport protein SEC22
				RN46E01	1	BX677917	putative protein
				RS33H04	1	BX679783	putative protein
				RS39E04	1	BX680131	putative protein
	12_sotaSL				1		
		na			**1**		
			na		1		
				RN61E12	1	CR393546	putative protein
	13_sotaSL				1		
		na			**1**		
			na		1		
				RS67C08	1	BX681954	37 kDa inner envelope membrane protein
	14_sotaSL				3		
		na			**3**		
			na		3		
				Ge093A05	1	BX254079	Inositol oxygenase (Myo-inositol oxygenase) (Aldehyde reductase-like 6) , p
				RN58E05	1	CR393324	Epoxide hydrolase (EC 3.3.2.3)
				RS49G03	1	BX680855	Alcohol dehydrogenase (EC 1.1.1.1)
	16_sotaSL				1		
		na			**1**		
			na		1		
				RN28G03	1	CR392243	putative protein
	18_sotaSL				11		
		na			**11**		
			6_sotaSC		3		
				NXCI_50_B7	1	BE643924	S-adenosyl-Met synthase
				NXNV_125_E12	1	AW870208	Aspartate transaminase
				NXSI_104_E11	1	BG039825	Ornitine transaminase
			na		8		
				NXCI_150_A7	1	BF220737	Arginine decarboxylase
				NXNV_117_F2	1	AW784085	glyceraldehyde-3-P-dehydrogenase
				NXNV_136_H4	1	AW985459	Aspartate transaminase
				NXSI_116_A9	1	BG040863	adenosine kinase (AK)
				NXSI_24_H1	1	BG039035	Cellulose synthase
				RS21A09	1	BX679178	Zinc finger protein 216
				RS45E04	1	BX680587	Unknown
				RS65H11	1	BX681865	Dihydropinosylvin synthase (EC 2.3.1.-)
	20_sotaSL				1		
		na			**1**		
			na		1		
				RS07C09	1	AL751302	Metallothionein-like protein
	na				17		
		na			**17**		
			1_sotaSC		7		
				RN09A04	1	BX666044	putative protein
				RN38C03	1	BX677392	putative protein
				RN42G05	1	BX677655	Flavonoid 3',5'-hydroxylase (EC 1.14.-.-)
				RS05A04	1	AL751090	Protein kinase (EC 2.7.1.-)
				RS05F11	1	AL751155	putative protein
				RS55D09	1	BX681214	putative protein
				RS57H02	1	BX681370	Unknown
			2_sotaSC		3		
				Ge075H01	1	BX252971	putative protein
				Ha_PP758	1	non soumis	_
				RS50E05	1	BX680905	Fumarate hydratase, mitochondrial precursor (EC 4.2.1.2)
			5_sotaSC		2		
				Ge053H04	1	BX251681	Germin(Oxalate oxidase)-like protein, partial (24%)
				NXSI_48_D6	1	BF609706	malate dehydrogenase
			6_sotaSC		5		
				NXCI_106_D10	1	BF049732	Fumarase
				NXCI_75_C7	1	BE761959	Glutamate synthase
				NXSI_99_F10	1	BG039488	S-adenosyl-Met decarboxylase
				RS36C01	1	BX679916	Metallophosphatase
				RS64D09	1	BX681771	60S ribosomal protein L34
Total					**41**		

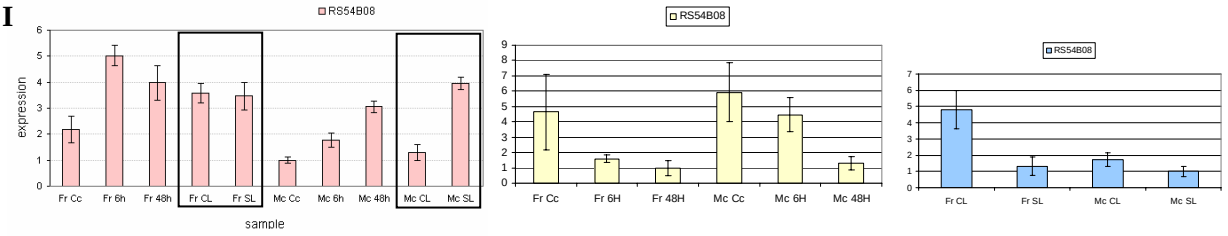
Figures Ax-3 : Profils d'expressions des gènes différentiellement exprimés en stress court et en stress long. Données d'expression de PCR quantitative (à gauche). profils d'expression correspondants obtenus par les données d'hybridation pour l'expérience stress court (au milieu) et stress long (à droite).



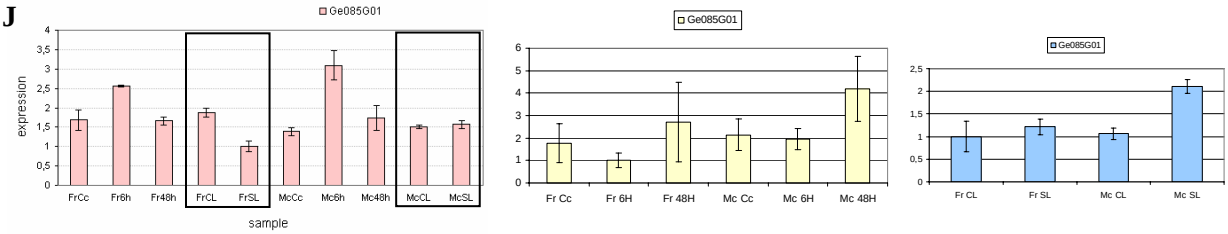
H


Ge110B02
expression
FrCc Fr6h Fr48h FrCL FrSL McCc Mc6h Mc48h McCL McSL
sample
Ge110B02
Fr Cc Fr 6H Fr 48H Mc Cc Mc 6H Mc 48H
Ge110B02
Fr CL Fr SL Mc CL Mc SL

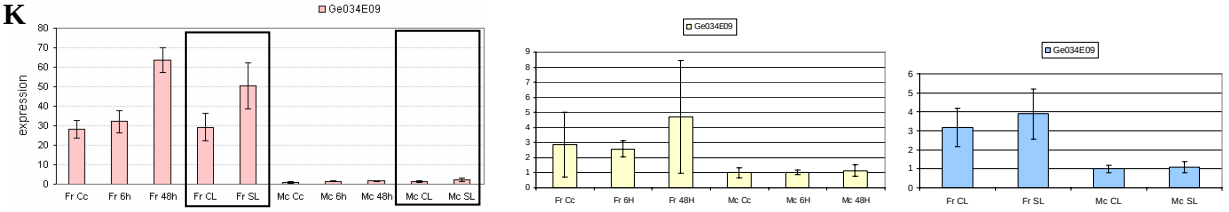
I


RS54B08
expression
Fr Cc Fr 6h Fr 48h Fr CL Fr SL Mc Cc Mc 6h Mc 48h Mc CL Mc SL
sample
RS54B08
Fr Cc Fr 6H Fr 48H Mc Cc Mc 6H Mc 48H
RS54B08
Fr CL Fr SL Mc CL Mc SL

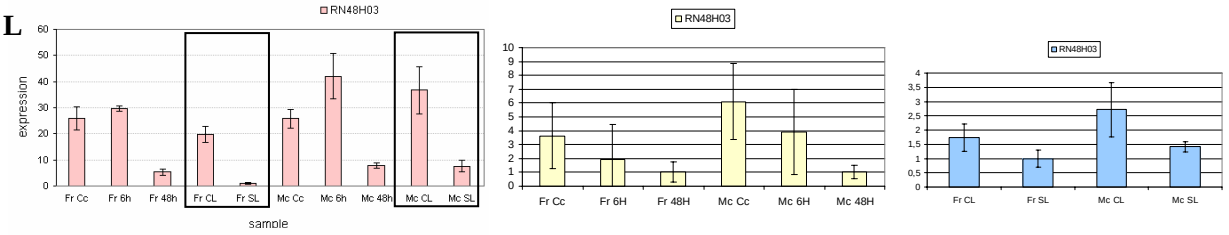
J


Ge085G01
expression
FrCc Fr6h Fr48h FrCL FrSL McCc Mc6h Mc48h McCL McSL
sample
Ge085G01
Fr Cc Fr 6H Fr 48H Mc Cc Mc 6H Mc 48H
Ge085G01
Fr CL Fr SL Mc CL Mc SL

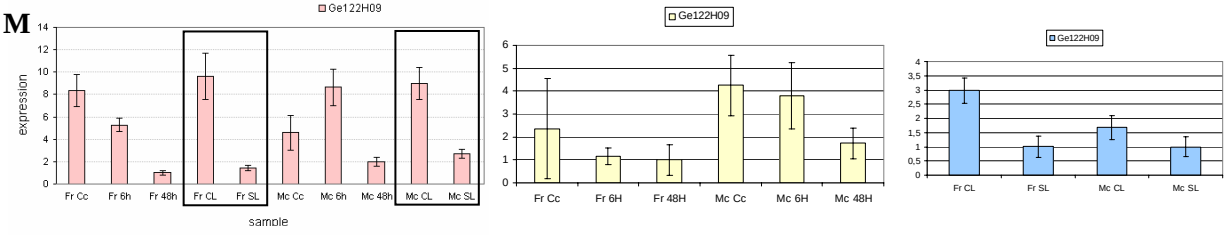
K


Ge034E09
expression
Fr Cc Fr 6h Fr 48h Fr CL Fr SL Mc Cc Mc 6h Mc 48h Mc CL Mc SL
Ge034E09
Fr Cc Fr 6H Fr 48H Mc Cc Mc 6H Mc 48H
Ge034E09
Fr CL Fr SL Mc CL Mc SL

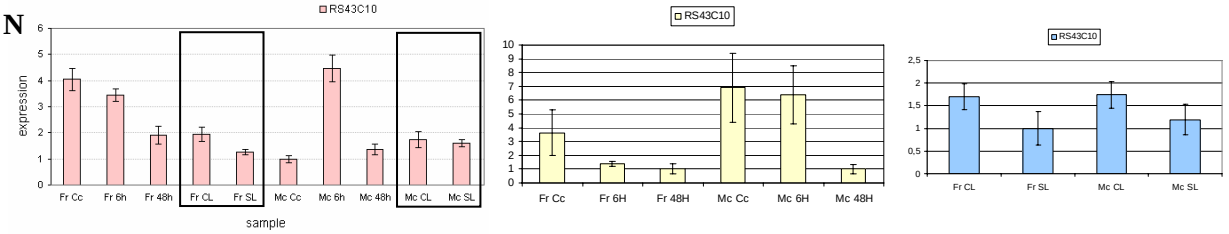
L


RN48H03
expression
Fr Cc Fr 6h Fr 48h Fr CL Fr SL Mc Cc Mc 6h Mc 48h Mc CL Mc SL
sample
RN48H03
Fr Cc Fr 6H Fr 48H Mc Cc Mc 6H Mc 48H
RN48H03
Fr CL Fr SL Mc CL Mc SL

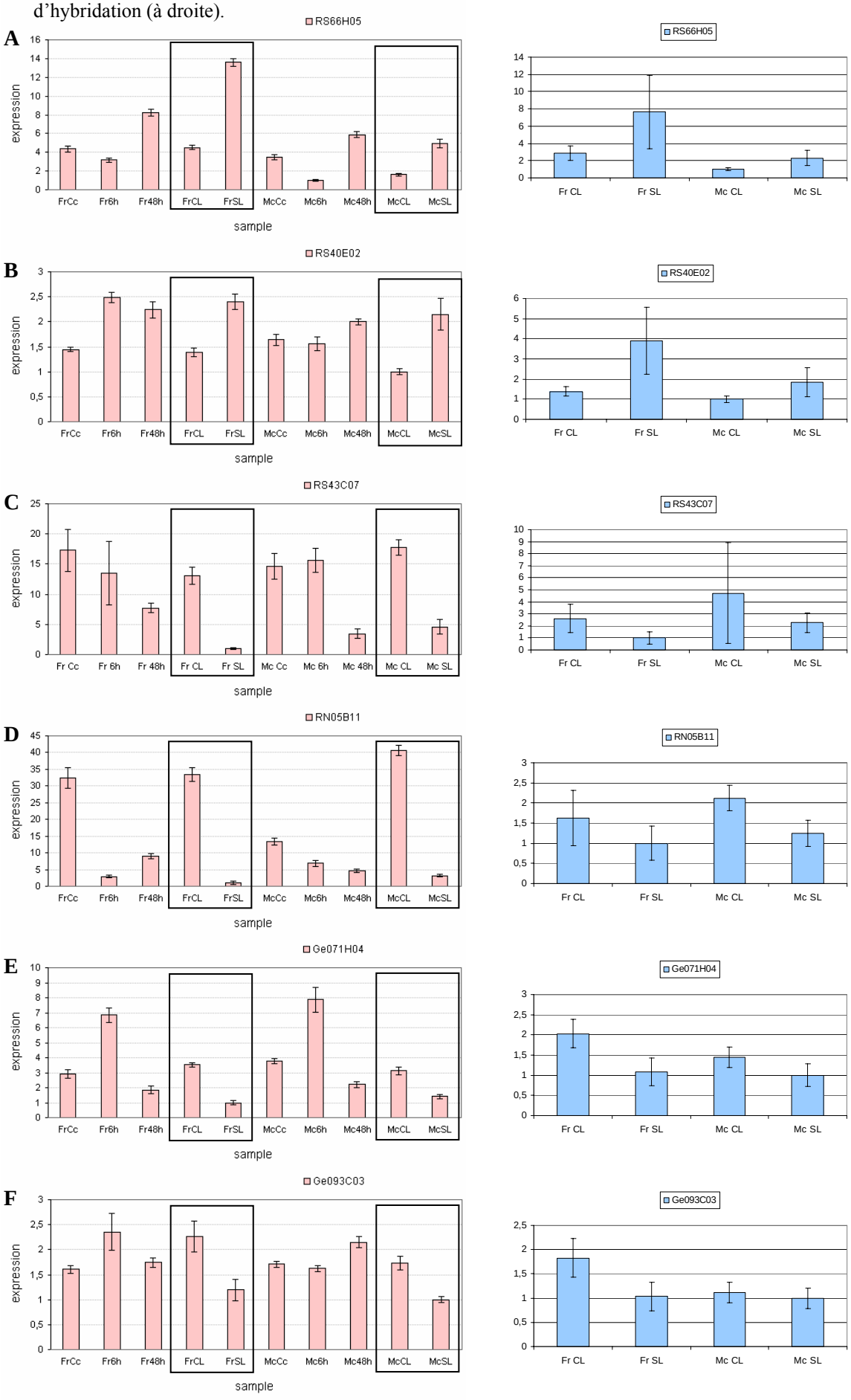
M


Ge122H09
expression
Fr Cc Fr 6h Fr 48h Fr CL Fr SL Mc Cc Mc 6h Mc 48h Mc CL Mc SL
sample
Ge122H09
Fr Cc Fr 6H Fr 48H Mc Cc Mc 6H Mc 48H
Ge122H09
Fr CL Fr SL Mc CL Mc SL

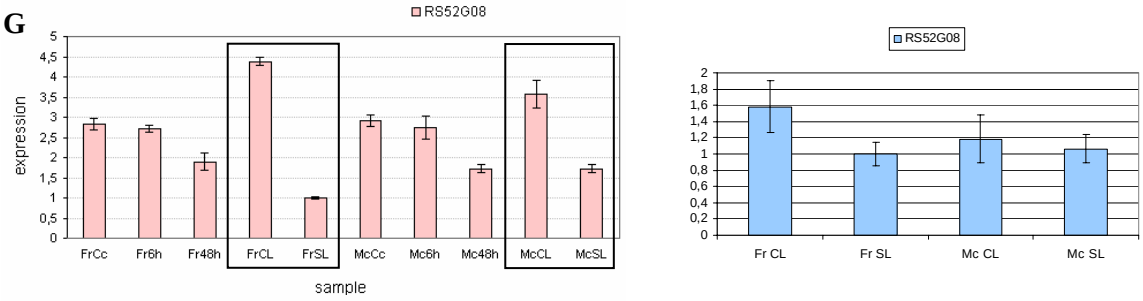
N


RS43C10
expression
Fr Cc Fr 6h Fr 48h Fr CL Fr SL Mc Cc Mc 6h Mc 48h Mc CL Mc SL
sample
RS43C10
Fr Cc Fr 6H Fr 48H Mc Cc Mc 6H Mc 48H
RS43C10
Fr CL Fr SL Mc CL Mc SL

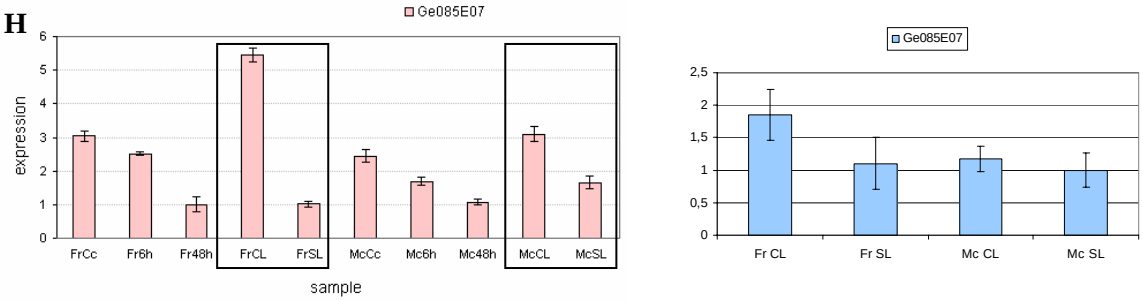
Figures Ax-4 : Profils d'expressions des gènes différentiellement exprimés et regroupés par Expander en « stress long » uniquement. Données d'expression de PCR quantitative (à gauche), d'hybridation (à droite).



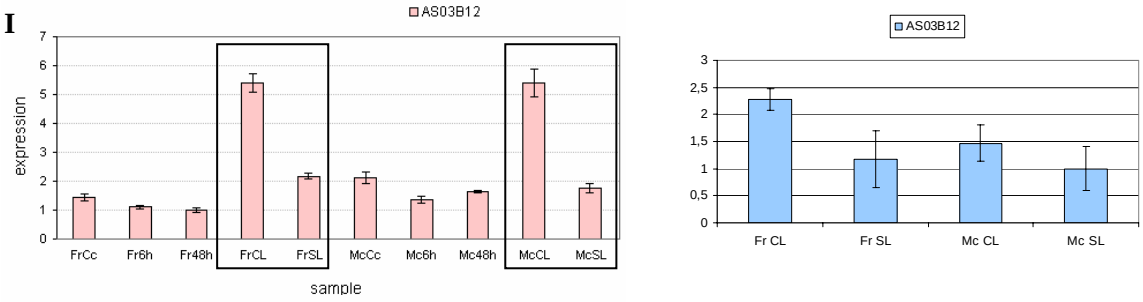
G


RS52G08
expression
FrCc Fr6h Fr48h FrCL FrSL McCc Mc6h Mc48h McCL McSL
sample
RS52G08
Fr CL Fr SL Mc CL Mc SL

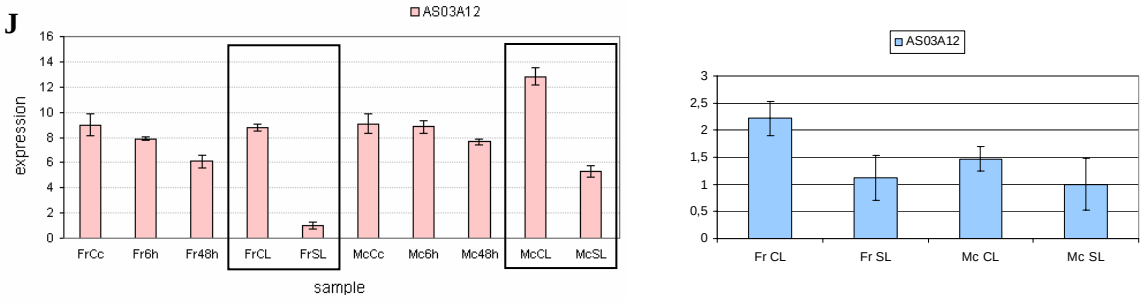
H


Ge085E07
expression
FrCc Fr6h Fr48h FrCL FrSL McCc Mc6h Mc48h McCL McSL
sample
Ge085E07
Fr CL Fr SL Mc CL Mc SL

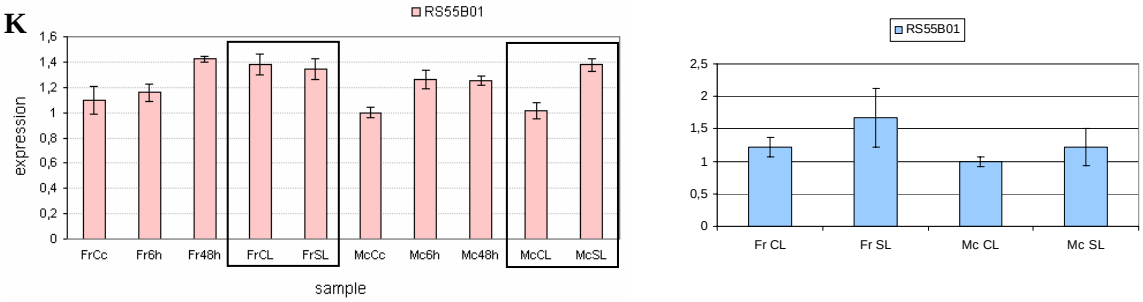
I


AS03B12
expression
FrCc Fr6h Fr48h FrCL FrSL McCc Mc6h Mc48h McCL McSL
sample
AS03B12
Fr CL Fr SL Mc CL Mc SL

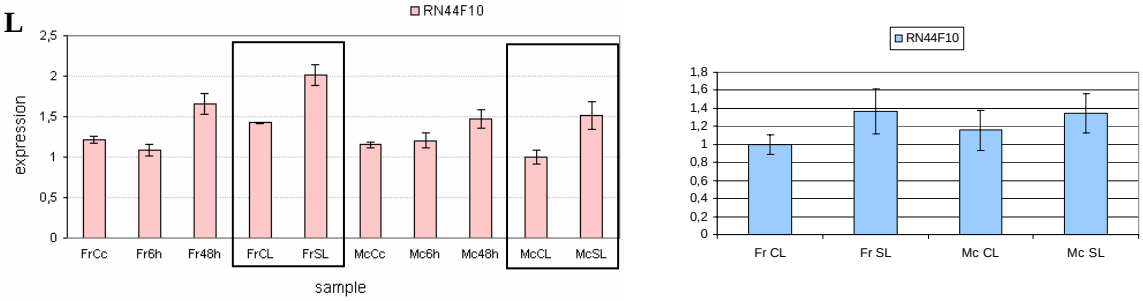
J

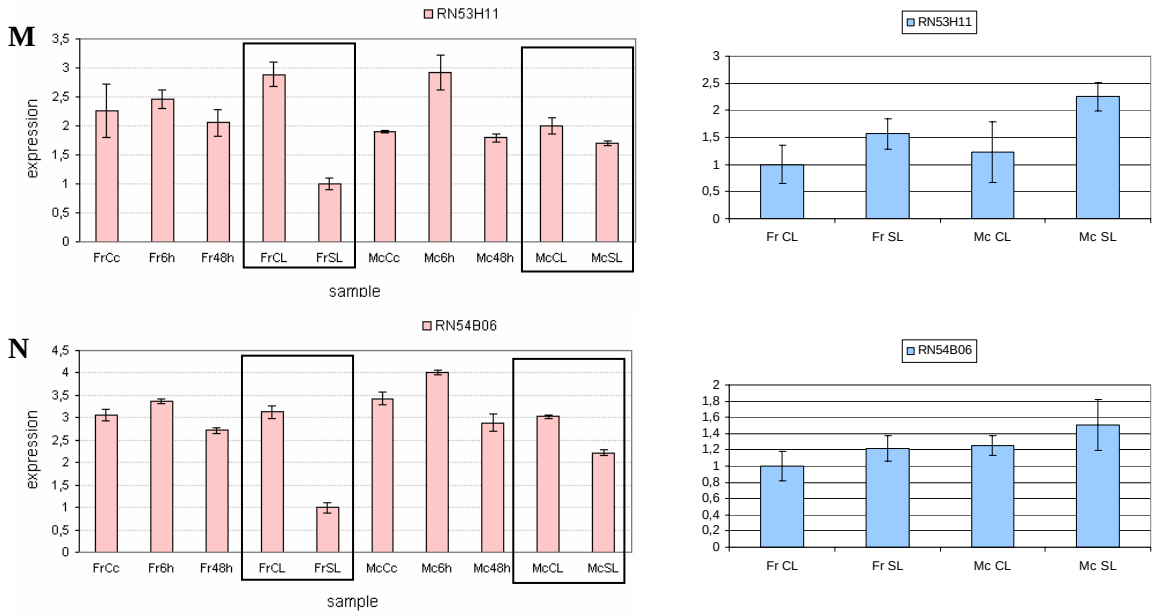

AS03A12
expression
FrCc Fr6h Fr48h FrCL FrSL McCc Mc6h Mc48h McCL McSL
sample
AS03A12
Fr CL Fr SL Mc CL Mc SL

K

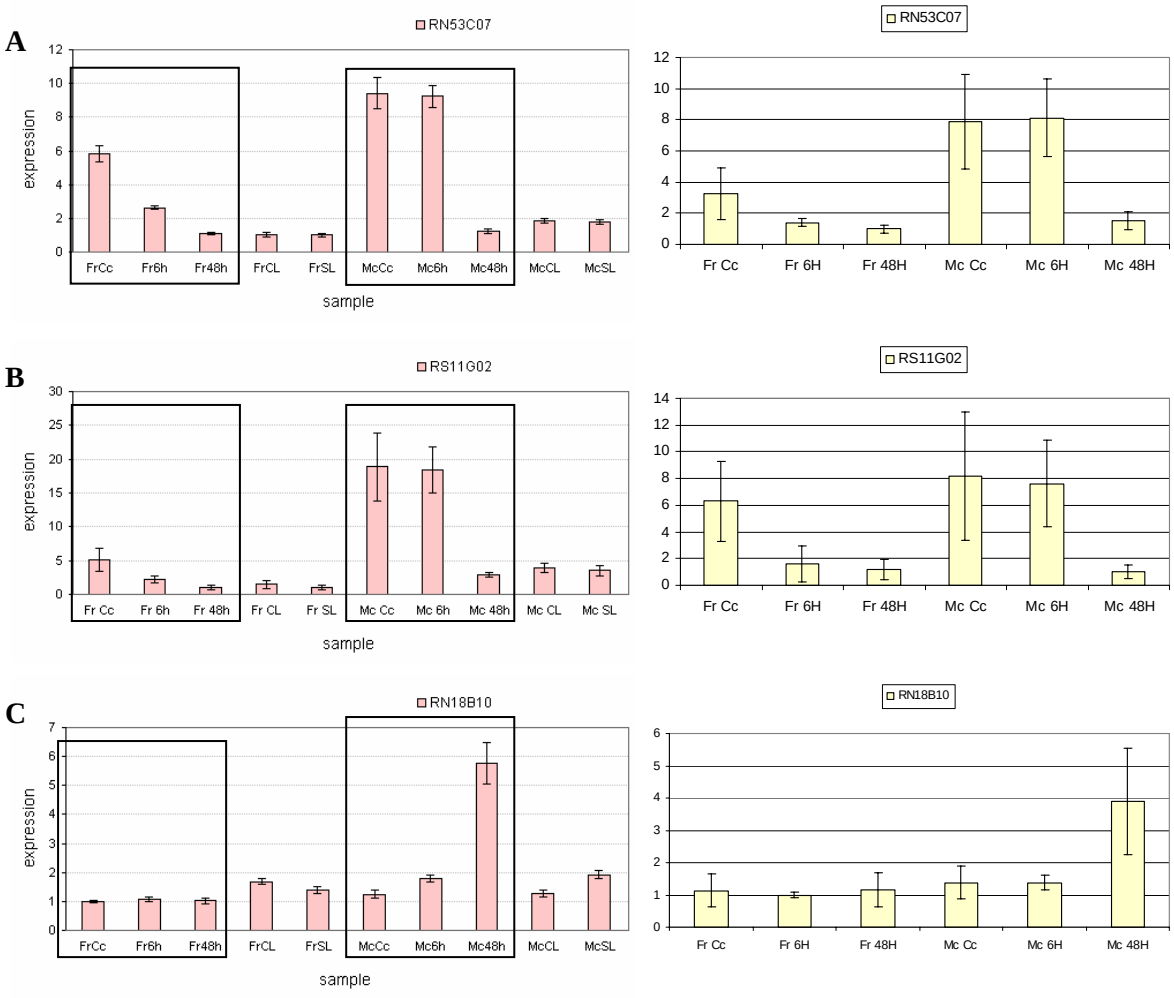

RS55B01
expression
FrCc Fr6h Fr48h FrCL FrSL McCc Mc6h Mc48h McCL McSL
sample
RS55B01
Fr CL Fr SL Mc CL Mc SL

L

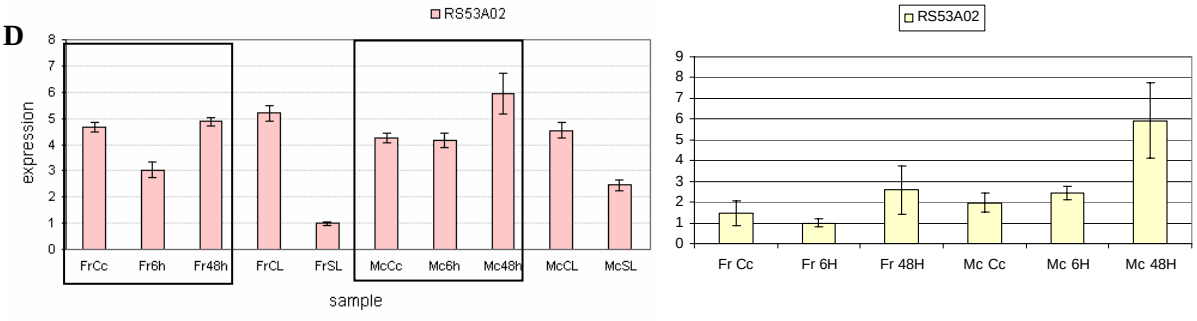

RN44F10
expression
FrCc Fr6h Fr48h FrCL FrSL McCc Mc6h Mc48h McCL McSL
sample
RN44F10
Fr CL Fr SL Mc CL Mc SL



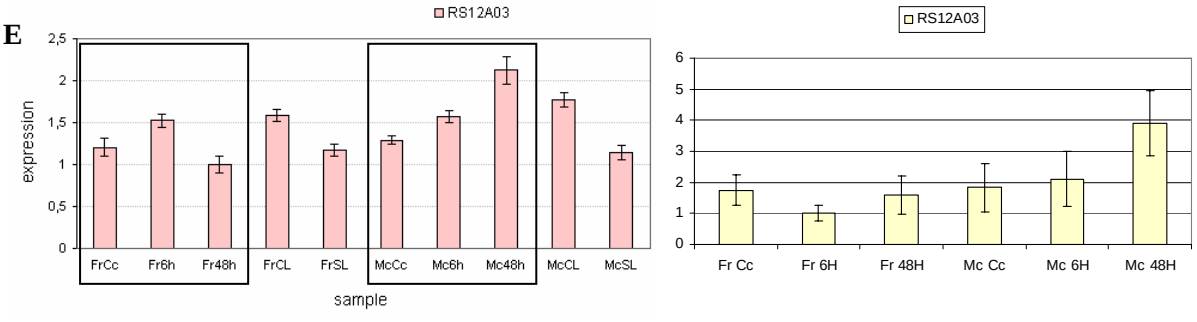
Figures Ax-5 : Profils d'expressions des gènes différentiellement exprimés et regroupés par Expander en « stress court » uniquement. Données d'expression de PCR quantitative (à gauche), d'hybridation (à droite).



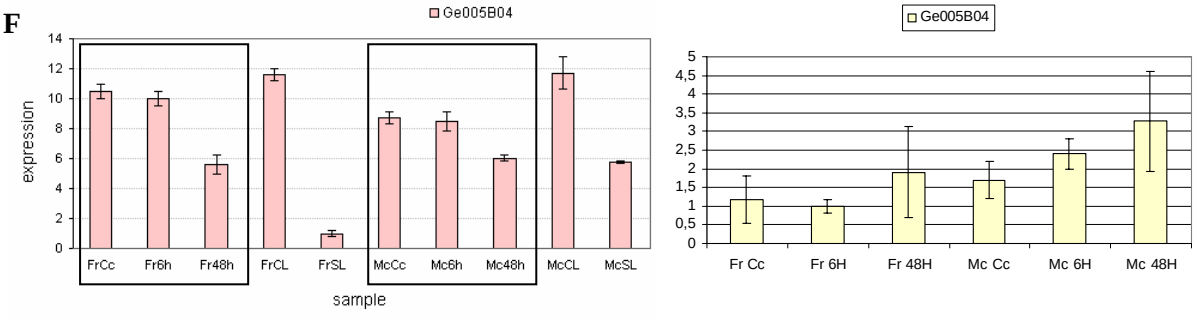
D


RS53A02
expression
FrCc
Fr6h
Fr48h
FrCL
FrSL
McCc
Mc6h
Mc48h
McCL
McSL
sample
RS53A02
Fr Cc
Fr 6H
Fr 48H
Mc Cc
Mc 6H
Mc 48H

E


RS12A03
expression
FrCc
Fr6h
Fr48h
FrCL
FrSL
McCc
Mc6h
Mc48h
McCL
McSL
sample
RS12A03
Fr Cc
Fr 6H
Fr 48H
Mc Cc
Mc 6H
Mc 48H

F


Ge005B04
expression
FrCc
Fr6h
Fr48h
FrCL
FrSL
McCc
Mc6h
Mc48h
McCL
McSL
sample
Ge005B04
Fr Cc
Fr 6H
Fr 48H
Mc Cc
Mc 6H
Mc 48H

G

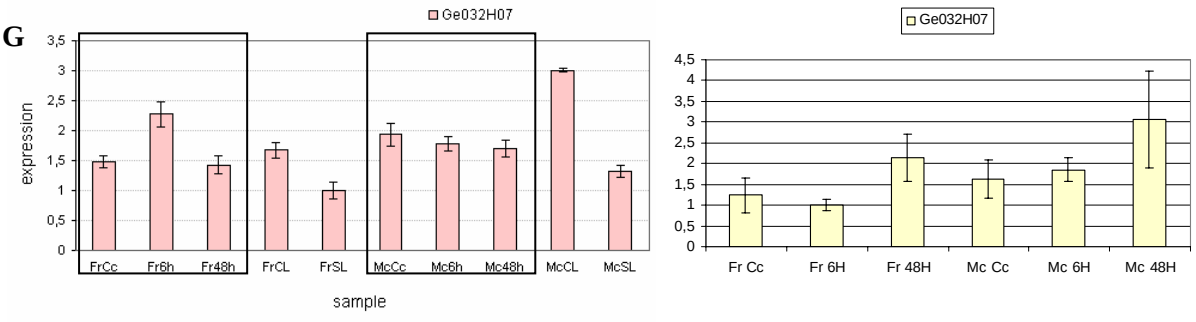

Ge032H07
expression
FrCc
Fr6h
Fr48h
FrCL
FrSL
McCc
Mc6h
Mc48h
McCL
McSL
sample
Ge032H07
Fr Cc
Fr 6H
Fr 48H
Mc Cc
Mc 6H
Mc 48H

Tableau Ax-12 : Liste des protéines identifiées et différentiellement exprimées.

Spot ID	obs. pI/MW	peptide number (%coverage)	TC or singleton	Annotation	theor. MW	theor. pI	SwissProt accession	organism matched	functional category (from KEGG)	significant effect	R2 Prov	R2 Trait	R2 Inter	behavior	Acc spottée
Aminoacids metabolism (1)															
485	6.15/54825	2(25%)	TC75909	Alanine aminotransferase EC 2.6.1.2	52682 Da	6.41	P34106	Panicum miliaceum	Glutamate metabolism, Alanine and aspartate metabolism	I	7,43	0,03	86,4	- L / + M	AL749625
101	6.07/107941	4(14%)	TC62025	Glutaminyl-tRNA synthetase EC 6.1.1.1	90483Da	7.03	P52780	Lupinus luteus	Glutamate metabolism, Alanine and aspartate metabolism	P/I	39,4	6,4	38,3	- L / + M	pas pinaster
149	5.79/78032	13(16%)	TC46379	5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase EC 2.1.1.14	84857 Da	6.78	Q42699	Catharanthus roseus	Methionine metabolism	I	4,6	3,6	58,0	+ L / - M	non sélectionné
150	5.89/77681	28(36%)	TC46379	5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase EC 2.1.1.14	84857 Da	6.78	Q42699	Catharanthus roseus	Methionine metabolism	P/T/I	48,1	19,8	27,3	+ L / = M	non sélectionné
154	4.85/77681	9(14%)	TC46379	5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase EC 2.1.1.14	84857 Da	6.78	Q42699	Catharanthus roseus	Methionine metabolism	I	12,5	1,4	69,4	+ L / - M	non sélectionné
155	5.99/75976	25(31%)	TC46379	5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase EC 2.1.1.14	84857 Da	6.78	Q42699	Catharanthus roseus	Methionine metabolism	I	9,7	20,1	54,7	+ L / - M	non sélectionné
521	6.20/52927	10(27%)	TC46264	S-adenosylmethionine synthetase EC 2.5.1.6	43168 Da	6.08	P50300	Pinus banksiana	Methionine metabolism, Selenoamino acid metabolism	P/T/I	65,7	8,4	20,9	+ L / - M	AL750448
533	6.20/51791	9(21%)	TC46264	S-adenosylmethionine synthetase EC 2.5.1.6	43168 Da	6.08	P50300	Pinus banksiana	Methionine metabolism, Selenoamino acid metabolism	P/T	37,5	36,2	13,4	- L / - M	AL750448
549	5.93/51140	5(8%)	TC46255	S-adenosylmethionine synthetase 2 EC 2.5.1.6	43004 Da	6.01	Q96552	Catharanthus roseus	Methionine metabolism, Selenoamino acid metabolism	P/I	49,9	2,6	32,2	+ L / - M	AL751321
724	4.85/42261	5(10%)	TC67153	Spermidine synthase 2 EC 2.5.1.16	33917 Da	5.22	O48659	Hyoscyamus niger	Urea cycle and metabolism of amino groups, Arginine and proline metabolism, beta-Alanine metabolism	P/T/I	33,2	33,2	33,2	- M (specific MT)	pas pinaster
1345	5.43/42784	6(21%)	TC52194	Adenosine kinase 2 EC 2.7.1.20	37846 Da	5.38	Q9LZG0	Arabidopsis thaliana	Purine metabolism	P/T/I	32,9	32,9	32,9	+ L (specific LS)	non sélectionné
1332	5.38/34078	4(7%)	TC47547	Chorismate mutase, chloroplast EC 5.4.99.5	37594 Da	6.31	P42738	Arabidopsis thaliana	Phenylalanine, tyrosine and tryptophan biosynthesis	P/T/I	29,3	29,3	29,3	+ L (specific LS)	BX254028
1244	5.77/44431	10(28%)	TC46477	Glutamine synthetase nodule isozyme EC 6.3.1.2	39104 Da	6.30	P32289	Vigna aconitifolia		P/T/I	63,1	17,9	17,9	- L	BX252140
Secondary metabolites biosynthesis (2)															
673	6.03/44980	10(19%)	TC51968	Leucoanthocyanidin reductase EC 1.3.1.77	37906 Da	6.51	Q9SEV0	Arabidopsis thaliana	Flavonoid biosynthesis	P/T/I	32,5	32,5	32,5	- M (specific MT)	CR393786
688	5.84/43849	6(17%)	TC73225	Leucoanthocyanidin dioxygenase EC 1.14.11.19	40332 Da	6.41	P51091	Malus domestica	Flavonoid biosynthesis	P/T/I	6,5	82,9	6,5	- L / - M (specific T)	BX678001
1344	5.65/39462	3(9%)	TC66871	Geranyl diphosphate synthase	39432 Da	7.04	Q43133	Sinapis alba	Terpenoid biosynthesis	P/T/I	32,5	32,5	32,5	+ L (specific LS)	CR392964 / BX254083
1351	5.35/62822	11(16%)	TC53438	Pinene synthase, chloroplast EC 4.2.3.14	71506 Da	5.87	O24475	Abies grandis	Monoterpenoid biosynthesis	P/T/I	65,9	15,9	15,9	+ L (specific LS)	AL750955
323	5.30/63497	9(13%)	TC53438	Pinene synthase, chloroplast EC 4.2.3.15	71506 Da	5.87	O24475	Abies grandis	Monoterpenoid biosynthesis	P/T/I	23,6	26,4	37,7	- L / + M	AL750955
Carbohydrates metabolism (3)															
790	5.91/37800	10(22%)	TC73537	Malate dehydrogenase, glyoxysomal EC 1.1.1.37	37385 Da	8.11	Q42972	Oryza sativa	Citrate cycle (TCA cycle), Pyruvate metabolism, Glyoxylate and dicarboxylate metabolism	I	18,7	21,8	42,2	++ L / - M	BX254527
1251	5.49/55073	4(17%)	TC48808	Xylose isomerase EC 5.3.1.5	53614 Da	5.63	Q40082	Hordeum vulgare	Fructose and mannose metabolism	P/I	54,9	3,3	31,8	- L / + M	pas pinaster
288	5.84/65043	10(12%)	TC72976	Pyruvate decarboxylase isozyme 1 EC 4.1.1.1	65428 Da	7.17	P28516	Zea mays	Glycolysis / Gluconeogenesis	P/T/I	12,8	52,7	28,7	++ L / + M	pas pinaster
341	5.15/62487	16(45%)	TC47765	Glucose-6-phosphate isomerase EC 5.3.1.9	57819 Da	5.69	Q8YY05	Anabaena sp	Glycolysis / Gluconeogenesis, Pentose phosphate pathway, Starch and sucrose metabolism	I	19,3	0,5	49,1	+ L / - M	pas pinaster
271	5.87/66508	17(30%)	TC46978	NADP-dependent malic enzyme EC 1.1.1.40	65227 Da	6.75	P51615	Vitis vinifera	Pyruvate metabolism	P/T/I	90,1	5,1	3,8	= L / + M	BX677862
272	5.03/66448	9(16%)	TC46978	NADP-dependent malic enzyme EC 1.1.1.40	65227 Da	6.75	P51615	Vitis vinifera	Pyruvate metabolism	P/T/I	6,1	6,1	85,5	+ L / - M	BX677862
1385	6.06/45060	6(10%)	TC60605	Putative 1-aminocyclopropane-1-carboxylate deaminase EC 3.5.99.7	34756 Da	7.72	Q9WY68	Thermotoga maritima	Propanoate metabolism	P/T/I	32,4	32,4	32,4	+ L (specific LS)	BX252805
418	5.22/58137	20(35%)	TC46852	UTP--glucose-1-phosphate uridylyltransferase EC 2.7.7.9	51920 Da	6.19	P57751	Arabidopsis thaliana		I	5,2	7,4	60,9	- L / + M	BX251183
450	5.34/56740	8(16%)	TC48861	Sucrose-phosphatase EC 3.1.3.24	48486 Da	6.30	Q850K9	Pinus taeda		P	90,9	1,0	1,0	specific M	pas pinaster
227	5.29/70097	12(36%)	TC41956	Phosphoglucomutase, cytoplasmic 1 EC 5.4.2.2	63097 Da	5.86	P93804	Zea mays	Glycolysis / Gluconeogenesis, Pentose phosphate pathway, Starch and sucrose metabolism,Galactose metabolism	P/T/I	53,1	29,5	10,6	++ L / + M	BX252576
Lipids metabolism (4)															
172	5.92/75138	10(22%)	TC47550	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase EC 1.17.4.3	73895 Da	7.54	Q6MD85	Parachlamydia sp.	Biosynthesis of steroids	I	13,0	17,4	46,7	= L / - M	pas pinaster
535	6.24/51791	3(11%)	TC43004	3-oxoacyl-[acyl-carrier-protein] synthase I, chloroplast EC 2.3.1.41	50413 Da	8.25	P52410	Arabidopsis thaliana	Fatty acid biosynthesis	I	3,0	3,0	75,2	+ L / - M	non sélectionné
Energetic metabolism (5)															
419	5.27/57827	15(19%)	TC57318	ATP synthase beta chain EC 3.6.3.14	59856 Da	6.60	P17614	Nicotiana plumbaginifolia	Oxidative phosphorylation, ATP synthesis, Photosynthesis	T/I	0,3	33,3	58,3	- L / ++ M	BX251974
Hydrolases															
456	4.81/55976	9(19%)	TC46505	Serine carboxypeptidase III EC 3.4.16.5	56362 Da	6.52	P21529	Hordeum vulgare	Hydrolases	T/I	4,2	43,8	42,3	- L / = M	BX249187
671	5.27/45020	5(19%)	TC41290	Cysteine proteinase 2 EC 3.4.22.-	39199 Da	7.53	Q10717	Zea mays	Hydrolases	P/T/I	3,2	63,4	32,5	--L / - M	pas pinaster
Nucleotides metabolism															
917	5.04/24927	5(17%)	TC41741	Adenine phosphoribosyltransferase 1 EC 2.4.2.7	26396 Da	8.71	P31166	Arabidopsis thaliana		P/I	38,6	1,1	48,9	+ L / - M	BX250460
793	5.32/37535	3(9%)	TC52267	Thioredoxin reductase 1 EC 1.8.1.9	35314 Da	6.41	Q39243	Arabidopsis thaliana	Pyrimidine metabolism	P/T	49,1	21,9	17,2	- L / = M	pas pinaster
Metabolism (1) and (5)															
795	5.34/37305	4(11%)	TC53481	Cysteine synthase EC 2.5.1.47	41656 Da	8.24	P47999	Arabidopsis thaliana	Cysteine metabolism, Selenoamino acid metabolism, Sulfur metabolism	P/I	56,8	0,4	32,3	+ L / - M	AL750413

Metabolism (1) et (3)																
488	5.66/54381	3(8%)	TC57052	Enolase 1 EC 4.2.1.11	47830 Da	5.96	Q9LEJ0	Hevea brasiliensis		I	0,1	0,1	91,4	+ L / - M	non sélectionné
712	4.99/42298	5(27%)	TC42332	Pyruvate dehydrogenase E1 component beta subunit, mitochondrial EC 1.2.4.1	39176 Da	5.98	Q38799	Arabidopsis thaliana	Glycolysis / Gluconeogenesis, Valine, leucine and isoleucine biosynthesis, Pyruvate metabolism, Butanoate metabolism	P/T/I	19,5	51,3	19,9	+ L / ++ M	AL750582
Metabolism (1) (2) (3)																
1230	5.25/64581	5(18%)	TC53576	Cyanogenic beta-glucosidase EC 3.2.1.21	48325 Da	8.38	P26205	Trifolium repens		P/T/I	20,9	16,2	58,5	- L / + M	pas pinaster
Metabolism (1) et (4)																
808	5.24/36303	5(11%)	TC41552	Phenylcoumaran benzylic ether reductase	33852 Da	6.92	P52578	Solanum tuberosum		P/I	40,7	3,6	50,5	- L / + M	pas pinaster
807	5.96/36367	5(11%)	TC40413	Phenylcoumaran benzylic ether reductase	34654 Da	5.92	P52579	Nicotiana tabacum		P/I	81,6	0,0	16,3	+ L / - M	BX682918
Metabolism (1) (3) (4)																
611	6.48/47833	10(25%)	TC65961	Alcohol dehydrogenase 3 EC 1.1.1.1	41143 Da	6.60	P14675	Solanum tuberosum	Glycolysis / Gluconeogenesis, Fatty acid metabolism, Tyrosine metabolism, Glycerolipid metabolism, Pyruvate metabolism	P/T/I	30,2	34,0	26,5	+ L / = M	BX680855 / BX681260
Cytoskeleton																
600	5.32/47833	14(30%)	TC51794	Actin	41666 Da	5.62	O81221	Gossypium hirsutum		I	17,7	16,9	52,7	--L / + M	AL750018
1464	5.00/56179	6(11%)	TC46352	Tubulin alpha chain	49527 Da	5.16	P33629	Prunus dulcis		T/I	4,2	27,9	50,5	- L / + M	BX249241
1042	4.91/57203	8(22%)	TC40992	Tubulin beta-1 chain	49993 Da	4.96	Q9ZPP0	Eleusine indica		P/I	26,8	7,0	49,9	+ L / - M	BX249627
PR proteins (Pathogenesis Related)																
918	5.08/24687	4(13%)	TC57064	Thaumatin-like protein	23999 Da	7.76	P33679	Zea mays		P/T/I	29,7	37,1	32,2	= L / + M	pas pinaster
971	4.82/15215	5(12%)	TC48205	Pathogenesis-related protein Pseml	17420 Da	6.05	P43180	Betula verrucosa		P/T/I	23,2	44,6	20,4	-- L / - M	non sélectionné
Stress response proteins																
384	5.79/59978	8(28%)	TC53907	Putative selenium-binding protein	54057 Da	5.84	O23264	Arabidopsis thaliana		P/T/I	24,0	45,5	25,6	+ L / ++ M	BX250616
397	5.67/59129	9(30%)	TC53907	Putative selenium-binding protein	54057 Da	5.84	O23264	Arabidopsis thaliana		I	16,6	3,0	62,0	- L / + M	BX250616
631	4.72/46794	12(12%)	TC64985	Heat shock cognate 70 kDa protein	71227 Da	5.32	P09189	Petunia hybrida		P/T	52,2	35,5	3,4	+ L / + M	BX255035
243	5.20/68675	10(22%)	TC47544	Heat shock 70 kDa protein, mitochondrial	72537 Da	6.38	Q01899	Phaseolus vulgaris		I	20,7	13,9	41,7	- L / + M	BX680662
637	4.80/46507	3(3%)	TC40629	Heat shock cognate 70 kDa protein	71227 Da	5.32	P09189	Petunia hybrida		P/T/I	31,2	31,2	31,2	- M (specific MT)	BX255035
638	4.92/46712	13(13%)	TC58306	Heat shock protein 81-2	80064 Da	5.14	P55737	Arabidopsis thaliana		P/T	40,3	56,3	0,5	+ L / + M	CR393344
639	5.10/46671	11(10%)	TC58306	Heat shock protein 81-2	80064 Da	5.14	P55737	Arabidopsis thaliana		P/T	60,4	33,8	0,0	+ L / + M	CR393344
1137	5.20/72704	13(18%)	TC65892	Heat shock 70 kDa protein	70880 Da	5.59	P26413	Glycine max		P/T/I	17,8	47,0	32,6	++ L / + M	BX251816
1210	4.88/47540	9(6%)	TC40629	heat shock cognate protein 70	71436 Da	5.32	Q8GSN3	Cucurbita maxima		P/T/I	30,3	30,3	30,3	- L (specific LT)	BX255035
1211	4.77/47248	8(6%)	TC40629	heat shock cognate protein 70	71436 Da	5.32	Q8GSN3	Cucurbita maxima		P/T/I	31,8	31,8	31,8	- L (specific LT)	BX255035
231	6.05/69785	10(15%)	TC66522	Heat shock protein STI	63585 Da	6.39	Q43468	Glycine max		P/T/I	33,0	33,0	33,0	- M (specific MT)	1 non sélectionné
237	5.06/69475	5(11%)	TC53412	Tubulin-specific chaperone C-like protein	62888 Da	5.84	Q6H7Q8	Oryza sativa		I	4,6	4,6	56,0	+ L / - M	BX676878
702	5.45/42972	11(23%)	TC73649	Auxin-induced protein PCNT115	33857 Da	7.69	P40691	Nicotiana tabacum		P/T/I	2,9	21,2	74,1	-- L / + M	CR392243
726	6.08/41928	7(16%)	TC73613	Auxin-induced protein PCNT115	33857 Da	7.69	P40691	Nicotiana tabacum		T/I	1,5	32,5	45,3	= L / + M	BX678549 / BX678993
890	5.82/26883	2(11%)	TC73079	ASR (Abscisic stress ripening) protein	13130 Da	7.68	Q08655	Lycopersicon esculentum		P/T/I	73,1	13,6	12,6	+ L / = M	BX251094 / BX682324
915	5.97/24992	2(20%)	TC73079	ASR (Abscisic stress ripening) protein	13130 Da	7.68	Q08655	Lycopersicon esculentum		P/I	53,3	6,0	35,0	+ L / - M	BX251094 / BX682324
897	5.86/26368	2(5%)	TC73461	Water deficit inducible protein LP3	17291 Da	6.56	O24314	Pinus taeda		P/T/I	27,6	46,7	24,6	-- L / - M	pas pinaster
1095	5.97/26023	2(14%)	TC73497	Water deficit inducible protein LP3	17291 Da	6.56	O24314	Pinus taeda		P/T/I	19,6	59,1	19,6	-- L / - M	BX253505
451	4.70/56484	6(16%)	TC41572	Osr40g3 protein	22822 Da	8.07	O24213	Oryza sativa		P/T/I	22,9	5,4	70,0	- L / + M	BX251186
433	4.68/57670	7(20%)	TC41572	Osr40g3 protein	22822 Da	8.07	O24213	Oryza sativa		P/T/I	44,6	39,7	9,9	++ L / + M	BX251186
Oxydative stress																
969	6.24/15498	2(9%)	TC46675	Superoxide dismutase [Cu-Zn] EC 1.15.1.1	15376 Da	6.53	P24669	Pinus sylvestris		P/I	27,37	35,2	24,3	= L / + M	BX249851
907	4.99/25368	9(25%)	TC65955	L-ascorbate peroxidase, cytosolic EC 1.11.1.11	27062 Da	6.00	P48534	Pisum sativum	Ascorbate and aldarate metabolism	P/T	36,6	45,6	6,5	+ L / ++ M	BX677114
Other categories																
815	4.33/35204	4(18%)	TC47197	Elongation factor 1-beta	24366 Da	4.66	P93447	Pimpinella brachycarpa		I	4,8	2,8	70,1	- L / + M	BX255307
816	4.28/35111	3(11%)	TC59049	Elongation factor 1-beta	24366 Da	4.66	P93447	Pimpinella brachycarpa		I	1,0	1,0	86,1	+ L / - M	BX255307
392	6.34/59499	11(23%)	TC41101	Ribulose bisphosphate carboxylase large chain EC 4.1.1.38	52713 Da	6.93	P26962	Pinus balfouriana	Photosynthesis, respiration	P/T/I	23,7	52,4	17,9	-- L / - M	pas pinaster
410	6.48/58709	15(29%)	TC41101	Ribulose bisphosphate carboxylase large chain EC 4.1.1.39	52713 Da	6.93	P26962	Pinus balfouriana	Photosynthesis, respiration	P/T/I	37,6	29,5	21,5	+ L / - M	pas pinaster
1206	5.07/31821	7(21%)	TC47183	Oxygen-evolving enhancer protein 1, chloroplast	34869 Da	6.86	O49079	Fritillaria agrestis	Photosynthesis	P	88,7	2,1	2,1	specific L	AL749654
845	5.65/31710	8(22%)	TC46528	Phosphosulfolactate synthase EC 4.4.1.19	29409 Da	5.08	O27710	Methanobacterium thermoautotrophicum	Lyase	T/I	3,3	34,4	43,3	+ L / - M	BX252228
1139	5.45/153833	3(13%)	TC56156	Lon protease homolog 2, mitochondrial EC 3.4.21.-	103930 Da	5.74	P93655	Arabidopsis thaliana		I	0,9	0,9	86,1	+ L / - M	CR392686
1110	4.72/79687	12(15%)	TC66389	Endoplasmin homolog	92917 Da	5.05	P36183	Hordeum vulgare		P/T/I	32,3	32,3	32,3	- M (specific MT)	pas pinaster
1177	4.77/65566	4(7%)	TC41594	Protein disulfide-isomerase EC 5.3.4.1	55561 Da	5.18	Q43116	Ricinus communis		I	17,1	17,1	47,6	- L / + M	AL750759
1341	5.41/40945	7(12%)	TC53473	Serine/threonine protein phosphatase PP1 isozyme 2 EC 3.1.3.16	35924 Da	5.90	P48481	Acetabularia cliftonii		P/T/I	32,7	32,7	32,7	LS pas pinaster	
1435	5.35/57411	8(13%)	TC74921	Anthranilate N-benzoyltransferase protein 1 EC 2.3.1.144	50038 Da	7.08	O24645	Dianthus caryophyllus		P	71,7	5,6	5,6	specific L	BX679860
91	5.23/112117	19(30%)	TC46746	Cell division cycle protein 48 homolog	89770 Da	5.42	P54774	Glycine max		P/T/I	30,1	30,1	30,1	- M (specific MT)	BX251155
94	5.28/111175	26(37%)	TC46746	Cell division cycle protein 48 homolog	89770 Da	5.42	P54774	Glycine max		I	8,8	20,8	45,4	+ L / - M	BX251155
116	5.70/99629	5(10%)	TC42016	ATP-dependent Clp protease ATP-binding subunit clpA homolog CD4B	102241 Da	6.33	P31542	Lycopersicon esculentum		P/I	65,9	16,8	13,3	+ L / = M	CR393594

Figure Ax-6 : Alignement des séquences de riz, épicéa et pins homologues à l'ACC désaminase

```
                        ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                                10         20         30         40         50         60         70
oryza XM_468034.1       -~~--------- ---------- ---------- ---------- ---------- ---------- ----------
spruce CO484002         -~~--------- ---------- ---------- ---------- ----GGAACC CTGATGGGTT CTGCGTCGTT
spruce TC8605           -~~--------- ---------- ---------- ---------- ---------- ---------- ----------
pinus BE662323          -~~--------- ---------- ---------- ---------- ---------- ---------- GGGCGTCGTT
pinus TC44655           -~~--------- ---------- ---------- ---------- ---------- ---------- ----------
pinaster BX252805       -~~--------- ---------- ---------- ---------- ---------- ---------- ----------
pinaster Consensus      -~~--------- ---------- ---------- ---------- ---------- ---------- ----------
TC60605                 GACGCTAAAG AATCCAGTTG AAAAATATAA TATAGTTGAC ATCCTGGAAC CCTATGGGTT CGGCGTCGTT
TC60606                 -~~--------- ---------- ---------- ---------- ---------- ---------- ----------


                        ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                                80         90        100        110        120        130        140
oryza XM_468034.1       ---------- ---------- ---------- ---------- GTGGGTTTTG GACTGTAGCT GCTGCCCCGA
spruce CO484002         AATTTCATAT AGGATTTGGC TGCCTGCCGA TTGCAAACGT TTCAGTATTG G-CAGTAGCT GCCATT-GGA
spruce TC8605           ---------- ---------- ---------- ---------- ---------- ---------- ----------
pinus BE662323          CATTTCATAT AGGATTAGGC TGACTCCCGA TTGCAAACAT TTCGGCTTTA G-CAGTGGCT GCCATT-GGA
pinus TC44655           ---------- ---------- ---------- ---------- ---------- ---------- ----------
pinaster BX252805       ---------- ---------- ---------- ---------- ---------- ---------- ----------
pinaster Consensus      ---------- ---------- ---------- ---------- ---------- ---------C TGCATT-GGA
TC60605                 CATTTCATAT AGGATTAGGC TGACTCCCGA TTGCAAACAT TTCGGCTTTA G-CAGTGGCT GCCATT-GGA
TC60606                 ---------- ---------- ---------- ---------- ---------- ---------- ----------


                        ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                               150        160        170        180        190        200        210
oryza XM_468034.1       CGCGCTGCTC -GCTGCCCCA CTCTCTCCCC -ATTCCCC-T CCACGCCGCC -GCCGCCGCC -GCCGCG-TG
spruce CO484002         TTCGTGTTAC ATCTGCATTG GGTTATAGGA AATTACAGGT GAAAGACAAG AGCATCAGAG CGAAGACAAC
spruce TC8605           ---------- ---------- ---------- ---------- ---------- ---------- ----------
pinus BE662323          TTGGTTTTAC ATCTCCATGG TGCTATAGTA AATTACAGGT GAAAGGCAAG AGCATCAGAG CGAAGACAAC
pinus TC44655           ---------- ---------- ---------- ---------- ---------- ---------- ----------
pinaster BX252805       ---------- ---------- ---------- ---------- ---------- ---------- ----------
pinaster Consensus      TTGGTTTTAC ATCTCCATTG GGCTATAGTA AATTACAGGG GAAAGGCAAG AGCATCAGAG CGAAGACAAC
TC60605                 TTGGTTTTAC ATCTCCATGG TGCTATAGTA AATTACAGGT GAAAGGCAAG AGCATCAGAG CGAAGACAAC
TC60606                 ---------- ---------- ---------- ---------- ---------- ---------- ----------


                        ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                               220        230        240        250        260        270        280
oryza XM_468034.1       GATGGCCGGA G---TTTCCG CCGCCTCCGC CGCCGG-GAA GA-----TCG GGAGCTTCCT CTCCAAGAGG
spruce CO484002         AATGGCCCAC GGGCATAGTT CAGAGCGAGG GGAAGGAGAA GGCATTTCCA GAGCCTTTCT GAGCAAATTT
spruce TC8605           ---------- ---------- ---------- ---------- ---------- ---------- ----------
pinus BE662323          AATGGCTCAC GGGCTTCGTG CAGAGCGAGT GGAAGGAGAA GGCAGTTCTG GAGCCTTTCT GAGCAAATTT
pinus TC44655           ---------- ---------- ---------- ---------- ---------- ---------- ----------
pinaster BX252805       ---------- ---------- ---------- ---------- ---------- ---------- ----------
pinaster Consensus      AATGGCTCAC GGGCTTCGTG CAGAGCGAGT GGAAGGAGAA GGCAGTTGTG GAGCCTTTCT GAGCAAATTT
TC60605                 AATGGCTCAC GGGCTTCGTG CAGAGCGAGT GGAAGGAGAA GGCAGTTCTG GAGCCTTTCT GAGCAAATTT
TC60606                 ---------- ---------- ---------- ---------- ---------- ---------- ----------


                        ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                               290        300        310        320        330        340        350
oryza XM_468034.1       CCGTACGCGC CGCCGTCCTG GGCCTCGCAC CTGTCCCCCG CCCCCTCGCA GACCTTCTCG CTCGGCCATT
spruce CO484002         CCATACACAC CTCCTCCCTG GGCAACTTAT CTTCATCCGA TTCCTTCTCA TTTCTATTCT CTTGGACATC
spruce TC8605           ---------- ---------- ---------- ---------- ---------- ---------- ----------
pinus BE662323          TCATACACAC CTCCCTCCTG GGCGACAGTG CTTCATCCGA TTCCGTCCCA TTTCTATTCT CTCGGACATC
pinus TC44655           ---------- ---------- ---------- ---------- ---------- ---------- ----------
pinaster BX252805       ---------- ---------- ---------- ---------- ---------- ---------- ----------
pinaster Consensus      TCATACACAC CTCCCTCCTG GGCGACAGAG CTTCATCCGA TTCCGTCCCA TTTTTATTCT CTCGGACATC
TC60605                 TCATACACAC CTCCCTCCTG GGCGACAGAG CTTCATCCGA TTCCGTCCCA TTTCTATTCT CTCGGACATC
TC60606                 ---------- ---------- ---------- ---------- ---------- ---------- ----------


                        ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                               360        370        380        390        400        410        420
oryza XM_468034.1       TCCCGACGCC GATCCACAAG TGGAATCTGC CCAATTTGCC GAATGGCACG GAGGTGTGGA TCAAGCGAGA
spruce CO484002         TCCCTACGCC AATTCATCAA TGGAATCTTC CAGGCTTACC AACAAACACT GAGGTCTGGA TAAAGCGGGA
spruce TC8605           ---------- ---------- ---------- ---------- ---------- ---------- ----------
pinus BE662323          TCCCCACACC AATTCATCGA TGGAATCTTC CGGGCTTACC ATCAAACACT GAAGTCTGGA TAAAGCGTGA
pinus TC44655           ---------- ---------- ---------- ---------- ---------- ---------- ----------
pinaster BX252805       ---------- ---------- ---------- ---------- ---------- ---------- ----------
pinaster Consensus      TCCCTACACC AATTCATCGA TGGAATCTTC CGGGCTTACC ATCAAACACT GAAGTCTGGA TAAAGCGTGA
TC60605                 TCCCCACACC AATTCATCGA TGGAATCTTC CGGGCTTACC ATCAAACACT GAAGTCTGGA TAAAGCGTGA
TC60606                 ---------- ---------- ---------- ---------- ---------- ---------- ----------


                        ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                               430        440        450        460        470        480        490
oryza XM_468034.1       CGACATCTCA GGCATGCAGT TGAGCGGGAA CAAGGTCCGG AAGCTCGAGT TCCTGATGGC AGATGCCGTC
spruce CO484002         TGACTTATCA GGTATGCAAC TTAGTGGGAA CAAGGTGCGG AAGCTGGAAT TTTTGATGGC TGATGCAAAG
```

```
spruce TC8605          ---------- ---------- ---------- ---------- ---------- ---------- ----------
pinus BE662323         TGACTTATCA GGTATGCAAC TT-....... .......... .......... .......... ..........
pinus TC44655          ---------- ---------- ---------- ---------- ---------- ---------- ----------
pinaster BX252805      ---------- ---------- ---------- ---------- ---------- ---------- ----------
pinaster Consensus     TGACTTATCA GGTATGCAAC TTAGTGGGAA CAAGGTGCGG AAGCTGGAAT TTTTAATGGC TGATGCAAAG
TC60605                TGACTTATCA GGTATGCAAC TTAGTGGGAA CAAGGTGCGG AAGCTGGAAT TTTTAATGGC TGATGCAAAG
TC60606                ---------- ---------- ---------- ---------- ---------- ---------- ----------

                       ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                               500        510        520        530        540        550        560
oryza XM_468034.1      GCGCAGGGCG CTGACTGCGT TATAACTGTA GGTGGCATCC AGAGCAATCA CTGCCGTGCC ACCGCAGTGG
spruce CO484002        GCACAAGGTG CAGATTGTGT TATCACAATA GGAGGCATTC AAAGTAACCA CTGCCGTGCT ACTGCAGTGG
spruce TC8605          ---------- ---------- ---------- ---------- ---------- ---------- ----------
pinus BE662323         .......... .......... .......... .......... .......... .......... ..........
pinus TC44655          ---------- ---------- ---------- ---------- ---------- ---------- ----------
pinaster BX252805      ---------- ---------- ---------- ---------- ---------- ---------- ----------
pinaster Consensus     GCGCAAGGTG CAGATTGTGT TATCACAATA GGAGGCATTC AAAGTAACCA CTGTCGTGCT ACTGCAGTGG
TC60605                GCGCAAGGTG CAGATTGTGT TATCACAATA GGAGGCATTC AAAGTAACCA CTGTCGTGCT ACTGCAGTGG
TC60606                ---------- ---------- ---------- ---------- ---------- ---------- ----------

                       ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                               570        580        590        600        610        620        630
oryza XM_468034.1      CTGCAAAGTA TATAAATCTT GATTGTTATC TGATACTGCG CACATCCAAG CTTCTTGTGG ATAAGGACCC
spruce CO484002        CAGCTCGATA CTTGAATCTT GACTGCTATC TCATATTGCG AACTACCAGG GTGCAAGTCG ATGAAGATCC
spruce TC8605          ---------- ---------- ---------- ---------- ---------- ---------- ----------
pinus BE662323         .......... .......... .......... .......... .......... .......... ..........
pinus TC44655          ---------- ---------- ---------- ---------- ---------- ---------- ----------
pinaster BX252805      ---------- ---------- ---------- ---------- ---------- ---------- ----------
pinaster Consensus     CAGCTCGATA CTTGAATCTT GACTGCTATC TCATATTACG AACTACCAGG GCACAAGTCG ATGAAGATCC
TC60605                CAGCTCGATA CTTGAATCTT GACTGCCATC TCATATTACG AACTACCAGG GCACAAGTCG ATGAAGATCC
TC60606                ---------- ---------- ---------- ---------- ---------- ---------- ----------

                       ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                               640        650        660        670        680        690        700
oryza XM_468034.1      TGGTTTGGTT GGGAATCTCC TTGTTGAGAG ATTGGTGGGA GCGCATATTG ATCTTGTTTC AAAAGAAGAA
spruce CO484002        TGGCTTGACT GGGAATCTCC TTGTTGAGCG GATGGTTGGT GCGAACATAG AGCTTGTGTC GAANNAGGAG
spruce TC8605          ---------- ---------- ---------- ---------- ---------- ---------- ----------
pinus BE662323         .......... .......... .......... .......... .......... .......... ..........
pinus TC44655          ---------- ---------- ---------- ---------- ---------- ---------- ----------
pinaster BX252805      ---------- ---------- ---------- ---------- ---------- ---------- ----------
pinaster Consensus     TGGCTTGACT GGGAATCTTC TTGTGGAGCG GATGGTTGGT GCGAACATAG AGCTTGTGTC CAAGGAGGAA
TC60605                TGGCTTGACT GGGAATCTTC TTGTGGAGCG GATGGTTGGT GCGAACATAG AGCTTGTGTC CAAGGAGGAG
TC60606                ---------- ---------- ---------- ---------- ---------- ---------- ----------

                       ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                               710        720        730        740        750        760        770
oryza XM_468034.1      TATGGAAAAA TT-GGCAGTG TGGCTTTAGC GGACTTGCTG AAAAAGAAGC TTTTGGAAGA AGGCCGAAAA
spruce CO484002        TATGCACGAG TT-GGAAGTG TGGCTCTTNG ANNATTGTTG GANNNNNNNN NGTTAAAGNN NNNCAGGANG
spruce TC8605          ---------- ---------- ---------- ---------- ---------- ---------- ----------
pinus BE662323         .......... .......... .......... .......... .......... .......... ..........
pinus TC44655          ---------- ---------- ---------- ---------- ---------- ---------- ----------
pinaster BX252805      ---------- ---------- ---------- ---------- ---------- ---------- ----------
pinaster Consensus     TATGCACAAA TT-GGAAGTG TGGCTCTTGG AAAATTCTTG GAGGAAAGGT TGTTAAAGGA AGGCAAGAAG
TC60605                TATGCACAAA TTTGGAAGTG TGGCTCTTGG AAAATTCTTG GAGGAAAGGT TGTTAAAGGA AGGCAAGAAG
TC60606                ---------- ---------- ---------- ---------- ---------- ---------- ----------

                       ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                               780        790        800        810        820        830        840
oryza XM_468034.1      CCATATGTTA TTCCTGTTGG TGGATCAAAC TCTCTAGGAA CTTGGGGATA TATAGAAGCA ATAAGGGAGA
spruce CO484002        NNTTATGTTA TTCCTGTTGG TGGCTCNNNC TCAT-..... .......... .......... ..........
spruce TC8605          ----ATGNNA NNNCTGTTGG TNNCTCNNNC TCATTGGGAA CATGGGGATA CATAGAAGCG ATGAGAGAGA
pinus BE662323         .......... .......... .......... .......... .......... .......... ..........
pinus TC44655          ---------- ---------- ---------- ---------- ---------- ---------- ----------
pinaster BX252805      ---------- ---------- ---------- ---------- ---------- ---------- ----------
pinaster Consensus     CCTTATGTTA TTCCTGTTGG TGGATCAAAC TCATTGGGAA CATGGGGCTA CATAGAGGCA GTGAGGGAGA
TC60605                CCTTATGTTA TTCCTGTTGG TGGCTCAAAC TCATTGGGAA CATGGGGGTA CATAGAGGCA GTGAGGGAGA
TC60606                ---------- ---------- ---------- ---------- CATGGGGGTA CATAGAGGCA GTGAGGGAGA

                       ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                               850        860        870        880        890        900        910
oryza XM_468034.1      TTGAACATCA AATTCA---- -----GATAT CAGGTGATGT TCAGTTTGAC GACATTGTTG TTGCATGTGG
spruce CO484002        .......... .......... .......... .......... .......... .......... ..........
spruce TC8605          TAGAGCAACA AGTGCAAGAA AAGAAGGTGC CCGGTATTAC TTTTTTTGAT GACATTGTGG TTGCTTGTGG
pinus BE662323         .......... .......... .......... .......... .......... .......... ..........
pinus TC44655          ---------- ---------- ---------- ---------- ---------- ---------- ----------
pinaster BX252805      ---------- ---------- ---------- ---------- ---------- ---------- ----------
pinaster Consensus     TAGAGCAACA ACTGCAAGAA AAGAAGGTGC CTGGTATTAC TTCTTTTGAT GACATTGTGG TTGCATGTGG
TC60605                TAGAGCAACA ACTGCAAGAA AAGAAGGTGC CTGGTATTAC TTCTTTTGAT GACATTGTGG TTGCATGTGG
TC60606                TAGAGCAACA ACTGCAAGAA AAGAAGGTGC CTGGTATTAC TTCTTTTGAT GACATTGTGG TTGCATGTGG
```

```
                    ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                            920        930        940        950        960        970        980
oryza XM_468034.1   CAGCGGTGGA ACCATTGCTG GTCTTGCTTT AGGATCAAAA TTGAGTAGCT TAAAGGCAAA AGTCCATGCT
spruce CO484002     .......... .......... .......... .......... .......... .......... ..........
spruce TC8605       AAGTGGAGGA AGTATAGCTG GCTTGTCTCT AGGATCATAT TTAAGCAACT TAAAGGCAAA AGTGCATGCA
pinus BE662323      .......... .......... .......... .......... .......... .......... ..........
pinus TC44655       ---------- ---------- ---------- ---------- ---------- ---------- ----------
pinaster BX252805   ---------- ---------- ---------- ---------- ---------- ---------- ----------
pinaster Consensus  AAGTGGAGGA AGTATAGCTG GCTTGTCCCT AGGATCACAT TTAAGCAACT TAAAGGCAAA AGTGCATGCA
TC60605             AAGTGGAGGA AGTATAGCTG GCTTGTCCCT AGGAGCACAT TTAAGCAACT TAAAGGCAAA AGTGCATGCA
TC60606             AAGTGGAGGA AGTATAGCTG GCTTGTCCCT AGGAGCACAT TTAAGCAACT TAAAGGCAAA AG--------

                    ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                            990       1000       1010       1020       1030       1040       1050
oryza XM_468034.1   TTCTCTGTTT GTGATGATCC TGGATATTTC CATTCCTATG TTCAAGACCT GATTGATGGA CTTCATTCTG
spruce CO484002     .......... .......... .......... .......... .......... .......... ..........
spruce TC8605       TTTGCTGTTT GTGATGATCC TGACTACTTT TATGATTACA CTCAAGGCCT TTTGGATGGC CTCAATGCTG
pinus BE662323      .......... .......... .......... .......... .......... .......... ..........
pinus TC44655       ---------- ---------- ---------- ---------- ---------- ---------- -------CTG
pinaster BX252805   ---------- ---------- ---------- ---------- ---------- ---------- -------CTG
pinaster Consensus  TTTTCTGTTT GTGATGATCC TGACTACTTT TATGATTACA CTCAAGGCCT TTTGGATGGC CTCAATGCTG
TC60605             TTTTCTGTTT GTGATGATCC TGACTACTTT TATGATTACA CTCAAGGCCT TTTGGATGGC CTCAATGCTG
TC60606             ---------- ---------- ---------- ---------- ---------- ---------- ----------

                    ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                            1060       1070       1080       1090       1100       1110       1120
oryza XM_468034.1   ATTTGCGTTC ACATGATTTA GTCAACATTG AAAATGCTAA GGGCTTAGGT TATGCCATGA ACACAGCTGA
spruce CO484002     .......... .......... .......... .......... .......... .......... ..........
spruce TC8605       GTTTGAACTC TCGTGACTTG GTCAATATTA TAAATGCAAA AGGCCTTGGC TATGCAATGA GCACAGCAGA
pinus BE662323      .......... .......... .......... .......... .......... .......... ..........
pinus TC44655       GATTCAACTC CCGTGACTTG ATCACTATTA TAAATGCAAA AGGCCTTGGC TACGCAATGA GCACGGCAGA
pinaster BX252805   GATTCAACTC CCGTGACTTG ATCACTATTA TAAATGCAAA AGGCCTTGGC TACGCAATGA GCACGGCAGA
pinaster Consensus  GATTCAACTC CCGTGACTTG ATCACTATTA TAAATGCAAA AGGCCTTGGC TACGCAATGA GCACGGCAGA
TC60605             GATTCAACTC CCGTGACTTG ATCACTATTA TAAATGCAAA AGGCCTTGGC TACGCAATGA GCACGACAGA
TC60606             ---------- ---------- ---------- ------CAAA AGGCCTTGGC TACGCAATGA GCACGACAGA

                    ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                            1130       1140       1150       1160       1170       1180       1190
oryza XM_468034.1   GGAGCTTAAG TTTGTTAAAG ATATAGCTAC AGCAACAGGC ATTGTCCTTG ATCCAGTCTA CAGTGGAAAG
spruce CO484002     .......... .......... .......... .......... .......... .......... ..........
spruce TC8605       GGAGCTTAAA TGTGTGACTG AAATTGCAGA AACAACAGGA ATTATCCTCG ACCCTGTTTA CAGTGGGAAG
pinus BE662323      .......... .......... .......... .......... .......... .......... ..........
pinus TC44655       GGAGCTTAAA TGTGTGATTG AAATTGCAGA AACAACAGGA ATTATCCTAG ACCCTGTTTA CAGTGGGAAG
pinaster BX252805   GGAGCTTAAA TGTGTGATTG AAATTGCAGA AACAACAGGA ATTATCCTAG ACCCTGTTTA CAGTGGGAAG
pinaster Consensus  GGAGCTTAAA TGTGTGATTG AAATTGCAGA AACAACAGGA ATTATCCTCG ACCCTGTTTA CAGTGGGAAG
TC60605             GGAGCTTAAA TATGTGCTTG AAATTGCAGA AACAACAGGA GTTATCCTCG ACCCTGTTTA CAGTGGGAAG
TC60606             GGAGCTTAAA TATGTGCTTG AAATTGCAGA AACAACAGGA GTTATCCTCG ACCCTGTTTA CAGTGGGAAG

                    ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                            1200       1210       1220       1230       1240       1250       1260
oryza XM_468034.1   GCAGCATATG GAATGCTGAA AGACATGGGT GCTAATCCAG CTAAGTGGGA AGGGCGAAAA ATTCTTTTTG
spruce CO484002     .......... .......... .......... .......... .......... .......... ..........
spruce TC8605       GCAATATATC AAATGCTTAA AGATATCATG GAGAACCCAT CCAAGTGGGG AGGAAGAAGA ATTCTCTTCA
pinus BE662323      .......... .......... .......... .......... .......... .......... ..........
pinus TC44655       GCAATATATC AAATGTGTAA AGATATCATG GAGAATCCAT CCAAGTGGGA AGGAAGAAGA ATTCTTTTCA
pinaster BX252805   GCAATATATC AAATGTGTAA AGATATCATG GAGAATCCAT CCAAGTGGGA AGGAAGAAGA ATTCTTTTCA
pinaster Consensus  GCAGTATATC AAATGTGTAA AGATATCATG GAGAATCCAT CCAAGTGGGA AGGAAGAAGA ATTCTCTTCA
TC60605             GCAGTATATC AAATGTGTAA AGATATCATG GAGAATCCAT CCAAGTGGGA AGGAAGAAGA ATTCTCTTCA
TC60606             GCAGTATATC AAATGTGTAA AGATATCATG GAGAATCCAT CCAAGTGGGA AGGAAGAAGA ATTCTCTTCA

                    ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                            1270       1280       1290       1300       1310       1320       1330
oryza XM_468034.1   TACATACAGG TGGTCTTCTT GGTTTGTATG ATAAGGTTGA TGAGTTATCA TCTTTGAGTG GGAGTTGGCG
spruce CO484002     .......... .......... .......... .......... .......... .......... ..........
spruce TC8605       TTCACACTGG GGGGCTTCTT GGGATGTTTG ACAAAGTGCA GCAACTGCAA CCCCTGATTG GAAAATGGCA
pinus BE662323      .......... .......... .......... .......... .......... .......... ..........
pinus TC44655       TTCACTCTGG AGGACTTTTG GGGATGTTTG ACAAAGTGCA GCAGCTCCAG CCCCTGATTG GAAAATGGCA
pinaster BX252805   TTCACTCTGG AGGACTTTTG GGGATGTTTG ACAAAGTGCA GCAGCTCCAG CCCCTGATTG GAAAATGGCA
pinaster Consensus  TTCACTCTGG AGGACTTTTG GGGATGTTTG ACAAAGTGCA GCAGCTCCAG CCCCTGATTG GAAAATGGCA
TC60605             TTCACTCTGG AGGACTTTTG GGGATGTTTG ACAAAGTGCA GCAGCTCCAG CCCCTGATTG GAAAATGGCA
TC60606             TTCACTCTGG AGGACTTTTG GGGATGTTTG ACAAAGTGCA GCAGCTCCAG CCCCTGATTG GAAAATGGCA

                    ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                            1340       1350       1360       1370       1380       1390       1400
oryza XM_468034.1   CAGAATGGAT CTTGAAGAAT CTGTTCCACG CAAAGATGGA ACTGGTAAGA TGTTCTGATT GCTATGCT-C
spruce CO484002     .......... .......... .......... .......... .......... .......... ..........
spruce TC8605       GAGACTTAAG GTGGACGAGT CTATGTTTCA GGCAGATGGA ATAGGAAAAA TGTTTTAGAA ATAAAACAAT
```

```
pinus BE662323       .......... .......... .......... .......... .......... .......... ..........
pinus TC44655        GAGATTCAAG GTGGACGAAT CTATGTTTCA GGCAGATGGA ATAGGAAAAA AGTTTTAGAA ATAAAACATT
pinaster BX252805    GAGATTCAAG GTGGACGAAT CTATGTTTCA GGCAGATGGA ATAGGAAAAA TGTTTTAGAA ATAAAACATT
pinaster Consensus   GAGATTCAAG GTGGACGAAT CTATGTTTCA GGCAGATGGA ATAGGAAAAA TGTTTTAGAA ATAAAACATT
TC60605              GAGATTCAAG GTGGACGAAT CTATGTTTCA GGCAGATGGA ATAGGAAAAA TGTTTTAGAA ATAAAACATT
TC60606              GAGATTCAAG GTGGACGAAT CTATGTTTCA GGCAGATGGA ATAGGAAAAA AGTTTTAGAA ATAAAACATT

                     ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                             1410       1420       1430       1440       1450       1460       1470
oryza XM_468034.1    CAAAATTGCT TCCATTGTTG TTAGA-GGTT TCAGACATCT -CATTGTCTA AAAAGAGTAT TCGCCTCAAT
spruce CO484002      .......... .......... .......... .......... .......... .......... ..........
spruce TC8605        TAACAGTTCT TATTTAACTC TCTTT-GATA GGAAAGAACA ATAGTTAATG GATTGCACAA TCAAGTGCAT
pinus BE662323       .......... .......... .......... .......... .......... .......... ..........
pinus TC44655        TAACAATTCT TATTTAACTC TCTTTTGATA GGAAAGAGCA ATAGCTAGTG GATTACACAA TCAAATGCAT
pinaster BX252805    TAACAATTCT TATTTAACTC TCTTT-GATA GGAAAGAGCA ATAGCTAGTG GATTACACAA TCAAATGTAT
pinaster Consensus   TAACAATTCT TATTTAACTC TCTTTGATAG GAAAGAGCAA TAGTTAGTGG ATTACACAAT CAAATGCATT
TC60605              TAACAATTCT TATTTAACTC TCTTT-GATA GGAAAGAGCA ATAGTTAGTG GATTACACAA TCAAGTGCAT
TC60606              TAACAATTCT TATTTAACTC TCTTTTGATA GGAAAGAGCA ATAGTTAGTG GATTACACAA TCAAGTGCAT

                     ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                             1480       1490       1500       1510       1520       1530       1540
oryza XM_468034.1    TTGTACGA-C AAAATTCAAA GCAGAT--GG TTAATGTAAA ACACT-GTTT GTTATATTTA AAGAATAAAG
spruce CO484002      .......... .......... .......... .......... .......... .......... ..........
spruce TC8605        TTGTTGGATC AAAATGCCTT GGGAATTTGA CTTGTTTCAT ACTTTAGTTT GTACTATTTT TCTTGCCAAC
pinus BE662323       .......... .......... .......... .......... .......... .......... ..........
pinus TC44655        TTGTTGGCTC AAAACACCTT GGGACATTGG CTTGCTTCAT ACTTTAGTTT GTACTATTCT GCTTCCCAAA
pinaster BX252805    TTGTTGGCTC AAAACACCTT GGGACTTTGG CTTGCTTCAT ACTTTAGTTT GTACTATTCT GCTTCCCAAA
pinaster Consensus   TGTTGGCTCA AAACACCTTG GGACTTTGGC TTGCTTCATA CTTTAGTTTG TACTATTCTG CTTCCCAAAT
TC60605              TTGTTGGCTC AAAACACCTT GGGACATTGG CTTGCTTCAT ACTTTAGTTT GTACTATTCT GCTTCCCAAA
TC60606              TTGTTGGCTC AAAACACCTT GGGACATTGG CTTGCTTCAT ACTTTAGTTT GTACTATTCT GCTTCCCAAA

                     ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                             1550       1560       1570       1580       1590       1600       1610
oryza XM_468034.1    CATATATTTC TGCCATTTGG ---------- ---------- ---------- ---------- ----------
spruce CO484002      .......... .......... .......... .......... .......... .......... ..........
spruce TC8605        TCACCATCAA TGCTATTGGA AAATGGCAGA GATTTAAGGT GGACGAATCT ATGTTTCAGG AAAATGGAAT
pinus BE662323       .......... .......... .......... .......... .......... .......... ..........
pinus TC44655        TCACCATCAA TTCTATT--- -AATGGCAGA GATTTAAGGT TGACGAATCT ACATTTCAGG CAAATGGAAT
pinaster BX252805    TCACCATCAA TTCTATT--- -AATGGCAGA GATTTAAGGT TGACGAATCT ACATTTCAGG CAAATGGAAT
pinaster Consensus   CACCATCAAA ANN....... .......... .......... .......... .......... ..........
TC60605              TCACCATCAA TTCTATT--- -AATGGCAGA GATTTAAGGT TGACGAATCT ACATTTCAGG CAAATGGAAT
TC60606              TCACCATCAA TTCTATT--- -AATGGCAGA GATTTAAGGT TGACGAATCT ACATTTCAGG CAAATGGAAT

                     ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                             1620       1630       1640       1650       1660       1670       1680
oryza XM_468034.1    ---------- ---------- ---------- ---------- ---------- ---------- ----------
spruce CO484002      .......... .......... .......... .......... .......... .......... ..........
spruce TC8605        AGGAAAAATG TTTTAGAAAT AATACAATTA ACAGTTCTGA TTTAACTCTC TTTGATGGCA AAGAACGATA
pinus BE662323       .......... .......... .......... .......... .......... .......... ..........
pinus TC44655        AAGGAAATTG TTTTAGAAAT AAAACAATTA ACAGTTCTTA GATA-CTCTT TTTGATTGTA AAGAACATTA
pinaster BX252805    AGGGAAATTG TTTTAGAAAT AAAACAATTA ACAGTTCTTA GATA-CTCTT TTTGATTGTA AAGAACATTA
pinaster Consensus   .......... .......... .......... .......... .......... .......... ..........
TC60605              AGGGAAATTG TTTTAGAAAT AAAACAATTA ACAGTTCTTA GATA-CTCTT TTTGATTGTA AAGAACATTA
TC60606              AAGGAAATTG TTT-AGAAAT AAAACAATTA ACAGTTCTTA TATA-CTCTT TTTGATTGTA AAGAACATTA

                     ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                             1690       1700       1710       1720       1730       1740       1750
oryza XM_468034.1    ---------- ---------- ---------- --------.. .......... .......... ..........
spruce CO484002      .......... .......... .......... .......... .......... .......... ..........
spruce TC8605        GTTAATAGAT T-GCATGATC AAGTGCATTT GTTGGCTTG. .......... .......... ..........
pinus BE662323       .......... .......... .......... .......... .......... .......... ..........
pinus TC44655        ATTAATGGAT TTGCACAATC AAGTTTTTGT TGGCTCAAAC ACCATGGGAA TTTGACTTGT TTCATACTCT
pinaster BX252805    ATTGATGGAT TTGCACAATC AAGTTTTTGT TGGCTCAAAC ACCATGGGAA TTTGACTTGT TTCATACTCT
pinaster Consensus   .......... .......... .......... .......... .......... .......... ..........
TC60605              ATTGATGGAT TTGCACAATC AAGTTTTTGT TGGCTCAAAC ACCATGGGAA TTTGACTTGT TTCATACTCT
TC60606              GTTAATGGAT TTGCACAATC AAGTTTTTGT TGGCTC.... .......... .......... ..........

                     .

oryza XM_468034.1    .
spruce CO484002      .
spruce TC8605        .
pinus BE662323       .
pinus TC44655        A
pinaster BX252805    A
pinaster Consensus   .
TC60605              A
TC60606              .
```

Annexe Px-1

EPA : integrated system for EST collection Processing

Garcia, V., Barre, A., Mohellibi, N., Jacob, D., Chaumeil, P., Frigerio, JM., Plomion, C., de Daruvar, A.

Soumis à NAR.

EPA: integrated system for EST collection Processing

Virginie Garcia[1], Aurélien Barre[2], Nacer Mohellibi[2,3], Daniel Jacob[2], Philippe Chaumeil[3], Jean-Marc Frigerio[3], Christophe Plomion[3], Antoine de Daruvar[2*]

[1]UMR Physiologie et Biotechnologie Végétale, INRA - Université V. Segalen Bordeaux 2 - Université Bordeaux 1, 71 avenue E. Bourleaux, 33883 Villenave d'Ornon Cedex, France,
[2]Centre de Bioinformatique Bordeaux, UMR 5162 CNRS - Université V. Segalen Bordeaux 2, 146 rue Léo Saignat, 33076 Bordeaux Cedex, France,
[3]UMR BIOGECO, INRA - Université Bordeaux 1, Equipe de génétique, 69 route d'Arcachon, 33610 Cestas Cédex, France

* To whom correspondence should be addressed.

ABSTRACT

Summary: EPA is a Web-based environment for assembly and annotation of collections of Expressed Sequence Tags (EST). The user defines and customizes an analysis strategy by selecting and parameterizing a set of sequence analysis tools. Jobs are then automatically launched in order to process the sequences resulting from transcript reconstruction. The annotation interface integrates the output of the various tools, providing an information-rich, interactive environment. The system handles assembly updates due to expansions of the EST collection over time. It also provides the means to automatically take into account updates of the public databases that are used for sequence analysis.

Availability: http://cbi.labri.fr/outils/EPA/
Contact: antoine.daruvar@u-bordeaux2.fr
Supplementary information*:* http://cbi.labri.fr/outils/EPA/help/help.php

INTRODUCTION

Large scale Expressed Sequence Tag (EST) sequencing projects generate partial sequences for a large fraction of genes from a given organism (Adams *et al.*,1991). This is a cost-efficient and widely used strategy for gene discovery from organisms whose genome sequence is not yet available (Rudd, 2003). In a typical analysis, raw data (trace files) are converted into a non-redundant set of sequences which are then annotated. This process requires a succession of steps: raw sequence treatment, clustering, assembly and various *in silico* analyses necessary for the annotation.

Dedicated software components are publicly available for performing each step individually. Additional tools have been developed that provide a first level of integration of those components. This is the case of stackPACK™ (Miller *et al*., 1999; Burke *et al*., 1999) which combines the clustering and assembly steps and provides an infrastructure for data storage and retrieval. However, publicly available fully integrated systems remain scarce. PipeOneline (Ayoubi *et al.* 2002) and ESTAnnotator (Hotz-Wagenblatt *et al*., 2003) are good examples of such integrated systems. They allow fully automatic processing and annotation of EST data.

Automated annotation is undoubtedly convenient for a high throughput processing, but presents various drawbacks due to blind transfer of annotation based on an arbitrary fixed similarity threshold. Conversely, fully manual annotation is time-consuming and not adapted to large data set analysis.

EPA combines a high level of automation for the processing of sequences with a high quality annotation based on expert decision. Our system embeds several publicly available software, especially StackPACK™ which is used for transcript reconstruction and whose database scheme was extended for the storage and query of annotations. EPA handles the analysis of a collection of ESTs that can be expanded by adding batches of new sequences

over time. Since functional annotation largely relies on homology based inference from sequences present in public databases, the system also takes into account updates of those databases.

EPA Workflow

Starting a project

In EPA, data access, processing and annotation are structured around projects. A project starts with a collection of ESTs that can be progressively extended by successive addition of new sequences. Submission and access to the data are protected by a login/password system in order to ensure confidentiality and integrity of the data. Login and password can be obtained from http://cbi.labri.fr/outils/EPA/

Transcript reconstruction

ESTs are submitted either as trace files or FASTA formatted raw sequences with the corresponding quality files. In the former case an initial processing of the sequences (base-calling and trimming, vector masking, etc.) is launched automatically in order to produce the sequence and quality files. Masking of repeated sequences, clustering and assembly are then chained automatically to generate consensus sequences using stackPACK™ software. A gene index comprising the consensus sequences and singletons (ESTs which could not be joined with other sequences during the clustering process) is produced and the user receives by email a summary of the results of this assembly process. All sequences, consensi and singletons, then undergo two further steps: analysis and annotation.
The Figure 1, which shows the diagram of sequences status transition, provides an overview of EPA workflow.

Analysis step

The analysis is an automatic process using various bioinformatics methods to gather structural and functional evidence on the sequences in order to allow the next step of annotation *per se*. The user defines the analysis strategy interactively by selecting a set of bioinformatics analysis tools. The selection is made from a list of methods ranging from similarity searches to functional and structural domain predictions. Depending on the selected methods, the user is prompted to specify parameters such as the databases to be browsed in case of a similarity search. Once the analysis strategy is defined, all the corresponding jobs can be launched at once on all the sequences that need to be processed. Initially, this is done for all the consensi and singletons. Over time, this is typically necessary only for a subset of the sequences depending on various update events as described in Figure 1. When all the jobs are completed, the user receives an email inviting him to proceed with the annotation step.

Annotation step

The user assigns a functional annotation interactively relying on the results of the analysis which are brought together and displayed in a graphical interface. A mouse click allows the selection of a putative annotation out of a list extracted from the results of homology searches. This annotation can then be further edited. In addition, the interface provides access to the history of the previous annotations (changed or deleted) and alerts the user when inconsistency, potentially due to ESTs or misassembled consensi, is detected between a consensus and its ESTs.
For each annotation a confidence rate must be chosen between high, when the sequence is considered to be definitively annotated, or low, when the annotation remains to be confirmed. Existing annotations can be visualized, modified or deleted.
Free text annotation can be completed by a structured section based on normalized vocabularies such as Gene Ontology (Ashburner *et al.*, 2000), MIPS Functional Catalogue (Ruepp *et al.*, 2004), or KEGG Metabolic Pathway definition (Kanehisa, M. and Goto, S. 2000).
In case of an assembly update, the annotation is kept for the consensi that were not affected and discarded otherwise.

Project management

The entry page of each project displays dynamically computed information about the overall progress of the analysis and annotation processes. The user interface allows the querying of annotations according to keywords and the visualization of the distribution of consensi according to normalized vocabularies. If appropriate, the owner of the project can open a public access to the data: a link to the

list of public projects is available from EPA home page at http://cbi.labri.fr/outils/EPA/ Initially designed and developed for the analysis of ESTs from maritime pine, *Pinus pinaster* Ait., (Le Dantec *et al*, 2004), EPA was used for several projects on various organisms.

ACKNOWLEDGEMENTS

The Centre de Bioinformatique de Bordeaux belongs to the « Plate-forme Génomique Fonctionnelle Bordeaux » (http://www.pgfb.u-bordeaux2.fr/eng/index.html) which is supported by Région Aquitaine (grant n° 20030304002FA) and the Fonds Européen de Développement Régional (FEDER n° 2003227).

REFERENCES

Adams, M.D., et al. (1991) Complementary DNA sequencing: expressed sequence tags and human genome project. *Science*, **252,** 1651-1656.

Ashburner, M., et al.(2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet.*, **25**, 25-29.

Ayoubi, P., et al. (2002) PipeOnline 2.0: automated EST processing and functional data sorting. *Nucleic Acids Res*, **30,** 4761-4769.

Burke, J., Davison, D. and Hide, W. (1999) D2_cluster: A Validated Method for Clustering EST and Full-length cDNA Sequences. *Genome Res.*, **9,** 1135-1142.

Hotz-Wagenblatt, A., et al. (2003) ESTAnnotator: A tool for high throughput EST annotation. *Nucleic Acids Res.*, **31**, 3716-3719.

Kanehisa, M., Goto, S., (2000) KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.*, **28**, 27-30.

Le Dantec, D. Chagne, D. Pot, O. Cantin, P. Garnier-Géré, F. Bedon, J. M. Frigerio, P. Chaumeil, P. Léger, V. Garcia, F. Laigret, A. De Daruvar. and C. Plomion (2004) Plant Mol Biol, 54, 461-470

Miller, R.T., et al. (1999) A comprehensive approach to clustering of expressed human gene sequence: The Sequence Tag Alignment and Consensus Knowledgebase. *Genome Res.*, **9,** 1143-1155.

Rudd, S. (2003) Expressed sequence tags: alternative or complement to whole genome sequences? *Trends Plant Sci.*, **8,** 321-329.

Ruepp, A., et al. (2004) The FunCat, a functional annotation scheme for systematic classification of proteins from whole genomes. *Nucleic Acids Res.*, **32**, 5539-5545.

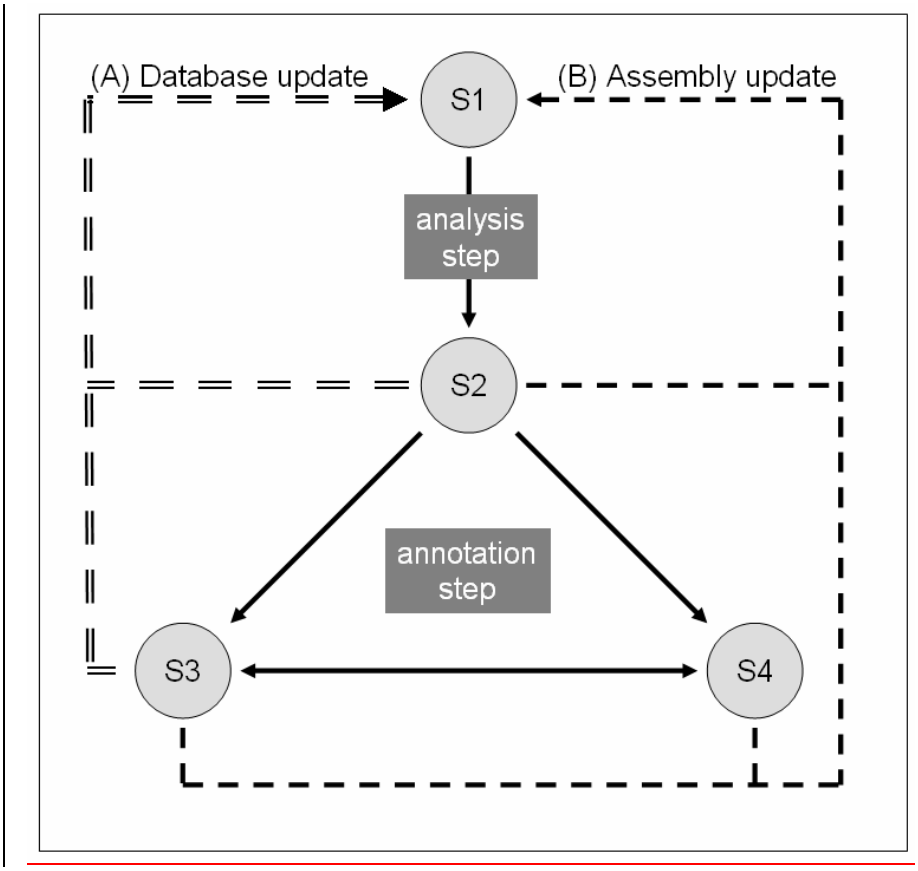


Figure 1: Diagram of sequences status transition. At any time, the status of a sequence (consensus and singleton) can be one of the followings: S1 (sequence needs to undergo the analysis step), S2 (sequence is analyzed and needs to undergo the annotation step), S3 (sequence is annotated with a low confidence rate) or S4 (sequence is annotated with high confidence rate). At the end of the assembly step, all the sequences are in S1 status and when all of them are annotated, they are in S3 or S4 status. Standard transitions between those status (plain line) are determined by the analysis step (transition from S1 to S2) or by the annotation step (transitions between S2, S3 and S4). The status can also be affected by two external events. (A) Public databases update (double dotted line): All the sequences in S2 and S3 go back to S1 in order to undergo the analysis phase taking advantage of the availability of new data. Because of the high confidence rate, sequences in S4 are not concerned and keep their annotation. (B) Assembly update resulting from the addition of a new batch of ESTs (single dotted line): All new singletons, new consensi and modified consensi (consensi whose EST composition has been modified) go back to S1 status. In this latter case, a history of annotation is saved.

Annexe Px-2

Automated SNP detection in expressed sequenced tags: statistical considerations and application to maritime pine sequences

Le Dantec, L., Chagné, D., Pot, D., Cantin, O., Garnier-Géré, P., Bedon, F., Frigerio, JM., Chaumeil, P., Léger, P., Garcia, V., Laigret, F., de Daruvar, A., Plomion, C.

Plant Molecular Biology 54 : 461-470, 2004.

Plant Molecular Biology **54:** 461–470, 2004.
© 2004 *Kluwer Academic Publishers. Printed in the Netherlands.*

Automated SNP detection in expressed sequence tags: statistical considerations and application to maritime pine sequences

Loïck Le Dantec[1,2,†], David Chagné[3,†,*], David Pot[3], Olivier Cantin[1,2], Pauline Garnier-Géré[3], Frank Bedon[3], Jean-Marc Frigerio[3], Philippe Chaumeil[3], Patrick Léger[3], Virginie Garcia[4], Frédéric Laigret[1], Antoine de Daruvar[2] and Christophe Plomion[3,*]

[1]*Unité de Recherche sur les Espèces Fruitières et la Vigne, INRA, 71 avenue Edouard Bourlaux, BP 81, 33883 Villenave d'Ornon Cedex, France;* [2]*Centre de Bioinformatique de Bordeaux, Université V. Segalen Bordeaux 2, 146 rue Léo Saignat, 33076 Bordeaux Cédex, France;* [3]*UMR 1202 BIOGECO, INRA, Equipe de Génétique, 69 route d'Arcachon, 33612 Cestas Cedex, France; (*authors for correspondence; e-mails: chagne@pierroton.inra.fr, plomion@pierroton.inra.fr);* [4]*UMR 619 Physiologie et Biotechnologie Végétales, INRA, IBVM, 71 Avenue Edouard Bourlaux, BP 81, 33883 Villenave d'Ornon Cedex, France;* [†]*These authors have contributed equally*

Received 7 January 2004; accepted in revised form 8 March 2004

Key words: EST, *in silico* detection, maritime pine, SNP

Abstract

We developed an automated pipeline for the detection of single nucleotide polymorphisms (SNPs) in expressed sequence tag (EST) data sets, by combining three DNA sequence analysis programs: *Phred*, *Phrap* and *PolyBayes*. This application requires access to the individual electrophoregram traces. First, a reference set of 65 SNPs was obtained from the sequencing of 30 gametes in 13 maritime pine (*Pinus pinaster* Ait.) gene fragments (6671 bp), resulting in a frequency of 1 SNP every 102.6 bp. Second, parameters of the three programs were optimized in order to retrieve as many true SNPs, while keeping the rate of false positive as low as possible. Overall, the efficiency of detection of true SNPs was 83.1%. However, this rate varied largely as a function of the rare SNP allele frequency: down to 41% for rare SNP alleles (frequency $< 10\%$), up to 98% for allele frequencies above 10%. Third, the detection method was applied to the 18 498 assembled maritime pine (*Pinus pinaster* Ait.) ESTs, allowing to identify a total of 1400 candidate SNPs, in contigs containing between 4 and 20 sequence reads. These genetic resources, described for the first time in a forest tree species, were made available at http://www.pierroton.inra/genetics/Pinesnps. We also derived an analytical expression for the SNP detection probability as a function of the SNP allele frequency, the number of haploid genomes used to generate the EST sequence database, and the sample size of the contigs considered for SNP detection. The frequency of the SNP allele was shown to be the main factor influencing the probability of SNP detection.

Introduction

Single nucleotide polymorphisms (SNPs) are the most frequent form of DNA variation in the human genome (Sachidanandam *et al.*, 2001). They have recently become the marker of choice in genetic analyses due to their abundance and stability compared to tandem repeat sequences (microsatellites) (Gray *et al.*, 2000). SNPs have been mainly developed and used in human genetics for complex diseases mapping and in association studies (Kruglyak, 1997; Collins *et al.*, 1999; Emahazion *et al.*, 2001), pharmacogenomics (Riley *et al.*, 2000), forensic studies (Wilson *et al.*, 1995; Andreasson *et al.*, 2002) and inference of population history (Brumfield *et al.*, 2003). In plants,

information on the frequency, nature, and distribution of SNPs is still in its infancy. Only in recent years, SNPs have been developed for a handful of species such as maize (Thornsberry *et al*., 2001; Ching *et al*., 2002), Arabidopsis (Cho *et al*., 1999; Nordborg *et al*., 2002), sugarcane (Grivet *et al*., 2003) and wheat (Somers *et al*., 2003).

The most classical way to identify SNP polymorphism is direct sequencing of amplicons of candidate genes from a set of individuals that represent the diversity in the population of interest. The drawback of this approach is the time and money investments. An alternative method takes advantage of the redundancy of gene sequences generated by expressed sequence tag (EST) sequencing programs (Picoult-Newberg *et al*., 1999; Useche *et al*., 2001; Batley *et al*., 2003; Kim *et al*., 2003). The development of ESTs has been initiated for many agronomically important plant species (Rounsley *et al*., 1998) and has generated a huge amount of data. It is now possible to use these data to search for DNA variation, providing enough individuals are included in the cDNA libraries. Mining SNPs from EST sequences generally involves the following steps: treatment of raw EST sequences (trace files), clustering, assembly, and detection of SNPs in aligned sequence data. Several bioinformatic tools have been developed for *in silico* SNP detection. They are based on sequence trace file analysis to filter out sequence errors. Among them, *PolyBayes* (Marth *et al*., 1999) is a software that applies a Bayesian inference to calculate the probability of a given site being polymorphic within aligned EST sequences, and *PolyPhred* (Nickerson *et al*., 1997) is a program that automatically detects the presence of SNP in heterozygotes by fluorescence-based sequencing of PCR products.

Other tools for SNP detection are based on a conservative approach, and use redundancy of SNPs in EST sequences combined with the co-segregation of haplotypes to differentiate between ''true'' SNPs and sequencing errors (Barker *et al*., 2003; Kota *et al*., in press).

Here, we present a bioinformatic approach to the automated detection of SNPs in EST sequences using the combination of *Phred*, *Phrap* and *PolyBayes* softwares. Rates of true SNP recovery and false positive sites were estimated using a reference set of true SNPs obtained from the sequencing of 13 maritime pine gene fragments representing a total of 6.6 kb. SNPs were then searched in the maritime pine ESTs, providing a valuable resource of DNA markers for this economically and ecologically important forest tree species in south-western Europe.

Materials and methods

SNP discovery pipeline

An automated system for SNP discovery was developed using a combination of three publicly available programs: *Phred* (Ewing and Green 1998; Ewing *et al*., 1998), *Phrap* (www.phrap.org) and *PolyBayes* (Marth *et al*., 1999). *Phred* makes the base identification and assigns a quality score to each position in a sequence. *Phrap* assembles the reads into contigs. *PolyBayes* is an engine using Bayesian inference methods, which determines, for each site within a multiple alignment, the probability of the site being polymorphic. This process takes into account the depth of coverage, the base quality values of the sequences, and the *a priori* expected rate of polymorphic sites in the considered species. Prior to its use for SNP mining in an EST dataset, the accuracy of the automatic detection method was tested with a reference set of true SNPs.

Development of a reference set of true SNPs

Sixty-five SNPs (Table 1) were used as a reference set of true SNPs. These SNPs were deduced from the expert visual inspection of maritime pine (*Pinus pinaster* Ait.) nucleotide sequences, resulting from the amplification of 13 DNA fragments in 30 megagametophytes (haploid tissue surrounding the embryo) representing the whole natural range of the species. Details about plant material, DNA extraction, primer design, PCR amplification, DNA sequencing and electrophoregram analysis are presented elsewhere (Le Provost *et al*., 2003). The use of megagametophytes lowered the risk of confusing polymorphism at a unique locus with differences between paralogous loci. With haploid tissue, amplification of two or even more members of a gene family would have been easily detected by the observations of double peaks in the sequences. The risk of confusing true SNPs with sequencing errors was also minimized by a systematic visual inspection (referred to as method A

Table 1. Result of SNP detection in the 13 gene fragments of the reference set using the visual inspection (A) and the automated method (B): gene identification, sequenced fragment size, comparison between the two methods, efficiency of the automated method and number of false positives (SNP detected by the method B only).

Gene identification	Fragment size (bp)	True SNPs detected by method A	dbSNP accession	SNPs detected by automated method B	Shared SNPs between A and B	Number of true SNPs not detected by method B	Number of false positive
			ss16208982; 8985–8987; 8991;8994–8998;9000				
C4H	539	10	8998;9000	10	10	0	0
CAD-F1	537	4	ss16209001; 9005;9008;9010	3	3	1	0
CAD-F2	523	4	ss16209011;9012;9015; 9016	4	3	1	1
PAL	547	9	ss16209020–9025;9028; 9029;9032	7	7	2	0
Korrigan	937	5	ss12709589–ss12709593	7	5	0	2
GRP	479	9	ss12709575–ss12709585	8	8	1	0
MYB-like TF	494	2	ss12709586–ss12709587	3	2	0	1
ACC oxydase	270	1	ss12709588	1	1	0	0
CesA7	553	2	ss12709573–ss12709574	2	2	0	0
CesA4	489	1	ss12709572	0	0	1	0
CesA3	490	7	ss12709565–ss12709571	6	6	1	0
CCoAOmT	492	7	ss16209076–ss16209082	3	3	4	0
AGP	321	4	ss16209070;9071;9073; 9074	4	4	0	0
Total	6671	65		58	54	11	4

Gene ID are as follows: C4H: *trans*-cinnamate-4-monooxygenase; CAD: cinnamyl alcohol dehydrogenase (two fragments F1 and F2); PAL: phenylalanine ammonia lyase; Korrigan: endo-1,4-β-D-glucanase; GRP: glycine-rich protein; MYB-like TF: MYB-like transcription factor; ACC oxidase: amino-cyclopropane-carboxylic acid oxidase; $CesA_n$: cellulose synthase; CCoAOMT: caffeoyl CoA *O*-methyltransferase; AGP: arabinogalactan protein.

in Table 1) of the electrophoregrams using *Sequencher* (GeneCode, Ann Arbor, MI, USA). This reference set was used as a positive control to test the accuracy of the automated method (referred to as method B in the next sections).

Recovery of true SNPs using the automated method

For each gene fragment, a file containing all sequence traces was used for *Phred-Phrap* analysis through a web interface generated using the *PISE* package (Letondal, 2001). Output files were accessible from a dynamically generated web page. *Polybayes* was then run on the ACE file generated by *Phrap*. *PolyBayes* generated two output files: a result file in a text format and a modified ACE file with polymorphism 'tags' that were viewed by *Consed* (Gordon *et al*., 1998, http://www.phrap.org/consed/consed.html). *Consed* displays the assembly and the associated tags along with the multiple alignment of the chromatograms. *Xml* files used by *PISE* as well as *Perl* scripts to create the directory structures required for the functioning of *Phred*/*Phrap*/*PolyBayes*/*Consed* are available upon request to the first author.

The efficiency of the pipeline to detect true SNPs was tested using the reference set described above. Several parameters of *Phrap* (i.e. retain exact duplicate reads, relaxed assembly stringency set to 10) and *Polybayes* (i.e. no paralog filtering, see Appendix A) were optimised by comparing the automated pipeline output of the reference set to the known true sets of SNPs. This combination of parameters allowed a maximum of true SNPs to be detected while keeping the rate of false positive sites as low as possible.

SNP detection in the maritime pine expressed sequence tag dataset

For SNP discovery in maritime pine ESTs with method B, a total of 28 128 plasmid clones were

available. These were derived from the following five cDNA libraries: one cDNA library constructed using differentiating xylem collected on four genotypes of the Corsican provenance (Cantón *et al*., in press), and four cDNA libraries constructed with root and needle tissues of 240 drought stressed and 120 well watered seedlings of the Aquitaine provenance (Frigerio *et al*., in press). EST data were assembled into contigs using a semi-automatic pipeline for EST processing and annotation (V. Garcia, unpublished), mainly based on *StackPACK* (Christoffels *et al*., 2001). For each contig, a directory structure (directory with the name of the contig and the following subdirectories: edit_dir, chromat_dir, phd_dir, poly_dir) required for the functioning of *Phred-Phrap* and *PolyBayes* was created. Chromatograms for all ESTs in each contig were first copied from the chromatogram repository into the associated *chromat_dir* directory.

The ACE files resulting from *Phred-Phrap* analysis were then passed to the SNP detection program *PolyBayes*. Based on the reference SNP dataset, one SNP every 100 bp was used as input value. *PolyBayes* output files were parsed to collect the most relevant SNP information. Functional annotation of the consensus sequence for each contig analysed was extracted from the EST database. Finally, the data were displayed on a dynamically generated HTML page with links to the *Phred-Phrap* and *PolyBayes* output files for further examination and SNP detection results for each contig (i.e. functional annotation, nucleotide variation, SNP probability P_{SNP}, SNP position on the consensus sequence, number of SNPs with $P_{SNP} > 0.99$ and contig length). A brief summary displayed the total number of SNP with $P_{SNP} > 0.99$, the number of contigs with no SNP detected and the total length of analysed contigs. Several home made PHP (*PHP Hypertext Processor*) scripts (available upon request from the first author) were used to combine *Phred-Phrap* and *Polybayes* and also to make interfaces for user query, EST database access, output files parsing and generate HTML result pages.

SNP detection probability in EST data

To estimate the probability of SNP detection in an already assembled EST data set, two main parameters have to be taken into account: (i) the number of haploid genomes 'g' segregating in the EST database (e.g. $g = 10$ for 5 diploid genotypes used to construct the cDNA library), and (ii) the number of ESTs 'e' in the contigs considered for SNP detection. Let X_A be the random variable corresponding to the number of allele A in the 'e' ESTs, Y_A the random variable corresponding to the number of allele A in the 'g' genome, and P_A the frequency of one the SNP allele in the population in which these genotypes are sampled (SNPs are considered as diallelic; Brookes, 1999). The probability of SNP detection (i.e. the two alleles detected in a given contig) is equal to $1 - [P(X_A = 0,\ or,\ X_A = e)] = 1 - P(X_A = 0) - P(X_A = e)$, with $P(X_A = e)$ the probability that all ESTs carries the A allele, and $P(X_A = 0)$ the probability that all ESTs carries the alternative allele. To compute these two probabilities we used the Bayse theorem as follows:

$$P(X_A = 0) = \sum_{y_A=1}^{g} P(Y_A = y_A) \times P(X_A = 0/Y_A = y_A)$$

and

$$P(X_A = e) = \sum_{y_A=1}^{g} P(Y_A = y_A) \times P(X_A = e/Y_A = y_A)$$

with

$$P(Y_A = y_A) = \beta(y_A, g, P_A) = \frac{g! P_A^{y_A} (1 - P_A)^{g - y_A}}{(g - y_A)! y_A!} \text{ and}$$

$$P(X_A = x_A/Y_A = y_A) = \beta\left(x_A, e, \frac{y_A}{g}\right)$$

i.e.,

$$P(X_A = 0/Y_A = y_A) = \left[\frac{1 - y_A}{g}\right]^e \quad \text{and}$$

$$P(X_A = e/Y_A = y_A) = \left[\frac{y_A}{g}\right]^e$$

The probability distribution β corresponds to the binomial distribution.

The probability of SNP detection in a contig was finally computed for different values of the three parameters P_A, g and e. It should be noted that the probability of SNP detection as estimated above depends on several assumptions regarding: (i) the quality of the sequences, (ii) the constitution of the contigs (mixture of paralogs, sequence

length homogeneity), and (iii) the efficiency of the SNP detection algorithm. All were assumed optimal, i.e. high sequence quality, absence of paralog within each contig, DNA fragment with identical length, polymorphism detection algorithm independent of the frequency of the SNP alleles in the EST dataset.

Results and discussion

Validation of the automated method for SNP detection

Sixty-five SNPs could be detected by visual inspection of 6671 bp in 13 gene fragments, resulting in a frequency of one SNP every 102.6 bp, which is comparable to the results obtained in other plant species, such as maize where one SNP per 100 bp were detected in contigs comprising at least 20 ESTs (Batley *et al*., 2003). In the reference sequence dataset 58 SNPs were detected using the automated system (method B, Table 1). These 58 SNPs were checked by visualizing the sequence chromatograms using *Consed*, and compared to the 65 true SNPs previously detected by visual inspection of aligned sequences (method A, Table 1). While 54 SNPs corresponded to true SNPs, indicating an efficiency of discovery of true SNPs of 83.1%, 4 SNPs among 58 (6.9%) corresponded to false positive sites. Visual inspection clearly showed that they were due to weaker sequence quality inducing wrong base calling. Because this occurred at the 3′ and 5′ ends of the sequences it is anticipated that most of these artifacts could be circumvented using trimmed EST sequences based on optimized *Phred* scores.

More critical was the inability of the pipeline to detect 11 true SNPs. Interestingly, most of them had a low allele frequency ($P < 10\%$) in the reference dataset. Table 2 shows the results obtained with both the manual and the automated methods, for SNP allele frequencies above or below 10%: 17 among the 65 true SNPs had rare alleles ($P < 10\%$), from which only 7 were found using the automated method, indicating a 41% efficiency rate for rare allele SNP detection, whereas 98% of the SNPs with allele frequency greater than 10% were found by the automated method. This result clearly shows the limits of the automated method for detecting rare SNP alleles, and this should be taken into account for the use of the pipeline in EST database.

Table 2. Efficiency of the automated method for the detection of true SNPs in relation with the SNP allele frequency (P_A).

Rare allele frequency	True SNPs (A)	Shared SNPs (A and B)	% recovery of true SNPs
$P < 10\%$	17	7	41
$P > 10\%$	48	47	98
Total	65	54	83.1

A and B referred to the automated and visual method for SNP detection, respectively.

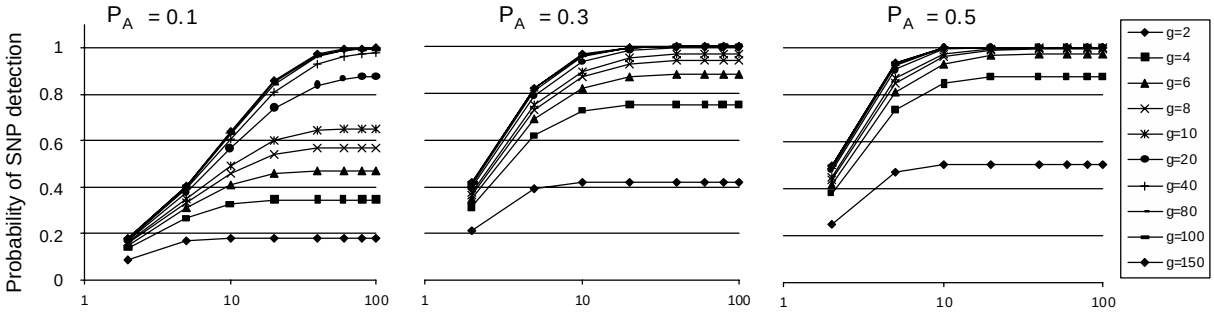


Figure 1. SNP detection probability in relation to (i) the number of alleles in the database ($2 < g < 150$), (ii) the number of sequences in a contig ($2 < e < 100$) and (iii) the rare SNP allele frequency in the population ($0.1 < P_A < 0.5$).

SNP detection probability

Figure 1 shows the probability of SNP detection according to (i) the number of ESTs '*e*' in a contig considered for SNP detection ($e = 2$–100), (ii) the frequency '*P*' of one SNP allele in the population ($P = 0.1$, 0.3 and 0.5), and (iii) the number of haploid genomes '*g*' segregating in the EST database ($g = 2$–150). A 95% SNP detection probability was achieved with the following combination of parameters: $P = 0.1$, $g > 80$, $e > 40$, $P = 0.3$, $g > 20$ and $e > 20$, $P = 0.5$, $g > 8$, $e > 10$, again showing the difficulty to detect low frequency SNP alleles.

These three parameters were taken as independent variables in a multiple linear regression analysis in order to test their effect on SNP detection probabilities. All parameters had a strong effect on the ability to recover SNPs (overall determination coefficient of 90%; Table 3). The factor explaining the highest proportion of variation (43%) in the probability of SNP detection was the SNP allele frequency; the number of ESTs in the contig still explaining 33%, and finally the number of haploid genomes segregating in the database was the less limiting factor.

Recovery of true SNPs from a reference data set using ESTs

For three genes of the reference dataset, AGP, CCoAOMT and GRP, 19, 19 and 21 maritime pine ESTs were available, respectively. This made it possible to validate the efficiency of the pipeline to detect true SNPs in EST data. The four true SNPs detected in AGP were found using the automated method. For CCoAOMT and GRP, one SNP out of three, and six SNPs out of nine were detected in the EST sequences, respectively. Interestingly, the missing SNPs corresponded to rare alleles in the Aquitaine provenance that was the main source of mRNA for the construction of cDNA libraries. Again, this result demonstrates the limitation of detecting rare SNP alleles in EST datasets, even in contig size containing about 20 reads.

Automated SNP detection in maritime pine ESTs

Untreatable chromatograms and stringent quality checking led us to eliminate 9630 sequences from the whole set of ESTs, resulting in a total of 18 498 ESTs used in the clustering and assembly steps. These ESTs had an average length of 480 bp and included 89% of bases with a quality score greater or equal to 20 (99% per base accuracy according to *Phred* definition). As a result of clustering and assembly, 13 447 sequences were members of one of the 2697 multi-sequence clusters found, whereas 5051 ESTs remained as singletons (Table 4). Thus a gene index of 7748 different sequences was created, assuming that singletons corresponded to unique transcripts.

Considering that '*g*' was not a limiting factor in our EST dataset (see the Materials and methods section), the SNP detection probability, P_{SNP}, corresponding to a contig size of $e = 20$ ESTs was thus 0.86, 1 and 1 for a rare SNP allele frequency of $P = 0.1$, 0.3 and 0.5, respectively. If 10 reads were considered, P_{SNP} dropped to 0.65 for $P = 0.1$. In the whole EST database, the number of contigs containing at least one SNP was recorded for e values increasing from 2 to 100 (Figure 2). SNP abundance increased with increasing contig sample size and almost all contigs were found to contain at least one SNP for a number of sequence reads $e \geq 20$. More interesting was the mean number of SNPs per contig in relation to the contig size. The abundance of SNPs per contig remained low for $e < 20$ and then increased exponentially with e (Figure 3, dotted line), reaching a value of one SNP every 11 bp in the two contigs containing more than 100 reads.

Table 3. Multiple regression on SNP detection probability. P_A, *g* and *e* were used as independant variables.

Source of variation	df	SS	*F*-value	*P*-value	r^2 (%)
p	1	3.711962	82.66597	0	42.62
g	1	2.121951	47.25617	5.51E−11	24.36
e	1	2.876731	64.06526	5.44E−14	33.02

P_A: SNP allele frequency; *g*: number of haploid genomes in the database; *e*: number of ESTs in a contig; r^2 is the determination coefficient.

Table 4. Cluster profile of maritime pine ESTs resulting from the *StackPACK* analysis.

Number of ESTs per contig	Number of contigs	Total length (bp)
4	286	251 108
5	166	152 166
6	113	112 073
7	71	74 002
8	49	54 314
9	41	43 579
10	31	30 482
11	29	31 958
12	23	27 779
13	17	17 766
14	14	14 102
15	9	11 915
16	15	17 252
17	7	7306
18	6	5751
19	9	11 387
20	12	15 539
21	5	6021
22	4	5872
23	6	5449
24	6	8178
25	5	5262
26	2	1490
27	1	1583
28	3	4130
29	5	5602
31	1	524
32	3	2754
33	3	2196
37	1	1331
39	1	1597
40	1	1759
41	1	3264
44	1	1763
49	1	1586
50	1	1353
51	1	2460
53	1	2535
55	1	607
59	1	863
60	1	2300
65	1	2155
78	1	728
80	1	1392
154	1	740
179	1	765

Such disproportionate number of SNPs was compared with the expected number of SNPs obtained using the SNP frequency observed in the reference set (i.e. one SNP per 100 bp) and the total length in bp used for the detection (Figure 3, plain line). To explain the striking difference between the observed and expected abundance of SNPs for contig size above 20 ESTs, it could be first hypothesized that rare SNP alleles were only detected in contigs containing a high number of sequence reads. However, since the detection of SNPs in *Polybayes* is based on sequence quality, it is more likely that such difference resulted from the accumulation of reverse transcription errors (Bebenek *et al*., 1993) rather than sequencing errors *per se*. Indeed, reverse transcriptase reaction always lead to good quality (high phred score) sequence, compared to sequencing errors. Taking into account either multiple occurrences of detected SNPs (as proposed by Batley *et al*. (2003)) or the co-segregation of SNP alleles in the sequence alignment (as proposed by Kota *et al*., in press) would certainly help to increase the accuracy of true SNP detection and distinguish between SNPs and errors. Besides this problem, the presence of paralogous sequences within large contigs (Batley *et al*., 2003), can not be excluded. This might be particularly the case for pine genes whose structure on complex families has been described (Kinlaw and Neale, 1997).

As a conclusion, it was decided to ignore those contigs with more than 20 sequence reads for the automatic detection of SNPs in the maritime pine EST dataset. On the other hand, a minimum of 4 reads was considered as a minimum for SNP mining. It should be noted that rare SNP alleles were probably overlooked for contigs represented by a small number of sequences (Figure 1).

The results of the automated SNP search are summarized in Table 5 and were obtained within 54 min on a 3.06 GHz Intel Pentium IV PC with 512 MB RAM running Mandrake Linux 9.1). A total of 1400 SNPs (*Polybayes* $P_{SNP} > 0.99$) were found in a set of 940 contigs containing between 4 and 20 reads and representing 942 216 bp in total length. Five hundred and sixty-eight contigs (60%) did not have any SNP. All the detected SNPs were biallelic, and more transitions (818) than transversions (582) were obtained, which agrees with other results obtained in previous SNP discovery programs (Picoult-Newberg *et al*., 1999; Batley *et al*., 2003). Overall, SNP abundance was one SNP per 660 bp, ranging from SNP frequencies of one per 1457 bp of aligned sequences for 4-read contigs, to one per 137 bp of aligned sequences for 20-read contigs.

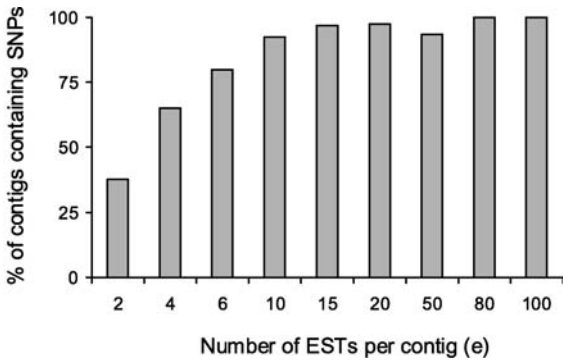


Figure 2. Proportion of contigs containing at least one SNP, in relation with the number of ESTs per contig $e = 2$–100.

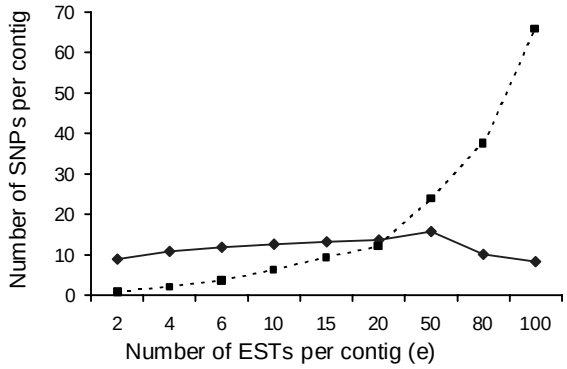


Figure 3. Comparison between the quantity of SNP detected by the automated method and the quantity expected considering the SNP frequency in the reference set (1 SNP/90 bp) and the total length of the sequences used.

Table 5. SNP detection in maritime pine ESTs with a minimum and a maximum number of 4 and 20 ESTs per contig: variant type (transition vs. transversion) and SNP frequency.

	Number of contigs $4 < e < 20$	940
	Total length of the sequences	924 216 bp
	Number of contig with no SNP	568
	SNP frequency	1 SNP/660 bp
	Number of SNP detected ($P_{SNP} > 0.99$)	1400
	More than 3 alleles	0
Transition	ag	374
	ct	444
	Total	818
Transversion	ac	148
	at	167
	cg	122
	gt	145
	Total	582

Conclusion

This study demonstrates the usefulness of bioinformatics-based sequence analysis tools to extend the scope of EST data in the field of marker development. A bioinformatic pipeline was developed and allowed for the first time the detection of SNPs in a EST dataset of a forest tree species. By comparing automated and visual methods for SNP mining on a reference set of true SNPs, we showed the intrinsic power of the developed engine: i.e. 83.1% of all the SNPs and 98% of the "non-rare" SNP alleles could be detected. From the calculation of SNP detection probabilities, it was clearly shown that the efficiency to uncover true SNPs dropped drastically for rare SNP alleles ($P < 0.1$). It was also shown that considering more than 20 reads per contig considerably increased the number of false SNPs, likely resulting from random errors accumulating during the sequencing step and/or the mixture of paralogous loci.

Finally, most of the SNPs detected in this study, for contig size ranging from 4 to 20, were found in known function genes (data not shown). Such 'functional SNPs' will provide a valuable source of molecular markers for genetic mapping (Chagné *et al*., 2002), comparative mapping in pines (Chagné *et al*., 2003), nucleotide diversity analysis, and association studies, four topics that are currently explored in the frame of the maritime pine genomics program.

Acknowledgements

LLD was funded by the Conseil Régional d'Aquitaine, DC was funded by the French Ministry for Research. Maritime pine ESTs were produced with the support of the Aquitaine Région, INRA (Lignome project) and European Union (GEMINI project, QLK5-CT1999-00942). The reference SNPs were generated within a project funded by DERF (no. 61.45.49/01). We thank Drs Gabor T. Marth, Mark D. Yandell, Ian F. Korf and Warren R. Gish who gave us access to the Polybayes software, and Brigitte Mangin (INRA Toulouse) for help in statistics.

References

Andreasson, H., Asp, A., Alderborn, A., Gyllensten, U. and Allen, M. 2002. Mitochondrial sequence analysis for forensic

identification using pyrosequencing technology. Biotechniques 32: 124–126.
Barker, G., Batley, J., O'Sullivan, H., Edwards, K.J. and Edwards, D. 2003. Redundancy based detection of sequence polymorphisms in expressed sequence tag data using autoSNP. Bioinformatics 19: 421–422.
Batley, J., Barker, G., O'Sullivan, H., Edwards, K.J. and Edwards, D. 2003. Mining for single nucleotide polymorphisms and insertions/deletions in maize expressed sequence tag data. Plant Physiol. 132: 84–91.
Bebenek, K., Abbotts, J., Wilson, S.H. and Kunkel, T.A. 1993. Error-prone polymerization by HIV-1 reverse transcriptase. J. Biol. Chem. 268: 10324–10334.
Brookes, A.J. 1999. The essence of SNPs. Gene 234: 177–186.
Brumfield, R.T., Beerli, P., Nickerson, D.A. and Edwards, S.V. 2003. The utility of single nucleotide polymorphisms in inferences of population history. Trends Ecol. Evol. 18: 249–256.
Cantón, F.R., Le Provost, G., Garcia, V., Barré, A., Frigério, J.-M., Paiva, J., Fevereiro, P., Ávila, C., Mouret, J.-F., de Daruvar, A., Cánovas, F.M. and Plomion C. in press. Transcriptome analysis of wood formation in maritime pine. In: Sustainable Forestry, Wood products & Biotechnology, BIOFOR Proceeding.
Chagné, D., Lalanne, C., Madur, D., Kumar, S., Frigerio, J.-M., Krier, C., Decroocq, S., Savoure, A., Bou-Dagher, K.-M., Bertocchi, E., Brach, J. and Plomion, C. 2002. A high density genetic map of maritime pine based on AFLPs. Ann. For. Sci. 59: 627–636.
Chagné, D., Brown, G., Lalanne, C., Madur, D., Pot, D., Neale, D. and Plomion, C. 2003. Comparative genome and QTL mapping between maritime and loblolly pines. Mol. Breeding 12: 185–195.
Ching, A., Caldwell, K.S., Jung, M., Dolan, M., Smith, O.S., Tingey, S., Morgante, M. and Rafalski, A.J. 2002. SNP frequency, haplotype structure and linkage disequilibrium in elite maize inbred lines. BMC Genet. 3: 1–19.
Cho, R.J., Mindrinos, M., Richards, D.R., Sapolsky, R.J., Anderson, M., Drenkard, E., Dewdney, J., Reuber, T.L., Stammers, M., Federspiel, N., Theologis, A., Yang, W.H., Hubbell, E., Au, M., Chung, E.Y., Lashkari, D., Lemieux, B., Dean, C., Lipshutz, R.J., Ausubel, F.M., Davis, R.W. and Oefner, P.J. 1999. Genome-wide mapping with biallelic markers in *Arabidopsis thaliana*. Nat. Genet. 23: 203–207.
Christoffels, A., van Gelder, A., Greyling, G., Miller, R., Hide, T. and Hide, W. 2001. STACK: sequence tag alignment and consensus knowledgebase. Nuc. Ac. Res. 29(1): 238–238.
Collins, F.S., Guyer, M.S. and Charkravarti, A. 1997. Variations on a theme: cataloging human DNA sequence variation. Science 278: 1580–1581.
Collins, A., Lonjou, C. and Morton, N.E. 1999. Genetic epidemiology of single-nucleotide polymorphisms. Proc. Natl. Acad. Sci. USA 96: 15173–15177.
Emahazion, T., Feuk, L., Jobs, M., Sawyer, S.L., Fredman, D., St Clair, D., Prince, J.A. and Brookes, A.J. 2001. SNP association studies in Alzheimer's disease highlight problems for complex disease analysis. Trends Genet. 17: 407–413.
Ewing, B. and Green, P. 1998. Base calling of automated sequencer traces using *Phred*. II. Error probabilities. Genome Res. 8: 186–194.
Ewing, B., Hiller, L.D., Wendl, M.C. and Green, P. 1998. Base calling of automated sequencer traces using *Phred*. II. Accuracy assessment. Genome Res. 8: 175–185.
Frigerio, J.-M., Dubos, C., Chaumeil, P., Salin, F., Garcia, V., Barré, A. and Plomion, C. in press. Using transcriptome analysis to identify osmotic stress candidate genes in maritime pine (*Pinus pinaster* Ait.). In: Sustainable Forestry, Wood products & Biotechnology, BIOFOR Proceeding.
Gordon, D., Abajian, C. and Green P. 1998. *Consed*: a graphical tool for sequence finishing. Genome Res. 8: 195–202.
Grivet, L., Glaszmann, J.-C., Vincentz, M., da Silva, F. and Arruda, P. 2003. ESTs as a source for sequence polymorphism discovery in sugarcane: example of the Adh genes. Theor. Appl. Genet. 106: 190–197.
Gray, I.C., Campbell, D.A. and Spurr, N.K. 2000. Single nucleotide polymorphisms as tools in human genetics. Hum. Mol. Genet. 9: 2403–2408.
Kim, *et al*. 2003.
Kinlaw and Neale, 1997.
Kota, R., Rudd, S., Facius, A., Kolesov, G., Thiel, T., Zhang, H., Stein, N., Mayer, K. and Graner, A. in press. Snipping polymorphisms from large EST collections in barley (*Hordeum vulgare* L.). Mol. Gen. Genome.
Kruglyak, L. 1997. The use of a genetic map of biallelic markers in linkage studies. Nat. Genet. 17: 21–24.
Le Provost, G., Paiva, J., Pot, D., Brach, J. and Plomion, C. 2003. Seasonal variation in transcript accumulation in wood forming tissues of maritime pine (*Pinus pinaster* Ait.) with emphasis on a cell wall glycin rich protein. Planta 217: 820–830.
Letondal, C. 2001. A Web interface generator for moleular biology programs in Unix, Bioinformatics 17: 73–82.
Marth, G.T., Korf, I., Yandell, M.D., Yeh, R.T., Gu, Z., Zakeri, H., Stitziel, N.O., Hillier, L., Kwok, P. and Gish, W.R. 1999. A general approach to single-nucleotide polymorphism discovery. Nat. Genet. 23: 452–456.
Nickerson, D.A., Tobe, V.O. and Taylor, S.L. 1997. *PolyPhred*: automating the detection and genotyping of single nucleotide substitutions using fluorescence-based resequencing. Nuc. Ac. Res. 25: 2745–2751.
Nordborg, M., Borevitz, J.O., Bergelson, J., Berry, C.C., Chory, J., Hagenblad, J., Kreitman, M., Maloof, J.N., Noyes, T., Oefner, P.J., Stahl, E.A. and Weigel, D. 2002. The extent of linkage disequilibrium in *Arabidopsis thaliana*. Nat. Genet. 30: 190–193.
Picoult-Newberg, L., Ideker, T.E., Pohl, M.G., Taylor, S.L., Donaldson, M.A., Nickerson, D.A. and Boyce-Jacino, M. 1999. Mining SNPs from EST databases. Genome Res. 9: 167–174.
Rounsley, S., Xiaoying, L. and Ketchum, K.A. 1998. Large-scale sequencing of plant genomes. Curr. Opin. Plant Biol. 1: 136–141.
Riley, J.H., Allan, C.J., Lai, E. and Roses, A. 2000. The use of single nucleotide polymorphisms in the isolation of common disease genes. Pharmacogenomics 1: 39–47.
Sachidanandam, R., Weissman, D., Schmidt, S.C., Kakol, J.M., Stein, L.D., Marth, G., Sherry, S., Mullikin, J.C., Mortimore, B.J., Willey, D.L., Hunt, S.E., Cole, C.G., Coggill, P.C., Rice, C.M., Ning, Z., Rogers, J., Bentley, D.R., Kwok, P.Y., Mardis, E.R., Yeh, R.T., Schultz, B., Cook, L., Davenport, R., Dante, M., Fulton, L.,

Hillier, L., Waterston, R.H., McPherson, J.D., Gilman, B., Schaffner, S., Van Etten, W.J., Reich, D., Higgins, J., Daly, M.J., Blumenstiel, B., Baldwin, J., Stange-Thomann, N., Zody, M.C., Linton, L., Lander, E.S. and Atshuler, D. 2001. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. Nature 409: 928–933.

Somers, D.L., Kirkpatrick, R., Moniwa, M. and Walsh, A. 2003. Mining single-nucleotide polymorphisms from hexaploid wheat ESTs. Genome 49: 431–437.

Thornsberry, J.M., Goodman, M.M., Doebley, J., Kresovich, S., Nielsen, D. and Buckler, E.S. 2001. Dwarf polymorphisms associate with variation in flowering time. Nat. Genet. 28: 286–289.

Useche, F.J., Gao, G., Harafey, M. and Rafalski, A. 2001. High-throughput identification, database storage and analysis of SNPs in EST sequences. Genome Inform. Ser. Workshop Genome Inform. 12: 194–203.

Wilson, M.R., Di Zinno, J.A., Polanskey, D., Replogle, J. and Budowle, B. 1995. Validation of mitochondrial DNA sequencing for forensic casework analysis. Int. J. Legal Med. 108: 68–74.

Appendix

Phred-Phrap parameters

$ctid -ace -maxgap 30 -repeat_stringency 0.95 -node_space 4 -node_seg 8 -tags -default_qual 15 -qual_show 20 -retain_duplicates -max_subclone_size 5000 -trim_penalty -2 -trim_score 20 -trim_qual 13 -confirm_length 8 -confirm_trim 1 -confirm_penalty -5 -confirm_score 30 -indexwordsize 10 -trim_start 0 -forcelevel 10 -bypasslevel 1 > phredphrap_log.

Polybayes parameters

-maskAmbiguousMatches -readPhdFiles -phdFilePathIn ../phd_dir -inputFormat ace -aceIn $ctid.fasta.screen.ace.1 -nofilterParalogs -noconsiderAnchor -noconsiderTemplateConsensus -preScreenSnpsMinimumBaseQuality 30 -priorPoly 0.01106 -priorPoly2 0.99666 -priorPoly3 0.00333 -priorPoly4 0.00001 -priorPolyAC 0.1666 -priorPolyAG 0.1666 -priorPolyAT 0.1666 -priorPolyCG 0.1666 -priorPolyACG 0.25 -priorPolyACT 0.25 -thresholdSnp 0.1 -maxTerms 50 -screenSnps -nowriteAlignment -nodisplayQuality -showUnAligned -writeAce -aceOut $ctid.polybayes.ace -lineLength 50 -nodebug -nomonitor -writeReport -reportOut $ctid.polybayes.report. out.

Résumé :

L'alimentation en eau constitue le principal facteur limitant la croissance, voire la survie des plantes. Les modèles climatiques prévoient pour les 50 à 100 années à venir une baisse des précipitations et des températures estivales accrues dans la moitié sud de la France. La durée de vie d'une forêt de pin maritime, de sa plantation jusqu'à la coupe d'exploitation est justement de 50 ans. Il est donc important de savoir si ces organismes pourront faire face à ces brusques changements climatiques ; en d'autres termes si les variétés améliorées plantées aujourd'hui pourront maintenir le niveau actuel de productivité dans un milieu plus pauvre en eau, et tolérer des épisodes de sécheresse intense. La capacité de ces organismes à faire face à ces perturbations brutales dépendra à la fois de leur plasticité phénotypique et de leur diversité génétique. Dans le cadre de cette thèse, nous avons étudié la plasticité moléculaire du système racinaire de jeunes plants de pin maritime élevés en milieu hydroponique et soumis à un stress osmotique par ajout de polyéthylène glycol. Un plan factoriel croisant deux écotypes (France et Maroc) par cinq niveaux de stress a permis d'analyser les réponses du transcriptome et du protéome à court et long terme. Nos investigations ont porté sur l'accumulation des transcrits de 7000 gènes et de 1200 protéines. L'analyse statistique des données a permis d'identifier des gènes dont la plasticité moléculaire est génétiquement contrôlée, révélant des stratégies de réponse différentes de chaque écotype. La valeur adaptative de ces gènes pourra alors être confirmée par l'interprétation des patrons de diversité nucléotidique de ces gènes candidats.

Mots-clefs : stress hydrique, sécheresse, *Pinus pinaster* Ait., microarrays, transcriptome, Expressed Sequence Tag, assemblage bioinformatique, annotation, expression des gènes

---

Summary:

Among abiotic stresses, water deficit is probably the most severe limiting factor of plant growth and development. In South-Western France climatologists predict a significant increase of temperature combined with lower rainfall during the growing period for the next century. Such modifications will place forest trees in the presence of new and sudden environmental constraints. It is therefore essential to know if these organisms will be able to face such climatic changes. In other words, if plantation of improved varieties will maintain the present level of productivity with lower water availability as well as sustain severe drought episodes. The capacity of these organisms to adapt to such perturbations will depend on both their phenotypic plasticity and their genetic diversity. In the frame of this thesis, we studied molecular plasticity of maritime pine seedlings grown in hydroponics and exposed to osmotic stress using polyethylene glycol as osmoticum. A factorial design crossing two ecotypes (France & Morocco) by different levels of stress (short & long term) was used to study the transcriptomic and proteomic responses of the root system. Our investigations focused on accumulation of 7,000 transcripts and 1,200 proteins. Statistical analysis revealed differentially expressed genes according to ecotype and condition effects. More interesting was genes whose molecular plasticity was genetically controlled, revealing ecotype dependent strategy to response to the imposed stresses. The adaptive significance of these genes could be confirmed by analysing their patterns of nucleotidic diversity with the main objective of detecting molecular signature of natural selection.

Keywords : water deficit, drought stress, *Pinus pinaster* Ait., microarrays, transcriptome, ESTs, bioinformatic clustering, annotation, gene expression