HAL
open science

Comparaison moléculaire des graines orthodoxes de Medicago truncatula et récalcitrantes de Castanospermum australe : une nouvelle approche pour comprendre l'acquisition de la tolérance à la dessiccation

Julien Delahaie

To cite this version:

Julien Delahaie. Comparaison moléculaire des graines orthodoxes de Medicago truncatula et récalcitrantes de Castanospermum australe : une nouvelle approche pour comprendre l'acquisition de la tolérance à la dessiccation. Biologie végétale. Université d'Angers, 2013. Français. NNT : . tel-00961568

HAL Id: tel-00961568

https://theses.hal.science/tel-00961568

Submitted on 20 Mar 2014

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.

Thèse de Doctorat

Julien DELAHAIE

Mémoire présenté en vue de l'obtention du
grade de Docteur de l'Université d'Angers
sous le label de L'Université Nantes Angers Le Mans

École doctorale : VENAM

Discipline : 66
Spécialité : Biologie cellulaire et moléculaire végétale
Unité de recherche : UMR 1345 Institut de Recherche en Horticulture et Semences

Soutenue le 27 novembre 2013

Thèse N° 1364

Comparaison moléculaire des graines orthodoxes de *Medicago truncatula* et récalcitrantes de *Castanospermum australe* : une nouvelle approche pour comprendre l'acquisition de la tolérance à la dessiccation

JURY

Rapporteurs :	**Mme. Françoise CORBINEAU**, Professeur, Université Pierre et Marie Curie (Paris 6) **M. Richard THOMPSON**, Directeur de recherches, INRA Dijon UMR 1347
Examinateurs :	**M. Thierry JOET,** Chargé de recherches, IRD Montpellier **M. Azeddine DRIOUICH**, Professeur, Université de Rouen **M. Anis LIMAMI**, Professeur, Université d'Angers
Directrice de Thèse :	**Julia BUITINK**, Directrice de recherches, UMR 1345 Institut de Recherche en Horticulture et Semences
Co-encadrant de Thèse :	**Jérôme BOVE**, Maître de conférences, Université d'Angers

L'Université Nantes Angers Le Mans

Remerciements

Tout d'abord un grand merci à mes encadrants bien sûr qui ont permis à cette thèse de se dérouler au mieux. ***Jérôme*** : mon père spirituel de la biologie moléculaire ! Tout a commencé en 2009 en master 1 avec une PCR dégénérée (déjà pour cloner ce fameux ABI3 de Castano). Plus de 4 ans après on a parcouru pas mal de chemin, on a revu ensemble je ne sais pas combien de manips, je ne sais pas combien de rapport, je ne sais pas combien de powerpoint et à chaque fois le travail en était bonifié. Même si tu as suivi de plus loin la fin de thèse je n'oublierai pas tous les précieux conseils que tu m'as apportés.

Julia : toujours disponible et surtout quand tu as pris la relève après le départ de Jérôme. Je ne sais pas comment tu as trouvé le temps de relire cette thèse mais tu l'as fait ! Je suis fier du travail accompli ici et je pense que c'est en grande partie parce que tu as su à la fois m'inciter à donner le meilleur et en même temps me rassurer et me redonner confiance quand j'en avais besoin.

Je remercie également les membres du jury qui ont accepté d'évaluer ces travaux de thèse ***Françoise Corbineau***, ***Richard Thompson***, ***Azeddine Driouich***, ***Thierry Joët*** et ***Anis Limami***. Je n'oublie pas non plus les membres du comité de suivi de thèse qui ont suivi l'évolution de ces travaux et permis de nous concentrer sur les objectifs principaux ***Maïté Vicré-Gibouin***, ***Stéphane Dussert*** et ***Jean Pierre Renou***. Je remercie enfin ***Angers Loire métropole*** qui a financé ce projet de thèse.

Passons maintenant à tous les amis de l'ARES. Il faut bien le dire, la thèse n'est pas une expérience facile mais elle devient beaucoup plus agréable en étant bien entouré. Heureusement, les gens qui peuplent l'ARES sont bien plus chaleureux que la façade du bâtiment.

Manu : ça fait quand même 4 ans qu'on se suit entre le stage de master 2 et la thèse et après tous ces repas au RU ensemble en tête à tête c'est une grosse page qui va se tourner ! Je garderai des souvenirs de franche rigolade (pas tous avouables sur une page de remerciement de thèse…) à l'ARES ou en Afrique du Sud. On est passé tous les deux par l'ascenseur émotionnel de la thèse et je pense qu'on s'est bien soutenu dans les moments de doute. Merci aussi pour le bon plan monitorat à l'IUT qui m'a fait découvrir les joies de l'enseignement ! Pas de crasse le dernier jour je ne veux pas voir une poule débarquer dans mon bureau sinon je te fais bouffer une poire aux amandes !

Boby : mon compagnon du bureau des espèces exotiques de l'ARES. Tu es arrivé avec l'étiquette du mec SIMPA et tout le temps passé ensemble a confirmé que tu étais encore mieux que ça ! Ton asperge aura marqué les esprits… autant que tes chansons venus d'ailleurs (on passera sur toutes les vieilles choses qu'on a pu chanter dans ce bureau).

Karima : tes clochettes et tes boulottes, ton accent chantant vont me manquer. By the way, la langue française est vraiment plus sympa dans ta bouche Karima ! Merci pour les données de puce Medicago !

David : mon partenaire Castano et ABI3 ! Tu l'auras eu finalement ce gène, enfin plutôt ces gènes… Merci pour ton aide, notamment sur ces fameuses transfos de racines, le chapitre IV serait resté bien creux sans toi. N'abuse pas trop du sucre maintenant que je ne serai plus là pour finir les gâteaux avec toi ! Merci aussi pour les données de puce *abi3* !

Pascale et ***Aurélia***, votre bureau est certainement l'endroit le plus convivial du labo. Vous savez redonner le sourire dans les moments pas très drôles d'une vie de thésard et ça c'est super. Mme Satour garde ta joie de vivre, ton âme de « djeuns »… et ton rire. Aurélia nos petites discussions en mode « ont refait le monde » vont me manquer. Je te promets d'aller chercher Clisson dans la Suisse normande !
Sandra : #Ceci est un script de remerciement à lancer dans R studio
>/Sandra/Merci_pour_toute_l'aide_bio-informatique
--<J'ai appris à $lancer$ des blasts [sous Ubuntu] tout seul 'comme un grand' en /ligne de commande. Si on m'avait dit ça un jour…moi qui détestais les ordinateurs quand j'étais plus jeune.
--<Éclate-toi bien avec tes scripts, tes macros et même ton Python maintenant !

Merci à l'équipe de choc du sous-sol ***Joseph***, ***Daniel*** et ***Benoit***, c'était un plaisir de s'occuper des plantes avec vous et de discuter de toute sorte de choses sauf de sciences. J'ai appris plein de choses sur les Hmongs ! Un grand merci à toi ***Joseph*** pour ton boulot avec les transfos d'Arabido quand je ne n'avais plus le temps de les refaire, toi aussi tu as grandement contribué au chapitre IV. Continue à venir en vélo, je te laisse ma place de parking ! Merci aussi à ***Abdel***, le sportif et le veilleur de l'ARES, vas-y doucement avec le basket dans la salle de sport d'Agrocampus. Une pensée pour ***Thierry***, mon dernier voisin de bureau tout aussi sympathique que les précédents. Je te pardonne tes quelques infidélités au GEVES et je te laisse en cadeau mon « petit » chlorophytum (tu vas être content).

Je n'oublie pas bien sûr les anciens : ***Adrien*** et ***Chloé*** avec qui j'ai passé de bons moments dans le bureau, notamment encore des délires musicaux, ***Aurélie*** le petit basmati, ***Johann*** notre aventurier australien un peu fêlé, ***Sophie*** et son rire mythique fusant à travers tout l'ARES et bien sûr ***Michaela.*** Tu as touché un peu à tout ici et j'ai été content de partager le bureau avec toi (même si tu te moquais de mes fonds d'écran fleuris…). Merci de m'avoir fait découvrir la protéomique, le carrot cake et le top du top… la polka finlandaise (Korpiklaani bien sûr). Une pensée aussi pour ***Imad***, mon premier stagiaire, on a quand même dégoussé ensemble plus de 8 550 graines de Medicago d'après mes calculs ! Bon vent à toi qui commence tout juste une thèse.

Un grand merci également aux Sud Africains pour leur accueil et leur travail sur Castano : ***Jill***, ***Keren***, ***Sershen***, ***Syd***, ***Udisha***, ***Dani*** et à ***Sylvain Gaillard*** pour les quelques conseils bio-informatiques… et oui soumettre des données de séquençage à NCBI ça ne s'improvise pas.

Bon courage à ceux qui restent : la « petite » ***Julia*** (d'ailleurs merci pour le bon coup de main avec les transfos de racines), ***Gael*** et ***Chvan*** qui prennent la relève des doctorants et tous les autres membres de l'ARES pour les petits moments de convivialité passés ensemble. J'espère que vous quitterez ce vieux bâtiment un peu lugubre un jour… en 2015 il paraît !

Enfin merci à mon rayon de soleil et ma famille pour votre soutien tout au long de ces trois années et surtout sur la fin ! Pas sûr que vous ayez bien compris tout ce que j'ai fait pendant ce temps là mais ce n'est pas grave vous avez maintenant un beau livre de chevet à votre disposition pour rattraper tout ça…

Sommaire

Liste des figures, tableaux et annexes 8

Liste des abréviations 11

Introduction : contexte et objectifs de l'étude 13

Chapitre I : Bibliographie 15

1-Développement des graines sensibles et tolérantes à la dessiccation 16

1.1-Les graines orthodoxes 16

1.1.1-Les phases du développement 16

1.1.2-Développement de la graine de *M. truncatula* 18

1.1.3- Les régulateurs majeurs de la maturation des graines 20

a)-La régulation hormonale : le rôle majeur de l'ABA 20

b)-Les facteurs de transcription majeurs 24

LEC1 24

ABI3 26

FUS3 et LEC2 27

ABI5 et ABI4 29

Autres protéines impliquées dans le développement des graines 30

Relations entre ces facteurs pour le contrôle de la maturation de la graine 32

Contrôle de la transition maturation/germination de la graine 32

1.2-Les graines récalcitrantes 35

1.2.1- Caractéristiques morphologiques et écologiques 35

1.2.2-Les phases du développement 37

1.2.3-Régulation du développement 39

2-Tolérance à la dessiccation 40

2.1-Définition et distribution dans le monde vivant 40

2.2-Dommages causés par la dessiccation 42

2.3-Les mécanismes impliqués dans la TD 43

2.3.1-Adaptations structurales 43

2.3.2-Ralentissement du métabolisme 44

2.3.3-LEA 45

a)-Définition et classification 45

b)-Caractéristiques biochimiques, structurales 47

Déhydrines 47

LEA_5 48

LEA_4 48
LEA_1 48
SMP 49
LEA_2 et LEA_3 49
Autres protéines associées aux LEA 49
Les protéines LEA identifiées chez *M. truncatula* 51
c)-Localisation cellulaire 51
d)- Fonctions proposées chez les plantes 51
e)-Mécanismes d'action 53
2.3.4-HSP 54
2.3.5-Sucres solubles 56
2.3.6-Etat vitreux 59
2.3.7-ROS et antioxydants 60
a) Définition des ROS et impact dans les graines 60
b) Antioxydants enzymatiques 62
c) Antioxydants non enzymatiques 63
2.3.8- Oleosines 64
2.4-Régulation de l'acquisition de la tolérance à la dessiccation 64
3-Modèle d'étude 65
3.1- Intérêt de la comparaison orthodoxe-récalcitrante 65
3.2- Intérêt des Fabacées et de *M. truncatula* 67
3.3- Choix de *C. australe* comme modèle de graine récalcitrante 67

Chapitre II : LEA polypeptide profiling of recalcitrant and orthodox legume seeds reveals ABI3-regulated LEA protein abundance linked to desiccation tolerance 70
Abstract 71
Introduction 72
Materials and methods 74
Results 79
Physiological description of *C. australe* seeds 79
Sequencing of the *C. australe* seed transcriptome and identification of LEA contigs 81
Identification of the heat-stable proteome of *C. australe* 83
Comparative analysis of the LEA proteome between *C. australe* and *M. truncatula* 83
Characterization of the LEA proteome of the desiccation sensitive *Mtabi3* mutant seeds and comparison with *C. australe* 85
ABI3 regulation of identified LEA proteins in relation to desiccation tolerance 87
Discussion 87
Supplementary data 92

Données supplémentaires à l'article : compléments sur l'étude du phénotype de *Mtabi3* 94
Etude de la dormance des graines *Mtabi3* 94

Chapitre III : Analyse transcriptomique des graines immatures et matures de *C. australe* et comparaison avec le développement des graines de *M. truncatula* 95
Approche et principaux résultats 96
1-Caractérisation du transcriptome de graines de *C. australe* 98
1.1-Séquençages et assemblage 98
1.2-Assemblage des lectures des deux séquençages 100
1.3-Annotation des contigs 100
1.3.1-Description par similarité : blast 100
1.3.2-Annotation fonctionnelle par Gene Ontology 102
1.4-Construction de la puce spécifique des graines de *C. australe* 104
2-Comparaison transcriptomique des graines de *C. australe* et de *M. truncatula*..... 106
2.1-Association des données transcriptomiques des deux espèces 106
2.2-Comparaison globale des différents stades de développement des graines de *C. australe* et *M. truncatula*
2.2.1-Les intensités d'expression des gènes sont-elles globalement comparables entre *C. australe* et *M. truncatula* ? 108
2.2.2-Analyse en composantes principales 110
2.3-Expression chez *C. australe* des régulateurs majeurs de la maturation des graines orthodoxes 112
2.4-ABI3 et ses cibles : un réseau sous-exprimé dans les graines matures de *C. australe* par rapport à *M. truncatula* 115
2.4.1-Les gènes codant pour les protéines LEA (Late Embryogenesis Abundant) 115
2.4.2-Les gènes cibles régulés par MtABI3 117
2.5-Expression des mécanismes liés à l'acquisition de la tolérance à la dessiccation : comparaison entre les graines de *C. australe* et de *M. truncatula* 121
2.5.1- Gènes liés à la réponse antioxydante 121
2.5.2- Métabolisme et expression des enzymes de voie de biosynthèse des RFO 123
2.6-La graine mature de *C. australe* exprime des gènes de réponse à l'acide abscissique et au stress mais également aux gibbérellines 126
2.6.1-Les gènes liés à la biosynthèse et à la signalisation de l'ABA 126
2.6.2-Les gènes liés à la réponse au stress 128
2.6.3-Les gènes liés aux gibbérellines 130
Discussion 131

Chapitre IV : Analyse fonctionnelle des homologues *CaABI3* et *CaABI3-like* : comparaison avec *MtABI3* et *AtABI3* 137
Approche et principaux résultats 138

1-Clonage des homologues du gène ABI3 chez *C. australe* 140
2-Expression ectopique de *CaABI3* et *CaABI3-like* 142
2.1-Principe de l'expression ectopique dans les racines de *M. truncatula* 142
2.2-Profil d'expression des cibles ABI3 142
3-Complémentation du mutant *abi3-5* d'*A. thaliana* 143
3.1- Dégradation de la chlorophylle 145
3.2- Dormance 145
3.3- Sensibilité à l'ABA 145
Discussion 146

Chapitre V : Discussion générale 150

Conclusion et perspectives 162

Chapitre VI : Matériel et méthodes 164
1-Matériel végétal 166
1.1-*Medicago truncatula* 166
1.2-*Castanospermum australe* 166
1.3-*Arabidopsis thaliana* 166
2-Traitements 167
2.1-Test de germination 167
2.2-Mesure de tolérance à la dessiccation 167
2.2.1-*M. truncatula* 167
2.2.2-*C. australe* 167
2.3-Mesure de sensibilité à l'ABA 168
3-Analyse bi-dimensionnelle du protéome stable à la chaleur 168
3.1-Extraction des protéines stables à la chaleur 168
3.2-Isoélectrofocalisation 168
3.3-Electrophorèse sur gel de polyacrylamide en condition dénaturante 170
3.4-Coloration des gels et quantification des spots 170
3.5-Identification des spots 170
4-Détermination de la teneur en sucres solubles des cotylédons de *C. australe* 171
5-Séquençage du transcriptome de *C. australe* 171
5.1-Extraction des ARN et séquençage des banques d'ADNc 171
5.2-Assemblage des lectures en contigs 173
5.3-Annotations des contigs 173
5.3.1-Description par blast 173
5.3.2-Annotation par Blast2GO 173

6-Détermination des gènes exprimés au cours du développement des graines de *M. truncatula* 175
7-Analyse transcriptomique du développement des graines de *C. australe* 175
7.1-Amplification des ARN 175
7.2-Marquage des ADNc et précipitation des cibles 177
7.3-Plan d'hybridation 177
7.4-Hybridation des puces et scan des lames 177
7.5-Traitement statistiques 178
8-Analyse fonctionnelle de *CaABI3* 180
8.1-Clonage des séquences complètes des gènes *MtABI3* et *CaABI3* 180
8.2-Construction des vecteurs d'expression pour la surexpression des gènes *ABI3* et la complémentation du mutant *abi3-5* d'*A. thaliana* 180
8.2.1-Les vecteurs binaires et le clonage Gateway 180
8.2.2-Construction des vecteurs d'expression 182
8.3-Sur-expression dans les racines de *M. truncatula* 184
8.3.1-Transformation des racines 184
8.3.2-Détermination des cibles de *MtABI3* et *CaABI3* par analyse transcriptomique 185
8.4-Complémentation du mutant *abi3-5* d'*A. thaliana* 185
8.4.1-Transformation des plantes 185
8.4.2-Sélection et caractérisation phénotypique des plantes transformées 185

Références bibliographiques 187

Annexes 215

Liste des figures, tableaux et annexes

Liste des figures

Figure 1.1. Evolution du poids frais, du poids sec, de la teneur en eau et acquisition des différents aspects de la qualité physiologique au cours des différentes phases du développement des graines orthodoxes. *17*
Figure 1.2. Evolution du poids sec, de la teneur en eau, de la composition des réserves, acquisition de la tolérance à la dessiccation et de la longévité au cours des différentes phases du développement des graines de *M. truncatula* génotype A17. ..... *19*
Figure 1.3. Voie de biosynthèse de l'ABA et modèle de voie de signalisation dépendante de l'ABA. *21*
Figure 1.4. Profil d'expression des gènes codant pour les régulateurs majeurs de la maturation des graines d'*A. thaliana*. *23*
Figure 1.5. Structure des facteurs de transcription majeurs de la maturation des graines et modèle présentant leur mode de régulation. *25*
Figure 1.6. Modèle représentatif des régulations au sein du réseau contrôlant la maturation des graines. *31*
Figure 1.7. Modèle représentatif de l'implication des hormones et des régulateurs majeurs des graines orthodoxes dans les différents processus du développement. *33*
Figure 1.8. Evolution du poids frais, du poids sec et de la teneur en eau au cours du développement des graines orthodoxes et récalcitrantes. *38*
Figure 1.9. Niveaux d'hydratation, processus de détérioration et de protection dans la graine au cours de la dessiccation. *41*
Figure 1.10. Structure des sucres non réducteurs impliqués dans la tolérance à la dessiccation et voie de biosynthèse du saccharose et des oligosaccharides de la famille du raffinose (RFO). *55*
Figure 1.11. Mécanismes de protection des protéines et membranes par les sucres selon le niveau d'hydratation. *57*
Figure 1.12. Différentes voies d'élimination des espèces réactives de l'oxygène (ROS) par action enzymatique. *61*
Figure 1.13. Classification phylogénétique de la famille des Fabacées. *66*
Figure 1.14. Morphologie des plantes (ou arbres), fleurs, gousses et graines de *M. truncatula* et *C. australe*. *68*

Figure 2.1. Determination of the critical water content of mature *Castanospermum australe* cotyledons. *73*
Figure 2.2. Reference map of the heat-stable proteome of mature *Castanospermum australe* cotyledons. *80*
Figure 2.3. Comparative analysis of LEA polypeptides in cotyledons of *Castanospermum australe* and *Medicago truncatula* seed heat stable proteome. *82*
Figure 2.4. Relative abundance of the different LEA polypeptides identified in cotyledons of *Castanospermum australe* and *Medicago truncatula* seeds. *84*
Figure 2.5. Characterization of *abscisic acid insensitive3* (*Mtabi3*) mutants of *Medicago truncatula*. *86*
Figure 2.6. The LEA protein profile from desiccation-sensitive cotyledons of *Castanospermum australe* and seeds of *Mtabi3* mutants compared to desiccation-tolerant *Medicago truncatula* wild type seeds. *88*
Figure S2.7. Caractérisation du niveau de dormance des graines de *M. truncatula* sauvages R108 et mutantes *Mtabi3-1*, *Mtabi3-2* en fonction du stade de développement. *93*

Figure 3.1. Nombre de contigs décrits par blast contre les différentes bases de données généralistes ou spécifiques de *M. truncatula*. *99*
Figure 3.2. Organisation en graphique acyclique dirigé (DAG) de la classification Gene Ontology (GO). *99*
Figure 3.3. Annotation Gene Ontology (GO) du transcriptome de graines de *C. australe*. *101*
Figure 3.4. Comparaison des processus biologiques associés aux gènes exprimés dans les graines de *C. australe* et *M. truncatula*. *103*
Figure 3.5. Association entre les sondes des puces Nimblegen spécifiques de *C. australe* et de *M. truncatula*. *105*

Figure 3.6. Evolution de la valeur de l'intensité médiane d'expression des gènes au cours du développement des graines de *C. australe*, *M. truncatula* A17, R108 et *Mtabi3-2*. *107*
Figure 3.7. Analyse en composantes principales (ACP) de l'expression des gènes préférentiellement exprimés dans les graines de *C. australe*, *M. truncatula* R108 et *Mtabi3-2*. *109*
Figure 3.8. Intensité d'expression des gènes codant pour les régulateurs majeurs de la maturation au cours du développement des graines de *C. australe*, *M. truncatula* A17, R108 et *Mtabi3-2*. *111*
Figure 3.9. Intensité et ratio d'expression des gènes codant pour les protéines LEA au cours du développement des graines de *C. australe*, *M. truncatula* A17, R108 et *Mtabi3-2.* *114*
Figure 3.10. Intensité et ratio d'expression des gènes cibles de MtABI3 identifié dans le transcriptome de graines de *C. australe* au cours du développement des graines de *C. australe* et *M. truncatula* A17. ..... *116*
Figure 3.11. Intensité et ratio d'expression des gènes cibles de MtABI3 non identifié dans le transcriptome de graines de *C. australe* au cours du développement des graines de *M. truncatula* A17. ..... *118*
Figure 3.12. Intensité d'expression des gènes codant pour des protéines impliquées dans la réponse au stress oxydatif au cours du développement des graines de *C. australe* et *M. truncatula* A17. ..... *120*
Figure 3.13. Teneurs en sucres solubles et intensité d'expression des gènes impliqués dans la synthèse des oligosaccharides de la famille du raffinose (RFO) au cours du développement des graines de *C. australe* et *M. truncatula* A17. ..... *122*
Figure 3.14. Intensité d'expression des gènes impliqués dans la biosynthèse et la signalisation de l'ABA au cours du développement des graines de *C. australe* et *M. truncatula* A17. ..... *125*
Figure 3.15. Intensité d'expression des gènes codant pour des protéines et des facteurs de transcription liés à la réponse au stress au cours du développement des graines de *C. australe* et *M. truncatula* A17. ..... *127*
Figure 3.16. Intensité et ratio d'expression des gènes impliqués dans le métabolisme et la signalisation des gibbérellines au cours du développement des graines de *C. australe* et *M. truncatula* A17. ..... *129*

Figure 4.1. Alignement des séquences protéiques de MtABI3 (génotype R108), CaABI3 et CaABI3-like. *139*
Figure 4.2. Profil d'expression des gènes cibles de MtABI3 en réponse à la sur-expression de *CaABI3g*, de *MtABI3g*, de $MtABI3_{CDS}$ dans les racines R108 et de *CaABI3-like* et $MtABI3_{CDS}$ dans les racines de M*tabi3-2*. *141*
Figure 4.3. Couleur, dormance et sensibilité à l'ABA de graines matures sauvages, mutantes *abi3-5* et mutantes complémentées par *AtABI3*, *CaABI3* et le vecteur vide. *144*

Figure 5.1. Modèle possible du développement des graines de *C. australe* déduit de l'analyse transcriptomique. *151*

Figure 6.1. Marquage des fleurs pour la récolte des graines au cours de la maturation. *165*
Figure 6.2. Différents stades de développement des gousses de *C. australe* utilisés dans cette étude. *165*
Figure 6.3. Procédure détaillée et paramètres appliqués pour les étapes de séquençage, assemblage et annotations du transcriptome de graines de *C. australe*. *172*
Figure 6.4. Détermination par RT-PCR du seuil d'intensité associé à une accumulation de transcrit sur la puce Affymetrix de *M. truncatula*. *174*
Figure 6.5. Plan d'hybridation des échantillons pour l'analyse du transcriptome de graines de *C. australe*. *176*
Figure 6.6. Principe de la PCR inverse utilisée pour le clonage des séquences complètes des gènes *MtABI3* et *CaABI3*. *179*
Figure 6.7. Caractéristiques des vecteurs de destination pK7WG2D, pKGW et principe du clonage Gateway monosite et multisite. *181*
Figure 6.8. Racines transformées sur-exprimant la GFP. *183*

Liste des tableaux

Tableau 1.1. Conditions optimales et délétères de stockage de graines récalcitrantes. *36*
Tableau 1.2. Différentes classifications des protéines LEA. *46*
Tableau 1.3. Gènes codant pour des protéines LEA identifiés chez *M. truncatula*. *50*

Table 2.1. Contigs features from 454 and Illumina RNAseq of *Castanospermum australe* seeds and annotation of the corresponding transcriptome. *76*
Table 2.2. LEA transcripts identified in the *Castanospermum australe* seed transcriptome. *78*
Table 2.3. Evidence for ABI3-dependent regulation of LEA homologues for which protein abundance is reduced or absent in desiccation-sensitive tissues (*C. australe* and *Mtabi3*) (group A) or unaffected or increased (group B). *90*

Tableau 3.1. Caractéristiques des contigs issus de l'assemblage des données de séquençage 454, Illumina et hybride 454/Illumina du transcriptome de graines de *C. australe*. *97*
Tableau 3.2. Comparaison des technologies de séquençage haut-débit 454 et Illumina. *132*

Tableau 4.1. Pourcentage de similarité entre les séquences protéiques de MtABI3, AtABI3, CaABI3 et CaABI3-like. *139*

Tableau 5.1. Protéines potentiellement impliquées dans la tolérance à la dessiccation ou la réponse au stress. *153*

Tableau 6.1. Composition des gels de polyacrylamide 12% 24 cm pour analyse protéomique bi-dimensionnelle. *169*
Tableau 6.2. Composition du milieu Fahräeus additionnée de kanamycine et fongicide pour la culture des racines transformées. *183*

Liste des annexes

Annexe 1. Séquences des amorces et températures d'hybridation utilisées pour le chapitre II. *216*
Annexe 2. Contigs de *C. australe* homologues des gènes codant pour les protéines LEA identifiées chez *M. truncatula*. *217*
Annexe 3. Description des spots identifiés dans le protéome stable à la chaleur de cotylédons de *C. australe*. *219*
Annexe 4. Intensité normalisée des spots associés aux LEA dans les protéomes stables à la chaleur de graines matures de *M. truncatula* et *C. australe* (moyenne de trois gels avec 150 µg de protéines de chaque espèce). *222*
Annexe 5. Intensité normalisée des spots associés aux protéines LEA dans les protéomes stables à la chaleur de graines matures de *M. truncatula* R108 (type sauvage), *Mtabi3-1* et *Mtabi3-2*. *223*
Annexe 6. Contigs Ca associés aux gènes Mt codant pour les régulateurs et répresseurs majeurs de la maturation des graines. *224*
Annexe 7. Alignement des séquences protéiques d'AtABI3, CaABI3 et CaABI3-like. *225*
Annexe 8. Séquences des amorces et températures d'hybridation utilisées pour le chapitre IV. *226*

Liste des abréviations

AAO3 : abscissique aldehyde oxydase-3

ABA : acide abscissique

ABF : ABRE binding factor

ABI3 : ABA insensitive-3

ABI4 : ABA Insensitive-4

ABI5 : ABA Insensitive-5

ABRE : ABA responsive element

AGL15 : agamous-like15

AIA : auxine

AIP2 : ABI3 interacting protein E3 ligase

AP2 : Apetala2

AREB : ABA Responsive Element Binding

ARF : auxine response factor

ASIL1 : Arabidopsis 6b-interacting protein-1 like-1

bHLH : basic Helix Loop helix

BOB1 : bobber1

bZIP : basic leucine zipper

CDS : séquence codante

CE : coupling element

Chip : immuprécipitation de chromatine

CK : cytokinine

DAF : jours après pollinisation

DREB : dehydration responsive element binding

EAR : ERF-associated amphiphilic repression

EM1 : early methionin labelled 1

EM6 : early methionin labelled 6

FUS3 : FUSCA3

GA : gibbérelline

GL2 : glabra2

HDA : histone deacetylase

HSI2 : high level expression of sugar inducible gene 2

HSL1 : HSI2 like 1

KEG : keep on going

L1L : LEC1-like

LEA : late embryogenesis abundant

LEC1 : leafy cotyledon-1

LEC2 : leafy cotyledon-2

LIMMA : Linear Models for Microarray data

MoCo : molybdène cofacteur

MS : matière sèche

NCED : 9-cis-epoxycarotenoïde dioxygenase

NF-YB : nuclear factor Y-B

PcG : polycomb group

PIL5 : phytochrome interacting factor-3 like-5

PKL : PICKLE

PP2C : 2C-type protein phosphatase

PRC2 : polycomb repressive complex 2

PYL1 : PYR1-like

PYR1 : pyrabactine resistance-1

RAV : related to ABI3/VP1

RBP : RNA-binding protein

RCAR : regulatory components of the ABA receptor

RFO : oligosaccharides de la famille du raffinose

SCR : seed coat ratio

SDR1 : déshydrogénase réductase des chaînes courtes

SID : seed information database

SIP : seed imbibition protein

SNF4b : sous-unité β du complexe 'sucrose non-fermenting-related kinase'

SNP : single nucleotide polymorphism

SnRK2 : SNF1-related protein Kinase 2

SSP : protéines de réserves

SUA : suppressor of *abi3-5*

TD : tolérance à la dessiccation

VAL : VP1/ABI3-like

VP1 : viviparous1

WRI1 : Wrinkled1

ZEP : zeaxanthine epoxidase

Comparaison moléculaire des graines orthodoxes de *Medicago truncatula* et récalcitrantes de *Castanospermum australe* : une nouvelle approche pour comprendre l'acquisition de la tolérance à la dessiccation

Introduction : contexte et objectifs de l'étude

D'un point de vue économique, les graines occupent une place majeure dans notre agriculture puisqu'elles représentent près de 70 % de l'alimentation mondiale (Bewley *et al.*, 2013). Pour assurer un rendement élevé, la qualité physiologique de ces graines, c'est-à-dire leur aptitude à la conservation et à une germination rapide et homogène, est primordiale. La tolérance à la dessiccation (TD) est une propriété remarquable de la majorité des graines (dites orthodoxes). Elle est la à la base de notre système agricole moderne (Alpert & Oliver, 2002) et de la conservation de la biodiversité dans les banques de semences (Walters *et al.*, 2005a) car la TD est la condition indispensable au stockage à long terme des graines. Cependant certaines graines, dites récalcitrantes, sont sensibles à la dessiccation et ne peuvent pas être stockés à l'état sec. Leur conservation à long terme est cruciale pour la sauvegarde de la biodiversité mais reste actuellement problématique et coûteuse (Li & Pritchard, 2009 ; Walters *et al.*, 2013). Une meilleure compréhension des mécanismes cellulaires impliqués dans la TD permettrait d'améliorer la conservation des graines récalcitrantes. De plus, d'autres applications sont possibles comme le développement de cultures et de plantes d'ornement plus tolérantes à la sécheresse (Bartels & Sunkar, 2005) et même la conservation de poches sanguines ou de cellules humaines (Crowe *et al.*, 2005).

L'acquisition de la TD des graines est difficile à étudier car elle se produit durant la phase de maturation, concomitamment à d'autres processus physiologiques. Pour « disséquer » moléculairement l'acquisition de la TD, il faut réussir à l'isoler des autres processus et cela ne peut être fait autrement que par une analyse comparative. C'est pourquoi l'approche développée dans ce mémoire consiste en une comparaison moléculaire entre une graine qui acquiert la TD (orthodoxe) et une graine qui reste sensible à la dessiccation tout au long de son développement (récalcitrante). L'objectif est d'observer leurs différences aux niveaux protéomique et transcriptomique car celles-ci révèleront des mécanismes spécifiques de l'acquisition de la TD. Cette étude constitue également la première caractérisation moléculaire d'une graine récalcitrante et ouvre des voies intéressantes pour comprendre l'origine de la sensibilité de ces graines à la dessiccation.

Le premier chapitre de cette thèse consiste en une introduction bibliographique qui aborde dans sa première partie le développement des graines orthodoxes (en particulier celui de *M. truncatula*), les régulateurs majeurs de leur maturation, ainsi que les connaissances sur les graines récalcitrantes et les caractéristiques de leur développement. Les conséquences de la dessiccation ainsi que les mécanismes liés à la TD sont décrits

dans la deuxième partie. La troisième partie apporte des précisions sur l'intérêt de l'approche comparative entre graines orthodoxes et récalcitrantes et le choix des deux espèces *Medicago truncatula* et *Castanospermum australe* comme modèles d'étude. Le deuxième chapitre se présente sous la forme d'un article publié dans Journal of Experimental Botany. Il porte sur l'analyse comparative de l'accumulation de l'ensemble des protéines LEA (Late Embryogenesis Abundant) entre les graines de *C. australe*, de *M. truncatula* et du mutant *Mtabi3*, dont les graines sons sensibles à la dessiccation. Le troisième chapitre décrit les résultats de l'analyse transcriptomique de graines matures et immatures de *C. australe* et la comparaison avec l'évolution du transcriptome au cours du développement de graines de *M. truncatula* et du mutant *Mtabi3*. Les analyses protéomique et transcriptomique ont fait ressortir l'importance du gène *ABI3* et posent la question du rôle de ce gène chez les plantes récalcitrantes. Une analyse fonctionnelle du gène *ABI3* de *C. australe* a donc été menée afin de vérifier si *CaABI3* est capable de contrôler les mêmes gènes cibles que *MtABI3.* Cette analyse constitue le quatrième chapitre de la thèse.

Chapitre I

Bibliographie

1-Développement des graines sensibles et tolérantes à la dessiccation

Les graines constituent l'étape clé du cycle de développement des Angiospermes. Elles forment un moyen efficace de dissémination, avantage sélectif qui a certainement grandement participé au succès de colonisation de ces dernières au cours de l'évolution. Les graines sont très diverses de par leur taille, leur forme, leur composition biochimique et présentent des caractéristiques (longévité, dormance, vigueur) adaptées à leur environnement. Trois catégories de graines ont été définies selon leur aptitude à la conservation en fonction de la température et de l'humidité relative : orthodoxes, intermédiaires, récalcitrantes (Roberts, 1973 ; Ellis *et al.*, 1990 ; Hong *et al.*, 1996). Dans un premier temps, Roberts (1973) a défini les graines orthodoxes comme celles pouvant être séchées sans dommage à 5 % de teneur en eau relative. De plus, la durée de conservation de ces graines augmente si l'humidité relative et la température du lieu de stockage sont abaissées. Les graines récalcitrantes forment une deuxième catégorie regroupant les graines ne pouvant être stockées à l'état sec car elles perdent leur viabilité à des teneurs en eau plus élevées (teneur en eau relative de 20 à 50 % selon les espèces). Plus récemment, une troisième catégorie a été proposée regroupant les graines dont la longévité augmente lorsque leur teneur en eau diminue jusqu'à être en équilibre avec l'air à 20°C et 40 à 50 % d'humidité relative mais diminue à des teneurs en eau plus basse (Ellis *et al.*, 1990 ; Hong *et al.*, 1996). Ces graines, dites intermédiaires, tolèrent moins bien la dessiccation que les graines orthodoxes mais mieux que les graines récalcitrantes. De plus, la longévité de ces graines est réduite si elles sont stockées à des températures froides (inférieures à 10°C) contrairement aux graines orthodoxes. Selon les estimations, de 75 à 95 % des 8 000 espèces spermaphytes étudiées auraient des graines orthodoxes, 1 à 15 % pourraient être classées comme intermédiaire et 4 à 10 % produiraient des graines récalcitrantes (Tweddle *et al.*, 2003 ; Walters *et al.*, 2013 ; Gaff & Oliver, 2013). La majorité des espèces cultivées ont des graines orthodoxes mais certaines espèces d'intérêt économique produisent des graines intermédiaires comme le café (Ellis *et al.*, 1990) ou la papaye (Ellis *et al.*, 1991) et même des graines récalcitrantes comme le cacaoyer, le cocotier ou l'hévéa (Chin & Roberts, 1980). Les graines orthodoxes et récalcitrantes diffèrent principalement par la capacité ou non à tolérer la dessiccation, nous allons par la suite détailler les caractéristiques du développement propres à ces deux groupes.

1.1-Les graines orthodoxes

1.1.1-Les phases du développement

Le développement des graines varie selon de nombreux paramètres environnementaux et génétiques mais pour chaque graine orthodoxe, trois phases peuvent être distinguées (Figure 1.1A) :

- La phase I correspond à l'embryogenèse et est caractérisée par les divisions et l'élongation cellulaires. Cette activité cellulaire intense s'accompagne d'une forte augmentation de la teneur en eau.

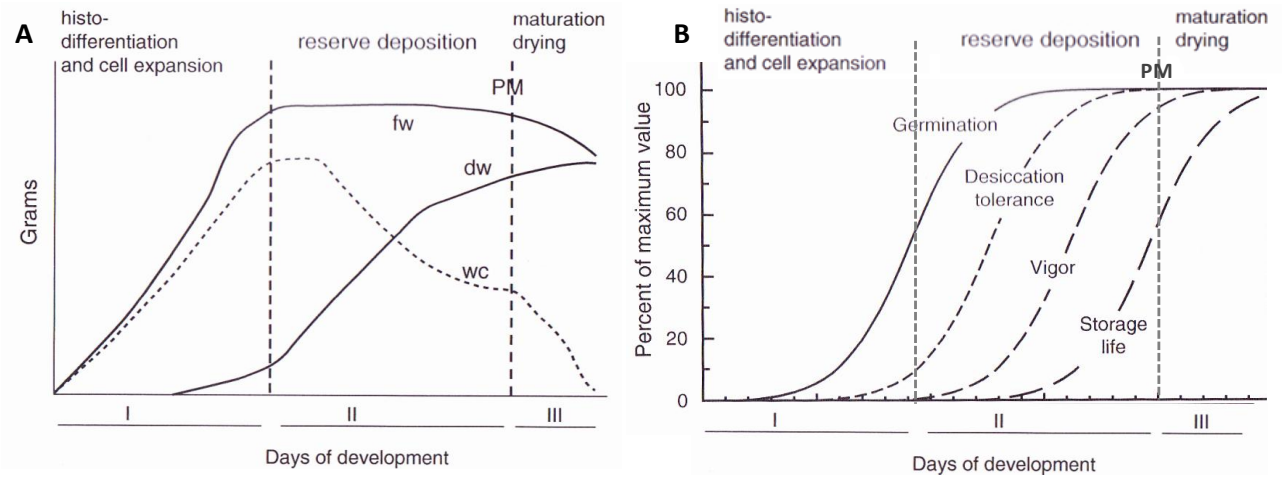


Figure 1.1. Evolution du poids frais, du poids sec, de la teneur en eau et acquisition des différents aspects de la qualité physiologique au cours des différentes phases du développement des graines orthodoxes (D'après Bewley *et al.*, 2013).

A) Trois phases peuvent être distinguées au cours du développement des graines orthodoxes. L'histo-différentiation (I) est caractérisée par une forte augmentation de la teneur en eau (WC exprimée en gH_2O/gDW) et du poids frais (FW). L'accumulation des réserves (II) se traduit par une forte augmentation du poids sec (DW) et une diminution de la teneur relative en eau par effet de concentration (pas de perte en eau). A la fin de ce stade, la maturité physiologique (PM) est atteinte. La maturation tardive (III) correspond à la dessiccation de la graine. **B)** Les différents aspects déterminant la qualité physiologique des graines sont acquis de façon séquentielle et chevauchante au cours de leur maturation. L'acquisition de la capacité germinative débute à la fin de l'embryogénèse et s'achève au début du remplissage. La tolérance à la dessiccation et la vigueur (mesurée par la vitesse de germination et la tolérance au stress) sont acquises durant le remplissage mais la longévité n'est acquise qu'après la maturité physiologique (PM), au cours de la dessiccation de la graine.

- La phase II correspond au remplissage, stade au cours duquel les cellules contenant les réserves entrent en expansion. L'accumulation de réserves se traduit par une augmentation importante du poids sec au cours de ce stade et une diminution de la concentration en eau, bien que la graine ne perde pas d'eau à ce stade.
- La phase III correspond à la maturation tardive. A ce stade, la graine n'accumule plus de réserves et perd de l'eau. Cette dernière étape est située entre la maturité physiologique (fin de remplissage) et la maturité de récolte (abscission du fruit). Elle peut même se poursuivre au-delà de l'abscission et est fondamentale pour assurer à la graine sa qualité physiologique et notamment sa longévité (Figure 1.1B).

La maturation de la graine débute après l'histodifférenciation et inclut les étapes de remplissage et de maturation tardive (phases II et III). C'est une étape complexe du développement au cours de laquelle plusieurs processus sont mis en place : capacité germinative, remplissage, acquisition de la TD, de la vigueur et de la longévité, mise en place de la dormance. Ces processus déterminent la qualité physiologique de la graine et sont acquis de façon séquentielle au cours de la maturation (Figure 1.1B). Selon les espèces, la proximité entre l'acquisition de ces propriétés peut être plus ou moins proche. Les graines acquièrent d'abord la capacité à germer, mais des régulateurs empêchent la viviparité (germination dans le fruit sur la plante mère) (cf 1.1.3). Puis la TD est acquise au cours du remplissage ; elle permet à la graine de faire face à la phase III et de survivre après perte totale de l'eau cellulaire. Ensuite vient l'acquisition de la vigueur germinative ; cette propriété définit la capacité des graines à germer de façon rapide et homogène et à tolérer des conditions environnementales peu favorables lors du développement de la plantule. En fin de développement, la dormance est mise en place parallèlement à l'acquisition de la longévité pour un certain nombre d'espèces. La longévité (ou aptitude à la conservation) correspond à la capacité de stockage de la graine, c'est-à-dire au temps pendant lequel la graine peut survivre à l'état sec sans perte de sa viabilité. La dormance est définie comme une inaptitude temporaire à germer dans des conditions favorables (Bewley, 1997). Elle permet de retarder la germination et de la répartir dans le temps, propriété avantageuse pour la pérennité de la descendance dans la nature. Cependant l'hétérogénéité de germination qui en résulte a été contre-sélectionnée lors de la domestication des espèces cultivées.

A l'abscission, la graine n'est pas encore à l'état sec. La dessiccation finale a lieu au cours des quelques jours qui suivent la séparation de la plante mère.

1.1.2-Développement de la graine de *M. truncatula*

M. truncatula est une plante autogame de la famille des Fabacées. Deux génotypes sont particulièrement étudiés dans notre laboratoire : A17 et R108. Chez le génotype A17 (Jemalong), dans des conditions d'éclairement et de température optimales (23°C 16h de jour et 21°C 8h de nuit), le développement des graines s'étale sur 44 jours de la floraison à l'abscission (Figure 1.2A). Il est légèrement plus rapide chez R108 puisque l'abscission survient 40 jours après floraison dans ces conditions. Durant les 12 premiers jours après floraison (DAF) se déroule la phase I d'embryogenèse (non représentée sur la figure 1.2A). La gousse en hélice est formée au bout de 3 jours et subit ensuite une série de divisions cellulaires. Le stade pro-embryon (précurseur de l'embryon) est visible dès 3 DAF et le tégument de la graine est déjà différencié. L'embryon

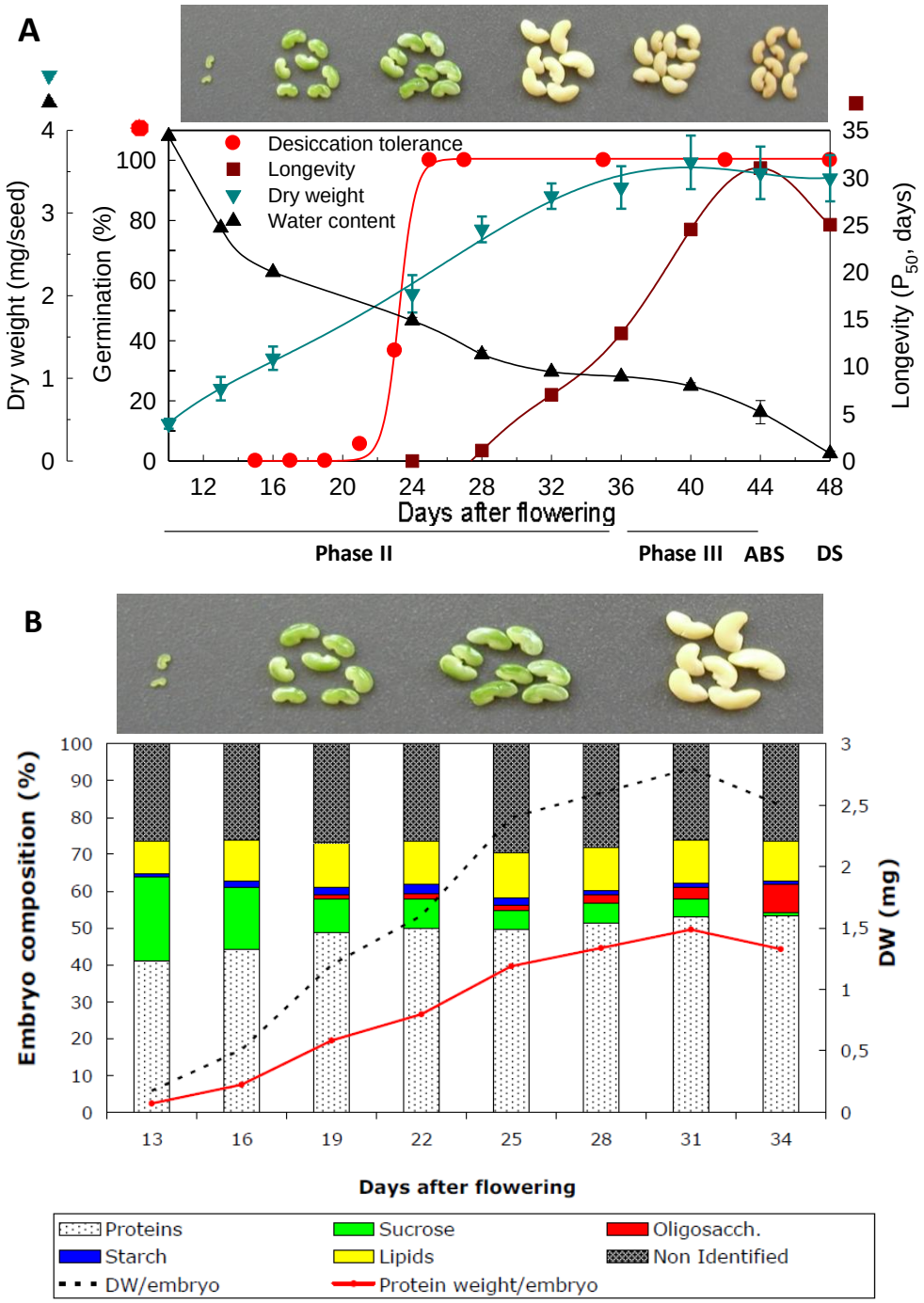


Figure 1.2. Evolution du poids sec, de la teneur en eau, de la composition des réserves, acquisition de la tolérance à la dessiccation et de la longévité au cours des différentes phases du développement des graines de *M. truncatula* génotype A17 (D'après Buitink *et al.*, 2006 ; Gallardo *et al.*, 2006).

A) L'évolution du poids sec (en bleu) et de la teneur en eau (en noir) est figurée au cours des phases II et III du développement des graines de *M. truncatula* A17. L'abscission (ABS) survient 44 jours après floraison (Days After Flowering ou DAF) et à 48 DAF la graine est sèche (DS). La tolérance à la dessiccation (en rouge) est acquise entre 20 et 24 DAF et la longévité (en marron) à partir de 28 DAF. L'évolution de la morphologie et de la couleur de la graine au cours du développement est figurée au dessus du graphique. **B)** L'évolution de la composition biochimique des réserves ainsi que l'augmentation du poids sec (courbe noire pointillée) et de la teneur en protéine (courbe rouge) de l'embryon au cours du remplissage sont figurées.

évolue rapidement au stade globulaire (6 DAF) puis au stade torpille (8-10 DAF) (Wang & Grusak, 2005 ; Abirached-Darmency *et al.*, 2005). Le rythme des divisions cellulaires diminue autour de 12-14 DAF. A ce stade, l'albumen occupe encore une place importante au sein de la graine, il se réduira progressivement au cours du remplissage au profit des cotylédons (Gallardo *et al.*, 2006). L'augmentation du poids sec montre que l'accumulation des réserves protéiques, lipidiques et glucidiques dans les cotylédons (phase II) a lieu entre 12 et 36 DAF (Figure 1.2A). Selon les conditions de cultures, Djemel *et al.* (2005) ont déterminé que la graine mature du génotype A17 contient entre 3 et 4 mg de matière sèche (MS) constituée de protéines (30 à 39 % MS), de lipides (7 à 9 % MS) et de sucres solubles (6 à 8 % MS). Parmi les grands types de protéines de réserve identifiées chez *M. truncatula*, les vicillines sont accumulées les premières (14 DAF), puis les légumines (16 DAF), suivi par les convicillines (18 DAF) (Gallardo *et al.*, 2003). Au cours du développement, la composition de ces réserves évolue même si les protéines restent le composé majoritaire (Figure 1.2B). L'amidon représente moins de 1 % de la MS dans la graine mature mais une accumulation transitoire est visible dans l'embryon et le tégument (mais pas dans l'albumen) entre 16 et 22 DAF (Djemel *et al.*, 2005). Sa dégradation pourrait fournir des squelettes carbonés nécessaires à d'autres molécules de réserve. Le profil des sucres solubles est totalement remanié entre le début et la fin de la maturation. Alors que le glucose, le fructose et surtout le saccharose sont les sucres solubles accumulés dans les graines au début de remplissage, leur teneur diminue progressivement au profit des oligosaccharides de la famille du raffinose (RFO) (Djemel *et al.*, 2005 ; Rosnoblet *et al.*, 2007). Ces RFO s'accumulent en fin de maturation principalement sous la forme de stachyose qui représente 79 % des sucres solubles contre 13 % pour le saccharose (Rosnoblet *et al.*, 2007).

L'acquisition de l'ensemble des paramètres liés à la qualité physiologique de la graine de *M. truncatula* a été étudiée au cours du développement. La capacité germinative est acquise à la fin de l'embryogenèse, autour de 14 DAF (Buitink *et al.*, 2006) et la vigueur germinative augmente fortement entre 14 et 22 DAF (Gallardo *et al.*, 2003). L'acquisition de la TD a lieu durant le remplissage, entre 20 et 28 DAF, et précède l'acquisition de la longévité qui débute à 28 DAF (Chatelain *et al.*, 2012). La longévité augmente jusqu'à l'abscission de la gousse, et diminue durant la dessiccation post-abscission. La chlorophylle est dégradée dans les graines à la fin du remplissage (autour de 36 DAF) (Chatelain *et al.*, 2012, Figure 1.2A). A l'abscission, la teneur en eau est d'environ 0,6 g H_20/g MS mais après 4 jours dans la gousse détachée de la plante mère, les graines sont desséchées et atteignent une teneur en eau à l'équilibre avec l'atmosphère de 0,1 g H_20/g MS (Chatelain *et al.*, 2012).

1.1.3- Les régulateurs majeurs de la maturation des graines

a)-La régulation hormonale : le rôle majeur de l'ABA

Parmi les phytohormones régulant le développement des graines se distinguent l'auxine (AIA), les cytokinines (CK), les gibbérellines (GA) et l'acide abscissique (ABA) (Bewley *et al.*, 2013). Les deux premières contrôlent respectivement l'élongation et les divisions cellulaires, elles sont donc principalement impliquées dans l'embryogenèse. L'étude de mutant déficient en synthèse de GA a montré l'importance de ce régulateur dans le développement, notamment la croissance, des graines de tomate (Groot *et al.*, 1987) et de pois (Swain *et al.*, 1997). De plus, les GA jouent un rôle majeur dans la germination en activant la levée de dormance et la mobilisation des réserves (Hilhorst & Karssen, 1992). Cette hormone joue un rôle antagoniste à l'ABA et le ratio

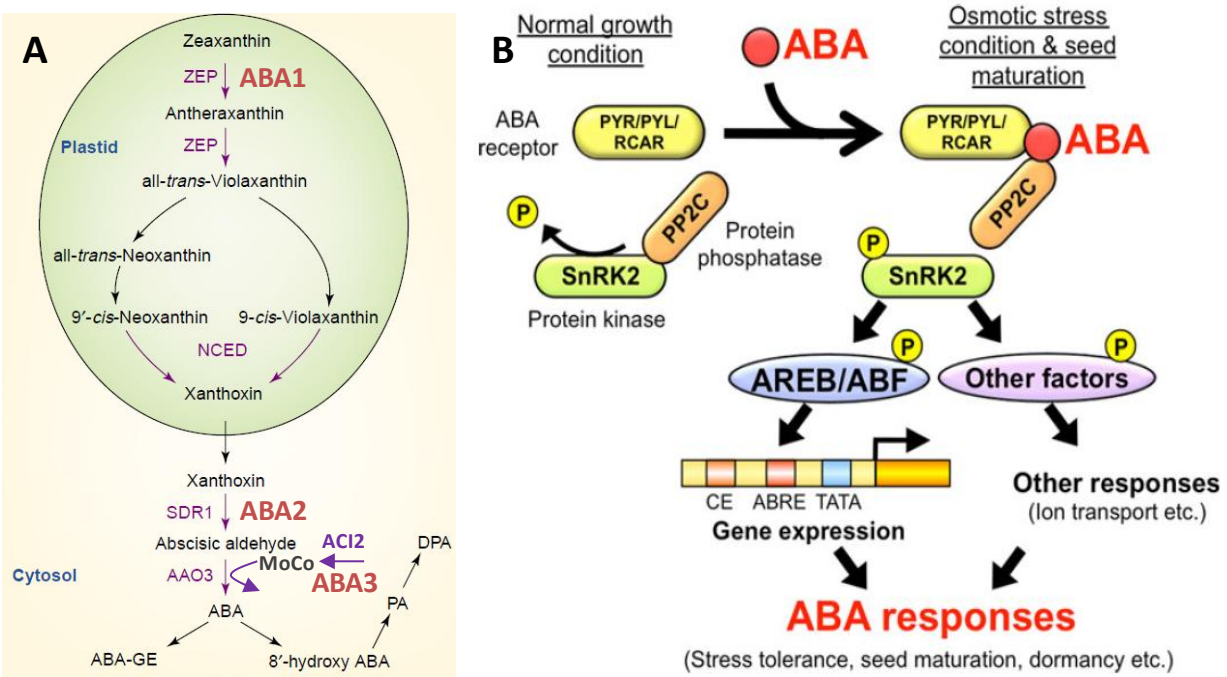


Figure 1.3. Voie de biosynthèse de l'ABA et modèle de voie de signalisation dépendante de l'ABA (D'après Nambara & Marion-Poll, 2003, 2005 ; Nakashima & Yamaguchi-Shinozaki, 2013).

A) Les caroténoïdes, dont la zeaxanthine fait partie, sont les précurseurs de l'ABA. Dans les plastes, ce précurseur est converti en trans-violaxanthine en deux temps par la zeaxanthine epoxidase (ZEP ou ABA1). La dernière étape de la conversion en xanthoxine (premier précurseur à 15 carbones comme l'ABA) est réalisée par la 9-cis-epoxycaroténoïde dioxygénase (NCED). La xanthoxine est exportée dans le cytosol et convertie en abscissique aldéhyde par une déshydrogénase réductase des chaînes courtes (SDR1 ou ABA2), lui-même oxydé en ABA par l'abscissique aldéhyde oxydase (AAO3). Cette dernière enzyme contient un cofacteur molybdène (MoCo) activé par une molybdène cofacteur sulfurase (ACI2 ou ABA3). L'ABA peut être hydroxylé par la CYP707A puis dégradé en acide phaséique (PA) et acide dihydro-phaséique (DPA) ou bien se conjuguer à d'autres molécules pour former par exemple l'ABA glycosyl ester (ABA-GE).
B) En condition normale sans ABA, la phosphatase PP2C (2C-type Protein Phosphatase) est associée à la kinase SnRK2 et l'inactive par déphosphorylation. L'ABA présent dans le milieu est fixé par le récepteur PYR/PYL/RCAR (PYrabactine Resistance-1 / PYR1-like / Regulatory Components of the ABA Receptor) qui interagit avec PP2C et inactive son activité phosphatase. La kinase SnRK2 est donc libérée sous sa forme active phosphorylée et active à son tour des facteurs de transcription de type ABF (ABRE Binding Factor), AREB (ABA Responsive Element Binding) ou d'autres protéines de réponse à l'ABA par phosphorylation. Les facteurs ABF et ABRE reconnaissent les motifs ABRE (ABA Responsive Element) des promoteurs de gène de réponse à l'ABA et activent leur transcription.

ABA/GA est le régulateur majeur du développement des graines : la maturation est favorisée par un ratio élevé alors que la germination est induite par un ratio faible (McCarty, 1995 ; White *et al.*, 2000). En effet, la teneur en ABA augmente de façon transitoire au cours de la maturation des graines. Selon les espèces, un (tomate) ou deux (*A. thaliana*, colza, coton) pics de synthèse d'ABA sont observés (Bewley *et al.*, 2013). Le premier est principalement d'origine maternel (Karssen *et al.*, 1983 ; Frey *et al.*, 2004) mais pas exclusivement (Kanno *et al.*, 2010 ; Boursiac *et al.*, 2013) et empêche la germination précoce de l'embryon. Des transporteurs d'ABA appartenant à la famille ABC (Kang *et al.*, 2010 ; Kuromori *et al.*, 2010) ou de type NRT-1 (Kanno *et al.*, 2012) ont été récemment identifiés et permettent la diffusion de l'ABA de la plante mère à la graine. Le deuxième pic est produit par l'embryon lui-même et induit notamment la synthèse des protéines et lipides de réserves, l'acquisition de la TD et la mise en place de la dormance (Finkelstein *et al.*, 2002 ; Lefevbre *et al.*, 2006). Les études ayant permis d'arriver à ces conclusions et les principaux gènes mis en jeu dans la signalisation de l'ABA seront détaillées au paragraphe 1.1.3.b.

L'ABA, qui appartient à la famille des terpénoïdes, est synthétisé chez les plantes à partir des caroténoïdes par une voie indirecte impliquant des enzymes plastidiques puis cytosoliques (Figure 1.3A). Cinq enzymes majeures ont été identifiées par l'analyse de mutants déficients en synthèse d'ABA (revu par Nambara & Marion-Poll, 2005). La zeaxanthine epoxidase (ZEP ; enzyme déficiente chez le mutant *aba1*) et la 9-cis-epoxycarotenoïde dioxygenase (NCED) participent à la conversion des caroténoïdes en xanthoxine dans les plastes. Puis dans le cytosol ce précurseur terpénoïdique est converti en ABA par l'intervention de trois enzymes : une déshydrogénase réductase des chaînes courtes (SDR1 ; enzyme déficiente chez le mutant *aba2*), une Molybdène Cofacteur (MoCo) sulfurase (ACI2 ; enzyme déficiente chez le mutant *aba3*) et l'abscissique aldehyde oxydase-3 (AAO3). Tous les gènes codant pour ces enzymes excepté pour SDR1 sont activés en réponse aux stress abiotiques (Qin & Zeevart, 1999 ; Thompson *et al.*, 2000 ; Iuchi *et al.*, 2001 ; Xiong *et al.*, 2001 ; Cheng *et al.*, 2002a ; Zhang *et al.*, 2008 ; Huo *et al.*, 2013). La NCED appartient à une famille multigénique mais seulement cinq des neuf gènes identifiés chez *A. thaliana* seraient réellement impliqués dans la synthèse d'ABA (Schwartz *et al.*, 2003). L'activité de cette enzyme semble être le régulateur majeur de la teneur en ABA dans les graines, plus que l'activité de la ZEP (Chono *et al.*, 2006 ; Frey *et al.*, 2006 ; Gutierrez *et al.*, 2007). L'ABA est inactivé par conjugaison à d'autres molécules ou par hydroxylation via une enzyme codée par le gène *CYP707A*. L'expression de ce gène clé de la dégradation de l'ABA est également finement régulé par d'autres hormones, notamment les GA (Saito *et al.*, 2004).

La majorité des gènes de réponse à l'ABA est caractérisée par la présence dans leur promoteur d'un élément cis conservé ACGTG(G/T)C appelé motif ABRE (ABA Responsive Element ou G-box) (Hattori *et al.*, 2002 ; Maruyama *et al.*, 2012). Une activation efficace de ses gènes nécessite plusieurs motifs ABRE ou une combinaison d'un motif ABRE avec un autre élément cis appelé CE (Coupling Element) (Hobo *et al.*, 1999). Pour activer ces gènes, l'ABA est d'abord perçu par des récepteurs : PYR1 (PYrabactine Resistance-1), PYL (PYR1-like) et RCAR (Regulatory Components of the ABA receptor) (Ma *et al.*, 2009 ; Park *et al.*, 2009). Ce complexe inhibe des phosphatases de type PP2C (2C-type Protein Phosphatase) ce qui permet la libération de kinases de type SnRK2 (SNF1-related protein Kinase 2) sous une forme active phosphorylée (Figure 1.3B). Ces kinases activent à leur tour des facteurs de transcription ABF (ABRE Binding Factor) ou AREB (ABA Responsive Element Binding) par phosphorylation. Ces facteurs se fixent sur les motifs ABRE des promoteurs pour réguler

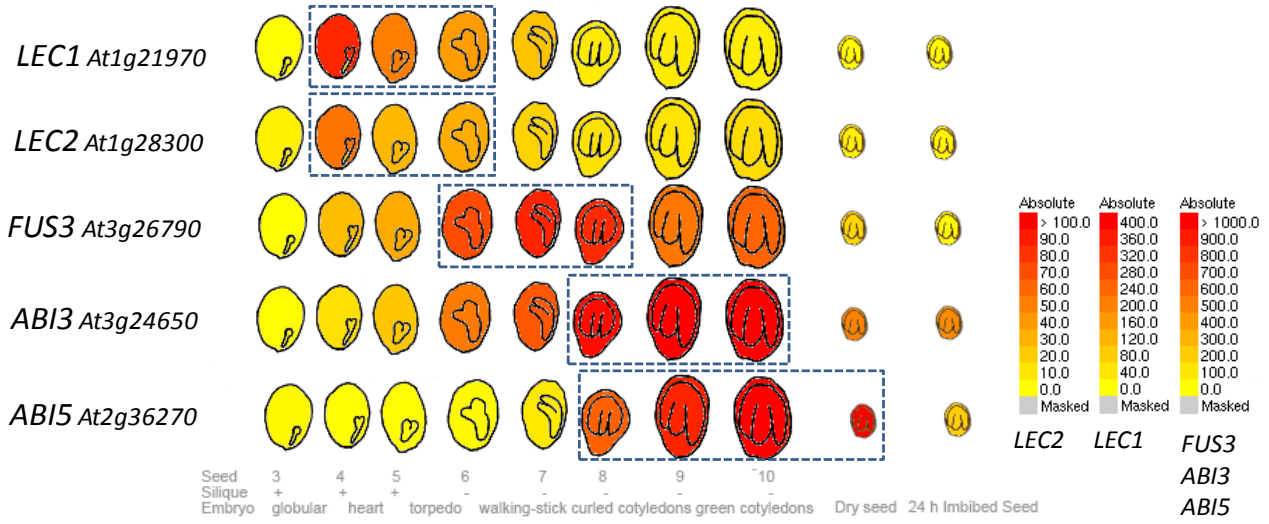


Figure 1.4. Profil d'expression des gènes codant pour les régulateurs majeurs de la maturation des graines d'*A. thaliana* (D'après eFP browser www.bar.utoronto.ca Winter *et al.*, 2007).

La valeur absolue de l'intensité d'expression des gènes *Leafy Cotyledon-1* (*LEC1*), *Leafy Cotyledon-2* (*LEC2*), *FUSCA3* (*FUS3*), *ABA Insensitive-3* (*ABI3*), et *ABA Insensitive-5* (*ABI5*) obtenue à partir de données transcriptomiques est représentée selon un gradient de couleur au cours de l'embryogenèse et de la maturation des graines d'*A. thaliana* ainsi que dans la graine sèche et 24h après imbibition. Trois gradients différents ont été utilisés selon l'intensité d'expression des gènes pour visualiser au mieux les différences d'expression au cours du développement : *LEC2* a une intensité d'expression plus faible que *LEC1*, lui-même exprimé plus faiblement que *FUS3*, *ABI3* et *ABI5*. La zone encadrée correspond au pic d'expression du gène au cours du développement d'*A. thaliana*.

positivement l'expression des gènes de réponses. L'importance des kinases SnRK2 dans le contrôle de la TD par l'ABA est montrée par la sensibilité à la dessiccation du triple mutant *snrk2* (Nakashima *et al.*, 2009). D'autres éléments cis que le motif ABRE sont impliqués dans la signalisation par l'ABA comme les éléments CE, RY/Sph, MYB ou MYC (Finkelstein *et al.*, 2002). Ils sont spécifiquement reconnus par des familles de facteurs de transcription (bZIP, B3, MYB, MYC). Les principaux facteurs en lien avec la maturation de la graine et notamment l'acquisition de la TD sont détaillés dans les paragraphes suivants.

b)-Les facteurs de transcription majeurs

LEC1

Le gène *Leafy cotyledon-1* (*LEC1*) code pour un facteur de transcription homologue à la sous-unité HAP3 des facteurs liant les CCAAT-box, sous-unité également appelé Nuclear Factor Y-B (NF-YB) (Lotan *et al.*, 1998 ; Laloum *et al.*, 2013). *LEC1* est exprimé spécifiquement dans l'embryon et l'albumen des graines lors de l'embryogenèse (Lotan *et al.*, 1998 ; Junker & Bäumlein, 2012) et éteint au cours de la maturation (Figure 1.4). Les graines du mutant *lec1* restent sensibles à la dessiccation à maturité et sont occasionnellement vivipares. Elles sont généralement insensibles à l'ABA, bien que la réponse à cette hormone soit très variable selon les graines, ce qui suggère un rôle mineur de LEC1 dans la perception de l'ABA. Les cotylédons accumulent des anthocyanines et présentent des caractéristiques de feuille (présence de trichomes et d'un système vasculaire plus développé que sur des cotylédons de type sauvage) (Meinke, 1992 ; Meinke *et al.*, 1994 ; Parcy *et al.*, 1997). Ces nombreux effets pléiotropiques sur le mutant *lec1* et l'induction de structures embryonnaires dans des feuilles sur-exprimant ce gène suggèrent un rôle majeur de LEC1 dans l'arrêt de la division cellulaire, la morphogenèse de l'embryon et la différenciation cellulaire du cotylédon (Lotan *et al.*, 1998 ; Raz *et al.*, 2001). L'auxine, qui joue un rôle majeur lors de l'embryogenèse somatique, stimule d'ailleurs l'expression de *LEC1* (Ledwon & Gaj, 2011). Malgré son expression précoce au cours de l'embryogenèse, LEC1 semble également contrôler le remplissage des graines tant au niveau des protéines de réserves que des lipides de réserve (Parcy *et al.*, 1997 ; Mu *et al.*, 2008 ; Shen *et al.*, 2010). Ce contrôle pourrait être indirect via la régulation des autres régulateurs majeurs ABI3, FUS3, LEC2 (Kagaya *et al.*, 2005 ; Mu *et al.*, 2008) ou également par le recrutement du facteur de transcription bZIP67 (Yamamoto *et al.*, 2009) (Figure 1.5B). Un domaine central conservé (domaine B sur la figure 1.5A) permettrait l'interaction avec l'ADN et d'autres protéines (Li *et al.*, 1992). Une liste de 356 gènes potentiellement cibles de LEC1 a été récemment proposée après induction de ce gène dans des plantules d'*A. thaliana* (Junker *et al.*, 2012). Un autre gène appartenant à la même famille de facteur de transcription et dont la séquence est très proche de *LEC1* a été caractérisé. Il est capable de complémenter le mutant *lec1* et a donc été nommé *LEC1-like* (*L1L*) (Kwong *et al.*, 2003). *L1L* est également exprimé au cours de l'embryogenèse mais le profil d'expression est légèrement plus précoce que celui de *LEC1*. De même, le développement de l'embryon est arrêté au stade globulaire chez le mutant *l1l* alors qu'il atteint un stade plus avancé chez le mutant *lec1* (Kwong *et al.*, 2003). Ce profil d'expression légèrement décalé ainsi que les phénotypes des mutants *lec1* et *l1l* confirme un rôle complémentaire de ces gènes au cours de l'embryogenèse.

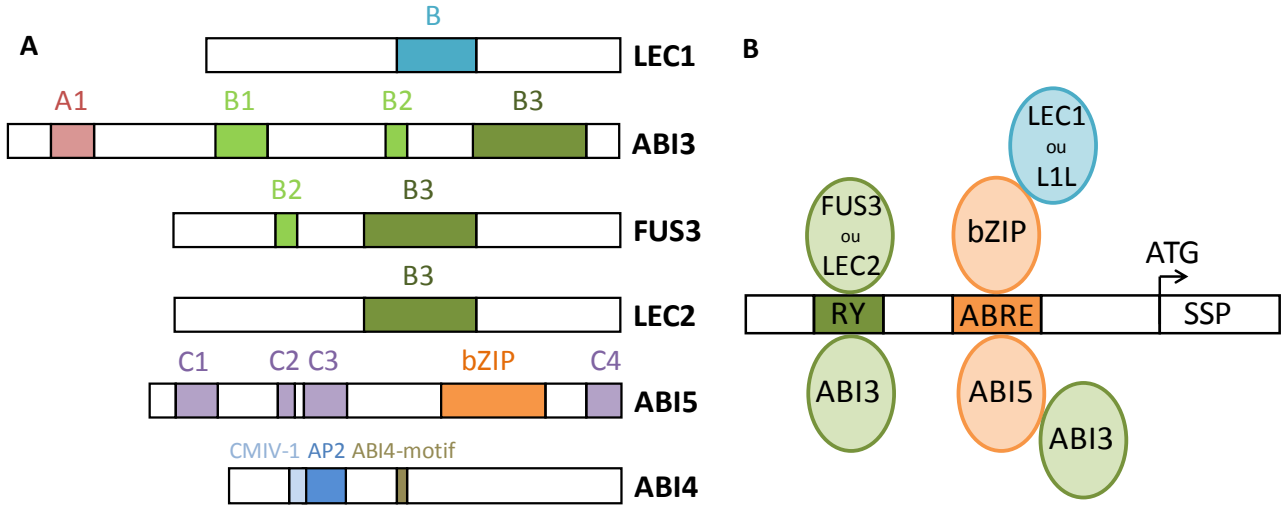


Figure 1.5. Structure des facteurs de transcription majeurs de la maturation des graines et modèle présentant leur mode de régulation.

A) Les domaines conservés des facteurs Leafy Cotyledon-1 (LEC1), Leafy Cotyledon-2 (LEC2), FUSCA3 (FUS3), ABA Insensitive-3 (ABI3) ABA Insensitive-5 (ABI5) et ABA Insensitive-4 (ABI4) sont figurés (D'après Li *et al.*, 1992 ; Li *et al.*, 2010 ; Wind *et al.*, 2011 ; Tezuka *et al.*, 2013). **B)** Les facteurs B3 (LEC2, FUS3 ou ABI3) sont capables de se fixer sur le motif RY des promoteurs tandis que les facteurs de la famille bZIP (Basic Leucine Zipper) comme ABI5 se fixent sur les éléments ABRE (ABA response element). ABI3 interagit avec ABI5 pour permettre la fixation de ce dernier sur le motif ABRE. D'autres facteurs bZIP peuvent être recrutés par LEC1 ou LEC1-like (L1L). Ce modèle peut représenter la régulation de gène codant pour des protéines de réserve (SSP) dont les promoteurs contiennent des motifs RY et ABRE.

ABI3

Le gène *ABA Insensitive-3* (*ABI3*) et son homologue *Viviparous1* (*VP1*) chez le maïs ont été respectivement identifiés à partir de mutants insensibles à l'ABA et mutants vivipares (Giraudat *et al.*, 1992 ; McCarty, 1995). Depuis, des homologues ont été retrouvés chez un grand nombre d'espèces dicotylédones, monocotylédones (Santos Mendoza *et al.*, 2008) et même un gymnosperme (Lazarova *et al.*, 2002) et une mousse (Marella *et al.*, 2006). Toutes ces séquences présentent des domaines très conservés malgré la distance génétique entre espèces au point que la surexpression de *VP1* du maïs (monocotylédone) est capable d'activer la majorité des cibles d'ABI3 chez le mutant *abi3-6* d'*A. thaliana* (dicotylédone) (Suzuki *et al.*, 2003). Quatre domaines caractérisent l'ensemble des homologues ABI3 identifiés à ce jour (Figure 1.5A). Le domaine acide A1, le moins conservé d'entre eux, est présent en position N-terminale et a une fonction d'activateur transcriptionnel (McCarty *et al.*, 1991). Trois domaines basiques très conservés assurent également les fonctions clés de ce facteur. B1 interagit spécifiquement avec le facteur de transcription ABI5 (Nakamura *et al.*, 2001), B2 serait essentiel à la localisation nucléaire et jouerait également un rôle de transactivateur (Marella & Quatrano, 2007) et B3 est capable de se fixer sur des séquences spécifiques de promoteurs appelées élément RY/Sph dont la séquence conservée est CATGCA (Suzuki *et al.*, 1997 ; Monke *et al.*, 2004) (Figure 1.5B). Ce domaine B3 d'environ 120 acides aminés est retrouvé chez d'autres facteurs de transcription tels que LEC2 (Leafy Cotyledon-2) FUS3 (FUSCA3), les ARF (Auxine Response Factor), les VAL (VP1/ABI3-like) et RAV (Related to ABI3/VP1) (revu dans Santos Mendoza *et al.*, 2008).

Plusieurs mutants alléliques *abi3* ont été identifiés chez *A. thaliana* avec des phénotypes plus ou moins forts selon la mutation (Bies-Ethève *et al.*, 1999). Dans le mutant *abi3-5*, utilisé pour notre étude, la délétion d'une base engendre un décalage du cadre de lecture et l'apparition d'un codon stop prématuré entre le domaine B1 et le domaine B2 (Bies-Ethève *et al.*, 1999). Le phénotype de ce mutant est très marqué : alors que la germination des graines de type sauvage est inhibée par 10 µM d'ABA, les graines mutantes sont majoritairement insensibles à 1 mM d'ABA. De plus elles ne sont pas dormantes et la chlorophylle n'est pas dégradée en fin de développement (Ooms *et al.*, 1993 ; Bies-Ethève *et al.*, 1999). Les graines matures sont sensibles à la dessiccation mais une acquisition transitoire de la TD à pourtant été mise en évidence en milieu de maturation (Ooms *et al.*, 1993). Les graines matures *abi3-5* n'accumulent pas de raffinose ni de stachyose (oligosaccharides de la famille des raffinose = RFO) contrairement aux graines de type sauvage mais contiennent plus de saccharose (Ooms *et al.*, 1993). De plus, les mutants sévères *abi3* accumulent moins de protéines de réserves et de protéines LEA (Late Embryogenesis Abundant) EM6 et EM1 que le type sauvage (Nambara *et al.*, 1992 ; Bies-Ethève *et al.*, 1999). Cette différence est également visible aux niveaux de l'expression des gènes codant pour ces protéines (Parcy *et al.*, 1994 ; Bies-Ethève *et al.*, 1999).

ABI3 est donc un régulateur majeur du remplissage et de la maturation tardive de la graine (acquisition de la TD, de la longévité et de la dormance) mais n'est pas impliqué dans l'embryogenèse comme LEC1 (Raz *et al.*, 2001). L'expression du gène *ABI3* est d'ailleurs plus tardive que celle de *LEC1* et se poursuit dans la graine mature (Parcy *et al.*, 1994) (Figure 1.4). Récemment 98 gènes cibles d'ABI3 ont été identifiés par des analyses Chip-chip, transcriptomiques et des activations transitoires de promoteur chez *A. thaliana* (Mönke *et al.*, 2012). La nature des cibles renforce le rôle majeur d'ABI3 dans les processus de maturation et surtout d'acquisition de TD dans les graines (cf 2.4 pour plus de détails). L'action d'ABI3 sur ces cibles peut être directe via l'élément

RY ou indirecte par association à d'autres facteurs de transcription. En effet, ABI3 et HSFA1 (Heat Shock Factor A1) ont une action synergique dans l'activation de l'expression du gène *HSP17.7 G4* (Rojas *et al.*, 1999). De même, la co-expression d'*ABI3* avec *bZIP10* et *bZIP25* est nécessaire à l'activation du promoteur du gène *At2S1* codant pour une protéine de réserve (Lara *et al.*, 2003). Plus récemment, Yotsui *et al.* (2013) ont montré une interaction physique entre les homologues chez *Physcomitrella patens* des facteurs ABI3 et NF-YC (facteur appartenant à la même famille que LEC1). Ce complexe active l'expression du gène *LEA1* par l'intermédiaire d'un nouvel élément cis ACTT-core.

Le gène *ABI3* est formé de 6 exons et 5 introns et un épissage alternatif a été mis en évidence chez le blé (McKibbin *et al.*, 2002), le pois (Gagete *et al.*, 2009), *A. thaliana* (Sugliani *et al.*, 2010), la tomate (Gao *et al.*, 2013) et *M. truncatula* (données non publiées). L'expression de l'isoforme *SlABI3-T* chez la tomate est stimulée par l'apport exogène d'ABA, de GA et d'auxine contrairement à l'isoforme SlABI3-F mais, d'après des expériences de double hybrides, cette dernière interagit plus fortement avec le facteur de transcription ABI5 (Gao *et al.*, 2013). L'épissage alternatif d'ABI3 est une modification post-transcriptionnelle régulée au cours du développement de la graine d'*A. thaliana* par le facteur d'épissage SUA (Suppressor of *abi3-5*) et conduit à l'accumulation d'un ARN codant une protéine tronquée non fonctionnelle dans la graine mature (Sugliani *et al.*, 2010). Ces résultats suggèrent des rôles spécifiques pour les différentes isoformes d'*ABI3*. Une régulation post-traductionnelle a également été mise en évidence par la protéine AIP2 (ABI3 Interacting Protein E3 ligase) qui se fixe sur ABI3 et l'adresse au 26S protéasome par ubiquitination (Zhang *et al.*, 2005). Ce mécanisme semble conservé puisqu'un homologue d'AIP2 interagissant avec ABI3 en double hybride a été récemment identifié chez le cyprès de Nootka (Gymnospermes) (Zeng *et al.*, 2013). Ces régulations permettent une inactivation rapide du programme de maturation dépendant d'ABI3 dans les graines matures pour la mise en place du programme de germination.

Bien qu'*ABI3* soit préférentiellement exprimé dans les graines, trois homologues ont été identifiés chez *Physcomitrella patens* (Marella *et al.*, 2006), espèce appartenant à la famille des Bryophytes et qui par conséquent ne produit pas de graines. Cette mousse est naturellement sensible à la dessiccation mais elle peut survivre à des teneurs en eau inférieures à 0,1 gH_2O/gMS suite à un prétraitement à 50 µM d'ABA (Koster *et al.*, 2010). Si l'apport exogène d'ABA est indispensable à l'acquisition de la TD chez *P. patens*, l'analyse de mutants pour lesquels l'expression d'*ABI3* a été supprimée montre que ce facteur est également essentiel à la survie à l'état sec de cette mousse (Khandelwal *et al.*, 2010). ABI3 pourrait y activer les mêmes cibles que dans les graines durant leur maturation, comme c'est le cas pour la protéine LEA1 (Sakata *et al.*, 2010 ; Yotsui *et al.*, 2013). De plus, ABI3 à d'autres fonctions dans le développement des tissus végétatifs : son expression est importante pour le développement des bourgeons de peuplier (Rohde *et al.*, 2002), la différentiation des plastes (Rohde *et al.*, 2000) et le développement des racines latérales (Brady *et al.*, 2003).

FUS3 et LEC2

Les gènes *FUS3* et *LEC2* codent également pour deux facteurs de transcription de la famille B3 (Luerssen *et al.*, 1998 ; Stone *et al.*, 2001), tout comme *ABI3*, et forment avec lui le groupe des facteurs AFL (ABI3-FUS3-LEC2) de la famille B3. Ils sont capables de fixer *in vitro* l'élément RY des promoteurs (Reidt *et al.*, 2000 ; Braybrook *et al.*, 2006) (Figure 1.5B). Cependant, alors que des homologues d'*ABI3* ont été identifiés

dans des espèces à travers tout le règne végétal, les homologues de *FUS3* semblent restreint aux Angiospermes et ceux de *LEC2* aux dicotylédones (Li *et al.*, 2010). Si d'un point de vue structural ces deux facteurs sont classés dans la même famille qu'ABI3, l'analyse des mutants correspondant chez *A. thaliana* a révélé des points communs avec *lec1*. En effet, la présence de trichomes sur les cotylédons des mutants *fus3* et *lec2* a conduit au regroupement de ces gènes avec *LEC1* pour former le groupe *LEC* impliqués dans la différenciation des cotylédons (Meinke *et al.*, 1994 ; Harada *et al.*, 2001). Les graines du mutant *fus3* présentent quelques similitudes supplémentaires avec celles de *lec1* et *abi3*: absence de dormance, diminution de l'accumulation des protéines de réserves et des RFO (Baumlein *et al.*, 1994 ; Keith *et al.*, 1994 ; Parcy *et al.*, 1997 ; Kroj *et al.*, 2003 ; Tiedemann *et al.*, 2008). Cependant, la sensibilité à la dessiccation des graines de *fus3* est variable selon le mutant considérée. Les graines matures des mutants obtenus par mutagenèse chimique ont été déterminées comme sensibles à la dessiccation sans pour autant préciser la méthode utilisée pour arriver à cette conclusion (Baumlein *et al.*, 1994 ; Meinke *et al.*, 1994). En revanche, les graines matures du mutant d'insertion de transposon *fus3-T* sont viables après dessiccation et sont seulement affectées dans leur longévité, comme le montre des expériences de vieillissement accéléré à 37°C (Tiedemann *et al.*, 2008). Cette différence phénotypique peut s'expliquer par la nature des mutations. Chez le mutant chimique la production d'un transcrit aberrant serait responsable de nombreux effets pléiotropiques alors qu'aucun transcrit *FUS3* n'est détectable après l'insertion du transposon chez le mutant *fus3-T*. Ce mutant perte de fonction étant plus spécifique, FUS3 n'est vraisemblablement pas impliqué dans l'acquisition de la TD.

Les graines du mutants *lec2* présentent un phénotype nettement moins altéré que celles de *lec1 abi3* et *fus3*. Les graines matures accumulent moins d'anthocyanines, sont tolérantes à la dessiccation et leur viabilité est seulement légèrement altérée après quelques semaines de stockage, suggérant un rôle mineur de LEC2 dans la longévité (Meinke *et al.*, 1994 ; Stone *et al.*, 2001). *LEC2* est exprimé principalement au cours de l'embryogenèse de la graine alors que l'expression de *FUS3* se poursuit durant le développement et diminue seulement en toute fin de maturation (Stone *et al.*, 2001 ; Kroj *et al.*, 2003) (Figure 1.4). Pourtant ces gènes sont tous les deux impliqués dans l'accumulation de protéines de réserves (SSP) où ils contrôlent des voies partiellement redondantes mais différentes de celles d'ABI3 (Kroj *et al.*, 2003). Tout comme LEC1, LEC2 contrôle la biosynthèse des lipides via le facteur de transcription Wrinkled1 (WRI1) (Baud *et al.*, 2007) et régule des gènes impliqués dans l'embryogenèse somatique (Braybrook *et al.*, 2006 ; Braybook & Harada, 2008 ; Stone *et al.*, 2008).

FUS3 et LEC2 ne sont pas impliqués dans la perception de l'ABA, comme l'atteste la sensibilité des mutants à cette hormone (Meinke *et al.*, 1994 ; Parcy *et al.*, 1997) mais occupent une place centrale dans le contrôle hormonal. En effet l'auxine influence l'expression de *FUS3* qui à son tour stimule la biosynthèse d'ABA (Gazzarrini *et al.*, 2004) alors que l'expression d'*AtGA3ox2*, codant pour une enzyme clé de l'activation des GA, est réprimée par LEC2 et FUS3 (Curaba *et al.*, 2004). Les GA exercent également un contrôle sur FUS3 via un mécanisme post-traductionnel impliquant une dégradation protéique par le protéasome, alors même que le gène *FUS3* reste fortement exprimé dans la graine (Lu *et al.*, 2010).

Les cibles de FUS3 ont été récemment identifiées par immunoprécipitation de chromatine. Après recoupement des résultats avec des données transcriptomiques dans lesquelles *FUS3* est sur- ou sous-exprimé, une liste de 366 cibles directes potentielles a été publiée (Wang & Perry, 2013). Concernant les cibles de LEC2,

une liste de 718 gènes sur-exprimés suite à l'induction constitutive de *LEC2* chez *A. thaliana* a été publiée, parmi lesquels 420 sont exprimés au cours du développement de la graine (Braybrook *et al.*, 2006).

ABI5 et ABI4

Les gènes *ABA Insensitive-4* (*ABI4*) et *ABA Insensitive-5* (*ABI5*) codent pour deux facteurs de transcription appartenant à deux familles différentes. ABI4 contient un domaine Apetala2 (AP2) (Finkelstein *et al.*, 1998) tandis qu'ABI5 est un membre de la famille des Basic Leucine Zippers (bZIP) (Finkelstein & Lynch, 2000). Ces gènes ont été identifiés à partir de mutants insensibles à l'ABA, *abi4* et *abi5*, et présentent quelques similitudes avec le mutant *abi3* comme la faible expression du gène *AtEM6* codant pour une LEA (surtout pour *abi5*) (Finkelstein, 1994). Cependant, leur phénotype est nettement moins marqué que les mutants sévères *abi3* : l'insensibilité à l'ABA est moins forte et la dormance comme la tolérance à la dessiccation ne sont pas affectées par ces mutations (Finkelstein *et al.*, 1998 ; Finkelstein & Lynch, 2000). L'expression d'*ABI4* et d'*ABI5* n'est pas spécifique des graines même si le niveau d'expression dans les tissus végétatifs est plus faible (Finkelstein *et al.*, 1998 ; Finkelstein & Lynch, 2000). Cependant, l'expression d'*ABI5* est induite par l'ABA et par la sur-expression d'*ABI3* (Finkelstein & Lynch, 2000) contrairement à *ABI4* (Söderman *et al.*, 2000).

ABI5 est caractérisé par un domaine bZIP et 4 domaines conservées (C1 à C4) (Tezuka *et al.*, 2013) (Figure 1.5A). Il est capable de se fixer sur les motifs ABRE des promoteurs de ses cibles via le motif bZIP et d'activer leur expression, comme par exemple les gènes codant pour les LEA *AtEM1* et *AtEM6* (Carles *et al.*, 2002). D'autres facteurs de transcription reconnaissent le domaine ABRE et participent à la réponse au stress dépendante de l'ABA (comme les facteurs ABF ou AREB). Mais le niveau d'expression d'*ABI5* dans les graines suggèrent un rôle spécifique dans la maturation de la graine (Finkelstein *et al.*, 2005) principalement par l'interaction physique avec ABI3 au niveau des domaines conservées C2 et C3 (Nakamura *et al.*, 2001). Cette interaction est très importante pour l'action d'ABI3 dans l'activation des gènes codant pour des protéines de réserves, dont le promoteur contient à la fois des éléments RY et un motif ABRE (Kroj *et al.*, 2003 ; Lara *et al.*, 2003) (Figure 1.5B). Par analyse double hybride, Marella *et al.* (2006) ont montré une faible interaction entre PpABI3A (codée par un des trois gènes identifiés chez *P. patens*) et ABI5. Cette faible interaction pourrait expliquer la mauvaise complémentation du mutant *abi3-6* concernant les gènes *AtEM6* et *AtEM1*, dont l'expression est régulée par ABI3 et ABI5 (Marella *et al.*, 2006). Cette complémentarité entre ABI3 et ABI5 est renforcée par leur co-expression (figure 1.4) et par l'importance d'ABI5 en aval d'ABI3 dans l'arrêt de croissance des radicules au cours de la germination en réponse à l'ABA (Lopez-Molina *et al.*, 2001 ; 2002).

Trois domaines conservés sont mis en évidence par alignement des homologues d'ABI4 clonés à ce jour (Wind *et al.*, 2013) (Figure 1.5A) : un domaine AP2 flanqué d'un motif CMIV-1 présent chez les facteurs DREB group IV (Dehydration Responsive Element Binding) et la séquence de 8 acides aminés LRPLLPRP, spécifique des facteurs ABI4 des Angiospermes. En plus d'être impliqué dans la voie de signalisation de l'ABA, ABI4 joue également un rôle majeur dans la signalisation des sucres (Huijser *et al.*, 2000). En effet, ce facteur se fixe au motif CE1 (Coupling Element 1) présent dans les promoteurs de plusieurs gènes de réponse à l'ABA (*Rab17*) et aux sucres (*ADH1*) (Niu *et al.*, 2002). Si aucune interaction physique n'a été mise en évidence entre ABI4 et ABI3 ou ABI4 et ABI5 (Nakamura *et al.*, 2001), ces facteurs ont des cibles communes et leur action serait synergique (Reeves *et al.*, 2011 ; Kong *et al.*, 2013). De même, l'étude de double mutants a montré

qu'ABI4 et ABI5 interagissent avec FUS3 et LEC1 même si aucune interaction physique n'a été mise en évidence par double hybride (Brocard-Gifford *et al.*, 2003).

Les régulations post-transcriptionnelles et post-traductionnelles régulent l'activité de ces deux facteurs. En effet, ABI5 est activée par phosphorylation via la kinase SnRK2 (Fujii *et al.*, 2007) (cf 1.1.3.a). Comme pour *ABI3*, un épissage alternatif a été mis en évidence pour *ABI5* chez le riz aboutissant à une fixation plus ou moins forte avec VP1 (homologue d'ABI3 chez le riz) selon les isoformes (Zou et al., 2007). De plus, la protéine ABI5 est rapidement dégradée par le protéasome après ubiquitylation par la protéine KEG (Keep on Going) (Stone *et al.*, 2006) ou sumoylation (Miura *et al.*, 2009). En fusionnant *ABI4* à la GFP ou au GUS dans des lignées transgéniques, Finkelstein *et al.* (2011) ont montré que la protéine ABI4 est rapidement dégradée, au moins en partie par le protéasome. Ces résultats suggèrent une régulation au niveau post traductionnel par un mécanisme non identifié à ce jour qui pourrait impliquer le domaine AP2 (Finkelstein *et al.*, 2011).

Autres protéines impliquées dans le développement des graines

Tous ces facteurs majeurs régulent l'expression de facteurs secondaires comme par exemple AGL15 (Agamous-like15) lié au métabolisme des GA (Wang *et al.*, 2004), GL2 (Glabra2) ou WRI1 impliqués dans l'accumulation des lipides de réserve (Shen *et al.*, 2006 ; Baud *et al.*, 2007). Ces facteurs vont eux-mêmes activer leurs propres cibles, ce qui complexifie encore le réseau de régulation de la maturation (Gutierrez *et al.*, 2007 ; Santos-Mendoza *et al.*, 2008). D'autres protéines semblent interagir avec les régulateurs majeurs LEC. C'est le cas de la protéine TANMEI/EMB2757 qui joue un rôle majeur dans l'embryogenèse et la maturation des graines d'*A. thaliana*, comme le montre l'analyse du mutant *tan* déficient pour cette protéine (Yamagishi *et al.*, 2005). Les graines mutantes présentent des similitudes avec celles des mutants *lec1*, *lec2* et *fus3* : malformation du méristème apical, trichomes sur les cotylédons, sensibilité à la dessiccation, moindre accumulation de protéines et de lipides de réserve. D'ailleurs, l'étude de double mutants *tan/lec1* et *tan/lec2* confirment une interaction génétique entre ces protéines (Yamagishi *et al.*, 2005). Bien que TANMEI/EMB2757 possède sept domaines répétés WD, domaines connus pour être impliqués dans les interactions protéines-protéines (Nocker & Ludwig, 2003), à ce jour aucune interaction directe avec LEC1 ou LEC2 n'a été mise en évidence.

Hormis ABI5, d'autres facteurs bZIP sont exprimés au cours de la maturation de la graine et pourraient jouer un rôle complémentaire à ABI5. AtbZIP10 et AtbZIP25 ont été précédemment cités comme nécessaires à l'activation du promoteur du gène *At2S1* codant pour une protéine de réserve (Lara *et al.*, 2003). Plus récemment il a été montré que AtbZIP53 pouvait former des hétérodimères avec ces deux facteurs et activer l'expression du gène *At2S2* (Alonso *et al.*, 2009). De plus, l'expression de gènes codant pour des protéines de réserve est activée par AtbZIP67 en interaction avec LEC1 (Yamamoto *et al.*, 2009). Le facteur AtbZIP12 (appelé également EEL) a été identifié comme antagoniste d'ABI5. Ce répresseur entre en compétition avec ABI5 pour fixer le domaine ABRE du promoteur de *AtEM1* et moduler ainsi le niveau d'expression de ce gène (Bensmihen *et al.*, 2002 ; 2005).

D'autres gènes impliqués dans le contrôle de la balance ABA/GA joue un rôle majeur dans la maturation de la graine comme PIL5 (Phytochrome Interacting factor-3 like-5) et SOMNUS. PIL5 est un facteur de transcription de la famille bHLH (basic Helix Loop helix) qui active la biosynthèse d'ABA et la dégradation des

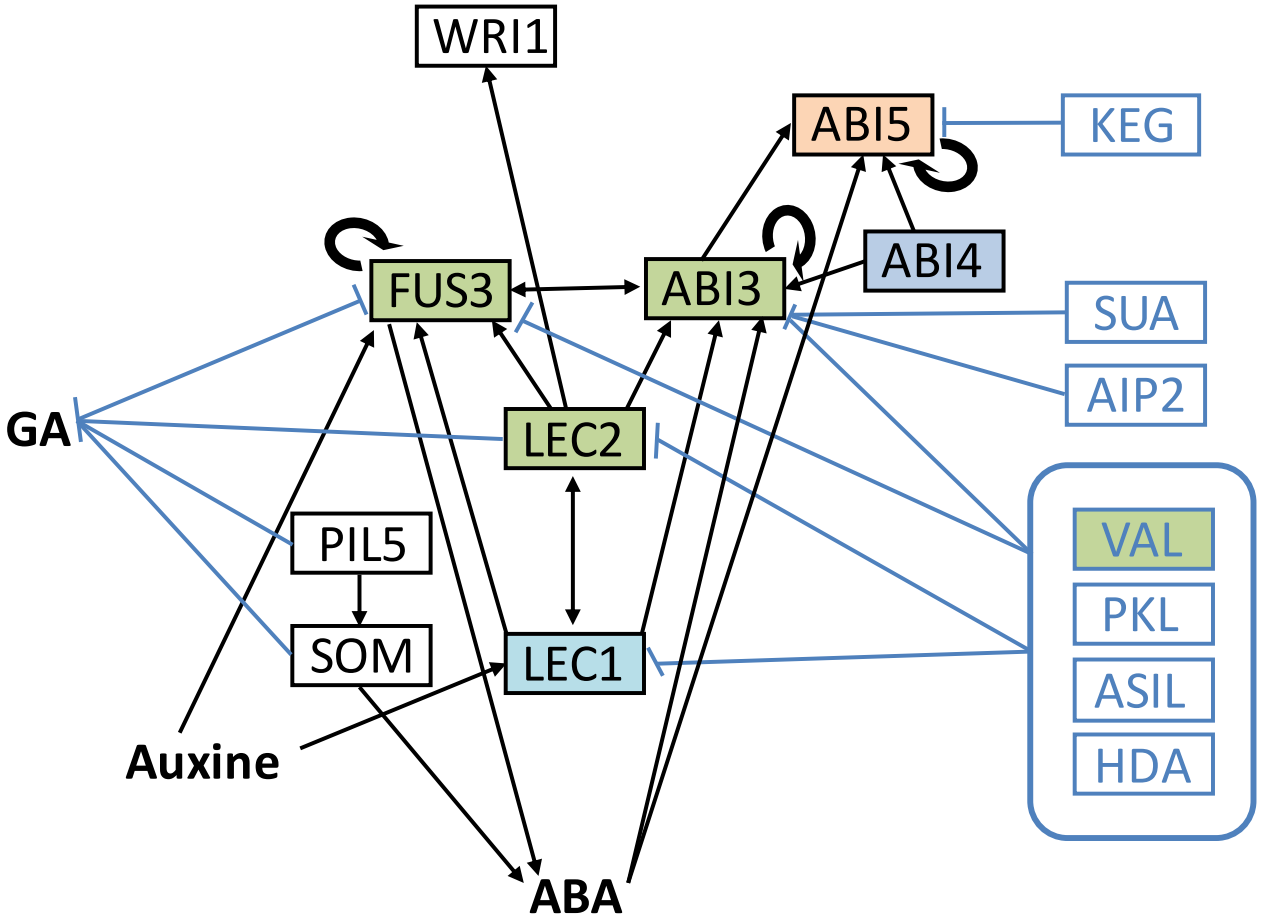


Figure 1.6. Modèle représentatif des régulations au sein du réseau contrôlant la maturation des graines.
Les trois hormones les plus importantes sont représentées en gras : gibbérellines (GA) ; acide abscissique (ABA) ; auxine. Leafy Cotyledon-1 (LEC1) est en amont du réseau et contrôle l'expression des gènes B3 du réseau AFL : *ABA-insensitive-3* (*ABI3*), *Leafy Cotyledon-2* (*LEC2*), *FUSCA3* (*FUS3*). LEC2 contrôle l'expression de *Wrinkle-1* (*WRI1*). PIL5 (Phytochrome interacting factor-3 like-5) et SOMNUS (SOM) régulent la balance hormonale ABA/GA au profit de l'ABA. Plusieurs inter-régulations et autorégulations ont été mises en évidence entre ABI3, ABI4 et ABI5. Les couleurs des facteurs de transcription font référence aux domaines conservés présentés sur la figure 1.5, Viviparous1-ABI3-Like (VAL) qui possède un domaine B3 est donc également représenté en vert. Les activations sont représentées par des flèches noires et les répressions en bleu. VAL, Pickle (PKL), ASIL (Arabidopsis 6b-interacting protein-1 Like-1) et les histones deacétylases (HDA) agissent ensemble pour réprimer l'ensemble des gènes *LEC1*, *LEC2*, *FUS3* et *ABI3*. Le facteur ABI3 est spécifiquement dégradé par l'intervention de SUA (suppressor of *abi3-5*) et AIP2 (ABI3-interacting protein E3 ligase) alors que la protéine ABI5 est dégradée par l'intermédiaire de KEG (Keep on Going).

GA (via le contrôle de *NCED6*, *NCED9*, *GA2ox2*) et réprime la dégradation d'ABA et la biosynthèse des GA (via le contrôle de *CYP707A2*, *GA3ox1* et *GA3ox2*) (Oh *et al.*, 2006 ; 2007). Son action est indirecte et passe en partie par l'activation du gène *SOMNUS* codant pour une protéine nucléaire de type CCCH zinc finger (Kim *et al.*, 2008). Les phytochromes participent à la dégradation de PIL5 par le 26S proteasome, ce qui permet le déclenchement de la germination (Oh *et al.*, 2006).

Relations entre ces facteurs pour le contrôle de la maturation de la graine

La régulation des gènes codant pour les facteurs LEC1, FUS3, ABI3 et LEC2 est complexe et de nombreuses inter-régulations ou autorégulations ont été mises en évidence entre eux (Meinke *et al.*, 1994 ; Parcy *et al.*, 1997 ; Kagaya *et al.*, 2005 ; Santos-Mendoza *et al.*, 2005 ; To *et al.*, 2006). Le nombre très important d'effets pléiotropiques sur le mutant *lec1* comparé aux autres mutants suggèrent que ce facteur est en amont des facteurs B3 (Meinke *et al.*, 1994). De même l'expression de *FUS3* et *ABI3* permet de restaurer une grande partie du phénotype de *lec2*, ce qui suggère que ces deux facteurs agissent en aval de *LEC2* (To *et al.*, 2006) (Figure 1.6). L'expression d'*ABI3* est régulée par LEC1, LEC2 et FUS3 au cours du développement des cotylédons d'*A. thaliana* mais ces inter-régulations sont tissus spécifiques et FUS3 ou LEC2 ne semblent pas impliqués dans la régulation d'ABI3 dans l'axe embryonnaire (To *et al.*, 2006). Plusieurs études ont également montré un réseau de régulation entre ABI3, ABI4 et ABI5 (Söderman *et al.*, 2000 ; Bossi *et al.*, 2009). Ces trois facteurs agissent dans la même voie de signalisation et l'expression d'*ABI5* est régulée par ABI4 et ABI3 (Lopez-Molina *et al.*, 2002 ; Bossi *et al.*, 2009) (Figure 1.6).

Le déclenchement et l'arrêt du programme de maturation des graines orthodoxes sont donc finement régulés. De plus, l'interaction entre plusieurs de ces facteurs est nécessaire à l'expression de chaque processus du développement de la graine orthodoxe (Figure 1.7). Les gènes *LEC* (*LEC1*, *LEC2* et *FUS3*) régulent tous les trois l'embryogenèse (par l'intermédiaire, au moins en partie, d'autres facteurs secondaires non figurés ici). *FUS3* et *LEC2* sont également impliqués dans l'accumulation de lipides (via *WRI1*) et protéines de réserve (SSP). Concernant les SSP, les facteurs ABI3 et ABI5 jouent également un rôle majeur ainsi que d'autres facteurs bZIP non figurés ici (bZIP10 et bZIP25). L'accumulation des protéines LEA est principalement régulée par le réseau ABI3, ABI4, ABI5. ABI3 est le régulateur majeur de l'acquisition de la TD et de la dormance, même si FUS3 est également impliqué dans le contrôle de la dormance. La majorité du contrôle de LEC1 sur la maturation est indirecte étant donné son profil d'expression précoce et sa position très en amont dans le réseau de régulation.

Contrôle de la transition maturation/germination de la graine

Plusieurs types de répresseurs de l'expression des régulateurs majeurs de la maturation des graines ont été identifiés. On retrouve les ligases (AIP2, KEG) et le facteur d'épissage SUA cités précédemment. Hormis ces répresseurs spécifiques d'un facteur donné, des répresseurs généraux de type facteurs de transcription ou facteurs chromatiniens sont exprimés en fin maturation pour permettre la transition vers la germination.

VAL1, VAL2 et VAL3 sont trois membres de la famille de facteur de transcription VAL (VP1/ABI3-like) caractérisés par la présence du domaine B3 et d'un domaine EAR (ERF-associated Amphiphilic Repression) (Tsukagoshi *et al.*, 2005). Ce motif conservé de 7 acides aminés suffit à conférer une activité répressive aux

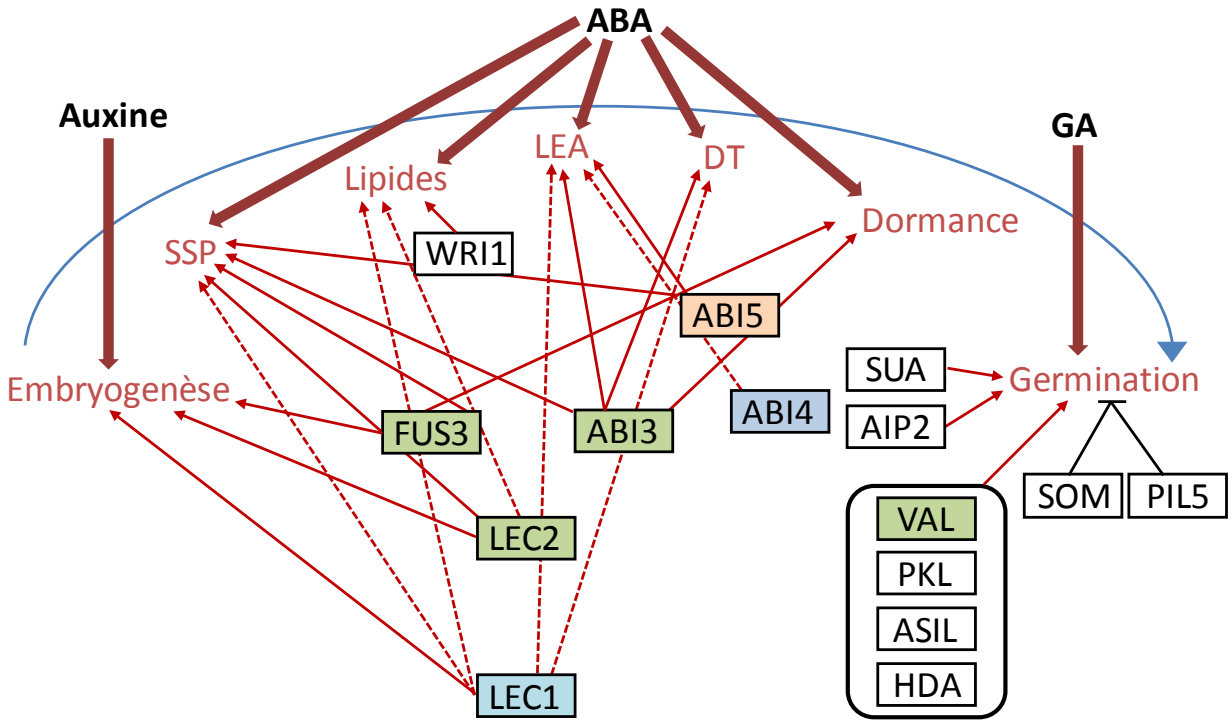


Figure 1.7. Modèle représentatif de l'implication des hormones et des régulateurs majeurs des graines orthodoxes dans les différents processus du développement.

L'auxine active l'embryogenèse alors que l'acide abscissique (ABA) est l'activateur majeur des processus de maturation : accumulation des protéines (SSP) et des lipides de réserve, des protéines LEA (Late Embryogenesis Abundant), mise en place de la tolérance à la dessiccation (DT) et de la dormance. Les gibbérellines (GA) stimulent la germination. Une flèche pleine indique une implication directe du facteur de transcription dans le processus alors qu'une flèche en pointillé signifie que l'action est vraisemblablement indirecte via l'activation d'autres facteurs de transcription (cf figure 1.6). Une action inhibitrice est représentée en noir. ABI3 = ABA-insensitive-3 ; ABI4 = ABA-insensitive-4 ; ABI5 = ABA-insensitive-5 ; AIP2 = ABI3-interacting protein E3 ligase ; ASIL = Arabidopsis 6b-interacting protein-1 Like-1 ; FUS3 = FUSCA3 ; HDA = histone deacétylase ; LEC1 = Leafy Cotyledon-1 ; LEC2 = Leafy Cotyledon-2 ; PIL5 = Phytochrome interacting factor-3 like-5 ; PKL = Pickle ; SOM = SOMNUS ; SUA = Suppressor of *abi3-5* ; VAL = Viviparous1-ABI3-Like ; WRI1 = Wrinkle-1.

facteurs de transcription les contenant (Ohta *et al.*, 2001). VAL1 et VAL2 sont également appelés HSI2 (High level expression of Sugar Inducible gene 2) et HSL1 (HSI2 Like 1) car l'étude du double mutant associé a montré que ces facteurs répriment des gènes de maturation de la graine liés à la signalisation par le sucre (Tsukagoshi *et al.*, 2007). L'absence de cette répression dans les mutants *val1*/*val2* conduit à l'expression de caractères embryonnaires dans les plantules et à une inhibition du gène *AtGA3ox1*, lié à la biosynthèse de GA (Suzuki *et al.*, 2007 ; Tsukagoshi *et al.*, 2007). Parmi les gènes réprimés par VAL1 et VAL2 on retrouve les régulateurs majeurs de la maturation *LEC1*, *L1L*, *ABI3*, *FUS3* (Suzuki *et al.*, 2007) (Figure 1.6). Cette répression pourrait être directe car les facteurs VAL se fixent sur les éléments RY présents dans les promoteurs de ces facteurs via leur domaine B3 (Guerriero *et al.*, 2009). Le facteur de transcription ASIL1 (Arabidopsis 6b-interacting protein-1 Like-1) joue le même rôle que VAL en réprimant dans les plantules l'expression de nombreux gènes spécifiques du développement de la graine dont *LEC1*, *LEC2*, *ABI3*, des oléosines et des protéines de réserve (Gao *et al.*, 2009) (Figure 1.6). D'après un modèle proposé par Guerriero *et al.* (2009), la répression du programme de maturation, dont ABI3 est un acteur central, pourrait se faire en deux temps au cours de la transition maturation/germination. Dans un premier temps, ces auteurs proposent une répression passive en fin de maturation au cours de laquelle l'activité d'ABI3 serait inhibée par déphosphorylation et perdrait son activité de transactivation du domaine RY. Puis une répression active s'établirait lors de la germination via l'activité des répresseurs VAL capable de fixer les éléments RY des promoteurs de gènes de maturation. La forme active de répression est plus permanente que la forme passive et pourrait être un préalable à l'intervention des facteurs de type chromatiniens (Guerriero *et al.*, 2009).

L'état de méthylation de la chromatine joue un rôle crucial dans le contrôle de l'expression des gènes, comme par exemple la triméthylation de l'histone H3 chez *A. thaliana*, responsable de la répression d'environ 4 400 gènes (Zhang *et al.*, 2007). Ces méthylations sont reconnus par des complexes PcG (Polycomb Group) capables de réprimer l'expression des gènes en formant de plus gros complexe comme PRC2 (Polycomb Repressive Complex 2) (revu dans Hennig & Derkacheva, 2009). Il a été montré que *FUS3* était ciblé par une catégorie de PcG (Makarevich *et al.*, 2006). De plus, le gène *BMI1* codant pour un PcG participe à initier la répression des gènes de maturation chez *A. thaliana* conjointement avec les gènes *VAL* (Yang *et al.*, 2013). Par Suzuky & McCarty (2008) proposent donc qu'un faible degré de méthylation H3K27me3 pour les gènes *VAL* couplé à une augmentation de cette méthylation sur les gènes *AFL* puisse expliquer la transition maturation/germination. D'autres facteurs de remodelage de la chromatine ont été mis en évidence dans la répression de la maturation des graines, comme PICKLE (PKL) qui a été particulièrement étudié. Ce facteur est exprimé lors de la germination pour éteindre les caractères embryonnaires dans la radicule (Ogas *et al.*, 1997 ; 1999) en réprimant notamment l'expression de *LEC1*, *LEC2*, *FUS3* (Dean Rider *et al.*, 2003), *ABI3* et *ABI5* (Perruc *et al.*, 2007) (Figure 1.6). La modification du degré de méthylation et d'acétylation des histones nécessite l'intervention d'histones deactylases. Chose intéressante, Tanaka *et al.* (2008) ont montré que l'expression de *LEC1*, *FUS3* et *ABI3* est réprimée par les histones déacétylases HDA6 et HDA19 (Figure 1.6). De plus, des expériences double hybride et de co-immunoprécipitation ont récemment montré que HDA19 interagit directement avec VAL2 (Zhou *et al.*, 2013). Alors que les facteurs VAL semblent impliqués dans l'initiation de la répression des gènes de maturation (Yang *et al.*, 2013), toutes les études sur les facteurs chromatiniens tels que PKL, PcG ou HDA confirment l'importance des mécanismes épigénétiques dans le maintien de la répression

au cours de la germination et de la croissance des plantules (Jia *et al.*, 2013). Ces mécanismes favorisent donc la germination (Figure 1.7).

1.2-Les graines récalcitrantes

1.2.1- Caractéristiques morphologiques et écologiques

Comme indiqué en début de chapitre, la majorité des graines sont orthodoxes. Cependant, un minimum de 281 espèces produisent des graines récalcitrantes (Gaff & Oliver, 2013), ce chiffre ne prenant en compte que les 7056 espèces pour lesquels le comportement des graines face à la dessiccation a été répertorié dans la base de données « Seed Information Database » du Kew Gardens. Ce chiffre est même deux fois plus important si on considère les espèces pour lesquelles la récalcitrance des graines est supposée mais pas démontrée (SID Kew). En extrapolant les données actuelles à l'ensemble des espèces connues, certains auteurs estiment que 20 000 espèces produiraient des graines récalcitrantes, soit 8 % de la flore mondiale (Daws *et al.*, 2006), et 70 % des arbres tropicaux produiraient des graines non orthodoxes (Berjak & Pammenter, 2000).

On retrouve des graines récalcitrantes dans toutes les familles végétales suggérant que la sensibilité à la dessiccation des graines est un caractère postérieur à la tolérance. La TD aurait été perdue plusieurs fois au cours de l'évolution (Tweddle *et al.*, 2003 ; Li & Pritchard, 2009). Les graines récalcitrantes forment donc un groupe hétérogène ayant pour point commun de ne pas survivre à l'état sec (0,1 g H_20/g MS) mais elles vont perdre leur viabilité à des teneurs en eau très variables. Ainsi sur 64 espèces récalcitrantes testées, Sun *et al.* (1999) ont déterminé que la teneur en eau critique variait de 0,1 à 1,4 g H_20/g MS. Berjak & Pammenter (2002) parle ainsi de continuum allant des graines orthodoxes jusqu'aux graines extrêmement sensibles à la dessiccation. Parmi les espèces les plus sensibles on retrouve *Avicienna marina,* arbre caractéristique des mangroves (Farrant *et al.*, 1997). Greggains *et al.* (2001) ont déterminé qu'aucune des graines produites par cet arbre n'était capable de germer après séchage à une teneur relative en eau de 47 % (soit 0,9 g H_2O/g MS) mais la capacité germinative est déjà altérée dès 57 % de teneur relative (1,3 g H_2O/g MS). A l'inverse, les glands de *Quercus robur* sont des graines récalcitrantes moins sensibles à la déshydratation puisque la capacité germinative de 50 % des graines est maintenue à des teneurs relatives en eau située entre 25 % et 32 % (soit 0,3 à 0,5 g H_2O/g MS) selon les années de récolte (Finch Savage & Blake 1994). Il est cependant difficile de donner des valeurs absolues de teneur en eau critique pour chaque espèce car cette valeur varie selon le tissu considéré et son stade de développement, la méthode utilisée pour la déterminer, la température et la vitesse de déshydratation appliquée lors de la dessiccation (Berjak & Pammenter, 2002). En effet, un séchage rapide limite la durée d'exposition des cellules à des niveaux intermédiaires d'hydratation particulièrement délétères et permet de maintenir la viabilité du tissu à des teneurs en eau plus basses que lors d'un séchage lent. Du fait de la taille importante des graines récalcitrantes, le séchage rapide est particulièrement efficace pour abaisser la teneur en eau tout en maintenant la survie des axes embryonnaires isolés.

Si aucune phylogénie claire n'a pu être déterminée concernant les graines récalcitrantes, des caractéristiques communes ont été mises en évidence à partir d'études corrélatives sur la masse, la forme, l'habitat, la teneur en eau à l'abscission et la vitesse de germination (Hong & Ellis, 1998 ; Dussert *et al.*, 2000 ; Tweddle *et al.*, 2003 ; Pritchard *et al.*, 2004). Un modèle prédictif basé sur la masse des graines et le SCR (Seed Coat Ratio) a même été déterminé à partir de l'étude de 104 espèces appartenant à 37 familles

Tableau 1.1. Conditions optimales et délétères de stockage de graines récalcitrantes (Bewley *et al.*, 2013).

Espèces	**Longévité (% germin)**	**Conditions stockage optimal**	**Conditions délétères**
Avicennia marina	Quelques jours	humide	Séchage
Threoboma cacao	8 à 10 semaines	21-27°C dans cabosses recouvertes de fongicide	Séchage ; < 13°C
Hevea brasiliensis	4 mois (3%)	7-10°C dans sciure humide dans des sacs perforés en polyéthylène	Teneur en eau relative < 20%
Castanea crenata	6 mois	0-3°C dans sacs de polyethylène ou boîtes ventilées	Excès d'eau ou séchage < 0°C
Cocos nucifera	16 mois	T°C ambiante forte humidité relative	Séchage
Quercus borealis	20 mois (50%)	5°C dans boîtes scellées	Teneur en eau relative < 20%-40%

différentes (Daws *et al.*, 2006). Il en ressort que ces graines sont généralement grosses et entourées d'un fin tégument (valeur faible du SCR). Les graines récalcitrantes sont disséminées à une teneur en eau élevée mais ce critère n'est pas fiable car certaines graines orthodoxes le sont également dans l'échantillonnage de Daws *et al.* (2006).

D'un point de vue écologique, les graines récalcitrantes sont souvent d'origine tropicale humide mais pas exclusivement (Tweddle *et al.*, 2003 ; Pritchard *et al.*, 2004). Cependant, la sensibilité de ces graines à la dessiccation pourrait être liée à l'habitat et au climat, celles des régions tempérées étant visiblement moins sensibles que celles des régions tropicales (Farrant *et al.*, 1996). Daws *et al.* (2004) ont d'ailleurs montré une influence de l'environnement sur la sensibilité à la dessiccation des graines récalcitrantes de marronnier (*Aesculus hippocastanum*). Les graines écossaises s'étant développé dans les conditions les plus froides sont moins tolérantes à la dessiccation que celles issues d'arbres grecs vivant sous des températures plus élevées (Daws *et al.*, 2004). Cette différence pourrait s'expliquer par un développement plus abouti en Grèce, visible par une MS plus importante et une teneur en eau à l'abscission plus basse. D'autres facteurs environnementaux influencent la viabilité des glands de *Quercus Ilex* en conditions naturelles (capacité germinative et développement normal des plantules) comme la quantité de précipitations pendant l'hiver, la température, l'ombre et le sol (Joët *et al.*, 2013). Bien que peu de graines récalcitrantes soient dormantes (Tweddle *et al.*, 2003), certaines espèces peu sensibles à la déshydratation comme le marronnier disséminent des graines dormantes à l'automne qui ne germeront qu'au printemps suivant après une période de froid (Pritchard *et al.*, 1996).

Les possibilités de stockage à moyen et long terme sont actuellement limitées pour les graines récalcitrantes et n'excède pas deux ans pour les espèces les moins sensibles (Berjak & Pammenter, 2008). Les conditions optimales de stockage sont variables d'une espèce à l'autre mais généralement une atmosphère humide et les températures supérieures à 10°C sont favorables. Dans ces conditions, les graines des espèces tropicales peuvent être conservées de quelques jours pour *A. marina* à quelques mois pour *Hevea brasilensis* voire plus d'un an pour *Cocos nucifera* (Tableau 1.1). Une température supérieure à 20°C favorise le développement de champignons et l'utilisation de fongicides est nécessaire dans certains cas, comme pour les graines de *Threoboma cacao*. Pour les graines d'origine tempérée moins sensibles à la dessiccation, la température peut être abaissée et le stockage prolongée à près de deux ans chez *Quercus borealis*. La cryo-préservation d'axes embryonnaires a été développé avec succès chez certaines espèces récalcitrantes et permet d'envisager un stockage à long terme dans des banques (Li & Pritchard, 2009). A titre de comparaison, 95 % d'un lot de graines orthodoxes de pois ont germé après 65 ans de stockage à 5°C (40 ans) puis -18°C (25 ans) (Walters *et al.*, 2005a).

1.2.2-Les phases du développement

Le développement des graines récalcitrantes diffère principalement par l'absence de phase de maturation tardive (Berjak & Pammenter, 2000) (Figure 1.8). L'accumulation des réserves et l'activité métabolique se poursuivent jusqu'à l'abscission et les graines sont disséminées à des teneurs en eau généralement très élevées comparées aux graines orthodoxes : de 0,4 à 4,4 g H_20/g MS selon les espèces (Berjak & Pammenter, 2008). Elles restent sensibles à la dessiccation tout au long de leur développement

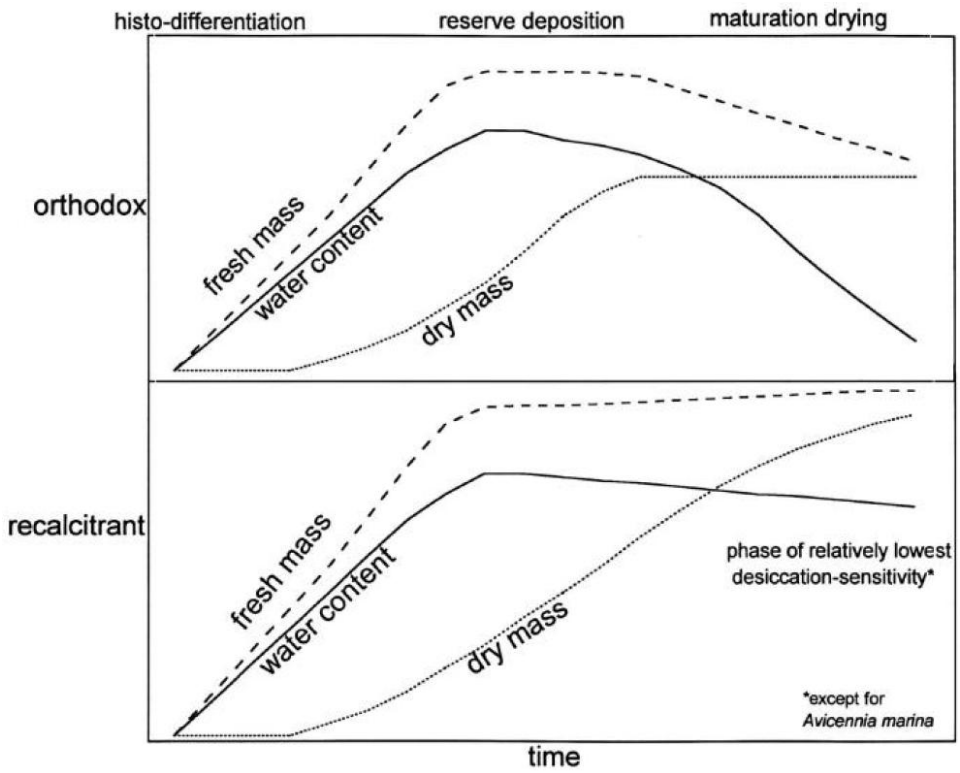


Figure 1.8. Evolution du poids frais, du poids sec et de la teneur en eau au cours du développement des graines orthodoxes et récalcitrantes (Pammenter & Berjak, 2000).

Le développement des graines orthodoxes est caractérisé par une phase de maturation tardive après l'accumulation des réserves, au cours de laquelle la teneur en eau diminue. Chez les graines récalcitrantes, cette phase n'existe pas et l'accumulation des réserves se poursuit jusqu'à l'abscission de la gousse. L'absence de phase de dessiccation dans le développement engendre une dissémination à des teneurs en eau plus élevées que chez les graines orthodoxes.

même si une diminution de la sensibilité a été observée dans les graines matures de *Quercus robur* (Finch Savage, 1992), *Camellia sinensis* (Berjak *et al.*, 1993) ou *A. hippocastanum* (Tompsett & Pritchard, 1993). Cependant la notion de graine mature est difficile à généraliser chez les graines récalcitrantes car, selon les années, les graines de *Quercus robur* ne sont pas disséminées à la même teneur en eau et celles qui ont la teneur en eau la plus faible sont plus tolérantes à la dessiccation (Finch Savage & Blake, 1994). De plus, les graines de certaines espèces de mangrove germent dans la continuité du développement comme *A. marina* (Farrant *et al.*, 1993a et b) et d'autres sont même semi-vivipares (germination initiée avant l'abscission) comme *Rhizophora mucronata* (Ismail *et al.*, 2010). Ces différences de développement au sein des graines récalcitrantes ont amené certains auteurs à proposer deux catégories de graines sensibles à la dessiccation (Farrant *et al.*, 1997). Le développement de celles des régions tempérées pourrait être assimilé à un développement tronqué de graines orthodoxes, elles sont disséminées avant acquisition complète des mécanismes liés à la TD et restent donc sensible à la dessiccation (Finch Savage & Blake, 1994). En revanche, les graines produites par les espèces tropicales n'entrent pas dans la phase d'acquisition de la TD mais se prépare à germer, elles restent donc très sensibles à la dessiccation (Farrant *et al.*, 1997).

1.2.3-Régulation du développement

La teneur en hormones, et particulièrement en ABA, a été étudiée chez certaines graines récalcitrantes plus ou moins sensibles à la dessiccation. Chez le cacaoyer, un pic de teneur en ABA (de 1 à 3 µg d'ABA/g MF) a été mesuré en début de maturation (Pence *et al.*, 1991). Les glands de *Quercus robur* présentent dans les axes et les cotylédons un pic de teneur en ABA semblable à ce qui est observé durant le développement des graines orthodoxes de *Phaseolus vulgaris* (Finch Savage & Blake, 1994). Cependant, chez d'autres espèces tropicales, la teneur en ABA libre reste basse tout au long du développement, suggérant une faible synthèse ou un turn-over important (Farrant *et al.*, 1993c ; Garello & Le Page-Degivry, 1995 ; Farnsworth & Farrant, 1998 ; Farnsworth, 2000). Farrant *et al.* (1996) ont observé un lien intéressant entre la teneur en ABA et l'origine géographique des espèces récalcitrantes. Une teneur en ABA inférieure à 50 ng/g MS a été mesurée dans les axes matures des graines d'espèces tropicales alors que chez les espèces tempérées, généralement moins sensibles à la dessiccation, cette teneur est supérieure à 250 ng/g MS. De plus, seules les espèces présentant une teneur en ABA élevée accumulent des LEA (Farrant *et al.*, 1996). Cependant, la faible teneur en ABA ne suffit pas à expliquer l'absence de TD car la sensibilité à l'ABA est également à prendre en compte. En effet, un apport exogène d'ABA n'induit pas une forte expression des déhydrines dans les graines de *Zizania palustris* (Bradford & Chandler, 1992) et n'inhibe pas la germination des graines de *Hopea odorata* (Garello & Le Page-Degivry, 1995).

Aucune donnée moléculaire n'est disponible concernant l'expression des gènes homologues aux gènes codant les facteurs de transcription majeurs des graines orthodoxes au cours du développement des graines récalcitrantes. Seule une étude a montré que l'expression du gène *ABI3* dans les graines récalcitrantes de *Magnolia ovata* n'augmentait pas en réponse au séchage (José *et al.*, 2009). Un des objectifs des études menées dans le cadre de cette thèse est de combler ce manque de connaissance.

2-Tolérance à la dessiccation

2.1-Définition et distribution dans le monde vivant

Le potentiel hydrique des plantes est très élevé au point qu'on estime que l'humidité relative de l'air environnant doit approcher 99,6 % pour que les potentiels hydriques de la plante et de l'air soient à l'équilibre (Wood & Jenks, 2007). Cette valeur d'humidité relative de l'air n'est jamais atteinte (généralement elle varie entre 50 % et 90 %) ce qui explique l'efflux permanent d'eau des feuilles vers l'atmosphère. En cas de condition limitante en eau, les plantes peuvent rapidement se retrouver en situation de déficit hydrique (ou sécheresse) et adapter leur métabolisme et leur développement en conséquence. Les mécanismes mis en jeu permettent de limiter la perte en eau et s'assimilent à une stratégie d'évitement afin de maintenir une teneur en eau cellulaire supportable pour la plante (Walters *et al.*, 2002). La dessiccation est une notion bien différente du déficit hydrique. Elle est définie comme la perte totale de l'eau libre cellulaire au profit de l'atmosphère et l'organisme atteint une teneur en eau de 0,1 g H_2O/g MS (Alpert, 2005 ; Wood & Jenks, 2007). Un organisme est dit tolérant à la dessiccation s'il est capable de sécher en équilibre avec l'air et de reprendre une activité métabolique normale après réhydratation (Bewley, 1979). La tolérance à la dessiccation (TD) n'est donc pas une stratégie d'évitement du déficit hydrique mais bien une capacité à survivre à l'état sec et des organismes tolérant à la sécheresse ne sont pas forcément tolérant à la dessiccation (Walters *et al.*, 2002).

La TD est une propriété très importante permettant aux organismes vivants de pallier au manque d'eau durant les périodes peu favorable à la vie. Elle est également considérée comme un facteur majeur ayant permis l'évolution des plantes terrestres suite à la sortie des algues hors de l'eau (Oliver *et al.*, 2000). On retrouve des organismes tolérants à la dessiccation parmi une grande variété d'espèces vivantes comme les bactéries (cyanobactéries), les champignons, les invertébrés (tardigrades, rotifères, nématodes) et les plantes (Alpert, 2005 ; Berjak, 2006). Au sein de ces dernières, la TD a été identifiée dans les parties végétatives d'espèces appartenant à toutes les grandes classes, à l'exception des Gymnospermes (Wood & Jenks, 2007). Par exemple, 210 Bryophytes (1 % des espèces testées) (Wood, 2007) et 330 plantes vasculaires (Porembski & Barthlott, 2000) ont été recensées. Chez les Angiospermes, la TD est assez rare dans les tissus végétatifs, environ 135 espèces appelées plantes reviviscentes, mais très courante dans les organes de dissémination comme les graines et les pollens (Gaff & Oliver, 2013).

Toutes ces espèces et/ou organes tolérants à la dessiccation n'ont pas développé les mêmes mécanismes pour survivre à l'état sec. Alors que les bryophytes et lichens tolèrent une dessiccation très rapide sans contrôle du flux d'eau, les structures plus complexes comme les plantes vasculaires et les graines ont besoin d'une dessiccation plus lente afin de mettre en place les mécanismes de tolérance (Proctor & Tuba, 2002 ; Wood & Jenks, 2007). La TD des bryophytes est vraisemblablement un caractère primitif tandis que celles des plantes vasculaires pourrait être le fruit de plusieurs évolutions indépendantes dérivant de celles des graines (Oliver *et al.*, 2000 ; Proctor & Tuba 2002 ; Illing *et al.*, 2005 ; Oliver & Gaff, 2013). Il n'est donc pas surprenant de retrouver des mécanismes communs entre l'acquisition de la TD des graines et des plantes reviviscentes (Farrant & Moore, 2011). Cependant une des spécificités des graines est de programmer l'acquisition de la TD comme une partie intégrante du développement et non comme une réponse à des contraintes environnementales (Illing *et al.*, 2005 ; Berjak, 2006, Berjak *et al.*, 2007). La suite de cette

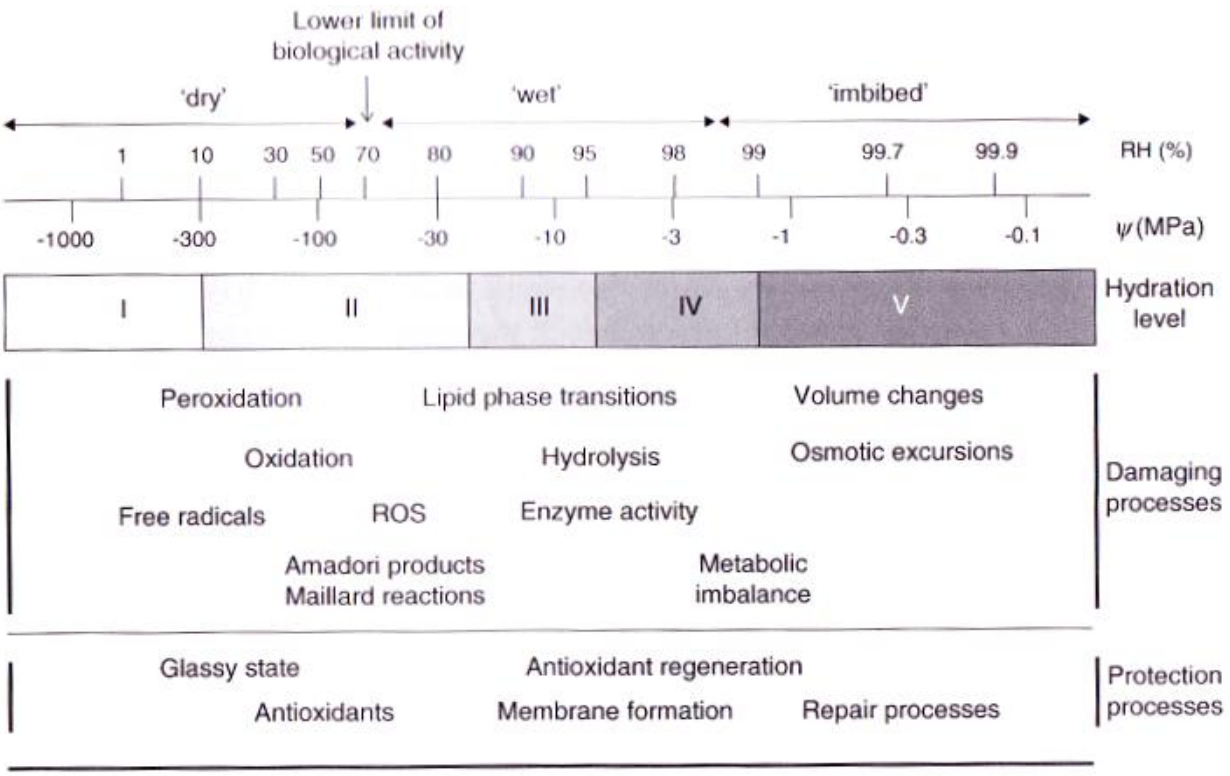


Figure 1.9. Niveaux d'hydratation, processus de détérioration et de protection dans la graine au cours de la dessiccation (Bewley *et al.*, 2013).

Les différentes échelles du haut indiquent l'humidité relative (RH), le potentiel hydrique (Ψ), et le stade d'hydratation correspondant. Les processus de détérioration et de protection mis en place à ces différents niveaux d'hydratation sont figurés en dessous. Aucune activité métabolique et biochimique n'ayant lieu en dessous de 70% RH une distinction entre des processus « secs » et des processus « humides » a été ajoutée. « imbibed » correspond aux processus nécessitant un contact avec l'eau liquide. ROS = espèces réactives de l'oxygène.

introduction se focalisera principalement sur la TD des graines, mais quelques résultats intéressants chez les plantes reviviscentes seront présentés pour certains mécanismes.

2.2-Dommages causés par la dessiccation

A l'échelle cellulaire et moléculaire, l'eau joue de multiples rôles : (1) structurel en assurant la turgescence, la conformation des macromolécules et la mobilité intracellulaire ; (2) protecteur en évitant l'interaction entre des groupements réactifs de différentes macromolécules ; (3) métabolique en tant que réactif ou produit de nombreuses réactions. Pour toutes ces raisons, le retrait de l'eau entraîne de fortes perturbations néfastes pour l'intégrité cellulaire (Walters *et al.*, 2002). Des études thermodynamiques concernant les propriétés de l'eau ont permis de définir cinq stades d'hydratation différentes au cours de la dessiccation des graines allant du stade V (graine imbibée) au stade I (état sec avec une teneur en eau inférieure à 0,08 g H_2O/g MS) (Vertucci, 1990 ; Walters *et al.*, 2005b) (Figure 1.9). Au cours des premières phases de la dessiccation (stade IV correspondant à une teneur en eau comprise entre 0,7 et 0,45 g H_2O/g MS), la graine se retrouve dans des conditions semblables à un tissu végétatif soumis à un déficit hydrique. Cette phase est peu étudiée chez les graines mais il est vraisemblable que des mécanismes généraux de réponses aux stress abiotiques soient mis en place pour faire face à cette déshydratation (Berjak *et al.*, 2007). C'est également à ce stade que le stress mécanique débute, compensé par l'accumulation de réserves dans des vacuoles spécifiques (Farrant & Moore, 2011). Le stade III (teneur en eau comprise entre 0,45 et 0,25 g H_2O/g MS) constitue un stade intermédiaire au cours duquel le métabolisme, bien que limité, est toujours actif mais les mécanismes de détérioration sont de plus en plus présents (Walters *et al.*, 2005b ; Farrant & Moore, 2011 ; Bewley *et al.*, 2013). Tous les processus métaboliques ne sont pas ralentis de façon simultanée et ce déséquilibre engendre la formation d'espèces réactives de l'oxygène (ROS). Les réactions de Maillard sont des réactions de dégradation complexes initiées par l'attaque d'un groupement aldéhyde de sucres réducteurs (monosaccharides) sur le groupement amine de protéines. Les conditions hydriques du stade III favorisent la mise en place de ces réactions car la teneur en eau est encore assez élevée pour permettre la réaction hydrolytique initiale mais la proximité entre espèces réactives est importante (Sun & Leopold, 1995 ; Walters & Koster, 2007). Par conséquent, une exposition prolongée à ce stade réduit la longévité des graines car tous les mécanismes de tolérance à la dessiccation ne sont pas encore acquis alors que le risque oxydatif est maximum (Walters *et al.*, 2005b ; Berjak *et al.*, 2007). A des teneurs en eau inférieures (stades II et I), la mobilité cellulaire est fortement réduite et les graines atteignent progressivement l'état sec. La plupart des graines orthodoxes voient leur longévité réduite au stade I, considéré comme ultra sec (teneur en eau inférieure à 0,08 g H_2O/g MS). Le stade II correspond donc à l'optimum de conservation des graines orthodoxes (teneur en eau autour de 0,1 g H_2O/g MS) (Walters *et al.*, 2005b. Il est intéressant de noter que la viabilité des graines récalcitrantes est perdue aux stades III ou IV selon la sensibilité de l'espèce alors que les graines intermédiaires meurent généralement au stade II (Walters *et al.*, 2002).

En résumé, l'ensemble des dégradations auxquelles la cellule doit faire face au cours de sa dessiccation peuvent être regroupées en trois catégories : (1) les dommages structuraux dus à la réduction du volume cellulaire ; (2) les dommages liés au désordre métabolique que rencontre la cellule au cours de la perte en eau ; (3) la dénaturation des macromolécules consécutive au retrait de l'eau recouvrant leur surface

(Pammenter & Berjak, 1999 ; Walters *et al.*, 2002). Si maintenir l'intégrité cellulaire et moléculaire au cours de la déshydratation est le premier défi à relever pour les organismes tolérants à la dessiccation, il faut encore survivre à l'état sec et maintenir l'intégrité au moment de la réhydratation. Pour répondre à toutes ces exigences, une large panoplie de mécanismes, qui commencent à être mis en évidence, est requise.

2.3-Les mécanismes impliqués dans la TD

2.3.1-Adaptations structurales

La dessiccation s'accompagne inévitablement d'une réduction du volume cellulaire et la paroi joue alors un rôle majeur dans le maintien de l'intégrité structurale lors de la perte en eau mais également de la réhydratation. Chez les graines orthodoxes comme chez les plantes reviviscentes des observations microscopiques ont révélé des repliements caractéristiques de la paroi permettant de maintenir l'association avec le plasmalesmme (Webb & Arnott, 1982 ; Vicré *et al.*, 1999). La structure du repliement varie selon les espèces et la composition biochimique des parois (Webb & Arnott, 1982). Peu d'études de l'évolution de la paroi au cours de la dessiccation ont été réalisées chez les graines. Koch *et al.* (1999) ont montré que l'activité de la pectinesterase augmentait d'un facteur 5 au cours de la dessiccation dans les cotylédons de *Glycine max*. Cette activité est à l'origine d'une diminution du pourcentage de pectine methylesterification et donc d'une plus grande plasticité de la paroi. Cette propriété est très importante pour faire face aux variations de volume cellulaire au cours de la déshydration et de l'imbibition. Davantage d'études ont été menées chez les plantes reviviscentes à ce niveau. Chez *Crasterostigma wilmsii* certaines modifications biochimiques de la paroi ont pu être mises en évidence entre feuille sèches et hydratées : diminution de la teneur en glucose dans la fraction hemicellulosique des parois et augmentation des substitutions en galactose dans les xyloglucanes (Vicré *et al.*, 2004a). De même la concentration en ion calcium augmente significativement dans les parois de *C. wilmsii* mais pas dans celles de *Saintpaulia ionantha*, espèce sensible à la dessiccation. Cette modification pourrait jouer un rôle majeur dans l'architecture des parois à l'état sec en stabilisant les pectines (Vicré *et al.*, 2004b).

Le degré de vacuolisation et la proportion d'endomembranes diminuent au cours du développement des graines orthodoxes de *Phaseolus vulgaris*, permettant de réduire l'exposition de membranes susceptibles d'être détériorées par le retrait d'eau. Au contraire, ce degré de vacuolisation reste élevé tout au long du développement des graines très récalcitrantes d'*Avicennia marina* (Farrant *et al.*, 1997). Le cytosquelette subit lui aussi des modifications au cours de la dessiccation. Faria *et al.* (2005) ont montré que les microtubules sont entièrement dissociés sous forme de tubuline dans les radicules de *M. truncatula* à l'état sec et que dès la germination les filaments de microtubules sont de nouveaux visibles. Chez l'espèce récalcitrante *Inga vera*, ces microtubules sont visibles dans les cellules de graines matures fraichement récoltées mais l'abaissement de la teneur en eau dans ces graines de 1,8 à 0,75 g H_2O/g MS conduit à leur dissociation irréversible, y compris après réhydratation (Faria *et al.*, 2004). D'autres travaux étaient arrivés à la même conclusion concernant les filaments d'actine avec les graines récalcitrantes de *Quercus robur* (Mycock *et al.*, 2000 cité dans Berjak & Pammenter, 2000) et *Trichilia dregeana* (Gumede *et al.*, 2003 cité dans Berjak & Pammenter, 2008).

2.3.2-Ralentissement du métabolisme

L'activité métabolique est incompatible avec une faible teneur en eau mais certaines voies métaboliques sont plus sensibles que d'autres à la diminution de la teneur en eau. Par exemple, la croissance et la synthèse protéique sont ralenties à des teneurs en eau encore élevées alors que la respiration, bien que ralentie, se poursuit jusqu'à des teneurs en eau correspondant au stade III (Leprince & Hoekstra, 1998 ; Walters *et al.*, 2002). Les gènes codant des enzymes majeures du métabolisme primaire et énergétique sont réprimés au cours de l'acquisition de la TD (Buitink *et al.*, 2006). Il semble que la diminution de l'activité respiratoire chez les graines orthodoxes n'est pas qu'une conséquence de la diminution de la teneur en eau, mais bien une adaptation mise en place avant la perte en eau (Leprince *et al.*, 2000). L'augmentation de la viscosité cellulaire limite la diffusion du O_2 et la demande en ATP doit diminuer en conséquence pour maintenir un métabolisme équilibré et éviter l'hypoxie durant la dessiccation (Leprince & Hoekstra, 1998). A teneur en eau égale, les radicules non-germées de concombre qui sont tolérantes à la dessiccation émettent deux fois moins de CO_2 que les radicules germées, qui sont devenues sensibles. Les tissus tolérants à la dessiccation réduisent donc leur demande en ATP avant la perte en eau ce qui n'est pas le cas des tissus sensibles, ces derniers se retrouvent donc en hypoxie au cours du séchage (Leprince *et al.*, 2000). Une étude comparative entre espèces orthodoxe (*Phaseolus* vulgaris) et récalcitrante (*Avicennia marina*) a confirmé que l'activité respiratoire restait élevée tout au long du développement de la graine sensible alors qu'il était ralenti en parallèle de l'acquisition de la TD chez *P. vulgaris* (Farrant *et al.*, 1997). La régulation du métabolisme est un moyen de limiter la production de ROS au cours des stades intermédiaires de la dessiccation (stade III) et certaines études ont montré qu'elle était couplée à une dédifférenciation des mitochondries permettant ainsi de réduire les surfaces membranaires (Farrant *et al.*, 1997 ; Pammenter & Berjak, 1999).

En dehors du métabolisme énergétique, d'autres processus sont stoppés en lien avec l'acquisition de la TD chez les graines orthodoxes, comme le cycle cellulaire. Au cours du séchage des graines orthodoxes de *Vicia faba*, la réplication de l'ADN est un des premiers processus à être stoppé et les cellules restent bloquées en phase G1 (Brunori, 1967 cité dans Pammenter & Berjak, 1999). Chez *A. marina* la réplication de l'ADN n'est arrêtée que quelques heures autour de l'abscission (Boubriak *et al.*, 2000). L'activité métabolique élevée des graines récalcitrantes permet un développement rapide des plantules mais peut aussi expliquer leur sensibilité à la dessiccation.

Cependant, il ne faut pas envisager la graine au cours du séchage comme un organe sans la moindre activité moléculaire. Par exemple, la phase de post-maturation décrite au cours du développement des graines orthodoxes est cruciale pour préparer la graine en vue d'une phase de quiescence et de sa germination (Angelovici *et al.*, 2010). Bien que le métabolisme soit ralenti, il n'est pas stoppé et une activité transcriptionnelle semble se poursuivre (Bove *et al.*, 2005, Cadman *et al.*, 2006). Parmi les gènes sur-exprimés les processus biologiques associés ne sont pas forcément liés à la dessiccation mais sont caractéristiques de la germination des graines (machinerie traductionnelle, protéines du métabolisme énergétique et de la réparation de l'ADN) (Angelovici *et al.*, 2010). Cependant, ces hypothèses sur l'activité transcriptionnelle durant la post-maturation sont débattues et certains auteurs ont montré qu'une oxydation sélective de certains transcrits durant le stockage des graines pouvait expliquer la différence apparente d'expression des gènes (Bazin *et al.*, 2011).

Même si des modifications structurales et métaboliques permettent de limiter l'exposition des macromolécules au manque d'eau, l'impact de la dessiccation sur les structures restantes est majeur : dénaturation et agrégation des protéines, oxydation et rupture des membranes phospholipidiques. Des mécanismes protecteurs sont donc également mis en place, comme les protéines chaperonnes, la réduction de la mobilité cellulaire ou les défenses antioxydantes. Ils sont détaillés dans les paragraphes suivants.

2.3.3-LEA

a)-Définition et classification

L'abondance de la littérature concernant les protéines LEA (Late Embryogenesis Abundant) montre bien que ces molécules sont très étudiées en lien avec la TD. Comme leur nom l'indique, les ARNm codant ces protéines ont été initialement identifiés comme fortement exprimés durant la maturation tardive des cotylédons de coton par rapport aux autres stades de développement (Dure *et al.*, 1981 ; Galau & Dure, 1981). Les premiers travaux ont révélé que tous ces ARNm étaient induits précocement dans des embryons excisés cultivés en présence d'ABA (Dure *et al.*, 1981) et que 13 des 18 ARNm clonés étaient induits en réponse à l'ABA exogène (Galau *et al.*, 1986). Par la suite, un nombre croissant de protéines homologues ont été identifiées chez d'autres espèces en dehors du règne végétal, complexifiant la définition et la nomenclature de cette famille (Tunnacliffe & Wise, 2007 ; Battaglia *et al.*, 2008). Ces protéines ont été identifiées en réponse à de nombreux stress abiotiques et à la dessiccation et partagent un certain nombre de caractéristiques structurales : ce sont de petites protéines généralement hydrophiles, stables à la chaleur, comportant des zones de faible complexité (motifs répétés) et présentant, pour la plupart, une structure majoritairement désordonnée en condition hydratée (Bray, 1993 ; Garay Arroyo *et al.*, 2000 ; Wise & Tunnacliffe, 2004 ; Tunnacliffe & Wise, 2007 ; Battaglia *et al.*, 2008).

La nomenclature et la classification des protéines LEA ont subi de nombreuses modifications depuis leur mise en évidence. Tout d'abord, Dure *et al.*, (1989) ont classé les protéines identifiées chez plusieurs espèces végétales en trois groupes en se basant sur la présence de motifs dans les séquences d'acides aminés. Puis, toujours en se basant sur la similarité entre séquences, trois nouveaux groupes ont été créés et renommés pour donner les groupes : D-7, D-11, D-19, D-29, D-34, D-113 (cités dans Bray, 1993 ; Wise, 2003). Certaines LEA ont également été nommées déhydrines en référence à leur surexpression en réponse à un stress hydrique dans des plantules de maïs et d'orge (Close *et al.*, 1989). En 2003, Wise répertorie 112 protéines LEA dans les bases de données internationales Swissprot et SpTrEMBL et montre que l'assignation de toutes ces protéines dans les groupes existant est de plus en plus difficile et parfois incohérente. Le développement d'outils bioinformatiques tels que l'analyse de profil peptidique (POPP) (Wise, 2002) et les domaines Pfam (Protein Family database) (Bateman *et al.*, 2002) a permis de repenser la classification. L'analyse POPP (Protein or Oliginucleotide Probability Profile) permet de classer les séquences en cluster selon leur composition en acides aminés plutôt qu'en fonction de la similarité de séquences et se révèle particulièrement pertinente pour les protéines LEA qui possèdent des zones de faible complexité (Wise & Tunnacliffe, 2004). En effet, ces zones posent des problèmes pour l'alignement local des séquences et donc pour la comparaison de leur similarité. Ainsi, toutes les LEA ont pu être reclassées dans de nouveaux groupes en fonction de la sur- ou sous-représentation de certains acides aminés (Wise, 2003). Par ailleurs, deux analyses complètes et indépendantes

Tableau 1.2. Différentes classifications des protéines LEA.
D'après Hundertmark & Hincha, 2008 ; Wise, 2003 ; [a] Battaglia *et al.*, 2008)

Pfam	**Hundertmark & Hincha 2008**	**Wise 2003**	**Cuming 1999**	**Bray 1993**	**Dure *et al.* 1989**
PF00257	déhydrines	IIa ; IIb	2	groupe 2	D-11
PF03760	LEA_1	IIa ; III	2 ; 3	groupe 4	D-113
PF03168	LEA_2	VI	5	/	D95
PF03242	LEA_3	V	/	/	D73
PF02987	LEA_4	III	3	groupe 3 ; groupe 5	D-7 ; D-29
PF00477	LEA_5	I	1	groupe 1	D-19
PF04927	SMP	IV	5	groupe 6	D-34
PF10714 [a]	PvLEA18	/	/	/	/
PF02496 [a]	AtM	/	/	/	/

du génome d'*A. thaliana* ont permis d'identifier 51 (Hundertmark & Hincha, 2008) et 50 (Bies-Etheve *et al.*, 2008) gènes codant pour des LEA classés en neuf groupes. Dans la première étude, l'ensemble de ces gènes ont été classés en neuf groupes sur la base des domaines Pfam identifiés dans les séquences. Dans la suite du mémoire, nous suivrons cette classification Pfam pour les LEA identifiées chez *M. truncatula* et *C. australe* et la correspondance entre toutes les nomenclatures est donnée dans le tableau 1.2 pour permettre de faciliter les références aux travaux précédents. Cependant, la classification des protéines LEA n'est toujours pas figée et des évolutions continuent à être proposées. Par exemple, une étude récente et exhaustive des 710 séquences codant pour des protéines LEA répertoriées dans la base de données LEAPdb a permis de distinguer des sous groupes au sein des familles Pfam et d'arriver à 12 classes (Jaspard *et al.*, 2012). Les auteurs se basent sur les similarités et sur les propriétés biochimiques des différentes séquences protéiques.

b)-Caractéristiques biochimiques, structurales

Une des caractéristiques majeures des protéines LEA est leur hydrophilicité élevée, visible par leur profil d'hydropathie ou leur GRAVY (Grand Average of Hydropathy) très négatifs (Baker *et al.*, 1988 ; Dure *et al.*, 1989 ; Tunnacliffe & Wise, 2007 ; Hundertmark & Hincha, 2008 ; Battaglia *et al.*, 2008). Cette caractéristique s'explique notamment par la richesse en acide aminés polaires et est à la base des propriétés structurales et biochimiques comme la structure désordonnée à l'état hydratée. Cependant certaines caractéristiques sont propres à chaque famille comme la composition en acide aminées, des motifs particuliers ou la présence de structure native ordonnée chez certaines LEA (Jaspard *et al.*, 2012). Nous allons donc faire un bref aperçu des données disponibles famille par famille.

Dehydrines

Ces LEA font partie des plus hydrophiles. Elles sont riches en acides aminés polaires et chargés et généralement dépourvus d'acides aminés aromatiques et de cystéine (Wise & Tunnacliffe, 2004 ; Battaglia *et al.*, 2008). Une des raisons pouvant expliquer l'absence de cystéine est la capacité de cet acide aminé à créer des ponts disulfures, un élément très structurant pour les protéines et qui est incompatible avec la structure désordonnée des déhydrines. Au contraire la richesse en histidine observées chez les déhydrines semble être une caractéristique des protéines désordonnées (Jaspard *et al.*, 2012). Ces LEA sont caractérisées par un motif de 15 acides aminés EKKGIMDKIKEKLPG à proximité de l'extrémité C-terminale (segment-K), d'une succession de résidus sérine (segment-S) et d'un motif conservé T/VDEYGNP près de l'extrémité N-terminale (segment-Y) (Close *et al.*, 1993 ; Close, 1997). Le segment-K est toujours présent mais pas forcément les deux autres, ce qui a conduit à la création de sous catégories selon la présence ou l'absence et la position des segments S et Y. Le segment-K est répété de une à 11 fois et le segment-Y jusqu'à 35 fois selon les protéines. Le segment-S est parfois phosphorylé comme démontré pour RAB-17 chez le maïs (Vilardell *et al.*, 1990 ; Plana *et al.*, 1991) ou chez *C. plantagineum* (Röhrig *et al.*, 2006) ce qui montre l'importance des régulations post traductionnelle dans l'activité de ces protéines. La forte proportion d'acides aminés chargés permet de diviser les déhydrines en deux groupes : acide et neutre/basique, nous verrons par la suite que les déhydrines acides ont des propriétés particulières. L'analyse structurale par dichroïsme circulaire a montré que ces protéines sont en grande partie désordonnées à l'état hydratée (Ismail *et al.*, 1999 ; Hara *et al.*, 2001 ; Soulages *et al.*, 2003 ; Mouillon *et al.*, 2006, Shih *et al.*, 2012) mais peuvent adopter une structure en hélice α au cours du séchage (Shih *et al.*,

2012), favoriser l'apparition de structure poly-L-proline (PII) aux basses températures (Soulages *et al.*, 2003) ou rester désordonnée en présence de sels stabilisants, d'ions métalliques ou de température élevée (Mouillon *et al.*, 2006).

LEA_5

Ces LEA sont également très hydrophiles, riches en acides aminés chargés et en glycine (environ 18 %). Elles contiennent un motif hydrophile de 20 acides aminés GGQTRREQLGEEGYSQMGRK (Galau *et al.*, 1992) qui est répété de une à quatre fois selon les protéines. Chez les plantes, deux autres motifs conservés ont été identifiés (Battaglia *et al.*, 2008 ; Bies-Etheve *et al.*, 2008). D'un point de vue structural, les LEA-5 partagent beaucoup de caractéristiques avec les déhydrines. Plusieurs études structurales par NMR (Eom *et al.*, 1996), dichroïsme circulaire (Soulages *et al.*, 2002) ou FTIR (Boudet *et al.*, 2006 ; Shih *et al.*, 2012) ont confirmé que ces protéines sont majoritairement désordonnées à l'état natif en milieu hydraté mais qu'elles se structurent au cours de la dessiccation. Par exemple MtEM6 voit sa proportion de structure désordonnée diminuer de 64 % à 32 % au profit des hélices α au cours du séchage, qu'il soit lent ou rapide (Boudet *et al.*, 2006). Soulages *et al.* (2002) ont montré que l'abaissement de la température induisait une transition structurale chez GmD-19 semblable à GmDHN1 (déhydrine) qui se traduit par une proportion plus importante en structure PII. Cette structure particulière favorise l'exposition des acides aminés polaires et donc l'interaction de la protéine avec l'eau (Soulages *et al.*, 2002).

LEA_4

Ce groupe est le plus important en nombre et rassemble des LEA assez variées, exprimées chez les plantes mais aussi chez les autres procaryotes, nématodes et rotifères (Battaglia *et al.*, 2008). Dure *et al.* (1989) ont rapidement mis en évidence des particularités dans les séquences peptidiques de ces LEA et notamment la répétition d'un motif de 11 acides aminés TAQAAKEKAXE susceptible de former une hélice α amphiphile. Mais ce motif, bien que présent dans les premières LEA identifiées, n'est pas retrouvé dans les 18 protéines d'*A. thaliana* classées dans ce groupe (Hundertmark & Hincha, 2008). En analysant 65 séquences protéiques appartenant à ce groupe, Battaglia *et al.* (2008) ont mis en évidence quatre autres motifs permettant de diviser ce grand groupe en deux sous groupes en fonction de la présence ou non de ces motifs et de leur position. Certains acides aminés (Lys, Glu, Asp, Ala) sont périodiquement répétés dans les séquences de ce groupe (Tunnacliffe & Wise, 2007). L'analyse structurale par dichroïsme circulaire a montré des résultats comparables à ceux obtenus pour les déhydrines : une structure native désordonnée capable d'adopter une structure majoritairement en hélice α au cours d'un séchage (Goyal *et al.*, 2003 ; Tolleter *et al.*, 2007). La vitesse de séchage et la présence d'autres molécules comme le saccharose peuvent influer sur la conformation obtenue après dessiccation (Wolkers *et al.*, 2001 ; Shih *et al.*, 2012). De plus ces changements de conformation sont parfaitement réversibles (Wolkers *et al.*, 2001 ; Tolleter *et al.*, 2007).

LEA_1

Les protéines appartenant au groupe LEA_1 présentent une zone assez conservées de 70 à 80 acides aminés en position N-terminale mais pendant longtemps aucun motif clair n'avait été identifié et la pertinence de maintenir ce groupe a même été discuté (Wise, 2003). Plus récemment, Battaglia *et al.* (2008) ont proposé

jusqu'à cinq motifs identifiés chez la majorité d'entre elles tandis que Bies-Etheve *et al.* (2008) mettait en évidence le motif SAKEKIS dans les cinq séquences protéiques de ce groupe chez *A. thaliana*. Ces LEA sont présentes à travers tout le règne végétal mais ont été assez peu étudiées d'un point de vue structural. Shih *et al.* (2004 ; 2010) ont montré que GmPM16, GmPM1 et GmPM28, bien que majoritairement désordonnée, possède des structures en hélice α à l'état hydratée. Ces auteurs ont également montré qu'un séchage rapide induisait une modification réversible de conformation de la protéine plus ou moins prononcée selon les protéines avec 90 % d'hélice α pour GmPM16 (Shih *et al.*, 2004) contre 49 % pour GmPM1 et 42 % pour GmPM28 (Shih *et al.*, 2010).

SMP

Ce groupe rassemble des LEA caractérisées par une forte proportion d'acides aminés hydrophobes et des propriétés physico-chimiques indiquant que ces protéines ne sont pas majoritairement désordonnées à l'état natif (Jaspard *et al.*, 2012). Certains auteurs ont suggéré que ces différences amenaient à exclure ces protéines du groupe des LEA (Tunnacliffe & Wise, 2007). Boudet *et al.*, (2006) ont étudié par FTIR la structure de PM25 chez *M. truncatula* et ont montré que 33 % de la protéine est structurée en hélice α, 18% en feuillet β et 49 % est donc désordonnée à l'état natif. La part de structure désordonnée serait même supérieure d'après des études ultérieures par dichroïsme circulaire (Boucher *et al.*, 2010), ce qui va à l'encontre des prédictions structurales *in silico* proposées pour les protéines de ce groupe (Tunnacliffe & Wise, 2007). Une fois encore la conformation évolue de façon réversible en condition de séchage, avec davantage de structure en hélice α et moins de régions désordonnées (Boudet *et al.*, 2006).

LEA_2 et LEA_3

Les groupes LEA_2 et LEA_3 regroupent également des LEA atypiques, plus hydrophobes (Hundertmark & Hincha, 2008). Au même titre que les protéines du groupe SMP, leur appartenance au groupe des LEA est discutée (Tunnacliffe & Wise, 2007). Contrairement à la majorité des LEA, la structure de LEA14, qui appartient à ce groupe, a été déterminée par NMR comme étant ordonnée en hélice α et feuillet β en solution (Singh *et al.*, 2005). Ces protéines fortement structurées pourraient avoir une autre fonction que les LEA décrites précédemment. En effet, elles sont plus rapidement fonctionnelles et pourraient donc être induites pour les stress modérés alors que les protéines désordonnées seraient plus spécifiques des stress sévères (Jaspard *et al.*, 2012).

Autres protéines associées aux LEA

Deux autres groupes de protéines ont été récemment mis en évidence présentant plusieurs caractéristiques communes avec les LEA. Le groupe PvLEA18, dont les premiers membres ont été identifiés chez le haricot, est constitué de petites protéines hydrophiles induites en condition de stress hydrique et en présence d'ABA (Colmenero-Flores *et al.*, 1997). Quatre motifs, dont deux très conservés, sont retrouvés parmi ces protéines (Battaglia *et al.*, 2008) et leur structure est prédite comme étant majoritairement désordonnée (Garay Arroyo *et al.*, 2000). Le deuxième groupe est formé par AtM17 et AtM10, deux protéines très

Tableau 1.3. Gènes codant pour des protéines LEA identifiés chez *M. truncatula*.

Les protéines sont classées d'après les domaines Pfam. Les domaines Interpro associés et la classification de Dure sont rappelés pour chacun de ces groupes. Pour chaque gène, l'identifiant MtGI11 (*Medicago Truncatula* Gene Index) ainsi que les identifiants IMGAG 3.5 (International Medicago Genome Annotation Group) et AGI (Arabidopsis Gene Index) sont précisés. Les données d'expression sont déduites de l'analyse *in silico* de Chatelain *et al.* (2012). Les gènes codant des protéines identifiées dans le protéome stable à la chaleur de graine mature (Chatelain *et al.*, 2012 ; Boudet *et al.*, 2006 pour BudCar5) sont indiqués en rouge ; || = non disponible.

Pfam name	Pfam domain	InterPro	Dure	Protein name	MtGI11	IMGAG 3.5	AGI	Expression (Chatelain *et al.*, 2012)
dehydrin	PF00257	IPR000167	D11	DHN3	TC175037	||	||	Graine / racine stress salin
				PM12	TC176019	Medtr3g117190	||	Graine
				DHN-cognate	TC174308	Medtr3g117290	AT1G76180	||
				BudCar5	TC183811	Medtr7g086340	||	Ubiquiste
				SLTI66	TC182121	||	AT1G54410	Ubiquiste
				RAB21	TC198429	Medtr8g106140	||	Graine / racine stress salin
LEA_5	PF00477	IPR000389	D19	EM6	TC185907	Medtr4g016960	AT2G40170	Graine
				EM1	AJ498523	||	||	Graine
LEA_4	PF02987	IPR004238	D7	SBP65	TC185888	Medtr4g079690	AT2G42560	Graine
				PM10	TC174929	||	AT2G18340	Graine / racine stress salin
				PM18	TC183861	||	AT2G36640	Graine
				MP2	TC187312	Medtr1g061730	AT2G36640	Graine / racine stress salin
				LEAM	TC179172	Medtr2g014040	AT5G44310	Graine
				CapLEA.I	TC175990	||	AT1G52690	Graine / racine stress salin
				CapLEA.II	TC194635	||	AT1G52690	Graine / racine stress salin
				CapLEA.III	TC178278	||	AT1G52690	Graine / racine stress salin
LEA_1	PF03760	IPR005513	D113	PM1	TC174780	Medtr7g093170	AT5G06760	Graine / racine stress salin / fleur / feuille
				D-113.I	TC183667	Medtr7g093140	AT5G06760	Ubiquiste
				D113.II	TC178511	Medtr7g093160	AT5G06760	Graine / racine stress salin
				PM29	TC183357	||	AT1G32560	||
SMP	PF04927	IPR007011	D34	D-34.I	TC176519	Medtr1g072090	AT3G22500	Graine
				PM25	TC174777	||	AT3G22490	Graine
				D-34.II	TC183570	||	AT3G22490	Graine
				D-34.III	TC178467	Medtr2g076230	AT3G22490	Graine
				ECP31	TC189170	Medtr3g034660	AT1G03120	Ubiquiste
				D-34.IV	TC196118	||	AT3G22490	||
LEA_2	PF03168	IPR004864	D95	LEA14-like c	BQ124186	||	AT1G01470	Graine
				LEA14	TC179700	Medtr7g079180	AT1G01470	Ubiquiste
				LEA-like a	TC179542	Medtr5g088850	AT2G44060	Ubiquiste
				LEA-like b	TC174481	||	AT2G44060	Ubiquiste
LEA_3	PF03242	IPR008082	D73	LEA5	TC186221	Medtr4g120040	||	Graine
				LEA-like protein	NP7252631	||	||	||
				ARG-2	TC176432	||	||	||
				ARG-2-like	TC187764	||	||	||

hydrophiles excepté à leur extrémité N-terminale, présentant un domaine constitué de six cystéines au sein d'un motif de 47 acides aminés (Raynal *et al.*, 1999).

Les protéines LEA identifiées chez *M. truncatula*

Trente quatre gènes codant pour des protéines LEA ont été répertoriés chez *M. truncatula* par Chatelain *et al.* (2012), ils sont présentés dans le tableau 1.3. Selon la nomenclature Pfam, ils sont classés en sept groupes différents car aucun membre des groupes AtM et PvLEA18 n'a été identifié. Sur les 28 pour lesquels une analyse de l'expression *in silico* a pu être effectuée, 12 sont spécifiques des graines (43 %), neuf sont exprimés dans les graines et dans les racines en réponse à un stress salin (32 %) et sept ont une expression ubiquiste (25 %) (Chatelain *et al.*, 2012). Les protéines codées par 18 de ces gènes, principalement classées dans les groupes LEA_5, LEA_4, LEA_1 et SMP, ont été identifiées dans le protéome stable à la chaleur de graines matures (Boudet *et al.*, 2006 ; Chatelain *et al.*, 2012). La moitié de ces gènes sont exprimés spécifiquement dans les graines.

c)-Localisation cellulaire

Les études de localisation subcellulaire des protéines LEA menées à ce jour ont montré que ces protéines ne sont pas membranaires mais plutôt réparties au sein des différents organites (chloroplaste, mitochondrie, réticulum endoplasmique, noyau, cytoplasme) (revue dans Tunnacliffe & Wise, 2007). Il n'est pas surprenant que ces protéines ne soient pas membranaires étant donné leur caractère fortement hydrophile. Une analyse exhaustive de la localisation des 51 LEA identifiées chez *A. thaliana* a été récemment entreprise par transformation de protoplastes avec les gènes codant ces LEA couplés à la GFP puis visualisation au microscope confocal. Elle confirme que ces protéines sont présentes dans tous les compartiments cellulaires même si la grande majorité d'entre elles (39) sont cytosoliques et/ou nucléaires (Candat, 2012). Excepté pour la grande famille LEA_4 dont les 18 membres ont été localisés dans 6 compartiments cellulaires différents, il y a une certaine homogénéité dans la localisation cellulaire des LEA au sein d'une même famille. Par exemple, les protéines des groupes LEA_1 ; LEA_2 ; LEA_5 ; SMP ; PvLEA18 et les déhydrines sont toutes cytosoliques et/ou nucléaires. Les deux membres de la famille AtM ont été localisées au niveau de la vacuole et la famille LEA_3 regroupe 4 protéines dont 3 sont mitochondriales (Candat, 2012).

d)- Fonctions proposées chez les plantes

Dès leur découverte, il a été établi que les LEA étaient principalement exprimées dans les graines en fin de maturation et étaient induites en réponse à de nombreux stress abiotiques (revu par Tunnacliffe & Wise 2007 ; Battaglia *et al.*, 2008), conditions où l'ABA joue un rôle de régulateur majeur. Cependant, deux études indépendantes publiées la même année ont montré que tous les gènes codant pour des LEA chez *A. thaliana* ne répondent pas à un traitement à l'ABA. En effet, 27 des 51 gènes identifiés par Hundertmark & Hincha (2008) sont sur-exprimés dans des plantules de sept jours en présence de 10 µM d'ABA et seulement 10 des 45 gènes identifiés par Bies-Ethève *et al.* (2008) le sont dans des jeunes plantules mises au contact de 50 µM d'ABA. D'après des données de PCR quantitative, 33 des 51 gènes sont exprimés dans les graines (pas nécessairement ceux induits par l'ABA) mais 22 sont fortement exprimés dans d'autres tissus. Plus de la moitié de ces gènes

exprimés en dehors de la graine (12) est fortement induite en réponse au froid, au stress salin ou à la sécheresse (Hundertmark & Hincha, 2008). Seulement 10 gènes sont à la fois exprimés dans les graines et dans les tissus végétatifs en condition de stress suggérant qu'un groupe de LEA est impliqué dans la réponse aux stress abiotiques tandis qu'un autre groupe jouerait un rôle spécifique dans les graines. Cette hypothèse a été renforcée par des analyses transcriptomique et protéomique chez *M. truncatula*. L'analyse des données de puce à ADN pour 28 des 34 gènes codant des protéines LEA chez cette espèce a révélé que 12 d'entre eux étaient spécifiquement exprimés dans les graines (Chatelain *et al.*, 2012). Ces LEA, principalement retrouvées au sein des familles LEA_4, LEA_5 et SMP chez *M. truncatula*, pourraient jouer un rôle spécifique dans la TD. En effet, l'accumulation de cinq d'entre elles a été spécifiquement corrélée à l'acquisition de la TD chez *M. truncatula* alors que d'autres LEA non spécifiques des graines (notamment des déhydrines) sont accumulées en réponse au stress osmotique (Boudet *et al.*, 2006). Pour identifier les protéines induites spécifiquement en réponse à la dessiccation, les auteurs ont tiré partie de la possibilité de ré-induire la TD dans des radicules de taille inférieure à 2,8 mm en présence de polyethylene glycol (PEG). Les protéines induites uniquement dans des radicules de 5 mm en présence de PEG (sensibles à la dessiccation mais stressées) par rapport aux radicules dans l'eau (sensibles à la dessiccation) répondent à un stress osmotique tandis que celles induites uniquement dans les radicules de 2,8 mm sont spécifiquement liées à la TD. D'autres études associent l'expression de protéines LEA à la dessiccation. Par exemple, Buitink *et al.* (2006) ont montré que presque tous les gènes codant pour des LEA sont sur-exprimés lors de la réinduction de la TD dans des radicules. De plus, l'expression de gènes codant pour des LEA a été détectée chez les plantes reviviscentes au cours de leur déshydratation (Ditzer *et al.*, 2001 ; Collett *et al.*, 2004 ; Illing *et al.*, 2005). Cependant les déhydrines ont été identifiées dans les graines récalcitrantes de plusieurs espèces (Finch-Savage *et al.*, 1994; Farrant *et al.*, 1996; Han *et al.*, 1997; Panza *et al.*, 2007; Sunderlikova *et al.*, 2009; Ismail *et al.*, 2010; Lee *et al.*, 2012). L'expression de gènes codant pour les déhydrines dans des graines récalcitrante de *Quercus robur* est induite par l'ABA (Finch-Savage *et al.*, 1994). Si la présence ou l'absence de déhydrines ne semble pas liée à la TD, l'expression des autres familles de LEA n'a pas été étudiée chez les graines récalcitrantes.

Ces études corrélatives ne suffisent pas à comprendre les fonctions biologiques et biochimiques de ces protéines. Par conséquent, plusieurs études fonctionnelles ont été menées par sur-expression d'un ou plusieurs gènes codant pour des LEA. Dans la majorité des cas, une augmentation de la tolérance au stress hydrique (Sivamani *et al.*, 2000 ; Chandra Babu *et al.*, 2004 ; Amara *et al.*, 2013) et salin (Xu *et al.*, 1996 ; Cheng *et al.*, 2002b ; Rohila *et al.*, 2002 ; Park *et al.*, 2005, 2011) ou au froid (Houde *et al.*, 2004 ; Xing *et al.*, 2011 ; Wu *et al.*, 2012) a été observée. Cependant, aucun effet n'est visible après sur-expression des déhydrines RAB18 et COR47 chez *A. thaliana* (Puhakainen *et al.*, 2004 concernant le stress hydrique). Le même résultat est obtenu en sur-exprimant simultanément chez le tabac trois protéines induites par la dessiccation chez *Craterostigma plantagineum* (Iturriaga *et al.*, 1992). Ces trois protéines sont en réalité deux déhydrines et une LEA appartenant à la famille LEA_4 (Tunnacliffe & Wise, 2007). L'analyse phénotypique de mutants déficients pour l'expression d'une LEA donnée est assez délicate du fait de la redondance fonctionnelle entre certaines de ces protéines qui pourraient compenser cette mutation (Tunnacliffe & Wise 2007). Néanmoins, le mutant d'Arabidopsis déficient en EM6 (LEA_5) présente une maturation altérée de la graine, avec notamment une déshydratation plus rapide en fin de développement (Manfre *et al.*, 2006 ; 2009). Olvera-Carillo *et al.*, (2010)

ont inactivé l'expression des trois protéines LEA appartenant au groupe LEA_1 chez Arabidopsis, et montré chez cette lignée mutante une tolérance réduite au stress osmotique dans les plantules, et au stress hydrique dans les plantes adultes par rapport aux plantes sauvages. De plus, en condition hydrique optimale, la production de graines est altérée chez le mutant avec une masse sèche inférieure à celle mesurée chez l'espèce sauvage. Par ailleurs, la longévité des graines et la germination en condition de stress salin sont réduites dans une lignée d'*A. thaliana* sous-exprimant fortement trois déhydrines (Hundertmark *et al.*, 2011). Ces graines restent néanmoins tolérantes à la dessiccation.

Si le rôle des LEA dans la réponse aux stress et à la dessiccation est de plus en plus avérée, l'expression importante de LEA dans les tissus végétatifs non stressées (Bies-Ethève *et al.*, 2008 ; Hundertmark & Hincha, 2008) permet de supposer que ces protéines jouent de nouvelles fonctions, mal connues à ce jour, dans le développement des plantes (Battaglia *et al.*, 2008 ; Olvera-Carillo *et al.*, 2010). Afin de mieux appréhender l'ensemble des fonctions jouées par ces protéines il est important de comprendre leur mécanisme d'action à l'échelle cellulaire.

e)-Mécanismes d'action

La modélisation des structures tri dimensionnelles a permis d'émettre des hypothèses pour comprendre par quels mécanismes les protéines LEA peuvent jouer un rôle dans la TD. Par exemple, le segment-K des déhydrines, qui se structure en hélice α, peut interagir avec les domaines hydrophobes d'autres molécules et éviter leur agrégation (Close, 1997). Depuis, de nombreuses études in vitro ont été menées pour mettre en évidence les propriétés des protéines LEA. Elles jouent un rôle de molécule chaperonne en préservant l'activité enzymatique de la lactate déshydrogenase, malate déshydrogenase, la catalase, la rhodanese, la fumarase ou la citrate synthase contre le gel (Wisniewski *et al.*, 1999 ; Hara *et al.*, 2001 ; Bravo *et al.*, 2003) ou la dessiccation (Grelet *et al.*, 2005 ; Reyes *et al.*, 2005 ; Goyal *et al.*, 2005 ; Boucher *et al.*, 2010). Si ces enzymes sont mises au contact d'albumine de sérum bovin (protéine contrôle sans activité chaperonne), elles sont dénaturées et s'agrègent dans ces conditions extrêmes. Goyal *et al.* (2005) ont montré que l'activité chaperonne des protéines EM de blé (LEA_5) et AaLEA1 (LEA_4) du nématode tolérant à la dessiccation *Aphelenchus avenae* est spécifique des stress limitant la disponibilité en eau (dessiccation et gel). Les LEA n'interagissent pas avec leurs protéines cibles de la même manière que les autres chaperonnes comme les Heat Shock Proteins (HSP). Les LEA ne coprécipitent pas avec leurs cibles contrairement aux HSP, suggérant qu'elles ne forment pas d'étroits complexes mais plutôt qu'elles agissent comme des boucliers moléculaires limitant les interactions entre molécules (Chakrabortee *et al.*, 2011).

Il a également été mis en évidence que COR15a (dehydrine) (Steponkus *et al.*, 1998) et LEAM (LEA_4) (Tolleter *et al.*, 2007, 2010) pouvaient s'associer respectivement aux membranes chloroplastiques et mitochondriales et participer à leur stabilisation en réponse au froid et à l'état sec. Le segment-K des déhydrines joue un rôle majeur dans l'association de ces LEA avec les phospholipides, comme le démontre Koag *et al.* (2009) en induisant des délétions dans ce motif chez DHN1 d'*A. thaliana*. Cette association est bien la conséquence du changement de conformation induit par la déshydratation car les interactions entre déhydrines et membranes ne sont pas visibles à l'état natif (Koag *et al.*, 2003). De plus, la déhydrine wCOR410 s'accumulent préférentiellement dans les tissus les plus exposées au stress (Danyluk *et al.*, 1998).

De part leur forte teneur en acides aminés chargés, les déhydrines (particulièrement celles du groupe acide) sont capables de fixer des ions Ca^{2+} (Heyen *et al.*, 2002) et des ions métalliques Fe^{3+} Cu^{2+} Ni^{2+} Co^{2+} Zn^{2+} (Hara *et al.*, 2005). Cette propriété a d'ailleurs été utilisée pour purifier les déhydrines par chromatographie (Svensson *et al.*, 2000). Une région riche en histidine pourrait être à l'origine de la liaison de ces protéines aux ions métalliques (Hara *et al.*, 2005). Afin de fixer des ions Ca^{2+} les déhydrines acides doivent être activées par phosphorylation (Alsheikh *et al.*, 2005). Ces ions peuvent être à l'origine de radicaux hydroxyles et leur fixation aurait donc une fonction antioxydante (Hara *et al.*, 2004 ; Liu *et al.*, 2011).

D'autres fonctions ont été proposées pour les déhydrines comme tampon d'hydratation. Cette propriété permet de conserver une activité enzymatique au niveau de poche d'eau dans les graines de bouleau (Rinne *et al.*, 1999). La même fonction a été proposée pour EM6 afin de réduire la vitesse de déshydratation dans les derniers stades de maturation des graines d'*A. thaliana* (Manfre *et al.*, 2006 ; Boucher *et al.*, 2010). De plus, les LEA jouent un rôle majeur dans la mise en place de l'état vitreux que nous détaillerons au point 2.3.6.

2.3.4-HSP

Les Heat Shock Proteins (HSP) ont été initialement mises en évidence en réponse aux stress thermiques de façon ubiquiste au sein des organismes vivants. Plusieurs groupes de ces molécules chaperonnes ont été décrites et classées selon leur poids moléculaire approximatif, leur localisation cellulaire et leur similarité de séquences (revue par Vierling, 1991). Parmi les petites HSP (sHSP) (15 à 42 kDa), 6 classes ont été définies : trois classes cytosoliques ou nucléaires (I, II et III), une classe chloroplastique, une classe mitochondriale et une classe endomembranaire. Ces différentes classes partagent au moins un domaine conservé de 90 acides aminés à l'extrémité carboxy-terminale : le domaine α-crystallin (ou « heat shock domain ») (revue par Waters *et al.*, 1996 ; Sun *et al.*, 2002). L'expression de sHSP cytosoliques au cours de la maturation a été démontrée dans les graines de pois (De Rocher & Vierling, 1994), de tournesol (Coca *et al.*, 1994), d'*A. thaliana* (Wehmeyer *et al.*, 1996, Sun *et al.*, 2001) et de *Jatropha curcas* (Omar *et al.*, 2011). Leur expression est régulée par le développement et non par l'environnement. De plus, ces protéines ont la particularité de persister à l'état sec. Des analyses fonctionnelles *in vitro* ont confirmé qu'elles jouent un rôle de chaperon moléculaire, principalement sous la forme de complexes oligomériques (Waters *et al.*, 1996, Sun *et al.*, 2001). Les sHSP cytosoliques de pois PsHSP18.1 et PsHSP17.7 stimulent le repliement de la citrate synthase et de la lactate déshydrogenase après dénaturation chimique de ces enzymes et empêchent l'agrégation de la malate déshydrogenase et de la glyceraldehyde 3-phosphate déshydrogenase (Lee *et al.*, 1995 ; Waters *et al.*, 1996). De même, 60 % de l'activité de la citrate synthase est récupérée après dénaturation en présence de AtHSP17.6 (Sun *et al.*, 2001). *In vivo*, la luciférase est protégée d'une inactivation thermique par les HSP17.7 et HSP17.3 de tomate dans des protoplastes d'*A. thaliana* (Löw *et al.*, 2000). D'autres études *in vivo* ont montré que leur sur-expression dans des plantes transgéniques induit une augmentation de la tolérance aux stress hydrique et salin chez *A. thaliana* (Sun *et al.*, 2001) et au stress hydrique chez le riz (Sato & Yokoya, 2008). Le mécanisme d'action des sHSP est différent de celui des protéines LEA : elles ne forment pas de boucliers moléculaires mais interagissent avec les protéines substrats pour éviter leur agrégation. Les sHSP font partie des protéines chaperonnes indépendante de l'ATP (au même titre que les protéines LEA), même si certaines études ont montré que la présence d'ATP pouvait stimuler (Muchowski &

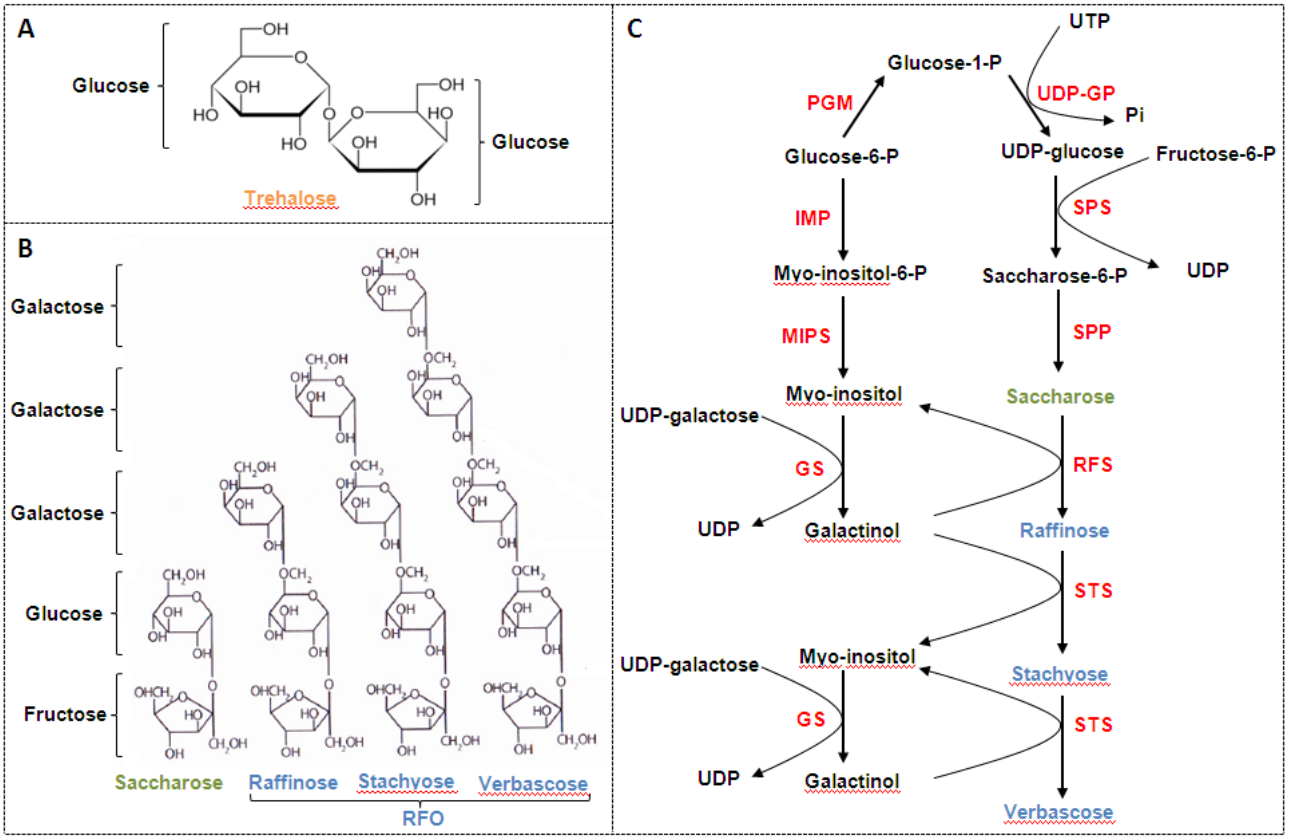


Figure 1.10. Structure des sucres non réducteurs impliqués dans la tolérance à la dessiccation et voie de biosynthèse du saccharose et des oligosaccharides de la famille du raffinose (RFO).
A) Le tréhalose est un disaccharide formé de deux glucoses. **B)** Les RFO sont formés par ajout d'un (raffinose), deux (stachyose) ou trois (verbascose) résidus galactose sur un saccharose, disaccharide constitué d'un glucose et d'un fructose (D'après Bewley *et al.*, 2013). **C)** Voie de biosynthèse du saccharose (vert) et des RFO (bleu). Les enzymes impliquées sont notées en rouge : PGM = phosphoglucomutase ; UDP-GP = UDP-glucose pyrophosphorylase ; SPS = saccharose phosphate synthase ; SPP = saccharose phosphate phosphorylase ; RFS = raffinose synthase ; STS = stachyose synthase ; IMP = myo-inositol monophosphatase ; MIPS = myo-inositol phosphate synthase ; GS = galactinol synthase (Adapté de Verdier *et al.*, 2013).

Clarck, 1998) ou au contraire inhiber (Smykal *et al.*, 2000) leur activité chaperonne. Cependant, elles fonctionnent en collaboration avec d'autres familles de chaperons moléculaires dépendant de l'ATP, comme les HSP70, DnaK ou DnaJ (Nakomoto & Vigh, 2007).

Il existe une corrélation entre le niveau de sHSP dans les graines et la TD ou la longévité des graines. Bettey & Finch Savage (1998) ont montré que les lots de graines de *Brassica oleracea* avec les teneurs en HSP17.6 les plus élevées étaient moins sensibles au vieillissement accéléré, suggérant un impact positif des sHSP sur la longévité. De même la sur-expression du gène de tournesol codant pour le Heat Shock Factor HSFA9 dans des plantes de tabac stimule l'expression des HSP (dont les sHSP de classe I et II) et augmente la longévité des graines transgéniques en condition de vieillissement accéléré (Prieto-Dapena *et al.*, 2006). La teneur en HSP17.4 est réduite ou nulle chez les mutants d'*A. thaliana* sensibles à la dessiccation : *abi3-6*, *lec1-2* (Wehmeyer & Vierling, 2000) et *tan* (Yamagishi *et al.*, 2005). Cependant, elle reste inchangée chez d'autres mutants tolérants à la dessiccation : *lec2-1* et *emb266* (Wehmeyer & Vierling, 2000). Ces auteurs ont également montré par expression GUS que le promoteur du gène codant pour HSP17.4 est activé par un stress thermique, y compris chez les mutants *lec1-2* et *abi3-6*. Ces travaux suggèrent donc que deux voies de régulation indépendantes contrôlent l'expression de ce gène au cours du développement et en réponse à un stress thermique (Wehmeyer & Vierling, 2000).

Les études sur les plantes reviviscentes confirment l'importance de ces protéines pour la TD. En effet, des homologues de HSP17.6 et HSP17.9 sont également exprimés de façon constitutive dans les tissus végétatifs tolérants à la dessiccation de *C. plantagineum*. Ces polypeptides n'ont pas été identifiés dans les cals sensibles à la dessiccation en absence de stress mais l'ajout exogène d'ABA dans ces cals induit l'acquisition de la TD et l'expression de ces sHSP (Alamillo *et al.*, 1995).

Cependant, une HSP homologue aux sHSP cytosolique de classe I identifiée dans les graines orthodoxes s'accumule fortement au cours du développement des graines récalcitrantes de châtaigner (Collada *et al.*, 1997). *In vitro*, cette sHSP est tout à fait fonctionnelle en augmentant la renaturation de la citrate synthase et en empêchant son inactivation thermique. Ces auteurs ont testé leur anticorps contre les extraits protéiques de graines orthodoxes ou récalcitrantes de 13 autres espèces d'arbres et un signal a pu être détecté chez toutes les espèces, suggérant que l'expression de ces protéines n'est pas suffisante pour induire l'acquisition de la TD.

2.3.5-Sucres solubles

L'accumulation de sucres non réducteurs tels que le saccharose, les oligosaccharides de la famille du raffinose (RFO) (raffinose, stachyose et verbascose) et le tréhalose ont été corrélés à l'acquisition de la TD chez les végétaux et animaux tolérants à la dessiccation (Koster & Leopold, 1988 ; Crowe *et al.*, 2002). La structure de ces sucres ainsi que la voie de biosynthèse des RFO sont représentées sur la figure 1.10. Koster & Leopold (1988) ont montré que les teneurs en saccharose et en RFO diminuent fortement dans des axes embryonnaires de pois, de soja et de maïs au cours de leur germination, parallèlement à la perte de la TD de ces tissus. De plus, la teneur en saccharose et en stachyose augmente fortement dans des axes immatures de soja séchées lentement (Blackman *et al.*, 1992), tout comme la teneur en raffinose et stachyose lors de l'acquisition de la TD de graines de haricot (Bailly *et al.*, 2001) ou de pois (Corbineau *et al.*, 2000). La même observation est faite avec la teneur en raffinose et l'expression d'un gène codant pour une galactinol synthase (enzyme de

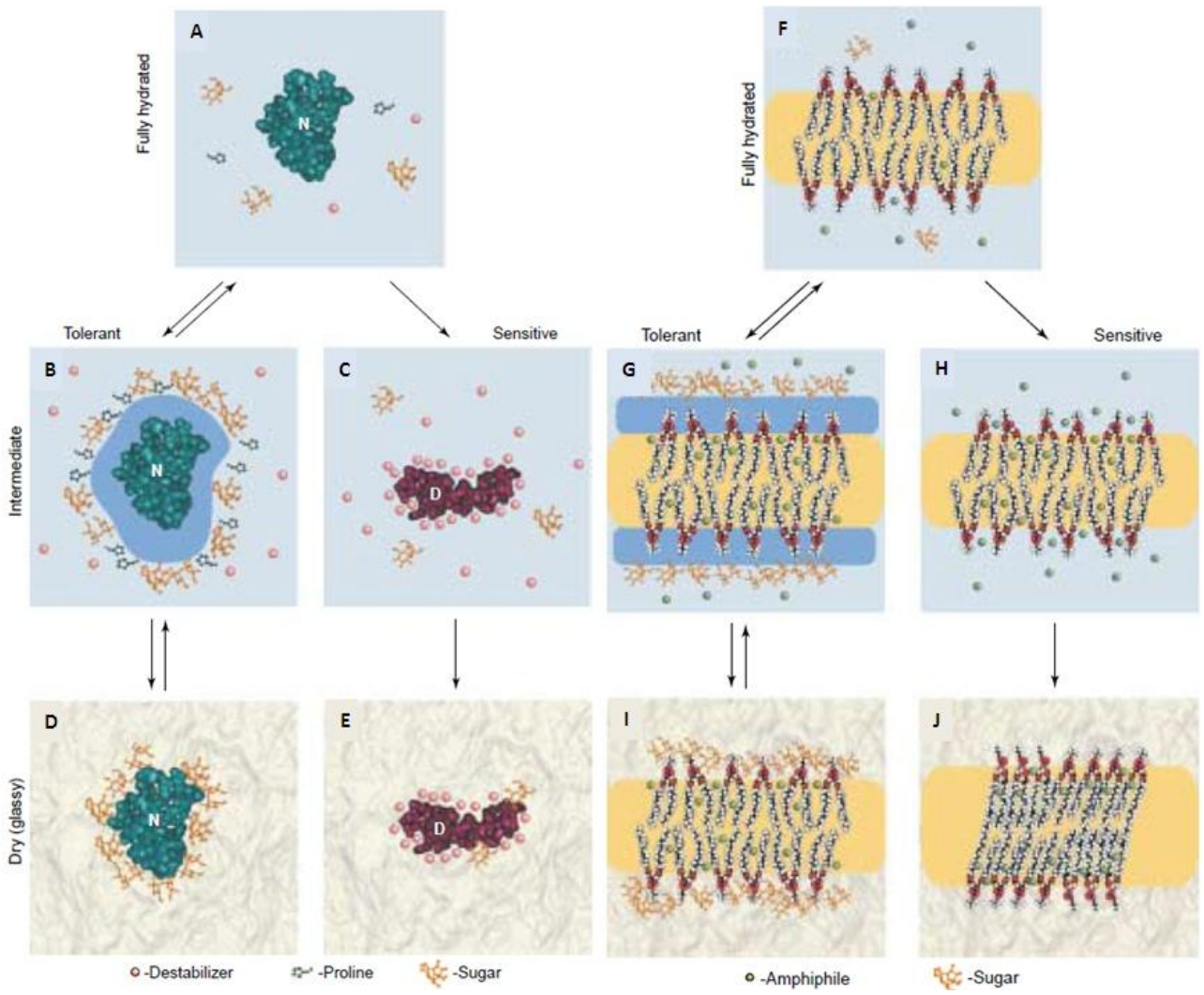


Figure 1.11. Mécanismes de protection des protéines (A-E) et membranes (F-J) par les sucres selon le niveau d'hydratation (D'après Hoekstra *et al.*, 2001).

A l'état hydraté **(A et F)**, les protéines (noté N) sont en solution dans leur conformation fonctionnelle et les membranes dans leur état cristallin liquide. Au cours de la dessiccation, la cellule connait un état d'hydratation intermédiaire et la concentration des solutés non compatibles (destabilizer) et des molécules amphiphiles augmente. Dans un tissu sensible, ces solutés se fixent aux protéines **(C)** ce qui aboutit à leur dénaturation (noté D). Les molécules amphiphiles s'intercalent entre les phospholipides **(H)** ce qui déstabilise les membranes. Dans un tissu tolérant, les solutés compatibles comme la proline et les sucres non réducteurs maintiennent une couche d'eau autour des protéines **(B)**, évitant ainsi leur dénaturation par hydratation préférentielle. Les sucres non réducteurs jouent le même rôle autour des membranes **(G)**. En dessous de 0,3 gH_2O/gMS, la viscosité augmente et les macromolécules sont progressivement immobilisées dans un état vitreux à partir de 0,1 gH_2O/gMS. Dans un tissu tolérant, les sucres non réducteurs remplacent les molécules d'eau en créant des liaisons hydrogènes avec la protéine ou les phospholipides **(D ; I)**. La structure native des protéines et l'espace entre les phospholipides sont conservés, les protéines ne sont donc pas dénaturées et les membranes restent dans leur état cristallin liquide. Dans un tissu sensible **(E ; J)**, les protéines sont irrémédiablement dénaturées et les phospholipides fusionnent, à l'origine de transition de phase. Une double flèche indique un processus réversible.

biosynthèse du raffinose) dans les graines de colza (Li *et al.*, 2011). Bailly *et al.*, (2001) montre également que les graines sèches sont riches en saccharose et pauvres en monosaccharides (glucose et fructose), profil retrouvé lorsque la TD est réinduite dans des radicules de concombre (Avelange-Macherel *et al.*, 2006). L'accumulation de stachyose est plus précoce et plus importante dans des graines de *Brassica campestris* soumises à un stress hydrique (irrigation arrêtée en début de maturation), ce qui renforce l'importance de ce sucre lorsque la disponibilité en eau diminue (Sinniah *et al.*, 1998).

La composition en sucres et notamment le ratio saccharose/oligosaccharide est également utilisé comme marqueur de longévité des graines. Ainsi d'après Horbowicz & Obendorf (1994) un ratio < 1 caractérise des graines à longévité importante (> 10 ans). Bien que la composition en sucre varie fortement d'une espèce à l'autre et entre les tissus d'une même graine, Steadman *et al.* (1996) ont noté que la teneur en raffinose et stachyose est globalement plus faible chez les espèces récalcitrantes. D'ailleurs, chez les espèces récalcitrantes *Inga vera* (Caccere *et al.*, 2013) et *Quercus rubra* L. (Sun *et al.*, 1994), aucun sucre de la famille des RFO n'est détecté au cours de la maturation.

Le tréhalose est un disaccharide formé de deux glucoses. Il est rarement détecté dans les graines, excepté dans une étude portant sur le développement des graines d'*A. thaliana* où une forte accumulation est mise en évidence lors de la transition entre accumulation des réserves et dessiccation (Fait *et al.*, 2006). Cependant, ce sucre est fortement accumulé chez les animaux anhydrobiotes tels que nématodes et *Artemia salina* ou les levures tolérantes à la dessiccation (Crowe *et al.*, 2002). Le tréhalose est très efficace pour former des états vitreux (Crowe *et al.*, 1996), protéger les protéines solubles (Crowe *et al.*, 1987 ; Kaushik & Bhat, 2003) et les membranes au cours du séchage (Crowe *et al.*, 1996 ; Ricker *et al.*, 2003). Une étude a même montré que des cellules humaines sur-exprimant des gènes codant pour des enzymes de biosynthèse de ce sucre deviennes tolérantes à la dessiccation pendant 5 jours (Guo *et al.*, 2000). Le saccharose est également capable de protéger les protéines (Crowe *et al.*, 1987 ; Leslie *et al.*, 1995) et les membranes (Bryant *et al.*, 2001 ; Halperin & Koster, 2006), en limitant leur fusion et en abaissant leur température de transition de phase, mais avec une moindre efficacité que le tréhalose (Crowe *et al.*, 1996 ; Sun & Davidson, 1998).

Cependant, toutes les études ne montrent pas de lien entre accumulation de ces sucres non réducteurs et la survie à l'état sec. Par exemple, aucune corrélation n'est observée entre l'accumulation de raffinose et l'acquisition de la TD chez le blé (Black *et al.*, 1999) ou entre la teneur en sucres solubles et la longévité en se basant sur une analyse QTL avec plusieurs accessions d'*A. thaliana* (Bentsink *et al.*, 2000). La corrélation entre la teneur en RFO et la longévité n'est pas retrouvée en étudiant l'effet du priming (qui diminue la longévité) et de traitement post-priming (qui restaure la longévité) sur les graines de plusieurs espèces différentes (Gurusinghe & Bradford, 2001). De plus, la comparaison entre une espèce sensible et une espèce tolérante à la dessiccation du genre *Selaginella* montre que le tréhalose est plus abondant chez l'espèce sensible (Yobi *et al.*, 2012). Ce sucre est même absent chez des rotifères pourtant tolérant à la dessiccation (Lapinski & Tunnacliffe, 2003).

Deux mécanismes ont été proposés pour expliquer le mécanisme de protection des membranes et des protéines par les sucres : l'exclusion préférentielle (ou hydratation préférentielle) et le remplacement de l'eau (Hoekstra *et al.*, 2001) (Figure 1.11). Selon le premier mécanisme, les sucres maintiennent une couche d'eau autour des macromolécules, évitant ainsi leur dénaturation. Cette hypothèse est privilégiée lorsque la teneur en

eau n'est pas encore trop basse dans la cellule (Hoekstra *et al.*, 2001). Dans le deuxième mécanisme, les sucres se substituent aux molécules d'eau pour interagir avec les têtes polaires des phospholipides et maintenir suffisamment d'espace entre eux (Hoekstra *et al.*, 1997 ; Golovina *et al.*, 2009).

D'autres fonctions ont été proposées pour les sucres non réducteurs. Le raffinose pourrait jouer un rôle antioxydant en réagissant avec les radicaux hydroxyles (Nishizawa *et al.*, 2008), tout comme le saccharose (Ende & Valluru, 2009). Les RFO sont dégradés très rapidement durant l'imbibition des graines (Gorecki *et al.*, 1997 ; Rosnoblet *et al.*, 2007), leur mobilisation pourrait servir de substrat respiratoire durant la germination (Blöchl *et al.*, 2007) et la croissance hétérotrophe de la plantule (Vandecasteele *et al.*, 2011). Enfin, les RFO pourraient être impliqués dans la réponse aux stress abiotiques dans les tissus végétatifs. Une accumulation de raffinose et de galactinol (précurseur du raffinose) a été mise en évidence dans des plantes d'*A. thaliana* soumises à différents stress abiotiques (Taji *et al.*, 2002) et dans des plantules de *Vicia villosa* soumises à un stress osmotique (Lahuta & Gorecki, 2011).

2.3.6-Etat vitreux

Par opposition à une structure cristalline, l'état vitreux peut être assimilé à un solide, thermodynamiquement instable, sans structure régulière et présentant une viscosité élevée (Walters, 1998 ; Buitink & Leprince, 2004). Dans cette matrice, la mobilité moléculaire est fortement réduite ce qui pourrait expliquer la corrélation entre la formation de l'état vitreux et l'allongement de la durée de stockage des graines (Sun, 1997, Buitink *et al.*, 1998). En effet, dans ces conditions, les structures cellulaires ne sont pas altérées au cours de la déshydratation et les radicaux libres, nocifs pour les membranes, sont fortement ralentis (Sun & Leopold, 1995). L'état vitreux se forme lors des derniers stades de déshydratation de tous les tissus tolérants à la dessiccation : pollen (Buitink *et al.*, 1996), graine orthodoxe et plante reviviscente (Buitink & Leprince, 2008). Cependant, la teneur en eau à partir de laquelle il se forme dépend de la température. A 20°C, la viscosité augmente de façon progressive dans les graines de pois à partir d'une teneur en eau inférieure à 0,3 g H_2O/g MS et le cytoplasme devient vitreux à partir de 0,1 g H_2O/g MS (Buitink & Leprince, 2004) alors qu'il est vitreux dans des embryons de maïs dès 0,15 g H_2O/g MS en condition de stockage standard (-5°C) (Bruni & Leopold, 1992). La température de transition à laquelle une molécule passe de l'état vitreux à l'état liquide (Tg) est plus élevée pour les substances de poids moléculaire important (Buitink *et al.*, 2000). Par conséquent, la composition biochimique influe également sur les conditions de formation de l'état vitreux. Par exemple, un état vitreux constitué de saccharose (faible poids moléculaire) se formera à des teneurs en eau plus basse qu'un état vitreux constitué de raffinose ou stachyose (poids moléculaire plus élevé). Les premières études sur la formation de l'état vitreux dans les graines en lien avec la TD ont proposé que les sucres étaient responsables de leur formation (Koster & Leopold, 1988 ; Koster, 1991) mais d'autres molécules pourraient également jouer un rôle (Sun & Leopold, 1993 ; Walters, 1998), notamment les protéines LEA (Berjak, 2006 ; Buitink & Leprince, 2008). En effet, une quantité importante de sucres reste fortement liée aux protéines LEA extraites dans une fraction stable à la chaleur (Walters *et al.*, 1997). Wolkers *et al.* (2001) ont également montré que l'association entre une LEA pollinique et du saccharose produisait un état vitreux plus stable qu'en présence de saccharose seul. L'interaction entre protéines LEA et les sucres non réducteurs a depuis été démontrée chez le soja (Shih *et al.*, 2010 et 2012). Peu d'études ont été menées pour comprendre la synergie entre les sucres et

les LEA dans la formation de l'état vitreux. Il a été proposé que les hélices α formées à l'état sec chez AavLEA1 pourraient constituer un réseau de filaments renforçant la matrice de sucres formant l'état vitreux (Goyal *et al.*, 2003). Les travaux de Bryant *et al.* (2001) montrent que les petits solutés (comme le saccharose) s'immiscent à proximité des membranes pour limiter les interactions physiques entre membranes, propriété non observée pour les gros solutés. Ces travaux amènent Berjak (2006) à proposer une répartition hétérogène des LEA (gros solutés) et des sucres au sein de l'état vitreux, ces derniers occuperaient les espaces inter membranaires étroits et joueraient un rôle complémentaire des LEA dans l'état vitreux. Il est probable que les autres composants cellulaires (sels, ions, acides aminés et composés organiques) participent également à la formation de l'état vitreux (Buitink & Leprince, 2004)

Si la formation d'état vitreux est indispensable à la survie à l'état sec, certaines observations montrent que son absence ne suffit pas à expliquer la sensibilité à la dessiccation. En effet, un état vitreux a été observé dans le pollen de maïs, tissus sensible à la dessiccation (Buitink *et al.*, 1996). De plus, les graines récalcitrantes meurent généralement à des teneurs en eau bien supérieures à celles nécessaires la formation de l'état vitreux (Buitink & Leprince, 2008 ; Berjak & Pammenter, 2008).

2.3.7-ROS et antioxydants

a) Définition des ROS et impact dans les graines

Les espèces réactives de l'oxygène (ROS) regroupent les radicaux dérivées du dioxygène présentant des électrons non appareillés comme l'oxygène singlet ($^1O_2$), l'anion superoxyde ($O_2^{\cdot -}$), le radical hydroxyle ($^{\cdot}OH$) et l'oxyde nitrique ($NO^{\cdot}$) auxquels il faut ajouter le peroxyde d'hydrogène (H_2O_2) qui n'est pas un radical (Kranner & Birtic, 2005). Ces espèces sont instables (à l'exception du H_2O_2) et hautement réactives, elles ont de nombreux effets délétères sur les structures cellulaires : peroxydation des membranes, oxydation des protéines et de l'ADN. Les oxydases impliquées dans la chaine respiratoire sont une source importante de production de $O_2^{\cdot -}$ par transfert d'électrons libres sur le dioxygène ($^3O_2$) (Kranner & Birtic, 2005). Cet anion $O_2^{\cdot -}$ est peu impliqué directement dans la dégradation des structures mais réagit avec lui-même pour former le H_2O_2 (Kranner & Birtic, 2005). Le radical hydroxyle ($^{\cdot}OH$), le plus réactif des ROS, se forme à partir des deux espèces précédentes (H_2O_2 et $O_2^{\cdot -}$) d'après la réaction de Haber-Weiss mais est aussi produit dans la paroi par des peroxydases. L'oxygène singlet ($^1O_2$) est principalement produit au niveau des photosystèmes, suite à un transfert d'énergie lumineuse trop important sur le $^3O_2$. Il peut également se former à partir du $^{\cdot}OH$ et du $O_2^{\cdot -}$ dans les tissus non photosynthétiques et il a été montré qu'il initie la peroxydation des lipides. Les peroxysomes, les chloroplastes et les mitochondries sont les sources les plus importantes de ROS (Foyer & Noctor, 2005).

La production de ROS fait partie intégrante du métabolisme et joue même un rôle majeur de signalisation et de régulateur du développement : ces molécules sont impliquées dans la réponse aux stress biotiques et abiotiques, la mort cellulaire programmée, la mitose et les tropismes (revu par Mittler *et al.*, 2004, 2011 ; Foyer & Noctor, 2005 ; Suzuki & Mittler 2006, Miller *et al.*, 2010). Un système antioxydant basal fonctionne donc en permanence au niveau cellulaire pour maintenir ce rôle de signalisation tout en évitant les effets nocifs de ces molécules. Cependant, la dessiccation peut entraîner la production de ROS (cf 2.2) et des adaptations du système antioxydant sont nécessaires pour limiter ou éviter les dégâts structuraux.

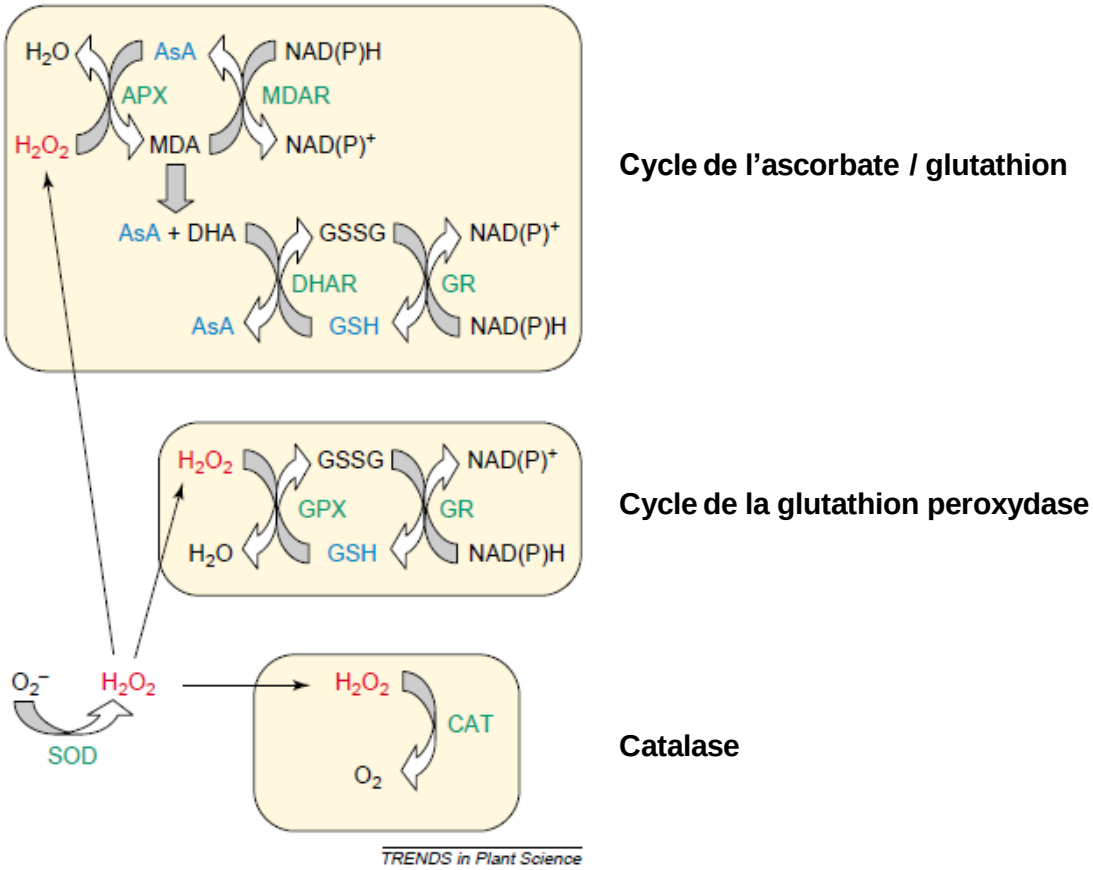


Figure 1.12. Différentes voies d'élimination des espèces réactives de l'oxygène (ROS) par action enzymatique (D'après Mittler, 2002).

La superoxyde dismutase (SOD) convertit l'ion superoxyde ($O_2^{\cdot -}$) en peroxyde d'hydrogène (H_2O_2). La catalase réduit directement le H_2O_2 en dioxygène dans les peroxysomes. L'ascorbate peroxydase (APX) et la glutathion peroxydase (GPX) réduisent le H_2O_2 en oxydant respectivement les formes réduites de deux ascorbates (AsA) ou d'un glutathion (GSH). Dans le cycle de la glutathion peroxydase, la forme oxydée du glutathion (glutathion disulphide ou GSSG) est régénéré sous sa forme réduite (GSH) par la glutathion réductase (GR), réaction consommant du pouvoir réducteur sous la forme d'un NAD(P)H. Dans le cycle de l'ascorbate/glutathion, l'ascorbate sous sa forme réduite (AsA) peut être régénéré directement par réduction du monodehydroascorbate (MDA) via la monodehydroascorbate réductase (MDAR), réaction consommant un NAD(P)H. Si le MDA n'est pas rapidement réduit il se dissocie en ascorbate (AsA) et dehydroascorbate (DHA). Le DHA est réduit en ascorbate (AsA) par l'oxydation d'un glutathion (GSH) via la dehydroascorbate réductase (DHAR). Le GSH est régénérée par la glutathion réductase (GR) comme indiqué pour le cycle de la glutathion peroxydase.

b) Antioxydants enzymatiques

Huit enzymes majeures participent directement ou indirectement à la dégradation des ROS en molécules non toxiques parmi lesquelles la superoxyde dismutase (SOD), la catalase (CAT), l'ascorbate peroxydase (APX), (mono)déhydroascorbate réductase (MDAR et DHAR) et les glutathion -réductase (GR) -peroxydase (GPX) -S-transférase (GST) (Noctor & Foyer, 1998). La SOD transforme $O_2^{\cdot -}$ en H_2O_2 qui est lui-même converti en eau et dioxygène par la CAT. Les peroxydases sont également capables de dégrader le H_2O_2 mais un réducteur est nécessaire (chez les plantes principalement l'ascorbate et le glutathion sous leur forme réduite) (Noctor & Foyer, 1998). Les enzymes MDHAR, DHAR et GR agissent donc indirectement dans la dégradation des ROS en permettant la réduction de l'ascorbate et du glutathion et donc la régénération des substrats pour les peroxydases APX et GPX, formant ainsi le cycle de l'ascorbate/glutathion (Bailly, 2004). L'ensemble de ces réactions est visible dans la figure 1.12. La CAT à une plus faible affinité pour le H_2O_2 que les peroxydases et est donc moins efficace à faible concentration de H_2O_2. Cependant son activité catalytique élevée la rend très efficace pour une détoxification rapide lorsque la concentration en H_2O_2 est élevée (Kranner & Birtic, 2005). La CAT est spécifiquement localisée dans les peroxysomes, contrairement aux autres enzymes retrouvées dans tous les compartiments cellulaires (Mittler *et al.*, 2002 ; Mhamdi *et al.*, 2012).

L'eau et l'énergie sont nécessaires aux activités enzymatiques, les mécanismes enzymatiques antioxydants présentés ci-dessus ne sont donc vraisemblablement pas opérationnels à l'état sec. Cependant, suffisamment d'eau est disponible au stade intermédiaire d'hydratation (stade III 0,45 et 0,25 g H_2O/g MS) pour permettre l'activation de certaines enzymes (Bewley *et al.*, 2013). Les enzymes à activité antioxydante peuvent donc jouer un rôle au cours de la maturation et de l'imbibition des graines. D'ailleurs, elles changent radicalement lors de l'acquisition de la TD au cours de la maturation de la graine de haricot (Bailly *et al.*, 2001). Les activités CAT et GR augmentent fortement tandis que les activités SOD et APX diminuent en lien avec l'acquisition de la TD dans les graines. Plusieurs études montrent l'importance de la CAT et de la GR en lien avec la TD car leur activité enzymatique augmente avant le séchage des graines (De Gara *et al.*, 2003 ; Bailly *et al.*, 2004). De plus, la CAT est impliquée dans la protection durant le vieillissement et dans les processus de réparation au cours du priming chez le tournesol (Kibinza *et al.*, 2011). En revanche, l'ascorbate est peu présent et l'activité APX est faible dans les graines sèches de trois espèces orthodoxes mais élevée dans les graines récalcitrantes de quatre espèces à l'abscission (Tommasi *et al.*, 1999). Les auteurs suggèrent un lien avec l'activité métabolique qui reste élevée dans les graines récalcitrantes, hypothèse confortée par la reprise de synthèse de l'ascorbate et de l'activité APX 10h après imbibition dans des embryons de blé (De Gara *et al.*, 1997). L'ensemble des données suggère que le système de détoxification par l'ascorbate n'est pas lié à l'acquisition de la TD mais que cette enzyme pourrait jouer un rôle de co-substrat pour les dioxygenases, enzymes clés de synthèse des phytohormones (De Tullio & Arrigoni, 2003 ; Bailly, 2004).

Dans des tissus sensibles à la dessiccation les activités antioxydantes sont insuffisantes pour contrôler la production de ROS. Farrant *et al.* (2004) ont montré, au cours du séchage de coléoptiles de blé sensibles et tolérants à la dessiccation, une accumulation plus forte de H_2O_2 et de malondialdehyde (indicateur d'oxydation des lipides) dans les tissus sensibles. Les activités de la SOD, de la GR et de la catalase sont altérées lors de la réhydratation de ces tissus sensibles après dessiccation (Farrant *et al.*, 2004) et après séchage de grains de

maïs (Leprince *et al.*, 1990). La perte de viabilité des graines récalcitrantes est corrélée à une augmentation de la peroxydation des lipides et une diminution des activités de la SOD et des peroxydases (Chaitanya & Naithani, 1994 ; Li & Sun, 1999 ; Parkhey *et al.*, 2012).

c) Antioxydants non enzymatiques

Un grand nombre de molécules ont été décrites comme ayant des propriétés antioxydantes directes sans l'intervention d'activité enzymatique. A l'état sec, les réactions enzymatiques sont très limitées, ces antioxydants non enzymatiques jouent donc certainement un rôle très important (Bailly, 2004). Parmi les exemples les plus documentés on retrouve l'ascorbate, le glutathion, les β-carotènes ou les tocophérols (vitamine E) (Noctor & Foyer, 1998 ; Munné-Bosch & Falk, 2004 ; Noctor *et al.*, 2012). Les deux premiers sont solubles dans l'eau tandis que les deux derniers sont liposolubles et donc spécifiques des membranes (Kranner & Birtic, 2005). Le glutathion et l'ascorbate ne sont pas seulement des co-facteurs enzymatiques, ils peuvent également jouer directement un rôle d'antioxydant (Noctor & Foyer, 1998). L'étude de mutants d'*A. thaliana* déficients dans la synthèse de tocophérols a confirmé le rôle de ces molécules dans la protection des lipides contre l'oxydation au cours du stockage à l'état sec, et de l'imbibition des graines (Sattler *et al.*, 2004). Par la même approche, Cairns *et al.* (2006) ont montré l'importance du glutathion dans la maturation de la graine d'*A. thaliana*, sans pour autant démontrer une implication précise dans l'acquisition de la TD. En effet, les graines du mutant homozygote *gsh1* ne sont pas viables mais peuvent être sauvées si mises à germer en présence de glutathion 10 jours après fécondation. Deux enzymes dépendante de l'ATP sont impliquées dans la biosynthèse du glutathion : la glutamate cystéine ligase (ou γ-glutamylcysteine synthetase) et la glutathion synthétase (Noctor & Foyer, 1998). La forme réduite du glutathion (GSH) joue le rôle d'antioxydant et se retrouve sous forme oxydée (GSSG) après réaction avec les ROS. Cette forme oxydée est recyclée en GSH par l'intervention de la glutathion réductase (GR) et consommation de pouvoir réducteur (NADPH) (Noctor *et al.*, 2012) (Figure 1.12). La régénération du GSH joue un rôle majeur dans la longévité des graines orthodoxes et des plantes tolérantes à la dessiccation, si bien que le potentiel redox du glutathion ($E_{GSSSG/2GSH}$) est utilisé comme marqueur de viabilité. Au dessus d'une certaine proportion de GSSG dans le pool total de glutathion, $E_{GSSSG/2GSH}$ augmente et initie un signal de mort cellulaire programmée (Kranner & Birtic, 2005).

D'autres molécules ont été proposées comme antioxydant d'après des études de réponse au stress dans les tissus végétatif ou d'après des études *in vitro* : flavonoïdes, polyols, proline, polyamines (revu par Smirnoff, 1993). Dans les graines, un certain type de protéines appelées peroxyredoxines a été identifié comme antioxydant potentiellement impliqué dans la TD. En effet, les 1-cys peroxyredoxines sont accumulées en quantité importante dans le noyau juste avant la dessiccation des grains d'orge (Stacy *et al.*, 1999). De plus, le profil d'expression du gène associé, régulée par l'ABA, est semblable à celui des gènes codant les LEA (diminution lors de l'imbibition en lien avec la perte de la TD) (Aalen, 1999). D'après Illing *et al.* (2005), seulement 4 des 71 gènes codant pour des gènes impliqués dans la réponse antioxydante sont spécifiques des graines (parmi lesquels celui codant pour la 1-cys peroxyredoxine), ce qui amènent ces auteurs à penser que les mécanismes antioxydants mis en place en réponse à la TD ne sont pas spécifiques chez les espèces tolérantes mais que, contrairement aux espèces sensibles, ils sont protégés et restent actifs aux faibles teneurs en eau.

2.3.8- Oleosines

Des réserves de lipides sont présentes au niveau cellulaire, elles sont entourées par des protéines amphiphiles nommées oléosines. La partie hydrophobe interagit avec les lipides et la partie hydrophile avec la phase aqueuse du cytosol permettant ainsi d'isoler des corps lipidiques au sein du cytoplasme (Huang, 1992). Les gènes codant pour ces protéines sont exprimés spécifiquement pendant le développement des graines sous le contrôle du facteur de transcription LEC2 (Che *et al.*, 2009). Plusieurs gènes codant pour des oléosines sont également des cibles directes d'ABI3 (Mönke *et al.*, 2012 ; Verdier *et al.*, 2013). Les oléosines jouent un rôle majeur dans la tolérance au froid des graines d'*A. thaliana* en évitant la fusion des corps lipidiques (Shimada *et al.*, 2008). Elles pourraient avoir la même fonction dans la tolérance à la dessiccation. En effet, ces protéines sont absentes des graines récalcitrantes de cacaoyer et de *Quercus rubra*, pourtant riches en lipides de réserve (Leprince *et al.*, 1997). De plus, une perte de l'intégrité cellulaire liée à la fusion des corps lipidiques a été observée lors de la réhydratation de graines artificiellement séchées de cacaoyer (Leprince *et al.*, 1997).

2.4-Régulation de l'acquisition de la tolérance à la dessiccation

L'ABA étant connu comme le régulateur majeur de la maturation des graines et de la réponse au stress (Ingram & Bartels, 1996) il n'est pas surprenant que la plupart des mécanismes impliqués dans l'acquisition de la TD des graines cités précédemment soient positivement régulés par l'ABA. L'expression des LEA (Galau *et al.*, 1986, Hundertmark & Hincha, 2008), des sHSP (Wehmeyer *et al.*, 1996, Wehmeyer & Vierling, 2000), des antioxydants (Larkindale & Knight, 2002 ; Larkindale & Huang, 2004) est en effet activé par l'ABA. Parmi les mutants déficients pour des facteurs de transcription majeurs de la maturation des graines (*lec1, lec2, fus3, abi3, abi5, abi4)* d'*A. thaliana*, les mutants *abi3* et *lec1* sont sensibles à la dessiccation (To *et al.*, 2006) et deviennent donc des outils intéressants pour l'étude de l'expression des mécanismes liés à l'acquisition de la TD.

L'analyse de ces mutants montre qu'ABI3 est l'activateur le plus important des protéines LEA exprimées dans les graines. En effet, 19 des gènes codant pour des LEA majoritairement spécifiques des graines chez le mutant *abi3* sont sous-exprimés (Bies-Ethève *et al.*, 2008). De plus, de nombreuses LEA ont été identifiées dans les cibles directes d'ABI3 chez *A. thaliana* (Mönke *et al.*, 2012) et *M. truncatula* (Verdier *et al.*, 2013) ainsi que dans les cibles d'ABI4 et d'ABI5 (Reeves *et al.*, 2011). Les séquences promotrices des gènes codant pour les LEA sont enrichies en élément RY (motif reconnu par ABI3) et en motif ABRE (reconnu par les facteurs bZIP comme ABI5) (Hundertmark & Hincha, 2008), deux boites majeures dans la signalisation dépendante de l'ABA. Elles sont également enrichies en motif DRE/LTRE reconnu par le facteur DREB1 en réponse aux stress abiotiques et dont la signalisation est indépendante de l'ABA (Hundertmark & Hincha, 2008). Cependant, la régulation des LEA est assez complexe car ABI3 réprime également 10 gènes codant pour des LEA principalement exprimés dans les tissus végétatifs. Certains gènes, comme *EM6* (LEA_5), sont sous-exprimés chez les mutants *abi5*, *abi4* et *fus3* pourtant tolérant à la dessiccation (Delseny *et al.*, 2001 ; Bies-Ethève *et al.*, 2008). *EM6* est également sous-exprimé dans le mutant *lec1* (Delseny *et al.*, 2001) bien qu'aucune LEA n'ai été identifiée parmi les cibles directes du facteur de transcription LEC1 (Junker *et al.*, 2012). L'action du facteur

LEC1 sur la TD se ferait donc plus en amont en régulant l'expression des autres facteurs tels que ABI3 (To *et al.*, 2006).

ABI3 est également un régulateur majeur des sHSP, comme le montre la faible teneur de ces protéines dans les mutants *abi3* d'*A. thaliana* (Wehmeyer *et al.*, 1996). L'analyse d'autres mutants suggère que LEC1 et FUS3 participent également à la régulation mais pas LEC2, mutant dont les graines restent tolérantes à la dessiccation (Wehmeyer & Vierling, 2000). L'expression des sHSP est régulée par des facteurs de transcription spécifiques : les Heat Shock Factor (HSF) (Scharf *et al.*, 2011 pour une revue récente). Parmi les 21 gènes codant des HSF chez *A. thaliana*, *HSFA9* est spécifiquement exprimé dans les graines en fin de développement sous le contrôle d'ABI3 (Kotak *et al.*, 2007). La sur-expression du facteur HSFA9 de tournesol dans le tabac augmente la longévité des graines (Prieto-Dapena *et al.*, 2006 ; Almoguera *et al.*, 2009). Cependant, l'inactivation de ce gène dans les graines de tabac n'a pas d'impact sur l'acquisition de la TD au cours du développement, même s'il conduit bien à une baisse de la teneur en sHSP et de la longévité des graines (Tejedor-Cano *et al.*, 2010).

Concernant les autres mécanismes présentés comme liés à la TD, des gènes codant pour les oléosines, 1-cys peroxyredoxine et superoxyde dismutase ont été identifiés dans le régulon d'ABI3 chez *A. thaliana* (Mönke *et al.*, 2012). Une raffinose synthase et une glutathion-S-transférase ont été également identifiées parmi les cibles d'ABI3 chez *M. truncatula* (Verdier *et al.*, 2013). Chez *A. thaliana*, Le facteur HSFA1-b est responsable de l'accumulation de raffinose dans les feuilles en réponse à un stress thermique via l'activation d'une galactinol synthase (Panikulangara *et al.*, 2004). D'après les données d'expression disponibles sur le site BAR-eFPbrowser, le gène codant HSFA1-b (At5g16820) est également fortement exprimé dans la graine sèche, où il pourrait jouer le même rôle.

3-Modèle d'étude

3.1- Intérêt de la comparaison orthodoxe-récalcitrante

La maturation est une étape complexe du développement de la graine au cours de laquelle la TD est acquise au milieu d'autres processus (cf 1.1.1). De plus, les régulateurs majeurs des graines orthodoxes exercent des contrôles partiellement redondants sur l'ensemble de ces processus (cf 1.1.2). Il est donc très difficile d'isoler spécifiquement les mécanismes moléculaires liés à l'acquisition de la TD en analysant simplement la maturation des graines orthodoxes et des approches comparatives sont nécessaires. La réinduction de la TD dans des radicules juste après germination a été utilisée précédemment chez *A. thaliana* (Maia *et al.*, 2011) et *M. truncatula* (Buitink *et al.*, 2003 ; 2006 ; Boudet *et al.*, 2006) pour dissocier l'acquisition de la TD du reste de la maturation. Dans cette étude, nous avons initié une autre approche en comparant sur le plan moléculaire le développement d'une graine orthodoxe avec celui d'une graine récalcitrante. L'objectif est de mettre en évidence des mécanismes présents chez l'espèce tolérante à la dessiccation et absents chez l'espèce sensible, ou inversement. Cette approche, suggérée et initiée il y a déjà plus de 15 ans (Farrant *et al.*, 1996, 1997 ; Kermode, 1997 ; Greggains *et al.*, 2000), se heurtait à la difficulté d'obtenir des données moléculaires sur des espèces non modèles. Le bond technologique du séquençage

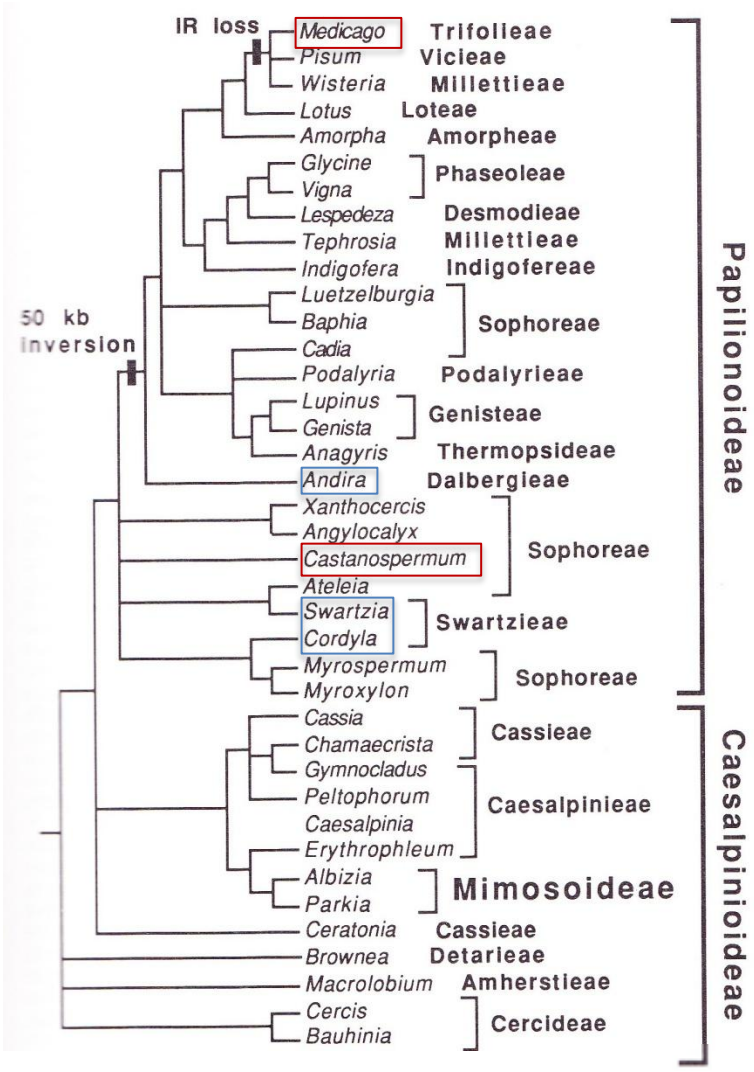


Figure 1.13. Classification phylogénétique de la famille des Fabacées (Doyle, 1995).
Synthèse des résultats d'une analyse par maximum de parcimonie basée sur les séquences complètes ou partielles du gène *rbcL* chez 54 espèces. Pour plus de clarté, seuls les noms de sous familles (gros caractères), de tribus (caractères intermédiaires) et de genres (petits caractères en italique) sont indiqués sur cette figure. Les sous famille des Papilionoideae et Mimosoideae sont monophylétiques alors que la sous famille des Caesalpinoideae est paraphylétique. Parmi les 17 espèces récalcitrantes identifiées dans la famille des Fabacées, quatre appartiennent à la même sous famille que *M. truncatula*, dont *C. australe.* Ces deux espèces sont encadrées en rouge dans l'arbre alors que les trois autres genres contenant des espèces récalcitrantes de la sous famille des Papilionoideae sont encadrées en bleu.

permet désormais l'acquisition de telles données. La comparaison de tissus végétatifs sensibles et tolérants à la dessiccation a déjà fourni des résultats intéressants chez les plantes reviviscentes (Oliver *et al.*, 2011 ; Gasulla *et al.*, 2013), les mousses (Yobi *et al.*, 2012) et les fougères (Farrant *et al.*, 2009).

3.2- Intérêt des Fabacées et de *M. truncatula*

Les Fabacées (anciennement Légumineuses) représentent la deuxième famille la plus importante du règne végétal, d'un point de vue économique, derrière les Graminées. Les cultures de Fabacées représentent environ un quart de la production mondiale, principalement pour l'alimentation humaine et animale de part la richesse en protéines et/ou lipides des graines (Graham & Vance, 2003). Parmi les espèces les plus importantes on retrouve le soja (*Glycine max*), le haricot (*Phaseolus vulgaris*), le pois (*Pisum sativum*) ou la cacahuète (*Arachis hypogea*). La particularité la plus intéressante des Fabacées est leur capacité à utiliser le diazote de l'air via une symbiose avec des bactéries du genre *Rhizobium* au niveau de nodosités racinaires. De Faria *et al.* (1989) ont estimé que 88 % des 3 395 espèces étudiées étaient capables de développer des nodosités dans des conditions favorables. Cette propriété renforce l'intérêt des Fabacées pour un développement d'une agriculture plus respectueuse de l'environnement. En effet, on estime qu'entre 40 et 65 millions de tonnes de diazote sont fixées chaque année par les cultures et les espèces sauvages de Fabacées (Graham & Vance, 2003). Par conséquent, des cultures de trèfle ou de luzerne en rotation avec des céréales permettent d'enrichir le sol en azote et de limiter l'apport d'engrais.

Medicago truncatula est une espèce proche de la luzerne cultivée qui a été choisi comme espèce modèle des Fabacées car elle présente des caractéristiques intéressantes. C'est une espèce diploïde, autogame avec un temps de génération relativement court et une production de graine conséquente (Cook, 1999). La phase de maturation de la graine s'étale sur environ 4 semaines (dont environ 10 jours de maturation tardive après remplissage) contre quelques jours pour *A. thaliana*. *M. truncatula* constitue donc un meilleur modèle qu'*A. thaliana* pour étudier la maturation des graines, notamment l'acquisition de la TD et de la longévité.

Son génome est relativement petit (environ 500 Mbases) et est presque entièrement séquencé. De plus, des méthodes de transformation de feuille par *Agrobacterium tumefaciens* et de racines par *Agrobacterium rhizogenes* ont été développées (Chabaud *et al.*, 2003 ; Araujo *et al.*, 2004 ; Crane *et al.*, 2006 ; Cheng *et al.*, 2011). Pour l'analyse fonctionnelle, plus de 19 000 lignées de mutants d'insertions ont été crées par la Samuel Roberts Noble Foundation. En intégrant les collections crées par des groupes européens, l'objectif est d'arriver à plus de 21 000 lignées couvrant 90 % de l'ensemble des gènes (Medicago-mutant.noble.org). Deux génotypes sont utilisés dans cette étude : A17 qui correspond au génotype séquencé et R108 pour lequel la régénération de plantes à partir de cals est plus efficaces (Hoffman *et al.*, 1997 ; Wright *et al.*, 2006). La banque de mutants d'insertion (comprenant *Mtabi3-1* et *Mtabi3-2* utilisés dans cette étude) a donc été produite dans le génotype R108 (Tadege *et al.*, 2005).

3.3- Choix de *C. australe* comme modèle de graine récalcitrante

Les Fabacées constituent également la troisième famille en taille du règne végétal avec environ 700 genres et 18 000 espèces (Graham & Vance, 2003), d'où une très grande diversité écologie et morphologique.

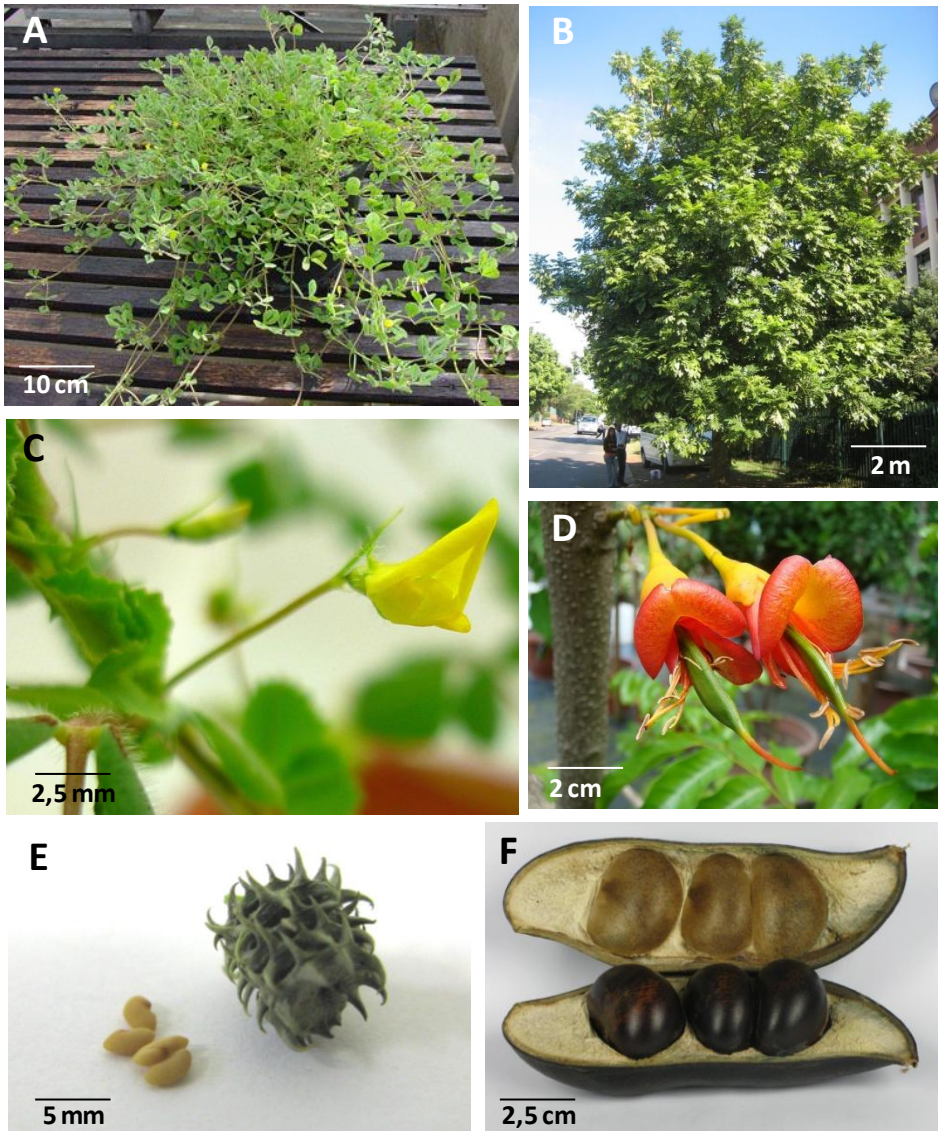


Figure 1.14. Morphologie des plantes (ou arbres), fleurs, gousses et graines de *M. truncatula* et *C. australe*. *M. truncatula* est une plante au port étalé **(A)** alors que *C. australe* est un arbre pouvant atteindre 45 m de haut **(B)**. Les fleurs de *M. truncatula* **(C)** et *C. australe* **(D)** sont zygomorphes (symétriques par rapport à un plan) ce qui caractérise l'appartenance à la sous famille des Faboidées. La taille et la forme des gousses et des graines de *M. truncatula* **(E)** et *C. australe* **(F)** diffèrent fortement. Les gousses de *M. truncatula*, enroulées en hélice, font moins d'un centimètre de long et contiennent de 8 à 10 graines. Celles de *C. australe*, de forme allongée, contiennent généralement de 2 à 6 graines et mesurent entre 10 et 25 cm. Une graine mature de *M. truncatula* est ovoïde, mesure environ 3 mm et pèse autour de 4 mg alors qu'une graine à l'abscission de *C. australe* est sphéroïde, mesure entre 3 et 5 cm et pèse entre 15 et 45 g. Les échelles de taille sont approximatives. (Photos : (A) INRA Montpellier ; (B) Jérôme Bove ; (C et E) Julien Delahaie ; (D) Tyrrenian sur Flickr ; (F) Benoît Ly Vu).

A ce jour, 17 espèces ont été clairement identifiées comme produisant des graines récalcitrantes chez les Fabacées (SID, Kew). Nous avons donc cherché l'espèce récalcitrante la moins éloignée de *M. truncatula* d'un point de vue phylogénétique afin de maximiser la similarité des séquences au niveau des transcrits. En se basant sur les classifications réalisées par Käss & Wink (1996) et Doyle (1995), la plus proche disponible est *Castanospermum australe* (Figure 1.13). Cette espèce, seule représentant du genre *Castanopermum* appartient à la même sous-famille que *M. truncatula* : les Faboidées (anciennement Papilionoidées). Il s'agit d'un arbre originaire d'Australie et planté comme arbre d'ornement en Afrique du Sud. Il est communément appelé châtaigner de Morton ou châtaigner d'Australie et peut mesurer jusqu'à 45 m en milieu naturel (environ 10 m en ornement). Les graines ne sont pas comestibles sauf après élimination des saponines et alcaloïdes qu'elles contiennent par une cuisson de plusieurs heures (Duke, 1989).

Cette espèce a été relativement peu étudiée en lien avec la récalcitrance mais la castanospermine, inhibiteur des glycosydases, a été proposée (puis abandonnée) comme candidat pour lutter contre les rétro-virus et notamment celui du VIH (Virus Immunodéficience Humaine) (Walker *et al.*, 1987). Les réserves accumulées dans la graine de *C. australe* sont principalement constituée de glucides (85 % de la matière sèche [MS]) (Harris, 1987) et notamment d'amidon (Thorburn *et al.*, 1987). La teneur en protéines est plus faible que chez *M. truncatula* (environ 8 % de la MS contre 30 à 39 % chez l'espèce orthodoxe) (Harris, 1987 ; Djemel *et al.*, 2005) et la teneur en lipide est de 5,8 % (Finkelstein & Grubb, 2002). Farrant *et al.* (1996) ont montré que la teneur en ABA était élevée dans ces graines (392 ng/g MS) comparée à d'autres espèces récalcitrantes d'origine tropicale. De plus, l'accumulation de déhydrines a été mise en évidence par Western blot dans les axes et cotylédons de graines mi-matures et matures (Farrant *et al.*, 1996 ; Han *et al.*, 1997). La taille du génome de *C. australe* a été évaluée à 554 Mb (Bennett *et al.*, 2000 ; Ohri *et al.*, 2004). A l'exception de deux séquences obtenues pour la réalisation d'arbre phylogénétique, aucune donnée n'est disponible dans GenBank. Les différences morphologiques et écologiques entre les plantes et les graines de *M. truncatula* et *C. australe* sont importantes (Figure 1.14).

Il est intéressant de noter que cette analyse comparative orthodoxe/récalcitrante n'aurait pas pu se faire avec *A. thaliana* car aucune espèce récalcitrante n'a été identifiée chez les Brassiccacées (Hong *et al.*, 1996).

Chapitre II

LEA polypeptide profiling of recalcitrant and orthodox legume seeds reveals ABI3-regulated LEA protein abundance linked to desiccation tolerance

LEA polypeptide profiling of recalcitrant and orthodox legume seeds reveals ABI3-regulated LEA protein abundance linked to desiccation tolerance

J. Exp. Bot. (2013) doi: 10.1093/jxb/ert274 First published online: September 16, 2013

Julien Delahaie[1], Michaela Hundertmark[1,2], Jerome Bove[1], Olivier Leprince[3], Hélène Rogniaux[4], Julia Buitink[5*]

[1]Université d'Angers, UMR 1345 Institut de Recherche en Horticulture et Semences, SFR 4207 QUASAV, 49045 Angers, France

[2]Current address: Vilmorin SA, Route du Manoir, 49250 La Ménitré, France

[3]Agrocampus Ouest, UMR 1345 Institut de Recherche en Horticulture et Semences, SFR 4207 QUASAV, 49045 Angers, France

[4]Institut National de la Recherche Agronomique, UR1268 Biopolymères, Interactions, Assemblages, Plate-forme Biopolymères - Biologie Structurale, 44316 Nantes, France

[5]Institut National de la Recherche Agronomique, UMR 1345 Institut de Recherche en Horticulture et Semences, SFR 4207 QUASAV, 49045 Angers, France

Abstract

In contrast to orthodox seeds that acquire desiccation tolerance during maturation, recalcitrant seeds are unable to survive drying. These desiccation sensitive seeds constitute an interesting model for comparative analysis with phylogenetically close species that are desiccation-tolerant. Considering the importance of LEA proteins as protective molecules both in drought-and desiccation tolerance, we characterized the heat stable proteome in cotyledons of the legume *Castanospermum australe* and compared it to the orthodox model legume *Medicago truncatula*. RNA sequencing identified transcripts of 16 homologues out of 17 LEA genes for which polypeptides are detected in *M. truncatula* seeds. We show that for 12 LEA genes, polypeptides were either absent or strongly reduced in *C. australe* cotyledons compared to *M. truncatula* seeds. Instead, osmotically responsive, non-seed specific dehydrins accumulated to high levels in the recalcitrant cotyledons compared to orthodox seeds. Next, we characterized *M. truncatula* mutants of the *ABSCISIC ACID INSENSITIVE3* (*ABI3*) gene. Mature *Mtabi3* seeds were found to be desiccation sensitive when dried below a critical water content of 0.4 g H2O g / dw. Characterization of the LEA proteome of the *Mtabi3* seeds revealed a subset of LEA proteins with severely reduced abundance that were also found reduced or absent in *C. australe* cotyledons. Transcripts of these genes were indeed shown to be *ABI3*-responsive. Our results highlight those LEA proteins that are critical to desiccation tolerance and suggest that comparable regulatory pathways responsible for their accumulation are missing in both desiccation-sensitive genotypes, revealing new insights into the mechanistic basis of the recalcitrant trait in seeds.

Keywords: *abi3*, *Castanospermum australe*, desiccation tolerance, late embryogenesis abundant proteins, *Medicago truncatula*, recalcitrant seed, RNAseq, proteomics

Introduction

Global agriculture and the conservation of plant biodiversity rely on seeds and their ability to be stored for long periods of time in dedicated national and international storage facilities (Li & Pritchard 2009; Walters *et al.*, 2013). The terms 'orthodox' and 'recalcitrant' are used to describe the storage behaviour of seeds. Orthodox seeds undergo maturation drying and are shed from the parent plant at low moisture contents. During maturation, they acquire desiccation tolerance, allowing them to be dried to moisture contents in the range of 1-5% without irreversible damage. Because of this ability, seeds can be stored for long periods in cold and dry vaults. Recalcitrant seeds, on the other hand, do not undergo maturation drying, and are shed at relatively high moisture contents. Such seeds are highly susceptible to desiccation injury, and thus are not storable under conditions suitable for orthodox seeds (reviewed in Farnsworth, 2000; Berjak & Pammenter, 2008; Li & Pritchard, 2009). The mechanisms by which recalcitrant seeds lose viability during drying and/or storage are not well understood, which poses a challenge to determine appropriate measures to better conserve these species.

In orthodox seeds, isolation and analysis of viviparous mutants and loss-of-function mutants impaired in embryogenesis and seed maturation resulted in the identification of master seed development regulator loci *lec1* and *abi3*, regulating partially and redundantly desiccation tolerance (Ooms *et al.*, 1993; Parcy *et al.*, 1997; To *et al.*, 2006). A third regulator, *FUSCA3,* appears to control seed longevity (Tiedemann *et al.*, 2008). In *Arabidopsis* and maize, some of the target genes of these activators are genes proposed to have a protective role in desiccation tolerance, such as small heat shock protein genes, genes with antioxidant functions, as well as Late Embryogenesis Abundant (LEA) genes (Kotak *et al.*, 2007; Bies-Etheve *et al.*, 2008; Mönke *et al.*, 2012).

LEA proteins are small hydrophilic, largely unstructured and thermostable proteins that are synthesized in orthodox seeds during mid- to late-maturation and in vegetative tissues upon osmotic stress. They are thought to have a range of protective functions against desiccation with different efficiencies, including ion binding, antioxidant activity, hydration buffering, membrane and protein stabilization (Tunnacliffe & Wise, 2007; Battaglia *et al.*, 2008; Amara *et al.*, 2012). Evidence of an *in vivo* role of these proteins in seed desiccation comes from *Arabidopsis thaliana em6* deficient mutants that show defects in maturation drying (Manfre *et al.*, 2009). A recent study in *A. thaliana* showed that down-regulation of seed-specific dehydrins (one of the LEA families) reduced seed survival in the dry state, although seeds did acquire desiccation tolerance (Hundertmark *et al.*, 2011). However, the precise role of LEA proteins in seed desiccation tolerance remains to be ascertained for the vast majority of them. Genomic studies to date have identified a large number of LEA genes whose expression is either restricted to seed tissues and/or up-regulated in response to biotic and abiotic stress in vegetative tissues (Illing *et al.*, 2005; Hundertmark & Hincha, 2008; Amara *et al.*, 2012; Chatelain *et al.*, 2012). Proteomic studies demonstrate that a subset of polypeptides accumulate during the acquisition of desiccation tolerance and/or longevity in orthodox seeds (Boudet *et al.*, 2006; Buitink *et al.*, 2006; Chatelain *et al.*, 2012). In desiccation-sensitive seeds of Arabidopsis mutants, transcript levels of several LEA genes were reduced, whereas other increased (*e.g.* dehydrins) (Bies-Etheve *et al.*, 2008).

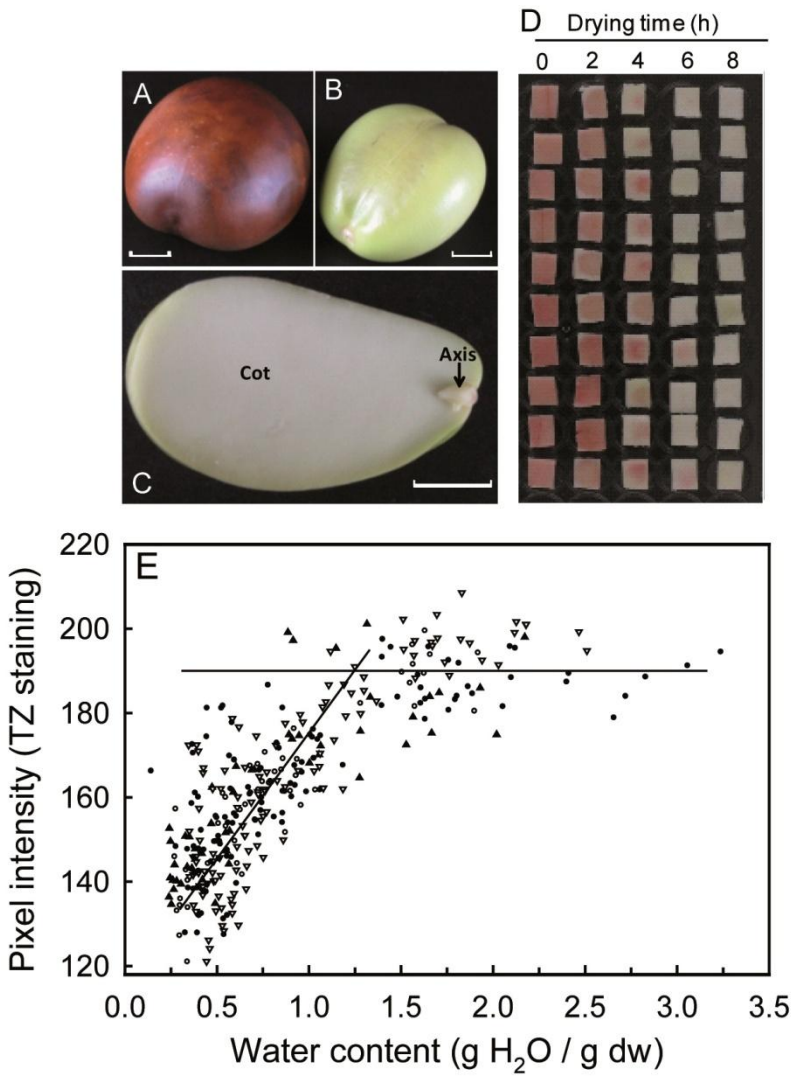


Figure 2.1. Determination of the critical water content of mature *Castanospermum australe* cotyledons.
C. australe seed after shedding with **(A)** and without seed coat **(B)**. **C)** Embryonic axis and cotyledons of mature seeds. Scale bar represents 1 cm. **D)** Tetrazolium staining (red being alive) of cubes (50 mm^3) taken from the core of mature cotyledons that were first dried at 44 % RH for the indicated time. **E)** The relation between water content during drying and pixel intensity of the tetrazolium (TZ) staining. Data for four independent experiments (represented by different symbols) are shown.

The situation regarding the occurrence and role of LEAs in recalcitrant seeds is ambiguous. Work has been constrained to detect members of the dehydrin family and showed that they are present in a range of species from different habitats, while apparently being absent from others. Several studies reported the presence of dehydrins in recalcitrant species of temperate origin, whereas these proteins could not be detected in some highly desiccation-sensitive seeds from certain tropical species (Finch-Savage *et al.*, 1994; Farrant *et al.*, 1996; Han *et al.*, 1997; Hinniger *et al.*, 2006; Panza *et al.*, 2007; Sunderlikova *et al.*, 2009; Ismail *et al.*, 2010; Lee *et al.*, 2012). In stored recalcitrant seeds of *Quercus robur* L., dehydrin mRNA can also be induced by ABA and limited drying treatments (Finch-Savage *et al.*, 1994). Whereas the presence/absence of dehydrins cannot explain the recalcitrant behavior of the species studied to date, several other families of LEA proteins exist in orthodox seeds that have not been studied in recalcitrant seeds.

A comparative analysis of recalcitrant and orthodox seed development is an interesting alternative to identify mechanisms involved in desiccation tolerance, especially if closely related species are compared (Kermode, 1997; Oliver *et al.*, 2011). Recently, the comparison of the metabolomic responses of drying leaves of two closely related grass species (sister group contrast), one being desiccation-tolerant and the other desiccation-sensitive, highlighted the metabolic predispositions associated with desiccation tolerance (Oliver *et al.*, 2011). In this study, we characterized a recalcitrant seed species of the *Papilionaceae* sub-family to allow comparison with our previous studies on orthodox seeds of the model legume *M. truncatula*. The phylogenetically closest recalcitrant species for which seeds can be currently obtained is *Castanospermum australe* A.Cunn ex Hook. (Doyle, 1995). *C. australe* is a tropical tree native of east Australia and now implanted in South Africa and Sri Lanka. In the seeds of this species, dehydrins were detected by western blot analysis (Han *et al.*, 1997). The absence of a sequenced genome of this species impedes the thorough molecular comparison of the entire LEA proteome with orthodox seeds. Thus, we used high throughput sequencing technology to obtain, assemble and annotate the transcriptome of these recalcitrant seeds. Whereas transcripts could be detected in *C. australe* for most LEA genes that are present in the desiccation-tolerant *M. truncatula* seeds, a comparative analysis of the LEA proteome profiles revealed that abundance for a number of seed-specific LEA proteins was severely affected in the recalcitrant seeds. In contrast, homologues of several dehydrins that are expressed in seedlings or non-seed tissues of *M. truncatula* submitted to osmotic stress accumulated to high levels in *C. australe* seeds. Comparison of the LEA proteome with desiccation-sensitive *abi3* mutants of *M. truncatula* showed a comparable reduction of a number of seed-specific LEA proteins.

Materials and methods

Plant material and treatments

Seeds of *M. truncatula* (A17) were produced as described in Chatelain *et al.* (2012). *C. australe* seeds were harvested during maturation and at shedding from trees growing in Pietermaritzburg (Kwazulu-Natal, South Africa) in 2009 and 2011. Within 48 h after collection, they were air-freighted to Angers (France) where there were immediately processed as indicated. From the 2009 harvest, embryos and cotyledons from immature and mature seeds were separated. Cotyledon tissues were used for the critical water content determination or dried one or three days over 75 % relative humidity (RH) NaCl before being frozen in liquid

nitrogen then stored at -80°C for RNA sequencing (Illumina) and proteomics. From the 2011 harvest, cotyledons and embryos were extracted from immature (green pods), mid-mature (yellow pods) and mature (brown pods) seeds and frozen fresh in liquid nitrogen then stored at -80°C. The 2011 harvest was used for analysis by 454 to improve the sequence assembly.

Desiccation sensitivity of mature *C. australe* cotyledons was determined on 3x5x3 mm cubes that were isolated from the inner part of the cotyledons. Cubes were dried for the indicated time intervals over a saturated salt solution at 75 % RH, after which they were divided in two halves. One half was used for water content determination, and the other half for viability assessment following incubation for 24h in a 1 % (w/v) tetrazolium solution (Sigma-Aldrich, France). Red color was quantified by pixel intensity on the image using ImageJ software (http://rsb.info.nih.gov/ij/). Water content was determined gravimetrically by weighing the seeds before and after drying in an oven for 48h at 96°C. Viability assays were performed on 4 independent drying experiments of 50-100 cubes.

Two *M. truncatula* mutants with *Tnt1* insertions in the *ABI3* gene (NF3185, hereafter referred to as *Mtabi3-1* and NF6003, *Mtabi3-2*) were obtained from the Samuel Noble Foundation (Oklahoma, USA). *Tnt1* insertions in two mutants were verified by PCR (Annex 1 for primers). Mutant and wild-type lines (R108) were multiplied in a growth chamber according to Chatelain *et al.* (2012), and lines were backcrossed once for the *Mtabi3-1* and twice for the *Mtabi3-2* mutants.

Desiccation tolerance was determined on seeds that were harvested at different time points during development. Two to five replicates of 30-50 seeds were rapidly dried to 0.09 g H_2O g DW-1 over an airflow of 43% RH, and rehydrated after two days on filter paper at 20°C in the dark. Seeds were considered desiccation-tolerant when they germinated, scored by the protrusion of the radicle through the seed coat. For ABA insensitivity assays, triplicates of 40-50 of freshly harvested seeds just prior to pod abscission (0.8-1.0 g H_20/g DW) were imbibed on filter paper on a range of ABA concentrations (mixed isomers, Sigma, St Louis, MO, USA) at 20°C. ABA was dissolved in methanol prior to dilution in water. Control seeds were imbibed in the MeOH concentration corresponding to the highest ABA concentration (0.5 % MeOH). Germination was scored after 14 days. For proteomic analysis, *Mtabi3-1* and *Mtabi3-2* seeds were harvested at the point of abscission, when seeds were still viable. For RT-PCR analysis, seeds were harvested at 24 DAP.

Cloning of *MtABI3*

To obtain the full-length sequence for *MtABI3*, genomic DNA was extracted using the Nucleospin Food kit (Macherey Nagel) from leaf material of *M. truncatula* A17. An inverse PCR was performed on 5 µg genomic DNA that was digested with EcoRI (25u/50µl final volume, Promega, Madison, WI, USA), and ligated using T4 DNA ligase (50u/450µL final volume, Fermentas, Vilnius, LT). The full length was amplified on the ligated DNA using the primers indicated in Annex 1 that were designed based on the *MtABI3*-fragment available in the EST database (TC97588, The DFCI *Medicago truncatula* Gene Index v8). The full-length genomic DNA fragment (3,458 bp) was cloned into pJet1.2 (CloneJET kit, Thermo Scientific, Bremen, Germany) and sequenced (Annex 1 for primers).

RNA extraction, and sequencing and assembly

Table 2.1. Contigs features from 454 and Illumina RNAseq of *Castanospermum australe* seeds and annotation of the corresponding transcriptome.

	454	Illumina	454 + Illumina
Total number or reads after sequencing	**626,225**	**7,784,004**	**8,410,229**
Total number of contigs	**36,767**	**18,483**	**48,334**
Average contig length	869	318	773
N50	1,001	345	1,020
Number of nucleotides in contigs	31,937,738	5,882,730	37,365,884
Total number of contigs annotated	**28,391**	**16,018**	**35,050**
Contigs annotated with MT3.5	19.885	11,987	25,615
Contigs annotated with MtGI11	4,847	3,207	6,138
Contigs annotated with Swissprot	1,470	414	739
Contigs annotated with NR (NCBI)	2,189	410	2,558

For *M. truncatula* seeds, total RNA was extracted using the nucleospin RNAplant kit (Macherey Nagel, Düren, Germany) and 10 µg of total RNA from each sample were DNAse treated (Turbo DNAse, Ambion) and purified (RNeasy MinElute Cleanup kit, Qiagen), according to manufacturer's instructions. Total RNA was extracted with phenol from cotyledons or embryonic axes of *C. australe* as described by Bove *et al.* (2005). Quantity, purity and integrity of RNA were checked using a NanoDrop ND-1000 UV-VIS spectrophotometer (NanoDrop Technologies) and a bioanalyzer (Experion, BioRad). From the 2009 RNA pool, a cDNA library was prepared, normalized and sequenced by GenXPro GmbH (Frankfurt am Main, Germany) using Illumina technology (Genome Analyser-IIx). From the 2011 RNA pool, a cDNA library was prepared, normalized and sequenced by Eurofins (Ebersberg, Germany) using the 454 GS FLX+ technology. Reads obtained from each sequencing were assembled *de novo* in two steps: first with MIRA 3.4.0 (Chevreux *et al.*, 2004) then with DNA Dragon (SequentiX, http://www.sequentix.de/software_dnadragon). The detailed procedure is described in figure 6.3.

Functional annotation and classification

Contig annotation to known sequences by sequence similarity was performed using two *M. truncatula* nucleic databases: MT3.5 from the International Medicago Genome Annotation Group (IMGAG) and MtGI11 from the Dana-Farber Cancer Institute (DFCI) Medicago Gene Index. Contigs that remained un-annotated after these two analyses were blasted using Blast2GO (version 2.6.0) (Götz *et al.*, 2008) against protein databases including all plant species: Swissprot and non redundant protein from NCBI. Next, classification of *C. australe* annotations in Gene Ontology (GO) was performed by Blast2GO. GO terms were retrieved from public databases and mapped to each contig, after which the most specific ones were selected by an annotation rule. The detailed annotation workflow is described in Figure 6.3.

Reverse transcriptase PCR

Two µg of *M. truncatula* wild type and *abi3-1* and *abi3-2* RNA was retro-transcribed according to the manufacturers' instructions (Thermo Scientific, Bremen, Germany). Resulting cDNA were diluted 1:3. Primers sequences and annealing temperatures are provided in Annex 1. PCR was performed with DreamTaq (Fermentas, Vilnius, Lithuania) according to manufacturers' instructions.

Protein extraction and 2D gel electrophoresis

Total soluble proteins were extracted in triplicate from 25 seeds of *M. truncatula* (A17, R108 or *Mtabi3-1* and *Mtabi3-2*) and 400 mg of cotyledons of mature *C. australe* from a minimum of three seeds for each replicate (2009 harvests) according to Boudet *et al.* (2006) and the heat stable proteins were recovered according to Chatelain *et al.* (2012). After centrifugation at 20 000 g at 4°C, the pellet was successively washed with 100 µL of 80 % acetone, 100 % acetone, 80 % ethanol and 100 % ethanol then resuspended in 300 µL of rehydration buffer for 36h according to Boudet *et al.* (2006). Protein concentration was assayed by Bradford (1976). Heat stable proteins fractions of *M. truncatula* and *C. australe* (150µg), as well as a 1:1 mix of both protein fractions (300 µg) were rehydrated and separated on 24 cm immobilized non linear pH 3–10 gradient strips (Bio-Rad, Hercules, CA, USA). Isoelectrofocusing was performed at 20°C, with 3h at 250 V, then 4h at 6

Table 2.2. LEA transcripts identified in the *Castanospermum australe* seed transcriptome.

Family (PFAM)	protein name	Blast database	*M. truncatula* IMGAG 3.5	*C. australe* contigs[1]	E-value	Aln length	% iden	% sim	Spot 2D gel
dehydrin	DHN3	MtGI11	TC175037	Ca_11990	2.E-09	198	47.2	52.0	3,4
	DHN-cognate	Mt3.5	Medtr3g117290	Ca_9276	2.E-10	120	53.3	65.0	5
	BudCar5	Mt3.5	Medtr7g086340	Ca_31427	5.E-06	119	62.2	77.7	1
LEA_5	EM6	Mt3.5	Medtr4g016960	Ca_14307	4.E-45	98	81.6	90.8	37
	EM1	MtGI11	AJ498523	Ca_2036	9.E-40	100	78.2	85	51
LEA_4	SBP65	Mt3.5	Medtr4g079690	Ca_7340	5.E-17	90	58.9	72.2	75
	PM10	MtGI11	TC174929	Ca_3035	4.E-43	274	59.9	74.8	ND
	PM18	MtGI11	TC183861	Ca_330	2.E-41	335	56.0	67.2	ND
	MP2	Mt3.5	Medtr1g061730	Ca_8462	3.E-31	226	57.5	67.1	9
	LEAm	Mt3.5	Medtr2g014040	Ca_7604	3.E-45	268	48.6	62.4	74
	CapLEA.I	MtGI11	TC175990	Ca_8841	1.E-31	141	66.0	85.8	18,19
LEA_1	PM1	Mt3.5	Medtr7g093170	Ca_8304	3.E-16	91	64.8	75.8	ND
	D113.II	Mt3.5	Medtr7g093160	Ca_8304	5.E-17	91	65.2	76.1	ND
SMP	PM25	MtGI11	TC174777	Ca_6007	2.E-83	238	75.6	84.4	92
	D-34.I	MtGI11	Medtr1g072090	ND					ND
	D-34.II	MtGI11	TC183570	Ca_25629	6.E-42	128	86.7	93	ND
	D-34.III	Mt3.5	Medtr2g076230	Ca_23377	3.E-06	24	76.0	87.5	ND

[1]*C. australe* contigs that were homologs to the 17 LEA gene products detected in *M. truncatula* seeds were identified based on the MtGI11 and Mt3.5 Medicago databases (E-value < e-06). ND = homologue not detected. Contigs were translated to calculate the percentage of identity (iden) and similarity (sim). Aln, alignment.

kV, followed by a gradual increase to 27 kVh at 6 kV/h and to 40 kVh at 8 kV/h) in a Bio-Rad Protean isoelectric focusing cell. Size separation of proteins was performed on vertical polyacrylamide gels (12 % (w/v) acrylamide) in a Ettan Daltsix Electrophoresis system (Amersham Biosciences, Orsay, France) according to Boudet *et al.* (2006) using a running buffer containing 15.6 mM Tris (pH 8.3), 120 mM Gly, 0.1 % (w/v) SDS. Gels were stained with 0.08 % (w/v) Brillant Blue G-Colloidal for 24 h, destained briefly in 5% (v/v) acetic acid and 25 % (v/v) methanol then in 25 % (v/v) methanol for 8 h. Stained gels were scanned at 63.5 x 63.5 resolution using a GS 800 scanner (Bio-Rad). At least three digitalized gels from three independent experiments (extraction, focalization and migration) were analyzed using the PDQuest 7.2.0 software (Bio-Rad). Spot intensities were normalized using the total quantity in valid spot method. A paired t-test was performed to analyze differences in intensity between *C. australe* and *M. truncatula* LEA proteins and between WT (R108) and *Mtabi3-1*/*Mtabi3-2* seeds.

Mass spectrometry and protein identification

Spots of interest were excised from the 2D gels and subjected to in-gel tryptic digestion according to Chatelain *et al.* (2012). Tryptic fragments were analysed by LC-ESI-MS/MS spectroscopy using a nanoscale HPLC (Famos-Switchos-Ultimate system, LC Packings, Dionex, San Francisco, CA, USA) coupled to a hybrid quadrupole orthogonal acceleration time-of-flight mass spectrometer (Q-TOF Global, Micromass-Waters, Manchester, UK) as described in Boudet *et al.* (2006). Mass data were analysed with the Protein Lynx Global Server software (Micromass-Waters). Protein identification was performed by comparing the data with the UniProt sequence databank (date of release: August 2010) or with TIGR Medicago EST databank (date of release: April 2010). For the *M. truncatula* heat stable proteome, spots linked to LEA polypeptides were identified according to the reference gel published by Chatelain *et al.* (2012).

Data submission

Raw sequence data from this article can be found in the Sequence Reads Archive database (NCBI) under BioProject PRJNA193308. The data on the ectopic expression of *MtABI3* in hairy roots discussed in this publication have been deposited in NCBI's Gene Expression Omnibus (Edgar *et al.*, 2002) and are accessible through GEO Series accession number GSE44291
(http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE44291).

Results

Physiological description of *C. australe* seeds

At reception, fresh weight and water content of mature, shedded *C. australe* seeds was 45.9 ± 14.9 g/seed and 1.94 ± 0.41 g H_20 g dry weight $(DW)^{-1}$, respectively. The embryo is composed of two prominent cotyledons and a comparatively small axis and surrounded by a thin brown testa (about 250 µm thickness) (Fig. 2.1A-2.1C). When planted, fresh seeds germinated at 100 % and produced healthy seedlings. The critical water content corresponding to the onset of loss of cell viability during rapid drying was determined on cotyledons using tetrazolium staining. During drying, the loss of viability as a function of water content showed the typical

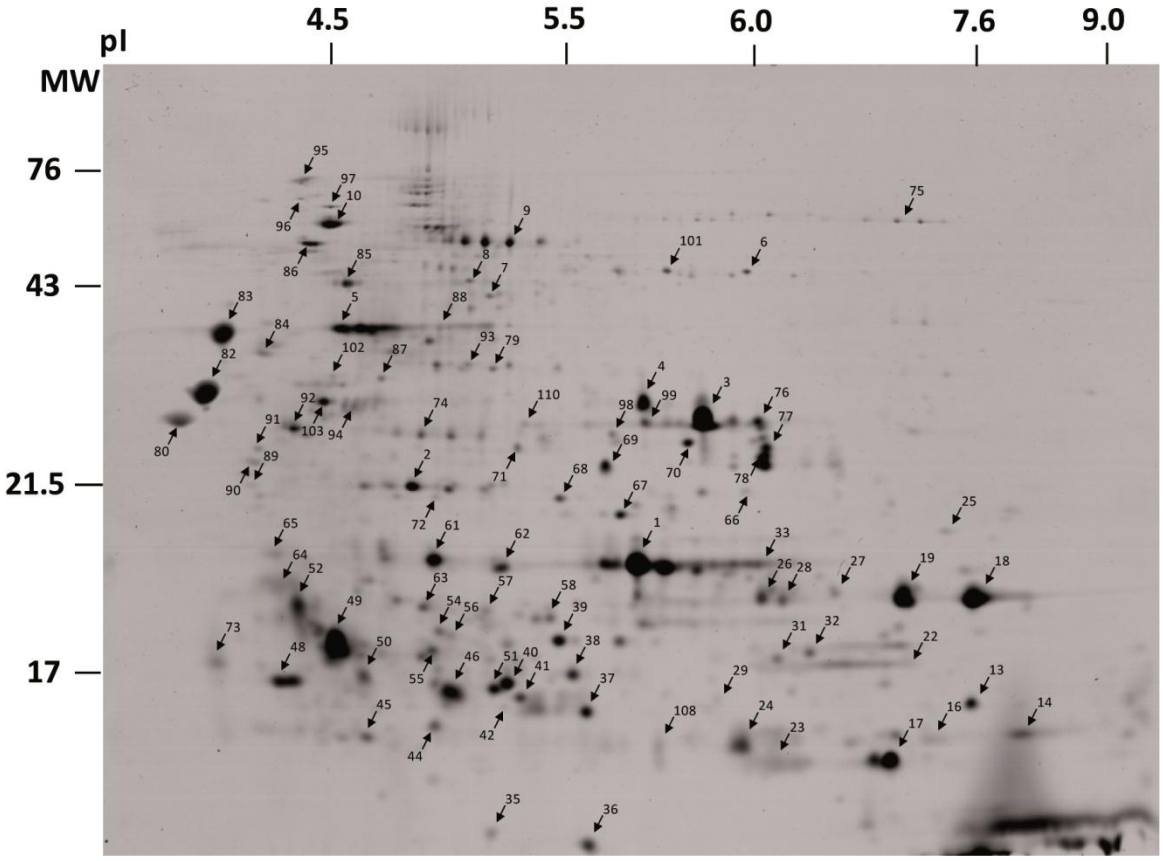


Figure 2.2. Reference map of the heat-stable proteome of mature *Castanospermum australe* cotyledons. Hundred-fifty µg of the heat-stable proteins was separated by 2D SDS-PAGE using 24 cm non-linear immobilized pH gradient strips (3–10). pI and molecular mass (in kDa) are indicated. Numbers indicate the polypeptides that were sequenced (see Annex 3 Table 2).

pattern found in other recalcitrant seeds (Fig. 2.1D and 2.1E). The intensity of the staining remained high and constant until a water content of *ca.* 1.5 g H_20 g DW $^{-1}$ was reached, after which the intensity decreased progressively with further drying (Fig. 2.1E). The critical water content, here defined as the water content corresponding to the break point for which cotyledon tissues begin to lose staining intensity, was estimated at 1.2 g H_20 g DW^{-1}. Tissues were completely absent of red staining when dried below 0.5 g H_20 g DW^{-1} (Fig. 2.1D), indicating total loss of viability. This value is consistent with that reported by Han *et al.* (1997) on isolated axes of *C. australe* during rapid drying using electrolyte conductivity as an indication of membrane damage. In contrast, complete drying and rehydration of 10h imbibed *M. truncatula* cotyledons in tetrazolium solution rendered the tissues red, indicating that viability was maintained (data not shown). For the proteomic study, we chose to focus on cotyledons as a model for desiccation-sensitive tissues, due to their high critical water content. Cotyledons are large and surround the axis, thereby possibly slowing rate of dehydration of the latter. In *M. truncatula*, few differences were found between LEA abundance and composition of the axes and cotyledons, except for EM6 (Chatelain *et al.*, 2012).

Sequencing of the *C. australe* seed transcriptome and identification of LEA contigs

To enable comparative analysis on a molecular level between the recalcitrant *C. australe* and its orthodox counterpart *M. truncatula*, sequence information on transcripts present in seeds was obtained from a range of tissues to capture the maximum variation of the transcriptome at harvest: intact isolated axis, cotyledons from three developmental stages, and partially dried yet alive cotyledon tissues (Figure 6.3). Using Illumina and 454 technologies, sequencing of the normalized cDNA libraries resulted respectively in 7,784,004 paired reads of 76 bp long and 626,225 reads with an average length of 376 bp (Table 2.1). The assembly resulted in 48,334 contigs varying between 200 bp and 14,334 bp long with an average length of 773 bp (Figure 6.3 ; Table 2.1). A total of 35,050 contigs (72.5 %) were annotated, of which 91 % were provided by the IMGAG 3.5 *M. truncatula* database and the MtGI11 version of the Medicago EST database. An additional 740 and 2,558 of the remaining contigs were identified respectively using Swissprot and NR databases related to other plant species (Figure 6.3). Annotations were classified according to Gene Ontology (GO) using the Plant-GO-slim version of Blast2GO (B2G). Enzyme Classification (EC) numbers were retrieved with the additional functionalities of B2G linked to KEGG pathways. A total of 23,637 contigs (48.9%) were annotated with 98,615 GO terms and 8,962 EC numbers. Distribution of the main biological processes (BP, 45,568 annotations), molecular functions (MF 35,110 annotations) and cellular components (CC 17,937 annotations) is shown in figure 3.3.

Using the annotated transcriptome of *C. australe*, the next step was to perform a comparative analysis between LEA sequences found in both legume species. In *C. australe*, contigs were found for 29 LEA genes that are identified in the *M. truncatula* genome (Annex 2). In mature seeds of *M. truncatula*, a proteome analysis led to the detection of polypeptides corresponding to 17 genes (Chatelain *et al.*, 2012). For 16 out of these 17 genes, at least one corresponding contig was detected in the *C. australe* transcriptome (Table 2.2). Amino acid sequence alignment displayed between 52 % to over 90 % similarity between the two species. The dehydrin family members were most divergent, with similarity that ranged only from 52 to 77 % (Table 2.2). Other

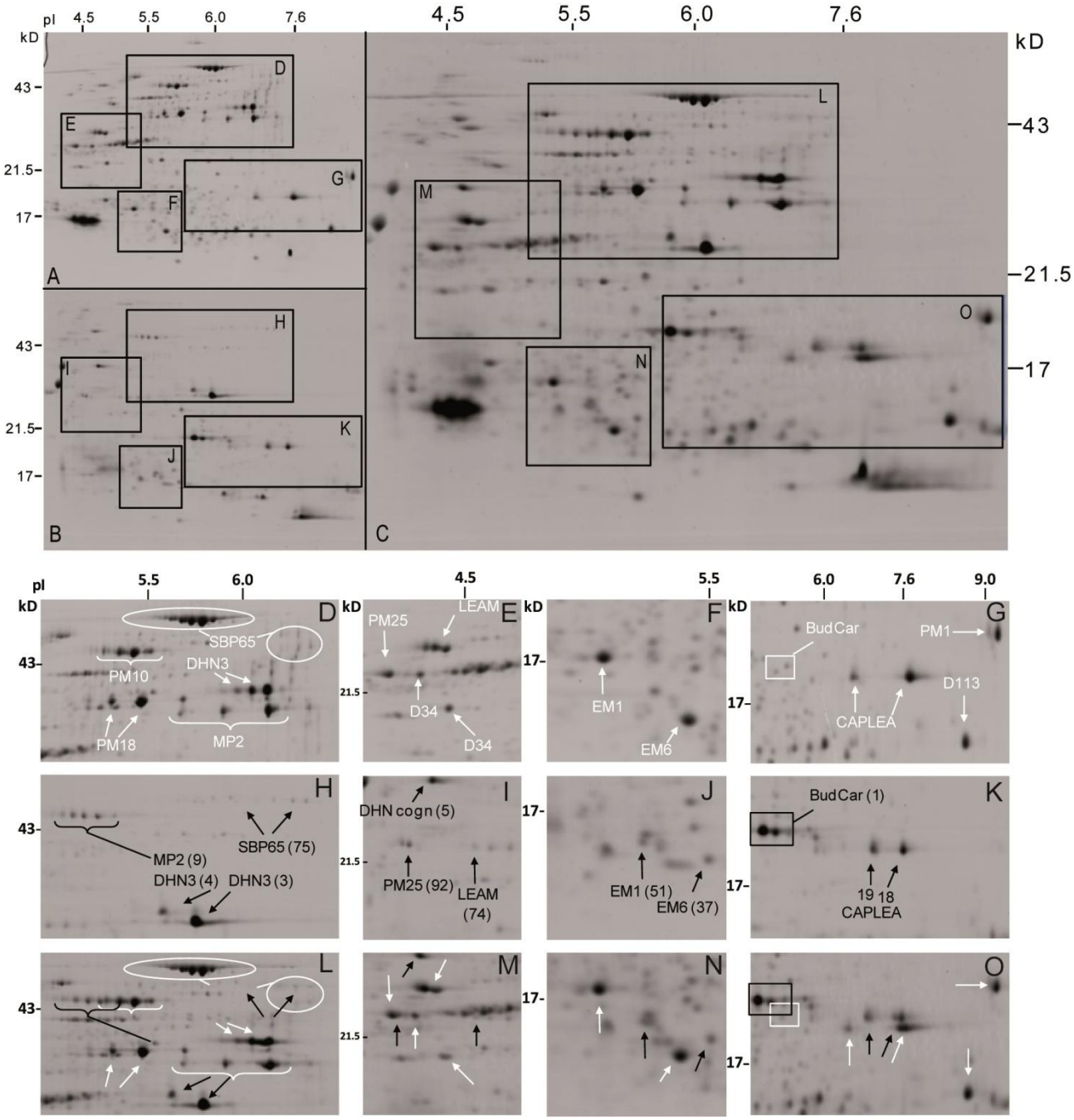


Figure 2.3. Comparative analysis of LEA polypeptides in cotyledons of *Castanospermum australe* and *Medicago truncatula* seed heat stable proteome.

Reference map of the heat stable proteome of *M. truncatula* **(A)** and *C. australe* seeds **(B)** and separation of a mixture of equal amounts of heat stable proteins (150 µg) from both species **(C)**. Panels D-O show details of different regions of 2D gels of the proteome of *M. truncatula* **(D-G)**, *C. australe* **(H-K)** and both species **(L-O)**. Indicated spots refer to Annex 3 and Annex 4.

families are much more conserved between the two species, such as the SMP and LEA_5 family (EM1 and EM6), showing 76-86 % identity and 85-91 % similarity with *M. truncatula,* respectively.

Identification of the heat-stable proteome of *C. australe*

Identification of polypeptides corresponding to the 16 LEA transcripts that were detected in *C. australe* transcriptome was carried out by separation of the heat-stable protein fraction on 2D gel electrophoresis (Fig. 2.2). This method has been successfully applied to characterize and quantify the entire LEA proteome of *M. truncatula* (Boudet *et al.*, 2006; Chatelain *et al.*, 2012). A total of 110 spots were sequenced using LC-ESI-MS/MS spectroscopy, out of which 82 spots were identified (Annex 3).

Polypeptides were detected for 10 of the 16 LEA genes identified from the *C. australe* sequence assembly (Table 2.2). These polypeptides include two highly abundant dehydrins (CaDHN3 (spot 3 and 4) and CaBudcar5 (spot 1)) and CaCapLEA-1 (spot 18 and 19). Other less abundant LEA polypeptides include one more dehydrin (CaDHN cognate, spot 5), the two LEA_5 members (CaEM1 (spot 51) and CaEM6 (spot 37)), three LEA_4 members (CaSBP65 (spot 75), CaMP2 (spot 9) and CaLEAm (spot 74)) and CaPM25 (spot 92). For six LEA contigs, no polypeptides were identified in *C. australe* proteome, despite the presence of their transcripts (Table 2.2) In *M. truncatula*, four of these LEA proteins are highly abundant in mature seeds and include two members of the LEA_5 family (CaPM10 and CaPM18) and both LEA_1 members (CaD113.I and CaPM1) (Chatelain *et al.*, 2012). The other two LEA proteins that were not identified are members of the SMP family (CaD34.II and III).

Next to LEA polypeptides, other abundant polypeptides were detected in the heat-stable *C. australe* proteome. They were identified as three pathogenesis related proteins (spots 26, 27 and 29) (Annex 3), five polypeptides corresponding to small heat shock proteins (spots 2, 56, 61, 62 and 93), and four polypeptides corresponding to superoxide dismutases (38, 54, 67 and 69). Furthermore, two desiccation-related polypeptides (spots 82 and 88) were detected with homology to Lb_13-62 and PCC13-62. These genes are up-regulated in the desiccation-tolerant resurrection plants *Craterostigma plantagineum* and *Lindernia brevidens* (Phillips *et al.*, 2008) and were also recently detected in floral nectar of the evergreen velvet bean (Mucuna sempervirens Hemsl) (Zha *et al.*, 2013).

Comparative analysis of the LEA proteome between *C. australe* and *M. truncatula*

The amount of heat stable proteins relative to the total soluble protein fraction was lower for *C. australe* (20 % ± 1.8) than for *M. truncatula* seeds (36 %, Chatelain *et al.*, 2012). Equal amounts of the heat stable protein fraction of *C. australe* or *M. truncatula* were separated by 2D gel electrophoresis and 2D profiles were compared. For most of the LEA polypeptides, the exact position on the gel differed slightly between both species (Fig. 2.3A and 2.3B). This made it possible to combine the two extracts and separate them on the same gel, allowing for an accurate comparative quantification of the polypeptides from both species and avoiding the drawbacks associated with variations due to polypeptide migration and gel staining (Fig. 2.3C, Annex 4). In both species, the dehydrin DHN3 (Fig. 2.3D, 2.3H and 2.3L) and LEA_4 CAPLEA (Fig. 2.3G, 2.3K and 2.3O) were present with high spot intensity. For six LEA polypeptides, spot intensity was much lower in *C. australe* compared to *M. truncatula*; SBP65 and MP2 (Fig. 2.3D, 2.3H and 2.3L), PM25 and LEAm (Fig. 2.3E, 2.3I and

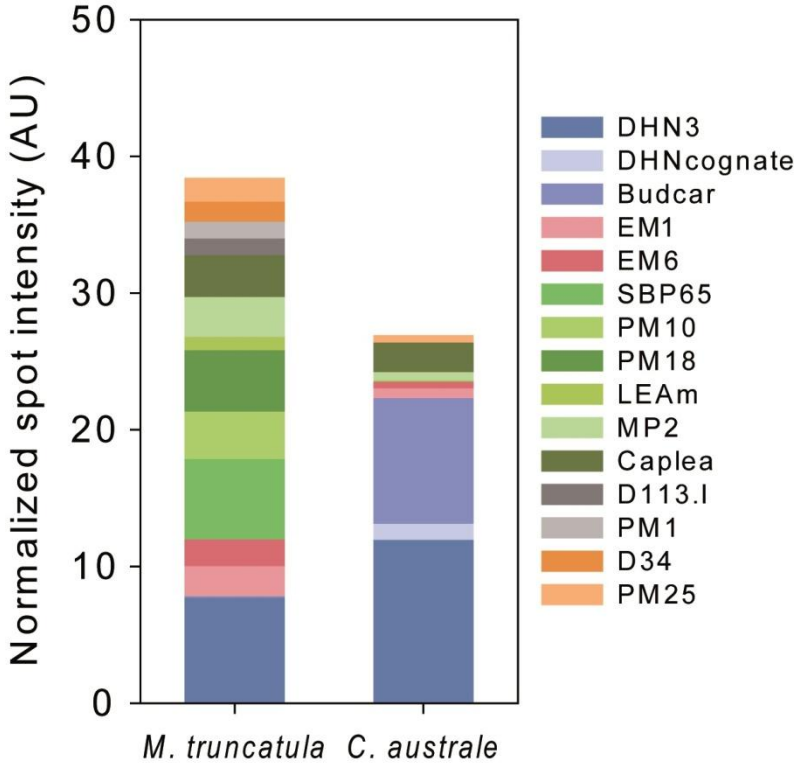


Figure 2.4. Relative abundance of the different LEA polypeptides identified in cotyledons of *Castanospermum australe* and *Medicago truncatula* seeds.

Abundances were calculated based on the spot intensities of 3 replicates of the gels shown in Fig. 3.

2.3M) and EM1 and EM6 (Fig. 2.3F, 2.3J and 2.3N) (Annex 4). The other two dehydrins, Budcar5 and DHN-cognate, are highly abundant in *C. australe*, whereas their homologues in *M. truncatula* are barely detectable.

A quantitative overview of the comparative analysis of the LEA proteome, based on relative spot intensity, is presented in Fig. 2.4. The LEA profile is strikingly different between the recalcitrant and orthodox seeds. In contrast to mature *M. truncatula* seeds, where dehydrins comprise 20 % of the LEA proteome, this family represents 83 % of the LEA proteome of *C. australe* cotyledons. Four LEA proteins (CaEM1, CaEM6, CaMP2 and CaPM25) were four-fold less abundant in the recalcitrant seeds compared to the orthodox *M. truncatula*, whereas CaLEAm and CaSBP65 relative abundance was reduced over twenty-fold. In addition, six LEA proteins were not detected in cotyledons of *C. australe* (CaPM1, CaD113.I, two CaD34 members, CaPM10 and CaPM18). CAPLEA was present in comparable amounts in both species.

Characterization of the LEA proteome of the desiccation sensitive *Mtabi3* mutant seeds and comparison with *C. australe*

To further investigate a cause-effect relationship between the lack of these LEA proteins and desiccation sensitivity, we examined the LEA proteome in an orthodox seed that was rendered desiccation sensitive by knocking out the *MtABI3* gene expression. First, we obtained two independent homozygous *Tnt1* insertion mutants (*Mtabi3-1* and *Mtabi3-2*) that were backcrossed once or twice, respectively (Fig. 2.5A). The Tnt1 insertions were located at 1,595 bp and 1,605 bp from the start-codon, respectively, just after the B2 domain (Fig. 2.5A). RT-PCR analysis confirmed the absence of transcripts in the two *Mtabi3* mutants (Fig. 2.5B). The resulting freshly harvested seeds were used for a physiological characterization. Like in *abi3* mutants of Arabidopsis (Ooms *et al.*, 1993), mature *Mtabi3* seeds retained their chlorophyll (Fig. 2.5C) and exhibited a strongly reduced sensitivity to ABA (Fig. 2.5D). During seed maturation between 24 and 32 DAP, the seed water content of *abi3* mutants remained at around 1,6 g H_2O g DW^{-1}, whereas in developing wild type seeds it decreased steadily from 1.2 to 1.0 g H_2O g DW^{-1} (Fig. 2.5E). Thereafter, in both *abi3* mutants and wild-type, the water content decreased until 40 DAP. During the latter stages of drying, when pods were detached, seeds of both genotypes exhibited a similar rate of water loss. Desiccation tolerance of harvested seeds was determined as a function of their water content at different stages during maturation and after enforced drying (at 40 DAP) (Fig. 2.5F). In contrast to wild-type, the seed population of *Mtabi3* mutants started to lose their viability when water content decreased below 1.0 g H_2O g DW^{-1} and decreased sharply below 0.5 g H_2O g DW^{-1} (Fig. 2.5F). At 0.2 g H_2O g DW^{-1}, all *abi3* seeds were dead. In contrast to wild type seeds, fully mature, dried seeds did not germinate, and tetrazolium tests showed no staining, indicating that viability was completely lost (data not shown).

Next, LEA polypeptide abundance was determined using 2D-gel electrophoresis on three replicates of *Mtabi3-1*, *Mtabi3-2* and wild type seeds (R108 background) (Annex 5). To be able to compare LEA profiles among *C. australe* and *Mtabi3* genotypes, polypeptide abundance was expressed as the relative difference with *M. truncatula* wild-type (A17 for the comparison with *C. australe*, and R108 for the comparison with *Mtabi3* mutants) (Fig. 2.6). The intensity of MtPM1 was highly variable amongst the samples, irrespective of the *M. truncatula* genotypes (Annex 5). This might be due to the very basic nature of this protein in the R108 genotype, placing it at the border of the gel where resolution is poor. Likewise, the intensity of MtLEAm and

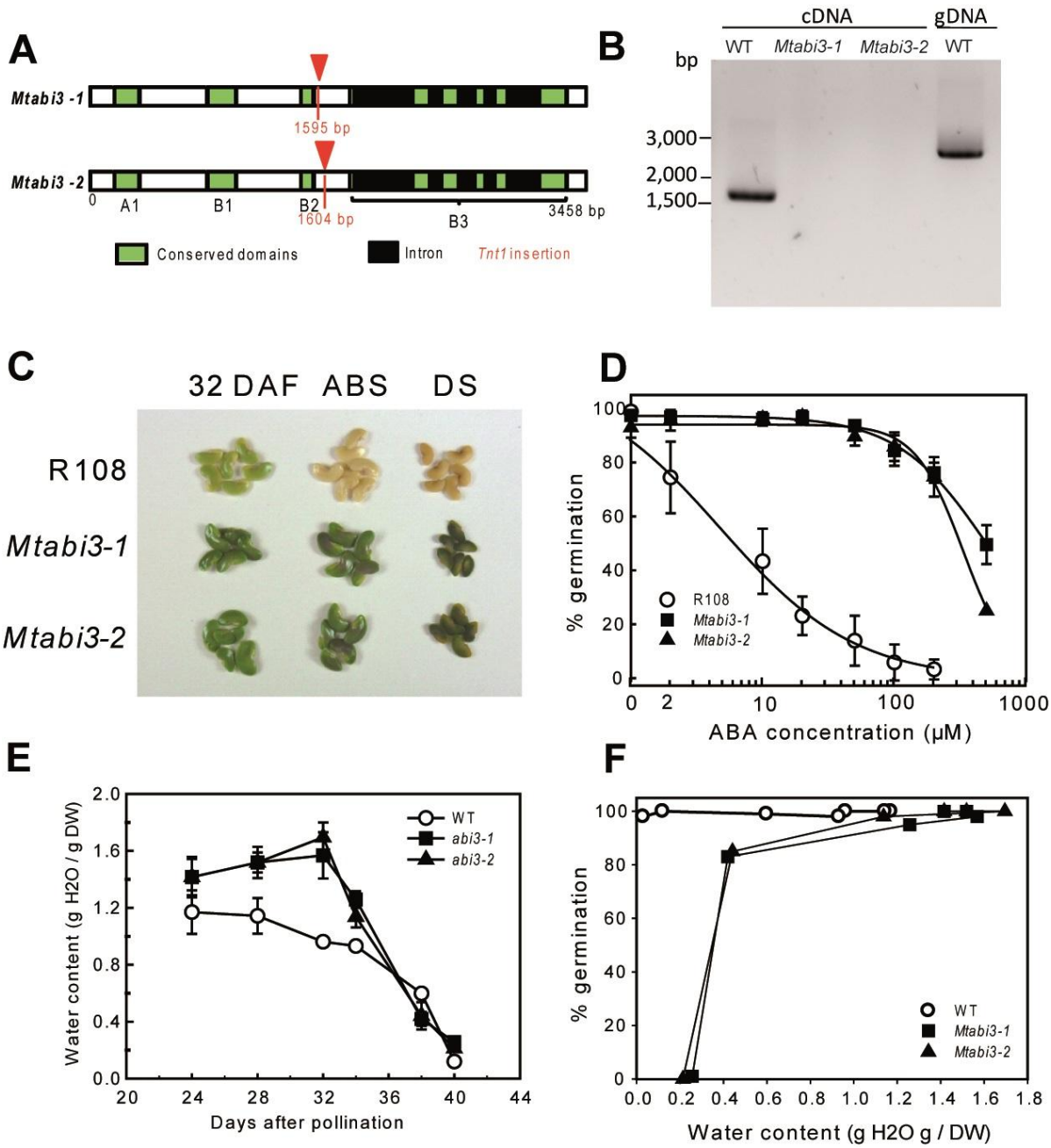


Figure 2.5. Characterization of *abscisic acid insensitive3* (*Mtabi3*) mutants of *Medicago truncatula*.
A) Gene structure, position of the A and B domains and *Tnt1* insertions within the *MtABI3* gene. **B)** Validation of absence of ABI3 transcripts in the *Mtabi3-1* and *Mtabi3-2* mutants. Using the same primer set, ABI3 was also amplified on genomic DNA. The increased size corresponds to the additional introns. **C)** Seed color phenotype of *Mtabi3-1* and *Mtabi3-2* and corresponding wild-type seeds (R108) at three stages of maturation: 32 days after pollination (DAP), at pod abscission (ABS, 38 DAP) and in dry seed (DS). **D)** ABA dose response analysis during germination of seeds collected at pod abscission. Germination was scored as emergence of the radicle. Data are the average of three replicates of 40-50 seeds ± SE. **E)** Changes in seed water content during development. Data are the average of three replicates of three seeds ± SE. F) Germination of *Mtabi3* and wild type seeds at different stages of development upon rehydration of 70-80 seeds. Data are significantly different when they differ by 18% or more (χ_2-test, $P<0.05$).

D34.II could not be determined correctly (Fig. 2.3, Annex 5). To avoid wrong interpretation of these data, they were omitted from further analysis. Overall, the abundance of the LEA proteome of the *Mtabi3* mutants compared to wild type seeds resembled that of desiccation-sensitive *C. australe* cotyledons when compared to *M. truncatula* (Fig. 2.6). Like in *C. australe*, abundance of nine LEA polypeptides from several families was decreased in the *Mtabi3* mutants, namely LEA_5 (MtEM1 and MtEM6); SMP (MtPM25 and MtD34), LEA_4 (MtSBP65, MtPM18, MtPM10 and MtMP2) and LEA_1 (D113.I). The dehydrin MtDHN3 was more abundant in seeds of *Mtabi3* mutants than in wild type seeds, which further underscored the similarity with *C. australe*. The relative amount of MtCAPLEA was slightly lower in seeds of the *Mtabi3-2* mutant compared to the wild-type seeds, whereas its amount was higher in the *Mtabi3-1* mutant (Fig. 2.6, Annex 5).

ABI3 regulation of identified LEA proteins in relation to desiccation tolerance

The reduction of LEA polypeptides in *Mtabi3* mutants raises the question whether the twelve reduced or absent LEA proteins in *C. australe* are regulated by ABI3 at the gene level. In *abi3* seeds of Arabidopsis, transcript levels of all LEA genes that are affected in the desiccation-sensitive tissues (group A) were decreased (Table 2.3, Bies-Etheve *et al.*, 2008). This was not the case for the LEA proteins that were not affected or more abundant in these tissues (group B). Transcript levels of the homologues of *DHNcognate* and *CAPLEA1* were even increased in the *abi3* mutants. In addition, to investigate whether MtABI3 regulates LEA targets, we took advantage of our recent transcriptome study on the effect of overexpressing *MtABI3* in the hairy roots of *M. truncatula* (GeOmnibus GS GSE44291). The advantage of this ectopic expression model is that it avoids the interfering effects of other B3 domain transcription factors such as FUS3 and LEC2 with ABI3 in seeds (Mönke *et al.*, 2012). Expression of nine out of eleven genes coding for group A LEA proteins was overexpressed *MtABI3* (Table 2.3). Moreover, *in silico* promoter analysis of eight out of twelve LEA genes for which promoter sequences could be retrieved indicated that seven LEA promoters contain both RY (CATGCA) and ABRE (ACGTG(G/T)C) cis-regulatory elements. The RY element is known to be bound by ABI3, while ABRE motifs are implicated in the binding of bZIP-TFs, known to interact with ABI3 (Busk *et al.*, 1997; Hattori *et al.*, 2002; Guerriero *et al.*, 2009). A study on the identification of the ABI3 regulon in Arabidopsis confirmed three LEA genes as direct targets by transient promoter activation assay or ChIP-chip analysis (Table 2.3). The other LEA proteins were identified as being ABI3-responsive gene products in *35S::ABI3-GR* seedlings (Mönke *et al.*, 2012). An analysis of the four other LEA proteins that were abundant in *C. australe* cotyledons (group B, Table 2.3) demonstrated that overexpressing *MtABI3* in *M. truncatula* roots induced *DHN3* trancripts and slightly activated *Budcar5,* although transcript levels were found to increase in *abi3* mutants (Table 2.3). No effect was found on transcript levels of *CapLEA-1*, and *DHN-cognate* transcripts even decreased significantly. No RY element was retrieved in promoter analysis of *BudCar5* and *DHN-cognate* coding genes, and neither gene was part of the ABI3 regulon identified by Monke *et al.* (2012).

Discussion

The aim of this study was to compare the seed LEA proteome of two legume species exhibiting orthodox and recalcitrant storage behaviour to gain further insights into the panoply of these protective

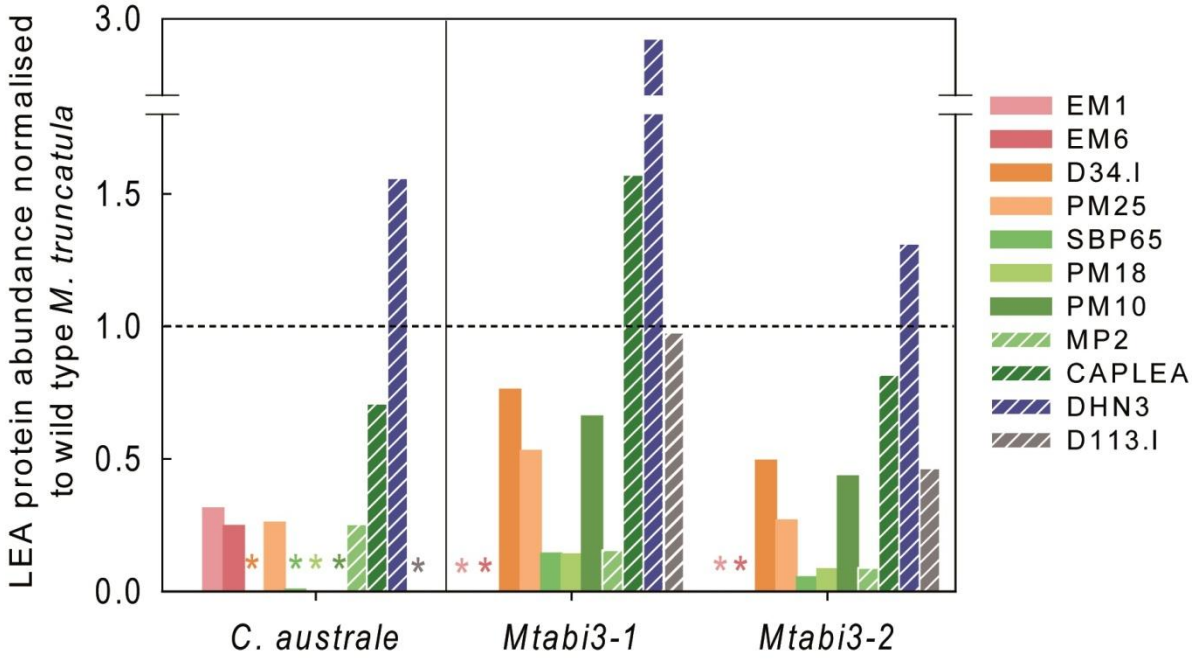


Figure 2.6. The LEA protein profile from desiccation-sensitive cotyledons of *Castanospermum australe* and seeds of *Mtabi3* mutants compared to desiccation-tolerant *Medicago truncatula* wild type seeds.
Abundance of LEA proteins in *C. australe* and *Mtabi3* (assessed as spot intensity, Annex 4 and Annex 5) was normalized against their respective value obtained for wild type *M. truncatula* seeds. A value of 1 corresponds to wild type values (*C. australe*/A17 and *Mtabi3*/R108). Hatched bars correspond to non-seed specific LEA proteins.

proteins necessary for desiccation tolerance. This work shows that *C. australe* and *M. truncatula*, both from the *Papilionaceae* subfamily of Fabaceae, are phylogenetically close enough to allow for a detailed sequence comparison between LEA accumulation in relation to desiccation tolerance. Assembly of a normalised sequencing library identified contigs with high similarity for 16 of the 17 *M. truncatula* LEA genes (Table 2.2) for which protein accumulation was shown in *M. truncatula* (Chatelain *et al.,* 2012). This comparison was further extended to a desiccation-sensitive *Mtabi3* mutant of *M. truncatula* that we obtained and characterized.

We believe this is the first report of full coverage of the identification of the LEA genes and their products (the 'LEAome') in cotyledons (the predominant tissue in this species) of a recalcitrant seed. To date, studies have been constrained to dehydrins using an antibody against the consensus sequence KIKEKLPG (Berjak & Pammenter, 2008). The comparison with *M. truncatula* revealed that 12 out of 16 LEA proteins are less abundant or not detected in the recalcitrant *C. australe* seed proteome (Figs. 2.4 and 2.6). *In silico* gene expression analysis of *M. truncatula* transcriptomes demonstrated that all but one (*MtMP2*) of these 12 genes are specifically expressed in seed tissues (Chatelain *et al.*, 2012). Further LEA proteome analysis of the *Mtabi3* mutants revealed that accumulation of the homologues of these LEA proteins was affected in these desiccation-sensitive seeds. Several of them (MtSBP65, MtPM25, MtEM6, MtPM18 and MtMP2) correlated with the re-induction of desiccation tolerance in germinated radicles of *M. truncatula* seeds (Boudet *et al.,* 2006). Fig. 2.5E and 5F shows that seeds of *Mtabi3* mutants can survive drying to 0.4 g H_2O/g DW and can be considered drought-tolerant. They lose their viability after they are shed from the mother plant. In *C. australe*, tissues did not survive drying down to 0.5 g H_2O/g DW. Collectively, these data suggest that these particular LEA proteins are needed once bulk water is removed. For most of them, their role in the dry state is not yet elucidated. *In vitro* studies of EM6, PM25 (Boudet *et al.,* 2006; Gilles *et al.*, 2007, Boucher *et al.*, 2010) and LEAm (Tolleter *et al.*, 2007) demonstrated multifunctional protective capacities with different efficiencies. These include membrane (LEAm) and enzyme protection (LEAm, EM6, PM25), anti-aggregation against thermo-mechanical stress (EM6 and PM25) and water-binding (EM6 and PM25).

This work offers a new model to study the regulatory and mechanistic pathways implicated in desiccation tolerance through comparative analysis of desiccation-sensitive cotyledon tissues from recalcitrant seeds with their orthodox counterparts. The proteome comparison with *Mtabi3* seeds suggests that comparable pathways leading to LEA accumulation are affected in both desiccation-sensitive orthodox and recalcitrant cotyledon tissues (Fig. 2.6). Consistent with this observation, homologous LEA genes in Arabidopsis and *M. truncatula* are ABI3-responsive (Table 2.3). It is unknown if the reduced LEA polypeptides in *C. australe* cotyledons are linked to reduced *CaABI3* activity or defective upstream or downstream signaling pathways. Interestingly, a *CaABI3* contig was detected in the RNA assembly, but its temporal and spatial expression, as well as its efficiency need to be assessed. In developing orthodox seeds, the RY cis-elements are elements that are crucial for transactivation through ABI3/VP1-like B3-domain proteins whereas conserved ABA responsive elements (ABREs; PyACGTGG/TC) mediate ABI3-related ABA signaling in conjunction with other transcription factors, such as ABI5 (Busk *et al.*, 1997; Hattori *et al.*, 2002; Guerrriero *et al.*, 2009). Most of the LEA promoters in Medicago for which protein abundance was decreased or absent in Castanospermum were found to contain both RY and ABRE elements (Table 2.3). The LEA genes for which protein abundance was not affected or even increased in *C. australe* compared to *M. truncatula* do not seem to be regulated by ABI3 (Table

Table 2.3. Evidence for ABI3-dependent regulation of LEA homologues for which protein abundance is reduced or absent in desiccation-sensitive tissues (*C. australe* and *Mtabi3*) (group A) or unaffected or increased (group B).

		Medicago truncatula					*Arabidopsis thaliana*		
protein name	LEA Group	IMGAG 3.5	Nimblegen probe	35S::ABI3 /control[1]	p-value	Cis-elements[2]	AGI	Transcript level in *abi3* seeds vs WT[3]	ABI3 targets[4]
EM6	A	Medtr4g016960	Medtr_v1_022627	3.87	1.92E-06	2 RY, 2 ABRE	AT2G40170	DOWN	-
EM1	A	AJ498523	Medtr_v1_072582	2.04	5.53E-02	ND	AT3G51810	DOWN	T,P
SBP65	A	Medtr4g079690	Medtr_v1_083614	4.18	3.07E-03	2 RY, 1 ABRE	AT2G42560	DOWN	T,P
PM10	A	Medtr8g134020	not present on slide			ND	AT5G44310	NA	T,P
LEAm	A	Medtr2g014040	Medtr_v1_009629	2.98	2.34E-03	1 RY, 2 ABRE	AT5G44310	NA	T,P
MP2	A	Medtr1g061730	Medtr_v1_005821	3.87	1.51E-05	0 RY, 2 ABRE	AT2G36640	DOWN	T,P
PM18	A	TC183861	Medtr_v1_076240	-0.02	8.65E-01	ND	AT2G36640	DOWN	T,P
PM1	A	Medtr7g093170	Medtr_v1_045826	4.59	7,89E-07	0 RY, 1 ABRE	AT5G06760	NA	C
D113.II	A	Medtr7g093160	Medtr_v1_045826	4.59	7,89E-07	2 RY, 1 ABRE	AT5G06760	NA	C
PM25	A	TC174777	Medtr_v1_082683	3.29	1.32E-02	ND	AT3G22490	DOWN	T,P,A
D-34.I	A	Medtr1g072090	Medtr_v1_006041	2.21	4.07E-02	5 RY, 0 ABRE	AT3G22490	DOWN	T,P,A
D-34.III	A	Medtr2g076230	Medtr_v1_012326	2.32	4.43E-02	0 RY, 0 ABRE	AT3G22490	DOWN	T,P,A
DHN3	B	TC175037	Medtr_v1_066754	3.32	1.10E-05	ND	ND	ND	ND
DHN-cognate	B	Medtr3g117290	Medtr_v1_020587	-1.47	4.91E-04	0 RY, 1 ABRE	AT1G76180	UP	-
BudCar5	B	Medtr7g086340	Medtr_v1_045277	1.57	3.70E-02	0 RY, 0 ABRE	ND	ND	ND
CapLEA-1	B	TC175990	Medtr_v1_085905	1.06	6.96E-02	ND	AT1G52690	UP	-

[1]Log ratio of transcript levels (and corresponding p-values) in hairy roots overexpressing *MtABI3* compared to control (empty plasmid), determined by trancriptome analysis using Nimblegen slides (GeOmnibus GSE44291).

[2]Number of RY (CATGCA) and ABRE (ACGTG(G/T)C) cis-regulatory motifs known to bind ABI3 was revealed by analyzing the 2kb promoter sequence of the *M. truncatula* genes.

[3]Relative level of transcripts in mature abi3 seeds of Arabidopsis compared to the wild type. Data are extracted from Bies-Etheve *et al.*, 2008.

[4]Identification of ABI3-responsive gene products in *35S::ABI3-GR* seedlings by array-based transcriptome analysis (T) or qRT-PCR (P) and confirmation as direct targets by transient promoter activation assay (A) or ChIP-chip analysis (C). Data are extracted from Mönke *et al.*, 2012.

ND, not detected; NA, not analyzed; WT, wild type.

2.3). Transcript levels of DHNcognate even increased in the *abi3* mutants of Arabidopsis and decreased in transgenic roots when overexpressing *MtABI3* (Table 2.3). Whether this gene is negatively regulated by ABI3 is unknown. Put together, these results strengthen the idea that only LEA proteins positively regulated by ABI3 are reduced in *C. australe* cotyledons. However, it is likely that additional regulatory pathways intervene in the accumulation of these desiccation-tolerant associated proteins because both in *C. australe* and the *Mtabi3* mutants, a number of LEA proteins were not absent but their levels were partially reduced. Other transcription factors that regulate LEA gene expression are *ABI4* and *ABI5* (Bies-Etheve *et al.*, 2008; Reeves *et al.*, 2011). ABI3 interacts with ABI5 to regulate expression of downstream genes whereas ABI4 controls the induction of ABI5 (Bossi *et al.*, 2009; Cutler *et al.*, 2010). However, in Arabidopsis, mature seeds of *abi4* and *abi5* mutants are desiccation tolerant. In addition, loss-of-function *lec1* mutants of Arabidopsis produce seeds that lose their viability during desiccation or during the first few weeks after harvest (Meinke, 1992). However, an analysis of direct targets of LEC1 did not reveal any LEA genes (Bäumlein & Junker, 2012; Wang & Perry, 2013). Considering that homologues of *ABI3, LEC1, ABI5, FUS3* were detected in the *C. australe* RNAseq assembly, the role of these transcription factors in seed development warrants further investigation, particularly in relation to its recalcitrant behaviour.

Sequencing of the *C. australe* transcriptome was performed by high throughput sequencing 454 and Illumina technologies on a normalised library. Library normalization improves the proportion of low abundant sequences and maximizes transcriptome coverage (Zhulidov *et al.*, 2004). Both technologies have been extensively used in the past few years to sequence transcriptomes of non model species without reference genome (reviewed in Schliesky *et al.*, 2012). Table 2.1 confirms that hybrid *de novo* assembly combining both sequencing technologies improves transcriptome coverage as suggested by Wall *et al.* (2009) and Garg *et al.* (2011). More than 72 % of the 48,334 contigs could be annotated by this approach and 91% of this annotation is provided by *M. truncatula* specific databases. This approach also enabled the discovery of almost all LEA transcripts for comparison with *M. truncatula*. However, a quantitative transcriptome analysis will be needed to reveal to which extent LEA polypeptide abundance is regulated at the transcriptional and/or posttranscriptional level in *C. australe*. Furthermore, there are many other molecular protective mechanisms that could be missing in this recalcitrant species such as antioxidant defences, non-reducing sugars, HSPs and/or induction of cell wall modifications (reviewed in Berjak & Pammenter, 2008; Leprince & Buitink, 2010). The sequence assembly from the normalised library will enable the construction of microarrays to further investigate molecular aspects of desiccation sensitivity in recalcitrant seeds.

A striking observation was the highly increased amount of dehydrins in the recalcitrant seeds compared to *M. truncatula* (Fig. 2.6). Two of them (Budcar5, DHN3) have also been identified in desiccation-sensitive seedlings of *M. truncatula* submitted to osmotic stress (Boudet *et al.*, 2006). Furthermore, i*n silico* analysis using the Medicago gene atlas shows that these dehydrins are expressed in many different organs other than seeds in stressful conditions (Benedito *et al.*, 2008; Chatelain *et al.*, 2012). One can speculate on the functional role for such proteins in recalcitrant seeds (Berjak & Pammenter, 2008). Most recalcitrant seeds are spheroid, with large cotyledons surrounding the axis. The synthesis of dehydrins in cotyledons can protect the axis from the dehydration stress that they will undergo after shedding. Furthermore, due to their size, dehydration is likely to be slow and thus requirement for protection against only mild water deficit stress should be sufficient

for maintenance of seed viability as a whole in seeds shed into their natural environmental habitat. Dehydrins are known to increase tolerance to osmotic stress, demonstrated by the overexpression of dehydrin *Rab17* and *Rab28* in *A. thaliana* plants and maize plants, respectively (Figueras *et al.*, 2004 ; Amara *et al.*, 2013).

In conclusion, the comparative analysis of the LEA proteome profiles of two unrelated desiccation-sensitive tissues (cotyledons of *C. australe* and seeds of *Mtabi3*) with the orthodox *M. truncatula* indicates that the developmental program leading to desiccation tolerance involves the synthesis of a variety of seed-specific LEA proteins that have been poorly characterized so far and partially involves ABI3. This developmental program is intertwined with the synthesis of additional LEA proteins such as dehydrins as an apparent need to retain some tolerance against mild osmotic stress during maturation.

Supplementary data

Figure 6.3. Sequencing, assembly and annotation workflow of *C. australe* seed transcriptome (cf Material & Methods section).

Figure 3.3. GO annotation of the sequence assembly of the transcriptome of *C. australe* seeds (cf chapter III).

Annex 1. Primer sequences used for PCR (cf Annex section).

Annex 2. Overview of contigs from *C. australe* transcriptome matching LEA coding genes of *M. truncatula* (cf Annex section).

Annex 3. Summary of identified spots from the reference gel of the heat-soluble protein fraction of *C. australe* cotyledons (cf Annex section).

Annex 4. Normalized intensity of polypeptides of the heat-stable proteome of *Medicago truncatula* and *Castanospermum australe* seeds (cf Annex section).

Annex 5. Normalized intensity of polypeptides of the heat-stable proteome of *Medicago truncatula* R108 (wild type), *Mtabi3-1* and *Mtabi3-2* seeds (cf Annex section).

Funding

This work was supported by a grant from the Région des Pays-de-la-Loire, France (QUALISEM 2009–2013), the bilateral Partenariat Hubert Curien (PHC) program France-South-Africa (Grant number 25903RE to J.D., O.L. and J.B.) and post-doctoral EU FP7 Marie Curie Individual Fellowship (Grant number 252822 to M.H.).

Acknowledgements

We thank our South-African colleagues (Prof. P. Berjak, Prof N. Pammenter, Prof J Farrant) for help with collecting the material. The *abi3* mutants of *Medicago truncatula* utilized in this research project, which are jointly owned by the Centre National de la Recherche Scientifique, were obtained from The Samuel Roberts Noble Foundation, Inc. and were created through research funded, in part, by a grant from the National Science Foundation (NSF# 703285). No conflict of interest is declared.

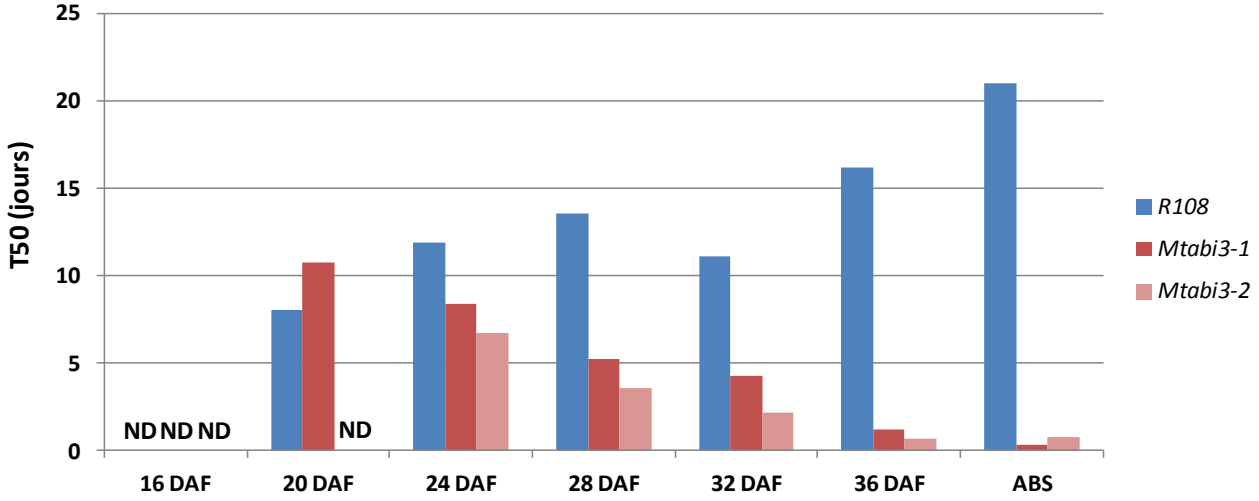


Figure S2.7. Caractérisation du niveau de dormance des graines de *M. truncatula* sauvages R108 et mutantes *Mtabi3-1*, *Mtabi3-2* en fonction du stade de développement.

Le pourcentage de germination de graines récoltées au cours de la maturation (16 DAF à ABS) et séchées quatre jours à température ambiante (DS) a été déterminé sur eau à 20°C et à l'obscurité. Les T50 correspondant à la durée nécessaire à la germination de 50 % des graines sont utilisés comme marqueur de dormance. Les valeurs présentées ici sont les moyennes de trois à cinq répétitions de 40 à 50 graines. ND = non déterminé car moins de 50 % des graines germent dans ces conditions.

Données supplémentaires à l'article : compléments sur l'étude du phénotype de *Mtabi3*

Etude de la dormance des graines *Mtabi3*

Chez *A. thaliana*, les graines des différents mutants *abi3* étudiés ne sont pas dormantes, à l'exception d'*abi3-1* pour lequel un certain degré de dormance (inférieur à celui mesuré pour des graines sauvages) est visible (Ooms *et al.*, 1993 ; Bies-Ethève *et al.*, 1999). Cet aspect du phénotype a également été étudié chez nos deux lignées de mutants *Mtabi3* et les résultats sont présentés dans la figure S2.7. Les graines *Mtabi3-1*, *Mtabi3-2* et R108 récoltées au cours de la maturation des graines ont été mises à germer sur eau sans prétraitement habituellement utilisé pour lever la dormance (stratification). La capacité germinative des graines R108 est acquise entre 16 DAF (18 % de germination) et 24 DAF (90 % de germination) (données non figurées). Elle est légèrement retardée chez *Mtabi3-1* et *Mtabi3-2* (respectivement 0 et 2 % à 16 DAF ; 73 et 70 % à 24 DAF ; 85 et 87 % à 28 DAF) (données non figurées). La durée nécessaire à la germination de 50 % des graines (T50) a été utilisée comme marqueur de dormance : plus cette valeur est importante, plus les graines sont dormantes. Cette valeur est comprise entre 11,1 et 13,6 jours pour les graines R108 entre 20 et 28 DAF. Elle augmente ensuite jusqu'à atteindre 21 jours à l'abscission (Figure S2.7), traduisant une mise en place de la dormance en fin de maturation pour les graines R108. En parallèle, la T50 diminue fortement au cours du développement des graines *Mtabi3-1* et *Mtabi3-2* jusqu'à atteindre moins d'une journée au stade abscission (Figure S2.7). Les premières graines *Mtabi3-1* et *Mtabi3-2* germent 14h après imbibition aux stades 36 DAF et abscission dans ces conditions. Ce profil est caractéristique de graines absolument non dormantes. La diminution de la T50 entre 20 et 36 DAF signifie une amélioration de la capacité germinative de ces graines en fin de maturation chez les mutants *Mtabi3*.

En conclusion, le phénotype décrit pour les mutants *Mtabi3* présente de nombreuses caractéristiques communes avec celui mis en évidence chez les mutants d'*A. thaliana* : absence de dégradation de la chlorophylle en fin de maturation, forte insensibilité à l'ABA, sensibilité à la dessiccation et absence de dormance.

Chapitre III

Analyse transcriptomique des graines immatures et matures de *C. australe* et comparaison avec le développement des graines de *M. truncatula*

Approche et principaux résultats

La maturation des graines de *M. truncatula* est bien caractérisée d'un point de vue transcriptomique (Verdier *et al.*, 2013) et des gènes potentiellement impliqués dans l'acquisition de la TD ont été mis en évidence (Buitink *et al.*, 2006). L'objectif est de confirmer ces résultats et de mettre en évidence de nouveaux gènes potentiellement impliqués dans la TD des graines en comparant l'expression des gènes au cours du développement d'une graine orthodoxe et d'une graine récalcitrante. Dans un premier temps, il est nécessaire d'obtenir pour les graines de *C. australe* des données transcriptomiques comparables à celles obtenues au cours du développement des graines de *M. truncatula*. Pour cela, deux banques d'ADNc de graines de *C. australe* ont été séquencées par les techniques haut-débit 454 et Illumina. Les transcrits potentiels obtenus après assemblage ont permis de créer une puce à ADN spécifique de *C. australe*. L'analyse comparative des données transcriptomiques avec *M. truncatula* a mis en évidence des résultats tout à fait nouveaux :

- Les régulateurs majeurs de la maturation des graines orthodoxes sont exprimés dans les graines immatures de *C. australe* comme chez *M. truncatula*. En revanche, l'expression du gène *ABI3* et de ses cibles chute fortement dans les graines matures de *C. australe* mais pas chez *M. truncatula*.
- Les gènes codant pour les protéines LEA (dont plusieurs sont des cibles d'ABI3) sont exprimés dans les graines immatures de *C. australe* mais pas dans les graines matures contrairement à *M. truncatula*. Ces résultats confortent la moindre accumulation de protéines LEA dans les graines matures de *C. australe* décrite dans le chapitre précédent. Cependant, la ressemblance entre les profils d'accumulation de protéines LEA entre les graines matures de *Mtabi3-2* et de *C. australe* ne se retrouve pas au niveau transcriptomique. Les gènes codant les LEA sont plus fortement exprimés dans les graines matures que dans les graines immatures de *Mtabi3-2*, contrairement à *C. australe*.
- Dans les graines matures de *C. australe*, plusieurs gènes impliqués dans la biosynthèse et la signalisation de l'ABA et dans la réponse au stress (antioxydants, chaperonnes) sont fortement exprimés alors que leur expression est faible chez *M. truncatula*.
- Les oligosaccharides de la famille de raffinose (RFO), accumulés très fortement dans des graines orthodoxes des Fabacéess, sont absents dans les graines de *C. australe*. En parallèle, les gènes codant pour des enzymes impliquées dans la biosynthèse des RFO, fortement exprimés chez *M. truncatula* en fin de maturation, sont très faiblement exprimés dans les graines matures de *C. australe*.
- L'analyse en composantes principales de l'expression des gènes préférentiellement exprimés dans la graine montre que le transcriptome d'une graine mature de *C. australe* ressemble plus au transcriptome d'une graine germée que d'une graine mature de *M. truncatula*. De plus, des gènes caractéristiques d'une graine prête à germer (répresseur de maturation, voie des GA) sont sur-exprimés dans les graines matures de *C. australe*.

Tableau 3.1. Caractéristiques des contigs issus de l'assemblage des données de séquençage 454, Illumina et hybride 454/Illumina du transcriptome de graines de *C. australe*.

Les lectures issus du séquençage Illumina ont été assemblés en contigs par GenXpro. Les lectures du séquençage 454 et l'ensemble des lectures 454/Illumina ont été assemblés de novo avec le programme MIRA puis avec DNA Dragon assembly. Seuls les contigs de l'assemblage hybride 454/Illumina de taille supérieure à 200 pb ont été conservés pour la suite de l'étude.

	454	Illumina	454 + Illumina
Nombre de lectures après séquençage	**626 225**	**7 784 004**	**8 410 229**
Nombre de bases séquencées	235 912 638	490 138 508	726 051 146
Nombre de contigs après le 1er assemblage (MIRA ou GenXpro)	**48 471**	**18 483**	**133 360**
Taille minimale	40	150	25
Taille maximale	10 450	3 185	13 889
Taille moyenne	771	318	418
N50	867	345	674
Nombre de contigs taille > N50	14 173 (29%)	5 216 (28%)	23 226 (17%)
Nombre total de bases consensus assemblées en contigs	37 379 064	5 882 730	55 802 891
Nombre de contigs après le 2ème assemblage (DNA Dragon assembly)	**37 977**	Voir ci-dessus car DNA Dragon n'a pu réassembler aucun des contigs formés par GenXpro	**48 334**
Taille minimale	41		200
Taille maximale	14 013		14 334
Taille moyenne	845		773
N50	997		1020
Nombre de contigs taille > N50	10 563 (28%)		11 392 (24%)
Nombre total de bases consensus assemblées en contigs	32 084 935		37 365 884
Nombre final de contigs > 200 pb	**36 767**	**12 876**	**48 334**
Taille maximale	14 013	3 185	14 334
Taille moyenne	869	382	773
N50	1001	401	1020
Nombre de contigs taille > N50	10 493 (29%)	3 898 (30%)	11 392 (24%)
Nombre total de bases consensus assemblées en contigs	31 937 738	4 914 157	37 365 884

1-Caractérisation du transcriptome de graines de *C. australe*

1.1-Séquençages et assemblage

La création d'une puce à ADN spécifique des graines de *C. australe* la plus complète possible repose sur l'identification du répertoire des gènes exprimés dans différents tissus de la graine, à différents stades de développement et sous différentes conditions. Les techniques actuelles de séquençage haut-débit permettent désormais d'atteindre cet objectif pour des espèces non modèles.

Dans un premier temps, des ARN de cotylédons et d'axes embryonnaires de *C. australe*, récoltés à deux stades de maturité différents en 2009, ont été extraits à partir de tissus frais, séchés un jour ou trois jours. Le séchage en présence de NaCl 75 % humidité relative induit un stress hydrique mais pas de dessiccation. Ce séchage des tissus avant extraction des ARN permet d'activer les gènes de réponse au stress et augmente ainsi la probabilité de séquencer les transcrits associés à ces gènes. De plus, la banque d'ADNc formée par le pool des huit échantillons a été normalisée avant séquençage par la technologie Illumina. La normalisation permet d'équilibrer la probabilité de séquençage entre des transcrits associés à des gènes fortement ou faiblement exprimés, en réduisant la proportion des transcrits très abondants au profit des transcrits plus rares. Ce premier séquençage a permis d'obtenir 7 784 004 lectures de 76 pb assemblés en 18 483 contigs par l'entreprise GenXpro GmbH (Francfort-sur-le-Main, Allemagne) (Tableau 3.1). Ces contigs sont en général de petite taille, comme le montre la taille moyenne de 318 pb et le N50 de 345 pb. Le N50 est équivalent à une médiane, il correspond à la taille du contig pour lequel 50 % des nucléotides de l'assemblage sont dans des contigs de taille supérieure et les 50 % restant sont dans des contigs de taille inférieure. Pour ce séquençage, 5 882 730 nucléotides ont été assemblés en contigs (somme cumulée de la taille de tous les contigs) soit 1,2 % du nombre total de nucléotides disponible dans les lectures. Deux possibilités peuvent expliquer cette faible valeur. Soit la proportion de lectures non utilisées pour l'assemblage est forte. Soit la profondeur de séquençage est importante, c'est-à-dire qu'un nombre important de lectures est associé à un même transcrit. Les données fournies par GenXpro ne nous permettent pas de trancher sur cette question mais nous avons estimé que ce premier séquençage n'était pas suffisant pour couvrir l'ensemble des gènes exprimés chez *C. australe*. De plus, la taille relativement faible des contigs pouvait constituer un obstacle à leur annotation par similarité.

Pour ces deux raisons et pour ajouter un stade plus précoce, un deuxième séquençage a été réalisé à partir d'une deuxième banque d'ADNc normalisée. Celle-ci a été formée à partir d'un pool d'ARN extraits de cotylédons et d'axes embryonnaires de graines récoltées en 2011 sur les mêmes arbres qu'en 2009. Pour cette banque, un stade plus précoce, appelé immature (ou gousse verte), a été ajouté aux stades mi-mature (gousse jaune) et mature (gousse marron). Pour ce deuxième séquençage, la technologie 454 a été privilégiée car elle très complémentaire de la technologie Illumina. En effet, les lectures obtenues sont moins nombreuses (seulement 626 225) mais plus longues (376 pb en moyenne) (Tableau 3.1). Elles présentent donc de plus grandes régions chevauchantes entre elles et facilitent l'assemblage en contigs comme le montre la taille moyenne des 48 471 contigs (771 pb) et le N50 de 867 pb. Cette fois, 37 379 064 nucléotides ont été

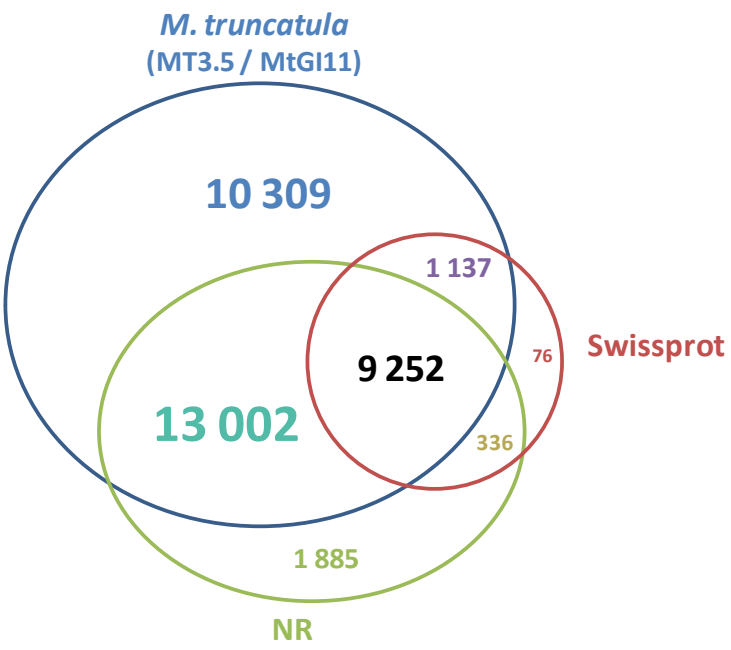


Figure 3.1. Nombre de contigs décrits par blast contre les différentes bases de données généralistes ou spécifiques de *M. truncatula*.
35 997 contigs sont décrits par blast contre au moins une des quatre bases de données utilisées pour l'analyse. Le schéma présenté ici montre le nombre de contigs décrits par au moins une des bases de données spécifiques de *M. truncatula* : MT3.5 et MtGI11 (cercle bleu) et par les bases de données plus générales : Swissprot (cercle rouge) et NR (cercle vert). La taille de chaque zone du diagramme dépend du nombre de contigs qu'elles représentent et la couleur du chiffre est panachée à partir de celle des bases de données qui se recoupent.

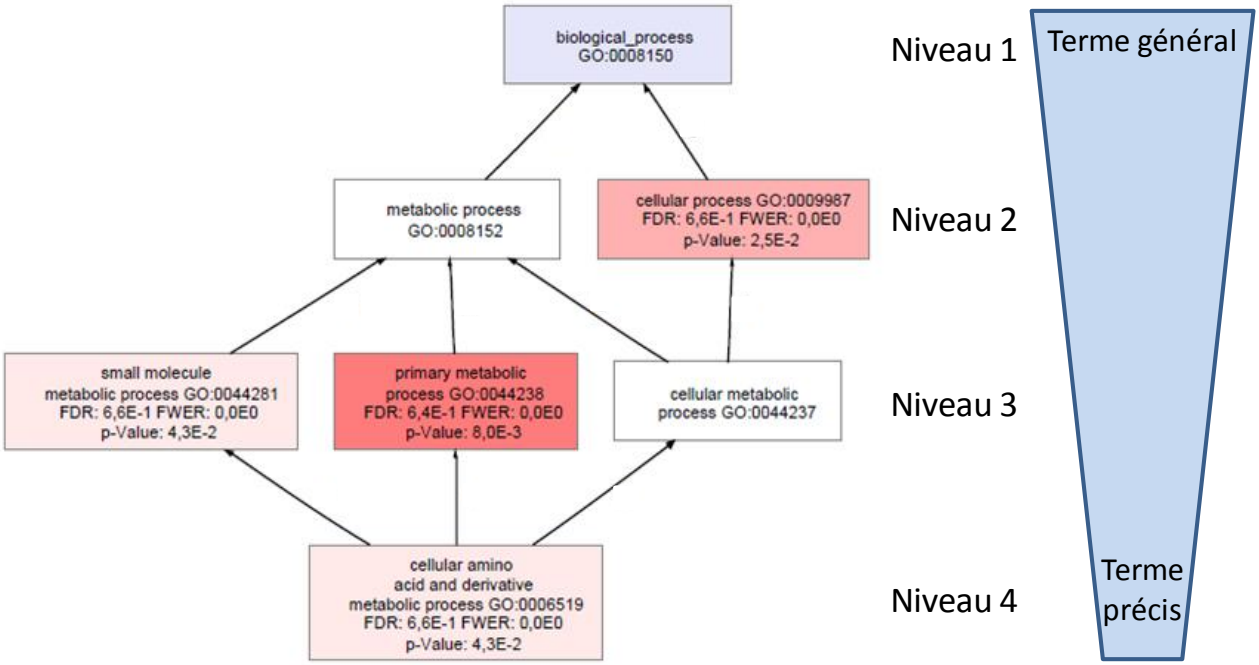


Figure 3.2. Organisation en graphique acyclique dirigé (DAG) de la classification Gene Ontology (GO).
Les termes GO sont définis selon plusieurs niveaux hiérarchiques du plus général (terme « parent ») au plus précis (terme « fils ») mais, contrairement à une classification hiérarchique classique, un terme « fils » peut avoir plusieurs parents. Dans cet exemple, le terme 'métabolisme cellulaire des acides aminés et dérivés' appartient à trois termes « parents » plus généraux : 'métabolisme des petites molécules', 'métabolisme primaire', 'métabolisme cellulaire'.

assemblés en contigs soit 15,8 % du nombre total de nucléotides obtenus après séquençage. La profondeur de séquençage est donc assez faible mais le nombre plus important de contigs obtenus suggère une meilleure couverture du transcriptome de graines de *C. australe* et la taille importante des contigs facilite leur annotation par similarité.

1.2-Assemblage des lectures des deux séquençages

La complémentarité des technologies 454 et Illumina nous a incité à assembler les lectures issues des deux séquençages pour profiter des avantages de chacune. Les lectures assez longues du 454 permettent de créer un squelette de contigs tandis que les lectures Illumina apportent de la profondeur pour l'assemblage. L'assemblage *de novo* de ces lectures a été réalisé à l'aide du programme MIRA mais le résultat obtenu n'était pas entièrement satisfaisant. En effet, un très grand nombre de contigs a été obtenu (133 360) (Tableau 3.1) mais la taille moyenne de 418 pb et le N50 de 674 pb sont inférieurs aux résultats obtenus avec le séquençage 454 seul. L'assemblage hybride *de novo* a donc généré un grand nombre de petits contigs non désirables pour le processus d'annotation. Le nombre de contigs hybrides (133 360) étant largement supérieur à la somme des contigs obtenus pour chacun des séquençages (48 471 + 18 483), il est fort probable qu'une redondance subsiste et que certains de ces contigs hybrides puissent être assemblés entre eux. Un deuxième assemblage moins stringent a donc été réalisé à partir des 71 239 contigs de plus de 200 pb à l'aide de l'assembleur DNA Dragon.

Alors que ce deuxième assemblage a eu un effet limité sur les contigs spécifiques 454 et Illumina, il a permis de réassembler un grand nombre de contigs hybrides. En effet, le nombre de contigs finaux a été abaissé à 48 334 et la taille moyenne et le N50 atteignent désormais respectivement 773 pb et 1020 pb (Tableau 3.1). Ces valeurs de taille sont proches de celles obtenues en assemblant uniquement les lectures 454 (869 pb pour la taille moyenne et 1001 pb pour le N50) mais le nombre de contigs est plus élevé.

1.3-Annotation des contigs

1.3.1-Description par similarité : blast

Les 48 334 contigs ont été comparés à quatre bases de données pour les annoter par similarité. Le processus d'annotation a débuté par un blast contre des bases de données spécifiques de *M. truncatula* (MT3.5 et MtGI11) puis contre des bases plus larges incluant des séquences de toutes les espèces végétales (Swissprot et NR). Une annotation par au moins une de ces bases de données a été obtenue pour 35 997 contigs (74,5 %). Un total de 91 % de ces contigs a reçu une annotation spécifique de *M. truncatula* : 33 141 par la base de données de séquences génomiques MT3.5 et 159 supplémentaires par la banque d'EST MtGI11. Respectivement 412 et 1 885 contigs supplémentaires ont été décrits par les bases de données protéiques Swissprot et NR. La figure 3.1 montre une redondance importante entre les annotations spécifiques de *M. truncatula* et celles des bases NR ou Swissprot. Cependant 10 309 contigs sont uniquement décrits par une base spécifique de *M. truncatula*, résultat vraisemblablement expliqué par la forte proximité génétique entre *C. australe* et *M. truncatula*.

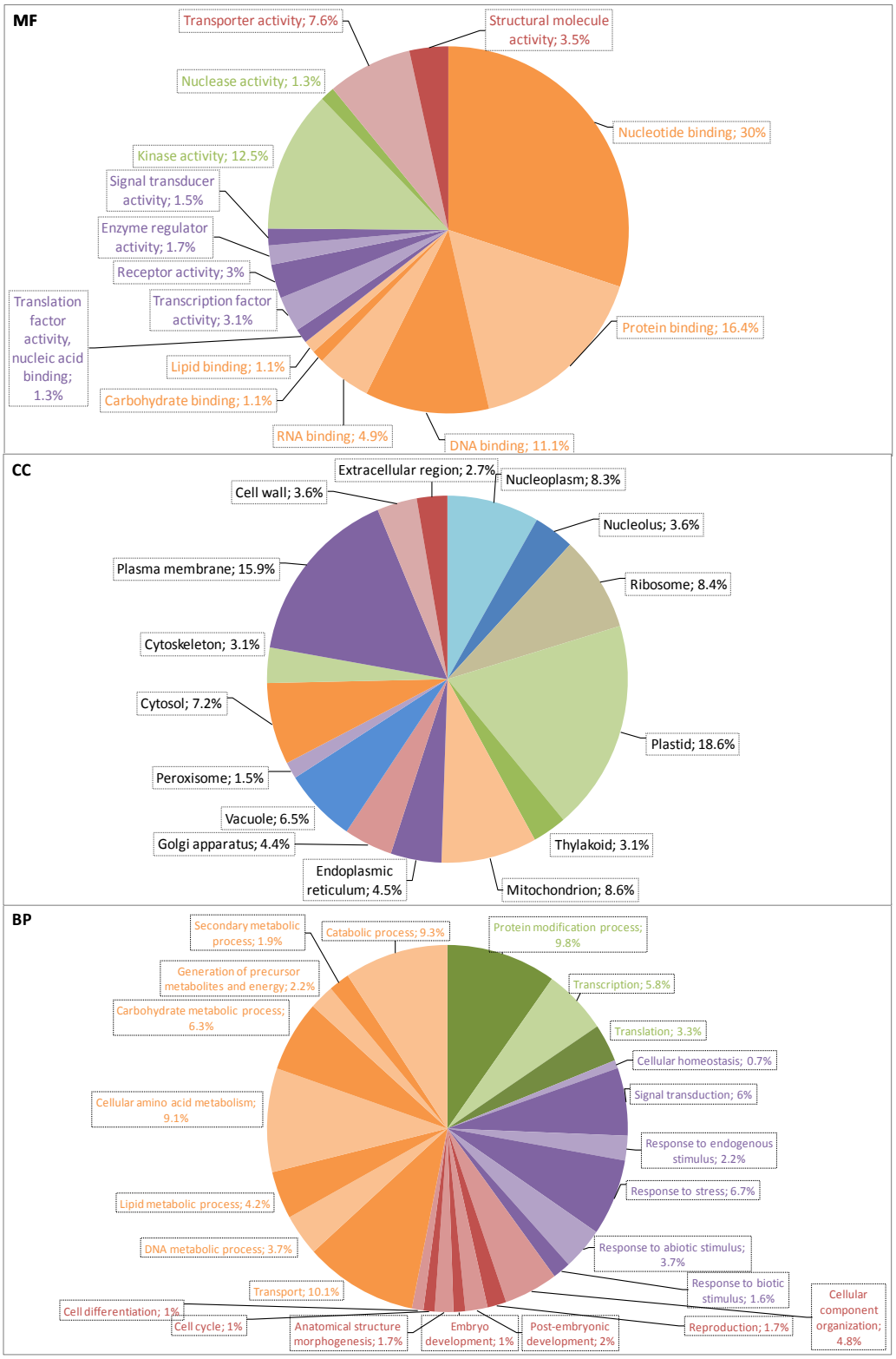


Figure 3.3. Annotation Gene Ontology (GO) du transcriptome de graines de *C. australe*.
Seuls les termes GO-slim les plus spécifiques (termes fils) représentant les fonctions moléculaires (MF), les compartiments cellulaires (CC) et les processus biologiques (BP) associés à plus de 120 contigs sont représentés. Pour les BP, les termes GO sont regroupés en catégories selon leur implication biologique : métabolisme (orange), développement (rouge), réponse aux signaux (violet) et expression des gènes (vert). Pour les MF, les termes GO sont regroupés selon leur activité : fixation à des molécules (orange), transduction du signal (violet), activité catalytique (vert).

1.3.2-Annotation fonctionnelle par Gene Ontology

Ces descriptions basées sur les résultats de blast ont été complétées par une annotation fonctionnelle. Ce type d'annotation est attribué à l'aide d'ontologies et permet de regrouper les gènes présentant des caractéristiques fonctionnelles proches en catégories plus larges. Dans cet objectif, la classification Gene Ontology (GO) a été mise en place par un consortium international pour décrire les produits de gènes selon leurs fonctions moléculaires (MF), les processus biologiques (BP) dans lesquels ils sont impliqués et les compartiments cellulaires (CC) dans lesquels ils sont exprimés (Ashburner *et al.*, 2000). Actuellement, la classification GO compte plus de 30 000 termes différents définissant des catégories très larges (par exemple GO:0008152 - métabolisme) ou très précises (par exemple GO:0000024 - biosynthèse du maltose). Ces trois ontologies parallèles sont de type acyclique, ce qui les différencie d'autres classifications dites hiérarchiques comme les catégories BIN définies par Mapman. Quelque soit le type de classification (acyclique ou hiérarchique), les termes sont définis selon plusieurs niveaux du plus général (terme « parent ») au plus précis (terme « fils »). Cependant, un terme « fils » peut avoir plusieurs parents dans la nomenclature acyclique GO (Figure 3.2), ce qui n'est pas le cas pour la nomenclature hiérarchique Mapman. La classification acyclique rend mieux compte de la complexité des interactions entre processus biologiques comparé à une classification hiérarchique.

Les termes GO des ontologies BP, MF et CC ont été associés aux contigs de *C. australe* à l'aide du logiciel Blast2GO (Conesa *et al.*, 2005 ; Götz *et al.*, 2008). Dans un premier temps, les descriptions de blast contre la base de données NR ont été utilisées pour retrouver les termes GO associés en interrogeant plusieurs bases de données (étape de mapping). Dans un deuxième temps, seuls les termes GO les plus pertinents ont été conservés pour chaque contig (étape d'annotation). Cette sélection suit une règle d'annotation modulable mais les paramètres par défaut ont été conservés pour notre analyse. La procédure de base a été complétée par l'utilisation des outils supplémentaires de Blast2GO : recherche de motifs IPS par InterProScan, attribution des codes d'enzymes selon la nomenclature internationale EC (Enzyme Commission), augmentation de l'annotation par l'outil Annex et simplification de la classification par l'outil 'plant GOslim'. Il s'agit d'une version simplifiée de l'annotation GO dans laquelle plusieurs termes spécifiques sont regroupés dans un terme plus générique. Par exemple la version GOslim adaptée aux plantes regroupe les 20 437 termes GO BP en seulement 47 catégories représentatives. Cet outil permet d'avoir une première caractérisation générale du transcriptome de *C. australe*.

Au final, 23 637 contigs (48,9 % du total) ont été annotés avec 98 615 termes et 8 962 identifiants EC. La distribution des principaux termes BP, MF et CC après GOslim est représentée sur la figure 3.3. Pour plus de clarté, seuls les termes les plus spécifiques associés à un minimum de 120 contigs sont figurés sur chacun des trois diagrammes. Ainsi, ces trois diagrammes représentent 17 349 annotations MF, 9 122 annotations CC et 25 529 annotations BP. Les fonctions moléculaires majoritaires sont la liaison aux nucléotides (30 %), aux protéines (16,4 %) et à l'ADN (11,1 %) ainsi que l'activité kinase (12,5 %). Les plastes (18,6 %) et la membrane plasmique (15,9 %) sont les composants cellulaires les plus représentés dans l'annotation des transcrits de graines de *C. australe*. Enfin l'ensemble des processus biologiques liés au métabolisme (termes orange) au développement (termes rouge), à la réponse aux signaux (termes violet) et à l'expression des gènes (termes vert) représentent respectivement 46,8 %, 13,3 %, 20,9 % et 18,9 % des annotations BP.

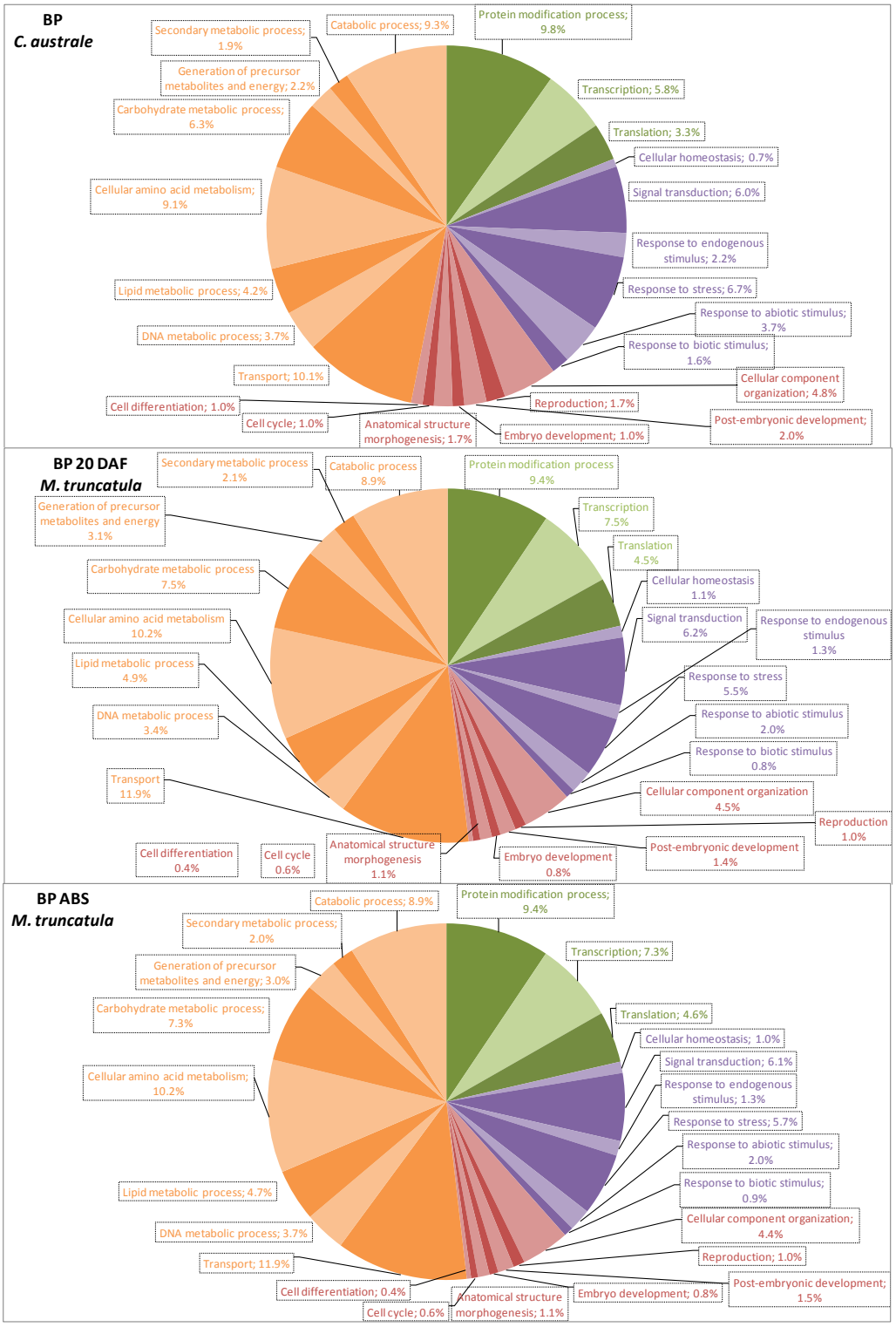


Figure 3.4. Comparaison des processus biologiques associés aux gènes exprimés dans les graines de *C. australe* et *M. truncatula*.

Deux stades de développement sont présentés pour *M. truncatula* : 20 DAF et abscission (ABS). Seuls les termes GO-slim les plus spécifiques (termes fils) représentant les processus biologiques (BP) associés à plus de 120 contigs sont représentés. Les termes GO sont regroupés en catégories selon leur implication biologique : métabolisme (orange), développement (rouge), réponse aux signaux (violet) et expression des gènes (vert).

Ces proportions ont été comparées à celles obtenues pour l'ensemble des gènes exprimés dans les graines de *M. truncatula* au stade 20 DAF et au stade abscission (ABS) (Figure 3.4). A partir de données transcriptomiques récemment publiées (Verdier *et al.*, 2013) issues de la puce Affymetrix, respectivement 32 655 et 33 641 gènes ont été déterminés comme exprimés aux stades 20 DAF et ABS sur 48 116 sondes (cf matériel et méthodes). 88 % des gènes exprimés au stade 20 DAF restent exprimés au stade ABS, confirmant que le développement des graines dépend du niveau d'expression des gènes plus que de la nature des gènes exprimés à différents stades. Par conséquent, les proportions des termes BP associés aux gènes exprimés pour ces deux stades de développement de la graine de *M. truncatula* sont très proches (Figure 3.4). Il est intéressant de constater que les différences de proportion entre *C. australe* et *M. truncatula* sont peu importantes. Ce résultat paraît logique car le répertoire des processus biologiques exprimés entre différentes graines ne peut pas être très différent. Il est donc plutôt rassurant de constater que notre répertoire de contigs représentant le transcriptome de graine de *C. australe* ne semble pas biaisé par rapport à celui de *M. truncatula*. On note cependant une légère sous représentation de tous les BP liés au métabolisme dans le transcriptome de *C. australe* par rapport aux gènes exprimés dans la graine de *M. truncatula* (Figure 3.4). En effet, les BP liés au métabolisme représentent 51,9 % des annotations pour *M. truncatula* contre 46,8 % chez *C. australe*. Cependant, les BP liés à la réponse aux stress et aux stimuli biotiques et abiotiques sont sur représentés chez *C. australe* (12 %) par rapport à *M. truncatula* (8,3 %).

C'est donc vraisemblablement le niveau d'expression des gènes qui traduit les différences entre le transcriptome des graines de *C. australe* et de *M. truncatula*. Des données quantitatives sont donc nécessaires.

1.4-Construction de la puce spécifique des graines de *C. australe*

Afin d'obtenir des données quantitatives sur l'expression des gènes au cours du développement des graines de *C. australe*, une puce à ADN spécifique de ces graines a été élaborée à partir du répertoire de contigs obtenu par séquençage haut-débit et décrit précédemment. Une nouvelle puce Nimblegen 102 k (102 123 sondes) spécifique de *M. truncatula* est désormais disponible au laboratoire. La même plate forme Nimblegen a donc été utilisée pour la puce spécifique de *C. australe* afin de faciliter la comparaison des données entre les deux espèces. A partir des séquences des 48 334 contigs, 48 125 sondes spécifiques de 60 nucléotides représentant 48 139 contigs ont pu être dessinées par NimbleGen. En effet, 14 de ces sondes reconnaissent deux contigs et ne sont pas totalement spécifiques. Les lames commandées pour cette étude accueillent 12 puces d'une capacité maximale de 135 k. Nous avons profité de la place disponible pour déposer sur chaque puce les 48 125 sondes en sens et en antisens, soit 96 250 sondes (Figure 3.5). L'ajout de ces sondes antisens sur la puce présente deux intérêts :

- Avec les techniques de séquençage utilisées, le contig obtenu peut représenter le transcrit (cas 1) ou son inverse complément (cas 2). L'expression du gène, correspondant à l'accumulation du transcrit, est quantifiée par la sonde 'sens' dans le cas 1 et par la sonde 'antisens' dans le cas 2.
- L'expression des antisens des gènes (sonde 'antisens' pour le cas 1 et sonde 'sens' pour le cas 2) peut se révéler intéressante. En effet, il est désormais admis que les antisens jouent un rôle majeur dans la

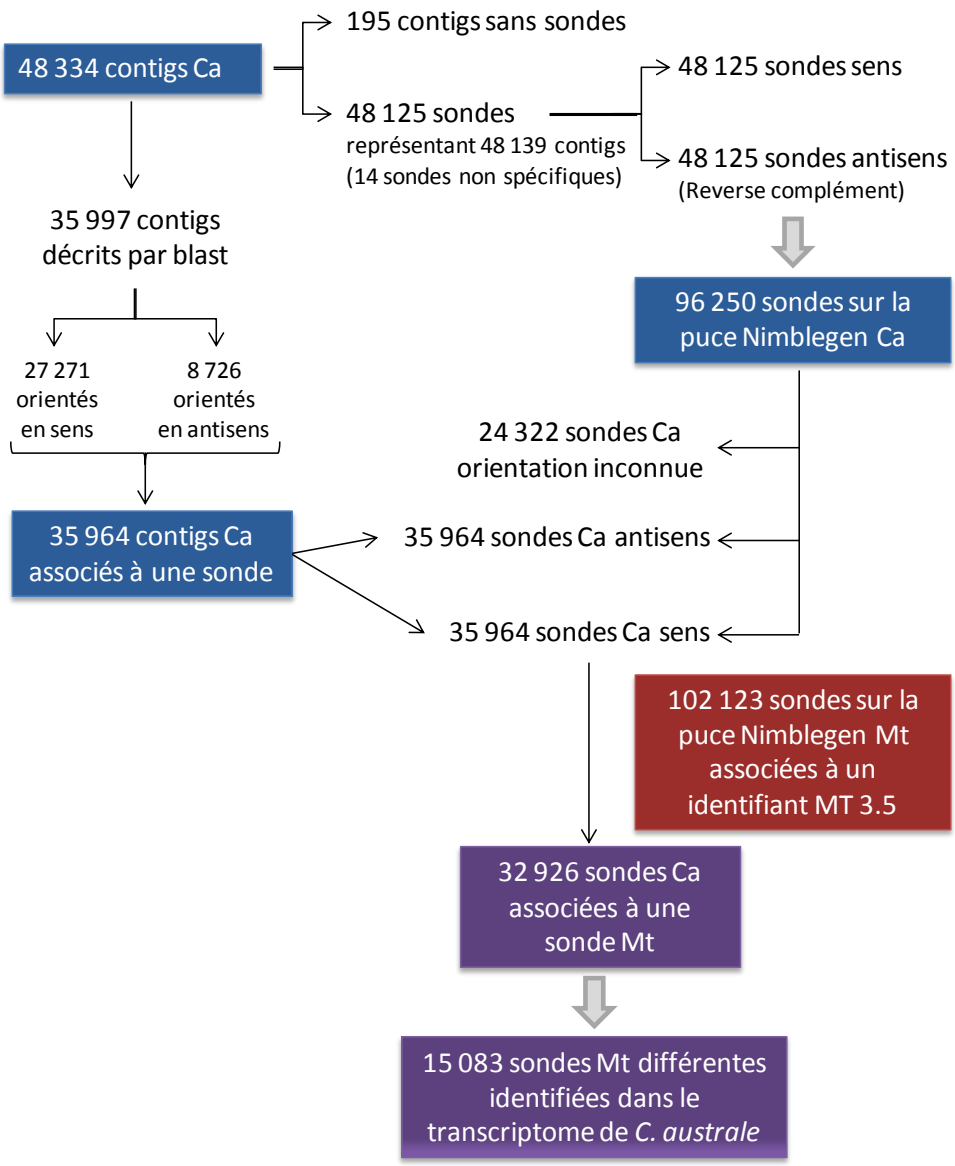


Figure 3.5. Association entre les sondes des puces Nimblegen spécifiques de *C. australe* et de *M. truncatula*.
A partir des 48 334 contigs obtenus après assemblage, 48 125 sondes spécifiques (sauf 14 d'entre elles qui reconnaissent deux contigs) ont pu être dessinées par Nimblegen. Ces sondes ainsi que leur reverse complément ont été placés sur la puce Nimblegen spécifique de *C. australe* (Ca), soit 96 250 sondes présentes sur chaque puce. 35 997 contigs (74% des contigs totaux) ont pu être décrits par au moins un des blast utilisé pour l'annotation ce qui permet de déterminer si ces contigs correspondent directement au transcrit (76% des cas) ou au reverse complément du transcrit (24% des cas). 35 964 de ces contigs sont associés à une sonde sens et une sonde antisens de la puce *C. australe*, ce qui permet à la fois de quantifier l'expression du gène et du transcrit antisens potentiel. En parallèle, la puce spécifique de *M. truncatula* (Mt) comporte 102 123 sondes associées à autant d'identifiants MT3.5. Il est donc possible d'associer 32 926 sondes Ca à une sonde Mt via l'annotation MT3.5 des contigs Ca. Au total 15 083 sondes Mt différentes, soit 15 083 gènes, ont été identifiées parmi ces 32 926 associations. Les cadres en bleu représentent Les données spécifiques à *C. australe*, celles spécifiques à *M. truncatula* sont représentées par un cadre rouge et les données communes à *C. australe* et *M. truncatula* sont représentées par des cadres violets.

répression de l'expression de nombreux gènes (Hamilton & Baulcombe, 1999 ; Vaucheret, 2006). Ils n'ont cependant pas été étudiés dans le cadre de cette thèse.

Seulement 195 contigs (0,4 %) ne sont pas représentés sur la puce ce qui signifie que l'expression de presque tous les gènes, et des antisens associés, exprimés dans les graines de *C. australe* est quantifiée par cet outil. Cependant, pour tirer partie de ces données d'expression il faut connaitre l'orientation du contig, c'est à dire savoir pour chaque contig si nous avons affaire au cas 1 ou au cas 2. Le seul moyen de déterminer l'orientation du contig est d'utiliser les alignements de blast. Si un contig a été aligné dans le même sens qu'un gène connu dans les bases de données, nous sommes dans le cas 1 où le contig représente directement le transcrit. Au contraire, un contig aligné en sens inverse avec un gène connu dans les bases de données signifie que le contig correspond à l'inverse complément du transcrit, nous sommes donc dans le cas 2.

D'après les résultats d'annotation par similarité présentés précédemment, 35 997 contigs ont pu être décrits par blast, 27 271 (76 %) dans le cas 1 et 8 726 (24 %) dans le cas 2. 35 964 d'entre eux sont associés à une sonde sur la puce ce qui signifie que les 96 250 sondes de la puce peuvent être réparties en trois catégories : 35 964 sondes 'sens' permettant de quantifier l'expression du gène, 35 964 sondes 'antisens' permettant de quantifier l'expression éventuelle de l'antisens du gène et 24 322 sondes non exploitables à l'heure actuelle car l'orientation du contig associée n'est pas connue (Figure 3.5).

2-Comparaison transcriptomique des graines de *C. australe* et de *M. truncatula*

2.1-Association des données transcriptomiques des deux espèces

Les 102 123 sondes de la puce Nimblegen spécifique de *M. truncatula* sont associées à un identifiant MT3.5. Cette annotation commune aux puces des deux espèces permet de relier les sondes de la puce spécifique *C. australe* (Ca) aux sondes de la puce spécifique *M. truncatula* (Mt). Ainsi, sur les 35 964 sondes Ca exploitables, 32 926 (92 %) ont pu être reliées à une sonde Mt via l'identifiant MT3.5. Dans certains cas plusieurs sondes Ca sont associées à une même sonde Mt. Ainsi 15 083 sondes Mt différentes sont liées à la puce spécifique de *C. australe* (Figure 3.5). Ces 15 083 gènes potentiels ont servi de base à la comparaison des données transcriptomiques des deux espèces en se basant sur l'intensité (en log) du signal après soustraction du bruit de fond pour chacune de ces sondes. Lorsqu'une seule sonde Ca est associée à une sonde Mt (43 % des 15 083 sondes) les intensités d'expression des gènes sont directement comparables entre les deux espèces. Lorsque plusieurs sondes Ca sont associées à une même sonde Mt, la moyenne des intensités de signal des sondes Ca a été calculée et cette valeur a été comparée avec l'intensité du signal de la sonde Mt. Ce calcul d'intensité moyenne n'est pas idéal d'un point de vue statistique mais constitue la seule façon de comparer l'expression de ces gènes entre les deux espèces. Dans la majorité des cas, seulement deux ou trois sondes Ca sont associées à une sonde Mt (respectivement 27 % et 15 % des 15 083 sondes). De plus, toutes les sondes Ca associées à une même sonde Mt suivent bien souvent le même profil d'intensité. La valeur moyenne d'intensité reflète donc bien l'expression du gène associé à toutes ces sondes. A titre d'exemple, l'association

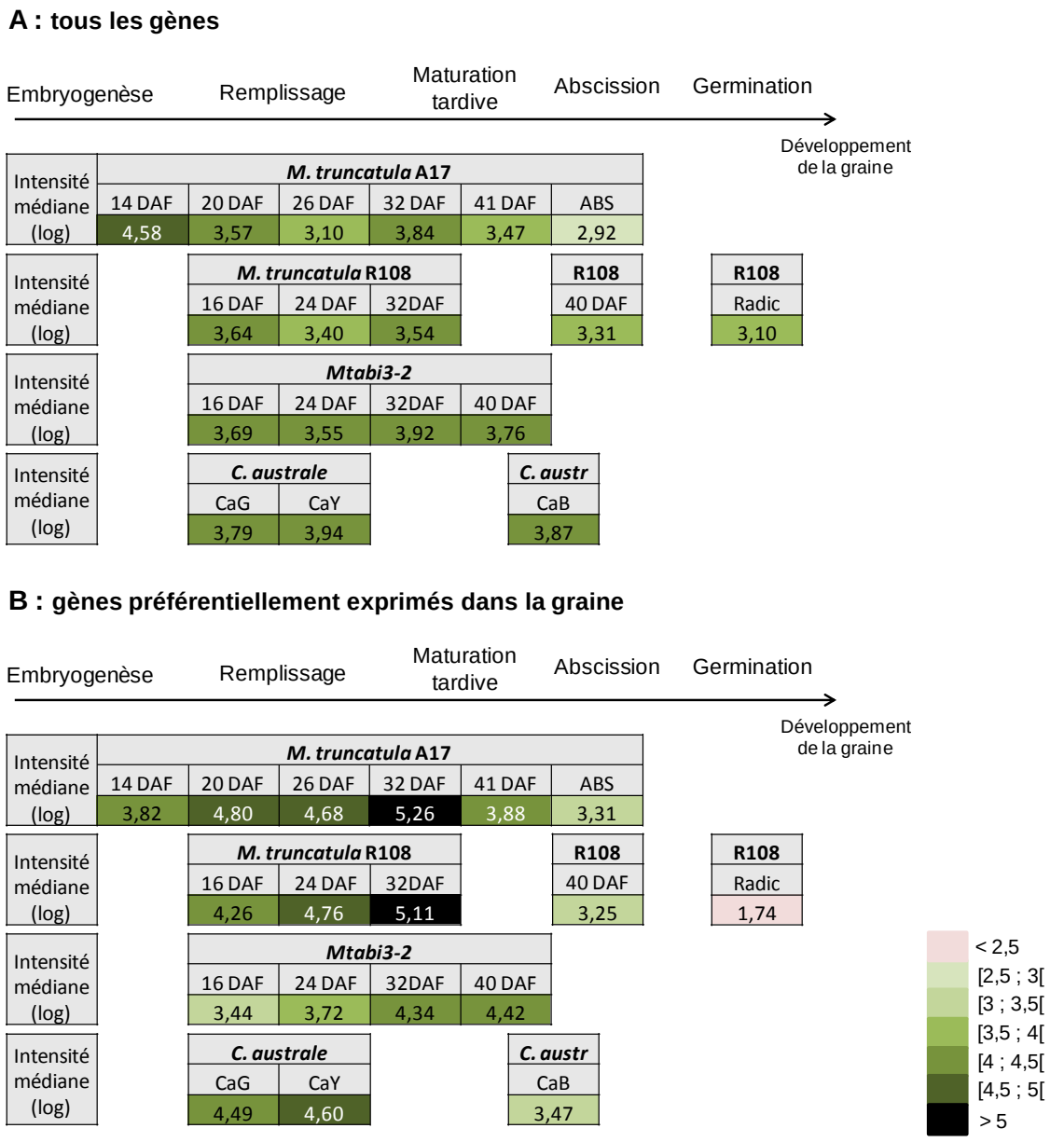


Figure 3.6. Evolution de la valeur de l'intensité médiane d'expression des gènes au cours du développement des graines de *C. australe*, *M. truncatula* A17, R108 et *Mtabi3-2*.

La valeur médiane a été calculée à partir des intensités d'expression des 15 083 gènes différents de *M. truncatula* qui ont pu être associés à *C. australe* **(A)** ou à partir des intensités d'expression des 273 gènes préférentiellement exprimés dans les graines **(B)**. Un gradient de couleur permet de visualiser les variations de cette valeur selon l'échantillon considéré. Les différents stades de développement de chacune des graines ont été placés le long d'un axe de développement de la graine allant de l'embryogenèse à la germination. Pour *C. australe*, les graines des deux premiers stades sont en phase de remplissage alors qu'au dernier stade les graines sont presque à l'abscission. DAF = jours après floraison ; ABS = abscission.

entre les gènes de *M. truncatula* et de *C. australe* codant pour les régulateurs majeurs de la maturation des graines est présentée dans l'annexe 6.

L'association des deux puces permet donc de comparer l'expression de 15 083 gènes potentiels entre tout échantillon de graines de *M. truncatula* et de *C. australe*. Nous avons à notre disposition au sein de l'équipe quatre jeux de données transcriptomiques chez *M. truncatula* :

- ➔ Le développement des graines génotype A17 en condition optimale de 20°C, avec 14 stades de la fin de l'embryogenèse (8 jours après floraison [DAF]) à la graine sèche.
- ➔ Le développement des graines mutantes *Mtabi3-2* en condition optimale à 20°C, avec quatre stades (16 DAF, 24 DAF, 32 DAF et 40 DAF).
- ➔ Le développement des graines sauvages R108 (fond génétique de *Mtabi3-2*), avec les quatre mêmes stades (16 DAF, 24 DAF, 32 DAF et 40 DAF).
- ➔ Le stade germination des graines R108 correspondant à la percée des téguments (radicule 2,7 mm).

Concernant les échantillons de graines de *C. australe*, les ARN issus des trois stades de développement de graines récoltés en 2011 et utilisés pour le séquençage 454 ont été hybridés sur la puce Castanospermum. La couleur de la gousse a servi de critère pour séparer ces trois stades qui peuvent être définis comme immature (gousse verte : CaG), mi-mature (gousse jaune : CaY) et mature (gousse marron : CaB). Une caractérisation physiologique plus précise de ces stades est en cours en Afrique du Sud. Pour une raison inconnue, l'extraction d'ARN à partir d'axes embryonnaires pose problème : les rendements d'extraction sont bons mais les tests d'intégrité des ARN sur gel et sur Experion révèlent une forte dégradation. Nous avons donc concentré notre analyse sur les cotylédons de graines de ces trois stades.

2.2-Comparaison globale des différents stades de développement des graines de *C. australe* et *M. truncatula*

2.2.1-Les intensités d'expression des gènes sont-elles globalement comparables entre *C. australe* et *M. truncatula* ?

L'objectif est de comparer les intensités d'expression des gènes entre les deux espèces pour tenter de rapprocher chaque stade de développement de *C. australe* d'un stade caractérisé chez *M. truncatula*. Comme les données transcriptomiques des deux espèces ont été obtenues sur des lames différentes et à partir de plans d'hybridation différents, au préalable, il est nécessaire de s'assurer que ces intensités sont globalement comparables entre *C. australe* et *M. truncatula* afin d'éviter des erreurs d'interprétation. La valeur médiane des intensités associées aux 15 083 sondes a donc été calculée pour les différents échantillons utilisés dans notre analyse (Figure 3.6A). Elle est du même ordre de grandeur pour presque tous les échantillons (log intensité compris entre 3,5 et 4). Elle est cependant plus élevée pour le stade 14 DAF (4,58) et plus basse pour le stade ABS (2,92) au cours du développement des graines A17. Un problème de normalisation n'est pas forcément à l'origine de ces différences. En effet, l'activité transcriptionnelle générale est très importante en fin d'embryogenèse/début de remplissage alors qu'elle diminue en fin de développement dans les graines d'*A. thaliana* (Le *et al.*, 2010). L'intensité médiane sur l'ensemble des gènes est donc logiquement plus élevée en

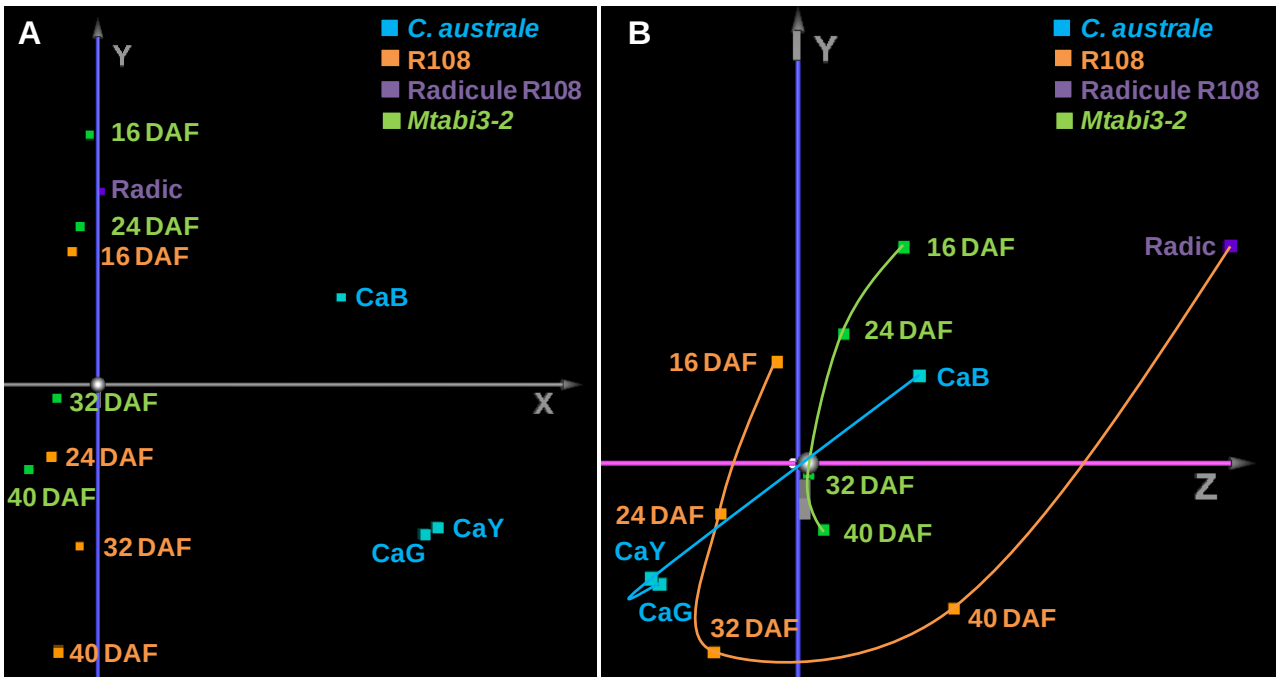


Figure 3.7. Analyse en composantes principales (ACP) de l'expression des gènes préférentiellement exprimés dans les graines de *C. australe*, *M. truncatula* R108 et *Mtabi3-2*.

Basé sur les données d'expression dans différents tissus de *M. truncatula*, 991 sondes (sur 102 123) ont été identifiées comme associées à des gènes préférentiellement exprimés dans les graines. Parmi ces sondes, 273 correspondent à des gènes dont des homologues ont été identifiés dans le transcriptome de *C. australe*. L'ACP présenté ici est basée sur l'expression de ces 273 sondes au cours du développement des graines de *C. australe* (bleu), *M. truncatula* R108 (orange) et *Mtabi3-2* (vert). Le stade radicule (violet) correspond aux gènes exprimés dans une radicule de 2,7 mm. La visualisation 3D est représentée sur deux plans : selon les axes X et Y (respectivement 35% et 28% de la variation totale des données) **(A)** puis selon les axes Y et Z (13% de la variation totale des données) **(B)**.

début de développement qu'en fin de développement. A titre comparatif, l'intensité maximale (en log) observée pour chaque échantillon est située entre 8 et 8,5.

Les intensités médianes des gènes exprimés aux différents stades de développement des graines A17, R108, *Mtabi3-2* et *C. australe* ont été placés le long d'un axe allant de l'embryogenèse à la germination (Figure 3.6A). Comme le développement des graines R108 est légèrement plus rapide que celui des graines A17, le stade 16 DAF R108 est proche du stade 20 DAF A17 tout comme le stade 24 DAF R108 est proche du stade 26 DAF A17. De même le stade 40 DAF R108 correspond à l'abscission de la graine et doit être rapproché du stade ABS de A17 plutôt que du stade 41 DAF. Le développement des graines *Mtabi3-2* suit la même chronologie que chez le sauvage R108, les différents stades *Mtabi3-2* ont donc été placés au même niveau que pour R108.

Après vérification de l'homogénéité des données transcriptomiques grâce au calcul de l'intensité médiane sur l'ensemble des sondes, nous avons analysé un jeu de données restreint dont le niveau d'expression est susceptible de varier au cours du développement des graines : les gènes préférentiellement exprimés dans les graines. Ces gènes ont été identifiés à partir de données transcriptomiques issues de différents tissus de *M. truncatula* (feuille, tige, racine, graine). Ils ont été définis comme étant les gènes au moins cinq fois plus exprimés dans les graines que dans les autres tissus. Sur les 15 083 gènes analysés, 273 sont préférentiellement exprimés dans les graines. La valeur médiane des intensités associées à ces gènes montre des variations très importante au cours du développement de *M. truncatula* (Figure 3.6B). Elle augmente progressivement au cours du remplissage jusqu'à atteindre son maximum à 32 DAF (log intensité médiane > 5) avant de diminuer en fin de développement pour les graines des deux génotypes. Comme attendu pour des gènes préférentiellement exprimés dans les graines, la valeur médiane est très faible (1,74) dans les autres tissus, comme les radicules R108 lors de la germination. L'évolution de la valeur médiane est différente pour les graines *Mtabi3-2*. Elle augmente plus tardivement que chez les graines sauvages R108 et ne diminue pas en fin de développement car elle atteint son maximum au stade 40 DAF. Cette différence majeure chez le mutant suggère un retard général du développement des graines en l'absence de l'expression du gène *ABI3*. La valeur médiane d'expression de ces 273 gènes préférentiellement exprimés dans les graines nous donne une indication pour placer les trois stades de *C. australe* sur cet axe de développement. En effet, pour les stades immature et mi-mature, elle est respectivement très proche des stades 16 DAF et 24 DAF des graines R108. Elle diminue ensuite fortement dans les graines matures pour atteindre une valeur proche de celle observée à l'abscission des graines A17 et R108. Cependant, ces graines ont été récoltées dans l'arbre avant l'abscission ce qui explique leur position légèrement en avant de l'abscission sur l'axe de développement.

2.2.2-Analyse en composantes principales

Pour aller plus loin dans le rapprochement des stades de développement entre les graines de *M. truncatula* et de *C. australe*, les valeurs d'intensité des 273 sondes liées aux gènes préférentiellement exprimés dans les graines ont été analysées en composantes principales (ACP). Pour plus de clarté, les stades de développement des graines A17 ne sont pas figurés sur les figures 3.7A et 3.7B. Le premier axe, qui explique 35 % de la variation, montre une séparation claire entre les échantillons de *M. truncatula* et ceux de *C. australe*. L'effet espèce est donc important (Figure 3.7A). Les axes Y et Z, qui expliquent respectivement 28 % et 13 % de la variation, mettent en évidence l'évolution au cours du développement des graines de *M.*

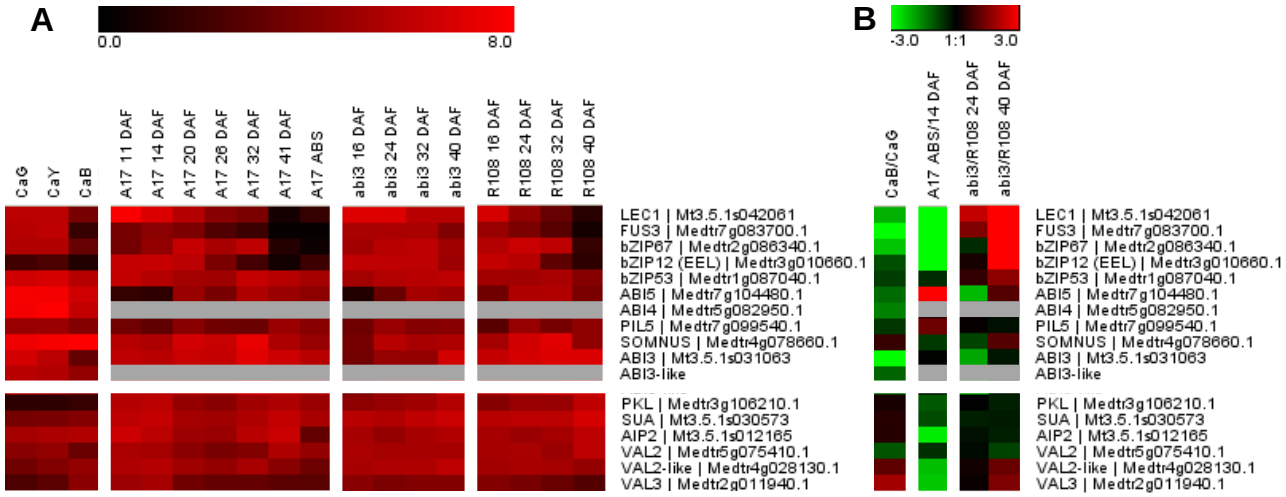


Figure 3.8. Intensité d'expression des gènes codant pour les régulateurs majeurs de la maturation au cours du développement des graines de *C. australe*, *M. truncatula* A17, R108 et *Mtabi3-2*.

A) L'évolution de l'expression des gènes codant pour les régulateurs majeurs de la maturation des graines est représentée au cours du développement : Leafy cotyledon-1 (LEC1), FUSCA3 (FUS3), basic Leucine Zipper67 (bZIP67), basic Leucine Zipper12 (bZIP12 appelé aussi EEL), basic Leucine Zipper53 (bZIP53), ABA-insensitive-5 (ABI5), ABA-insensitive-4 (ABI4), Phytochrome Interacting factor3 like5 (PIL5), SOMNUS et ABA-insensitive-3 (ABI3) ainsi que des répresseurs majeurs de la maturation Pickle (PKL), Suppressor of abi3-5 (SUA), ABI3-interacting protein (AIP2), VP1/ABI3-like2 (VAL2) et VP1/ABI3-like3 (VAL3). L'intensité d'expression est visualisée sous la forme d'un gradient de couleur allant de noir pour une expression nulle à rouge pour une expression maximale (intensité exprimée en log). Chez *M. truncatula* aucun gène *ABI3-like* n'a été mis en évidence et la sonde de la puce liée au gène *ABI4* ne permet pas de suivre son expression. L'absence de données d'expression pour ces deux gènes est donc remplacée par un carré gris. En revanche, deux gènes codant pour VAL2 ont été identifiés chez *M. truncatula* et *C. australe*, le deuxième a été nommé *VAL2-like* dans cette figure. **B)** Pour chacun de ces gènes, les ratios d'intensité entre stade mature (CaB et ABS) et immature (CaG et 14 DAF) des graines de *C. australe* et *M. truncatula* A17 sont figurés dans les deux premières colonnes. Les autres colonnes représentent la différence d'expression entre les graines mutantes *Mtabi3-2* et sauvages R108 à un stade donné (24 DAF pour la troisième colonne et 40 DAF pour la quatrième colonne). La différence d'expression est visualisée sous la forme d'un gradient de couleur allant de vert pour une sous expression chez la graine mature par rapport à la graine immature ou chez la graine mutante par rapport à la graine sauvage ; rouge pour une sur expression et noir si l'expression est comparable entre les deux conditions.

truncatula et de *C. australe* (Figure 3.7B). L'ACP confirme d'une part que le développement des graines *Mtabi3-2* semble retardé comparé à celui des graines R108. D'autre part, alors que les stades immatures et mi-matures de *C. australe* sont très proches entre eux, ils sont très différents du stade mature.

Sur la base de l'ACP, il est difficile d'associer chaque stade de développement des graines de *C. australe* à un stade de développement de *M. truncatula*. Cependant, il est intéressant de constater que l'évolution entre les stades mi-mature et mature de *C. australe* suit la même direction qu'une graine de *M. truncatula* en germination.

2.3-Expression chez *C. australe* des régulateurs majeurs de la maturation des graines orthodoxes

Après une première analyse globale basée sur le niveau d'expression de l'ensemble des gènes identifiés chez *M. truncatula* et *C. australe*, concentrons nous maintenant sur le profil d'expression de groupes de gènes spécifiques en commençant par les régulateurs majeurs de maturation des graines.

LEC1, *FUS3*, *bZIP67* (codant pour un facteur qui interagit avec LEC1 au cours de la maturation des graines [Yamamoto *et al.*, 2009]) et *EEL* (codant pour un facteur antagoniste d'ABI5 [Bensmihen *et al.*, 2002 ; 2005]) sont exprimés exclusivement au début du développement des graines de *M. truncatula* A17 et R108 (Figure 3.8A). Le même profil est observé au cours du développement des graines de *C. australe*, avec cependant un niveau d'expression plus faible pour *EEL* tout au long du développement. Ces quatre gènes sont donc sous-exprimés dans les graines matures par rapport aux graines immatures chez les deux espèces (Figure 3.8B colonnes 1 et 2). Les gènes *ABI5* et *bZIP53* (codant pour un facteur capable de former des hétérodimères avec ABI5 [Alonso *et al.*, 2009]) présentent également des profils d'expression similaires entre *M. truncatula* et *C. australe*. Le gène *bZIP53* est exprimé tout au long du développement des graines de *M. truncatula* alors que l'expression d'*ABI5* augmente en cours de maturation (20 DAF pour A17 et 24 DAF pour R108) avant de diminuer légèrement en fin de développement (Figure 3.8A). Chez *C. australe*, l'expression est forte aux stades immature et mi-mature et diminue également légèrement au stade mature. Le gène *ABI4* suit le même profil d'expression qu'*ABI5* dans les cotylédons de *C. australe* (Figure 3.8A). Aucun signal n'est détecté au cours du développement des graines de *M. truncatula* pour la sonde associée à ABI4. Cependant, il a été montré avec l'ancienne puce Affymetrix que ce gène était exprimé au cours du développement. Ce résultat serait donc plutôt du à un problème technique et n'est pas considéré ici.

D'autres gènes présentent des profils différents entre *M. truncatula* et *C. australe* au cours du développement des graines. Par exemple, l'expression de *PIL5* diminue légèrement dans les graines matures de *C. australe* (ratio = -0,7) alors qu'elle augmente légèrement dans les graines matures de *M. truncatula* A17 (ratio = 1,3) (Figures 3.8A et 3.8B colonnes 1 et 2). Le profil d'expression inverse est observé pour *SOMNUS*, avec une expression très élevée dans les graines immature et mature de *C. australe* alors qu'une diminution est visible dans les graines matures de *M. truncatula* A17 (Figure 3.8A). Mais la différence la plus importante entre les deux espèces concerne le profil d'expression du gène *ABI3*. Ce gène est fortement exprimé tout au long du développement des graines de *M. truncatula* A17 et R108 (Figure 3.8A). Deux gènes *ABI3* ont été identifiés chez *C. australe* (cf chapitre IV) et suivent tous les deux le même profil d'expression dans les cotylédons. Ils

sont exprimés aux stades immature et mi-mature (Figure 3.8A) mais fortement sous-exprimés au stade mature, contrairement à ce qui est observé chez *M. truncatula* (Figure 3.8B).

L'expression de la plupart de ces régulateurs est décalée au cours du développement des graines *Mtabi3-2* par rapport au type sauvage R108. Par exemple, les gènes *FUS3*, *LEC1*, *bZIP67*, *EEL* et *bZIP53* sont exprimés au début du développement des graines *Mtabi3-2* comme chez R108 mais reste exprimés en fin de maturation (Figure 3.8A). Ils sont donc fortement sur-exprimés par rapport aux graines sauvages R108 au stade 40 DAF (Figure 3.8B colonne 4).

L'expression d'autres gènes est retardée chez *Mtabi3-2* par rapport au type sauvage. Par exemple, le niveau d'expression du gène *ABI5* est bas en début de maturation mais augmente fortement au stade 32 DAF. chez le mutant *Mtabi3-2*. Ce gène est donc sous-exprimé au stade 24 DAF dans les graines mutantes mais le niveau d'expression est comparable au type sauvage en fin de maturation (40 DAF) (Figure 3.8A et 3.8B colonnes 3 et 4). Le même profil est observé pour *ABI3*.

Enfin les gènes *PIL5* et *SOMNUS* constituent une troisième catégorie de gènes pour lesquels le profil d'expression est semblable entre *Mtabi3-2* et R108 tout au long du développement (Figure 3.8A et 3.8B colonnes 3 et 4).

L'intensité d'expression des gènes codant pour les répresseurs majeurs de la maturation des graines PKL, SUA et AIP2 est forte et varie peu au cours du développement des graines de *M. truncatula* (excepté une forte diminution au stade abscission pour *AIP2*). Elle ne varie pas non plus dans les cotylédons de *C. australe* mais le niveau d'expression est plus faible que chez *M. truncatula* pour *PKL* et *SUA* (Figure 3.8A).

Concernant les gènes *VAL*, les différences d'expression sont plus marquées au cours du développement et entre espèces. *VAL2-like* et *VAL3* suivent le même profil d'expression : ils sont sur-exprimés dans les graines matures de *C. australe* mais sous-exprimés dans celles de *M. truncatula* A17 par rapport aux graines immatures (Figure 3.8B colonnes 1 et 2). En revanche, le gène *VAL2* suit un profil d'expression inverse. Il reste exprimé dans les graines matures de *M. truncatula* A17 et R108, alors qu'il est sous-exprimé dans celles de *C. australe* par rapport au stade immature (Figure 3.8A et 3.8B colonnes 1 et 2).

Le profil d'expression des répresseurs de la maturation est moins divergent entre les graines mutantes *Mtabi3-2* et les graines sauvages R108. *PKL*, *SUA* et *AIP2* sont fortement exprimés à un niveau comparable tout au long du développement des graines de R108 et *Mtabi3-2* (Figure 3.8A et 3.8B colonnes 3 et 4). Cependant, *VAL2-like* et *VAL3* restent exprimés tout au long du développement des graines *Mtabi3-2* alors qu'ils sont réprimés dans les graines matures de R108 au stade 40 DAF (Figure 3.9A). A l'inverse le gène *VAL2* est sous-exprimé dans les graines matures *Mtabi3-2* par rapport aux graines matures sauvages (Figure 3.8B colonne 4).

Finalement, concernant l'expression des régulateurs majeurs de la maturation, la différence majeure entre les graines de *C. australe* et de *M. truncatula* est la sous-expression d'*ABI3* dans les graines matures récalcitrantes, alors que ce gène est exprimé dans les graines immatures à un niveau comparable chez les deux espèces. Nous avons également mis en évidence une expression tardive de ce gène dans les graines *Mtabi3-2*. Ce résultat est surprenant car l'insertion du transposon TnT1 (5 334 pb) entre le domaine B2 et le domaine B3 du gène *ABI3* devrait empêcher son expression. Dans le chapitre II, l'absence de transcrit dans les mutants

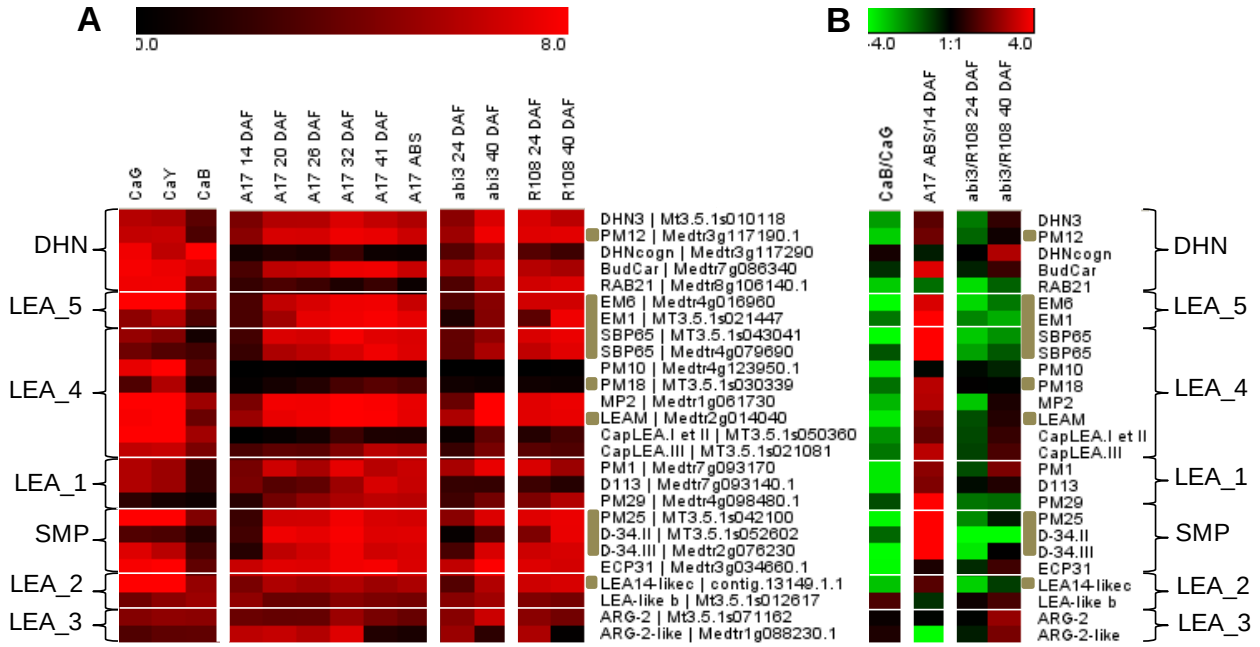


Figure 3.9. Intensité et ratio d'expression des gènes codant pour les protéines LEA au cours du développement des graines de *C. australe*, *M. truncatula* A17, R108 et *Mtabi3-2.*

Les gènes codant pour les protéines LEA sont classés par famille selon la nomenclature Pfam (DHN = déhydrines ; SMP = Seed Maturation Protein). **A)** L'intensité d'expression est visualisée sous la forme d'un gradient de couleur allant de noir pour une expression nulle à rouge pour une expression maximale (intensité exprimée en log). **B)** Les ratios d'intensité entre stade mature (CaB et ABS) et immature (CaG et 14 DAF) des graines de *C. australe* et *M. truncatula* A17 sont figurés dans les deux premières colonnes. Les autres colonnes représentent la différence d'expression entre les graines mutantes *Mtabi3-2* et sauvages R108 à un stade donné (24 DAF pour la troisième colonne et 40 DAF pour la quatrième colonne). La différence d'expression est visualisée sous la forme d'un gradient de couleur allant de vert pour une sous expression chez la graine mature par rapport à la graine immature ou chez la graine mutante par rapport à la graine sauvage ; rouge pour une sur expression et noir si l'expression est comparable entre les deux conditions. Les gènes dont l'expression est spécifique des graines d'après Chatelain *et al.* (2012) sont indiqués par un carré marron. Sur les puces à ADN Mt et Ca, deux sondes sont associées à des gènes codant pour la protéine SBP65. Les signaux de ces deux sondes sont représentés sur la figure. Comme présenté dans le chapitre précédent, seul le contig Ca_8304 a été identifié dans le transcriptome de graines de *C. australe* comme étant homologue des gènes PM1 et D113 de *M. truncatula*. La très forte similarité entre les séquences de ces deux gènes ne permet pas d'associer préférentiellement ce contig à *PM1* ou *D113*. Par conséquent, le signal associé au contig Ca_8304 est représenté ici pour l'expression de ces deux gènes chez *C. australe*.

Mtabi3-1 et *Mtabi3-2* avaient été vérifiée à partir d'ARN extraits au stade 24 DAF par RT-PCR avec des amorces ATG-stop (Figure 2.5). Cependant, la production en fin de maturation d'un transcrit tronqué pouvant s'hybrider sur la puce à ADN ne peut pas être écartée. La sonde de la puce spécifique d'*ABI3* hybride une région juste en aval du transposon (126 pb après le transposon). Quoi qu'il en soit, le profil d'expression d'*ABI3* est inversé entre les graines de *C. australe* et celles du mutant *Mtabi3-2*. Alors que l'expression est retardée chez le mutant par rapport aux graines sauvages (mais atteint un niveau élevé dans les graines matures), elle est éteinte au cours de la maturation des graines de *C. australe*. Nous nous sommes donc ensuite intéressés à l'expression des gènes connus pour être des cibles du facteur de transcription ABI3 afin de déterminer si les mêmes profils d'expression pouvaient être mis en évidence.

2.4-ABI3 et ses cibles : un réseau sous-exprimé dans les graines matures de *C. australe* par rapport à *M. truncatula*

2.4.1-Les gènes codant pour les protéines LEA (Late Embryogenesis Abundant)

De nombreuses études ont confirmé le rôle d'ABI3 dans l'activation de l'expression de la majorité des protéines LEA (Parcy *et al.*, 1994 ; Bies-Ethève *et al.*, 1999 ; Mönke *et al.*, 2012 ; Verdier *et al.*, 2013). Dans le chapitre précédent, 16 homologues sur les 17 gènes codant pour des LEA présentes dans le protéome de graines de *M. truncatula* ont été identifiés dans le transcriptome de *C. australe*. En ajoutant les protéines LEA identifiées dans le transcriptome de *M. truncatula* mais non détectées dans le protéome stable à la chaleur des graines, un total de 34 gènes codant pour des protéines LEA, répartis en 7 familles, a été mis en évidence chez *M. truncatula* (Chatelain *et al.*, 2012). Des homologues pour 25 d'entre eux ont été identifiés dans le transcriptome de *C. australe* et l'évolution de l'expression de ces gènes au cours du développement des graines de *M. truncatula* A17, R108, *Mtabi3-2* et *C. australe* est présentée dans la figure 3.9.

La majorité de ces 25 gènes sont exprimés (log intensité > 3) au cours du développement des graines de *M. truncatula* (20 pour A17 et 21 pour R108) et de *C. australe* (23). 14 gènes ont un niveau d'expression comparable dans les graines immatures de *C. australe* et de *M. truncatula* A17 (différence de log intensité < 1). Six gènes sont plus exprimés chez l'espèce récalcitrante (*DHN cognate*, *RAB21*, *PM10*, *PM18*, *CAPLEA-I* et *LEA14-likec*) et cinq chez l'espèce orthodoxe (*EM1*, *SBP65*, *PM29*, *D34-II* et *Arg2-like*) (Figure 3.10A). Parmi ces cinq gènes, trois ont été déterminés comme spécifiques des graines (*EM1*, *SBP65* et *D34-II*) tandis que l'expression des deux autres n'a pas été étudiée (Chatelain *et al.*, 2012). Le très faible niveau d'expression des gènes *PM10* et *PM18* chez *M. truncatula* A17 est surprenant car les protéines correspondantes sont fortement accumulées dans le protéome stable à la chaleur (Chapitre II). Une mauvaise spécificité de la sonde pourrait expliquer ce résultat.

Dans les graines de *M. truncatula* A17, la grande majorité des gènes exprimés (18/20) sont sur-exprimés dans les graines à l'abscission par rapport au stade 14 DAF (ratio >1) (Figure 3.9B colonne 1). Le profil d'expression inverse est observé chez *C. australe*. En effet, l'expression de 18 des 23 gènes exprimés est réprimée dans les graines matures (ratio < -1) (Figure 3.9B colonne 2). Les cinq gènes restant dont l'expression ne diminue pas dans les graines matures de *C. australe* sont membres des déhydrines (*BudCar* et

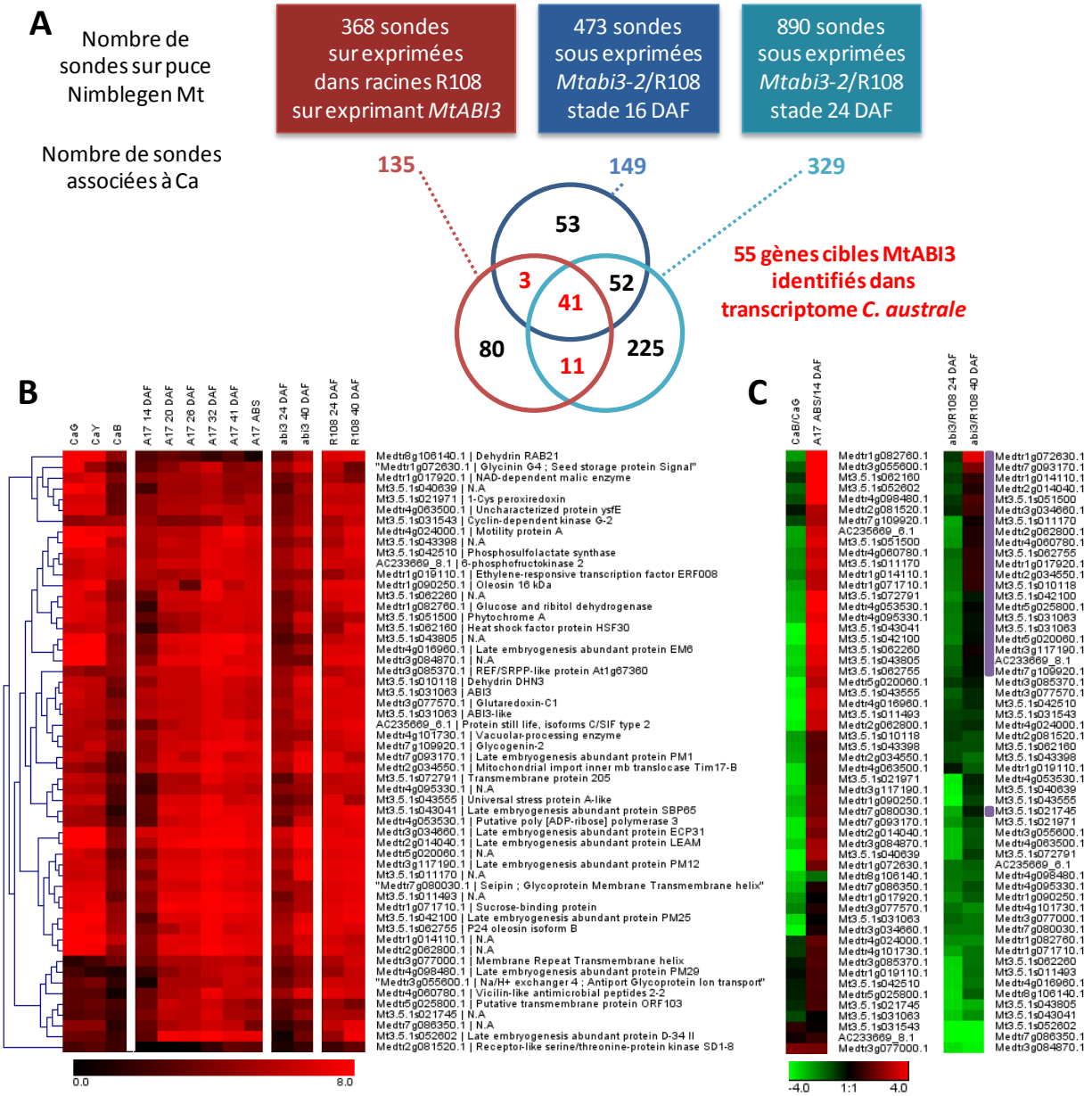


Figure 3.10. Intensité et ratio d'expression des gènes cibles de MtABI3 identifié dans le transcriptome de graines de *C. australe* au cours du développement des graines de *C. australe* et *M. truncatula* A17.

A) Les résultats de deux analyses transcriptomiques ont été utilisés pour déterminer les gènes cibles de MtABI3 : la sur expression de *MtABI3* dans les racines de *M. truncatula* R108 (Verdier *et al.*, 2013) et l'analyse transcriptomique du développement des graines du mutant *Mtabi3-2* (données non publiées). Un gène cible de MtABI3 a été défini comme étant à la fois sur-exprimé en réponse à la sur-expression de *MtABI3* dans les racines (cadre rouge et cercle rouge du diagramme de Venn) et sous-exprimé dans les graines de *Mtabi3-2* comparé à celles du type sauvage R108 au stade 16 DAF (cadre et cercle bleu) ou 24 DAF (cadre et cercle cyan). Le nombre de gènes associés à une sonde de la puce Nimblegen Ca parmi les trois listes de gènes identifiés sur la puce Nimblegen Mt est indiqué sous chaque cadre. 55 gènes cibles de MtABI3 ont ainsi été identifiés dans le transcriptome de *C. australe* (en rouge dans le diagramme de Venn). **B)** L'intensité de leur expression au cours du développement des graines de *C. australe*, de *M. truncatula* A17, R108 et *Mtabi3-2* est visualisée sous la forme d'un gradient de couleur allant de noir pour une expression nulle à rouge pour une expression maximale (intensité exprimée en log) ; N.A = non annoté. **C)** Les ratios d'intensité entre stade mature (CaB et ABS) et immature (CaG et 14 DAF) des graines de *C. australe* et *M. truncatula* A17 sont figurés dans les deux premières colonnes. Les autres colonnes représentent la différence d'expression entre les graines mutantes *Mtabi3-2* et sauvages R108 à un stade donné (24 DAF pour la troisième colonne et 40 DAF pour la quatrième colonne). La différence d'expression est visualisée sous la forme d'un gradient de couleur allant de vert pour une sous-expression chez la graine mature par rapport à la graine immature ou chez la graine mutante par rapport à la graine sauvage ; rouge pour une sur expression et noir si l'expression est comparable entre les deux conditions ; carré violet = gène sous-exprimé dans les graines immatures *Mtabi3-2* (24 DAF) par rapport à celles de R108 mais non sous-exprimé dans les graines matures (40 DAF).

DHN cognate) ou des familles LEA-2 (LEA-likeb) et LEA-3 (ARG2 et ARG2-like) regroupant des protéines LEA non spécifiques des graines (Figure 3.9B colonne 2).

La grande majorité des gènes codant pour des protéines LEA (17 sur 25) et la totalité de ceux spécifiquement exprimés dans les graines (9 sur 9 en excluant PM18) sont sous-exprimés dans les graines immatures de *Mtabi3-2* comparé aux graines immatures de R108 (ratio < -1) (Figure 3.9B colonne 3). Cependant, seuls 7 gènes sur 25 (5 sur 9 spécifiques des graines) sont encore sous-exprimés dans les graines matures de *Mtabi3-2* : *RAB21*, *EM6*, *EM1*, *SBP65*, *PM29*, *D34-II* et *LEA14-likec*. Certains sont même sur-exprimés dans les graines mutantes, comme *DHN cognate*, *CAPLEA-III*, *PM1*, *ARG2* et *ARG2-like* (Figure 3.9B colonne 4).

Que ce soit dans les graines de *M. truncatula*, de *Mtabi3-2* ou de *C. australe*, l'expression des gènes codant pour les protéines LEA suit le même profil qu'*ABI3* :

- Une forte expression dans les graines matures de *M. truncatula*.
- Une expression importante dans les graines immatures de *C. australe* suivie d'une forte répression dans les graines matures
- Un retard d'expression dans les graines de *Mtabi3-2*.

Cette conclusion doit être nuancée pour *Mtabi3-2* car sept gènes restent sous-exprimés dans les graines matures mutantes par rapport à celles du type sauvage R108 (Figure 3.9B colonne 4). Ces LEA sont-elles des cibles particulières d'ABI3 ? Qu'en est-il des autres gènes cibles d'ABI3 ne codant pas pour des protéines LEA ?

2.4.2-Les gènes cibles régulés par MtABI3

Les LEA font partie des cibles d'ABI3 mais d'autres gènes sont positivement régulés par ce facteur de transcription. Une liste de 368 gènes potentiellement cibles de *MtABI3* (ou régulés en aval de ce facteur) a été récemment déterminée par sur-expression de ce gène dans des racines de *M. truncatula* (Verdier *et al.*, 2013). Parmi ces gènes, 135 (37 %) font partie des 15 083 gènes retrouvés dans le transcriptome de *C. australe* (Figure 3.10A). Afin de conserver uniquement les cibles les plus spécifiques de MtABI3, ces données ont été recoupées avec l'analyse transcriptomique du développement des graines de *Mtabi3-2*, en partant du principe que les gènes cibles de MtABI3 sont sous-exprimés dans les graines mutantes par rapport aux graines sauvages. Seuls les stades 16 DAF et 24 DAF ont été considérés car en fin de maturation l'expression d'un transcrit *ABI3* est mise en évidence chez le mutant (Figure 3.8A). Respectivement, 473 et 890 gènes sont sous-exprimés dans les graines *Mtabi3-2* par rapport aux graines R108 (BH < 5 %) aux stades 16 DAF et 24 DAF. 149 (32 %) et 329 (37 %) de ces gènes ont été identifiés dans le transcriptome de graines de *C. australe* (Figure 3.10A). En croisant ces trois jeux de données, 55 gènes identifiés chez *C. australe* sont retenus comme cibles potentielles d'ABI3 (intersection en rouge du diagramme de Venn de la figure 3.10A). Presque la totalité de ces gènes est exprimée dans les graines de *M. truncatula* et dans les graines immatures de *C. australe* (Figure 3.10B). Cependant, comme pour les gènes codant pour les protéines LEA, l'expression de ces gènes cibles est fortement réprimée dans les graines matures de l'espèce récalcitrante alors qu'elle est au contraire activée dans les graines matures de l'espèce orthodoxe (Figure 3.10C colonne 1 et 2). Le profil d'expression de

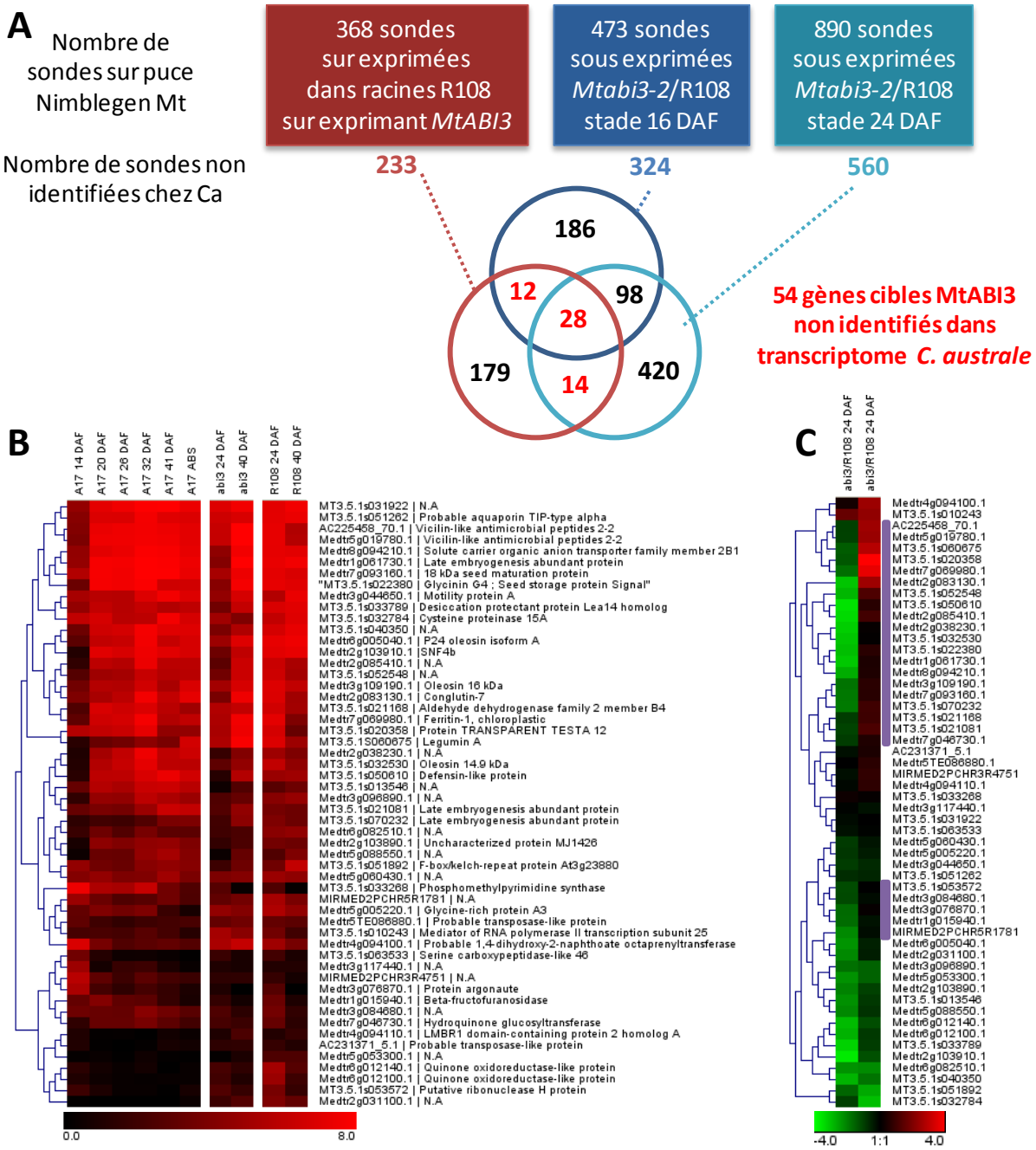


Figure 3.11. Intensité et ratio d'expression des gènes cibles de MtABI3 non identifié dans le transcriptome de graines de *C. australe* au cours du développement des graines de *M. truncatula* A17.

Dans la figure précédente, nous avions noté que moins de la moitié des gènes sur-exprimés en réponse à la sur-expression de *MtABI3* dans les racines (cadre rouge et cercle rouge du diagramme de Venn) et sous-exprimé dans les graines de *Mtabi3-2* comparé à celles du type sauvage R108 au stade 16 DAF (cadre et cercle bleu) ou 24 DAF (cadre et cercle cyan) avaient été identifiés dans le transcriptome de *C. australe*. Par conséquent, en appliquant la même démarche que pour la figure précédente, 54 gènes cibles de MtABI3 n'ont pas été identifiés dans le transcriptome de *C. australe* (en rouge dans le diagramme de Venn). L'intensité de leur expression au cours du développement des graines de *M. truncatula* A17, R108 et *Mtabi3-2* est visualisée sous la forme d'un gradient de couleur allant de noir pour une expression nulle à rouge pour une expression maximale (intensité exprimée en log) ; N.A = non annoté. Les ratios d'intensité entre graines mutantes *Mtabi3-2* et sauvages R108 à un stade donné (24 DAF pour la première colonne et 40 DAF pour la deuxième colonne) sont figurés. La différence d'expression est visualisée sous la forme d'un gradient de couleur allant de vert pour une sous expression chez la graine mature par rapport à la graine immature, rouge pour une sur expression et noir si l'expression est comparable entre les deux échantillons ; carré violet = gène sous-exprimé dans les graines immatures *Mtabi3-2* (24 DAF) par rapport à celles de R108 mais non sous-exprimé dans les graines matures (40 DAF).

ces 55 gènes cibles au cours du développement des graines *Mtabi3-2* est également semblable à ce qui est observé pour ceux codant les protéines LEA. 52 d'entre eux (95 %) sont sous-exprimés au stade 24 DAF par rapport aux graines sauvages R108 (BH < 5 %) (Figure 3.10C colonne 3) mais seulement 34 (62 %) le sont encore au stade 40 DAF. 12 gènes (22 %) sont mêmes sur-exprimés dans les graines *Mtabi3-2* par rapport aux graines sauvages à ce stade (BH < 5 %) (Figure 3.10C colonne 4).

Ces gènes cibles suivent donc parfaitement le profil d'expression du gène *ABI3* mis en évidence chez les deux espèces. Hormis ceux codant pour les protéines LEA, on retrouve parmi ces gènes des oléosines, un heat shock factor *HSF30*, des enzymes liés à la signalisation et au métabolisme des sucres (Protéine de liaison au saccharose, 6-phosphofructokinase 2, glycogenin 2), des protéines liées à la défense antioxydante (1-cys-peroxyredoxine, glutaredoxine) et le gène *ABI3* lui même.

Si 55 gènes cibles de MtABI3 ont pu être identifiés dans le transcriptome de *C. australe*, la majorité des gènes de *M. truncatula* sur exprimés dans les racines transformées et sous exprimés chez le mutant *Mtabi3-2* par rapport à R108 n'a pas trouvé d'homologues chez *C. australe*. En effet 233 des 368 gènes (63 %) mis en évidence par Verdier *et al.* (2013) ne font pas partie des 15 083 gènes identifiés sur les deux puces à ADN Mt et Ca. De même, 324 sur 473 (68 %) et 560 sur 890 (63 %) gènes sous-exprimés dans les graines *Mtabi3-2* par rapport aux graines R108 respectivement à 16 DAF et 24 DAF n'ont pas été identifiés dans les graines de *C. australe* (Figure 3.11A). En recoupant ces trois jeux de données comme précédemment, 54 gènes cibles de MtABI3 non identifiés dans le transcriptome de *C. australe* ont été mis en évidence (intersection en rouge du diagramme de Venn de la figure 3.11A). Ces gènes potentiellement non exprimés au cours du développement des graines de *C. australe* sont pour la plupart fortement exprimés au cours de la maturation des graines de *M. truncatula* A17 (Figure 3.11B). Une partie de ces gènes présentent des similitudes fonctionnelles avec d'autres gènes présents dans la figure 3.10 : gènes codant pour des oléosines, des protéines LEA, des protéines de réserve (legumine A, vicilline). Cependant, d'autres gènes jouant un rôle majeur dans la maturation des graines orthodoxes et pour lesquels aucun homologue n'est présent dans la figure 3.10 font également partie de ces 54 gènes : TT12 (Transparent Testa 12) (Debeaujon *et al.*, 2001) et SNF4b (Sucrose Non Fermenting related kinase subunit 4b) (Rosnoblet *et al.*, 2007 ; Bolingue *et al.*, 2010).

La majorité de ces 54 gènes (44 soit 81 %) est sous-exprimée au stade 24 DAF dans les graines *Mtabi3-2* par rapport aux graines sauvages R108 (BH < 5 %) (Figure 3.11C colonne 1) mais seulement 19 (35 %) le sont encore au stade 40 DAF. 19 autres gènes (35 %) sont mêmes sur-exprimés dans les graines *Mtabi3-2* par rapport aux graines sauvages à ce stade (BH < 5%) (Figure 3.11C colonne 2).

Le gène *ABI3* et les gènes cibles de ce facteur sont fortement sous-exprimés dans les graines matures de *C. australe*. ABI3 est impliqué dans la signalisation de l'ABA et dans la tolérance à la dessiccation (Ooms *et al.*, 1993 ; Parcy *et al.*, 1994 ; Bies-Ethève *et al.*, 1999). C'est également un régulateur majeur de la maturation de la graine et sa répression pourrait indiquer une préparation à la germination via l'activation de la voie de signalisation des GA. Les gènes impliqués dans la biosynthèse et la signalisation de l'ABA et des GA, la réponse au stress, le métabolisme des RFO ont été isolés parmi les 15 083 gènes et leur expression a été comparée au cours du développement des graines de *M. truncatula* et de *C. australe*. L'objectif est d'une part

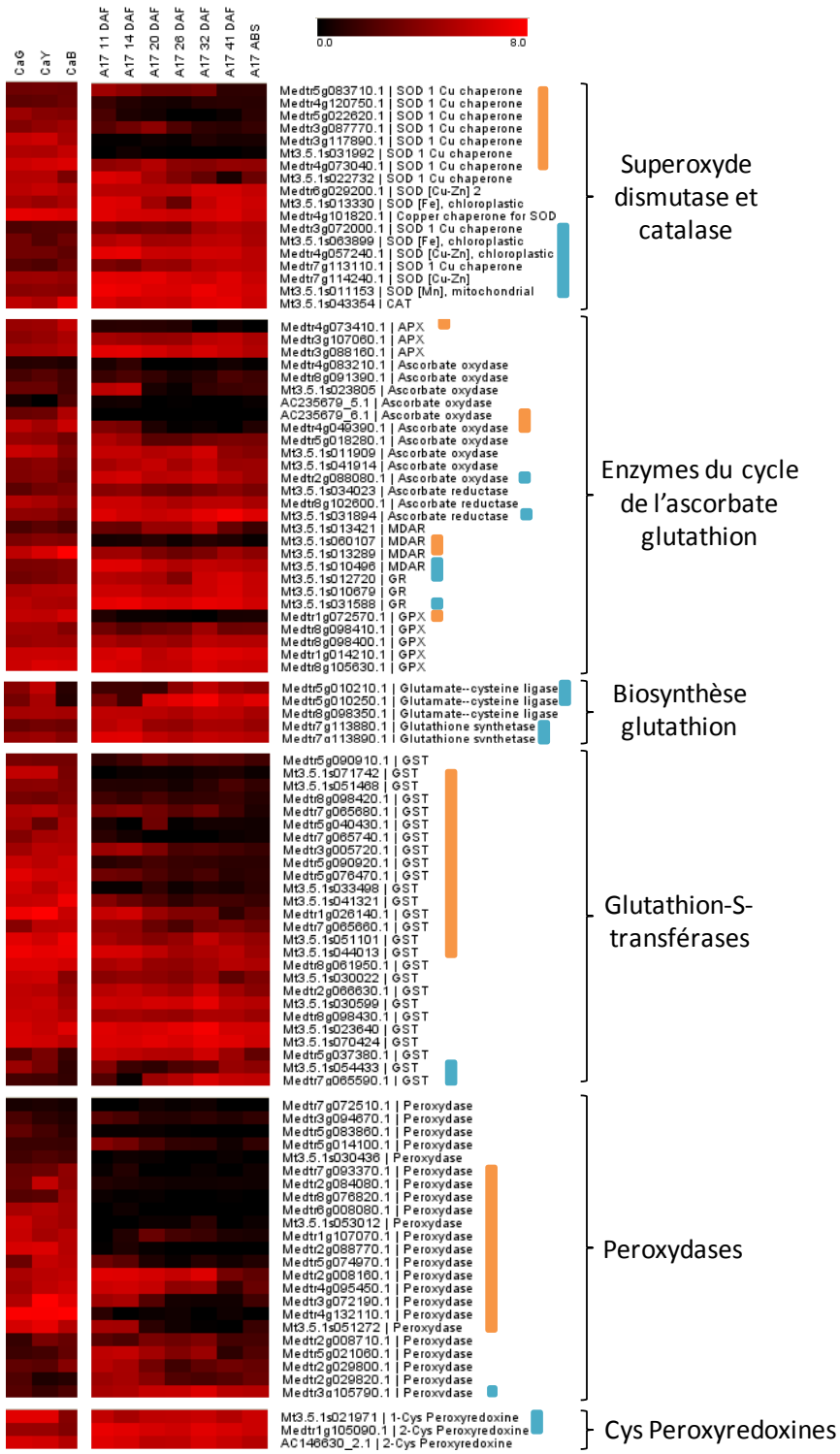


Figure 3.12. Intensité d'expression des gènes codant pour des protéines impliquées dans la réponse au stress oxydatif au cours du développement des graines de *C. australe* et *M. truncatula* A17.

Basé sur les annotations MT3.5, les gènes codant pour les enzymes et peroxyredoxines impliquées dans l'inactivation des espèces réactives de l'oxygène (ROS) ont été identifiées dans les transcriptomes de *C. australe* et *M. truncatula*. Ces gènes sont regroupés selon leur fonction [superoxyde dismutase (SOD) et catalase (CAT), enzymes impliquées dans le cycle de l'ascorbate/glutathion, Glutathion-S-transférase (GST), peroxydases ou peroxyredoxines] et par famille au sein de chaque fonction. L'intensité d'expression de ces gènes au cours du développement des graines de *C. australe* et *M. truncatula* A17 est visualisée sous la forme d'un gradient de couleur allant de noir pour une expression nulle à rouge pour une expression maximale (intensité exprimée en log). Un carré orange signifie que l'expression du gène est plus forte dans les graines matures de *C. australe* par rapport aux graines matures (41 DAF et ABS) de *M. truncatula.* A l'inverse, un carré bleu indique une expression plus forte dans les graines matures de *M. truncatula* par rapport à celles de *C. australe*. APX = ascorbate peroxydase ; MDAR = monodehydroascorbate réductase ; GR = glutathion réductase ; GPX = glutathion peroxydase.

d'identifier les mécanismes liés à la TD (autres que les LEA) exprimés chez *M. truncatula* mais sous-exprimés chez *C. australe* et d'autre part de caractériser les voies de signalisation activées dans la graine mature de l'espèce récalcitrante.

2.5-Expression des mécanismes liés à l'acquisition de la tolérance à la dessiccation : comparaison entre les graines de *C. australe* et de *M. truncatula*

2.5.1- Gènes liés à la réponse antioxydante

Les mécanismes antioxydants sont fortement mobilisés dans l'acquisition de la TD de la graine pour contrer la production croissante de ROS au cours de la dessiccation (revu dans Bailly, 2004 ; Kranner & Birtic, 2005). Les gènes codant pour des enzymes liées à la réponse antioxydante ont été identifiés parmi les 15 083 gènes communs à *M. truncatula* et *C. australe* en effectuant une recherche par mots clés dans les annotations MT3.5. Ces 103 gènes ont été regroupés en six catégories selon leur fonction ('SOD et CAT' ; 'cycle de l'ascorbate et glutathion' ; 'biosynthèse du glutathion' ; 'GST' ; 'peroxydases' ; 'cys-peroxyredoxine') et en famille au sein de chacune des catégories. L'expression de ces gènes au cours du développement des graines des deux espèces est présentée sur la figure 3.12. Les gènes fortement exprimés dans les graines matures de *C. australe* par rapport aux graines matures de *M. truncatula* (41 DAF et ABS) (différence log intensité > 2) sont identifiés par un carré orange. Un carré bleu est utilisé pour les gènes exprimés à un niveau plus élevé dans les graines matures de *M. truncatula*.

Parmi les 17 gènes codant pour une SOD, enzyme qui convertit l'ion superoxyde en peroxyde d'hydrogène (Kranner & Birtic, 2005), sept sont exprimés à un niveau plus élevé dans les graines matures de *C. australe* que dans celles de *M. truncatula* (différence log intensité > 2) et 6 suivent un profil inverse (Figure 3.12). Plusieurs co-facteurs métalliques différents sont associés à cette enzyme et dépendent de la localisation cellulaire (Bowler *et al.*, 1992). Ainsi les gènes codant pour les SOD cytosoliques (SOD [Cu]) sont principalement sur-exprimés dans les graines matures de *C. australe* tandis que ceux codant pour les SOD mitochondriales (SOD [Mn]) et chloroplastiques (SOD [Cu-Zn] et SOD [Fe]) sont exprimés plus fortement dans les graines matures de *M. truncatula*. Un seul gène codant pour une catalase, enzyme permettant la dégradation du peroxyde d'hydrogène, a été identifié. Il est exprimé à un niveau comparable dans les graines matures des deux espèces mais l'expression augmente fortement en fin de développement chez *C. australe* alors qu'elle est élevée tout au long du développement des graines de *M. truncatula* (Figure 3.12).

Le peroxyde d'hydrogène peut également être dégradé par le cycle de l'ascorbate et du glutathion qui implique l'activité de plusieurs enzymes. Aucune tendance forte ne se dégage en observant l'intensité d'expression des gènes impliqués dans ce cycle. Les gènes *Medtr4g073410.1* et *Medtr1g072570.1* codant respectivement pour une APX et une GPX sont exprimés plus fortement dans les graines matures de *C. australe* que dans celles de *M. truncatula* (Figure 3.12). Le même profil est observé avec les gènes codant pour les gènes *AC235679_6.1*, *Medtr4g049390.1* codant pour une ascorbate oxydase et les gènes *Mt3.5.1.s060107* *Mt3.5.1.s013289* codant pour une MDAR. Cependant, d'autres gènes codant pour ces enzymes suivent le profil inverse (*Medtr2g088080.1*, *Mt3.5.1.s031894* et *Mt3.5.1.s010496*) (Figure 3.12).

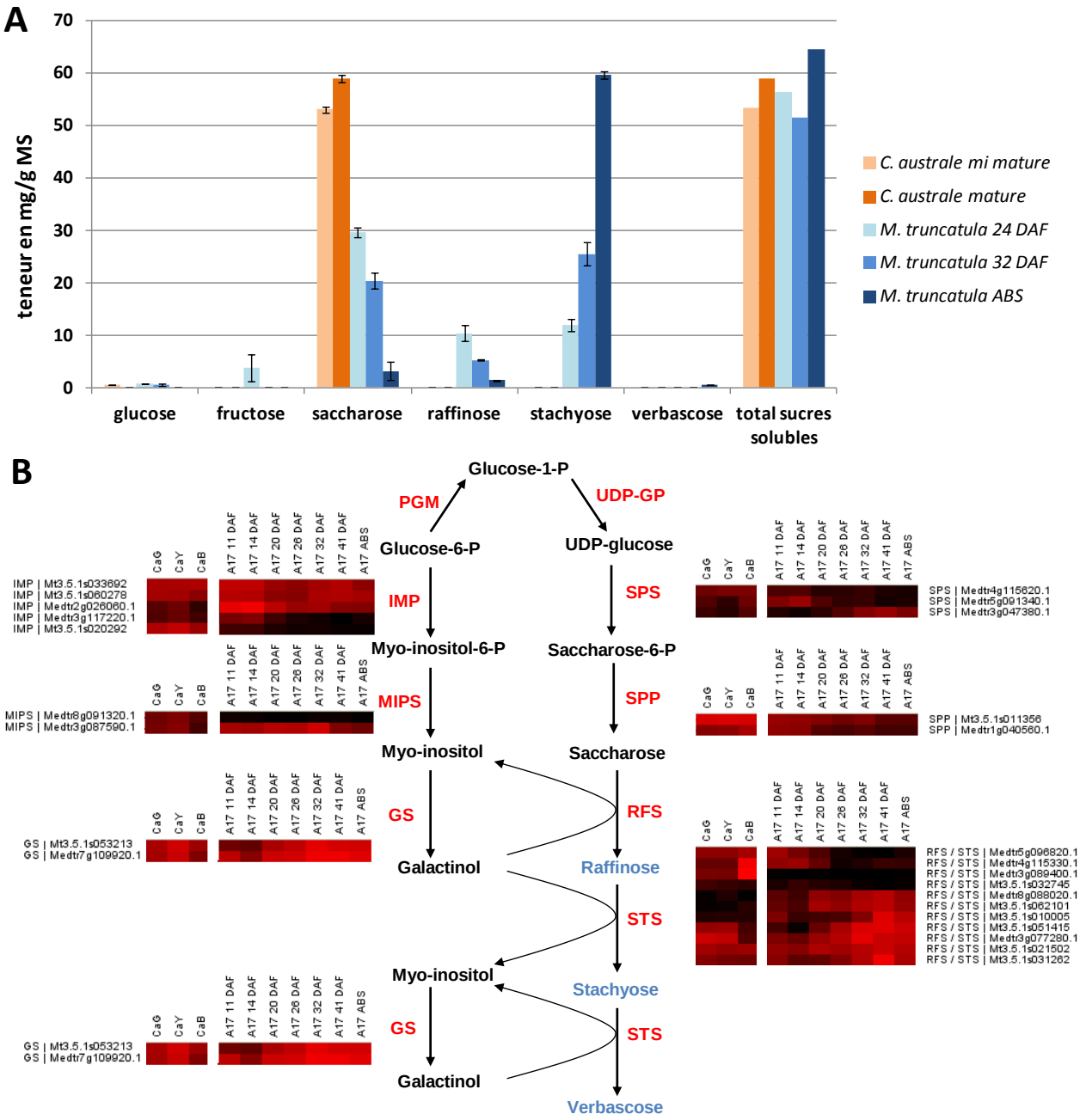


Figure 3.13. Teneurs en sucres solubles et intensité d'expression des gènes impliqués dans la synthèse des oligosaccharides de la famille du raffinose (RFO) au cours du développement des graines de *C. australe* et *M. truncatula* A17.

A) La teneur en glucose, fructose, saccharose et RFO (raffinose, stachyose et verbascose) a été mesurée dans des cotylédons de graines immatures et matures de *C. australe* et de *M. truncatula* (DAF = jours après floraison ; ABS = abscission) **B)** Sur cette voie métabolique, les RFO sont notés en bleu et les enzymes impliquées directement ou indirectement dans leur biosynthèse sont notées en rouge : PGM = phosphoglucomutase ; UDP-GP = UDP-glucose pyrophosphorylase ; SPS = saccharose phosphate synthase ; SPP = saccharose phosphate phosphorylase ; RFS = raffinose synthase ; STS = stachyose synthase ; IMP = myo-inositol monophosphatase ; MIPS = myo-inositol phosphate synthase ; GS = galactinol synthase. L'expression des gènes codant pour ces enzymes au cours du développement des graines de *C. australe* et *M. truncatula* est visualisée sous la forme d'un gradient de couleur allant de noir pour une expression nulle à rouge pour une expression maximale (intensité exprimée en log).

Notons que les gènes impliqués dans l'activité antioxydante dépendante du glutathion semblent moins activé dans les graines matures de *C. australe* que dans celles de *M. truncatula*. En effet, deux des trois gènes codant pour une GR et quatre des cinq gènes liés à la biosynthèse du glutathion sont plus exprimés dans les graines matures de *M. truncatula* (Figure 3.12). Le profil d'expression des gènes *Medtr5g010210.1* et *Medtr5g010250.1* codant pour une glutamate cystéine ligase est d'ailleurs très intéressant. Ces deux gènes, impliqués dans la biosynthèse du glutathion, sont exprimés dans les graines immatures de *C. australe* mais fortement réprimés dans les graines matures. Au contraire, ils sont progressivement exprimés au cours du développement des graines de *M. truncatula* jusqu'à être fortement exprimés dans les graines matures (Figure 3.12).

Deux familles de gènes codant pour des enzymes impliquées dans la réponse antioxydante sont fortement exprimées dans les graines matures de *C. australe* : les GST et les peroxydases. Concernant les GST, 15 gènes sur les 26 identifiés ont un niveau d'expression plus élevé dans les graines matures de *C. australe* que dans celles de *M. truncatula* alors que seulement deux suivent le profil inverse. Concernant les peroxydases, 13 gènes sur 23 sont plus exprimés chez *C. australe* et seulement un chez *M. truncatula* (Figure 3.12).

Enfin, parmi la dernière catégorie d'antioxydant étudié ici, un gène codant pour une 1-cys peroxyredoxine et un autre codant pour une 2-cys peroxyredoxine sont exprimés plus fortement dans les graines matures de *M. truncatula* que dans celles de *C. australe*.

La réponse antioxydante n'est pas complètement réprimée en fin de développement des graines de C. australe, comme en témoigne la forte expression de gènes codant pour les SOD cytosoliques, les GST et les peroxydases. Cependant, les gènes codant pour certaines catégories d'antioxydants liés au glutathion et aux 1- ou 2-cys peroxyredoxines sont peu exprimés dans les graines matures de *C. australe*.

2.5.2- Métabolisme et expression des enzymes de voie de biosynthèse des RFO

Une corrélation entre l'acquisition de la TD et l'accumulation des sucres solubles, particulièrement des RFO, a été démontrée dans plusieurs travaux (Koster & Leopold, 1988 ; Corbineau *et al.*, 2000 ; Bailly *et al.*, 2001). La teneur en sucres solubles a été mesurée au cours du développement des graines de *M. truncatula* (Rosnoblet *et al.*, 2007) et *C. australe* et est présentée dans la figure 3.13A. La teneur totale en sucres solubles est comparable entre les graines des deux espèces (comprise entre 51 et 65 mg / g MS selon les échantillons) mais la composition varie fortement au cours de la maturation des graines de *M. truncatula* et pas chez *C. australe*. Au stade 24 DAF, cinq sucres solubles différents ont été identifiés : le saccharose est majoritaire (52 % des sucres solubles totaux) suivi du stachyose (21 %) et du raffinose (19 %), des traces de fructose (7 %) et glucose (1 %) sont également détectées (Figure 3.13A). En fin de remplissage (32 DAF), les teneurs en saccharose (40 %) et raffinose (10 %) diminuent au profit du stachyose (49 %), des traces de glucose sont toujours présentes (1 %). A l'abscission, les graines contiennent presque exclusivement du stachyose (92 %) et des traces de saccharose (5 %), raffinose (2 %) et de verbascose (1 %). En revanche, quelque soit le stade de développement, le seul sucre soluble identifié dans les graines de *C. australe* est le saccharose (Figure 3.13A).

Une régulation transcriptionnelle pourrait expliquer l'absence de RFO dans les graines de *C. australe*. Pour vérifier cette hypothèse, les gènes codant pour les enzymes impliquées directement et indirectement dans la synthèse des RFO ont été identifiés parmi les 15 083 gènes communs à *M. truncatula* et *C. australe* par recherche de mots clé dans les annotations MT3.5 et TAIR10. Deux types d'enzymes sont présentés dans la figure 3.13B : (1) celles impliquées dans la synthèse de précurseur de RFO (saccharose, myo-inositol) et qui ne sont pas spécifiques de la synthèse des RFO ; (2) celles impliquées spécifiquement dans la synthèse de galactinol et de RFO.

Les gènes codant pour les enzymes SPS (Saccharose Phosphate Synthase) et SPP (Saccharose Phosphate Phosphorylase), impliquées dans la biosynthèse du saccharose sont exprimés dans les graines immatures et matures de *C. australe* (Figure 3.13B). Excepté pour le gène *Medtr3g047380.1*, l'expression des gènes codant les SPP et SPS décroit au cours du développement des graines de *M. truncatula*. Concernant la synthèse du myo-inositol, l'expression des gènes codant pour les enzymes IMP (myo-Inositol MonoPhosphatase) et MIPS (Myo-Inositol Phosphate Synthase) décroit au cours du développement des graines des deux espèces (Figure 3.13B). Les gènes présentés jusqu'à maintenant sont impliqués en amont dans la synthèse des RFO, qu'en est-il de l'expression des gènes spécifiquement impliqués dans la synthèse de ces molécules ?

Deux gènes codant pour la GS (Galactinol Synthase) ont été identifiés. Ils sont tous les deux sur-exprimés au cours de la maturation des graines de *M. truncatula* mais sous-exprimés dans les graines matures de *C. australe* (Figure 3.13B). Parmi les 11 gènes codant pour des RFS/STS (Raffinose Synthase / Stachyose Synthase) identifiés, 7 sont sur-exprimés au cours de la maturation des graines de *M. truncatula*, 2 sont sous-exprimés et 2 ne sont pas exprimés. Le profil d'expression de ces gènes est complètement inversé dans les cotylédons de *C. australe*. En effet, 5 des 7 gènes sur-exprimés chez *M. truncatula* (*Medtr8g088020.1*, *Mt3.5.1.s062101*, *Mt3.5.1.s010005*, *Mt3.5.1.s051415* et *Medtr3g077280.1*) ne sont pas exprimés ou sont sous-exprimés dans les graines matures de *C. australe* (Figure 3.13B). A l'inverse, parmi les 4 gènes non exprimés ou sous-exprimés au cours de la maturation de *M. truncatula*, 3 sont exprimés dans les graines matures de *C. australe* (*Medtr5g096820.1*, *Medtr4g115330.1* et *Medtr3g089400.1*). Les deux derniers sont mêmes fortement sur-exprimés dans les graines matures par rapport aux stades immatures (Figure 3.13B). L'activité des RFS/STS est réversible et certains de ces gènes pourraient être impliqués dans la dégradation de RFO, comme le gène *Medtr4g115330.1* annoté 'seed imbibition protein'.

Ces résultats transcriptomiques corrèlent très bien avec les dosages de sucres réalisés chez les deux espèces. Les gènes impliqués dans la synthèse du saccharose sont exprimés en fin de développement chez *C. australe* mais pas ceux impliqués dans la synthèse des RFO. Au contraire, ces derniers sont fortement sur-exprimés en fin de maturation chez *M. truncatula*.

Après s'être focalisé sur les mécanismes liés à l'acquisition de la TD, repartons de l'observation que la graine mature de *C. australe* présente des caractéristiques rappelant celle d'une graine prête à germer. La transition maturation/germination est gouvernée par le ratio ABA/GA et l'étude de l'expression des gènes

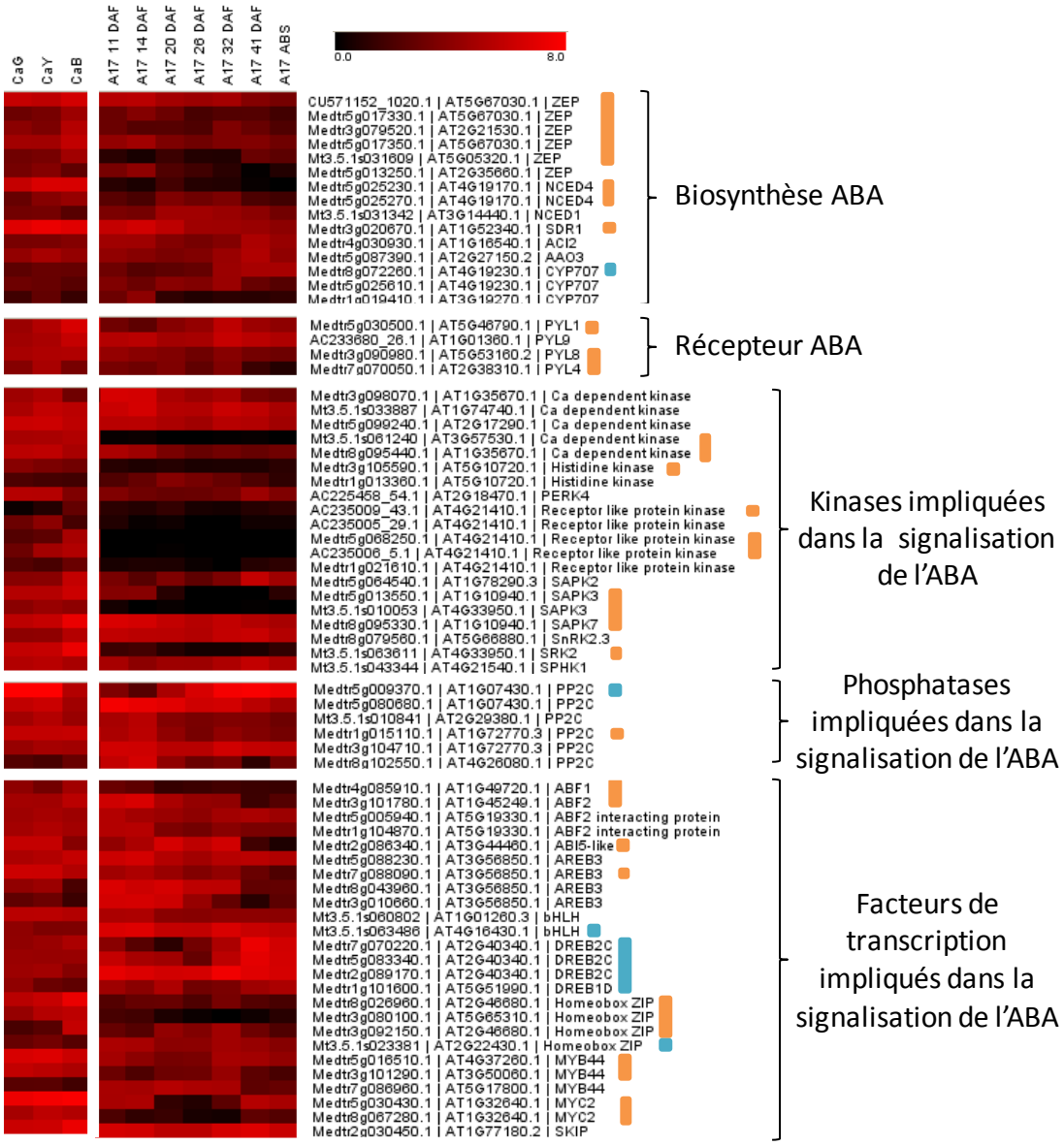


Figure 3.14. Intensité d'expression des gènes impliqués dans la biosynthèse et la signalisation de l'ABA au cours du développement des graines de *C. australe* et *M. truncatula* A17.

Basé sur le contenu des annotations MT3.5 et TAIR10, 70 gènes impliqués dans la biosynthèse ou la signalisation de l'ABA ont été identifiés dans les transcriptomes de *C. australe* et *M. truncatula*. Ces gènes sont regroupés selon leur fonction (biosynthèse, récepteur, kinase, phosphatase ou facteur de transcription) et par famille au sein de chaque fonction. L'intensité d'expression de ces gènes au cours du développement des graines de *C. australe* et *M. truncatula* A17 est visualisée sous la forme d'un gradient de couleur allant de noir pour une expression nulle à rouge pour une expression maximale (intensité exprimée en log). Un carré orange signifie que l'expression du gène est plus forte dans les graines matures de *C. australe* par rapport aux graines matures (41 DAF et ABS) de *M. truncatula* (différence log intensité > 2). A l'inverse, un carré bleu indique une expression plus forte dans les graines matures de *M. truncatula* par rapport à celles de *C. australe*. ZEP = zeaxanthin epoxydase ; NCED = 9-cis-epoxycarotenoïde dioxygenase ; SDR1 = dehydrogenase réductase des chaînes courtes ; ACI2 = molybdène cofacteur sulfurase ; AAO3 = abscissique aldehyde oxydase ; CYP707 = gène codant pour une ABA 8'-hydroxylase ; PYL = PYrabactine Resistance-1-Like ; PERK = Proline-rich receptor-like protein kinase ; SAPK = Stress Activated Protein Kinase ; SnRK2 ou SRK2 = Sucrose non-fermenting Related protein Kinase 2 ; SPHk = sphingosine kinase ; PP2C = protein phosphatase type 2C ; ABF = [ABA Responsive Element (ABRE)] Binding Factor ; AREB = ABA Responsive Element Binding protein ; bHLH = basic Helix Loop Helix ; DREB = Dehydration Responsive Element Binding ; Homeobox ZIP = Homeobox leucine zipper ; SKIP = Ski Interacting protein.

impliqués dans la voie de signalisation de ces deux phytohormones devrait nous permettre d'appuyer ou non cette hypothèse

2.6-La graine mature de *C. australe* exprime des gènes de réponse à l'acide abscissique et au stress mais également aux gibbérellines

2.6.1-Les gènes liés à la biosynthèse et à la signalisation de l'ABA

135 gènes liés à la biosynthèse et à la signalisation de l'ABA ont été identifiés parmi les 15 083 gènes communs à *M. truncatula* et *C. australe* en effectuant une recherche par mots clés dans les annotations MT3.5 et TAIR10. Ces 135 gènes ont été regroupés en catégories selon leur fonction et en famille au sein de chacune des catégories. Seuls les 70 gènes associés aux fonctions 'biosynthèse', 'récepteur', 'kinase', 'phosphatase' et 'facteur de transcription' sont représentés sur la figure 3.14 car le niveau d'expression des gènes appartenant à ces catégories présente le plus de variations entre les graines matures de *C. australe* et de *M. truncatula* A17. Les gènes fortement exprimés dans les graines matures de *C. australe* par rapport aux graines matures de *M. truncatula* (41 DAF et ABS) (différence log intensité > 2) sont identifiés par un carré orange. Un carré bleu est utilisé pour les gènes exprimés à un niveau plus élevé dans les graines matures de *M. truncatula*.

Concernant la biosynthèse d'ABA, le niveau d'expression des gènes codant pour les ZEP, NCED4 et SDR1 est plus élevé dans les graines matures de *C. australe* que dans celles de *M. truncatula*. Au contraire, un des trois gènes codant pour une CYP707, impliquée dans la dégradation de l'ABA, est exprimé plus fortement dans les graines matures de *M. truncatula* (Figure 3.14). Ces deux résultats indiquent que l'expression des gènes est compatible avec une production d'ABA dans les graines matures de *C. australe* mais pas dans celles de *M. truncatula*.

En aval de la production d'ABA, les gènes impliqués dans la perception et la signalisation de l'ABA indiquent également que ces voies sont activées dans les graines matures de *C. australe*, plus que dans celles de *M. truncatula*. En effet, trois des quatre gènes codant pour un récepteur d'ABA (PYL) ont un niveau d'expression supérieur dans les graines matures de *C. australe* (Figure 3.14). Plusieurs familles de kinases impliquées spécifiquement dans la signalisation de l'ABA sont également activées au niveau transcriptomique dans les graines matures de *C. australe* : certaines kinases dépendantes du calcium et histidine kinase mais surtout les 'receptor like protein kinase', et les SnRK2 (famille de kinase dans laquelle on retrouve SRK2 et les SAPK [Stress-Activated Protein Kinase]). Les deux gènes codant pour SAPK3 et ceux codant pour SAPK7 et SRK2 ont un niveau d'expression plus élevé dans les graines matures de *C. australe* que dans celles de *M. truncatula* (Figure 3.14).

Les phosphatases PP2C sont négativement impliquées dans la signalisation de l'ABA en empêchant l'activation des kinases de type SnRK2. Même si peu de différences sont visibles pour quatre des six gènes codant des PP2C, le profil d'expression du gène *Medtr5g009370.1* est inversé chez les deux espèces. Son expression augmente progressivement au cours du développement des graines A17 jusqu'à devenir très élevée dans les graines matures (Figure 3.14). A l'inverse, l'expression de ce gène est très élevée dans les graines immatures de *C. australe* mais est fortement réprimée dans les graines matures. Ce profil d'expression est tout

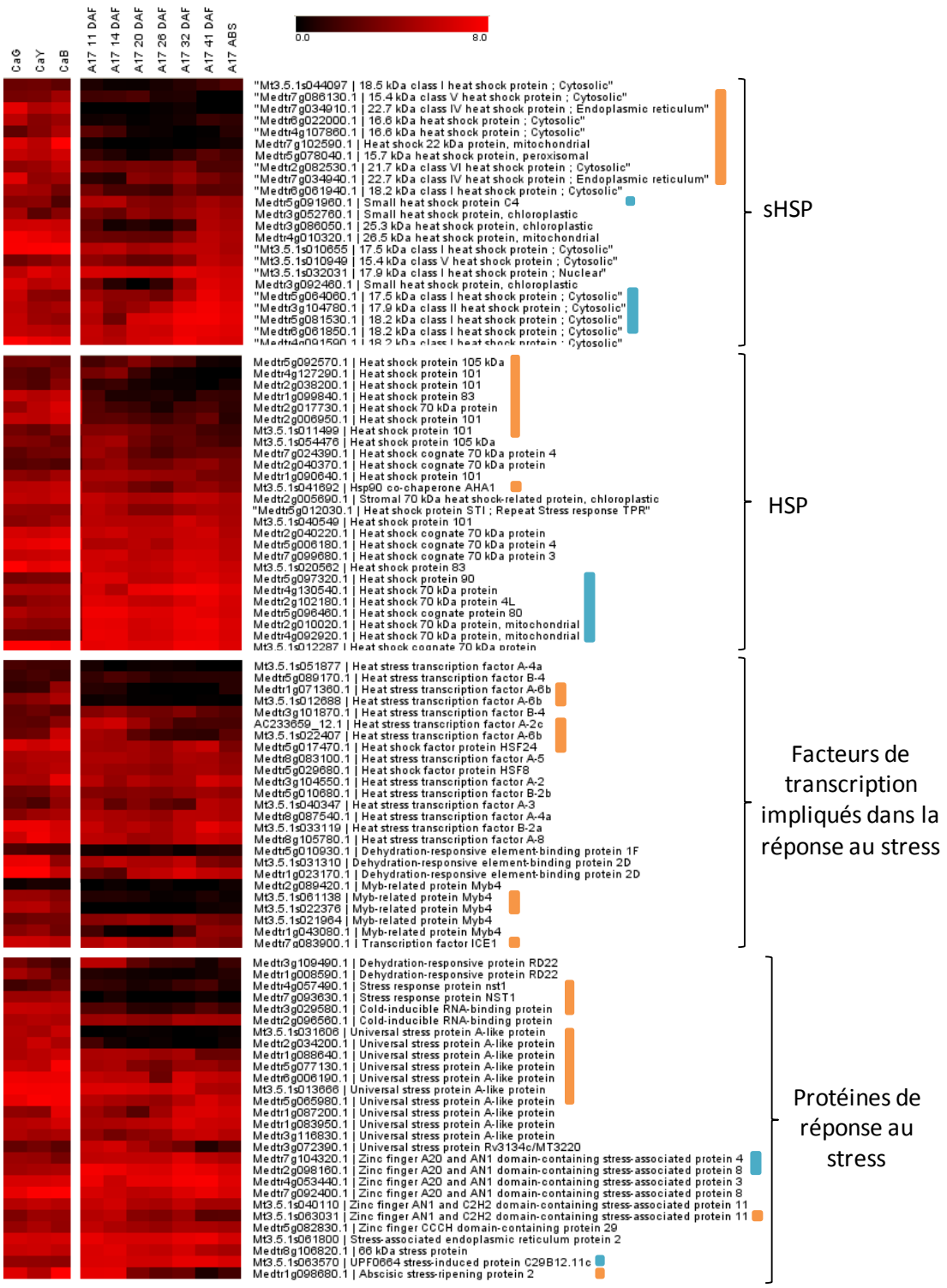


Figure 3.15. Intensité d'expression des gènes codant pour des protéines et des facteurs de transcription liés à la réponse au stress au cours du développement des graines de *C. australe* et *M. truncatula* A17.
Basé sur la présence du mot 'stress' dans les annotations MT3.5.1, les gènes codant pour les protéines de choc thermique (HSP et sHSP) et autres protéines ou facteurs de transcription liés à la réponse au stress ont été identifiées dans les transcriptomes de *C. australe* et *M. truncatula*. Les cibles ABI3 et les gènes liés à l'ABA citées précédemment n'ont pas été inclus dans cette analyse pour éviter des redondances. Ces gènes sont regroupés selon leur fonction (sHSP, HSP, facteur de transcription ou autres protéines) et par famille au sein de chaque fonction. L'intensité d'expression de ces gènes au cours du développement des graines de *C. australe* et *M. truncatula* A17 est visualisée sous la forme d'un gradient de couleur allant de noir pour une expression nulle à rouge pour une expression maximale (intensité exprimée en log). Un carré orange signifie que l'expression du gène est plus forte dans les graines matures de *C. australe* par rapport aux graines matures (41 DAF et ABS) de *M. truncatula.* A l'inverse, un carré bleu indique une expression plus forte dans les graines matures de *M. truncatula* par rapport à celles de *C. australe*.

à fait cohérent avec une activation de la signalisation liée à l'ABA dans les graines matures de *C. australe* et pas dans celles de *M. truncatula*.

Certaines familles de facteurs de transcription liés à la réponse à l'ABA ont plusieurs gènes exprimés à un niveau plus important dans les graines matures de *C. australe* par rapport à celles de *M. truncatula* : Homeobox ZIP (homologues d'AtHB5 et AtHB7 d'après les identifiants TAIR10), MYB44 et MYC2 (Figure 3.14). On peut ajouter à ces familles les gènes codant pour ABF1, ABF2, et ABI5-like. Tous ces facteurs semblent impliqués dans la réponse aux stress abiotiques (Söderman *et al.*, 1996 ; Jung *et al.*, 2008 ; Fujita *et al.*, 2011). Cependant, trois des quatre gènes codant pour AREB3 sont sous-exprimés dans la graine mature de *C. australe* par rapport au stade immature (ratio = -1,0 pour *Medtr3g010660.1* ; -1,1 pour *Medtr7g088090.1* ; -2,2 pour *Medtr8g043960.1*). De plus, les quatre gènes codant pour des facteurs de la famille DREB (Dehydration Responsive Element Binding) sont exprimés plus fortement dans les graines matures de *M. truncatula* que dans celles de *C. australe* (Figure 3.14).

Contrairement à ce que l'on pouvait attendre d'après le profil d'expression d'*ABI3*, les gènes impliqués dans la biosynthèse et la signalisation d'ABA sont plus fortement exprimés dans les graines matures de *C. australe* que dans celles de *M. truncatula*. L'ABA est un activateur majeur de la réponse générale au stress hydrique (Bartels & Sunkar, 2005). Les gènes codant pour des mécanismes impliqués dans la réponse générale au stress sont-ils activés en fin de développement chez *C. australe* ?

2.6.2-Les gènes liés à la réponse au stress

La recherche du mot 'stress' dans les annotations MT3.5 a permis de ressortir 266 gènes parmi les 15 083 gènes communs à *M. truncatula* et *C. australe*. Les cibles ABI3, protéines LEA, gènes de réponse à l'ABA ou impliqués dans la réponse antioxydante cités précédemment n'ont pas été inclus dans cette analyse pour éviter des redondances. Les 104 gènes codant pour des HSP, des sHSP, des facteurs de transcription et des protéines de réponse au stress sont présentés sur la figure 3.15 en conservant les mêmes types de classification (en catégories puis en famille) et de représentation que pour la figure précédente. Les gènes plus fortement exprimés dans les graines matures de *C. australe* par rapport aux graines matures de *M. truncatula* (41 DAF et ABS) (différence log intensité > 2) sont donc identifiés par un carré orange. Un carré bleu est utilisé pour les gènes exprimés à un niveau plus élevé dans les graines matures de *M. truncatula*.

Plusieurs gènes codant pour des sHSP et des HSP sont exprimés dans les graines matures de *C. australe* et de *M. truncatula* mais les plus fortement exprimés sont différents entre les deux espèces. Ainsi, huit gènes codant pour des sHSP non (ou très peu) exprimés chez *M. truncatula* le sont dans les cotylédons de *C. australe*. A l'inverse, les gènes *Medtr5g091960.1*, *Medtr5g064060.*1, *Medtr3g104780.1*, *Medtr5g081530.1* et *Medtr6g061850.1* codant pour des sHSP cytosoliques sont sur-exprimés en fin de développement dans les graines de *M. truncatula* alors qu'ils sont plus faiblement exprimés dans les graines matures de *C. australe* (Figure 3.15). Contrairement aux gènes codant pour les sHSP, ceux codant pour les HSP ne sont pas régulés au niveau transcriptionnel au cours du développement des graines.

Pour les autres catégories, plusieurs familles de gènes exprimés plus fortement dans les graines de *C. australe* que dans celles de *M. truncatula* ont pu être mises en évidence. C'est par exemple le cas de cinq

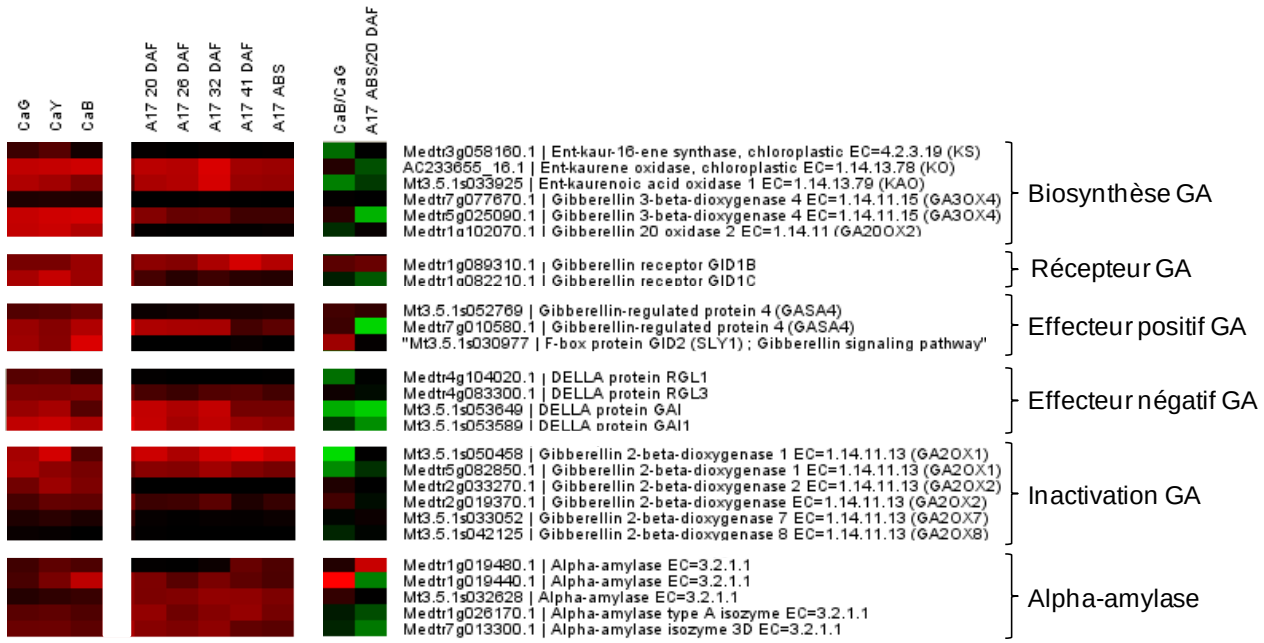


Figure 3.16. Intensité et ratio d'expression des gènes impliqués dans le métabolisme et la signalisation des gibbérellines au cours du développement des graines de *C. australe* et *M. truncatula* A17.

Basé sur la recherche de mots-clés dans les annotations MT3.5.1, les gènes codant pour les enzymes impliquées dans la biosynthèse et l'inactivation des gibbérellines (GA), la transduction du signal GA (récepteur, effecteurs positifs et négatifs) et des cibles connues des GA (alpha-amylase) ont été identifiées dans les transcriptomes de *C. australe* et *M. truncatula*. L'intensité d'expression de ces gènes au cours du développement des graines de *C. australe* et *M. truncatula* A17 est visualisée sous la forme d'un gradient de couleur allant de noir pour une expression nulle à rouge pour une expression maximale (intensité exprimée en log). Les ratios d'intensité entre stade mature (CaB et ABS) et immature (CaG et 20 DAF) des graines de *C. australe* et *M. truncatula* A17 sont également figurés. La différence d'expression est visualisée sous la forme d'un gradient de couleur allant de vert pour une sous expression chez la graine mature par rapport à la graine immature à rouge pour une sur expression et noir si l'expression est comparable entre les deux échantillons.

facteurs de transcription de la famille HSF (Heat Shock Factor ou Heat Stress Factor) de deux gènes codant pour un facteur Myb4 et pour le gène *Medtr7g083900.1* codant pour le facteur ICE1 (Inducer of CBF [Cold Binding Factor] Expression1) (Figure 3.15). Le même profil est observé pour 12 des 28 gènes codant pour des protéines de réponse au stress. On trouve en particulier, sept gènes codant pour la protéine 'Universal stress protein like A', deux codant pour la protéine de réponse au stress NST1, un codant pour une protéine ' cold inducible RNA binding', un codant pour une protéine de réponse au stress contenant un domaine 'Zinc finger AN1 and C2H2' et un codant pour une protéine 'abscisic stress ripening' (Figure 3.15).

En parallèle, seulement trois gènes sont plus exprimés dans les graines matures de *M. truncatula* que dans celles de *C. australe*. Deux d'entre eux codent pour une protéine de réponse au stress contenant un domaine 'Zinc finger AN1 and C2H2' (Figure 3.15). Aucun des 27 gènes codant pour un facteur de transcription impliqué dans la réponse au stress ne suit ce profil (aucun carré bleu pour cette catégorie sur la figure 3.15).

Cette analyse montre clairement que plusieurs gènes impliqués dans la réponse au stress (codant pour des facteurs de transcription, chaperonnes) sont exprimés plus fortement dans les graines matures de *C. australe* que dans celles de *M. truncatula*. Ce résultat est sans doute la conséquence du niveau élevé d'expression de la voie de signalisation de l'ABA dans les graines matures de *C. australe*. Dans ces conditions la biosynthèse et la voie de signalisation des GA sont-elles activées en fin de développement des graines de *C. australe* ?

2.6.3-Les gènes liés aux gibbérellines

Les gènes codant pour des enzymes impliquées dans la biosynthèse, la perception, la signalisation positive ou négative et le catabolisme des GA ont été identifiés parmi les 15 083 gènes communs à *M. truncatula* et *C. australe* par recherche de mots-clés dans les annotations MT3.5. Les profils d'expression associés sont présentés au cours du développement des graines de *C. australe* et *M. truncatula* dans la figure 3.16.

Concentrons-nous tout d'abord sur les effecteurs positifs de la biosynthèse et de la signalisation des GA. Aucun des gènes impliqués dans la biosynthèse de cette phytohormone n'est sur-exprimé dans les graines matures de *M. truncatula* et *C. australe* par rapport aux stades immatures. Cependant, le niveau d'expression des gènes *GA3ox4* et *GA20ox2*, codant respectivement une GA3-oxydase et une GA20-oxydase, est beaucoup plus élevé chez l'espèce récalcitrante tout au long du développement (Figure 3.16). Ces deux enzymes permettent l'activation des GA sous une forme active (GA_1, GA_3, GA_4 ou GA_7) à partir d'une forme inactive (GA_{12}) (Yamaguchi, 2008). Les gènes codant pour le récepteur GID1C (Gibberellin Insensitive Dwarf 1C) et pour les effecteurs positifs de la signalisation des GA GASA4 (Gibberellic Acid-Stimulated Arabidopsis 4) et GID2/SLY1 (Gibberellin Insensitive Dwarf 2 / Sleppy1) sont également fortement exprimés dans les graines matures de *C. australe* et pas dans celles de *M. truncatula* (Figure 3.16).

Les facteurs DELLA sont des répresseurs de la signalisation des GA (Peng *et al.*, 1997 ; Wen & Chang, 2002) tandis que les GA2-oxydases sont impliquées dans l'inactivation des GA (Thomas *et al.*, 1999). Les gènes codant ces protéines agissent donc négativement sur la signalisation des GA. Trois des quatre facteurs de transcription de la famille DELLA (RGL1 [Repressor of GA-Like1] ; GAI [GA-Insensitive] ; GAI1) et deux des six

gènes codant pour une GA2-oxydases (GA2ox1) sont réprimés dans les graines matures de *C. australe* (Figure 3.16). Dans les graines matures de *M. truncatula*, cette répression est également visible pour les gènes *GAI* et *GAI1* (*RGL1* n'est pas exprimé) mais pas pour *GA2ox1* (Figure 3.16). La répression de ces gènes dans les graines matures de *C. australe* est donc tout à fait compatible avec une activation du signal GA.

Les GA sont impliquées dans la mobilisation des réserves principalement par l'activation des α-amylases (Jacobsen & Beach, 1985 ; Gubler & Jacobsen, 1992 ; Lanahan *et al.*, 1992). Un des cinq gènes identifiés chez *C. australe* codant pour cette enzyme est fortement sur-exprimé dans la graine mature de l'espèce récalcitrante mais réprimé dans celle de l'espèce orthodoxe (*Medtr1g019440.1*) (Figure 3.16). Deux autres gènes sont réprimés dans les graines de *M. truncatula* et pas dans celles de *C. australe* (*Medtr1g026170.1* et *Medtr7g013300.1*).

Ces résultats transcriptomiques montrent que le niveau d'expression de plusieurs gènes est compatible avec une activation de la voie de signalisation liée aux GA dans les graines matures de *C. australe* et pas dans celles de *M. truncatula*.

Discussion

L'analyse comparative des transcriptomes de graines de *M. truncatula* et de *C. australe* repose sur la mise en place d'outils comparables entre les deux espèces : la puce à ADN ou le séquençage direct du transcriptome par des technologies haut-débit (RNA-seq). Le RNA-seq permet de quantifier et de comparer l'expression de l'ensemble des gènes entre différentes conditions pour des espèces non modèles, sans passer par les puces à ADN (Bräutigam & Gowik, 2010 ; Bourgis *et al.*, 2011). Cependant le développement récent d'une puce à ADN spécifique de *M. truncatula* au sein du laboratoire nous a incités à développer un outil équivalent pour *C. australe* afin de faciliter la comparaison avec l'espèce orthodoxe. En utilisant deux puces à ADN élaborées sur la même plate forme NimbleGen, il est possible de comparer les intensités d'expression des gènes entre différents stades de développement et pas seulement des ratios d'expression par rapport à un stade de référence. En effet, les valeurs médianes d'intensité d'expression sont du même ordre de grandeur entre les puces des deux espèces ce qui n'est pas le cas si on considère les données de la précédente puce à ADN Mt élaborée sur la plate forme Affymetrix.

Pour créer une puce à ADN spécifique de *C. australe* couvrant au maximum la diversité des gènes exprimés dans la graine, le transcriptome doit au préalable être caractérisé. Le séquençage par les technologies haut-débit (454 ou Illumina) de banques d'ADNc normalisées permet cette caractérisation, c'est-à-dire l'identification des différents gènes exprimés chez l'espèce d'intérêt. Cette approche est massivement utilisée pour caractériser les transcriptomes d'espèces non modèles (revu dans Schliesky *et al.*, 2012 ; Fu *et al.*, 2013 ; Liang *et al.*, 2013). La comparaison des données de séquençage entre banques d'ADNc normalisées et non normalisées a confirmé l'efficacité de la normalisation pour maximiser la diversité des transcrits et obtenir un répertoire de gènes le plus exhaustif possible (Bellin *et al.*, 2009 ; Yang *et al.*, 2010). Cependant, la normalisation empêche toute interprétation des niveaux d'expression des gènes. La caractérisation du transcriptome est donc qualitative mais pas quantitative. C'est pour cette raison que certains auteurs suggèrent d'augmenter la profondeur de séquençage plutôt que d'utiliser la normalisation (Vijay *et al.*, 2012). Les milliers

Tableau 3.2. Comparaison des technologies de séquençage haut-débit 454 et Illumina.
Etant donné le dynamisme de ces technologies, les performances sont en constante amélioration. Pour cette raison aucune valeur précise n'est donnée mais seulement une tendance pour percevoir les avantages de chaque technique. Plus le nombre de croix est important, plus la valeur symbolisée est importante. Tableau réalisé à partir de Shendure & Ji (2008) et Schliesky *et al.* (2012).

	454	**Illumina**
Nombre de lectures par run	x	xxxxx
Longueur des lectures	xxx	x
Nombre de bases séquencées par run	x	xxxxx
Coût par base	xxx	x
Type d'erreur	Insertion/délétion	Substitution

ou millions de lectures produites par le séquençage doivent ensuite être assemblées afin de reconstituer les transcrits. Plusieurs travaux ont montré qu'une combinaison des technologies 454 et Illumina permet d'améliorer l'assemblage du transcriptome (Wall *et al.*, 2009 ; Garg *et al.*, 2011 ; Wenger & Galliot, 2013). En effet, ces deux technologies présentent des caractéristiques complémentaires et, à l'heure actuelle, les qualités de l'une compensent les défauts de l'autre (Tableau 3.2). Le principal avantage de la technologie 454 est la longueur importante des lectures mais la technologie Illumina permet d'obtenir à moindre coût un nombre beaucoup plus élevé de lectures. Les données statistiques de notre étude confirment que le meilleur assemblage (contigs plus longs et plus nombreux) a été obtenu en combinant les lectures des deux technologies. Même si plusieurs études ont été menées pour tenter de définir les critères de qualité d'un assemblage, d'optimiser le choix des programmes et des paramètres d'assemblage (Kumar & Blaxter, 2010 ; Bräutigam *et al.*, 2011 ; Martin & Wang, 2011 ; Mundry *et al.*, 2012 ; O'Neil & Emrich, 2013), cette étape reste critique et doit encore être améliorée (Schliesky *et al.*, 2012). MIRA est un assembleur adapté aux lectures issues de banques normalisées (Kumar & Blaxter, 2010). Il est dit conservatif, c'est-à-dire qu'il n'associe pas des lectures légèrement divergentes et produit donc de nombreux contigs assez courts (Mundry *et al.*, 2012). Par conséquent le risque de contigs redondants (plusieurs contigs pour un même transcrit) est plus élevé que le risque de contigs chimériques (un contig regroupant plusieurs transcrits). Effectivement, une redondance importante des contigs a été mise en évidence après notre premier assemblage par MIRA, résultat observé également dans les travaux de Coppe *et al.* (2010). Nous avons donc ajouté une deuxième étape d'assemblage à partir des contigs formés par MIRA pour limiter cette redondance. Cette approche a montré son efficacité dans plusieurs études (Milan *et al.*, 2011 ; Coppe *et al.*, 2010, 2012).

L'assemblage des données issues des deux séquençages haut-débit nous a permis d'identifier 48 334 transcrits potentiels dans lesquels des sondes ont pu être dessinées pour constituer la puce à ADN spécifique de *C. australe*. 74,5 % d'entre eux ont pu être annotés par similarité contre des bases de données et ce travail constitue donc la première caractérisation moléculaire du transcriptome d'une graine récalcitrante, même s'il est difficile d'estimer le pourcentage du transcriptome réellement couvert par ces contigs. L'annotation spécifique de *M. truncatula* a été privilégiée afin de faciliter l'association entre les sondes des puces à ADN spécifiques de chaque espèce. L'expression de 15 083 gènes a ainsi pu être comparée au cours du développement des graines de *C. australe*, de *M. truncatula* et du mutant *Mtabi3-2*.

S'il est difficile d'associer chacun des trois stades de développement des graines de *C. australe* avec un stade de *M. truncatula*, l'analyse transcriptomique montre une très grande proximité entre les deux premiers stades de *C. australe* (gousses vertes et gousses jaunes). A ces stades, le niveau d'expression des gènes codant pour les régulateurs majeurs est comparable à celui observé au cours du début de maturation des graines de *M. truncatula*. En revanche, le stade mature (gousse marron) est caractérisé par une répression globale des gènes caractéristiques des graines, notamment *ABI3* et ses cibles. Cette étude est la première à caractériser l'expression des régulateurs majeurs de la maturation des graines orthodoxes chez une graine récalcitrante. Elle montre clairement que le profil d'expression d'*ABI3* est la différence majeure entre les graines matures de *M. truncatula* et de *C. australe*, même si d'autres gènes présentent des profils d'expression différents entre les deux espèces : *PIL5*, *VAL2-like* et *VAL3*.

L'ACP suggère que la graine mature de *C. australe* se rapproche d'une graine se préparant à la germination. La répression d'*ABI3*, de gènes codant pour les facteurs DELLA et pour les GA2-oxydases ainsi que la sur-expression des répresseurs de maturation *VAL2-like* et *VAL3*, d'un gène codant pour une α-amylase et le niveau d'expression élevé des gènes liés à la biosynthèse des GA et aux effecteurs positifs de cette phytohormone (GASA4) pourrait être une première étape dans ce processus. En effet, l'activation du signal GA marque la préparation à la germination (Holdsworth *et al.*, 2008) et la teneur en forme active de GA augmente au cours de l'imbibition de la graine, juste avant l'émergence de la radicule (Ogawa *et al.*, 2003). Le facteur RGL1 est négativement impliqué dans la germination des graines d'*A. thaliana* (Wen & Chang, 2002) et la sous-expression du gène associé dans les graines matures de *C. australe* conforte l'hypothèse d'une préparation à la germination. De même GASA4 favorise la germination en réponse aux GA (Rubinovic & Weiss, 2010) et AtHB5 est fortement exprimé 24h après imbibition chez *A. thaliana* (Johannesson *et al.*, 2003). Les gènes associés sont sur-exprimés dans la graine mature de *C. australe* mais pas dans celle de *M. truncatula*. De plus, le régulateur majeur de la maturation PIL5 intervient dans la diminution de la teneur en GA active en inhibant l'expression de *GA3ox2* et en activant celle de *GA2ox2* (Oh *et al.*, 2006 ; 2007). Le profil d'expression de *PIL5* dans les graines matures de *C. australe* (sous-expression par rapport au stade immature) et de *M. truncatula* (sur-expression par rapport au stade immature) est en accord avec une augmentation de la teneur en forme active de GA dans les graines matures de *C. australe* et pas dans celles de *M. truncatula*. VAL2 pourrait également indirectement contrôler la biosynthèse de forme active de GA comme le montre la forte sous expression de *AtGA3ox1* dans le double mutant *val1*/*val2* chez *A. thaliana* (Suzuki *et al.*, 2007). En conclusion, les profils d'expression de plusieurs gènes concordent pour émettre l'hypothèse que la voie des GA est activée dans la graine mature de *C. australe* et que ces profils sont caractéristiques d'une graine prête à germer. Cette hypothèse devra être validée par des mesures de teneur en GA dans les graines.

Cependant, certaines observations viennent nuancer cette conclusion. RGL2 est le facteur DELLA majeur impliqué dans le contrôle de la germination (Lee *et al.*, 2002) et aucun homologue n'a pu être identifié sur la puce à ADN Ca. La même remarque peut être faite concernant les gènes *GA3ox1* et *GA3ox2*, acteurs majeurs dans la biosynthèse de formes actives de GA lors de la germination des graines d'*A. thaliana* (Mitchum *et al.*, 2006). Il faut donc rester prudent en associant les profils d'expression des gènes de la figure 3.16 à une préparation à la germination. Par ailleurs, *ABI3* est réprimé seulement à la fin de la germination des graines de tomate et d'*A. thaliana* et non lors de l'imbibition (Bassel *et al.*, 2006). La répression d'*ABI3* ne serait donc pas spécifiquement impliquée dans la préparation à la germination. Enfin, même si les gènes codant pour VAL2-like et VAL3 sont sur-exprimés dans les graines matures de *C. australe*, d'autres gènes codant pour des répresseurs connus de la maturation ne le sont pas (PKL, SUA, AIP2). Le gène *AIP2* est exprimé constitutivement mais son expression est maximale lors de l'imbibition des graines d'*A. thaliana* (Zhang *et al.*, 2005). Le gène *PKL* suit le même profil d'expression, comme le montre l'accumulation de transcrit 36h après imbibition de graines d'*A. thaliana* alors même qu'elles n'ont pas encore germées (Henderson *et al.*, 2004).

Malgré l'activation de la voie des GA, la graine mature de *C. australe* exprime également de nombreux gènes de réponse à l'ABA et aux stress abiotiques par rapport à *M. truncatula*, au moins dans les cotylédons. Nous avons mis en évidence la sur-expression de gènes codant pour des SOD cytosoliques, des GST, des peroxydases, des HSP et des protéines de stress universelles dans les graines matures de *C. australe* par

rapport à celles de *M. truncatula*. Ces résultats confortent les observations faites dans le protéome stable à la chaleur de graines matures de *C. australe*. En effet, quatre polypeptides associés à une SOD et cinq polypeptides associés à des sHSP ont été identifiés (Chapitre II). De nombreuses études ont montré une tolérance accrue à divers stress chez des plantes sur-exprimant un ou plusieurs gènes codant pour une SOD (revu par Amudha & Balasubramani, 2010). De même, la sur-expression d'une GST conjointement avec une GPX améliore la germination et la croissance de plantules en condition de stress salin et froid chez le tabac (Roxas *et al.*, 1997). Des gènes codant pour des sHSP sont sur-exprimés en réponse au stress hydrique chez diverses espèces comme le tournesol (Coca *et al.*, 1996) ou l'amandier (Campalans *et al.*, 2001) et la sur-expression de HSP17.7 chez le riz augmente la tolérance à la sécheresse (Sato & Yokoya, 2008). L'expression de ces protéines de stress est activée par des facteurs de transcription de type HSF, également sur-exprimés dans les graines matures de *C. australe*. Les facteurs HSF ne régulent pas seulement l'expression des HSP mais jouent un rôle intégrateur dans la réponse au stress en général en activant par exemple l'expression de l'APX chez *A. thaliana* (Panchuk *et al.*, 2002). Ils pourraient même percevoir directement le stress oxydatif (Miller & Mittler, 2006). Hormis les HSF, deux autres facteurs de transcription ont été mis en évidence dans les graines matures de *C. australe* : ICE1 et Myb4. Les facteurs CBF (activés entre autres par ICE1) et Myb4 sont impliqués dans la réponse aux stress abiotiques, notamment thermique (Chinnusamy *et al.*, 2003 ; Coraggio & Tuberosa, 2004 ; Agarwal *et al.*, 2006). Dans le chapitre II, nous avons mis en évidence l'accumulation de déhydrines dans les cotylédons et proposé que ces protéines participent à la protection de l'axe embryonnaire contre le déficit hydrique après abscission de la graine. La sur-expression dans les cotylédons matures de *C. australe* d'une gamme assez large de gènes associés à la réponse aux stress abiotiques conforte cette hypothèse.

Le rôle de l'ABA dans la réponse au stress est bien établi (Iuchi *et al.*, 2001 ; Fujita *et al.*, 2011) et il n'est pas étonnant de retrouver de nombreux gènes impliqués dans la biosynthèse et la signalisation de l'ABA fortement exprimés dans les graines matures de *C. australe* comparé à celle de *M. truncatula*. Des études menées en collaboration en Afrique du Sud montrent que la teneur en ABA augmente après abscission dans les cotylédons et les axes embryonnaires de *C. australe* (Chathuri, données non publiées). Cependant, l'ABA est également la phytohormone majeure contrôlant la maturation des graines et réprimant la germination. De par son antagonisme avec les GA, il est surprenant de retrouver à la fois la voie de l'ABA et la voie des GA activées dans la graine mature de *C. australe*. Pour expliquer cette apparente contradiction on peut émettre l'hypothèse que la voie des GA serait activée pour permettre la dégradation de l'amidon (environ 75% des réserves de *C. australe* (Thorburn *et al.*, 1987) via les α-amylases pour assurer l'activité métabolique importante. L'activation de la biosynthèse de l'ABA et de la signalisation liée à cette phytohormone empêcherait la germination. Cependant, tous les gènes de réponse à l'ABA ne sont pas sur-exprimés chez la graine récalcitrante. En effet, les gènes spécifiques de la maturation (*ABI3*, *ABI5*, *PIL5*) sont sous-exprimés dans la graine mature par rapport au stade immature et seuls les gènes de réponse au stress induit par l'ABA sont sur-exprimés. On retrouve dans cette dernière catégorie des kinases de la famille SnRK2 telles que SAPK3 et SAPK7 (Kobayashi *et al.*, 2004 ; Kulik *et al.*, 2011) et des facteurs de transcription tels que ABF1, ABF2/AREB1, Homeobox ZIP HB7, MYB44 et MYC2 (Söderman *et al.*, 1996 ; Choi *et al.*, 2000 ; Abe *et al.*, 2003 ; Fujita *et al.*, 2005 ; Yoshida *et al.*, 2010 ; Seo *et al.*, 2012). Chose intéressante, trois des quatre gènes codant pour le facteur AREB3 sont sous-exprimés en fin de développement chez *C. australe*. Mais celui-ci est impliqué dans la maturation des

graines et non dans la réponse au stress contrairement aux autres facteurs AREB (Uno *et al.*, 2000 ; Benshimen *et al.*, 2005 ; Fujita *et al.*, 2013). Les facteurs de transcription de la famille DREB sont moins exprimés chez *C. australe* que chez *M. truncatula*. Bien que la mention 'Abscisic acid signaling pathway' soit indiquée dans l'annotation TAIR10, les facteurs DREB1 et DREB2 sont connus pour être activés en réponse au stress indépendamment de l'ABA (Shinozaki *et al.*, 2003). Des études plus récentes ont néanmoins montré que certains de ces gènes répondent à l'ABA (revu dans Agarwal & Jha, 2010 ; Fujita *et al.*, 2011) dont *DREB1D* et *DREB2C* identifiés dans notre étude (Knight *et al.*, 2004 ; Lee *et al.*, 2010). La raison de la faible expression de ces gènes chez *C. australe* reste à déterminer.

Les résultats transcriptomiques suggèrent donc que les graines de *C. australe* et de *M. truncatula* ne suivent pas les mêmes voies de régulation en fin de développement. Plusieurs indices nous laissent penser que les graines récalcitrantes se préparent à la germination, via la répression de gènes de maturation et l'activation de la voie des GA. Mais ces graines se préparent également au stress modéré à venir après l'abscission, via l'activation de gènes de réponse à l'ABA : facteurs de transcription (ABF, MYB44, MYC2), kinases (membres de la famille SnRK2), déhydrines, chaperonnes, peroxydases, GST. L'expression de ces gènes a été mise en évidence dans les cotylédons. Ce tissu, qui occupe plus de 95 % du volume de la graine, entoure l'axe embryonnaire et forme une première protection mécanique pour retarder sa déshydratation. L'activation des gènes de réponse au stress dans les cotylédons constitue donc une protection supplémentaire de l'axe embryonnaire qui vient s'ajouter à la protection mécanique.

Quoi qu'il en soit, la fin de développement des graines de *C. australe* est indépendante d'ABI3, contrairement à *M. truncatula*. D'autres voies de signalisation sont activées en fin de développement chez la graine récalcitrante, comme celles des GA. Nos résultats ne vont donc pas dans le sens d'un développement tronqué chez les graines de *C. australe* comme il a été suggéré chez des espèces récalcitrantes d'origine tempérée (Finch-Savage & Blake, 1994 ; Farrant *et al.*, 1997). Chez ces espèces (*Quercus robur* et *Aesculus hippocastanum*), le développement ressemble à celui des graines orthodoxes mais l'abscission a lieu avant la phase de maturation tardive. Cependant, il faut rester prudent en comparant ces travaux avec le notre car notre conclusion s'appuie essentiellement sur des données transcriptomiques non disponibles dans les travaux de Finch-Savage & Blake (1994) et Farrant *et al.* (1997).

Par ailleurs, cette analyse transcriptomique a permis de bien différencier le développement des graines du mutant *Mtabi3-2* de celui des graines de *C. australe*. Chez le mutant l'expression du gène *ABI3* et d'environ la moitié de ses cibles n'est pas éteinte en fin de développement mais au contraire retardée en début de maturation. Pourtant la durée du développement des graines de la pollinisation à l'abscission reste la même que pour les graines sauvages. Par conséquent, les gènes impliqués dans l'acquisition de la TD (*LEA* par exemple) sont exprimés trop tardivement pour que les polypeptides soient accumulés dans les graines matures. Ce retard global d'expression des gènes cibles d'ABI3, aboutissant à un développement tronqué, explique sans doute la sensibilité à la dessiccation des graines *Mtabi3-2*. Au contraire, chez *C. australe*, ces gènes cible sont exprimés au début du développement puis réprimés en fin de développement. Une acquisition partielle et transitoire de la TD pourrait être envisagée au stade mi-mature dans les cotylédons de *C. australe*.

Chapitre IV

Analyse fonctionnelle des homologues *CaABI3* et *CaABI3-like* : comparaison avec *MtABI3* et *AtABI3*.

Approche et principaux résultats

Dans le chapitre III, nous avons montré que l'expression du gène *ABI3* est réprimée chez les graines de *C. australe* en fin de développement, alors qu'il reste fortement exprimé tout au long de la maturation des graines de *M. truncatula*. L'objectif de ce chapitre est de déterminer si les homologues identifiés chez l'espèce récalcitrante ont les mêmes fonctions que ceux identifiés chez *M. truncatula* et *A. thaliana*. Deux approches ont été mises en place pour tester la fonctionnalité de *CaABI3* : la comparaison des cibles potentielles activées par expression ectopique de *CaABI3* et de *MtABI3* dans des racines de *M. truncatula* et la complémentation du mutant *abi3-5* d'*A. thaliana*. Les résultats majeurs mis en évidence peuvent être résumés en quatre points :

- Deux gènes très similaires entre eux et avec *MtABI3* ont été clonés chez *C. australe* : *CaABI3* et *CaABI3-like*.
- Les gènes sur-exprimés par CaABI3 et MtABI3 dans un système d'expression ectopique sont très divergents. Alors que la sur-expression de *MtABI3* dans les racines de *M. truncatula* active l'expression de 90 % des 109 cibles ABI3 de *M. truncatula* identifiées dans le chapitre précédent, la sur-expression de *CaABI3* dans les mêmes conditions n'active l'expression que de 17 % d'entre elles.
- Contrairement à ce qui est observé pour *CaABI3*, la sur-expression de *CaABI3-like* dans les racines de *M. truncatula* active la majorité des 109 cibles ABI3 de *M. truncatula*. Le profil d'expression obtenu est semblable à celui observé en sur-exprimant *MtABI3* dans les mêmes conditions.
- *CaABI3* ne complémente pas le mutant *abi3-5*. Les graines mutantes transformées par *CaABI3* restent vertes à maturité, ne sont pas dormantes et sont insensibles à l'ABA. Le phénotype sauvage (dégradation de la chlorophylle, sensibilité à l'ABA, et dormance) est restauré dans les graines mutantes complémentées par *AtABI3*.

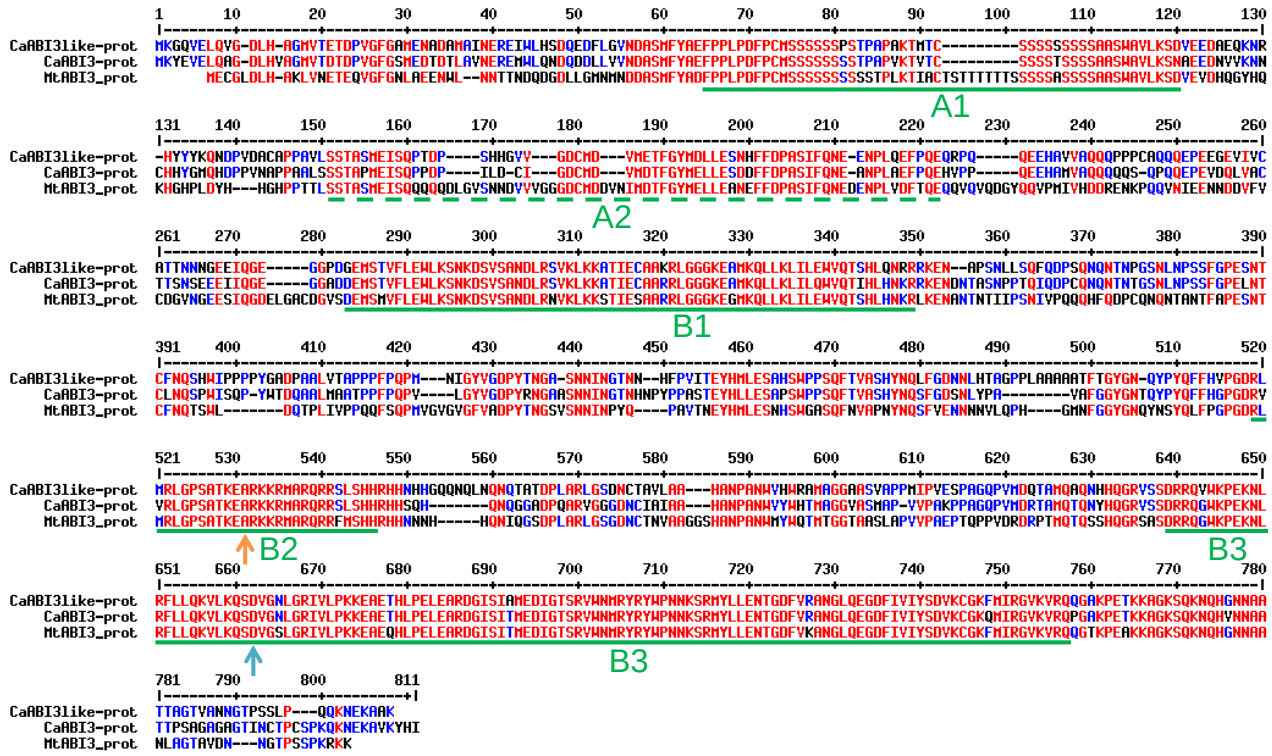


Figure 4.1. Alignement des séquences protéiques de MtABI3 (génotype R108), CaABI3 et CaABI3-like.
Les séquences génomiques et les ADNc des homologues *ABI3* chez *M. truncatula* et *C. australe* ont été clonés par PCR dégénérée. Les séquences protéiques intégrales ont été déduites de ces clonages et alignées par le logiciel MultAlin à l'aide de la matrice Blosum 62. Les domaines conservés A1, B1, B2 et B3 ont été positionnés. Un cinquième domaine conservé a été ajouté en pointillé. Ce domaine n'est pas décrit dans la littérature excepté chez le pois où un deuxième domaine acide (A2) a été mis en évidence (Gagete *et al.*, 2009). Chez le mutant *abi3-7*, l'alanine située au niveau de la flèche orange (position 531 sur cet alignement) est substituée par une thréonine. Chez le mutant *abi3-1*, l'aspartate situé au niveau de la flèche bleue (position 662 sur cet alignement) est substitué par une asparagine (Bies-Ethève *et al.*, 1999).

Tableau 4.1. Pourcentage de similarité entre les séquences protéiques de MtABI3, AtABI3, CaABI3 et CaABI3-like.
Les pourcentages ont été déterminés avec le programme MatGAT (Matrix Global Alignment Tool) (Campanella *et al.*, 2003) en utilisant la matrice Blosum50.

similarité en AA (%)	Protéine entière	Domaine A1	Domaine B1	Domaine B2	Domaine B3
CaABI3 / CaABI3-like	84,7	95,7	97	100	97,5
MtABI3 / CaABI3	74,3	76,8	94	96,4	98,3
MtABI3 / CaABI3-like	75,3	75	94	96,4	97,5
AtABI3 / CaABI3	61,5	80,8	91	89,3	97,5
AtABI3 / CaABI3-like	63,2	76,9	91	89,3	98,3
AtABI3 / MtABI3	59,8	80,4	82,5	92,9	97,5

1-Clonage des homologues du gène ABI3 chez *C. australe*

Dans la base de données Genbank, les séquences (complètes ou partielles) du gène *ABI3* de 18 espèces sont disponibles mais ce gène n'avait pas encore été séquencé chez *M. truncatula* au moment de l'initiation de ces travaux. Ces séquences ainsi que l'EST disponible chez *M. truncatula* (TC193312 MtGI11) ont été alignés pour dessiner des couples d'amorces dans les régions les plus conservées. Par PCR dégénérée, une portion des homologues *ABI3* chez *M. truncatula* et *C. australe* a ainsi été clonée à partir d'ADN génomique. Dans un deuxième temps, la séquence complète a été obtenue par PCR inverse à partir de cette région initiale. Alors qu'un gène de 3 459 pb a été cloné chez *M. truncatula* A17 et R108, deux gènes présentant une forte similarité ont été clonés chez *C. australe* : *CaABI3* (3 047 pb) et *CaABI3-like* (3 074 pb). Les CDS correspondant ont également été clonés à partir d'ARN de graines afin de déterminer la taille et la position d'éventuels introns. En effet, cinq introns ont été identifiés au niveau du domaine B3 chez plusieurs espèces comme *A. thaliana* (TAIR10), le pois (Gagete *et al.*, 2009), la tomate (Gao *et al.*, 2013). Ces cinq introns sont également présents chez *M. truncatula* et *C. australe*. Les CDS correspondant ont une taille de 2 313 pb (*MtABI3*), 2 265 pb (*CaABI3*) et 2 283 pb (*CaABI3-like*).

Les CDS de *CaABI3*, *CaABI3-like* et de *MtABI3* ont été traduit *in silico* pour quantifier la similarité entre ces trois protéines. Une très grande similarité est visible entre ces trois séquences en alignant les séquences protéiques déduites à l'aide du logiciel MultAlin, notamment au niveau des quatre domaines conservés (Figure 4.1). Un cinquième domaine très conservé est mis en évidence entre les positions 186 et 211. Il correspond à un deuxième domaine acide (A2) mis en évidence par Gagete *et al.* (2009) chez le pois. Le pourcentage de similarité global (séquence protéique entière) entre les deux protéines CaABI3 et CaABI3-like est de 84,7 % (Tableau 4.1). Il s'élève même respectivement à 95,7 %, 97 %, 100 % et 97,5 % en considérant uniquement les domaines conservés A1, B1, B2 et B3. Il apparait donc clairement que les gènes *CaABI3* et *CaABI3-like* codent tous les deux pour un facteur ABI3. Le pourcentage de similarité global avec MtABI3 est semblable pour CaABI3 et CaABI3-like, respectivement 74,3 % et 75,3 %. Il en est de même en considérant uniquement les domaines conservés : respectivement 76,8 % et 75 % pour A1 ; 94 % pour B1 ; 96,4 % pour B2 ; 98,3 % et 97,5 % pour B3 (Tableau 4.1). Les trois domaines basiques (B1, B2 et B3) sont extrêmement bien conservés entre les deux espèces. Le degré de conservation est moindre pour le domaine acide A1, notamment à cause d'une délétion du motif TSTTTTTTS chez *C. australe* (Figure 4.1). La similarité des séquences CaABI3 et CaABI3-like avec AtABI3 est globalement inférieure qu'avec MtABI3 (Annexe 6) mais la conservation au niveau des quatre domaines reste très élevée, notamment pour le domaine B3 (Tableau 4.1). Il est intéressant de noter que la similarité AtABI3/CaABI3 (61,5 %) ou AtABI3/CaABI3-like (63,2 %) est globalement plus importante que la similarité AtABI3/MtABI3 (59,8 %). Par ailleurs, les acides aminés substitués dans les mutants *abi3-1* ($D_{580}N$; phénotype faible) et *abi3-7* ($D_{580}N$ et $A_{458}T$; phénotype fort) (Bies-Ethève *et al.*, 1999) sont conservés chez *MtABI3*, *CaABI3* et *CaABI3-like* (Figure 4.1). Ces résultats suggèrent que la structure primaire des protéines CaABI3 et CaABI3-like est assez proche de celle de MtABI3 et AtABI3. Ces protéines ont-elles les mêmes fonctions ?

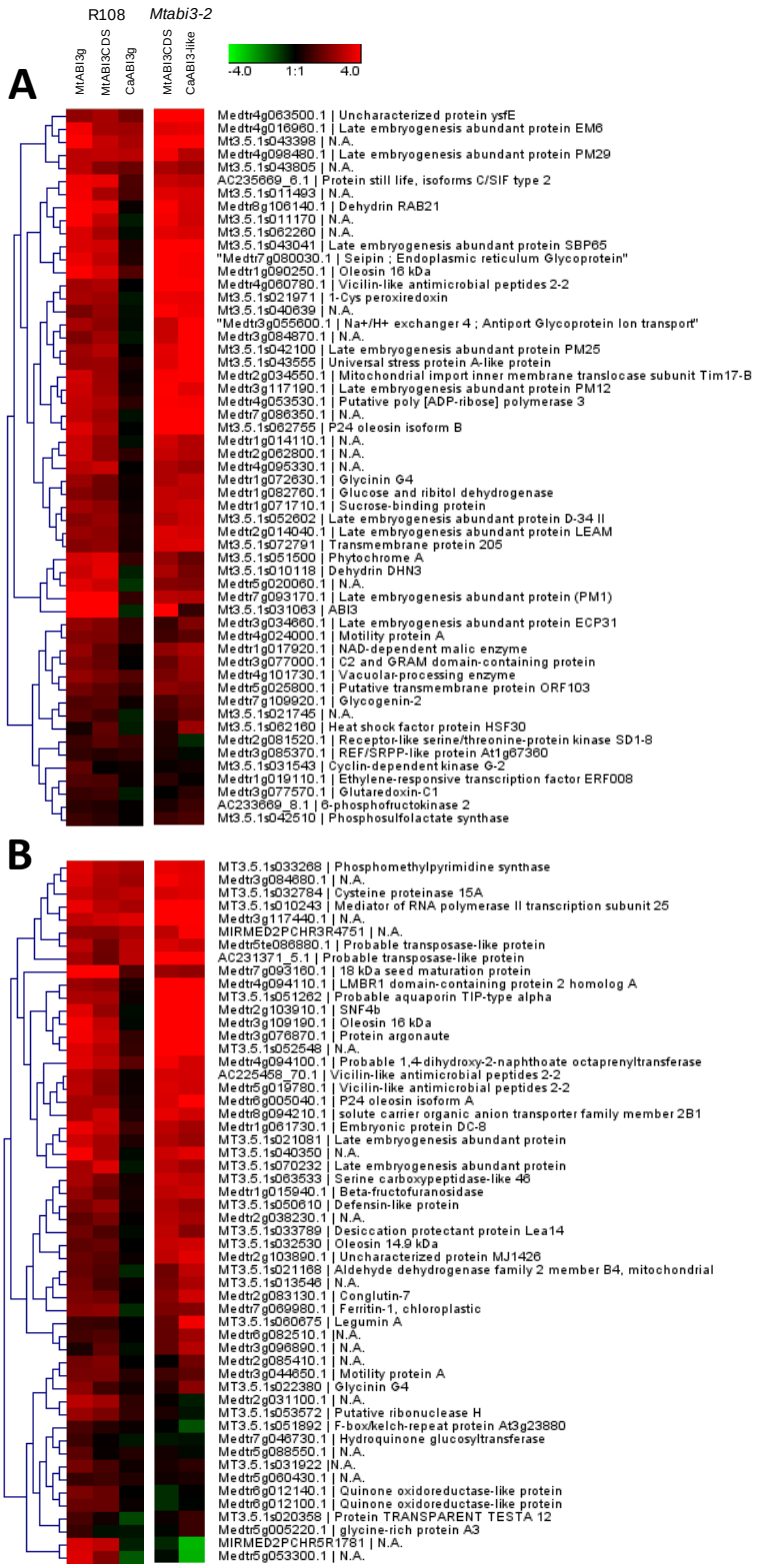


Figure 4.2. Profil d'expression des gènes cibles de MtABI3 en réponse à la sur-expression de *CaABI3g*, de *MtABI3g*, de *MtABI3$_{CDS}$* dans les racines R108 et de *CaABI3-like* et *MtABI3$_{CDS}$* dans les racines de M*tabi3-2*.
Les 55 gènes cibles de MtABI3 identifiés dans le transcriptome de *C. australe* **(A)** et les 54 gènes cibles de MtABI3 non identifiés sur la puce Ca **(B)** ont été classés par clustering hiérarchique (HCL) à partir des valeurs de ratios d'intensité entre racines R108 ou *Mtabi3-2* transformées par *ABI3* (*CaABI3g*, *CaABI3-like*, *MtABI3$_{CDS}$*, *MtABI3g*) et racines R108 ou *Mtabi3-2* transformées par le vecteur vide. La différence d'expression est visualisée sous la forme d'un gradient de couleur allant de vert pour une sous-expression dans la racine transformée par *ABI3* par rapport à la racine transformée par le vecteur vide ; rouge pour une sur-expression et noir si l'expression est comparable entre les deux échantillons. N.A. = non annoté.

2-Expression ectopique de *CaABI3* et *CaABI3-like*

2.1-Principe de l'expression ectopique dans les racines de *M. truncatula*

Chez *M. truncatula*, la transformation stable par *Agrobacterium tumefaciens* impliquant la régénération de plantes entières à partir de cals est longue et laborieuse. La transformation de racine par *Agrobacterium rhizogenes* permet d'obtenir beaucoup plus rapidement des plantes composites dans lesquelles le transgène est exprimé uniquement dans les racines (Boisson-Dernier *et al.*, 2001 ; Crane *et al.*, 2006). L'expression ectopique d'un gène dans les racines de *M. truncatula* a été utilisée avec succès pour valider rapidement la fonction de différents gènes (Elfstrand *et al.*, 2005 ; Pang *et al.*, 2008 ; Verdier *et al.*, 2013). Dans cet objectif, les séquences génomiques *CaABI3g*, *CaABI3-like* et *MtABI3g* ainsi que le CDS de *MtABI3* (nommé $MtABI3_{CDS}$ par la suite) ont été constitutivement sur-exprimés par le promoteur 35S dans des racines de *M. truncatula* par transformation avec *A. rhizogenes*. Des racines témoins ont également été transformées par le vecteur vide (nommé GFP par la suite). Les ARN des racines transformées sont extraits et hybridés sur la puce à ADN de Medicago (ABI3 contre GFP) afin d'identifier les gènes sur-exprimés et sous-exprimés en réponse aux différentes constructions. Pour s'affranchir de l'expression éventuelle du gène *MtABI3* endogène, des racines du mutant *Mtabi3-2* ont également été transformées par ces différentes constructions. Suite à des problèmes d'extraction d'ARN, toutes les conditions n'ont pas pu être hybridées sur la puce à ADN à l'heure actuelle. Nous disposons pour le moment des listes de gènes répondant à la sur-expression de *CaABI3g*, *MtABI3g* et $MtABI3_{CDS}$ dans les racines R108 et répondant à la sur-expression de *CaABI3-like* et $MtABI3_{CDS}$ dans les racines du mutant *Mtabi3-2*. La comparaison des résultats pour $MtABI3_{CDS}$ permet donc de vérifier un éventuel effet du fond génétique de la racine.

Trois répétitions biologiques de racines transformées ont été réalisées pour chaque construction. Elles ont toutes les trois été analysées pour l'obtention de la liste de gènes sous-exprimés et sur-exprimés par les gènes *MtABI3g*, *CaABI3-like* et $MtABI3_{CDS}$. Cependant, seules deux des trois répétitions biologiques pour *CaABI3g* ont pu être exploitées au niveau du scan. La puissance statistique de l'analyse a donc été affaiblie et seulement 65 gènes sur les 102 123 sondes de la puce Mt sont sur-exprimés ou sous-exprimés à un seuil FDR 5% en réponse à *CaABI3g*. Le seuil de stringence a donc été abaissé et un gène a été considéré comme sur-exprimé si le ratio *ABI3*/GUS est supérieur à 1 et si la p-value (non corrigée) est inférieure à 1 %. Ces résultats sont donc préliminaires et seront renforcés rapidement par d'autres répétitions.

2.2-Profil d'expression des cibles ABI3

109 sondes associées à des gènes potentiellement cibles de MtABI3 ont été définis dans le chapitre précédent, 55 identifiés dans le transcriptome de graine de *C. australe* et 54 absents de la puce à ADN Ca. Ce jeu de données constitue un bon moyen de contrôler la fonctionnalité des gènes *CaABI3* et *CaABI3-like* par rapport à *MtABI3*. Le profil d'expression de ces 109 gènes en réponse à la sur-expression de *MtABI3g*, *CaABI3g*, *CaABI3-like* et $MtABI3_{CDS}$ dans les racines de R108 et du mutant *Mtabi3-2* est présenté dans la figure 4.2.

Sans surprise, respectivement 98 et 96 de ces 109 gènes sont significativement sur-exprimés en réponse à la sur-expression de *MtABI3g* et de $MtABI3_{CDS}$ dans les racines R108. Cependant, ce résultat signifie

néanmoins qu'environ 10 % de ces gènes définis comme étant des cibles d'ABI3 ne sont pas sur-exprimés dans ce système ectopique. En comparant les résultats obtenus pour la sur-expression de *MtABI3*$_{CDS}$ dans les racines R108 et *Mtabi3-2* on constate que les profils d'expression des gènes cibles ABI3 sont comparables dans les deux fonds génétiques, même si 13 gènes ne sont pas significativement sur-exprimés dans les racines *Mtabi3-2* alors qu'ils le sont dans les racines R108. Le fond génétique de la racine a donc un effet limité et les données obtenues pour la sur-expression de *CaABI3-like* dans les racines *Mtabi3-2* peuvent être comparées à celles obtenues pour la sur-expression de *CaABI3g* et *MtABI3g* dans les racines R108.

Par rapport à *MtABI3g*, seulement 19 des 109 gènes potentiellement cibles de MtABI3 sont sur-exprimés par *CaABI3g* (Figure 4.2). Par contre, la sur-expression de *MtABI3*$_{CDS}$ et de *CaABI3-like* dans les racines du mutant *Mtabi3-2* aboutit à un profil comparable pour les deux gènes. En effet, respectivement 83 et 78 des 109 gènes cibles sont activés par MtABI3$_{CDS}$ et CaABI3-like avec 92 % de cibles communes entres les deux gènes (Figure 4.2A et 4.2B).

CaABI3 et CaABI3-like ont-ils des cibles spécifiques ? Si c'était le cas, nous pourrions imaginer que chacun des gènes active une partie des gènes cibles ABI3 et que l'expression des deux gènes permettrait d'activer l'ensemble du régulon ABI3 de façon complémentaire. En regardant plus en détails les 19 gènes sur-exprimés par CaABI3, parmi lesquels *EM6* et *PM29* codant pour des protéines LEA, on remarque que ces gènes sont également sur-exprimés par CaABI3-like (Figure 4.2A). Ce résultat nous permet donc d'exclure l'hypothèse d'une complémentarité fonctionnelle entre *CaABI3* et *CaABI3-like*. Par ailleurs, il est intéressant de noter que CaABI3-like est capable d'activer les gènes potentiellement cibles de MtABI3 non identifiés dans le transcriptome de *C. australe* (Figure 4.2B).

Le gène *ABI3* apparait fortement sur-exprimé dans les racines R108 et *Mtabi3-2* transformées avec les constructions *MtABI3g* et *MtABI3*$_{CDS}$ (Figure 4.2A). Ce résultat est tout à fait logique et traduit une sur-expression efficace du gène *MtABI3* précédé du promoteur 35S dans ces constructions. Il est également tout à fait logique que ce gène n'apparaissent pas sur-exprimé dans les racines transformées par les constructions *CaABI3g* et *CaABI3-like* car la sonde spécifique de la puce Mt hybride une région non conservée du transcrit *MtABI3*. Cette sonde n'est donc pas capable d'hybrider les transcrits *CaABI3* et *CaABI3-like* produits par sur-expression avec le promoteur 35S. Par RT-PCR, nous avons confirmé que ces deux gènes sont fortement exprimés dans les racines de *M. truncatula*.

En conclusion, l'expression ectopique de *MtABI3*, *MtABI3*$_{CDS}$ et *CaABI3-like* dans les racines de *M. truncatula* montre que ces gènes activent l'expression de la majorité des cibles *ABI3* et sont donc fonctionnels, ce qui n'est pas le cas de *CaABI3*.

3-Complémentation du mutant *abi3-5* d'*A. thaliana*

Le mutant *abi3-5* a été crée par mutagenèse chimique au diepoxybutanol dans l'écotype Landsberg (Ler) (Ooms *et al.*, 1993). La délétion d'une cytosine en position 1 588 induit la formation d'un codon stop prématuré entre le domaine B1 et B2. La protéine tronquée produite est dépourvue des domaines B2 et B3, il n'est donc

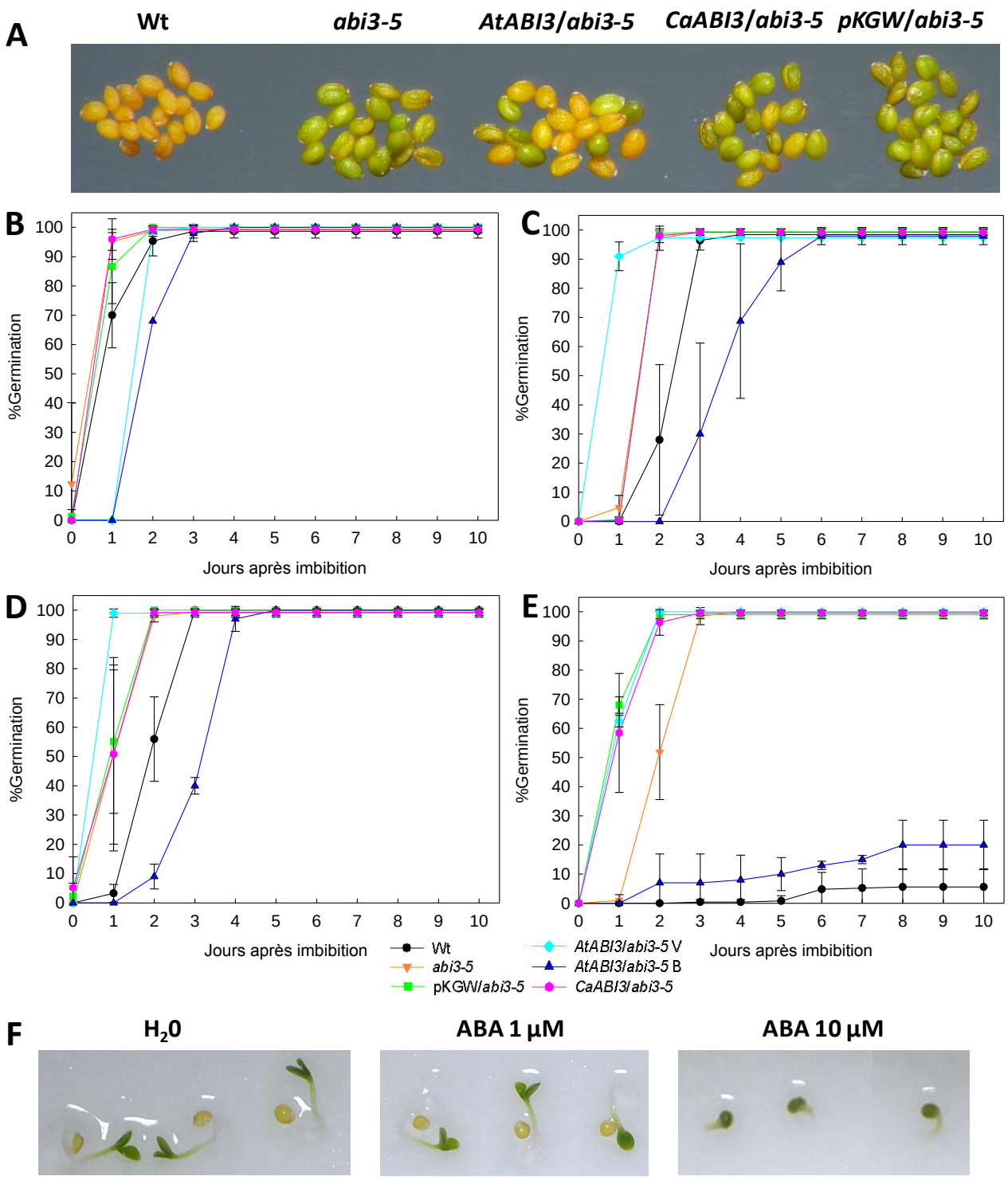


Figure 4.3. Couleur, dormance et sensibilité à l'ABA de graines matures sauvages, mutantes *abi3-5* et mutantes complémentées par *AtABI3*, *CaABI3* et le vecteur vide.

(A) Couleur de graines matures sauvages (Wt), mutantes *abi3-5* et mutantes complémentées par *AtABI3* (*AtABI3*/*abi3-5*), *CaABI3* (*CaABI3*/*abi3-5*) et le vecteur vide (pKGW/*abi3-5*) à l'ouverture de la silique. Les séquences codantes d'*AtABI3* et *CaABI3* sont placées sous le contrôle du promoteur d'*AtABI3* (1 883 pb en amont de la séquence codante) ; photo Joseph Ly-Vu. **(B - C)** Germination sur eau de ces différents lots de graines à 20°C et 16h de photopériode précédée **(B)** ou non **(C)** d'une stratification trois jours à 4°C. Les données présentées sont des moyennes de trois à cinq réplications de 40 à 50 graines. Pour *AtABI3/abi3-5*, les graines vertes (V) et beiges (B) ont été analysées séparément. **(D - E)** Germination sur acide abscissique (ABA) 1 µM **(D)** ou 10 µM **(E)** de ces différents lots de graines à 20°C et 16h de photopériode précédée d'une stratification trois jours à 4°C. Les données présentées sont des moyennes de trois à cinq réplications de 40 à 50 graines. Pour *AtABI3/abi3-5*, les graines vertes (V) et beiges (B) ont été analysées séparément. **(F)** Plantules de pKGW/*abi3-5* observées trois jours après imbibition sur eau, ABA 1 µM et 10 µM.

pas surprenant que le phénotype de ces graines soit fortement altéré dans de nombreux aspects physiologiques et moléculaires : sensibilité à l'ABA, tolérance à la dessiccation, dormance, dégradation de chlorophylle, accumulation de RFO, de protéines de réserves et de protéines LEA (Ooms *et al.*, 1993 ; Bies-Ethève *et al.*, 1999). Nous avons choisi certains aspects de ce phénotype pour tester la capacité de *CaABI3* à restaurer le phénotype sauvage. Ne disposant pas à l'heure actuelle de la séquence promotrice chez *C. australe*, la séquence codante *CaABI3* a été placée en aval de la séquence promotrice d'*AtABI3* (séquence de 1 883 pb en amont de la séquence codante d'*AtABI3*). La complémentation par *AtABI3* (séquence promotrice + séquence codante) et par le vecteur vide (pKGW) sont respectivement utilisées comme témoins positif et négatifs de complémentation.

3.1- Dégradation de la chlorophylle

Le gène *ABI3* contrôle la dégradation de la chlorophylle à la fin du développement des graines (Delmas *et al.*, 2013). Les graines matures sauvages d'*A. thaliana* (Wt) sont donc beiges à la déhiscence des siliques alors que celles du mutant *abi3-5* restent vertes (Figure 4.3A). Les graines complémentées par *AtABI3* (*AtABI3*/*abi3-5*) sont beiges pour la majorité d'entre elles. Cependant environ 10 à 15 % de graines vertes sont également présentes (Figure 4.3A). Des graines vertes et beiges sont retrouvées au sein d'une même silique. Ces deux types de graines ont été analysés séparément pour les autres aspects du phénotype. Les graines complémentées par *CaABI3* et pKGW (respectivement *CaABI3*/*abi3-5* et pKGW/*abi3-5*) sont toutes vertes (Figure 4.3A).

3.2- Dormance

La totalité des graines *abi3-5*, *CaABI3/abi3-5* et pKGW/*abi3-5* germent sur eau en 24h après trois jours de stratification à 4°C (Figure 4.3B). Sans stratification, il faut attendre deux jours pour obtenir 100 % de germination avec ce lot de graines (Figure 4.3C). De plus, la germination est très homogène puisque presque toutes les graines germent entre le premier et le deuxième jour (Figure 4.3C). Il y a donc peu de différences entre les graines stratifiés ou non, ce qui signifie qu'elles ne sont pas dormantes à la récolte. Les graines beiges *AtABI3*/*abi3-5* germent beaucoup plus lentement et de façon beaucoup plus hétérogène sans stratification. En effet, trois jours sont nécessaires pour faire germer la totalité des graines après trois jours de stratification à 4°C (Figure 4.3B). Sans stratification, les premières graines germent trois jours après imbibition et les dernières six jours après imbibition (Figure 4.3C). Ce profil est caractéristique de graines présentant un certain degré de dormance. En revanche, les graines vertes *AtABI3/abi3-5* se comportent comme les graines mutantes. Nos données suggèrent même que ces graines germent plus rapidement sans stratification (90 % de graines germées 24h après imbibition) (Figure 4.3C) qu'après stratification (la totalité des graines germent entre 24h et 48h après imbibition) (Figure 4.3B).

3.3- Sensibilité à l'ABA

Après 3 jours de stratification à 4°C, la totalité des graines sauvages (Wt) germent en deux jours sur eau (Figure 4.3B). Après stratification mais en présence de 1 µM d'ABA, il faut trois jours pour atteindre 100 %

de germination (Figure 4.3D) et la germination est complètement inhibée par 10 µM d'ABA (Figure 4.3E). Les graines de type sauvage sont donc très sensibles à l'ABA. Deux jours après imbibition, la totalité des graines *abi3*-5, *CaABI3/abi3-5* et pKGW/*abi3-5* a germé en présence de 1 µM et 10 µM d'ABA (Figure 4.3D et 4.3E). La germination est simplement légèrement retardée par rapport à la condition eau après stratification où la totalité des graines germent en une journée (Figure 4.3B). Ces graines sont donc très peu sensibles à l'ABA, au moins pour la germination. Les graines *AtABI3*/*abi3-5* vertes et beiges ont été analysées séparément. Les graines vertes sont insensibles à l'ABA comme les graines mutantes *abi3-5* mais les graines beiges se comportent comme les graines de type sauvage (Figure 4.3D et 4.3E). La germination est retardée en présence de 1 µM d'ABA et fortement inhibée en présence de 10 µM d'ABA (20 % de graines germées 10 jours après imbibition).

Si la germination des graines mutantes est très peu affectée par l'ABA, cette phytohormone ralentit fortement la croissance de l'hypocotyle (à 10 µM mais pas à 1 µM) comme l'atteste la figure 4.3F où sont présentées des graines pKGW/*abi3-5* trois jours après imbibition dans de l'eau ou de l'ABA 1 µM et 10 µM. Le même résultat est observé pour les graines *abi3-5* et *CaABI3/abi3-5*.

Discussion

A notre connaissance, *C. australe* serait la première espèce appartenant aux plantes supérieures (Angiospermes et Gymnospermes) dans laquelle deux gènes présentant les quatre domaines conservés caractéristiques d'*ABI3* ont pu être identifiés dans l'ADN génomique. Dans le chapitre II, nous avons mis en évidence que le profil d'expression des gènes cibles d'ABI3 suit parfaitement celui de *CaABI3* et *CaABI3-like* au cours du développement des graines de *C. australe*. Il est donc attendu qu'au moins un de ces deux gènes soit fonctionnel comme chez les graines orthodoxes. Deux approches complémentaires ont été développées pour tester la fonctionnalité de *CaABI3* et *CaABI3-like* : (1) l'expression ectopique dans les racines de *M. truncatula* suivie d'une analyse transcriptomique des racines pour détecter les gènes sur-exprimés ; (2) la complémentation du mutant *abi3-5* d'*A. thaliana* en testant la restauration du phénotype sauvage.

L'expression ectopique permet une étude précise des cibles potentielles pour chacun des gènes testés. Dans cette étude, deux génotypes différents de racine ont été utilisés : R108 et *Mtabi3-2*. L'utilisation des racines mutantes permet de s'affranchir de l'expression éventuelle de l'ABI3 endogène, qui est impliqué dans le développement racinaire (Brady *et al.*, 2003). La comparaison des résultats obtenus après sur-expression de $MtABI3_{CDS}$ dans les racines R108 et *Mtabi3-2* montre peu de différences. L'ABI3 endogène jouerait donc un rôle mineur. Notons que 10 % des 109 cibles potentielles d'ABI3 ne sont pas activées par MtABI3 dans les racines. Il a été démontré que l'activation de gènes cibles par ABI3 pouvait nécessiter la présence d'autres facteurs de transcription (Rojas *et al.*, 1999 ; Yotsui *et al.*, 2013), notamment de la famille bZIP (Nakamura *et al.*, 2001 ; Lara *et al.*, 2003). Certains de ces facteurs sont spécifiquement exprimés dans les graines comme ABI5 (Bensmihen *et al.*, 2002) et leur absence dans les racines pourrait expliquer le fait que certaines cibles d'ABI3 ne soient pas sur-exprimées dans cette expérience.

Pour des raisons techniques, nous n'avons pas pu comparer directement les gènes cibles activés par la sur-expression des séquences génomiques *CaABI3* et *CaABI3-like* dans les racines du même fond génétique de *M. truncatula*. Cependant, les comparaisons indépendantes avec *MtABI3* dans les racines R108 et *Mtabi3-2* ont révélé une grande divergence dans la fonctionnalité de *CaABI3* et *CaABI3-like*. Le même profil d'expression est

observé en sur-exprimant *CaABI3-like* et *MtABI3* dans les racines *Mtabi3-2* alors que CaABI3 n'active que 17 % des cibles ABI3 dans les racines R108. Un seul des deux gènes (*CaABI3-like*) serait donc l'homologue *ABI3* fonctionnel de *C. australe*. Chez *Physcomitrella patens*, espèce appartenant aux Bryophytes, trois gènes homologues d'ABI3 ont été clonés mais seul *PpABI3A* est capable d'activer l'expression du gène cible *PpLEA1* en présence d'ABA (Marella *et al.*, 2006). La sur-expression de ce gène est également capable d'activer le promoteur EM du blé dans *P. patens* et dans des cellules de la couche à aleurone de graines d'orge, même si le niveau d'activation est inférieur à ce qui est observé en sur-exprimant *AtABI3* et *ZmVP1* (Marella *et al.*, 2006). Les auteurs ont noté que les deux autres gènes codaient pour une protéine légèrement tronquée à son extrémité N-terminale et que l'absence du domaine acide d'activation devait être préjudiciable à la fonctionnalité de ces protéines. Dans notre cas les deux séquences protéiques CaABI3 et CaABI3-like sont très similaires. Un SNP (Single Nucleotide Polymorphism) serait-il à l'origine de l'incapacité de CaABI3 à activer les cibles d'ABI3 ? Bies-Ethève *et al.* (1999) ont montré que la substitution d'un nucléotide dans le domaine B2 altérait fortement la fonctionnalité d'ABI3 dans les graines d'*A. thaliana*. Cette mutation n'est pas présente dans nos deux séquences homologues chez *C. australe*. Si les domaines A1, B1, B2 et B3 sont très conservés entre ces deux protéines, des divergences plus importantes existent dans le reste de la séquence. Par exemple, deux délétions de huit et sept acides aminés sont visibles chez CaABI3 de part et d'autre du domaine B2 (respectivement en position 490-497 et 553-559 sur la figure 4.1). Ces délétions pourraient affecter la structure de la protéine.

Parmi les 109 cibles MtABI3 identifiées dans le chapitre II, la moitié (54) n'a pas été détectée dans la transcriptome de *C. australe*. Cette proportion importante nous laisse penser que ces gènes ne sont pas exprimés dans les graines de *C. australe*, alors qu'ils jouent un rôle dans la maturation des graines orthodoxes, comme *MtSNF4b* par exemple. Mais il est intéressant de noter qu'en expression ectopique, CaABI3-like est capable d'activer ces gènes au même titre que MtABI3. Si ces gènes ne sont pas exprimés dans les graines de *C. australe*, d'autres régulateurs exprimés chez *M. truncatula* mais absents chez l'espèce récalcitrante sont peut être nécessaires à leur activation. Par exemple, aucun transcrit homologue aux facteurs *bZIP10* et *bZIP25* n'a été identifié dans le transcriptome de *C. australe*. Ces deux facteurs participent à l'activation de gènes cibles d'ABI3 au cours de la maturation des graines d'*A. thaliana* (Lara *et al.*, 2003). D'autres raisons peuvent expliquer l'absence de ces gènes cibles d'ABI3 sur la puce à ADN Ca : (1) couverture insuffisante des séquençages, (2) annotation similaire de plusieurs contigs présentant une grande similarité entre eux alors qu'ils représentent des gènes différents (famille multigénique par exemple).

La complémentation de mutant est une stratégie intéressante pour valider les résultats obtenus par expression ectopique car elle replace la fonctionnalité du gène testé *in vivo*. Malheureusement, le clonage tardif de *CaABI3-like* n'a pas permis de tester sa capacité à complémenter le mutant *abi3-5* à l'heure actuelle. Ces travaux seront lancés rapidement. Pour la complémentation d'*abi3-5* avec CaABI3, trois aspects du phénotype ont été présentés : sensibilité à l'ABA, dégradation de la chlorophylle, et dormance. Les résultats confirment que *CaABI3* n'est pas capable de restaurer le phénotype sauvage sous le contrôle du promoteur d'*A. thaliana* dans les mutants *abi3-5*, contrairement à *AtABI3*. Par ailleurs, les graines du mutant *abi3-5* sont présentées

dans la littérature comme sensibles à la dessiccation, même si une acquisition transitoire a été mise en évidence au cours de la maturation (Ooms *et al.*, 1993). Cependant, dans nos conditions (48h K_2CO_3 44% RH) les graines *abi3-5*, complémentées ou non, sont tolérantes à la dessiccation (100% de germination deux jours après imbibition : données non montrées). Si les graines de ces mutants sont tolérantes à la dessiccation, elles sont peut-être affectées dans leur longévité. Des résultats préliminaires semblent confirmer cette hypothèse. La longévité est fortement réduite chez le mutant *abi3-5* non transformé ainsi que chez les graines complémentées par *CaABI3*. Ce résultat conforte celui de Clerkx *et al.* (2004) qui ont montré que les graines *abi3-5* étaient les plus affectées dans la longévité parmi 20 mutants testés (pas de germination après quatre ans de stockage à température ambiante). Des tests effectués au laboratoire ont montré qu'au-delà de trois mois de stockage à 4°C le pourcentage de germination de graines *abi3-5* est presque nul (données non présentées). Nous ne sommes pas encore en mesure de confirmer que la longévité des graines *abi3-5* complémentées par *AtABI3* retrouve un niveau comparable à ce qui est mesuré chez le sauvage.

Plusieurs études ont testé la fonctionnalité d'homologue d'*ABI3* par complémentation de mutant. Le gène *PpABI3A* (un des trois homologues de *P. patens*) complémente seulement en partie le mutant *abi3-6* sous le contrôle du promoteur d'*A. thaliana*. Les graines mutantes sont vertes, sensibles à la dessiccation et insensibles à 100 µM d'ABA. La complémentation par *PpABI3A* restaure la dégradation de la chlorophylle mais pas la tolérance à la dessiccation et seulement une légère sensibilité à l'ABA (60 % de graines germées à 100 µM d'ABA au lieu de 100 %) (Marella *et al.*, 2006). De plus, l'expression des gènes cible *EM1* et *EM6* est activée par *AtABI3* mais pas par *PpABI3A*. Ces auteurs ont montré par double-hybride qu'une interaction faible entre PpABI3A et AtABI5 pourrait être à l'origine de la mauvaise complémentation du mutant *abi3-6* (Marella *et al.*, 2006). En effet, ces deux facteurs sont nécessaires à l'activation de certains gènes cibles comme *EM1* et *EM6* (Nakamura *et al.*, 2001). Cependant, cette faible interaction s'explique sans doute par la faible conservation du domaine B1 (50,7% de similarité avec AtABI3), due à une distance génétique importante entre les deux espèces. La similarité étant plus importante entre CaABI3 et AtABI3 (91 %), cette explication n'est sans doute pas à privilégier pour expliquer l'absence de complémentation du mutant *abi3-5*. D'ailleurs, les gènes *CnABI3* (homologue du cyprès de Nootka ; Gymnospermes) et *VP1* (homologue du maïs : Monocotylédones), plus éloignés phylogénétiquement d'*AtABI3* que *CaABI3*, complémentent également partiellement le mutant *abi3-6* (Suzuki *et al.*, 2001 ; Zeng *et al.*, 2004). En regardant plus en détails les résultats, la dégradation de la chlorophylle et la TD sont complètement restaurés dans les graines mutantes complémentées par les deux gènes. Cependant, l'expression de gènes cibles comme les protéines de réserves et les oléosines dans les graines complémentées par *CnABI3* retrouve le niveau observé chez le sauvage, ce qui n'est pas le cas des graines complémentées par *VP1*. A l'inverse, la sensibilité à l'ABA n'est pas complètement restaurée dans les graines complémentées par *CnABI3* alors que celles complémentées par VP1 sont aussi sensibles que les graines sauvages (Suzuki *et al.*, 2001 ; Zeng *et al.*, 2004).

L'absence de complémentation du mutant *abi3-5* par *CaABI3* est à relier au résultat d'expression ectopique. Si CaABI3 n'est pas capable d'activer les gènes cibles d'ABI3, il est logique que le mutant ne soit pas complémenté. En ayant placé la séquence codante de *CaABI3* en aval du promoteur d'*AtABI3* pour la complémentation et sous le contrôle du promoteur 35S pour l'expression ectopique, nous pouvons éliminer une influence du promoteur de *CaABI3* pour expliquer l'absence de fonctionnalité de ce gène. Si le test de

complémentation d'*abi3-5* par *CaABI3-like* s'avérait positif, il faudrait rechercher une différence dans la séquence codante entre ces deux gènes ayant un impact majeur sur la structure pour expliquer la perte de fonction de CaABI3. Une étude plus approfondie des gènes sous-exprimés et sur-exprimés par CaABI3 dans les racines de *M. truncatula* pourrait aider à mieux comprendre le rôle de ce gène. Par exemple, les promoteurs des gènes cibles de CaABI3 pourraient être enrichis en éléments cis autres que l'élément RY, motif cible des domaines B3.

Chapitre V

Discussion générale

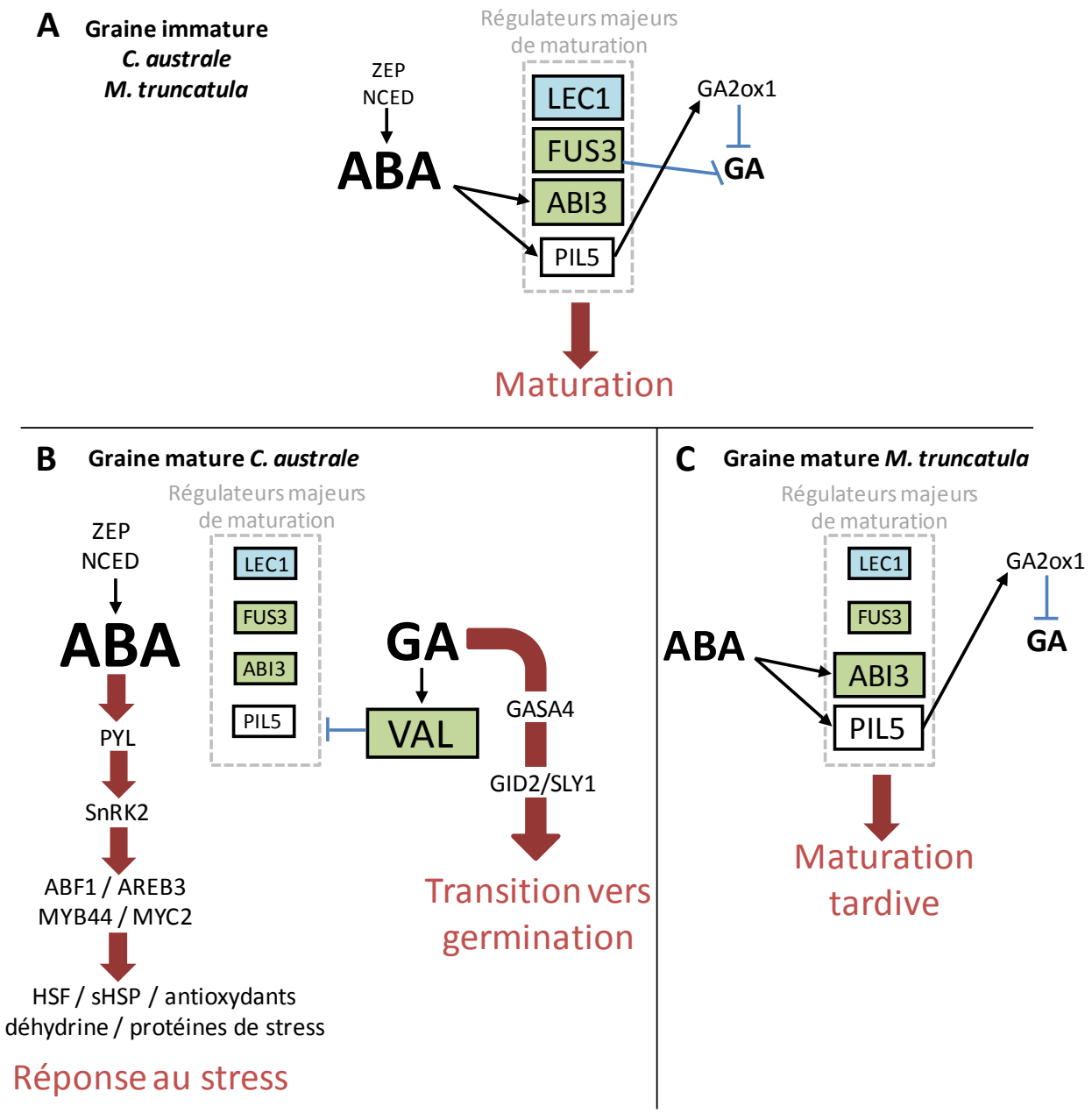


Figure 5.1. Modèle possible du développement des graines de *C. australe* déduit de l'analyse transcriptomique. Voies de régulation activées dans les graines immatures **(A)** et matures **(B)** de *C. australe* ainsi que les graines immatures **(A)** et matures **(C)** de *M. truncatula*. Les régulateurs majeurs de la maturation LEC1 (Leafy Cotyledon-1), ABI3 (ABA-insensitive-3), FUS3 (FUSCA3), VAL (Viviparous1-ABI3-Like) et PIL5 (Phytochrome interacting factor-3 like-5) sont représentés dans les cadres. Les couleurs des facteurs de transcription font référence aux domaines conservés présentés sur la figure 1.5 (vert pour les facteurs de la famille B3). Les phytohormones ABA et GA sont représentées en gras, les grands processus du développement en rouge. Plus la taille des caractères est importante, plus le facteur représenté est exprimé. Les gènes dont le profil d'expression a été utilisé pour établir ce modèle sont ajoutés en noir. Les relations entre différents acteurs de ce modèle sont représentées par une flèche noire pour une activation et un trait bleu pour une répression.

La comparaison moléculaire entre la maturation des graines de *C. australe* et de *M. truncatula* a mis en évidence des divergences majeures entre les graines matures des deux espèces. Le transcriptome de cotylédons de *C. australe* nous permet de proposer un modèle possible de régulation de la maturation qui est présenté sur la figure 5.1. Les graines immatures de *C. australe* sont caractérisées par un développement assez proche de ce qui est observé chez les graines immatures orthodoxes de *M. truncatula*. Les gènes codant pour les régulateurs majeurs de la maturation LEC1, FUS3 ABI3 et PIL5 sont fortement exprimés ainsi que leurs gènes cibles (comme les protéines LEA par exemple). Le ratio ABA/GA est probablement assez élevé car PIL5 et FUS3 contribue à inhiber la biosynthèse de GA (Curaba *et al.*, 2004 ; Oh *et al.*, 2006) (Figure 5.1A). Cependant, le profil d'expression de plusieurs gènes nous indique qu'une transition vers la germination semble enclenchée dans les graines matures de *C. australe*, contrairement à ce qui est observé dans les graines matures orthodoxes. En effet, les gènes codant pour GA2ox1, enzyme impliquée dans l'inactivation des GA, et pour les DELLA, effecteurs négatifs de la signalisation des GA, sont réprimés (Figure 5.1B). Des gènes codant pour des effecteurs positifs sont également activés, comme *GASA4* et *GID2/SLY1*. En parallèle, les gènes codant pour les régulateurs majeurs de la maturation (LEC1, ABI3, FUS3, PIL5) ainsi que l'ensemble des gènes cibles d'ABI3 sont réprimés. Chez les graines matures de *M. truncatula*, LEC1 et FUS3 sont réprimés mais ABI3 et PIL5 restent fortement exprimés (Figure 3.8). La voie des GA reste donc réprimée et le ratio ABA/GA est incompatible avec une transition vers la germination. Les gènes spécifiques de la maturation tardive, dont ceux impliqués dans la TD, restent donc exprimés (Figure 5.1C). Par ailleurs, l'analyse transcriptomique a révélé une deuxième différence majeure entre les graines matures de *M. truncatula* et de *C. australe*. Les gènes impliqués dans la biosynthèse (*ZEP*, *NCED*) et la signalisation de l'ABA (*SnRK2*, *ABF1*, *AREB3*, *MYB44*, *MYC2*) sont plus fortement exprimés chez l'espèce récalcitrante que chez l'espèce orthodoxe. Ils sont à l'origine d'une expression élevée des gènes codant pour des protéines de réponse au stress (HSF, sHSP, déhydrines, antioxydants et autres protéines de stress) (Figure 5.1B). L'activation de ces gènes peut être reliée à une préparation au stress hydrique à venir après l'abscission de ces graines. Ces gènes impliqués dans la réponse aux stress sont différents de ceux impliqués dans la TD chez les graines orthodoxes. Dans le premier cas, la voie de signalisation de l'ABA est activée alors que dans le second cas ABI3 est le régulateur majeur.

Ce point, largement discuté dans le chapitre III, vient répondre à un des deux objectifs de cette étude, à savoir la caractérisation moléculaire du développement d'une graine récalcitrante. Revenons désormais sur le deuxième objectif qui consiste à identifier les mécanismes majeurs dans l'acquisition de la TD des graines par comparaison d'une graine orthodoxe et récalcitrante.

Seize protéines LEA différentes appartenant à cinq familles ont été identifiées dans le protéome stable à la chaleur de graines de *M. truncatula*. En analysant celui de radicules de *M. truncatula* dans lesquelles la TD est ré-induite, Boudet *et al.* (2006) ont déterminé que PM25, EM6, SPB65, PM18 et une isoforme de DHN3 étaient spécifiquement liées à l'acquisition de la TD. Les trois premières sont respectivement 3,9, 4,1 et 161 fois moins abondantes dans les graines matures de *C. australe* et PM18 n'a même pas été détectée dans la graine récalcitrante. D'autres protéines LEA spécifiques des graines sont moins abondantes dans les graines matures de *C. australe* que dans celles de *M. truncatula* (EM1, LEAM) ou même absentes chez l'espèce récalcitrante (PM10, D34-I et D34-II). Nos résultats permettent donc d'élargir la conclusion de Boudet *et al.* (2006) en

Tableau 5.1. Protéines potentiellement impliquées dans la tolérance à la dessiccation ou la réponse au stress. Basé sur les analyses transcriptomique et protéomique, les différents membres des protéines LEA (Late Embryogensis Abundant), des sHSP (small Heat Shock Protein), des sucres non réducteurs et des enzymes impliquées dans la réponse antioxydante ont été classés selon leur fonction potentielle (marquage gris). Les mécanismes impliqués dans la tolérance à la dessiccation sont fortement sous-exprimés dans les graines matures de *C. australe* par rapport à celles de *M. truncatula*. Les mécanismes fortement exprimés dans les graines matures de *C. australe* sont reliés à la réponse au stress abiotique de manière plus générale. Pour les sucres non-réducteurs, le recoupement de nos résultats avec les autres travaux publiés ne permet pas d'aboutir à une conclusion claire, d'où le ' ?'. RFO = oligosaccharides de la famille du raffinose ; APX = ascorbate peroxydase ; MDAR = mono-déhydroascorbate réductase ; GST = glutathion-S transférase.

Famille	Membre	Tolérance à la dessiccation	Réponse au stress
LEA	LEA spécifiques des graines (EM6, EM1, SBP65, LEAM, PM18, D-34, PM25)		
	Déhydrines (DHN3, DHN cognate, BudCar5)		
	LEA non spécifiques des graines (CAPLEA, MP2)		
sHSP	HSP17.4		
	HSP17.6		
	HSP17.9		
	HSP18.2		
	HSP20-like		
	HSP21		
	HSP22		
	HSP22.7		
Sucres non réducteur	RFO	**?**	
	Saccharose	**?**	
Antioxydants	1-cys peroxyredoxine		
	liés au glutation (glutamate cystéine ligase, glutathion synthase)		
	liés à l'ascorbate (APX, MDAR, ascorbate oxydase)		
	Superoxyde dismutase		
	Catalase		
	GST		
	Peroxydases		

corrélant l'accumulation de l'ensemble des protéines LEA spécifiquement exprimées dans les graines à la TD. Des gènes codant pour des protéines présentant des homologies avec les protéines LEA sont également exprimés chez les plantes reviviscentes (Velasco *et al.*, 1998 ; Ditzer *et al.*, 2001 ; Bartels, 2005 ; Suarez-Rodriguez *et al.*, 2010) et les nématodes (Browne *et al.*, 2002) en réponse à la dessiccation. Ces LEA seraient donc impliquées dans l'acquisition de la TD de manière générale, et pas seulement au cours de la maturation des graines.

Cependant, les déhydrines (DHN3, BudCar, DHN cognate) et CAPLEA sont présentes à des niveaux comparables, voire plus élevé, dans les graines matures de *C. australe*. Ces protéines LEA ne sont pas spécifiques des graines et sont plutôt liées à la réponse aux stress abiotiques (Boudet *et al.*, 2006). L'identification de déhydrines dans les graines récalcitrantes de plusieurs espèces d'origine tempérée (Finch-Savage *et al.*, 1994; Farrant *et al.*, 1996) avait conduit à affaiblir l'importance de l'ensemble des protéines LEA dans l'acquisition de la TD. En réalité, nos résultats nous incitent à séparer les protéines LEA en deux catégories : (1) Les déhydrines et autres LEA non spécifiquement exprimées dans les graines (comme CAPLEA). Elles sont accumulées dans les graines matures de *C. australe* et sont donc impliquées dans la réponse aux stress abiotiques plutôt que dans l'acquisition de la TD. (2) Les protéines LEA spécifiquement exprimées au cours de la maturation des graines et absentes dans les tissus végétatifs. Elles sont peu accumulées, voire absentes des graines matures de *C. australe* et sont donc vraisemblablement impliquées dans l'acquisition de la TD (Tableau 5.1). Les résultats transcriptomiques montrent néanmoins dans les graines immatures de *C. australe* une expression importante des gènes codant pour les protéines LEA (y compris celles spécifiques des graines). Le niveau d'expression est comparable à celui observé pour les graines mi-matures de *M. truncatula*. Il serait intéressant de vérifier si ces protéines LEA s'accumulent en conséquence dans le protéome stable à la chaleur de graines mi-matures de *C. australe*. Elles seraient ensuite dégradées car la quantité détectée dans les cotylédons matures de *C. australe* est faible.

L'expression des gènes codant pour les LEA augmente en début de maturation, à partir de 12 DAF, dans les graines de *M. truncatula* et atteint son niveau maximal autour de 32 DAF (Verdier et al., 2013). L'accumulation des protéines correspondantes est pourtant visible seulement lors de la maturation tardive, de 36 DAF à 44 DAF (Chatelain *et al.*, 2012). Ce décalage explique vraisemblablement la faible accumulation de protéines LEA dans le protéome stable à la chaleur du mutant *Mtabi3*. En effet, les gènes codant pour les LEA sont exprimés tardivement chez le mutant par rapport aux graines sauvages (32 DAF à 40 DAF) et les protéines associées n'ont sans doute pas le temps de s'accumuler avant la dessiccation dans les graines matures *Mtabi3*. Par ailleurs, ce décalage de 18 jours entre l'accumulation des transcrits et des polypeptides démontré pour huit protéines LEA par Verdier *et al.* (2013) indique l'existence d'un mécanisme post-transcriptionnel régulant l'accumulation des protéines LEA chez *M. truncatula*. Ce mécanisme n'est peut être pas fonctionnel chez *C. australe*, ce qui pourrait expliquer la faible accumulation de ces protéines dans les graines malgré une expression importante des gènes associés. Dans cette hypothèse, la quantité de protéines LEA détectée dans les cotylédons mi-matures de *C. australe* serait faible, contrairement à l'hypothèse du paragraphe précédent impliquant une dégradation des protéines LEA dans les graines matures. Les RBP (RNA-Binding Protein) participent au contrôle de toutes les étapes de maturation des ARNm nécessaires à leur traduction (revu dans

Glisovic *et al.*, 2008). Ces protéines sont donc des acteurs majeurs de la régulation post-transcriptionnelle qui pourraient être étudiées pour valider cette hypothèse.

Les sHSP sont d'autres protéines stables à la chaleur régulièrement présentées comme impliquées dans l'acquisition de la TD des graines. Cinq spots associés à des sHSP ont été identifiés dans le protéome stable à la chaleur des graines de *C. australe*, ce qui nous indique que ces protéines ne sont pas spécifiques des graines tolérants à la dessiccation. Des sHSP avaient déjà été identifiées dans les graines récalcitrantes de châtaigner (Collada *et al.*, 1997). Cependant, les résultats transcriptomiques ont montré que les gènes codant pour les sHSP sur-exprimés dans les graines matures de *C. australe* n'étaient pas les mêmes que ceux exprimés en fin de développement dans les graines de *M. truncatula* (chapitre III). Ainsi HSP17.5, HSP17.9 et HSP18.2 pourraient être impliquées dans l'acquisition de la TD mais pas HSP15.4, HSP16.6, HSP22 ou HSP22.7 puisque les gènes codants ces dernières protéines sont exprimés dans les graines matures de *C. australe* (Figure 3.15). Cette conclusion vient confirmer les résultats de plusieurs études. En effet, une corrélation entre l'expression des homologues de HSP17.6 et HSP17.9 et la TD a été mise en évidence chez des graines de tournesol (Coca *et al.*, 1994). Les homologues de ces protéines ont été également liés à l'acquisition de la TD chez *C. plantagineum* (Alamillo *et al.*, 1995). L'homologue de *HSP17.4* chez *A. thaliana* est exprimé en lien avec l'acquisition de la TD de la graine et il n'est pas exprimé au cours du développement des graines mutantes *abi3-6* et *lec1-2* sensibles à la dessiccation (Wehweyer & Vierling, 2000). Enfin, l'analyse transcriptomique du développement des graines de *M. truncatula* a permis d'identifier un réseau de 220 gènes corrélés à l'acquisition de la TD (Verdier *et al.*, 2013) parmi lesquels on retrouve *HSP17.5* et *HSP18.2*.

Collada *et al.* (1997) identifient la protéine sHSP détectée dans les graines récalcitrantes de châtaigner comme une sHSP cytosolique de classe I de 20 kD. Parmi les cinq spots identifiés dans le protéome stable à la chaleur de *C. australe*, quatre identifiants UNIPROT différents ont été identifiés : B2D2G4 (spot 2), B4UW51 (spot 56), TC177423 (spots 61 et 62) et B4G0G4 (spot 93). En blastant ces séquences protéiques contre la base de données spécifique d'*A. thaliana* (blastp TAIR10) des annotations plus précises sont disponibles. Ainsi, B2D2G4 et TC177423 sont associées à une HSP21 (AT4G27670 ; e-value 7. 10^{-62} et 4. 10^{-66} respectivement), B4UW51 à une HSP17.6 (AT5G12020 ; e-value 1. 10^{-43}) et B4G0G4 à une HSP20-like BOB1 (Bobber1) (AT5G53400 ; e-value 2. 10^{-73}). Excepté HSP17.6, ces protéines n'ont pas été corrélées à l'acquisition de la TD des graines mais sont plutôt impliquées dans la réponse aux stress abiotiques (Vierling *et al.*, 1988 ; Perez *et al.*, 2009) et le développement des fleurs pour BOB1 (Perez *et al.*, 2009). C'est également le cas de HSP22 et HSP22.7 (Lenne *et al.*, 1995 ; Helm *et al.*, 1995 ; Banzet *et al.*, 1998), sHSP codées par des gènes fortement exprimés dans les graines matures de *C. australe* mais pas dans celles de *M. truncatula*. Comme pour les protéines LEA, nos résultats suggèrent que seuls certaines sHSP seraient impliquées dans l'acquisition de la TD (HSP17.4, HSP17.6, HSP17.9 et HSP18.2) (Tableau 5.1). Les autres sHSP (HSP20-like, HSP21, HSP22, HSP22.7) seraient plutôt des protéines de réponse au stress exprimées dans des tissus sensibles à la dessiccation comme les tissus végétatifs et les graines récalcitrantes (Figure 5.1). Néanmoins, chez le pois, HSP22 est une protéine mitochondriale plus fortement accumulée dans les tissus tolérants à la dessiccation (graines) que dans les tissus sensibles (épicotyles) (Stupnikova *et al.*, 2006).

La majorité des graines orthodoxes de la famille des Fabacées accumulent une quantité importante de RFO (raffinose, stachyose, verbascose) en fin de développement (Corbineau *et al.*, 2000 ; Bailly *et al.*, 2001 ; Rosnoblet *et al.*, 2007 ; revu dans Obendorf & Gorecki, 2012). Ces sucres non réducteurs ont d'ailleurs été corrélés à l'acquisition de la TD par certaines études (Koster & Leopold, 1988 ; Corbineau *et al.*, 2000 ; Bailly *et al.*, 2001). Aucun membre des RFO n'a été détecté dans les graines matures de *C. australe*. Cette absence de RFO a été observée chez une autre Fabacée produisant des graines récalcitrantes : *Inga vera* (Caccere *et al.*, 2013). D'autres études comparatives entre graines récalcitrantes et orthodoxes avaient déjà conclu à une teneur en RFO plus faible chez les graines récalcitrantes par rapport aux graines orthodoxes (Steadman *et al.*, 1996 ; Mello *et al.*, 2010).

La GS et la RFS/STS sont deux enzymes clés nécessaires à la synthèse des RFO. Les deux gènes codant les GS et sept des 11 gènes codant les RFS/STS sont clairement sur-exprimés au cours du développement des graines de *M. truncatula* (Figure 3.13). Liu *et al.* (1998) ont montré chez *A. thaliana* que le gène *wsi76*, précédemment décrit comme induit par la dessiccation indépendamment de l'ABA, code en réalité pour une GS. De plus, l'expression du gène *GolS* codant pour une GS a été corrélée à l'acquisition de la TD dans les graines de tomate (Downie *et al.*, 2003) et de colza (Li *et al.*, 2011). De même, le gène *MT3.5.1s053213*, codant également pour une GS, fait partie du réseau corrélé à l'acquisition de la TD chez *M. truncatula* (Verdier *et al.*, 2013). Tous ces gènes codant pour des GS et des RFS/STS sur-exprimés chez *M. truncatula* sont sous-exprimés dans les graines matures de *C. australe*. Nos résultats renforcent donc l'hypothèse d'une implication de ces enzymes dans l'acquisition de la TD. Au contraire les deux gènes identifiés codant pour une SPP (enzyme finale de biosynthèse du saccharose) sont sur-exprimés chez l'espèce récalcitrante. Les données transcriptomiques sont donc très cohérentes avec les dosages de sucres obtenus pour les graines matures des deux espèces : accumulation de stachyose chez *M. truncatula* et de saccharose chez *C. australe*.

Deux gènes codant pour une RFS/STS sont néanmoins fortement sur-exprimés dans les graines matures de *C. australe*. L'un d'entre eux est annoté 'Seed imbibition protein' (SIP). Ces protéines sont des α-galactosidases alcalines présentant une grande homologie avec les RFS/STS (Carmi *et al.*, 2003) mais sont impliquées dans la dégradation des RFO (Zhao *et al.*, 2006 ; Blöchl *et al.*, 2008). Chez le pois, l'activité enzymatique et l'expression du gène associé sont maximales après la germination (après la percée de la radicule) alors qu'elles sont faibles au cours de la maturation de la graine (Blöchl *et al.*, 2008). Par ailleurs, Zhao *et al.* (2006) ont montré que le saccharose activait l'expression du gène *ZmAGA1* codant pour une SIP chez le maïs. La sur-expression de ces gènes dans les graines matures de *C. australe* peut donc être reliée à la teneur élevée en saccharose. L'activation de gènes *SIP* dans des graines dépourvues de RFO semble paradoxale mais l'activité de ces protéines pourrait masquer un turn-over important, c'est-à-dire que des RFO seraient rapidement dégradés après leur synthèse, empêchant ainsi toute accumulation. Les RFO sont dégradés lors de l'imbibition des graines et fournissent l'énergie nécessaire à la germination et la croissance hétérotrophe de la plantule (Rosnoblet *et al.*, 2007). Ce processus rejoint l'hypothèse détaillée dans le chapitre III présentant la graine mature de *C. australe* comme une graine se préparant pour la germination.

Bien que l'absence de RFO dans la graine de *C. australe* et la sous-expression des gènes impliqués dans la biosynthèse de ces sucres non-réducteurs suggèrent une implication possible dans la TD, certaines études remettent en cause la corrélation entre accumulation de RFO et acquisition de la TD (Black *et*

al., 1999 ; Bentsink *et al.*, 2000). De plus, le stachyose est accumulé dans les axes de l'espèce hautement récalcitrante *A. marina* (Farrant *et al.*, 1993b) tout comme le raffinose dans ceux de *Quercus robur* (Finch-Savage & Blake, 1994). Les fonctions majeures proposées pour les RFO en lien avec la TD sont la participation à l'état vitreux du cytoplasme (Buitink & Leprince, 2004) et la stabilisation des membranes et protéines par remplacement de l'eau (Hoekstra *et al.*, 2001) Ces fonctions sont importantes à des teneurs en eau assez basses, bien plus faibles que la plupart des teneurs en eau critique de graines très récalcitrantes comme *A. marina*. Si ces sucres jouent un rôle dans la TD, ils ne sont vraisemblablement pas les acteurs majeurs expliquant la sensibilité à la dessiccation des graines récalcitrantes (Tableau 5.1). Chez les Fabacées, l'accumulation de RFO en fin de maturation est courante (Obendorf & Gorecki, 2012) et d'autres fonctions non liées à la TD sont également à considérer. Par exemple les RFO constituent une source de carbone rapidement mobilisable lors de la germination (Vandecasteele *et al.*, 2011). L'absence de ces sucres chez les Fabacées récalcitrantes *C. australe* et *Inga vera* (Caccere *et al.*, 2013) signifie que certains régulateurs importants du métabolisme de ces sucres ne sont pas exprimés. Par exemple, ABI3 et MtSNF4b (une sous-unité activatrice du complexe SnRK1) sont connus pour réguler la biosynthèse de RFO (Ooms *et al.*, 1993 ; Rosnoblet *et al.*, 2007). *ABI3* est sous-exprimé dans les graines matures de *C. australe* et aucun homologue de *SNF4b* n'a été identifié dans le transcriptome, suggérant que ce gène n'est pas exprimé chez *C. australe*. Le phénotype du mutant *Mtsnf4b* peut être rapproché de ce qui est observé dans les graines matures de *C. australe* : absence de RFO et teneur élevée en saccharose. En effet, la teneur en stachyose est réduite de 50 à 80 % et la teneur en saccharose augmentée de 60 % dans les graines matures mutantes par rapport aux graines sauvages (Rosnoblet *et al.*, 2007). Ce résultat appuie l'hypothèse d'une absence d'expression de *SNF4b* dans les graines de *C. australe*.

La sensibilité à la dessiccation de grains de maïs en germination (Leprince *et al.*, 1990) ou de graines récalcitrantes de *Q. robur* (Hendry *et al.*, 1992) s'explique, au moins en partie, par une protection inefficace face aux ROS. L'analyse transcriptomique des gènes impliqués dans la réponse antioxydante a révélé des différences intéressantes entre *M. truncatula* et *C. australe*, avec des groupes de gènes sur-exprimés chez l'espèce récalcitrante et d'autres sous-exprimés. Les gènes codant pour des enzymes présentant une activité antioxydante liée à l'ascorbate (APX, MDAR, ascorbate oxydase) ne sont pas différentes dans les graines matures de *C. australe* par rapport à celles de *M. truncatula*. Ce résultat n'est pas étonnant car plusieurs études ont suggéré que l'activité de ces enzymes était plus importante au cours de l'embryogenèse et du remplissage que pendant la dessiccation (Arrigoni *et al.*, 1992 ; Bailly *et al.*, 2001 ; De Gara *et al.*, 2003). De plus, la teneur en ascorbate et l'activité APX est plus élevée dans les graines récalcitrantes de *Ginkgo biloba*, *Quercus cerris*, *Aesculus hippocastanum* et *Cycas revoluta* à l'abscission que dans les graines sèches orthodoxes de *Vicia faba*, *Avena sativa* et *Pinus pinea* (Tommasi *et al.*, 1999). Nos données confortent l'idée que l'ascorbate ne serait pas impliqué dans la TD (Tableau 5.1), comme suggéré par De Tullio & Arrigoni (2003). Ces auteurs proposent des fonctions dans le contrôle du cycle cellulaire et la germination.

De même les gènes liés aux activités SOD et CAT sont fortement exprimés dans les graines matures de *C. australe*. Ces mécanismes sont vraisemblablement plutôt liés à la réponse au stress (Tableau 5.1). Des études précédentes avaient déjà suggéré que la SOD était impliquée dans la réponse aux stress abiotiques

plutôt que dans la TD (Bailly *et al.*, 2001 ; revu par Amudha & Balasubramani, 2010). En revanche l'activité de la CAT a été corrélée à la dessiccation au cours de la maturation de la graine de haricot et de tournesol (Bailly *et al.*, 2001 ; 2004). Chez le tournesol, l'augmentation d'activité CAT est liée à l'activation de la transcription du gène *CATA1* codant pour une sous-unité de cette enzyme (Bailly *et al.*, 2004). Guan *et al.* (2000) ont montré que l'expression du gène *CAT1* du maïs était activée par l'ABA via un élément ABRE présent dans le promoteur de ce gène. Etant donnée l'activation de la voie de signalisation de l'ABA dans les graines matures de *C. australe*, il n'est pas étonnant de voir une sur-expression du gène codant pour la CAT au cours du développement de ces graines. Nos résultats viennent donc nuancer la spécificité de l'activité CAT dans la TD puisque l'activation transcriptionnelle observée au cours du séchage des graines orthodoxes de tournesol est également visible dans des graines récalcitrantes. Il faut cependant rester prudent dans l'analyse de nos résultats puisqu'un seul gène codant pour la CAT a été identifié dans le transcriptome de *C. australe* alors que plusieurs gènes sont exprimés chez le tournesol, le maïs ou *A. thaliana* (Guan & Scandalios, 1998 ; Bailly *et al.*, 2004 ; Mhamdi *et al.*, 2012).

Au contraire, les gènes codant pour la biosynthèse du glutathion (glutamate cystéine ligase / glutathion synthétase) et ceux codant pour les 1-cys peroxyredoxines sont clairement sous-exprimés dans les graines matures de *C. australe* par rapport à celles de *M. truncatula*. L'expression du gène *PER1*, codant pour une 1-cys peroxyredoxine, est spécifique des graines chez l'orge et *A. thaliana*. Un rôle antioxydant au cours de la dessiccation et de la dormance des graines a été proposé pour ces protéines (Stacy *et al.*, 1996 ; 1999 ; Haslekas *et al.*, 1998 ; Aalen *et al.*, 1999). L'expression de *PER1* n'est pas induite dans les tissus végétatifs en réponse à l'ABA ou à un stress hydrique (Haslekas *et al.*, 1998) et est réprimée par les GA (Stacy *et al.*, 1996), trois facteurs supposés être importants dans la graine mature de *C. australe*. La sous-expression de ce gène observée chez *C. australe* est donc tout à fait cohérente et renforce le rôle potentiel des 1-cys peroxyredoxines dans l'acquisition de la TD (Tableau 5.1).

Bien que la teneur en glutathion n'ait pas été dosée dans les graines, les résultats transcriptomiques indiquent une probable forte diminution de sa biosynthèse dans les graines matures de *C. australe*, à l'inverse de ce qui est observé chez *M. truncatula*. Contrairement à l'ascorbate, le glutathion pourrait être un antioxydant majeur dans l'acquisition de la TD (Tableau 5.1). Cette hypothèse est appuyée par la présence de deux gènes codant une glutamate cystéine ligase (biosynthèse du glutathion), trois gènes codant une GST et un gène codant une GPX parmi les 220 gènes corrélés à l'acquisition de la TD chez *M. truncatula* (Verdier *et al.*, 2013). Cependant, nos résultats montrent que la teneur en glutathion serait plus importante que les activités GST, GPX et autres peroxydases pour l'acquisition de la TD (Tableau 5.1). En effet, les gènes codant pour ces enzymes restent fortement exprimés dans les graines matures de *C. australe* (à l'exception du gène *Medtr7g065590.1* codant pour une GST et corrélé à l'acquisition de la TD chez *M. truncatula*). En comparant les réponses antioxydantes des graines d'*Acer platanoides* (orthodoxes) et d'*Acer pseudoplatanus* (récalcitrantes) au cours de la dessiccation, Pukacka & Ratajczak (2007) ont montré un système de défense plus actif chez l'espèce orthodoxe, se traduisant par une teneur en glutathion plus élevé. En travaillant sur le même couple d'espèces, Greggains *et al.* (2000) avait montré précédemment au cours du développement de ces graines que la teneur en glutathion était deux fois plus importante dans la graine orthodoxe au début du séchage sur l'arbre (12,8 µmol/g MS) que dans la graine récalcitrante à l'abscission (6,3 µmol/g MS). Le même type de résultat a

été obtenu en comparant la teneur en glutathion dans les feuilles de *Sporobolus stapfianus* (plante tolérante à la desiccation) et *Sporobolus pyramidalis* (plante sensible à la desiccation) (Oliver *et al.*, 2011). La teneur en glutathion joue un rôle majeur dans la maturation de la graine d'*A. thaliana*, comme le montre le phénotype létal du mutant *gsh1* déficient en synthèse de glutathion (Cairns *et al.*, 2006).

D'autres gènes présentent des profils d'expression différents entre les graines de *M. truncatula* et de *C. australe* et peuvent être reliés à l'acquisition de la TD, comme SNF4b et les facteurs DREB. L'expression de *MtSNF4b* est activée lors de la réinduction de la TD dans des radicules de *M. truncatula* (Buitink *et al.*, 2004) alors que ce gène ne semble pas exprimé chez *C. australe*. Par ailleurs, les quatre gènes codant pour les facteurs de transcription DREB2C et DREB1D sont sur-exprimés en fin de maturation chez les graines orthodoxes mais restent faiblement exprimés chez les graines récalcitrantes. L'un d'entre eux (Mtr.31778.1.S1_at) a été corrélé à l'acquisition de la TD chez *M. truncatula* (Verdier *et al.*, 2013). Une analyse transcriptomique du développement des graines d'orge avait déjà souligné l'implication possible des facteurs DREB2A dans l'acquisition de la TD (Sreenivasulu *et al.*, 2006).

Tous ces mécanismes que nos analyses transcriptomique et protéomique ont permis de relier à l'acquisition de la TD partagent un point commun, ils sont régulés par ABI3. Parmi les cibles directes d'ABI3 chez *A. thaliana* et *M. truncatula*, on retrouve la plupart des gènes codant des LEA, le gène *PER1* (1-cys peroxyredoxine), le gène *SNF4b* et un gène codant pour une galactinol synthase (Mönke *et al.*, 2012 ; Verdier *et al.*, 2013). L'expression des gènes codant pour les LEA, pour PER1 et pour HSP17.4 est également réduite dans les graines mutantes *abi3* (Haslekas *et al.*, 1998 ; Wehmeyer & Vierling, 2000 ; Bies-Ethève *et al.*, 2008, Verdier *et al.*, 2013). Enfin la teneur en RFO des graines mutantes *abi3-5* et *abi3-4* est presque nulle alors que ces oligosaccharides représentent 11 % des sucres solubles dans les graines sauvages (Ooms *et al.*, 1993). Tous ces résultats ont permis progressivement de positionner ABI3 comme un régulateur majeur de l'acquisition de la TD chez les graines orthodoxes. La caractérisation de ce gène chez une graine récalcitrante est donc particulièrement intéressante.

Deux gènes *ABI3* ont été identifiés chez *C. australe* mais seul *CaABI3-like* semble être homologue à *MtABI3* d'un point de vue fonctionnel. Ces gènes sont tous les deux exprimés jusqu'au stade mi-mature avant d'être réprimés en fin de maturation. Cette répression conduit à la sous-expression de toutes les cibles d'ABI3 et est vraisemblablement à l'origine de la sensibilité à la dessiccation des graines de *C. australe*.

Deux nouvelles questions découlent de cette conclusion : quel est le rôle de *CaABI3* s'il n'active pas les cibles ABI3 ? Quel est le mécanisme à l'origine de la répression de *CaABI3* et *CaABI3-like* ?

Des études plus approfondies sont nécessaires pour mettre en évidence le rôle de *CaABI3* et nous ne sommes pas en mesure d'apporter des éléments de réponse à l'heure actuelle. Cependant, d'un point de vue évolutif, la non fonctionnalité de *CaABI3* est une observation très intéressante qui pourrait expliquer l'apparition d'un deuxième gène fonctionnel *CaABI3-like*. Cependant, quelle est l'utilité d'un homologue *ABI3* fonctionnel puisque ce gène est réprimé avant l'acquisition de la TD.

Concernant les mécanismes de répression de *CaABI3* et *CaABI3-like*, nous avons observé les profils d'expression des gènes codant pour les répresseurs connus d'*ABI3* chez les graines orthodoxes. Pour trois

d'entre eux, aucun homologue n'a pu être clairement identifié dans le transcriptome de graines de *C. australe* (*ASIL1*, *HDA6*, *HDA19*). De plus, le gène *PKL* ne semble pas exprimé au cours du développement de la graine de *C. australe* et le gène *SUA* est faiblement exprimé (Figure 3.8). En revanche, l'homologue d'*AIP2* est fortement exprimé dans les cotylédons de *C. australe*. Cependant, deux arguments nous laissent penser que ce gène n'est pas à l'origine de la répression des homologues *ABI3* de *C. australe*. Tout d'abord, le profil d'expression est constant tout au long du développement et n'est donc pas corrélé à la répression de *CaABI3* et *CaABI3-like*. Chez les deux espèces où l'expression d'*AIP2* a été étudiée dans la graine un pic d'expression a été mis en évidence. En effet, en couplant le promoteur de ce gène à un rapporteur GUS, Zhang *et al.* (2005) ont montré que l'expression d'*AIP2* est maximale dans des graines d'*A. thaliana* fraichement stratifiées, même si elle n'est pas nulle dans les autres tissus de la plante et les autres stades de développement de la graine. Chez le cyprès de Nootka (Gymnospermes), un homologue d'*AIP2* a été identifié et son expression est également maximale au cours de la levée de dormance (Zeng *et al.*, 2013). Le deuxième argument en défaveur d'une répression de *CaABI3* et *CaABI3-like* par AIP2 est la nature même de cette protéine. Il s'agit d'une E3 ubiquitine ligase qui réprime l'activité d'ABI3 au niveau protéique. Si ces auteurs ont montré une diminution de la quantité de protéine ABI3 en parallèle de l'activation d'AIP2, aucun effet de cette ligase sur l'expression du gène *ABI3* n'a été mis en évidence (Zhang *et al.*, 2005 ; Zeng *et al.*, 2013).

Les derniers répresseurs connus d'*ABI3* sont les facteurs de transcription de la famille VAL. Les homologues de *VAL3* et *VAL2-like* sont les seuls gènes codant pour des répresseurs de maturation dont l'expression est inversement corrélée à celle de *CaABI3* et *CaABI3-like*, profil cohérent avec un rôle répresseur (Figure 3.8). De plus, contrairement à AIP2, plusieurs travaux ont montré que les facteurs VAL modulent l'expression du gène *ABI3* et non l'abondance de la protéine. Tout d'abord, l'expression d'*ABI3* est huit fois plus importante dans des plantules du double mutant *val1/val2* (appelé aussi *hsl1/hsi2*) que chez le type sauvage Col-0 d'*A. thaliana* quatre jours après imbibition (Tsukagoshi *et al.*, 2007). Dans des plantules *val1/val2* de 7,5 jours, *ABI3* est même 50 fois plus exprimé que dans les plantules de type sauvage au même stade de développement (Suzuki *et al.*, 2007). En plus du domaine B3, ces facteurs possèdent des domaines spécifiques de type 'EAR', 'plant homeodomain' et 'CW' (Tsukagoshi *et al.*, 2005 ; Suzuki *et al.*, 2007). Le domaine EAR est associé à une activité répressive (Ohta *et al.*, 2001) et les deux derniers domaines sont connus pour interagir avec l'ADN (Berger, 2007 ; Hoppmann *et al.*, 2011). Ils confèrent aux facteurs VAL la capacité de réprimer l'expression des gènes par remodelage de la chromatine, notamment via la triméthylation de la lysine 27 de l'histone 3 (H3K27me3) (Veerappan *et al.*, 2012). Plus de la moitié des 655 gènes sur-exprimés plus de quatre fois dans le double mutant *val1/val2* par rapport au type sauvage possèdent au moins un élément cis RY/Sph dans leur promoteur ou le premier intron (dont *ABI3*). Cet enrichissement pour le motif reconnu par le domaine B3 laisse penser que ce domaine, présent chez les facteurs VAL, pourrait également être impliqué dans l'activité répressive de ces facteurs. Guerriero *et al.* (2009) ont montré *in vitro* par un système simple-hybride que VAL2 (mais pas VAL1) pouvait fixer l'élément RY/Sph. Cette interaction n'avait pas été mise en évidence auparavant (Suzuki & McCarthy, 2008) à cause d'un mécanisme d'auto-inhibition vraisemblablement du à la présence du domaine répressif. Pour contourner cette difficulté, Guerriero *et al.* (2009) ont utilisé un système chimérique avec FUS3 pour lequel il n'existe pas de problème d'auto-inhibition.

Il ressort de ces études et de nos résultats que l'activation des gènes *VAL2-like* et peut-être *VAL3* est la cause probable de la répression de *CaABI3* et *CaABI3-like* en fin de développement des graines de *C. australe* (Figure 5.1). Chose intéressante, le profil inverse est observé dans les graines matures de *M. truncatula* où *VAL2-like* et *VAL3* sont réprimés en parallèle d'une expression importante de *MtABI3* (Figure 3.8). La participation de VAL3 est à considérer avec précaution car le rôle de ce facteur semble mineur dans la répression des gènes de maturation (Suzuki *et al.*, 2007). Ces auteurs ont émis l'hypothèse que l'expression de *VAL2* pourrait être induite par les GA. Leur conclusion se base sur la présence du phénotype embryonnaire des plantules du mutant *val1* uniquement si ces dernières sont cultivées en présence d'inhibiteur de GA. Ils en déduisent que la synthèse de GA pourrait activer l'expression de *VAL2* pour compenser la déficience de *VAL1* dans le mutant et permettre un développement normal des plantules (Suzuki *et al.*, 2007). Cette hypothèse recoupe tout à fait nos observations transcriptomiques montrant une activation de la voie de biosynthèse et de signalisation des GA dans les graines matures de *C. australe*.

Conclusion et perspectives

Ces travaux de thèse ont porté sur l'analyse comparative des graines orthodoxes de *M. truncatula* et récalcitrantes de *C. australe* par des approches protéomique et transcriptomique. Ils ont été menés dans l'objectif de répondre à deux questions : mettre en évidence des mécanismes impliqués dans l'acquisition de la TD des graines mais également comprendre l'origine de la sensibilité à la dessiccation des graines récalcitrantes en caractérisant pour la première fois à l'échelle moléculaire le développement de l'une d'entre elles.

La caractérisation du protéome stable à la chaleur et du transcriptome de graines de *C. australe* par des techniques haut-débit a permis d'avoir une vue plus globale et plus précise sur le développement de ces graines par rapport aux données publiées jusqu'ici. Les conclusions de nos résultats viennent donc appuyer ou nuancer des résultats précédemment obtenus par des approches plus ciblées. Par exemple, des déhydrines ont été détectées par western-blot chez bon nombre d'espèces récalcitrantes, suggérant un rôle mineur de ces protéines dans la TD, mais aucune donnée n'était disponible concernant les autres familles de protéines LEA. Nos résultats confirment que les déhydrines sont fortement accumulées dans les cotylédons matures de *C. australe* mais ils montrent également que les autres protéines LEA spécifiques des graines sont nettement moins abondantes que chez *M. truncatula*, voire absentes.

Il est sans doute trop simple de considérer que la sensibilité à la dessiccation des graines récalcitrantes puisse être expliquée par l'absence complète d'un mécanisme majeur exprimé au cours de la maturation des graines orthodoxes : par exemple absence de toutes les protéines LEA, des sHSP ou de réponse antioxydante. La recherche de membres spécifiques au sein de ces grandes familles de molécules impliquées dans la TD est une voie plus réaliste. Effectivement, comme pour les protéines LEA, nos données transcriptomiques suggèrent que seules certaines sHSP et certains mécanismes antioxydants sont peu exprimés dans les cotylédons de *C. australe*. En recoupant ces données avec celles publiées sur d'autres modèles d'étude, nous avons pu ressortir des acteurs jouant probablement un rôle important dans l'absence de la TD de ces graines : HSP17.4, HSP17.9, HSP18.2, glutathion, 1-cys peroxyredoxine, LEA spécifique des graines.

Ces travaux apportent également des résultats tout à fait nouveaux concernant le développement des graines récalcitrantes. Pour la première fois, l'expression des gènes codant pour les régulateurs majeurs de la maturation des graines orthodoxes a été étudiée chez une graine récalcitrante. La forte répression des homologues d'*ABI3* et des gènes cibles de ce facteur de transcription dans les cotylédons matures de *C. australe* constitue la différence la plus remarquable avec le développement des graines de *M. truncatula*. Ce régulon ABI3 est pourtant fortement exprimé aux stades immatures des graines de *C. australe*. Etant donné le rôle majeur d'ABI3 dans la TD et dans la régulation des protéines LEA chez les graines orthodoxes, il serait intéressant d'étendre l'analyse du protéome stable à la chaleur à des graines immatures de *C. australe*. En l'absence de régulation post-transcriptionnelle, l'accumulation de protéines LEA et d'autres protéines cibles d'ABI3 pourrait être plus importante par rapport à ce que nous avons observé dans les graines matures. De même, une acquisition transitoire de la TD (ou tout du moins une moindre sensibilité à la dessiccation) des cotylédons immatures par rapport au stade mature constitue une hypothèse intéressante qui mériterait d'être testée. Toutes ces conclusions sont valables pour le cotylédon mais une étude de l'expression du régulon ABI3 et du profil d'accumulation des protéines LEA dans l'axe embryonnaire pourrait donner des résultats différents.

En effet, l'axe est plus tolérant à la dessiccation chez certaines espèces récalcitrantes comme *Castanea sativa* (Leprince *et al.*, 1999) et certaines protéines LEA (EM6, D113, D34-II) sont beaucoup plus abondantes dans l'axe que dans les cotylédons de graines matures de *M. truncatula* (Chatelain *et al.*, 2012).

La répression du régulon ABI3 couplée à une activation des gènes impliqués dans la biosynthèse et la signalisation des GA et des gènes codant pour les répresseurs de maturation VAL sont autant d'indices suggérant une transition maturation/germination dans les graines matures de *C. australe*. En se basant sur ce qui a été observé chez les graines orthodoxes, cette phase de transition devrait s'accompagner d'une augmentation du ratio GA/ABA dans les graines de *C. australe*. Cependant, le modèle est peut être plus complexe car les graines matures de *C. australe* sont également caractérisées par une activation de l'expression des gènes impliqués dans la biosynthèse et la signalisation de l'ABA ainsi que dans la réponse au stress. Les voies de l'ABA et des GA, connues pour leur antagonisme chez les graines orthodoxes, pourraient donc être activées toutes les deux en fin de développement chez les graines récalcitrantes de *C. australe*. Mesurer l'évolution de la teneur en GA et ABA au cours du développement des graines nous permettrait de confirmer cette hypothèse. A la suite de cette thèse, l'étude du développement des graines de *C. australe* sera poursuivie en collaboration avec l'équipe sud-africaine de Dr. Jill Farrant dans le cadre d'une thèse. L'analyse transcriptomique des cotylédons et des axes sera étendue à sept stades de développement et sera complétée par le dosage des phytohormones, une analyse métabolomique et une analyse cellulaire (microscopie électronique).

Le dernier chapitre de cette thèse s'est focalisé sur l'analyse fonctionnelle des deux gènes homologues *ABI3* clonés chez *C. australe*. Malgré une très forte similarité entre eux, seul CaABI3-like est capable d'activer les cibles ABI3 par sur-expression ectopique dans les racines. La fonction de *CaABI3*, qui n'active que très peu de cibles ABI3 et ne complémente pas le mutant *abi3-5* d'*A. thaliana*, reste à élucider. L'utilisation de prédicteurs de structures pourrait être intéressante pour essayer de comprendre pourquoi CaABI3 n'est pas fonctionnel alors que CaABI3-like, pourtant très proche au niveau similarité de séquences, l'est. La caractérisation fonctionnelle de ces deux gènes va être rapidement complétée par de nouvelles répétitions de sur-expression ectopique et de nouvelles complémentations de mutant par les gènes *MtABI3* et *CaABI3-like*, non encore testés à ce jour. La complémentation par *MtABI3* est une validation qui nous permettra d'exclure l'argument de la distance génétique entre espèce pour expliquer l'absence de complémentation du mutant d'*A. thaliana* par *CaABI3*. De plus, de nouvelles lignées de mutants complémentés par *AtABI3* et *CaABI3* doivent également être caractérisées pour confirmer les premiers résultats mis en évidence sur une seule lignée à l'heure actuelle.

D'autres perspectives sont envisageables concernant les mécanismes impliqués dans la TD, nos données transcriptomiques et protéomique ont été utilisées avec *a priori*. Nous avons spécifiquement recherché l'expression de mécanismes déjà proposés dans la littérature comme potentiellement impliqués dans la TD (LEA, sHSP, antioxydants, sucres non réducteurs). Cependant, nos données peuvent également être utilisées sans *a priori* en élargissant la comparaison entre *C. australe* et *M. truncatula* à l'ensemble des gènes sur-exprimés et sous-exprimés au cours du développement de la graine récalcitrante. Cette approche permettrait de proposer de nouveaux mécanismes potentiellement impliqués dans la TD des graines.

Chapitre VI

Matériel et méthodes

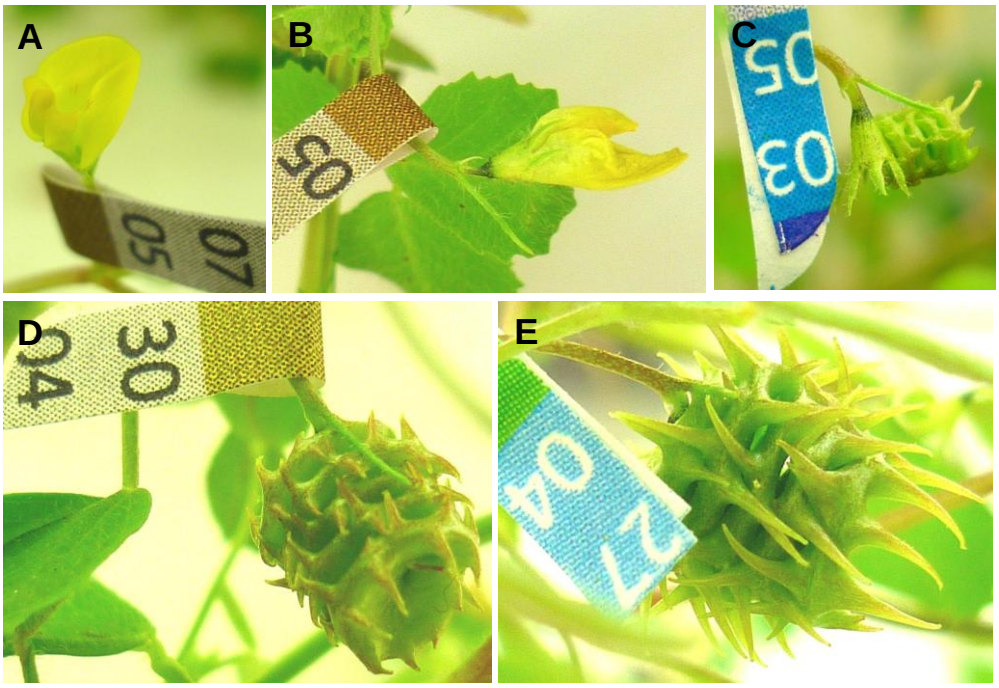


Figure 6.1. Marquage des fleurs pour la récolte des graines au cours de la maturation.
Les fleurs sont étiquetées lors de leur complète ouverture (jour 0) **(A)** et se referment rapidement après la fécondation, ici 1 jour après floraison (DAF) **(B)**. A partir de 4 DAF, la gousse est bien visible **(C)** et grossit rapidement au cours de l'embryogenèse 7 DAF **(D)** et 10 DAF **(E)**.

Figure 6.2. Différents stades de développement des gousses de *C. australe* utilisés dans cette étude.
Stade immature gousse verte-jaune (CaG) **(A),** mi-mature gousse jaune (CaY) **(B)** et mature gousse marron (CaB) **(C)** (Photos : Udisha Chathuri)**.**

1-Matériel végétal

1.1-*Medicago truncatula*

Deux génotypes de *M. truncatula* ont été utilisés au cours de cette thèse : le génotype A17 pour les analyses protéomiques et transcriptomiques des chapitres II et III et le génotype R108 pour l'analyse fonctionnelle des gènes *CaABI3* et *CaABI3-like* du chapitre IV. En effet, les deux lignées de mutants *M. truncatula* comportant une insertion *Tnt1* dans le gène *ABI3* (NF3185 et NF6003 respectivement appelées *Mtabi3-1* et *Mtabi3-2* dans le reste de l'étude) ont été obtenues par la Samuel Noble Foundation (Oklahoma, USA) dans le génotype R108. Les graines homozygotes mutantes ont été sélectionnées en se basant sur la couleur des graines qui restent vertes à maturité, phénotype récessif non visible chez les plantes homozygotes. Les insertions *Tnt1* dans les deux mutants ont été vérifiées par PCR (Annexe 1 pour les amorces) et les lignées ont été rétro-croisées avec le parent sauvage une fois pour *Mtabi3-1* et deux fois pour *Mtabi3-2*.

Les plantes de *M. truncatula* de tout génotype ont été cultivées en chambre climatique sous une photopériode de 16h à 24°C avec un éclairement de 200 µmol de photons m^{-2} s^{-2} et 8h à 21°C sous obscurité. Les fleurs ont été étiquetées au jour de floraison correspondant à leur ouverture totale (Figure 6.1) puis les graines ont été dégoussées à différents intervalles au cours de la maturation : de la fin de l'embryogenèse (11 jours après floraison DAF) à l'abscission. Les graines fraîches ont été immédiatement utilisées pour les manipulations physiologiques ou congelées dans l'azote liquide et stockées à -80°C en vue des analyses transcriptomiques, protéomiques et biochimiques. Etant donné l'impossibilité de stocker des graines homozygotes mutantes, des cultures de *Mtabi3-1* et *Mtabi3-2* ont été maintenues en permanence à partir de graines fraîchement récoltées.

1.2-*Castanospermum australe*

Les graines de *C. australe* ont été récoltées au cours de la maturation et à l'abscission sur des arbres poussant à Pietermaritzburg (Kwazulu-Natal, Afrique du Sud) en 2009 et 2011. Celles de la première récolte ont été ramenées à Angers par avion dans les 48h suivant la récolte où elles ont été immédiatement traitées. Les axes embryonnaires de graines immatures (récoltées en mai) et matures (récoltées en juillet) ont été séparés des cotylédons et congelés dans l'azote liquide puis stockés à -80°C en vue du séquençage Illumina. Les cotylédons frais ont été utilisés pour la détermination de la sensibilité à la dessiccation ou séchés un ou trois jours dans un dessiccateur NaCl à 75 % d'humidité relative (RH) avant d'être congelés dans l'azote liquide et stockés à -80°C pour le séquençage Illumina, les analyses protéomiques et biochimiques. Les ARN de cotylédons et d'axes embryonnaires de graines immatures (gousse verte-jaune ; CaG), mi-mature (gousse jaune ; CaY) et matures (gousse marron ; CaB) de la récolte de 2011 (Figure 6.2) ont été extraits en Afrique du Sud puis transportés par avion dans la carboglace jusqu'à Angers pour le séquençage 454.

1.3-*Arabidopsis thaliana*

Le mutant *abi3-5* d'*A. thaliana* a également été utilisé pour l'analyse fonctionnelle de *CaABI3*. Ce mutant a été obtenu dans l'écotype Landsberg par traitement EMS (méthanesulfonate d'éthyle) aboutissant à la délétion d'un nucléotide et la production d'une protéine tronquée non fonctionnelle (Bies-Ethève *et al.*, 1999). Des graines issues d'une plante hétérozygote nous ont été données par Thierry Joët de l'IRD Montpellier

(génération 0). Les graines homozygotes mutantes (vertes, insensibles à l'ABA) et sauvages (beiges, sensibles à l'ABA) ont été séparées dans la descendance de ces plantes en se basant sur la couleur des graines dans les siliques déhiscentes (génération 1) et en confirmant le phénotype vis-à-vis de la sensibilité à l'ABA dans la génération suivante (génération 2) (cf 2.3 pour le protocole de sensibilité à l'ABA). Comme pour *Mtabi3*, des cultures homozygotes d'*abi3-5* ont été maintenues en permanence à partir de graines fraîchement récoltées. Les cultures se sont faites en chambre climatique sous une photopériode de 16h à 24°C avec un éclairement de 200 µmol de photons m^{-2} s^{-2} et 8h à 21°C sous obscurité.

2-Traitements

2.1-Test de germination

Des graines de *C. australe* issues de la récolte 2009 ont été conservées un an dans une étuve à 15°C 60 % RH au terme duquel leur capacité germinative a été testée. Afin d'éliminer les contaminations fongiques, les téguments ont été régulièrement rincés à l'éthanol 70 %.

La capacité germinative de graines de *M. truncatula* fraîchement récoltées à différents stades de développement a été mesurée par percée de la radicule dans les 24 jours après imbibition dans l'eau sur papier filtre à 20°C et à l'obscurité.

2.2-Mesure de tolérance à la dessiccation

2.2.1-*M. truncatula*

La tolérance à la dessiccation des graines sauvages (R108) et mutantes (*Mtabi3-1* et *Mtabi3-2*) a été mesurée sur des graines récoltées à différents points au cours de la maturation. La teneur en eau (TE) associée à ces stades de développement a été mesurée sur trois à quatre répétitions de 5x3 graines par pesée des graines avant (MF) et après (MS) séchage dans un four à 96°C pendant 48h : TE (en g H_2O g MS^{-1}) = (MF-MS)/MS. Deux à cinq répétitions de 30 à 50 graines récoltées à chaque stade de développement ont été rapidement séchées dans un dessiccateur K_2CO_3 43 % RH (teneur en eau équilibrée à 0,09 g H20 g MS^{-1}) puis imbibées sur papier filtre et eau à 20°C à l'obscurité après 48h dans le dessiccateur. Les graines sont considérées tolérantes à la dessiccation lorsqu'elles germent dans les 24 jours suivant l'imbibition dans les conditions décrites au point 2.1.

2.2.2-*C. australe*

La sensibilité à la dessiccation des cotylédons de *C. australe* a été déterminée sur des cubes d'environ 3x5x3 mm isolés du cœur des cotylédons de graines récoltées en 2009. Les cubes sont séchés de 0h à 8h dans un dessiccateur contenant une solution saturée de NaCl 75 % RH puis coupés en deux moitiés. La teneur en eau après séchage est déterminée à partir de la première moitié (cf 2.2.1) tandis que la viabilité du tissu est évaluée par incubation de la deuxième moitié 24h dans une solution à 1 % (m/v) de tetrazolium (Sigma-Aldrich, France). Après incubation, les tissus vivants sont colorés en rouge contrairement aux tissus morts. Les cubes sont photographiés et la couleur rouge est quantifiée selon l'intensité des pixels par le programme ImageJ (http://rsb.info.nih.gov/ij/). La viabilité des tissus cotylédonaires a été mesurée sur quatre répétitions de séchage indépendants comportant chacun 50 à 100 cubes.

2.3-Mesure de sensibilité à l'ABA

La sensibilité à l'ABA de graines matures R108 (graines sèches post-maturées six mois), *Mtabi3-1* et *Mtabi3-2* (stade abscission) a été évaluée sur trois répétitions de 40 à 50 graines imbibées sur papier filtre à différentes concentrations en ABA (isomères mixés, Sigma, St Louis, MO, USA) à 20°C à l'obscurité. L'ABA est dissout dans le méthanol avant dilution dans l'eau à 500 µM puis dilution en cascade pour former une gamme décroissante. Les graines témoins ont été imbibées dans l'eau contenant 0,5 % de méthanol, ce qui correspond à la concentration de méthanol dans le test à 500 µM d'ABA. La germination est déterminée par percée de la radicule après 14 jours d'imbibition.

Pour les graines d'*A. thaliana*, plus sensibles à l'ABA, seules trois concentrations ont été testées : 0, 1 et 10 µM d'ABA.

3-Analyse bi-dimensionnelle du protéome stable à la chaleur

3.1-Extraction des protéines stables à la chaleur

Les protéines totales ont été extraites à partir de trois répétitions de 25 graines de *M. truncatula* (A17, R108, *Mtabi3-1* et *Mtabi3-2*) et de 400 mg de cotylédons issus d'au moins trois graines différentes de la récolte de 2009 de *C. australe* pour chaque répétition. L'extraction est réalisée selon le protocole de Boudet *et al.* (2006) dans 1 mL de tampon HEPES 50 mM / EDTA 1 mM / cocktail d'inhibiteurs de protéases Complete 1X (Roche), à partir d'échantillons broyés dans l'azote liquide avec mortier et pilon. Après 30 min sur glace et 15 min de centrifugation à 14000 rpm, le surnageant est récupéré et la concentration en protéines solubles totales est mesurée selon la méthode de Bradford (1976). Pour conserver les protéines stables à la chaleur selon la méthode utilisée par Chatelain *et al.* (2012), cette solution est diluée à 3 µg μL^{-1} dans le tampon d'extraction et incubée 10 min à 95°C sous agitation (300 rpm) dans un thermobloc (Thermomixer Comfort, Eppendorf). L'échantillon est incubé sur glace 15 min avant d'être centrifugé 15 min à 14000 rpm et 4°C. Les protéines stables à la chaleur, non précipitées par ce traitement, sont récupérées dans le surnageant et dosées par la méthode de Bradford (1976). Après 4h de précipitation en présence d'acide trichloroacétique (TCA) à 20% (v/v) sur glace et 20 min de centrifugation à 14000 rpm et 4°C, le culot est successivement lavé avec 100 µL d'acétone 80%, d'acétone 100%, d'éthanol 80% et d'éthanol 100% puis resuspendu dans 300 µL de tampon de réhydratation (Urée 6 M / Thiourée 2 M / CHAPS 4 % / dithiothreitol (DTT) 1 mM / biolytes 1 %). Après 36 h à 4°C, la concentration en protéines stables à la chaleur effectivement resuspendues est de nouveau dosée par la méthode de Bradford (1976) en réalisant la gamme témoin avec du tampon de réhydratation afin de prendre en compte la présence d'urée dans le tampon.

3.2-Isoélectrofocalisation

Pour l'analyse protéomique, 150 µg de protéines stables à la chaleur de graines de *C. australe* et de *M. truncatula* A17 ainsi qu'un mélange 1:1 des deux extraits (300 µg) ont été nécessaires. L'analyse a également été menée à partir de 150 µg de protéines stables à la chaleur de graines de *M. truncatula* R108, *Mtabi3-2*, *Mtabi3-1*. Les extraits sont ramenés à un volume final de 450 µL en ajoutant du tampon de réhydratation. La séparation selon le point isoélectrique (pI) des protéines est effectuée à l'aide de bandelettes de focalisation de

Tableau 6.1. Composition des gels de polyacrylamide 12 % 24 cm pour analyse protéomique bi-dimensionnelle. Protogel 40 % = solution Acrylamide / Bis (37,5:1) (Bio-Rad) ; APS = persulfate d'ammonium 40 % (m/v) ; Tampon de séparation 4X = Tris-HCl 1,5 M pH 8,8 / SDS 0,4 % (m/v) ; TEMED = Tetramethylethylenediamine

Réactif	**Concentration finale**	**Volume pour 1 gel**
Protogel 40 %	12 %	30 mL
Glycérol 80 % (v/v)	10 %	12,5 mL
Tampon de séparation 4X	1X	25 mL
APS 40 %	0,056 %	140 µL
TEMED 100 %	0,15 %	150 µL
H_2O	Pour volume final 100 mL	32,5 mL

24 cm sur lesquelles est immobilisé un gradient de pH 3-10 non linéaire (Bio-Rad, Hercules, CA, USA). Une fois imbibées de l'échantillon, les bandelettes sont recouvertes d'huile minérale pour éviter toute évaporation. L'isoélectrofocalisation (IEF) se déroule à 20°C dans une cuve d'isofocalisation (Protean IEF Cell, BIORAD) selon le programme suivant : réhydratation active de 3h à 250 V, puis une augmentation progressive du voltage à 6 kV pendant 4h. Ce voltage est maintenu jusqu'à cumuler 27 kVh (6 kVh^{-1} pendant 4,5h) avant d'augmenter à nouveau jusqu'à 8 kV. Ce voltage est ensuite maintenu jusqu'à cumuler 20 kVh (8 kVh^{-1} pendant 2,5h).

3.3-Electrophorèse sur gel de polyacrylamide en condition dénaturante.

Une fois séparées selon leur pI, les protéines doivent migrer selon leur poids moléculaire sur un gel de polyacrylamide en conditions dénaturantes (SDS-PAGE). Au préalable, les bandelettes sont dénaturées par une incubation de 15 min dans le tampon EQ1 [DTT 130 mM / Urée 8 M / Tris-HCl 375 mM pH 8,8 / Glycérol 20 % (v/v) / SDS 2 % (m/v)] puis une deuxième incubation de 15 min dans le tampon EQ2 [Iodoacétamide 250 mM / Urée 8 M / Tris-HCl 375 mM pH 8,8 / Glycérol 20 % (v/v) / SDS 2 % (m/v)]. Elles sont ensuite placées au sommet de gels de polyacrylamide 12 % (m/v), dont la composition est donnée dans le tableau 6.1, et recouverte d'une solution d'agarose 1 % (m/v) contenant du bleu de bromophénol afin de les maintenir et permettre la visualisation de la migration des protéines. La migration électrophorétique s'effectue à 10°C pendant 1h à 25 V et 20 h à 110 V dans le système Ettan DALTsix (Amersham Biosciences) avec un tampon de migration [Tris-HCl 15,6 mM / glycine 120 mM pH 8,0 / SDS 0,1 % (m/v)].

3.4-Coloration des gels et quantification des spots

Les protéines sont fixées dans le gel de polyacrylamide par une solution de fixation [acide acétique 2 % (v/v) / éthanol 40 % (v/v)] pendant 1 h sous agitation. Les gels sont ensuite colorés par incubation dans une solution de coloration [Brillant Blue G colloïdal concentrate (Sigma-Aldrich) 0,08 % (m/v) / méthanol 20 %] pendant 48h. Pour la décoloration, les gels sont tout d'abord brièvement placés dans un mélange acide acétique 5 % (v/v) / éthanol 25 % (v/v), puis dans plusieurs bains d'éthanol 25 % (v/v). Ils sont ensuite numérisés par le scanner GS 800 (Bio-Rad) à la résolution 63,5 x 63,5. Au moins trois gels issus de trois extractions indépendantes par condition sont analysés avec le logiciel PDQuest 7.2.0 software (Bio-Rad). Les intensités des spots sont normalisées selon la méthode de la quantité totale des spots valides. Un test-t a permis de révéler les différences significatives d'intensité entre les spots de *M. truncatula* et de *C. australe* et entre ceux de R108 et de *Mtabi3-1* / *Mtabi3-2*.

3.5-Identification des spots

110 spots d'intérêts ont été excisés du gel 2D pour être identifier. Tout d'abord, les spots sont rincés avec 100 µL de carbonate d'ammonium 25 mM pendant 1h et déshydratés avec 100 µL d'un mélange volume à volume d'acétonitrile et de carbonate d'ammonium 25 mM pendant 45 min à température ambiante. Les protéines sont ensuite réduites et alkylées 1h à 57°C en présence de DTT 10 mM puis d'iodoacétamide 55 mM pendant 45 min. Après un nouveau lavage au carbonate d'ammonium et une nouvelle déshydratation à l'acétonitrile 100 %, les spots sont réhydratés avec 10 µL de trypsine à 15 ng μL^{-1} (Sequencing grade, Promega, Madison, WI) dans du carbonate d'ammonium 25 mM. Après une nuit d'incubation à 37°C, les

fragments tryptiques sont extraits par un mélange d'acide formique 1 % dans de l'acétonitrile 70 % et analysés par un spectroscope LC-ESI-MS/MS avec HPLC (Famos-Switchos-Ultimate system, LC Packings, Dionex, San Francisco, CA, USA) couplé à un spectromètre de masse Q-TOF (hybrid quadrupole orthogonal acceleration time-of-flight) (Global, Micromass-Waters, Manchester, UK). Ces données générées sur la plate-forme BIBS (Biopolymères, Biologie Structurale ; INRA Nantes) sont ensuite traitées avec le programme Protein Lynx Global Server (Micromass-Waters). La comparaison avec les séquences des bases de données Uniprot (version août 2010) et TIGR Medicago EST permet l'identification des protéines. Pour le protéome stable à la chaleur de *M. truncatula*, les spots liés aux protéines LEA sont identifiés grâce au gel de référence publié par Chatelain *et al.* (2012).

4-Détermination de la teneur en sucres solubles des cotylédons de *C. australe*

La teneur en sucres solubles (raffinose, stachyose, verbascose, saccharose, fructose et glucose) a été mesurée dans les cotylédons de stade mi-mature et mature de *C. australe* récoltés en 2009. Après lyophilisation pendant 24 h, la masse sèche des échantillons est pesée avant broyage au broyeur à bille (TissueLyser, QUIAGEN) à une fréquence de 30 s^{-1} pendant 2 min. 40 µL d'une solution de mélézitose à 4 g/L, servant de témoin interne, et 1 mL d'une solution de méthanol 80 % sont ajoutés au broyat. Les échantillons sont ensuite incubés 15 min à 76°C, puis le méthanol est évaporé au speed-vac (Mi Vac quatro concentrator, GENEVAC) pendant 2 h à 40°C. Le culot est repris dans 1 mL d'eau ultra-pure et centrifugé 5 min à 4°C à 13500 g (R5417, Eppendorf). Le surnageant est dilué entre 25 et 50 fois. Les échantillons sont analysés par chromatographie ionique haute performance en phase liquide, en utilisant le système DIONEX (Dionex Corp., Sunnyvale, CA, USA), contenant une précolonne AS11 ion exchanger et une colonne Dionex CarboPac PA1 (2 x 250 mm). Les sucres sont détectés par ampérométrie après séparation sur une colonne CarboPac PA1 avec un débit constant à 0.25 mL/min de NaOH 0,1 M. La concentration en sucres des échantillons est déterminée par comparaison avec un échantillon standard dont les quantités de sucres sont connues. La concentration du verbascose est quant à elle calculée sur la base de la surface du pic chromatographique correspondant au stachyose du à l'absence de standard pur. Les teneurs en sucres sont exprimées en mg de sucre par mg de MS.

5-Séquençage du transcriptome de *C. australe*

5.1-Extraction des ARN et séquençage des banques d'ADNc

Un mélange volume à volume de phénol pH 4 et de tampon d'extraction TLES [Tris 100 mM pH 8 / LiCl 100 mM / EDTA 10 mM pH8 / SDS 1 %] est chauffé quelques minutes à 80°C et versé sur 150 mg de poudre de tissus congelés (cf 1.2 pour la description des tissus) obtenus par broyage dans l'azote liquide avec mortier et pilon. Après 30 secondes d'agitation vigoureuse au vortex, un volume de chloroforme / alcool isoamylique (24:1) est ajouté. Le mélange est vortexé vigoureusement puis centrifugé 10 min à 14000 rpm et 4°C. La phase aqueuse supérieure est récupérée et les ARN sont précipités en présence de LiCl 2 M pendant une nuit à 4°C. Après 30 min de centrifugation à 14000 rpm à 4°C, le culot est repris dans 250 µL d'H_20 DEPC (diethylpyrocarbonate) puis 25 µL d'acétate de sodium 3M pH 5,2 et 500 µL d'éthanol absolu. Après 2h à -20°C et 45 min de centrifugation à 14000 rpm à 4°C, le culot est lavé par 1 mL d'éthanol 70 % puis mis à sécher 10 min à température ambiante pour être finalement repris dans 40 µL d' H_20 DEPC. La concentration en ARN est

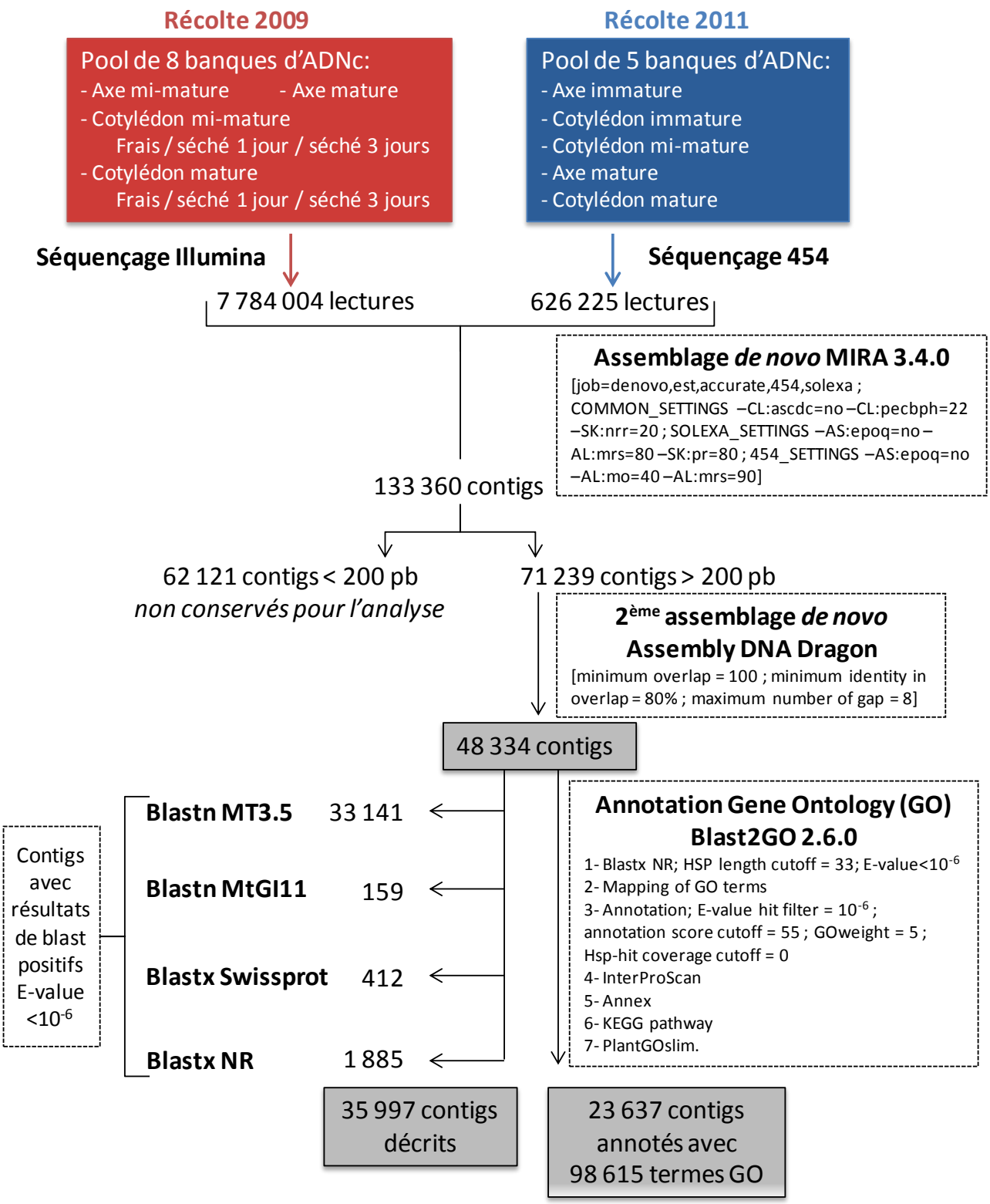


Figure 6.3. Procédure détaillée et paramètres appliqués pour les étapes de séquençage, assemblage et annotations du transcriptome de graines de *C. australe*.

mesurée par le spectrophotomètre NanoDrop ND-1000 UV-VIS (NanoDrop Technologies) et l'intégrité des ARN est vérifiée à l'aide du bioanalyseur Experion (Bio-Rad).

A partir des ARN extraits de cotylédons frais, séchés 1 ou 3 jours et d'axes embryonnaires récoltés en mai et juillet 2009 une banque d'ADNc représentative des gènes exprimés dans une graine de *C. australe* a été préparée, normalisée et séquencée par GenXPro GmbH (Francfort-sur-le-Main, Allemagne) avec la technologie Illumina Genome Analyser-IIx (Figure 6.3). Une autre banque d'ADNc a été réalisée à partir des ARN extraits des cotylédons et des axes immatures, mi-matures et matures de graines de *C. australe* récoltées en 2011 (cf 1.2). La normalisation et le séquençage ont été effectués par Eurofins MWG (Ebersberg, Allemagne) à l'aide de la technologie 454 GS FLX+ (Figure 6.3).

5.2-Assemblage des lectures en contigs

Les 8 410 229 lectures issues des deux séquençages ont été assemblées *de novo* en deux étapes : tout d'abord avec le programme MIRA 3.4.0 (Chevreux *et al.*, 2004) puis avec DNA Dragon (SequentiX, http://www.sequentix.de/software_dnadragon). La procédure détaillée ainsi que les paramètres utilisés pour l'assemblage est décrite dans la figure 6.3. Seuls les 48 334 contigs de taille supérieure à 200 pb ont été conservés pour l'annotation.

5.3-Annotations des contigs

5.3.1-Description par blast

Deux bases de données nucléotidiques spécifiques de *M. truncatula* ont été utilisées comme référence pour décrire les contigs de *C. australe* par similarité (blastn) : MT3.5v5 du groupe international d'annotation du génome de Medicago (IMGAG) http://jcvi.org/cgi-bin/medicago/download.cgi et MtGI11 (*Medicago truncatula* Gene Index version 11) hébergé par le DFCI (Dana-Farber Cancer Institute) http://compbio.dfci.harvard.edu/tgi/cgi-bin/tgi/gimain.pl?gudb=medicago. Deux versions de Mt3.5 ont été compilées pour augmenter le nombre de résultats : la version publique MT3.5v5 (14/10/2011) et une version mise à disposition par P. Gamas et J. Gouzy (Laboratoire des Interactions Plantes-Microorganismes, Toulouse), utilisant un assemblage plus récent complété par des données RNAseq. A l'aide du programme Blast2GO 2.6.0 (Götz *et al.*, 2008) http://www.blast2go.com/b2ghome, les contigs non décrits après ces deux blasts ont été comparés à deux bases de données protéiques incluant toutes les espèces végétales (blastx) : Swissprot et la base non redondante (NR) du NCBI. Seuls les alignements associés à une E-value inférieure à 10^{-6} ont été considérés comme significatifs (Figure 6.3).

5.3.2-Annotation par Blast2GO

Afin de regrouper les annotations au sein de grands processus biologiques (BP), fonctions moléculaires (MF) ou compartiments cellulaires (CC), les contigs ont également été décrits à l'aide de la classification Gene Ontology (GO) par le programme Blast2GO 2.6.0 (Conesa *et al.*, 2005 ; Götz *et al.*, 2008). A partir des résultats de blast contre la base de données NR, les termes GO disponibles dans les bases de données publiques ont été associés à chaque contigs et les plus spécifiques ont été sélectionnés par une règle d'annotation. L'ensemble du processus d'annotation et des paramètres utilisés est décrit dans la figure 6.3.

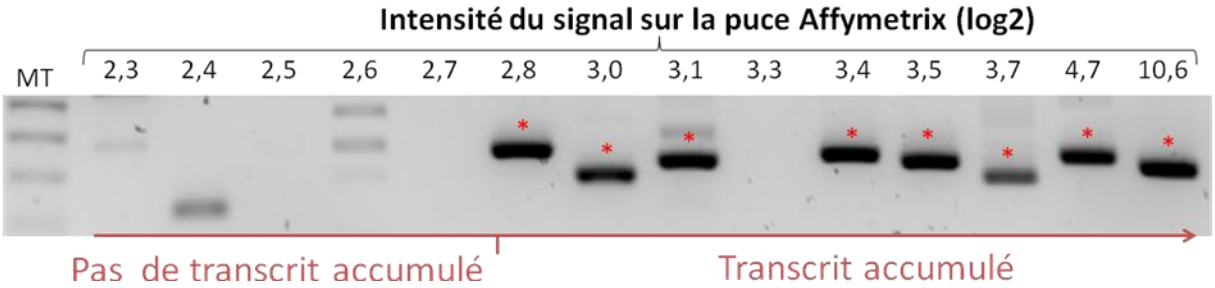


Figure 6.4. Détermination par RT-PCR du seuil d'intensité associé à une accumulation de transcrit sur la puce Affymetrix de *M. truncatula*.

Treize gènes associés à un signal d'intensité croissant sur la puce Affymetrix ont été sélectionnés. L'accumulation de transcrits est détectable après 35 cycles d'amplification pour les gènes associés à une intensité supérieure à 2,8. La présence de la bande attendue est signalée par une étoile rouge.

6-Détermination des gènes exprimés au cours du développement des graines de *M. truncatula*

Les gènes exprimés aux stades 20 DAF, 32 DAF et ABS du développement des graines de *M. truncatula* ont été déterminés à partir des données issues de la puce Affymetrix (*Benedito et al.*, 2008 ; Verdier *et al.*, 2013). Pour déterminer le seuil d'intensité pour lequel une accumulation de transcrit est détectable, une RT-PCR a été effectuée sur 13 gènes associés à signal d'intensité croissant. La bande attendue est observée uniquement pour une intensité supérieure à 2,8 (en log) (Figure 6.4). Respectivement, 32 655, 33 049, et 33 641 sondes sur les 48 116 que compte la puce remplissent ce critère. Les gènes associés à ces sondes ont été annotés par Blast2GO selon la même procédure que celle utilisée pour les contigs de *C. australe*. Ce jeu de données issu de la puce Affymetrix n'a été utilisé que pour l'analyse GO (Figure 3.4 et paragraphe 1.3.2 du chapitre III).

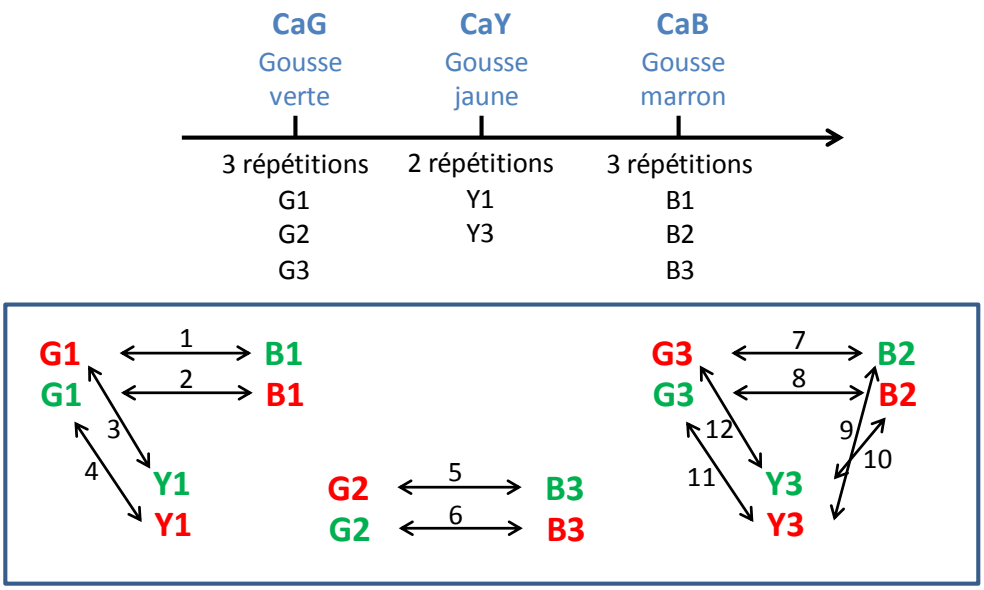
7-Analyse transcriptomique du développement des graines de *C. australe*

7.1-Amplification des ARN

Les ARNm ne représentent qu'une très faible proportion des ARN extraits, ils doivent donc être spécifiquement amplifiés. Cette étape a été réalisée en deux temps : transcription inverse des ARNm en ADNc puis transcription in vitro à partir de ces ADNc pour obtenir des ARN amplifiés (ARNa). Le kit Message AmpII aRNA amplification (Ambion) a été utilisé avec quelques modifications dans le protocole. Les concentrations des réactifs n'étant pas précisées dans le protocole du kit, les volumes utilisés plutôt que les concentrations finales des réactifs sont décrits dans ce paragraphe.

Les ARN ont été extraits de cotylédons récoltés à trois stades de développement différents en 2011 selon le protocole décrit au paragraphe 5.1. Huit échantillons ont été utilisés : trois extractions d'ARN différentes pour le premier stade (immature ; gousse verte-jaune ; CaG) et le dernier stade (mature ; gousse marron ; CaB) et deux extractions pour le stade intermédiaire (mi-mature ; gousse jaune ; CaY). Pour la première étape de l'amplification, 300 ng d'ARN et 0,25 µL de T7-oligo-dT (Ambion) sont incubés 10 min à 65°C dans un volume final de 3 µL puis 2h à 42°C après ajout de 2 µL d'un mix « first strand » [0,5 µL de tampon « first strand »10X (Ambion) / 1 µL de dNTP (Ambion) / 0,25 µL d'inhibiteur de Rnase (Ambion) / 0,25 µL de Reverse transcriptase Array Script (Ambion)]. 20 µL d'un deuxième mix « second strand » [2,5 µL de tampon « second strand » 10X (Ambion) / 1 µL de dNTP (Ambion) / 0,25 µL de Rnase H (Ambion) / et 0,5 µL de DNA polymérase (Ambion)] sont ajoutés pour la synthèse du deuxième brin d'ADNc pendant 2h à 16°C. Après ajout de 75 µL d'H_20 DEPC, les ADNc sont purifiés à l'aide du kit DNAclear (Ambion) selon les instructions du fabricant et l'élution finale est réalisée avec 2 x 5 µL d'H_20 DEPC préchauffée à 55°C.

La transcription *in vitro* est réalisée par ajout de 6 µL d'un mix [1 µL d'ATP, de GTP, de CTP et d'UTP 75 mM chacun (Ambion) / 1 µL de tampon de réaction T7 10X (Ambion) / 1 µL de T7 enzyme mix (Ambion)] et incubation pendant 2h à 37°C. Après ajout de 86 µL d'H_20 DEPC, les ARNa sont purifiés à l'aide du kit MEGAclear (Ambion) selon les instructions du fabricant et l'élution finale est réalisée avec 2 x 50 µL d'H_20 DEPC préchauffée à 55°C. La concentration en ARNa est mesurée par le spectrophotomètre NanoDrop ND-1000 UV-VIS (NanoDrop Technologies) tandis que leur intégrité est évaluée par dépôt de 500 ng sur gel d'agarose 1% TBE 0,5X.



Analyse en DYE SWAP : 12 puces sur la lame

→ 6 comparaisons graine immature (CaG) / graine mature (CaB) (puces 1/2/5/6/7/8)

→ 4 comparaisons graine immature (CaG) / graine mi-mature (CaY) (puces 3/4/11/12)

→ 2 comparaisons graine mi-mature (CaY) / graine mature (CaB) (puces 9/10)

Figure 6.5. Plan d'hybridation des échantillons pour l'analyse du transcriptome de graines de *C. australe*.
Les cotylédons de trois stades de développement de la graine de *C. australe* ont été utilisés pour l'analyse : un stade immature gousse verte-jaune (CaG), un stade mi-mature gousse jaune (CaY), un stade mature gousse marron (CaB). Pour le premier et le dernier stade, trois extractions d'ARN différentes ont pu être effectuées mais seulement deux pour le stade mi-mature. L'analyse a été faite en 'dye swap', c'est-à-dire que chaque comparaison entre deux répétitions a été faite deux fois avec inversion des marquages à la Cy5 (vert) et à la Cy3 (rouge). L'analyse complète entre cotylédons immatures et matures compte donc 2x3=6 comparaisons tandis que les analyses complètes avec les cotylédons mi-matures impliquent 2x2=4 comparaisons. Par manque de place sur la lame (qui ne compte que 12 puces), la comparaison Y1 / B1 n'a pas été effectuée pour le moment. Le numéro de la puce utilisée pour chaque comparaison est indiqué par les nombres de 1 à 12.

Les ARNa sont finalement concentrés par ajout de 1/100 de volume d'acrylamide (pour rendre le culot plus visible), de 1/10 de volume d'acétate de sodium 3 M pH 5,2 et de 2,5 volumes d'éthanol absolu. Après 1h à -20°C et 45 min de centrifugation à 14000 rpm et 4°C, le culot est lavé avec 1 mL d'éthanol 70 % puis repris dans le volume d'H_20 DEPC permettant d'obtenir une concentration d'environ 1 µg $µL^{-1}$.

7.2-Marquage des ADNc et précipitation des cibles

Le marquage est effectué par incorporation de cyanine rouge (Cy3) ou verte (Cy5) au cours d'une transcription inverse des ARNa. Un mélange de 5 µg d'ARNa, 2 µg de Random nonamères (Sigma) et 20 u d'inhibiteur de RNase (Ribolock, Thermo Scientific) estt mis à incuber à 65°C pendant 10 min dans un volume final de 10,5 µL puis 2h à 42°C après ajout de 20 µL d'un mix « marquage » [tampon « first strand » (Invitrogen) 1X final / DTT 10 mM final (Invitrogen) / mix [dDTP 500 µM final /dCTP 100 µM final] / Cy5 ou Cy3 selon l'échantillon 75 µM final (Interchim) / 200 u de Superscript revertase III (Invitrogen)]. Au bout de 1h à 42°C, 200 u de Superscript revertase sont rajoutés au mélange, soit une concentration finale utilisée pour cette étape de 2X. Après ajout de 79 µL d'H_20 DEPC, les ADNc marqués (les cibles) sont purifiés à l'aide du kit Nucleospin PCR Clean-up (Macherey-Nagel) selon les instructions du fabricant et l'élution finale est réalisée avec 2 x 21 µL d'H_20 DEPC préchauffée à 55°C. La concentration en ADNc (ng/µL) et en Cy5 (ou Cy3) (pmol/µL) est mesurée par le spectrophotomètre NanoDrop ND-1000 UV-VIS (NanoDrop Technologies).

L'analyse transcriptomique repose sur la compétition entre les ADNc d'un échantillon marqué à la Cy5 et ceux d'un deuxième échantillon marqués à la Cy3 pour l'hybridation des sondes sur une puce. Ces deux échantillons sont hybridés sur la puce en un seul dépôt et doivent donc être mélangés avant l'hybridation. Pour cela, 30 pmol de cibles marquées à la Cy5 (échantillon 1) et à la Cy3 (échantillon 2) sont précipitées ensemble selon le protocole décrit à la fin de l'amplification des ARN (cf 7.1).

7.3-Plan d'hybridation

Trois extractions d'ARN différentes ont été utilisées pour le premier et le dernier stade et deux extractions pour le stade intermédiaire. Ces échantillons ont été hybridés en 'dye-swap', approche adaptée aux analyses avec peu d'échantillons et plus robuste d'un point de vue statistique en augmentant le nombre de comparaisons (Figure 6.5) (Churchill, 2002). Le swap est un groupe d'échantillons correspondant à une seule comparaison biologique. Pour nos données sur *C. australe* trois swaps ont été étudiés : CaG/CaB (immature/mature) ; CaG/CaY (immature/mi-mature) ; CaY/CaB (mi-mature/mature). Chaque comparaison entre deux répétitions a été faite deux fois avec inversion des marquages à la Cy5 (vert) et à la Cy3 (rouge). L'analyse complète entre cotylédons immatures et matures compte donc 2x3=6 comparaisons tandis que les analyses complètes avec les cotylédons mi-matures impliquent 2x2=4 comparaisons. Par manque de place sur la lame (qui ne compte que 12 puces), la comparaison Y1 / B1 n'a pas été effectuée pour le moment. La comparaison entre graine mi-mature et mature est donc incomplète.

7.4-Hybridation des puces et scan des lames

Chaque culot contenant les ADNc marqués des deux échantillons hybridés sur la puce est resuspendu dans 2,2 µL de STC (Sample Tracking Control) (NimbleGen) puis 5,8 µL d'un mix de solution d'hybridation

[tubes 1, 2 et 3 du kit d'hybridation (NimbleGen)] avant d'être incubé 5 min à 95°C. 6,3 µL d'échantillon sont déposés sur chacune des puces via le mixer fixé sur la lame, puis cette dernière est placée 16-20h à 42°C dans la station d'hybridation. Le mixer est ensuite décollé de la lame dans une solution de lavage I [tampon lavage I (NimbleGen) / DTT 100 µM] préchauffée à 42°C puis la lame est successivement lavée 2 min avec la solution de lavage I, 1 min avec la solution de lavage II [tampon lavage II (Nimblegen) / DTT 100 µM] et 15 sec avec la solution de lavage III [tampon lavage III (NimbleGen) / DTT 100 µM].

Les lames 12-plex NimbleGen sont analysées par le scanner MS 200 (NimbleGen) et le logiciel MS 200 v1.2 avec une résolution de 2 µm. Le vert (532 nm) et le rouge (635 nm) sont scannés indépendamment, par conséquent deux fichiers TIFF sont générés pour une lame, respectivement nommés lame_puce_date_532.tif et lame_puce_date _635.tif.

7.5-Traitement statistiques

Les fichiers d'image des scans ont été traités par le logiciel DEVA v1.2.1 ; les signaux sont identifiés par alignement de la grille d'identification des sondes et convertis en valeurs numériques. Ces données sont exportées au format pair (données brutes) pour chaque couleur (532 nm et 635 nm) : "lame_puce_date_532.pair" et "lame_puce_date_635.pair". A l'aide des logiciels R v2.14 http://www.r-project.org, Rstudio http://www.rstudio.com et du script 'pair' développé au laboratoire, les fichiers de chaque couleur sont compilés en un fichier texte "lame_puce.txt" au format d'entrée pour une analyse par le package 'LIMMA' (Linear Models for Microarray data du projet Bioconductor http://www.bioconductor.org) pour l'analyse différentielle des signaux : liste pour chaque identifiant de sonde des intensités et du bruit de fond associé à 532 nm et 635 nm. Le script 'Anadiff' analyse les données et sort les résultats présentés sous différentes formes. Les valeurs d'intensité brutes subissent tout d'abord une transformation logarithmique puis sont linéarisées par l'application 'lmfit'. Elles sont ensuite normalisées par la méthode de Loess sans soustraction du bruit de fond (régression par intervalles). Pour l'analyse différentielle des intensités, le package 'LIMMA', permet d'utiliser un modèle Bayésiens d'évaluation statistique de la différence d'expression par modération des écart-types. C'est un test-t stable et puissant adapté aux expériences contenant un faible nombre de répétitions. Un contrôle des faux positifs est fait sur les p-value par le package kerfdr (AgroParisTech) utilisant la procédure de Benjamini-Hochberg (BH). Les sondes sont dites significativement sur ou sous-exprimées si la p-value corrigée (BH) est inférieur à 5 %.

Pour notre étude comparative entre *C. australe* et *M. truncatula* nous avons travaillé sur les valeurs d'intensités. Pour l'analyse des intensités, un calcul du bruit de fond est fait sur la moyenne des 10 000 intensités les plus faibles à laquelle on ajoute deux fois l'écart type. Ce bruit de fond est soustrait des intensités normalisées et les valeurs négatives sont remises à zéro. Ces valeurs d'intensités sont exportées en format texte : "norm_intensities_bkg_swap_txt". En parallèle de ces analyses, un fichier de traçabilité et d'assurance qualité est automatiquement généré. Il identifie entre autres le nombre de sondes pour lesquelles la valeur absolue du ratio est supérieure à 1 et la p-value supérieure à 1 %. Ces données sont écartées par l'analyse statistique car trop variables entre les différentes répétitions. Un autre contrôle qualité surveillé est la variance du swap : il doit se situer dans une fourchette de 0,04 à 0,1. Le swap est un groupe d'échantillons correspondant à une seule comparaison biologique. Pour nos données sur *C. australe* trois swaps ont été

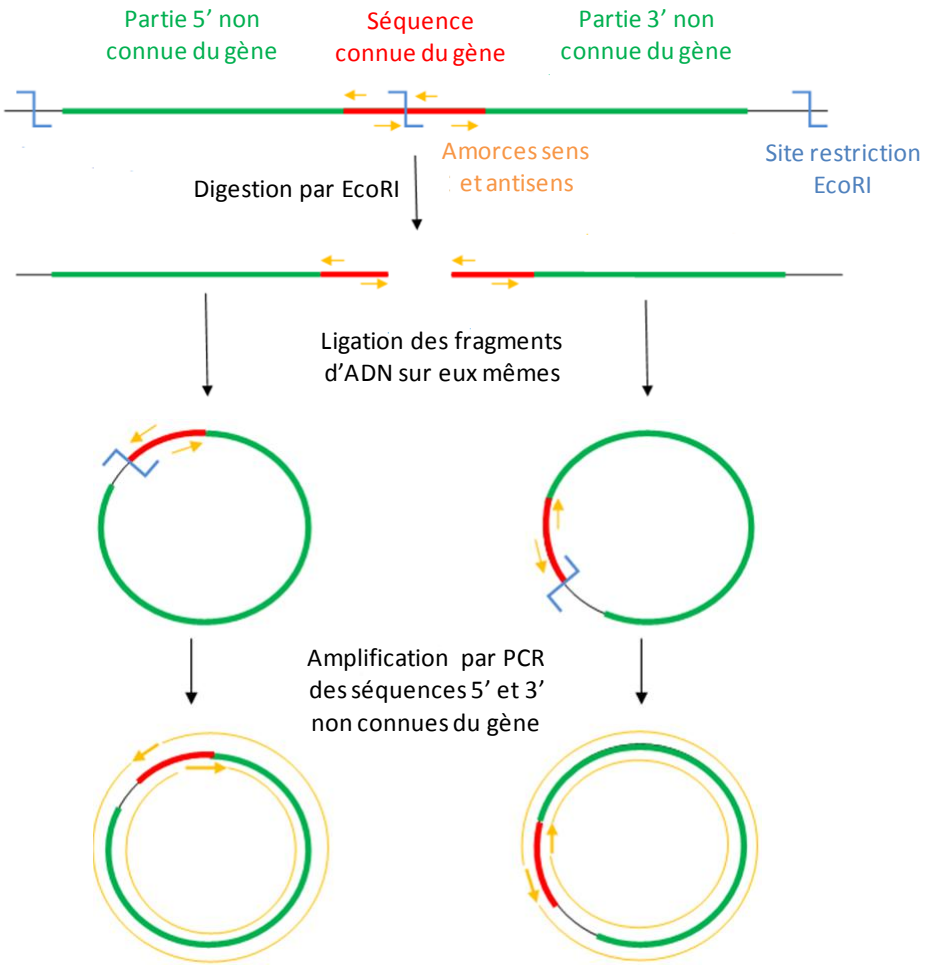


Figure 6.6. Principe de la PCR inverse utilisée pour le clonage des séquences complètes des gènes *MtABI3* et *CaABI3*.

La PCR inverse est utilisée pour cloner les régions 5' et 3' du gène (en vert) à partir d'une région centrale connue (en rouge). L'ADN génomique est digéré par une enzyme de restriction ayant un site de restriction (en bleu) dans la séquence connue. Puis deux fragments, contenant respectivement la région 5' et la région 3' du gène ainsi qu'une partie de la région centrale connue, sont ligués sur eux même dans un grand volume réactionnel contenant la T4 DNA ligase. Un couple d'amorces (en orange) est dessiné dans la région connue pour chacun de ces fragments. Les amorces sont orientées dans le sens inverse par rapport à une PCR classique et permettent d'amplifier les régions non connues par PCR.

étudiés : CaG/CaB (immature/mature) ; CaG/CaY (immature/mi-mature) ; CaY/CaB (mi-mature/mature). Le nombre de sondes écartées est respectivement de 141, 130 et 382 pour les swaps CaG/CaB , CaG/CaY et CaY/CaB, sur un total de 96 250 sondes. La variance est respectivement de 0,066, 0,031 et 0,044. Ces valeurs très faibles nous permettent d'analyser les intensités avec confiance.

8-Analyse fonctionnelle de *CaABI3*

8.1-Clonage des séquences complètes des gènes *MtABI3* et *CaABI3*

L'ADN génomique des deux espèces a été extrait avec le kit Nucleospin food (Macherey Nagel) à partir de feuilles de *C. australe* et de *M. truncatula* R108. La séquence du gène *MtABI3* était déjà partiellement disponible dans la base de données MtGI11 sous forme d'un EST (TC193312). Pour *CaABI3*, un premier fragment du gène a été cloné par PCR dégénérée. Pour cela, un couple d'amorces dégénérées a été dessiné dans les zones les plus conservées du gène, en se basant sur l'alignement des séquences *ABI3* de différentes espèces végétales disponibles dans Genbank (Annexe 8 pour les amorces). Le produit PCR a été cloné dans le vecteur pJet1.2 (CloneJET kit, Thermo Scientific, Bremen, Germany) et séquencé. Les séquences complètes ont été obtenues par inverse PCR à partir de 5 µg d'ADN génomique digéré par EcoRI (25 u / 50 µL volume final ; Promega, Madison, WI, USA) puis ligué par la T4 DNA ligase (50 u / 450 µL volume final ; Fermentas, Vilnius, Lituanie). La ligation s'est faite dans un grand volume pour favoriser la ligation de l'ADN sur lui-même et obtenir ainsi un fragment d'ADN génomique circularisé qui a servi de matrice pour la PCR inverse. A partir de la région connue des gènes *CaABI3* et *MtABI3*, de nouveaux couples d'amorces ont été dessinés pour amplifier les séquences complètes qui ont été clonées dans le vecteur pJet1.2 (CloneJET kit, Thermo Scientific, Bremen, Germany) et séquencés. Le principe de la PCR inverse est décrit dans la figure 6.6.

8.2-Construction des vecteurs d'expression pour la surexpression des gènes *ABI3* et la complémentation du mutant *abi3-5* d'*A. thaliana*

8.2.1-Les vecteurs binaires et le clonage Gateway

L'analyse fonctionnelle de *CaABI3* comprend d'une part une sur-expression du gène dans des racines de *M. truncatula* et d'autre part une complémentation du mutant *abi3-5* d'*A. thaliana*. Ces deux expérimentations nécessitent donc deux types de vecteurs d'expression différents. Pour la sur-expression dans les racines, les séquences ATG/stop (avec les introns) des gènes *CaABI3*, *CaABI3-like* et *MtABI3* ont été introduites en aval d'un promoteur constitutif 35S dans le vecteur de destination pK7WG2D (Karimi *et al.*, 2002) par le système Gateway (Invitrogen). Pour la complémentation, la séquence codante de *CaABI3* et *AtABI3* (contrôle) a été introduite en aval du promoteur de *AtABI3* (la séquence en amont de *CaABI3* n'ayant pas pu être clonée par PCR inverse) dans le vecteur de destination pKGW (Karimi *et al.*, 2002) par le système Gateway Multisite (Invitrogen). Les vecteurs de destination pK7WG2D et pKGW ainsi que le principe du clonage Gateway sont décrits dans la figure 6.7.

Le vecteur pK7WG2D est dit binaire, c'est-à-dire qu'une partie est exprimé dans la bactérie tandis que la région comprise entre les bordures gauche (LB) et droite (RB) est transférée à la plante. Cette région contient un promoteur 35S permettant de sur-exprimer constitutivement le gène *ABI3* inséré. Elle contient également un gène codant pour la GFP (Green Fluorescent Protein) afin d'identifier rapidement sous UV les

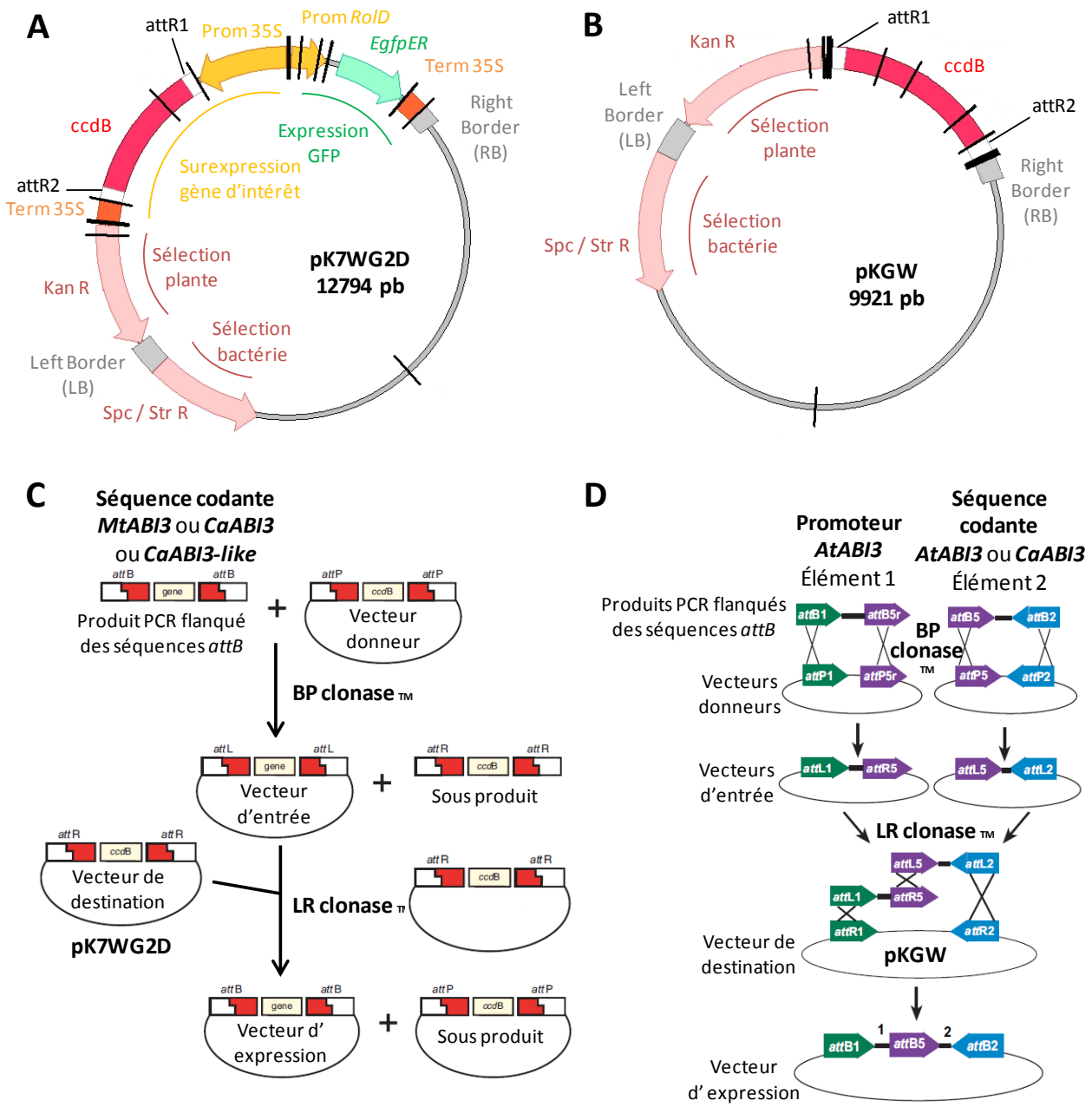


Figure 6.7. Caractéristiques des vecteurs de destination pK7WG2D, pKGW et principe du clonage Gateway monosite et multisite.

Les vecteurs de destination pK7WG2D **(A)** et pKGW **(B)** ont respectivement été utilisés pour la sur-expression du gène *ABI3* dans les racines de *M. truncatula* et pour la complémentation des mutants *abi3-5* d'*A. thaliana* (Karimi *et al.*, 2002). Prom 35S = promoteur constitutif 35S ; ccdB = gène codant pour la toxine « control of cell death » ; Term 35S = terminateur 35S ; Kan R = gène de résistance à la kanamycine ; Spc / Str R = Gène de résistance à la spectinomycine et à la streptomycine ; Prom *RolD* = promoteur permettant l'expression du gène codant pour la GFP (Green Fluorescent Protein) ; *EgfpER* = gène codant pour la GFP avec un signal d'adressage au réticulum endoplasmique (RE). Les systèmes Gateway **(C)** et Gateway Multisite **(D)** (Invitrogen) ont respectivement permis de créer les vecteurs d'expression pour la surexpression du gène *ABI3* dans les racines de *M. truncatula* et pour la complémentation des mutants *abi3-5* d'*A. thaliana*.

racines transformées. Le vecteur pKGW est également binaire et permet le transfert du gène *ABI3* sous le contrôle de son propre promoteur dans les plantes. Pour les vecteurs pKGW et pK7WG2D, le gène de résistance à la spectinomycine / streptomycine (Spc/Str R) permet la sélection des bactéries transformées alors que le gène de résistance à la kanamycine (Kan R) permet la sélection des racines ou plantules transformées (Figure 6.7A et 6.7B).

La BP clonase permet la recombinaison d'un produit PCR flanqué des sites de recombinaison attB avec un vecteur donneur contenant les sites de recombinaison attP pour former un vecteur d'entrée contenant le gène d'intérêt à la place du gène *ccdB* (toxique pour les bactéries). Notons que la recombinaison entre un site attB et un site attP donne un site attL dans le vecteur d'entrée. Dans un deuxième temps, la LR clonase permet d'insérer le gène d'intérêt dans un vecteur de destination, contenant les éléments nécessaires à sa sur-expression et les gènes de résistance, pour former le vecteur d'expression par recombinaison entre les sites attL du vecteur d'entrée et les sites attR du vecteur de destination (Figure 6.7C). Le vecteur d'expression pour la complémentation du mutant nécessite l'insertion de deux éléments distincts : une séquence promotrice et une séquence codante. Le clonage Gateway Multisite (Invitrogen) est basé sur le même principe que celui décrit précédemment. Deux vecteurs d'entrée possédant des sites attL différents ont été crées par la BP clonase : le premier contient la séquence promotrice et le deuxième la séquence codante. Ces deux séquences d'intérêt sont associées dans un même vecteur de destination pour former le vecteur d'expression par la LR clonase (Figure 6.7D).

8.2.2-Construction des vecteurs d'expression

Les séquences codantes des gènes *AtABI3*, *MtABI3*, *CaABI3* et *CaABI3-like* ainsi que la séquence promotrice de *AtABI3* (1 883 pb en amont du codon initiateur) ont été amplifiées par PCR à l'aide d'amorces sens positionnées au niveau du codon initiateur (ou 1 883 pb en amont pour le promoteur) et d'amorces antisens positionnées sur le codon stop (ou juste avant le codon initiateur pour le promoteur). Pour le clonage Multisite, ces amorces sont flanquées des sites attB1 / attB5r / attB5 ou attB2 comme indiqué sur la figure 6.7D (Annexe 8 pour les amorces). Pour le clonage Gateway classique, la réaction BP est facultative car il est possible de créer le vecteur d'entrée directement dans un vecteur pENTR-TOPO par clonage du produit PCR avec une amorce sens flanqué de la séquence CACC, cette deuxième option a été choisie. La polymérase 'proof reading' Phusion (Invitrogen) a été utilisée pour l'amplification PCR dans le milieu réactionnel suivant : tampon Phusion HF (contenant le $MgCl_2$) 1X final (Invitrogen)/ dNTP 200 µM final (Promega) / amorce sens 0,5 µM final (Eurogentec) / amorce antisens 0,5 µM final (Eurogentec) / 20 ng d'ADN génomique / Phusion polymerase 0,02 u final (Invitrogen).

Ces produits PCR ont été directement clonés dans les vecteurs donneurs pDONR (Gateway Multisite) ou pENTR-TOPO-D (Gateway) selon les recommandations du fabricant. Des bactéries compétentes *E. coli* One shot Top10 (Invitrogen) ont été transformées par ces vecteurs d'entrée à partir de 2 µL de produit de clonage par choc thermique (30 min sur glace / 30 sec à 42°C / 2 min sur glace). Les clones positifs sont sélectionnés sur LB kanamycine 50 µg/mL et vérifiés par PCR sur colonies. Les inserts sont séquencés à l'aide d'amorces internes pour s'assurer que les séquences protéiques prédites ne comportent aucune mutation par rapport aux séquences de référence (Annexe 8 pour les amorces).

Tableau 6.2. Composition du milieu Fahräeus additionnée de kanamycine et fongicide pour la culture des racines transformées.
Micro éléments = $MnSO_4$ $CuSO_4$ $ZnSO_4$ H_3BO_3 Na_2MoO_4 ; Agar HP696 est spécialement utilisé pour la culture *in vitro*. La solution a été autoclavée avant ajout du CaCl2 et de la kanamycine.

	Réactif	**Concentration finale**
Constituants du milieu Fahräeus classique	$MgSO_4$	0,5 mM
	KH_2PO4	0,7 mM
	Na_2HPO_4	0,8 mM
	Fe-EDTA	50 µM
	Micro élements	0,1 µg/mL
	$CaCl_2$	1 mM final
	HCl	5 µL/L
	Agar HP696	15 g/L
Composés spécifiques à cette étude	Kanamycine	25 µg/mL
	Fongicide (PPM)	1/1000^{e}

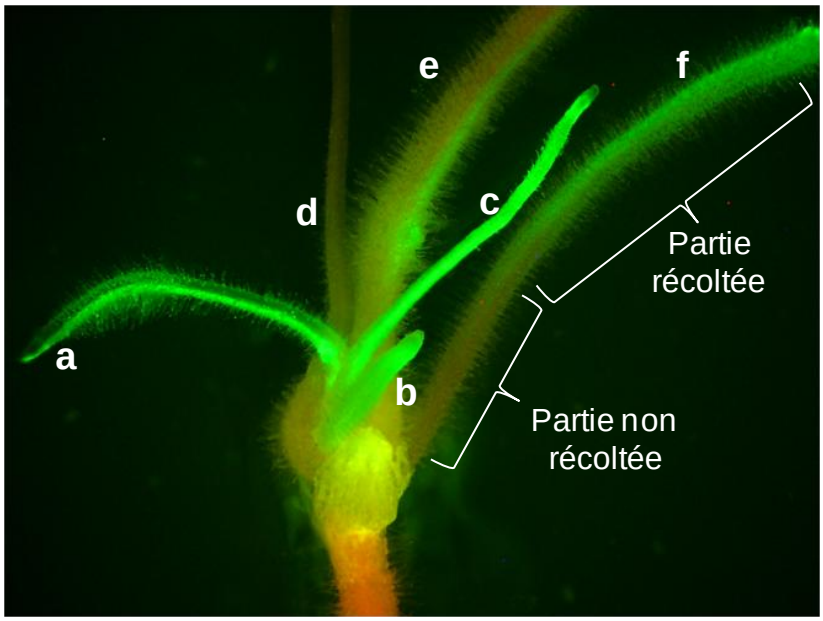


Figure 6.8. Racines transformées sur-exprimant la GFP.
Sous UV (488 nm), les racines sur-exprimant la GFP émettent une fluorescence verte (a, b, c). Cependant toutes les racines ne sont pas transformées (d et e) et dans le cas où une racine n'est pas entièrement verte (f), seule la partie fluorescente a été prélevée (photo : David Lalanne).

La LR clonase II plus enzyme mix (Invitrogen) a permis la construction des vecteurs d'expression à partir de 50 ng de chaque vecteur d'entrée et 150 ng de vecteur de destination suivant les instructions du fabricant. Des bactéries compétentes *E. coli* One shot Mach I (Invitrogen) sont transformées par ces vecteurs d'expression à partir de 2 µL de produit de clonage par choc thermique (30 min sur glace / 30 sec à 42°C / 2 min sur glace). Les clones positifs sont sélectionnés sur LB spectinomycine et streptomycine (50 µg/mL) et vérifiés par PCR sur colonies. Les inserts sont de nouveau séquencés pour s'assurer que les promoteurs sont bien en amont des séquences codantes (Annexe 8 pour les amorces).

Par choc thermique ([5 min azote liquide / 5 min 37°C] x 2), des *Agrobacterium rhizogenes* compétentes (souche Arqua-1) ont été transformées avec 1 µg de vecteur d'expression (*MtABI3*-pK7WG2D ou *CaABI3*-pK7WG2D ou *CaABI3-like*-pK7WG2D ou pK7WG2D vide). De même, des *Agrobacterium tumefaciens* compétentes (souche Agl1) ont été transformées avec 1 µg de vecteur d'expression (*At:AtABI3*-pKGW ou *At:CaABI3*-pKGW ou pKGW vide). Après 48h à 28°C, les bactéries transformées sont sélectionnées sur LB additionné de spectinomycine et streptomycine (50 µg/ml). Pour *A. tumefaciens* Agl1, souche naturellement résistante à la rifampicine (Rif) et à la carbénicilline (Carb), ces deux antibiotiques sont également ajoutés au milieu de sélection respectivement à la concentration de 20 µg/ml et 100 µg/ml. La présence des inserts attendus dans les clones est vérifiée par PCR sur colonies, les clones positifs sont conservés à -80°C dans le glycérol.

8.3-Sur-expression dans les racines de *M. truncatula*

8.3.1-Transformation des racines

Pour chacune des constructions, des racines de R108 (type sauvage) et de *Mtabi3-2* ont été transformées (Boisson-Dernier *et al.*, 2001). L'utilisation du mutant permet de s'affranchir de l'influence éventuelle de l'*ABI3* endogène qui peut être exprimé au cours du développement dans les racines. Après stérilisation à l'eau de javel 2,6 % 3 min, les graines (post-maturées pour R108 et au stade abscission pour *Mtabi3-2*) sont mises à germer sur eau à 20°C (pendant 30 h pour R108 et 48h pour *Mtabi3-2*) afin d'obtenir des radicules d'environ 1 cm. En parallèle, les clones d' *A. rhirogenes* contenant les vecteurs d'expression sont cultivées à 28°C afin d'obtenir un tapis bactérien (pré-culture 24 h dans 5 mL de LB liquide additionné de 50 µg/ mL de Spc et de Str sous agitation 200 rpm puis culture 24 h sur boîte LB additionné de 50 µg/ mL de Spc et de Str à partir de 450 µL de pré-culture).

Après excision d'environ 2 mm de racine et trempage dans le tapis bactérien, les plantules sont cultivées pendant environ 6 semaines à 20°C photopériode 16h sur milieu Fahräeus, dont la composition est donnée dans le tableau 6.2, additionné de 25 µg/mL de kanamycine (antibiotique de sélection) et d'un fongicide dilué 1/1000 (PPM pour Plant Preservative Mixture, Kalys Biotechnologies). Les plantules sont régulièrement repiquées sur de nouvelles boîtes. Seules les racines effectivement transformées exprimant le gène codant pour la GFP sont récoltées, elles fluorescent en vert sous UV (488 nm) (Figure 6.8). Après congélation immédiate dans l'azote liquide, elles sont stockées à -80°C.

8.3.2-Détermination des cibles de *MtABI3* et *CaABI3* par analyse transcriptomique

Les ARN de racines transformées (trois répétitions pour chaque construction ; environ 10 racines par répétition) sont extraits avec le kit Nucleospin RNA plant (Macherey Nagel). Le broyage des racines est effectué dans l'azote liquide avec mortier et pilon, puis 350 µL de tampon d'extraction RA3 additionné de B-mercaptoéthanol dilué 1/1000 (Sigma-Aldrich) et d'une pointe de spatule de PVP (polyvinylpyrrolidone) et PVPP (polyvinylpolypyrrolidone) sont ajoutés à la poudre obtenue. Les instructions du fabricant sont suivies pour le reste de l'extraction et les ARN sont élués dans 2 x 50 µL d'H_20 DEPC préchauffée à 55°C. Après concentration (cf 7.1), le culot d'ARN est repris dans 30 µL d'H_20 DEPC. L'analyse transcriptomique est ensuite réalisée sur les puces Nimblegen 102k *M. truncatula* selon le protocole décrit au paragraphe 7 afin de déterminer les gènes sur et sous-exprimés en réponse à la sur-expression des différents gènes.

8.4-Complémentation du mutant *abi3-5* d'*A. thaliana*.

8.4.1-Transformation des plantes

Les premières fleurs formées sur la hampe des mutants *abi3-5* ont été transformées par la technique du « floral dip » dans une solution d'*A. tumefaciens* contenant le vecteur d'expression. Trois constructions différentes ont été utilisées pour la transformation : *At:AtABI3*-pKGW, *At:CaABI3*-pKGW et pKGW vide (cf 8.2.2 pour la création de ces constructions et leur intégration dans *A. tumefaciens*). Une pré-culture de ces trois clones bactériens est réalisée dans 5 mL de LB liquide additionné des antibiotiques adéquats (Spc et Str 50 µg/ mL ; Rif 20 µg/ mL ; Carb 100 µg/mL) sous agitation 200 rpm à 28°C pendant 24h. Puis 300 mL de LB additionné des antibiotiques adéquats sont inoculés avec 1 mL de préculture et cultivés sous agitation jusqu'à atteindre une DO de 1,8 (environ 28 h). Les bactéries sont alors précipitées par centrifugation 5 min à 4000 rpm et resuspendues dans un tampon d'infiltration ($MgCl_2$ 10 mM / saccharose 5 % / Silwet 0,02 %). Les plantes sont immergées 30 secondes dans cette solution bactérienne puis placées sous cloche sur du papier imbibé pendant 48h pour conserver une humidité importante. Les graines sont récoltées au stade ouverture des siliques et stérilisées à l'éthanol 100 %.

8.4.2-Sélection et caractérisation phénotypique des plantes transformées

Les graines récoltées pour chacune des trois lignées *abi3-5* complémentées : *AtABI3*/*abi3-5*, *CaABI3*/*abi3-5* et pKGW/*abi3-5* sont mises à germer sur milieu de sélection (½ MS ; 0,8 % agar ; kanamycine [25 µg/mL] ; PPM [fongicide]). Après deux semaines, les plantules toujours vertes sont transférées sur terreau et un screening PCR est effectué pour confirmer la présence de la construction dans l'ADN génomique de ces plantes (Annexe 8 pour les amorces et le gel associé). Une seule lignée *CaABI3*/*abi3-5* a survécu à la sélection kanamycine et a été validée par PCR, deux lignées *AtABI3*/*abi3-5* ont survécu à la sélection kanamycine mais seule la première a été validée par PCR et les trois lignées pKGW/*abi3-5* sélectionnées sur kanamycine ont été validées par PCR.

Les graines sauvages, mutantes *abi3-5* non complémentées et complémentées *AtABI3*/*abi3-5* (lignée 1), *CaABI3*/*abi3-5* et pKGW/*abi3-5* (lignée 1) ont été caractérisées sur cinq aspects du phénotype : couleur de la graine à la déhiscence des siliques, dormance, sensibilité à l'ABA, tolérance à la dessiccation et longévité. La dormance est mise en évidence par comparaison des cinétiques de germination sur eau à 20°C 16h photopériode avec et sans stratification préalable 3 jours à 4°C obscurité. La sensibilité à l'ABA est évaluée par

germination sur ABA (isomères cis/trans, Sigma, St Louis, MO, USA) 1 µM et 10 µM à 20°C 16h photopériode. La TD des graines est déterminée par germination sur eau à 20°C 16h photopériode après 48h dans un dessiccateur K_2CO_3 44 % RH. La longévité des graines est testée par vieillissement accéléré à 35°C NaCl 75 % RH pendant 2, 4, 6, 8 et 10 jours.

Références bibliographiques

Aalen RB. **1999**. Peroxiredoxin antioxidants in seed physiology. *Seed Science Research* **9**: 285–295.

Abe H, Urao T, Ito T, Seki M, Shinozaki K, Yamaguchi-Shinozaki K. **2003**. Arabidopsis AtMYC2 (bHLH) and AtMYB2 (MYB) Function as Transcriptional Activators in Abscisic Acid Signaling. *The Plant Cell* **15**: 63–78.

Abirached-Darmency M, Abdel-gawwad MR, Conejero G, Verdeil JL, Thompson R. **2005**. In situ expression of two storage protein genes in relation to histo-differentiation at mid-embryogenesis in *Medicago truncatula* and *Pisum sativum* seeds. *Journal of Experimental Botany* **56**: 2019–2028.

Agarwal PK, Agarwal P, Reddy MK, Sopory SK. **2006**. Role of DREB transcription factors in abiotic and biotic stress tolerance in plants. *Plant Cell Reports* **25**: 1263–1274.

Agarwal PK, Jha B. **2010**. Transcription factors in plants and ABA dependent and independent abiotic stress signalling. *Biologia Plantarum* **54**: 201–212.

Alamillo J, Almoguera C, Bartels D, Jordano J. **1995**. Constitutive expression of small heat shock proteins in vegetative tissues of the resurrection plant Craterostigma plantagineum. *Plant Molecular Biology* **29**: 1093–1099.

Almoguera C, Prieto-Dapena P, Díaz-Martín J, Espinosa JM, Carranco R, Jordano J. **2009**. The HaDREB 2 transcription factor enhances basal thermotolerance and longevity of seeds through functional interaction with HaHSFA 9. *BMC Plant Biology* **9**: 75.

Alonso R, Oñate-Sánchez L, Weltmeier F, Ehlert A, Diaz I, Dietrich K, Vicente-Carbajosa J, Dröge-Laser W. **2009**. A pivotal role of the basic leucine zipper transcription factor bZIP53 in the regulation of Arabidopsis seed maturation gene expression based on heterodimerization and protein complex formation. *The Plant cell* **21**: 1747–1761.

Alpert P. **2005**. The Limits and Frontiers of Desiccation-Tolerant Life. *Integrative and Comparative Biology* **45**: 685–695.

Alpert P, Oliver MJ. **2002**. Drying without dying. In: *Desiccation and survival in plants: Drying without dying*, Black M, Pritchard HW, eds. Wallingford, UK, CABI Publishing. 3-43.

Alsheikh MK, Svensson JT, Randall SK. **2005**. Phosphorylation regulated ion-binding is a property shared by the acidic subclass dehydrins. *Plant, Cell & Environment* **28**: 1114–1122.

Amara I, Odena A, Oliveira E, Moreno A, Masmoudi K, Pagès M, Goday A. **2012**. Insights into Maize LEA Proteins: From Proteomics to Functional Approaches. *Plant and Cell Physiology* **53**: 312–329.

Amara I, Capellades M, Ludevid MD, Pagès M, Goday A. **2013**. Enhanced water stress tolerance of transgenic maize plants over-expressing LEA Rab28 gene. *Journal of Plant Physiology* **170**: 864–873.

Amudha J, Balasubramani G. **2011**. Recent molecular advances to combat abiotic stress tolerance in crop plants. *Biotechnology and Molecular Biology Reviews* **6**: 31–58.

Angelovici R, Galili G, Fernie AR, Fait A. **2010**. Seed desiccation: a bridge between maturation and germination. *Trends in Plant Science* **15**: 211–218.

Araújo S de S, Duque ASRLA, Santos DMMF dos, Fevereiro MPS. **2004**. An Efficient Transformation Method to Regenerate a High Number of Transgenic Plants Using a New Embryogenic Line of *Medicago truncatula* cv. Jemalong. *Plant Cell, Tissue and Organ Culture* **78**: 123–131.

Arrigoni O, Gara LD, Tommasi F, Liso R. **1992**. Changes in the Ascorbate System during Seed Development of Vicia faba L. *Plant Physiology* **99**: 235–238.

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, *et al.* **2000**. Gene Ontology: tool for the unification of biology. *Nature genetics* **25**: 25–29.

Avelange-Macherel M-H, Ly-Vu B, Delaunay J, Richomme P, Leprince O. **2006**. NMR metabolite profiling analysis reveals changes in phospholipid metabolism associated with the re-establishment of desiccation tolerance upon osmotic stress in germinated radicles of cucumber. *Plant, Cell & Environment* **29**: 471–482.

Bailly C, Audigier C, Ladonne F, Wagner MH, Coste F, Corbineau F, Côme D. **2001**. Changes in oligosaccharide content and antioxidant enzyme activities in developing bean seeds as related to acquisition of drying tolerance and seed quality. *Journal of Experimental Botany* **52**: 701 –708.

Bailly C. **2004**. Active oxygen species and antioxidants in seed biology. *Seed Science Research* **14**: 93–107.

Bailly C, Leymarie J, Lehner A, Rousseau S, Côme D, Corbineau F. **2004**. Catalase activity and expression in developing sunflower seeds as related to drying. *Journal of Experimental Botany* **55**: 475–483.

Baker J, dennSteele CV, Iii LD. **1988**. Sequence and characterization of 6 Lea proteins and their genes from cotton. *Plant Molecular Biology* **11**: 277–291.

Banzet N, Richaud C, Deveaux Y, Kazmaier M, Gagnon J, Triantaphylidès C. **1998**. Accumulation of small heat shock proteins, including mitochondrial HSP22, induced by oxidative stress and adaptive response in tomato cells. *The Plant Journal* **13**: 519–527.

Bartels D. **2005**. Desiccation Tolerance Studied in the Resurrection Plant *Craterostigma plantagineum*. *Integrative and Comparative Biology* **45**: 696–701.

Bartels D, Sunkar R. **2005**. Drought and Salt Tolerance in Plants. *Critical Reviews in Plant Sciences* **24**: 23–58.

Bassel GW, Mullen RT, Bewley JD. **2006**. ABI3 expression ceases following, but not during, germination of tomato and Arabidopsis seeds. *Journal of Experimental Botany* **57**: 1291–1297.

Bateman A, Birney E, Cerruti L, Durbin R, Etwiller L, Eddy SR, Griffiths-Jones S, Howe KL, Marshall M, Sonnhammer ELL. **2002**. The Pfam Protein Families Database. *Nucleic Acids Research* **30**: 276–280.

Battaglia M, Olvera-Carrillo Y, Garciarrubio A, Campos F, Covarrubias AA. **2008**. The Enigmatic LEA Proteins and Other Hydrophilins. *Plant Physiology* **148**: 6–24.

Baud S, Mendoza MS, To A, Harsco\ët E, Lepiniec L, Dubreucq B. **2007**. WRINKLED1 specifies the regulatory action of LEAFY COTYLEDON2 towards fatty acid metabolism during seed maturation in Arabidopsis. *The Plant Journal* **50**: 825–838.

Bäumlein H, Miséra S, Luerßen H, Kölle K, Horstmann C, Wobus U, Müller AJ. **1994**. The FUS3 gene of *Arabidopsis thaliana* is a regulator of gene expression during late embryogenesis. *The Plant Journal* **6**: 379–387.

Bazin J, Langlade N, Vincourt P, Arribat S, Balzergue S, El-Maarouf-Bouteau H, Bailly C. **2011**. Targeted mRNA Oxidation Regulates Sunflower Seed Dormancy Alleviation during Dry After-Ripening. *The Plant Cell* **23**: 2196–2208.

Bellin D, Ferrarini A, Chimento A, Kaiser O, Levenkova N, Bouffard P, Delledonne M. **2009**. Combining next-generation pyrosequencing with microarray for large scale expression analysis in non-model species. *BMC Genomics* **10**: 555.

Benedito VA, Torres-Jerez I, Murray JD, Andriankaja A, Allen S, Kakar K, Wandrey M, Verdier J, Zuber H, Ott T, *et al.* 2008. A gene expression atlas of the model legume *Medicago truncatula*. *The Plant Journal* **55**: 504–513.

Bennett MD, Bhandol P, Leitch IJ. **2000**. Nuclear DNA Amounts in Angiosperms and their Modern Uses—807 New Estimates. *Annals of Botany* **86**: 859–909.

Bensmihen S, Rippa S, Lambert G, Jublot D, Pautot V, Granier F, Giraudat J, Parcy F. **2002**. The homologous ABI5 and EEL transcription factors function antagonistically to fine-tune gene expression during late embryogenesis. *The Plant Cell* **14**: 1391–1403.

Bensmihen S, Giraudat J, Parcy F. **2005**. Characterization of three homologous basic leucine zipper transcription factors (bZIP) of the ABI5 family during Arabidopsis thaliana embryo maturation. *Journal of Experimental Botany* **56**: 597–603.

Bentsink L, Alonso-Blanco C, Vreugdenhil D, Tesnier K, Groot SPC, Koornneef M. **2000**. Genetic Analysis of Seed-Soluble Oligosaccharides in Relation to Seed Storability of Arabidopsis. *Plant Physiology* **124**: 1595–1604.

Berger SL. **2007**. The complex language of chromatin regulation during transcription. *Nature* **447**: 407–412.

Berjak P, Vertucci CW, Pammenter NW. **1993**. Effects of developmental status and dehydration rate on characteristics of water and desiccation-sensitivity in recalcitrant seeds of *Camellia sinensis*. *Seed Science Research* **3**: 155–166.

Berjak P, Pammenter NW. **2000**. What ultrastructure has told us about recalcitrant seeds. *Revista Brasileira de Fisiologia Vegetal* **12**: 22–55.

Berjak P, Pammenter NW. **2002**. Orthodox and recalcitrant seeds. In: *Tropical tree seed manual*, Vozzo JA ed, USDA Forest Service, Washington USA. 137-147

Berjak P. **2006**. Unifying perspectives of some mechanisms basic to desiccation tolerance across life forms. *Seed Science Research* **16**: 1–15.

Berjak P, Farrant J, Pammenter N.. **2007**. Seed desiccation tolerance mechanisms. In: Jenks MA and Wood AJ eds. *Plant Desiccation Tolerance*. Blackwell publishing, Iowa USA, 151-192.

Berjak P, Pammenter NW. **2008**. From Avicennia to Zizania: Seed Recalcitrance in Perspective. *Annals of Botany* **101**: 213 –228.

Bettey M, Finch-Savage WE. **1998**. Stress protein content of mature Brassica seeds and their germination performance. *Seed Science Research* **8**: 347–355.

Bewley JD. **1979**. Physiological aspects of desiccation tolerance. *Annual Review of Plant Physiology* **30**: 195–238.

Bewley J. **1997**. Seed Germination and Dormancy. *The Plant Cell* **9**: 1055–1066.

Bewley JD, Bradford KJ, Hilhorst HW, Nonogaki H. **2013**. Seeds physiology of development, germination and dormancy, 3rd edition. Springer, New York USA.

Bies-Etheve N, da Silva Conceicao A, Giraudat J, Koornneef M, Léon-Kloosterziel K, Valon C, Delseny M. **1999**. Importance of the B2 domain of the Arabidopsis ABI3 protein for Em and 2S albumin gene regulation. *Plant molecular biology* **40**: 1045–1054.

Bies-Etheve N, Gaubier-Comella P, Debures A, Lasserre E, Jobet E, Raynal M, Cooke R, Delseny M. **2008**. Inventory, evolution and expression profiling diversity of the LEA (late embryogenesis abundant) protein gene family in Arabidopsis thaliana. *Plant Molecular Biology* **67**: 107–124.

Black M, Corbineau F, Gee H, Côme D. **1999**. Water Content, Raffinose, and Dehydrins in the Induction of Desiccation Tolerance in Immature Wheat Embryos. *Plant Physiology* **120**: 463 –472.

Blackman SA, Obendorf RL, Leopold AC. **1992**. Maturation Proteins and Sugars in Desiccation Tolerance of Developing Soybean Seeds. *Plant Physiology* **100**: 225–230.

Blöchl A, Peterbauer T, Richter A. **2007**. Inhibition of raffinose oligosaccharide breakdown delays germination of pea seeds. *Journal of Plant Physiology* **164**: 1093–1096.

Blöchl A, Peterbauer T, Hofmann J, Richter A. **2008**. Enzymatic breakdown of raffinose oligosaccharides in pea seeds. *Planta* **228**: 99–110.

Boisson-Dernier A, Chabaud M, Garcia F, Bécard G, Rosenberg C, Barker DG. **2001**. *Agrobacterium rhizogenes* - Transformed Roots of *Medicago truncatula* for the Study of Nitrogen-Fixing and Endomycorrhizal Symbiotic Associations. *Molecular Plant-Microbe Interactions* **14**: 695–700.

Bossi F, Cordoba E, Dupré P, Mendoza MS, Román CS, León P. **2009**. The Arabidopsis ABA-INSENSITIVE (ABI) 4 factor acts as a central transcription activator of the expression of its own gene, and for the induction of ABI5 and SBE2.2 genes during sugar signaling. *The Plant Journal* **59**: 359–374.

Boubriak I, Dini M, Berjak P, Osborne DJ. **2000**. Desiccation and survival in the recalcitrant seeds of Avicennia marina: DNA replication, DNA repair and protein synthesis. *Seed Science Research* **10**: 307–315.

Boucher V, Buitink J, Lin X, Boudet J, Hoekstra FA, Hundertmark M, Renard D, Leprince O. **2010**. MtPM25 is an atypical hydrophobic late embryogenesis-abundant protein that dissociates cold and desiccation-aggregated proteins. *Plant, Cell & Environment* **33**: 418–430.

Boudet J, Buitink J, Hoekstra FA, Rogniaux H, Larré C, Satour P, Leprince O. **2006**. Comparative Analysis of the Heat Stable Proteome of Radicles of *Medicago truncatula* Seeds during Germination Identifies Late Embryogenesis Abundant Proteins Associated with Desiccation Tolerance. *Plant Physiology* **140**: 1418 –1436.

Bourgis F, Kilaru A, Cao X, Ngando-Ebongue G-F, Drira N, Ohlrogge JB, Arondel V. **2011**. Comparative transcriptome and metabolite analysis of oil palm and date palm mesocarp that differ dramatically in carbon partitioning. *Proceedings of the National Academy of Sciences* **108**: 12527-12532.

Boursiac Y, Léran S, Corratgé-Faillie C, Gojon A, Krouk G, Lacombe B. ABA transport and transporters. *Trends in Plant Science* **18**: 325–333.

Bove J, Lucas P, Godin B, Ogé L, Jullien M, Grappin P. **2005**. Gene expression analysis by cDNA-AFLP highlights a set of new signaling networks and translational control during seed dormancy breaking in *Nicotiana plumbaginifolia*. *Plant Molecular Biology* **57**: 593–612.

Bowler C, Montagu MV, Inze D. **1992**. Superoxide Dismutase and Stress Tolerance. *Annual Review of Plant Physiology and Plant Molecular Biology* **43**: 83–116.

Bradford MM. **1976**. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical biochemistry* **72**: 248–254.

Bradford KJ, Chandler PM. **1992**. Expression of' dehydrin-like' proteins in embryos and seedlings of Zizania palustris and Oryza sativa during dehydration. *Plant physiology* **99**: 488.

Brady SM, Sarkar SF, Bonetta D, McCourt P. **2003**. The ABSCISIC ACID INSENSITIVE 3 (ABI3) gene is modulated by farnesylation and is involved in auxin signaling and lateral root development in Arabidopsis. *The Plant Journal* **34**: 67–75.

Bräutigam A, Gowik U. **2010**. What can next generation sequencing do for you? Next generation sequencing as a valuable tool in plant research. *Plant Biology* **12**: 831–841.

Bräutigam A, Mullick T, Schliesky S, Weber APM. **2011**. Critical assessment of assembly strategies for non-model species mRNA-Seq data and application of next-generation sequencing to the comparison of C3 and C4 species. *Journal of Experimental Botany* **62**: 3093–3102.

Bravo LA, Gallardo J, Navarrete A, Olave N, Martínez J, Alberdi M, Close TJ, Corcuera LJ. **2003**. Cryoprotective activity of a cold-induced dehydrin purified from barley. *Physiologia Plantarum* **118**: 262–269.

Bray EA. **1993**. Molecular Responses to Water Deficit. *Plant Physiology* **103**: 1035–1040.

Braybrook SA, Stone SL, Park S, Bui AQ, Le BH, Fischer RL, Goldberg RB, Harada JJ. **2006**. Genes directly regulated by LEAFY COTYLEDON2 provide insight into the control of embryo maturation and somatic embryogenesis. *Proceedings of the National Academy of Sciences* **103**: 3468-3473.

Braybrook SA, Harada JJ. **2008**. LECs go crazy in embryo development. *Trends in plant science* **13**: 624–630.

Brocard-Gifford IM, Lynch TJ, Finkelstein RR. **2003**. Regulatory Networks in Seeds Integrating Developmental, Abscisic Acid, Sugar, and Light Signaling. *Plant Physiology* **131**: 78–92.

Browne J, Tunnacliffe A, Burnell A. **2002**. Anhydrobiosis: Plant desiccation gene found in a nematode. *Nature* **416**: 38–38.

Bruni F, Leopold AC. **1992**. Cytoplasmic glass formation in maize embryos. *Seed Science Research* **2**: 251–253.

Bryant G, Koster KL, Wolfe J. **2001**. Membrane behaviour in seeds and other systems at low water content: the various effects of solutes. *Seed Science Research* **11**: 17–25.

Buitink J, Walters-Vertucci C, Hoekstra FA, Leprince O. **1996**. Calorimetric Properties of Dehydrating Pollen (Analysis of a Desiccation-Tolerant and an Intolerant Species). *Plant Physiology* **111**: 235–242.

Buitink J, Claessens MMAE, Hemminga MA, Hoekstra FA. **1998**. Influence of Water Content and Temperature on Molecular Mobility and Intracellular Glasses in Seeds and Pollen. *Plant Physiology* **118**: 531–541.

Buitink J, van den Dries IJ, Hoekstra FA, Alberda M, Hemminga MA. **2000**. High Critical Temperature above Tg May Contribute to the Stability of Biological Systems. *Biophysical Journal* **79**: 1119–1128.

Buitink J, Ly Vu B, Satour P, Leprince O. **2003**. The Re-Establishment of Desiccation Tolerance in Germinated Radicles of *Medicago Truncatula* Gaertn. Seeds. *Seed Science Research* **13**: 273–286.

Buitink J, Leprince O. **2004**. Glass formation in plant anhydrobiotes: survival in the dry state. *Cryobiology* **48**: 215–228.

Buitink J, Thomas M, Gissot L, Leprince O. **2004**. Starvation, osmotic stress and desiccation tolerance lead to expression of different genes of the regulatory β and γ subunits of the SnRK1 complex in germinating seeds of *Medicago truncatula*. *Plant, Cell & Environment* **27**: 55–67.

Buitink J, Leger JJ, Guisle I, Vu BL, Wuillème S, Lamirault G, Bars AL, Meur NL, Becker A, Küster H, *et al.* 2006. Transcriptome profiling uncovers metabolic and regulatory processes occurring during the transition from desiccation-sensitive to desiccation-tolerant stages in *Medicago truncatula* seeds. *Plant Journal* **47**: 735–750.

Buitink J, Leprince O. **2008**. Intracellular glasses and seed survival in the dry state. *Comptes Rendus Biologies* **331**: 788–795.

Busk PK, Jensen AB, Pagès M. **1997**. Regulatory elements in vivo in the promoter of the abscisic acid responsive gene rab17 from maize. *The Plant Journal* **11**: 1285–1295.

Caccere R, Teixeira SP, Centeno DC, Figueiredo-Ribeiro R de CL, Braga MR. Metabolic and structural changes during early maturation of Inga vera seeds are consistent with the lack of a desiccation phase. *Journal of Plant Physiology* **170**: 791–800.

Cadman CSC, Toorop PE, Hilhorst HWM, Finch-Savage WE. **2006**. Gene expression profiles of Arabidopsis Cvi seeds during dormancy cycling indicate a common underlying dormancy control mechanism. *The Plant Journal* **46**: 805–822.

Cairns NG, Pasternak M, Wachter A, Cobbett CS, Meyer AJ. **2006**. Maturation of Arabidopsis seeds is dependent on glutathione biosynthesis within the embryo. *Plant Physiology* **141**: 446.

Campalans A, Pagès M, Messeguer R. **2001**. Identification of differentially expressed genes by the cDNA-AFLP technique during dehydration of almond (*Prunus amygdalus*). *Tree Physiology* **21**: 633–643.

Campanella JJ, Bitincka L, Smalley J. **2003**. MatGAT: an application that generates similarity/identity matrices using protein or DNA sequences. *BMC bioinformatics* **4**: 29.

Candat A. **2012**. Analyse de la localisation subcellulaire des protéines LEA (Late Embryogenesis Abundant) chez *Arabidopsis thaliana* par des approches de bioinformatique et de biologie cellulaire (thèse).

Carles C, Bies-Etheve N, Aspart L, Léon-Kloosterziel KM, Koornneef M, Echeverria M, Delseny M. **2002**. Regulation of *Arabidopsis thaliana* Em genes: role of ABI5. *The Plant Journal* **30**: 373–383.

Carmi N, Zhang G, Petreikov M, Gao Z, Eyal Y, Granot D, Schaffer AA. **2003**. Cloning and functional expression of alkaline α-galactosidase from melon fruit: similarity to plant SIP proteins uncovers a novel family of plant glycosyl hydrolases. *The Plant Journal* **33**: 97–106.

Chabaud M, de Carvalho-Niebel F, Barker DG. **2003**. Efficient transformation of *Medicago truncatula* cv. Jemalong using the hypervirulent *Agrobacterium tumefaciens* strain AGL1. *Plant cell reports* **22**: 46–51.

Chaitanya K, Naithani SC. **1994**. Role of superoxide, lipid peroxidation and superoxide dismutase in membrane perturbation during loss of viability in seeds of Shorea robusta Gaertn.f. *New Phytologist* **126**: 623–627.

Chakrabortee S, Tripathi R, Watson M, Schierle GSK, Kurniawan DP, Kaminski CF, Wise MJ, Tunnacliffe A. **2011**. Intrinsically disordered proteins as molecular shields. *Molecular BioSystems* **8**: 210–219.

Chandra Babu R, Zhang J, Blum A, David Ho T-H, Wu R, Nguyen HT. **2004**. HVA1, a LEA gene from barley confers dehydration tolerance in transgenic rice (*Oryza sativa* L.) via cell membrane protection. *Plant Science* **166**: 855–862.

Chatelain E, Hundertmark M, Leprince O, le Gall S, Satour P, Deligny-Penninck S, Rogniaux H, Buitink J. **2012**. Temporal profiling of the heat stable proteome during the late maturation of *Medicago truncatula* seeds identifies a restricted subset of late embryogenesis abundant proteins associated with longevity. *Plant, Cell & Environment* **35**: 1440–1455.

Che NY, Yang Y, Li YD, Wang LL, Huang P, Gao Y, An CC. **2009**. Efficient LEC2 activation of OLEOSIN expression requires two neighboring RY elements on its promoter. *Science in China Series C: Life Sciences* **52**: 854–863.

Cheng W-H, Endo A, Zhou L, Penney J, Chen H-C, Arroyo A, Leon P, Nambara E, Asami T, Seo M, *et al.* 2002a. A Unique Short-Chain Dehydrogenase/Reductase in Arabidopsis Glucose Signaling and Abscisic Acid Biosynthesis and Functions. *The Plant Cell* **14**: 2723–2743.

Cheng Z, Targolli J, Huang X, Wu R. **2002b**. Wheat LEA genes, PMA80 and PMA1959, enhance dehydration tolerance of transgenic rice (*Oryza sativa* L.). *Molecular Breeding* **10**: 71–82.

Cheng X, Wen J, Tadege M, Ratet P, Mysore KS. **2011**. Reverse genetics in *Medicago truncatula* using Tnt1 insertion mutants. In: *Methods in Molecular Biology. Plant Reverse Genetics*. Pereira A, ed., Humana Press, 179–190.

Chevreux B, Pfisterer T, Drescher B, Driesel AJ, Müller WEG, Wetter T, Suhai S. **2004**. Using the miraEST Assembler for Reliable and Automated mRNA Transcript Assembly and SNP Detection in Sequenced ESTs. *Genome Research* **14**: 1147–1159.

Chin HF, Roberts EH. **1980**. *Recalcitrant crop seeds.* Tropical Press, Kuala Lumpur, Malaysia.

Chinnusamy V, Ohta M, Kanrar S, Lee B, Hong X, Agarwal M, Zhu J-K. **2003**. ICE1: a regulator of cold-induced transcriptome and freezing tolerance in Arabidopsis. *Genes & Development* **17**: 1043–1054.

Choi H, Hong J, Ha J, Kang J, Kim SY. **2000**. ABFs, a Family of ABA-responsive Element Binding Factors. *Journal of Biological Chemistry* **275**: 1723–1730.

Chono M, Honda I, Shinoda S, Kushiro T, Kamiya Y, Nambara E, Kawakami N, Kaneko S, Watanabe Y. **2006**. Field studies on the regulation of abscisic acid content and germinability during grain development of barley: molecular and chemical analysis of pre-harvest sprouting. *Journal of Experimental Botany* **57**: 2421–2434.

Churchill GA. **2002**. Fundamentals of experimental design for cDNA microarrays. *Nature Genetics* **32**: 490–495.

Clerkx EJM, Blankestijn-De Vries H, Ruys GJ, Groot SPC, Koornneef M. **2004**. Genetic differences in seed longevity of various Arabidopsis mutants. *Physiologia Plantarum* **121**: 448–461.

Close TJ, Kortt AA, Chandler PM. **1989**. A cDNA-based comparison of dehydration-induced proteins (dehydrins) in barley and corn. *Plant Molecular Biology* **13**: 95–108.

Close TJ, Fenton RD, Moonan F. **1993**. A view of plant dehydrins using antibodies specific to the carboxy terminal peptide. *Plant Molecular Biology* **23**: 279–286.

Close TJ. **1997**. Dehydrins: A commonalty in the response of plants to dehydration and low temperature. *Physiologia Plantarum* **100**: 291–296.

Coca MA, Almoguera C, Jordano J. **1994**. Expression of sunflower low-molecular-weight heat-shock proteins during embryogenesis and persistence after germination: localization and possible functional implications. *Plant Molecular Biology* **25**: 479–492.

Coca MA, Almoguera C, Thomas TL, Jordano J. **1996**. Differential regulation of small heat-shock genes in plants: analysis of a water-stress-inducible and developmentally activated sunflower promoter. *Plant Molecular Biology* **31**: 863–876.

Collada C, Gomez L, Casado R, Aragoncillo C. **1997**. Purification and in Vitro Chaperone Activity of a Class I Small Heat-Shock Protein Abundant in Recalcitrant Chestnut Seeds. *Plant Physiology* **115**: 71–77.

Collett H, Shen A, Gardner M, Farrant JM, Denby KJ, Illing N. **2004**. Towards transcript profiling of desiccation tolerance in *Xerophyta humilis*: Construction of a normalized 11 k *X. humilis* cDNA set and microarray expression analysis of 424 cDNAs in response to dehydration. *Physiologia Plantarum* **122**: 39–53.

Colmenero-Flores JM, Campos F, Garciarrubio A, Covarrubias* AA. **1997**. Characterization of *Phaseolus vulgaris* cDNA clones responsive to water deficit: identification of a novel late embryogenesis abundant-like protein. *Plant Molecular Biology* **35**: 393–405.

Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M. **2005**. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **21**: 3674-3676.

Cook DR. **1999**. *Medicago truncatula* — a model in the making!: Commentary. *Current Opinion in Plant Biology* **2**: 301–304.

Coppe A, Pujolar J, Maes G, Larsen P, Hansen M, Bernatchez L, Zane L, Bortoluzzi S. **2010**. Sequencing, de novo annotation and analysis of the first *Anguilla anguilla* transcriptome: EeelBase opens new perspectives for the study of the critically endangered european eel. *BMC Genomics* **11**: 635.

Coppe A, Bortoluzzi S, Murari G, Marino IAM, Zane L, Papetti C. **2012**. Sequencing and Characterization of Striped Venus Transcriptome Expand Resources for Clam Fishery Genetics. *PLoS ONE* **7**: e44185.

Coraggio I, Tuberosa R. **2004**. Molecular bases of plant adaptation to abiotic stress and approaches to enhance tolerance to hostile environments. In: Christou P, Klee H, eds. *Handbook of plant biotechnology*. Wiley, Chichester, 413–468.

Corbineau F, Picard MA, Fougereux J-A, Ladonne F, Côme D. **2000**. Effects of dehydration conditions on desiccation tolerance of developing pea seeds as related to oligosaccharide content and cell membrane properties. *Seed Science Research* **10**: 329–339.

Crane C, Wright E, Dixon RA, Wang Z-Y. **2006**. Transgenic *Medicago truncatula* plants obtained from *Agrobacterium tumefaciens* -transformed roots and *Agrobacterium rhizogenes*-transformed hairy roots. *Planta* **223**: 1344–1354.

Crowe JH, Crowe LM, Carpenter JF, Aurell Wistrom C. **1987**. Stabilization of dry phospholipid bilayers and proteins by sugars. *Biochemical Journal* **242**: 1–10.

Crowe LM, Reid DS, Crowe JH. **1996**. Is trehalose special for preserving dry biomaterials? *Biophysical Journal* **71**: 2087–2093.

Crowe JH, Oliver AE, Tablin F. **2002**. Is There a Single Biochemical Adaptation to Anhydrobiosis? *Integrative and Comparative Biology* **42**: 497–503.

Crowe JH, Crowe LM, Wolkers WF, Oliver AE, Ma X, Auh J-H, Tang M, Zhu S, Norris J, Tablin F. **2005**. Stabilization of Dry Mammalian Cells: Lessons from Nature. *Integrative and Comparative Biology* **45**: 810–820.

Curaba J, Moritz T, Blervaque R, Parcy F, Raz V, Herzog M, Vachon G. **2004**. *AtGA3ox2*, a Key Gene Responsible for Bioactive Gibberellin Biosynthesis, Is Regulated during Embryogenesis by LEAFY COTYLEDON2 and FUSCA3 in Arabidopsis. *Plant Physiology* **136**: 3660–3669.

Cutler SR, Rodriguez PL, Finkelstein RR, Abrams SR. **2010**. Abscisic acid: emergence of a core signaling network. *Annual Reviews Plant Biology* **61**: 651–679.

Danyluk J, Perron A, Houde M, Limin A, Fowler B, Benhamou N, Sarhan F. **1998**. Accumulation of an Acidic Dehydrin in the Vicinity of the Plasma Membrane during Cold Acclimation of Wheat. *The Plant Cell Online* **10**: 623–638.

Daws MI, Garwood NC, Pritchard HW. **2006**. Prediction of Desiccation Sensitivity in Seeds of Woody Species: A Probabilistic Model Based on Two Seed Traits and 104 Species. *Annals of Botany* **97**: 667–674.

Daws MI, Lydall E, Chmielarz P, Leprince O, Matthews S, Thanos CA, Pritchard HW. **2004**. Developmental heat sum influences recalcitrant seed traits in *Aesculus hippocastanum* across Europe. *New Phytologist* **162**: 157–166.

Dean Rider S, Henderson JT, Jerome RE, Edenberg HJ, Romero-Severson J, Ogas J. **2003**. Coordinate repression of regulators of embryonic identity by PICKLE during germination in Arabidopsis. *The Plant Journal* **35**: 33–43.

Delmas F, Sankaranarayanan S, Deb S, Widdup E, Bournonville C, Bollier N, Northey JGB, McCourt P, Samuel MA. **2013**. ABI3 controls embryo degreening through Mendel's I locus. *Proceedings of the National Academy of Sciences* doi: 10.1073/pnas.1308114110.

Delseny M, Bies-Etheve N, Carles C, Hull G, Vicient C, Raynal M, Grellet F, Aspart L. **2001**. Late Embryogenesis Abundant (LEA) protein gene regulation during Arabidopsis seed maturation. *Journal of Plant Physiology* **158**: 419–427.

DeRocher AE, Vierling E. **1994**. Developmental control of small heat shock protein expression during pea seed maturation. *The Plant Journal* **5**: 93–102.

Ditzer A, Kirch H-H, Nair A, Bartels D. **2001**. Molecular characterization of two alanine-rich Lea genes abundantly expressed in the resurrection plant *C. plantagineum* in response to osmotic stress and ABA. *Journal of Plant Physiology* **158**: 623–633.

Djemel N, Guedon D, Lechevalier A, Salon C, Miquel M, Prosperi J-M, Rochat C, Boutin J-P. **2005**. Development and composition of the seeds of nine genotypes of the *Medicago truncatula* species complex. *Plant Physiology and Biochemistry* **43**: 557–566.

Downie B, Gurusinghe S, Dahal P, Thacker RR, Snyder JC, Nonogaki H, Yim K, Fukanaga K, Alvarado V, Bradford KJ. **2003**. Expression of a GALACTINOL SYNTHASE gene in tomato seeds is up-regulated before maturation desiccation and again after imbibition whenever radicle protrusion is prevented. *Plant Physiology* **131**: 1347–1359.

Doyle JJ. 1995. DNA data in legume phylogeny: a progress report. In: Crisp MD, Doyle JJ, eds. *Advances in legume systematics part 7. Phylogeny*. London UK: The Royal Botanic Gardens Kew, 11–30.

Duke JA. **1989**. *CRC handbook of nuts*. Boca Raton, Florida, CRC Press Incorporation.

Dure L, Greenway SC, Galau GA. **1981**. Developmental biochemistry of cottonseed embryogenesis and germination: changing messenger ribonucleic acid populations as shown by in vitro and in vivo protein synthesis. *Biochemistry* **20**: 4162–4168.

Dure LI, Crouch M, Harada J, Ho T-HD, Mundy J, Quatrano R, Thomas T, Sung ZR. **1989**. Common amino acid sequence domains among the LEA proteins of higher plants. *Plant Molecular Biology* **12**: 475–486.

Dussert S, Chabrillange N, Engelmann F, Anthony F, Louarn J, Hamon S. **2000**. Relationship between seed desiccation sensitivity, seed water content at maturity and climatic characteristics of native environments of nine *Coffea* L. species. *Seed Science Research* **10**: 293–300.

Edgar R, Domrachev M, Lash AE. **2002**. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Research* **30**: 207–210.

Elfstrand M, Feddermann N, Ineichen K, Nagaraj VJ, Wiemken A, Boller T, Salzer P. **2005**. Ectopic expression of the mycorrhiza-specific chitinase gene *Mtchit3-3* in *Medicago truncatula* root-organ cultures stimulates spore germination of glomalean fungi. *New Phytologist* **167**: 557–570.

Ellis RH, Hong TD, Roberts EH. **1990**. An Intermediate Category of Seed Storage Behaviour? I. Coffee. *Journal of Experimental Botany* **41**: 1167–1174.

Ellis RH, Hong TD, Roberts EH. **1991**. Effect of storage temperature and moisture on the germination of papaya seeds. *Seed Science Research* **1**: 69–72.

Ende WV den, Valluru R. **2009**. Sucrose, sucrosyl oligosaccharides, and oxidative stress: scavenging and salvaging? *Journal of Experimental Botany* **60**: 9–18.

Eom J, Baker WR, Kintanar A, Wurtele ES. **1996**. The embryo-specific EMB-1 protein of *Daucus carota* is flexible and unstructured in solution. *Plant Science* **115**: 17–24.

Fait A, Angelovici R, Less H, Ohad I, Urbanczyk-Wochniak E, Fernie AR, Galili G. **2006**. Arabidopsis Seed Development and Germination Is Associated with Temporally Distinct Metabolic Switches. *Plant Physiology* **142**: 839–854.

Faria JMR, Buitink J, Lammeren AAM van, Hilhorst HWM. **2005**. Changes in DNA and microtubules during loss and re-establishment of desiccation tolerance in germinating *Medicago truncatula* seeds. *Journal of Experimental Botany* **56**: 2119–2130.

De Faria SM, Lewis GP, Sprent JI, Sutherland JM. **1989**. Occurrence of nodulation in the Leguminosae. *New Phytologist*: 607–619.

Faria JMR, van Lammeren AAM, Hilhorst HWM. **2004**. Desiccation sensitivity and cell cycle aspects in seeds of *Inga vera* subsp. affinis. *Seed Science Research* **14**: 165–178.

Faria JMR, Buitink J, Lammeren AAM van, Hilhorst HWM. **2005**. Changes in DNA and microtubules during loss and re-establishment of desiccation tolerance in germinating *Medicago truncatula* seeds. *Journal of Experimental Botany* **56**: 2119–2130.

Farnsworth E. **2000**. The ecology and physiology of viviparous and recalcitrant seeds. *Annual Review of Ecology and Systematics* **31**: 107–138.

Farnsworth E, Farrant J. **1998**. Reductions in abscisic acid are linked with viviparous reproduction in mangroves. *Am. J. Bot.* **85**: 760.

Farrant JM, Berjak P, Pammenter NW. **1993a**. Studies on the development of the desiccation-sensitive (recalcitrant) seeds of *Avicennia marina* (Forssk.) Vierh.: the acquisition of germinability and response to storage and dehydration. *Annals of Botany* **71**: 405 –410.

Farrant JM, Pammenter NW, Berjak P. **1993b**. seed development in relation to desiccation tolerance: a comparison between desiccation-sensitive (recalcitrant) seeds of *Avicennia marina* and desiccation-tolerant types. *Seed Science Research* **3**: 1–13.

Farrant JM, Berjak P, Cutting JGM, Pammenter NW. **1993c**. The role of plant growth regulators in the development and germination of the desiccation-sensitive (recalcitrant) seeds of *Avicennia marina*. *Seed Science Research* **3**: 55–63.

Farrant JM, Pammenter NW, Berjak P, Farnsworth EJ, Vertucci CW. **1996**. Presence of dehydrin-like proteins and levels of abscisic acid in recalcitrant (desiccation sensitive) seeds may be related to habitat. *Seed Science Research* **6**: 175–182.

Farrant JM, Pammenter NW, Berjak P, Walters C. **1997**. Subcellular Organization and metabolic activity during the development of seeds that attain different levels of desiccation tolerance. *Seed Science Research* **7**: 135–144.

Farrant JM, Bailly C, Leymarie J, Hamman B, Côme D, Corbineau F. **2004**. Wheat seedlings as a model to understand desiccation tolerance and sensitivity. *Physiologia Plantarum* **120**: 563–574.

Farrant JM, Lehner A, Cooper K, Wiswedel S. **2009**. Desiccation tolerance in the vegetative tissues of the fern *Mohria caffrorum* is seasonally regulated. *The Plant Journal* **57**: 65–79.

Farrant JM, Moore JP. 2011. Programming desiccation-tolerance: from plants to seeds to resurrection plants. *Current Opinion in Plant Biology* **14**: 340–345.

Figueras M, Pujal J, Saleh A, Save R, Pages M, Goday A. **2004**. Maize *Rabl7* overexpression in Arabidopsis plants promotes osmotic stress tolerance. *Annals of Applied Biology* **144**: 251–257.

Finch-Savage WE. **1992**. Seed development in the recalcitrant species *Quercus robur* L.: germinability and desiccation tolerance. *Seed Science Research* **2**: 17–22.

Finch-Savage WE, Blake PS. **1994**. Indeterminate development in desiccation-sensitive seeds of *Quercus robur* L. *Seed Science Research* **4**: 127–133.

Finch-Savage WE, Pramanik SK, Bewley JD. **1994**. The expression of dehydrin proteins in desiccation-sensitive (recalcitrant) seeds of temperate trees. *Planta* **193**: 478–485.

Finkelstein RR. **1994**. Mutations at two new Arabidopsis ABA response loci are similar to the *abi3* mutations. *The Plant Journal* **5**: 765–771.

Finkelstein RR, Wang ML, Lynch TJ, Rao S, Goodman HM. **1998**. The Arabidopsis abscisic acid response locus abi4 encodes an APETALA2 domain protein. *The Plant Cell* **10**: 1043–1054.

Finkelstein RR, Lynch TJ. **2000**. The Arabidopsis abscisic acid response gene *ABI5* encodes a basic leucine zipper transcription factor. *The Plant Cell* **12**: 599–609.

Finkelstein RR, Gampala SS., Rock CD. **2002**. Abscisic acid signaling in seeds and seedlings. *The Plant Cell* **14**: suppl 1 S15-S45.

Finkelstein S, Grubb PJ. **2002**. Lipid concentration in the embryo-endosperm fraction of seeds of Australian tropical lowland rainforest trees: relevance to defence and dispersal. *Seed Science Research* **12**: 173–180.

Finkelstein R, Gampala SSL, Lynch TJ, Thomas TL, Rock CD. **2005**. Redundant and Distinct Functions of the ABA Response Loci *ABA-INSENSITIVE(ABI)5* and *ABRE-BINDING FACTOR (ABF)3*. *Plant Molecular Biology* **59**: 253–267.

Finkelstein R, Lynch T, Reeves W, Petitfils M, Mostachetti M. **2011**. Accumulation of the transcription factor ABA-insensitive (ABI)4 is tightly regulated post-transcriptionally. *Journal of Experimental Botany* **62**: 3971–3979.

Foyer CH, Noctor G. **2005**. Oxidant and antioxidant signalling in plants: a re-evaluation of the concept of oxidative stress in a physiological context. *Plant, Cell & Environment* **28**: 1056–1071.

Frey A, Godin B, Bonnet M, Sotta B, Marion-Poll A. **2004**. Maternal synthesis of abscisic acid controls seed development and yield in *Nicotiana plumbaginifolia*. *Planta* **218**: 958–964.

Frey A, Boutin J-P, Sotta B, Mercier R, Marion-Poll A. **2006**. Regulation of carotenoid and ABA accumulation during the development and germination of *Nicotiana plumbaginifolia* seeds. *Planta* **224**: 622–632.

Fu N, Wang Q, Shen H-L. **2013**. De novo assembly, gene annotation and marker development using illumina paired-end transcriptome sequences in celery (*Apium graveolens* L.). *PLoS ONE* **8**: e57686.

Fujii H, Verslues PE, Zhu J-K. **2007**. Identification of two protein kinases required for abscisic acid regulation of seed germination, root growth, and gene expression in Arabidopsis. *The Plant Cell* **19**: 485–494.

Fujita Y, Fujita M, Satoh R, Maruyama K, Parvez MM, Seki M, Hiratsu K, Ohme-Takagi M, Shinozaki K, Yamaguchi-Shinozaki K. **2005**. AREB1 is a transcription activator of novel ABRE-Dependent ABA signaling that enhances drought stress tolerance in Arabidopsis. *The Plant Cell* **17**: 3470–3488.

Fujita Y, Fujita M, Shinozaki K, Yamaguchi-Shinozaki K. **2011**. ABA-mediated transcriptional regulation in response to osmotic stress in plants. *Journal of Plant Research* **124:** 509-525.

Fujita Y, Yoshida T, Yamaguchi-Shinozaki K. **2013**. Pivotal role of the AREB/ABF-SnRK2 pathway in ABRE-mediated transcription in response to osmotic stress in plants. *Physiologia Plantarum* **147**: 15–27.

Gaff DF, Oliver M. **2013**. The evolution of desiccation tolerance in angiosperm plants: a rare yet common phenomenon. *Functional Plant Biology* **40**: 315–328.

Gagete AP, Riera M, Franco L, Rodrigo MI. **2009**. Functional analysis of the isoforms of an ABI3-like factor of Pisum sativum generated by alternative splicing. *Journal of experimental botany* **60**: 1703–1714.

Galau GA, Dure L. **1981**. Developmental biochemistry of cottonseed embryogenesis and germination: changing messenger ribonucleic acid populations as shown by reciprocal heterologous complementary deoxyribonucleic acid-messenger ribonucleic acid hybridization. *Biochemistry* **20**: 4169–4178.

Galau GA, Hughes DW, Iii LD. **1986**. Abscisic acid induction of cloned cotton late embryogenesis-abundant (Lea) mRNAs. *Plant Molecular Biology* **7**: 155–170.

Galau GA, Wang HY-C, Hughes DW. **1992**. Cotton *Lea4* (D19) and *LeaA2* (D132) Group 1 Lea Genes Encoding Water Stress-Related Proteins Containing a 20-Amino Acid Motif. *Plant Physiology* **99**: 783–788.

Gallardo K, Le Signor C, Vandekerckhove J, Thompson RD, Burstin J. **2003**. Proteomics of *Medicago truncatula* Seed Development Establishes the Time Frame of Diverse Metabolic Processes Related to Reserve Accumulation. *Plant Physiology* **133**: 664 –682.

Gallardo K, Lesignor C, Darmency M, Burstin J, Thompson R, Rochat C, Boutin JP, Küster H, Buitink J, Leprince O, *et al.* 2006. Seed biology of *Medicago truncatula*. *Medicago truncatula handbook*: 1–23.

Gao MJ, Lydiate DJ, Li X, Lui H, Gjetvaj B, Hegedus DD, Rozwadowski K. **2009**. Repression of seed maturation genes by a trihelix transcriptional repressor in Arabidopsis seedlings. *The Plant Cell* **21**: 54-71.

Gao Y, Liu J, Zhang Z, Sun X, Zhang N, Fan J, Niu X, Xiao F, Liu Y. Functional characterization of two alternatively spliced transcripts of tomato *ABSCISIC ACID INSENSITIVE3 (ABI3)* gene. *Plant Molecular Biology* **82**: 131-145.

De Gara L, de Pinto MC, Arrigoni O. **1997**. Ascorbate synthesis and ascorbate peroxidase activity during the early stage of wheat germination. *Physiologia Plantarum* **100**: 894–900.

De Gara L, Pinto MC de, Moliterni VMC, D'Egidio MG. **2003**. Redox regulation and storage processes during maturation in kernels of Triticum durum. *Journal of Experimental Botany* **54**: 249–258.

Garay-Arroyo A, Colmenero-Flores JM, Garciarrubio A, Covarrubias AA. **2000**. Highly Hydrophilic Proteins in Prokaryotes and Eukaryotes Are Common during Conditions of Water Deficit. *Journal of Biological Chemistry* **275**: 5668–5674.

Garello G, Le Page-Degivry M-T. **1995**. Desiccation-sensitive *Hopea odorata* seeds: Sensitivity to abscisic acid, water potential and inhibitors of gibberellin biosynthesis. *Physiologia Plantarum* **95**: 45–50.

Garg R, Patel RK, Jhanwar S, Priya P, Bhattacharjee A, Yadav G, Bhatia S, Chattopadhyay D, Tyagi AK, Jain M. **2011**. Gene discovery and tissue-specific transcriptome analysis in chickpea with massively parallel pyrosequencing and web resource development. *Plant Physiology* **156**: 1661–1678.

Gasulla F, vom Dorp K, Dombrink I, Zähringer U, Gisch N, Dörmann P, Bartels D. **2013**. The role of lipid metabolism in the acquisition of desiccation-tolerance in *Craterostigma plantagineum*: a comparative approach. *The Plant Journal* **75**: 726–741.

Gazzarrini S, Tsuchiya Y, Lumba S, Okamoto M, McCourt P. **2004**. The transcription factor FUSCA3 controls developmental timing in arabidopsis through the hormones gibberellin and abscisic acid. *Developmental Cell* **7**: 373–385.

Gilles GJ, Hines KM, Manfre AJ, Marcotte Jr. WR. **2007**. A predicted N-terminal helical domain of a Group 1 LEA protein is required for protection of enzyme activity from drying. *Plant Physiology and Biochemistry* **45**: 389–399.

Giraudat J, Hauge BM, Valon C, Smalle J, Parcy F, Goodman HM. **1992**. Isolation of the Arabidopsis *ABI3* gene by positional cloning. *The Plant Cell* **4**: 1251–1261.

Glisovic T, Bachorik JL, Yong J, Dreyfuss G. **2008**. RNA-binding proteins and post-transcriptional gene regulation. *FEBS Letters* **582**: 1977–1986.

Golovina EA, Golovin AV, Hoekstra FA, Faller R. **2009**. Water replacement hypothesis in atomic detail—factors determining the structure of dehydrated bilayer stacks. *Biophysical Journal* **97**: 490–499.

Górecki RJ, Piotrowicz-Cieślak AI, Lahuta LB, Obendorf RL. **1997**. Soluble carbohydrates in desiccation tolerance of yellow lupin seeds during maturation and germination. *Seed Science Research* **7**: 107–116.

Götz S, García-Gómez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ, Robles M, Talón M, Dopazo J, Conesa A. **2008**. High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic acids research* **36**: 3420-3435.

Goyal K, Tisi L, Basran A, Browne J, Burnell A, Zurdo J, Tunnacliffe A. **2003**. Transition from natively unfolded to folded state induced by desiccation in an anhydrobiotic nematode protein. *Journal of Biological Chemistry* **278**: 12977–12984.

Goyal K, Walton LJ, Tunnacliffe A. **2005**. LEA proteins prevent protein aggregation due to water stress. *Biochemical Journal* **388**: 151–157.

Graham PH, Vance CP. **2003**. Legumes: importance and constraints to greater use. *Plant Physiology* **131**: 872–877.

Greggains V, Finch-Savage WE, Quick WP, Atherton NM. **2000**. Putative desiccation tolerance mechanisms in orthodox and recalcitrant seeds of the genus Acer. *Seed Science Research* **10**: 317–327.

Greggains V, Finch-Savage WE, Atherton NM, Berjak P. **2001**. Viability loss and free radical processes during desiccation of recalcitrant *Avicennia marina* seeds. *Seed Science Research* **11**: 235–242.

Grelet J, Benamar A, Teyssier E, Avelange-Macherel M-H, Grunwald D, Macherel D. **2005**. Identification in pea seed mitochondria of a late-embryogenesis abundant protein able to protect enzymes from drying. *Plant Physiology* **137**: 157–167.

Groot SPC, Bruinsma J, Karssen CM. **1987**. The role of endogenous gibberellin in seed and fruit development of tomato: Studies with a gibberellin-deficient mutant. *Physiologia Plantarum* **71**: 184–190.

Guan L, Scandalios JG. **1998**. Effects of the plant growth regulator abscisic acid and high osmoticum on the developmental expression of the maize catalase genes. *Physiologia Plantarum* **104**: 413–422.

Guan LM, Zhao J, Scandalios JG. **2000**. Cis-elements and trans-factors that regulate expression of the maize Cat1 antioxidant gene in response to ABA and osmotic stress: H_2O_2 is the likely intermediary signaling molecule for the response. *The Plant Journal* **22**: 87–95.

Gubler F, Jacobsen JV. **1992**. Gibberellin-responsive elements in the promoter of a barley high-pI alpha-amylase gene. *The Plant Cell* **4**: 1435–1441.

Guerriero G, Martin N, Golovko A, Sundström JF, Rask L, Ezcurra I. **2009**. The RY/Sph element mediates transcriptional repression of maturation genes from late maturation to early seedling growth. *New Phytologist* **184**: 552–565.

Guo N, Puhlev I, Brown DR, Mansbridge J, Levine F. **2000**. Trehalose expression confers desiccation tolerance on human cells. *Nature Biotechnology* **18**: 168–171.

Gurusinghe S, Bradford KJ. **2001**. Galactosyl-sucrose oligosaccharides and potential longevity of primed seeds. *Seed Science Research* **11**: 121–134.

Gutierrez L, Van Wuytswinkel O, Castelain M, Bellini C. **2007**. Combined networks regulating seed maturation. *Trends in plant science* **12**: 294–300.

Halperin SJ, Koster KL. **2006**. Sugar effects on membrane damage during desiccation of pea embryo protoplasts. *Journal of Experimental Botany* **57**: 2303–2311.

Hamilton AJ, Baulcombe DC. **1999**. A species of small antisense rna in post-transcriptional gene silencing in plants. *Science* **286**: 950–952.

Han B, Berjak P, Pammenter N, Farrant J, Kermode AR. **1997**. The recalcitrant plant species, *Castanospermum australe* and *Trichilia dregeana*, differ in their ability to produce dehydrin-related polypeptides during seed maturation and in response to ABA or water-deficit-related stresses. *Journal of Experimental Botany* **48**: 1717–1726.

Hara M, Terashima S, Kuboi T. **2001**. Characterization and cryoprotective activity of cold-responsive dehydrin from Citrus unshiu. *Journal of Plant Physiology* **158**: 1333–1339.

Hara M, Fujinaga M, Kuboi T. **2004**. Radical scavenging activity and oxidative modification of citrus dehydrin. *Plant Physiology and Biochemistry* **42**: 657–662.

Hara M, Fujinaga M, Kuboi T. **2005**. Metal binding by citrus dehydrin with histidine-rich domains. *Journal of Experimental Botany* **56**: 2695–2703.

Harada JJ. **2001**. Role of Arabidopsis LEAFY COTYLEDON genes in seed development. *Journal of plant physiology* **158**: 405–409.

Harris D. **1987**. Aboriginal subsistence in a tropical rain forest environment: Food procurement, cannibalism, and population regulation in northeastern Australia. In: Harris M. and Ross EB. eds *Food and evolution: toward of theory of human food habits.* Temple University Press, Philadelphia, 357–385.

Haslekås C, Stacy RAP, Nygaard V, Culiáñez-Macià FA, Aalen RB. **1998**. The expression of a peroxiredoxin antioxidant gene, *AtPer1*, in Arabidopsis thaliana is seed-specific and related to dormancy. *Plant Molecular Biology* **36**: 833–845.

Haslekås C, Grini PE, Nordgard SH, Thorstensen T, Viken MK, Nygaard V, Aalen RB. **2003**. ABI3 mediates expression of the peroxiredoxin antioxidant *AtPER1* gene and induction by oxidative stress. *Plant Molecular Biology* **53**: 313–326.

Hattori T, Totsuka M, Hobo T, Kagaya Y, Yamamoto-Toyoda A. **2002**. Experimentally determined sequence requirement of ACGT-containing abscisic acid response element. *Plant and Cell Physiology* **43**: 136–140.

Helm KW, Schmeits J, Vierling E. **1995**. An endomembrane-localized small heat-shock protein from *Arabidopsis thaliana*. *Plant Physiology* **107**: 287–288.

Henderson JT, Li H-C, Rider SD, Mordhorst AP, Romero-Severson J, Cheng J-C, Robey J, Sung ZR, Vries SC de, Ogas J. **2004**. PICKLE acts throughout the plant to repress expression of embryonic traits and may play a role in gibberellin-dependent responses. *Plant Physiology* **134**: 995–1005.

Hendry GAF, Finch-savage WE, Thorpe PC, Atherton NM, Buckland SM, Nilsson KA, Seel WE. **1992**. Free radical processes and loss of seed viability during desiccation in the recalcitrant species *Quercus robur* L. *New Phytologist* **122**: 273–279.

Hennig L, Derkacheva M. **2009**. Diversity of Polycomb group complexes in plants: same rules, different players? *Trends in Genetics* **25**: 414–423.

Heyen BJ, Alsheikh MK, Smith EA, Torvik CF, Seals DF, Randall SK. **2002**. The calcium-binding activity of a vacuole-associated, dehydrin-like protein is regulated by phosphorylation. *Plant Physiology* **130**: 675–687.

Hilhorst HWM, Karssen CM. **1992**. Seed dormancy and germination: the role of abscisic acid and gibberellins and the importance of hormone mutants. *Plant Growth Regulation* **11**: 225–238.

Hinniger C, Caillet V, Michoux F, Amor MB, Tanksley S, Lin C, Mccarthy J. **2006**. Isolation and characterization of cdna encoding three dehydrins expressed during *Coffea canephora* (robusta) grain development. *Annals of Botany* **97**: 755–765.

Hobo T, Asada M, Kowyama Y, Hattori T. **1999**. ACGT-containing abscisic acid response element (ABRE) and coupling element 3 (CE3) are functionally equivalent. *The Plant Journal* **19**: 679–689.

Hoekstra FA, Wolkers WF, Buitink J, Golovina EA, Crowe JH, Crowe LM. **1997**. Membrane stabilization in the dry state. *Comparative Biochemistry and Physiology* **117**: 335–341.

Hoekstra FA, Golovina EA, Buitink J. **2001**. Mechanisms of plant desiccation tolerance. *Trends in Plant Science* **6**: 431–438.

Hoffmann B, Trinh TH, Leung J, Kondorosi A, Kondorosi E. **1997**. A new *Medicago truncatula* line with superior in vitro regeneration, transformation, and symbiotic properties isolated through cell culture selection. *Molecular Plant-Microbe Interactions* **10**: 307–315.

Holdsworth MJ, Bentsink L, Soppe WJJ. **2008**. Molecular networks regulating Arabidopsis seed maturation, after-ripening, dormancy and germination. *New Phytologist* **179**: 33–54.

Hong TD, Linington S, Ellis RH. **1996**. Seed Storage Behaviour: a Compendium. *International Plant Genetic Resources Institute*, Rome, Italy, 1-115

Hong TD, Ellis RH. **1998**. Contrasting seed storage behaviour among different species of Meliaceae. *Seed science and technology* **26**: 77–95.

Hoppmann V, Thorstensen T, Kristiansen PE, Veiseth SV, Rahman MA, Finne K, Aalen RB, Aasland R. **2011**. The CW domain, a new histone recognition module in chromatin proteins. *The EMBO Journal* **30**: 1939–1952.

Horbowicz M, Obendorf RL. **1994**. Seed desiccation tolerance and storability: Dependence on flatulence-producing oligosaccharides and cyclitols—review and survey. *Seed Science Research* **4**: 385–405.

Houde M, Dallaire S, N'Dong D, Sarhan F. **2004**. Overexpression of the acidic dehydrin WCOR410 improves freezing tolerance in transgenic strawberry leaves. *Plant Biotechnology Journal* **2**: 381–387.

Huang AHC. **1992**. Oil bodies and oleosins in seeds. *Annual Review of Plant Physiology and Plant Molecular Biology* **43**: 177–200.

Huijser C, Kortstee A, Pego J, Weisbeek P, Wisman E, Smeekens S. **2000**. The Arabidopsis *SUCROSE UNCOUPLED-6* gene is identical to *ABSCISIC ACID INSENSITIVE-4*: involvement of abscisic acid in sugar responses. *The Plant Journal* **23**: 577–585.

Hundertmark M, Hincha DK. **2008**. LEA (Late Embryogenesis Abundant) proteins and their encoding genes in *Arabidopsis thaliana*. *BMC genomics* **9**: 118.

Hundertmark M, Buitink J, Leprince O, Hincha DK. **2011**. The reduction of seed-specific dehydrins reduces seed longevity in *Arabidopsis thaliana*. *Seed Science Research* **21**: 165–173.

Huo H, Dahal P, Kunusoth K, McCallum CM, Bradford KJ. **2013**. Expression of 9-cis-EPOXYCAROTENOID DIOXYGENASE4 is essential for thermoinhibition of lettuce seed germination but not for seed development or stress tolerance. *The Plant Cell* **25**: 884–900.

Illing N, Denby KJ, Collett H, Shen A, Farrant JM. **2005**. the signature of seeds in resurrection plants: a molecular and physiological comparison of desiccation tolerance in seeds and vegetative tissues. *Integrative and Comparative Biology* **45**: 771–787.

Ingram J, Bartels D. **1996**. The molecular basis of dehydration tolerance in plants. *Annual Review of Plant Physiology and Plant Molecular Biology* **47**: 377–403.

Ismail AM, Hall AE, Close TJ. **1999**. Purification and partial characterization of a dehydrin involved in chilling tolerance during seedling emergence of cowpea. *Plant Physiology* **120**: 237–244.

Ismail FA, Nitsch LMC, Wolters-Arts MMC, Mariani C, Derksen JWM. **2010**. Semi-viviparous embryo development and dehydrin expression in the mangrove *Rhizophora mucronata* Lam. *Sexual Plant Reproduction* **23**: 95–103.

Iturriaga G, Schneider K, Salamini F, Bartels D. **1992**. Expression of desiccation-related proteins from the resurrection plant *Craterostigma plantagineum* in transgenic tobacco. *Plant Molecular Biology* **20**: 555–558.

Iuchi S, Kobayashi M, Taji T, Naramoto M, Seki M, Kato T, Tabata S, Kakubari Y, Yamaguchi-Shinozaki K, Shinozaki K. **2001**. Regulation of drought tolerance by gene manipulation of 9-cis-epoxycarotenoid dioxygenase, a key enzyme in abscisic acid biosynthesis in Arabidopsis. *The Plant Journal* **27**: 325–333.

Jacobsen JV, Beach LR. **1985**. Control of transcription of α-amylase and rRNA genes in barley aleurone protoplasts by gibberellin and abscisic acid. *Nature* **316**: 275–277.

Jaspard E, Macherel D, Hunault G. **2012**. Computational and statistical analyses of amino acid usage and physico-chemical properties of the twelve late embryogenesis abundant protein classes. *PLoS ONE* **7**: e36968.

Jia H, Suzuki M, McCarty DR. **2013**. Regulation of the seed to seedling developmental phase transition by the LAFL and VAL transcription factor networks. *Wiley Interdisciplinary Reviews: Developmental Biology*: DOI: 10.1002/wdev.126

Joët T, Ourcival J-M, Dussert S. **2013**. Ecological significance of seed desiccation sensitivity in *Quercus ilex*. *Annals of Botany* **111**: 693-701.

Johannesson H, Wang Y, Hanson J, Engström P. **2003**. The *Arabidopsis thaliana* homeobox gene *ATHB5* is a potential regulator of abscisic acid responsiveness in developing seedlings. *Plant Molecular Biology* **51**: 719–729.

José AC, Ligterink W, Davide AC, Silva E da, Hilhorst HW. **2009**. Changes in gene expression during drying and imbibition of desiccation sensitive *Magnolia ovate* (A. St.-Hil.) Spreng. Seeds. *Revista Brasileira de Sementes* **31**: 270–280.

Jung C, Seo JS, Han SW, Koo YJ, Kim CH, Song SI, Nahm BH, Choi YD, Cheong J-J. **2008**. Overexpression of *AtMYB44* enhances stomatal closure to confer abiotic stress tolerance in transgenic Arabidopsis. *Plant Physiology* **146**: 623–635.

Junker A, Bäumlein H. **2012**. Multifunctionality of the LEC1 transcription factor during plant development. *Plant Signaling & Behavior* **7**: 1718–1720.

Junker A, Mönke G, Rutten T, Keilwagen J, Seifert M, Thi TMN, Renou J-P, Balzergue S, Viehöver P, Hähnel U, *et al.* **2012**. Elongation-related functions of LEAFY COTYLEDON1 during the development of *Arabidopsis thaliana*. *The Plant Journal* **71**: 427–442.

Kagaya Y, Toyoshima R, Okuda R, Usui H, Yamamoto A, Hattori T. **2005**. LEAFY COTYLEDON1 Controls Seed Storage Protein Genes through Its Regulation of FUSCA3 and ABSCISIC ACID INSENSITIVE3. *Plant and Cell Physiology* **46**: 399–406.

Kang J, Hwang J-U, Lee M, Kim Y-Y, Assmann SM, Martinoia E, Lee Y. **2010**. PDR-type ABC transporter mediates cellular uptake of the phytohormone abscisic acid. *Proceedings of the National Academy of Sciences* **107**: 2355–2360.

Kanno Y, Jikumaru Y, Hanada A, Nambara E, Abrams SR, Kamiya Y, Seo M. **2010**. Comprehensive hormone profiling in developing arabidopsis seeds: examination of the site of aba biosynthesis, aba transport and hormone interactions. *Plant and Cell Physiology* **51**: 1988–2001.

Kanno Y, Hanada A, Chiba Y, Ichikawa T, Nakazawa M, Matsui M, Koshiba T, Kamiya Y, Seo M. **2012**. Identification of an abscisic acid transporter by functional screening using the receptor complex as a sensor. *Proceedings of the National Academy of Sciences* **109**: 9653–9658.

Karimi M, Inzé D, Depicker A. **2002**. GATEWAY™ vectors for Agrobacterium-mediated plant transformation. *Trends in Plant Science* **7**: 193–195.

Karssen CM, Swan DLCB der, Breekland AE, Koornneef M. **1983**. Induction of dormancy during seed development by endogenous abscisic acid: studies on abscisic acid deficient genotypes of *Arabidopsis thaliana* (L.) Heynh. *Planta* **157**: 158–165.

Käss E, Wink M. **1996**. Molecular evolution of the leguminosae: Phylogeny of the three subfamilies based on rbcL-sequences. *Biochemical Systematics and Ecology* **24**: 365–378.

Kaushik JK, Bhat R. **2003**. Why Is Trehalose an Exceptional Protein Stabilizer? an analysis of the thermal stability of proteins in the presence of the compatible osmolyte trehalose. *Journal of Biological Chemistry* **278**: 26458–26465.

Keith K, Kraml M, Dengler NG, McCourt P. **1994**. *fusca3*: a heterochronic mutation affecting late embryo development in Arabidopsis. *The Plant Cell* **6**: 589-600.

Kermode AR. **1997**. Approaches to elucidate the basis of desiccation-tolerance in seeds. *Seed Science Research* **7**: 75–96.

Khandelwal A, Cho SH, Marella H, Sakata Y, Perroud P-F, Pan A, Quatrano RS. **2010**. Role of ABA and ABI3 in desiccation tolerance. *Science* **327**: 546–546.

Kibinza S, Bazin J, Bailly C, Farrant JM, Corbineau F, El-Maarouf-Bouteau H. **2011**. Catalase is a key enzyme in seed recovery from ageing during priming. *Plant Science* **181**: 309–315.

Kim DH, Yamaguchi S, Lim S, Oh E, Park J, Hanada A, Kamiya Y, Choi G. **2008**. SOMNUS, a CCCH-type zinc finger protein in arabidopsis, negatively regulates light-dependent seed germination downstream of PIL5. *The Plant Cell* **20**: 1260–1277.

Knight H, Zarka DG, Okamoto H, Thomashow MF, Knight MR. **2004**. Abscisic acid induces CBF gene transcription and subsequent induction of cold-regulated genes via the CRT promoter element. *Plant Physiology* **135**: 1710–1717.

Koag M-C, Fenton RD, Wilkens S, Close TJ. **2003**. The binding of maize DHN1 to lipid vesicles. Gain of structure and lipid specificity. *Plant Physiology* **131**: 309–316.

Koag M-C, Wilkens S, Fenton RD, Resnik J, Vo E, Close TJ. **2009**. The K-Segment of maize DHN1 mediates binding to anionic phospholipid vesicles and concomitant structural changes. *Plant Physiology* **150**: 1503–1514.

Kobayashi Y, Yamamoto S, Minami H, Kagaya Y, Hattori T. **2004**. Differential activation of the rice sucrose nonfermenting1–related protein kinase2 family by hyperosmotic stress and abscisic acid. *The Plant Cell* **16**: 1163–1177.

Koch JL, Horbowicz M, Obendorf RL. **1999**. Methanol, pectin and pectinesterase changes during soybean seed maturation. *Seed Science Research* **9**: 311–320.

Kong Y, Chen S, Yang Y, An C. ABA-insensitive (ABI) 4 and ABI5 synergistically regulate DGAT1 expression in Arabidopsis seedlings under stress. *FEBS Letters* **587**: 3076–3082.

Koster KL, Leopold AC. **1988**. Sugars and desiccation tolerance in seeds. *Plant Physiology* **88**: 829-832.

Koster KL. **1991**. Glass formation and desiccation tolerance in seeds. *Plant Physiology* **96**: 302–304.

Koster KL, Balsamo RA, Espinoza C, Oliver MJ. **2010**. Desiccation sensitivity and tolerance in the moss *Physcomitrella patens*: assessing limits and damage. *Plant Growth Regulation* **62**: 293–302.

Kotak S, Vierling E, Baumlein H, Koskull-Doring P. **2007**. A novel transcriptional cascade regulating expression of heat stress proteins during seed development of Arabidopsis. *The Plant Cell* **19**: 182-195.

Kranner I, Birtić S. **2005**. A modulating role for antioxidants in desiccation tolerance. *Integrative and Comparative Biology* **45**: 734–740.

Kroj T, Savino G, Valon C, Giraudat J, Parcy F. **2003**. Regulation of storage protein gene expression in Arabidopsis. *Development* **130**: 6065-6073.

Kulik A, Wawer I, Krzywińska E, Bucholc M, Dobrowolska G. **2011**. SnRK2 protein kinases—key regulators of plant response to abiotic stresses. *OMICS: A Journal of Integrative Biology* **15**: 859–872.

Kumar S, Blaxter ML. **2010**. Comparing de novo assemblers for 454 transcriptome data. *BMC Genomics* **11**: 571.

Kuromori T, Miyaji T, Yabuuchi H, Shimizu H, Sugimoto E, Kamiya A, Moriyama Y, Shinozaki K. **2010**. ABC transporter AtABCG25 is involved in abscisic acid transport and responses. *Proceedings of the National Academy of Sciences* **107**: 2361–2366.

Kwong RW, Bui AQ, Lee H, Kwong LW, Fischer RL, Goldberg RB, Harada JJ. **2003**. LEAFY COTYLEDON1-LIKE defines a class of regulators essential for embryo development. *The Plant Cell* **15**: 5–18.

Lahuta LB, Górecki RJ. **2011**. Raffinose in seedlings of winter vetch (*Vicia villosa* Roth.) under osmotic stress and followed by recovery. *Acta Physiologiae Plantarum* **33**: 725–733.

Lanahan MB, Ho TH, Rogers SW, Rogers JC. **1992**. A gibberellin response complex in cereal alpha-amylase gene promoters. *The Plant Cell* **4**: 203–211.

Lapinski J, Tunnacliffe A. **2003**. Anhydrobiosis without trehalose in bdelloid rotifers. *FEBS Letters* **553**: 387–390.

Lara P, Oñate-Sánchez L, Abraham Z, Ferrándiz C, Díaz I, Carbonero P, Vicente-Carbajosa J. **2003**. Synergistic activation of seed storage protein gene expression in arabidopsis by ABI3 and two bZIPs Related to OPAQUE2. *Journal of Biological Chemistry* **278**: 21003–21011.

Larkindale J, Knight MR. **2002**. Protection against heat stress-induced oxidative damage in Arabidopsis involves calcium, abscisic acid, ethylene, and salicylic acid. *Plant Physiology* **128**: 682–695.

Larkindale J, Huang B. **2004**. Thermotolerance and antioxidant systems in *Agrostis stolonifera*: Involvement of salicylic acid, abscisic acid, calcium, hydrogen peroxide, and ethylene. *Journal of Plant Physiology* **161**: 405–413.

Lazarova G, Zeng Y, Kermode AR. **2002**. Cloning and expression of an *ABSCISIC ACID-INSENSITIVE 3* (*ABI3*) gene homologue of yellow-cedar (*Chamaecyparis nootkatensis*). *Journal of Experimental Botany* **53**: 1219–1221.

Le BH, Cheng C, Bui AQ, Wagmaister JA, Henry KF, Pelletier J, Kwong L, Belmonte M, Kirkbride R, Horvath S, *et al.* 2010. Global analysis of gene activity during Arabidopsis seed development and identification of seed-specific transcription factors. *Proceedings of the National Academy of Sciences* **107**: 8063-8070.

Ledwoń A, Gaj MD. **2011**. *LEAFY COTYLEDON1*, *FUSCA3* expression and auxin treatment in relation to somatic embryogenesis induction in Arabidopsis. *Plant Growth Regulation* **65**: 157–167.

Lee GJ, Pokala N, Vierling E. **1995**. Structure and in vitro molecular chaperone activity of cytosolic small heat shock proteins from pea. *Journal of Biological Chemistry* **270**: 10432–10438.

Lee S, Cheng H, King KE, Wang W, He Y, Hussain A, Lo J, Harberd NP, Peng J. **2002**. Gibberellin regulates Arabidopsis seed germination via *RGL2*, a *GAI/RGA-like* gene whose expression is up-regulated following imbibition. *Genes & Development* **16**: 646–658.

Lee S, Kang J, Park H-J, Kim MD, Bae MS, Choi H, Kim SY. **2010**. DREB2C interacts with ABF2, a bZIP protein regulating abscisic acid-responsive gene expression, and its overexpression affects abscisic acid sensitivity. *Plant Physiology* **153**: 716–727.

Lee AK, Slovin JP, Suh JK. **2012**. Dehydration intolerant seeds of *Ardisia* species accumulate storage and stress proteins during development. *Horticulture, Environment, and Biotechnology* **53**: 530–538.

Lefebvre V, North H, Frey A, Sotta B, Seo M, Okamoto M, Nambara E, Marion-Poll A. **2006**. Functional analysis of Arabidopsis *NCED6* and *NCED9* genes indicates that ABA synthesized in the endosperm is involved in the induction of seed dormancy. *The Plant Journal* **45**: 309–319.

Lenne C, Block MA, Garin J, Douce R. **1995**. Sequence and expression of the mRNA encoding HSP22, the mitochondrial small heat-shock protein in pea leaves.

Leprince O, Deltour R, Thorpe PC, Atherton NM, Hendry G a. F. **1990**. The role of free radicals and radical processing systems in loss of desiccation tolerance in germinating maize (*Zea mays* L.). *New Phytologist* **116**: 573–580.

Leprince O, van Aelst AC, Pritchard HW, Murphy DJ. **1997**. Oleosins prevent oil-body coalescence during seed imbibition as suggested by a low-temperature scanning electron microscope study of desiccation-tolerant and -sensitive oilseeds. *Planta* **204**: 109–119.

Leprince O, Hoekstra FA. **1998**. The responses of cytochrome redox state and energy metabolism to dehydration support a role for cytoplasmic viscosity in desiccation tolerance. *Plant Physiology* **118**: 1253–1264.

Leprince O, Harren FJM, Buitink J, Alberda M, Hoekstra FA. **2000**. Metabolic dysfunction and unabated respiration precede the loss of membrane integrity during dehydration of germinating radicles. *Plant Physiology* **122**: 597–608.

Leprince O, Buitink J. **2010**. Desiccation tolerance: From genomics to the field. *Plant Science* **179**: 554–564.

Leslie SB, Israeli E, Lighthart B, Crowe JH, Crowe LM. **1995**. Trehalose and sucrose protect both membranes and proteins in intact bacteria during drying. *Applied and Environmental Microbiology* **61**: 3592–3597.

Li X-Y, Mantovani R, Huijsduijnen RH van, Andre I, Benoist C, Mathis D. **1992**. Evolutionary variation of the CCAAT-binding transcription factor NF-Y. *Nucleic Acids Research* **20**: 1087–1091.

Li C, Sun WQ. **1999**. Desiccation sensitivity and activities of free radical-scavenging enzymes in recalcitrant Theobroma cacao seeds. *Seed Science Research* **9**: 209–217.

Li D-Z, Pritchard HW. **2009**. The science and economics of ex situ plant conservation. *Trends in Plant Science* **14**: 614–621.

Li Y, Jin K, Zhu Z, Yang J. **2010**. Stepwise origin and functional diversification of the AFL subfamily B3 genes during land plant evolution. *Journal of Bioinformatics and Computational Biology* **8**: 33–45.

Li X, Zhuo J, Jing Y, Liu X, Wang X. **2011**. Expression of a GALACTINOL SYNTHASE gene is positively associated with desiccation tolerance of *Brassica napus* seeds during development. *Journal of Plant Physiology* **168**: 1761–1770.

Liang C, Liu X, Yiu S-M, Lim BL. **2013**. De novo assembly and characterization of *Camelina sativa* transcriptome by paired-end sequencing. *BMC Genomics* **14**: 146.

Liu J-JJ, Krenz DC, Galvez AF, de Lumen BO. **1998**. Galactinol synthase (GS): increased enzyme activity and levels of mRNA due to cold and desiccation. *Plant Science* **134**: 11–20.

Liu G, Xu H, Zhang L, Zheng Y. **2011**. Fe binding properties of two soybean (*Glycine max* L.) LEA4 proteins associated with antioxidant activity. *Plant and Cell Physiology* **52**: 994–1002.

Lopez-Molina L, Mongrand S, Chua NH. **2001**. A postgermination developmental arrest checkpoint is mediated by abscisic acid and requires the ABI5 transcription factor in Arabidopsis. *Proceedings of the National Academy of Sciences* **98**: 4782-4787.

Lopez-Molina L, Mongrand S, McLachlin DT, Chait BT, Chua N-H. **2002**. ABI5 acts downstream of ABI3 to execute an ABA-dependent growth arrest during germination. *The Plant Journal* **32**: 317–328.

Lotan T, Ohto M, Yee KM, West MA., Lo R, Kwong RW, Yamagishi K, Fischer RL, Goldberg RB, Harada JJ. **1998**. Arabidopsis LEAFY COTYLEDON1 is sufficient to induce embryo development in vegetative cells. *Cell* **93**: 1195–1205.

Löw D, Brändle K, Nover L, Forreiter C. **2000**. Cytosolic heat-stress proteins Hsp17.7 class I and Hsp17.3 class II of tomato act as molecular chaperones in vivo. *Planta* **211**: 575–582.

Lu QS, Dela Paz J, Pathmanathan A, Chiu RS, Tsai AY-L, Gazzarrini S. **2010**. The C-terminal domain of FUSCA3 negatively regulates mRNA and protein levels, and mediates sensitivity to the hormones abscisic acid and gibberellic acid in Arabidopsis. *The Plant Journal* **64**: 100–113.

Luerssen H, Kirik V, Herrmann P, Miséra S. **1998**. FUSCA3 encodes a protein with a conserved VP1/ABI3-like B3 domain which is of functional importance for the regulation of seed maturation in *Arabidopsis thaliana*. *The Plant Journal* **15**: 755–764.

Ma Y, Szostkiewicz I, Korte A, Moes D, Yang Y, Christmann A, Grill E. **2009**. Regulators of PP2C phosphatase activity function as abscisic acid sensors. *Science* **324**: 1064–1068.

Maia J, Dekkers BJW, Provart NJ, Ligterink W, Hilhorst HWM. **2011**. The re-establishment of desiccation tolerance in germinated *Arabidopsis thaliana* seeds and its associated transcriptome. *PLoS ONE* **6**: e29123.

Makarevich G, Leroy O, Akinci U, Schubert D, Clarenz O, Goodrich J, Grossniklaus U, Köhler C. **2006**. Different Polycomb group complexes regulate common target genes in Arabidopsis. *EMBO reports* **7**: 947–952.

Manfre AJ, Lanni LM, Marcotte WR. **2006**. The Arabidopsis group 1 Late Embryogenesis Abundant Protein ATEM6 is required for normal seed development. *Plant Physiology* **140**: 140–149.

Manfre AJ, LaHatte GA, Climer CR, Marcotte WR. **2009**. Seed dehydration and the establishment of desiccation tolerance during seed maturation is altered in the *Arabidopsis thaliana* mutant A*tem6-1*. *Plant and Cell Physiology* **50**: 243–253.

Marella HH, Sakata Y, Quatrano RS. **2006**. Characterization and functional analysis of *Abscisic Acid Insensitive3-like* genes from *Physcomitrella patens*. *The Plant Journal* **46**: 1032–1044.

Marella HH, Quatrano RS. **2007**. The B2 domain of VIVIPAROUS1 is bi-functional and regulates nuclear localization and transactivation. *Planta* **225**: 863–872.

Martin JA, Wang Z. **2011**. Next-generation transcriptome assembly. *Nature Reviews Genetics* **12**: 671–682.

Maruyama K, Todaka D, Mizoi J, Yoshida T, Kidokoro S, Matsukura S, Takasaki H, Sakurai T, Yamamoto YY, Yoshiwara K, *et al.* 2012. Identification of cis-acting promoter elements in cold- and dehydration-induced transcriptional pathways in arabidopsis, rice, and soybean. *DNA Research* **19**: 37–49.

McCarty DR. **1995**. Genetic control and integration of maturation and germination pathways in seed development. *Annual Review of Plant Physiology and Plant Molecular Biology* **46**: 71–93.

McCarty DR, Hattori T, Carson CB, Vasil V, Lazar M, Vasil IK. **1991**. The *Viviparous-1* developmental gene of maize encodes a novel transcriptional activator. *Cell* **66**: 895–905.

McKibbin RS, Wilkinson MD, Bailey PC, Flintham JE, Andrew LM, Lazzeri PA, Gale MD, Lenton JR, Holdsworth MJ. **2002**. Transcripts of *Vp-1* homeologues are misspliced in modern wheat and ancestral species. *Proceedings of the National Academy of Sciences* **99**: 10203–10208.

Meinke DW. **1992**. A homoeotic mutant of arabidopsis thaliana with leafy cotyledons. *Science* **258**: 1647–1650.

Meinke DW, Franzmann LH, Nickle TC, Yeung EC. **1994**. Leafy cotyledon mutants of Arabidopsis. *The Plant Cell* **6**: 1049–1064.

Mello JI de O, Barbedo CJ, Salatino A, Figueiredo-Ribeiro R de CL. **2010**. Reserve carbohydrates and lipids from the seeds of four tropical tree species with different sensitivity to desiccation. *Brazilian Archives of Biology and Technology* **53**: 889–899.

Mhamdi A, Noctor G, Baker A. **2012**. Plant catalases: Peroxisomal redox guardians. *Archives of Biochemistry and Biophysics* **525**: 181–194.

Milan M, Coppe A, Reinhardt R, Cancela L, Leite R, Saavedra C, Ciofi C, Chelazzi G, Patarnello T, Bortoluzzi S, *et al.* 2011. Transcriptome sequencing and microarray development for the Manila clam, *Ruditapes philippinarum*: genomic tools for environmental monitoring. *BMC Genomics* **12**: 234.

Miller G, Mittler R. **2006**. Could heat shock transcription factors function as hydrogen peroxide sensors in plants? *Annals of Botany* **98**: 279–288.

Miller G, Suzuki N, Ciftci-Yilmaz S, Mittler R. **2010**. Reactive oxygen species homeostasis and signalling during drought and salinity stresses. *Plant, Cell & Environment* **33**: 453–467.

Mitchum MG, Yamaguchi S, Hanada A, Kuwahara A, Yoshioka Y, Kato T, Tabata S, Kamiya Y, Sun T. **2006**. Distinct and overlapping roles of two gibberellin 3-oxidases in Arabidopsis development. *The Plant Journal* **45**: 804–818.

Mittler R. **2002**. Oxidative stress, antioxidants and stress tolerance. *Trends in Plant Science* **7**: 405–410.

Mittler R, Vanderauwera S, Gollery M, Van Breusegem F. **2004**. Reactive oxygen gene network of plants. *Trends in Plant Science* **9**: 490–498.

Mittler R, Vanderauwera S, Suzuki N, Miller G, Tognetti VB, Vandepoele K, Gollery M, Shulaev V, Van Breusegem F. **2011**. ROS signaling: the new wave? *Trends in Plant Science* **16**: 300–309.

Miura K, Lee J, Jin JB, Yoo CY, Miura T, Hasegawa PM. **2009**. Sumoylation of ABI5 by the Arabidopsis SUMO E3 ligase SIZ1 negatively regulates abscisic acid signaling. *Proceedings of the National Academy of Sciences* **106**: 5418–5423.

Mönke G, Altschmied L, Tewes A, Reidt W, Mock H-P, Bäumlein H, Conrad U. **2004**. Seed-specific transcription factors ABI3 and FUS3: molecular interaction with DNA. *Planta* **219**: 158–166.

Mönke G, Seifert M, Keilwagen J, Mohr M, Grosse I, Hähnel U, Junker A, Weisshaar B, Conrad U, Bäumlein H, *et al.* 2012. Toward the identification and regulation of the Arabidopsis thaliana ABI3 regulon. *Nucleic Acids Research*.

Mouillon J-M, Gustafsson P, Harryson P. **2006**. Structural investigation of disordered stress proteins. comparison of full-length dehydrins with isolated peptides of their conserved segments. *Plant Physiology* **141**: 638–650.

Mu J, Tan H, Zheng Q, Fu F, Liang Y, Zhang J, Yang X, Wang T, Chong K, Wang XJ, *et al.* 2008. LEAFY COTYLEDON1 is a key regulator of fatty acid biosynthesis in Arabidopsis. *Plant physiology* **148**: 1042.

Muchowski PJ, Clark JI. **1998**. ATP-enhanced molecular chaperone functions of the small heat shock protein human αB crystallin. *Proceedings of the National Academy of Sciences* **95**: 1004–1009.

Mundry M, Bornberg-Bauer E, Sammeth M, Feulner PGD. **2012**. Evaluating characteristics of de novo assembly software on 454 transcriptome data: a simulation approach. *PLoS ONE* **7**: e31410.

Munné-Bosch S, Falk J. **2004**. New insights into the function of tocopherols in plants. *Planta* **218**: 323–326.

Nakamoto H, Vígh L. **2007**. The small heat shock proteins and their clients. *Cellular and Molecular Life Sciences* **64**: 294–306.

Nakamura S, Lynch TJ, Finkelstein RR. **2001**. Physical interactions between ABA response loci of Arabidopsis. *The Plant Journal* **26**: 627–635.

Nakashima K, Fujita Y, Kanamori N, Katagiri T, Umezawa T, Kidokoro S, Maruyama K, Yoshida T, Ishiyama K, Kobayashi M, *et al.* 2009. Three Arabidopsis SnRK2 protein kinases, SRK2D/SnRK2. 2, SRK2E/SnRK2. 6/OST1 and SRK2I/SnRK2. 3, involved in ABA signaling are essential for the control of seed development and dormancy. *Plant and Cell Physiology* **50**: 1345-1363.

Nakashima K, Yamaguchi-Shinozaki K. **2013**. ABA signaling in stress-response and seed development. *Plant Cell Reports* **32**: 959–970.

Nambara E, Marion-Poll A. **2003**. ABA action and interactions in seeds. *Trends in Plant Science* **8**: 213–217.

Nambara E, Marion-Poll A. **2005**. Abscisic acid biosynthesis and catabolism. *Annual Review of Plant Biology* **56**: 165–185.

Nambara E, Naito S, McCourt P. **1992**. A mutant of Arabidopsis which is defective in seed development and storage protein accumulation is a new *abi3* allele. *The Plant Journal* **2**: 435–441.

Nishizawa A, Yabuta Y, Shigeoka S. **2008**. Galactinol and raffinose constitute a novel function to protect plants from oxidative damage. *Plant Physiology* **147**: 1251–1263.

Niu X, Helentjaris T, Bate NJ. **2002**. Maize ABI4 Binds coupling element1 in abscisic acid and sugar response genes. *The Plant Cell* **14**: 2565–2575.

Nocker S van, Ludwig P. **2003**. The WD-repeat protein superfamily in Arabidopsis: conservation and divergence in structure and function. *BMC Genomics* **4**: 50.

Noctor G, Foyer CH. **1998**. Ascorbate and glutathione: keeping active oxygen under control. *Annual Review of Plant Physiology and Plant Molecular Biology* **49**: 249–279.

Noctor G, Mhamdi A, Chaouch S, Han Y, Neukermans J, Marquez-Garcia B, Queval G, Foyer CH. **2012**. Glutathione in plants: an integrated overview. *Plant, Cell & Environment* **35**: 454–484.

O'Neil ST, Emrich SJ. **2013**. Assessing de novo transcriptome assembly metrics for consistency and utility. *BMC Genomics* **14**: 465.

Obendorf RL, Górecki RJ. **2012**. Soluble carbohydrates in legume seeds. *Seed Science Research* **22**: 219–242.

Ogas J, Cheng J-C, Sung ZR, Somerville C. **1997**. Cellular differentiation regulated by gibberellin in the Arabidopsis thaliana pickle mutant. *Science* **277**: 91–94.

Ogas J, Kaufmann S, Henderson J, Somerville C. **1999**. PICKLE is a CHD3 chromatin-remodeling factor that regulates the transition from embryonic to vegetative development in Arabidopsis. *Proceedings of the National Academy of Sciences* **96**: 13839–13844.

Ogawa M, Hanada A, Yamauchi Y, Kuwahara A, Kamiya Y, Yamaguchi S. **2003**. Gibberellin Biosynthesis and Response during Arabidopsis Seed Germination. *The Plant Cell* **15**: 1591–1604.

Oh E, Yamaguchi S, Kamiya Y, Bae G, Chung W-I, Choi G. **2006**. Light activates the degradation of PIL5 protein to promote seed germination through gibberellin in Arabidopsis. *The Plant Journal* **47**: 124–139.

Oh E, Yamaguchi S, Hu J, Yusuke J, Jung B, Paik I, Lee H-S, Sun T, Kamiya Y, Choi G. **2007**. PIL5, a phytochrome-interacting bHLH protein, regulates gibberellin responsiveness by binding directly to the GAI and RGA promoters in Arabidopsis seeds. *The Plant Cell* **19**: 1192–1208.

Ohri D, Bhargava A, Chatterjee A. **2004**. Nuclear DNA amounts in 112 species of tropical hardwoods - new estimates. *Plant Biology* **6**: 555–561.

Ohta M, Matsui K, Hiratsu K, Shinshi H, Ohme-Takagi M. **2001**. Repression domains of class II ERF transcriptional repressors share an essential motif for active repression. *The Plant Cell* **13**: 1959–1968.

Oliver MJ, Tuba Z, Mishler BD. **2000**. The evolution of vegetative desiccation tolerance in land plants. *Plant Ecology* **151**: 85–100.

Oliver MJ, Guo L, Alexander DC, Ryals JA, Wone BW, Cushman JC. **2011**. A sister group contrast using untargeted global metabolomic analysis delineates the biochemical regulation underlying desiccation tolerance in *Sporobolus stapfianus*. *The Plant Cell* **23**: 1231–1248.

Olvera-Carrillo Y, Campos F, Reyes JL, Garciarrubio A, Covarrubias AA. **2010**. Functional analysis of the group 4 Late Embryogenesis Abundant proteins reveals their relevance in the adaptive response during water deficit in Arabidopsis. *Plant Physiology* **154**: 373 –390.

Omar SA, Fu Q-T, Chen M-S, Wang G-J, Song S-Q, Elsheery NI, Xu ZF. **2011**. Identification and expression analysis of two small heat shock protein cDNAs from developing seeds of biodiesel feedstock plant *Jatropha curcas*. *Plant Science* **181**: 632–637.

Ooms J, Leon-Kloosterziel KM, Bartels D, Koornneef M, Karssen CM. **1993**. Acquisition of desiccation tolerance and longevity in seeds of arabidopsis thaliana (a comparative study using abscisic acid-insensitive *abi3* mutants). *Plant Physiology* **102**: 1185–1191.

Pammenter NW, Berjak P. **1999**. A review of recalcitrant seed physiology in relation to desiccation-tolerance mechanisms. *Seed Science Research* **9**: 13–37.

Pammenter NW, Berjak P. **2000**. Aspects of recalcitrant seed physiology. *Revista Brasileira de Fisiologia Vegetal* **12**: 56–69.

Panchuk II, Volkov RA, Schöffl F. **2002**. Heat stress- and heat shock transcription factor-dependent expression and activity of ascorbate peroxidase in Arabidopsis. *Plant Physiology* **129**: 838–853.

Pang Y, Peel GJ, Sharma SB, Tang Y, Dixon RA. **2008**. A transcript profiling approach reveals an epicatechin-specific glucosyltransferase expressed in the seed coat of *Medicago truncatula*. *Proceedings of the National Academy of Sciences* **105**: 14210–14215.

Panikulangara TJ, Eggers-Schumacher G, Wunderlich M, Stransky H, Schöffl F. **2004**. Galactinol synthase1. a novel heat shock factor target gene responsible for heat-induced synthesis of Raffinose Family Oligosaccharides in Arabidopsis. *Plant Physiology* **136**: 3148–3158.

Panza V, Distéfano AJ, Carjuzaa P, Láinez V, Del Vas M, Maldonado S. **2007**. Detection of dehydrin-like proteins in embryos and endosperm of mature *Euterpe edulis* seeds. *Protoplasma* **231**: 1–5.

Parcy F, Valon C, Raynal M, Gaubier-Comella P, Delseny M, Giraudat J. **1994**. Regulation of gene expression programs during Arabidopsis seed development: roles of the ABI3 locus and of endogenous abscisic acid. *The Plant Cell* **6**: 1567–1582.

Parcy F, Valon C, Kohara A, Misera S, Giraudat J. **1997**. The ABSCISIC ACID-INSENSITIVE3, FUSCA3, and LEAFY COTYLEDON1 loci act in concert to control multiple aspects of Arabidopsis seed development. *The Plant Cell* **9**: 1265 –1277.

Park S-Y, Fung P, Nishimura N, Jensen DR, Fujii H, Zhao Y, Lumba S, Santiago J, Rodrigues A, Chow TF, *et al.* **2009**. Abscisic acid inhibits type 2c protein phosphatases via the PYR/PYL family of START proteins. *Science* **324**: 1068–1071.

Park B-J, Liu Z, Kanno A, Kameya T. **2005**. Genetic improvement of Chinese cabbage for salt and drought tolerance by constitutive expression of a B. napus LEA gene. *Plant Science* **169**: 553–558.

Park S-C, Kim Y-H, Jeong JC, Kim CY, Lee H-S, Bang J-W, Kwak S-S. **2011**. Sweetpotato *late embryogenesis abundant 14* (*IbLEA14*) gene influences lignification and increases osmotic- and salt stress-tolerance of transgenic calli. *Planta* **233**: 621–634.

Parkhey S, Naithani SC, Keshavkant S. ROS production and lipid catabolism in desiccating *Shorea robusta* seeds during aging. *Plant Physiology and Biochemistry* **57**: 261-267.

Pence VC. **1991**. Abscisic acid in developing zygotic embryos of *Theobroma cacao*. *Plant physiology* **95**: 1291-1293.

Peng J, Carol P, Richards DE, King KE, Cowling RJ, Murphy GP, Harberd NP. **1997**. The Arabidopsis GAI gene defines a signaling pathway that negatively regulates gibberellin responses. *Genes & Development* **11**: 3194–3205.

Perez DE, Hoyer JS, Johnson AI, Moody ZR, Lopez J, Kaplinsky NJ. **2009**. BOBBER1 is a noncanonical arabidopsis small heat shock protein required for both development and thermotolerance. *Plant Physiology* **151**: 241–252.

Perruc E, Kinoshita N, Lopez-Molina L. **2007**. The role of chromatin-remodeling factor PKL in balancing osmotic stress responses during Arabidopsis seed germination. *The Plant Journal* **52**: 927–936.

Phillips JR, Fischer E, Baron M, Van Den Dries N, Facchinelli F, Kutzer M, Rahmanzadeh R, Remus D, Bartels D. **2008**. *Lindernia brevidens*: a novel desiccation-tolerant vascular plant, endemic to ancient tropical rainforests. *The Plant Journal* **54**: 938–948.

Plana M, Itarte E, Eritja R, Goday A, Pagès M, Martínez MC. **1991**. Phosphorylation of maize RAB-17 protein by casein kinase 2. *Journal of Biological Chemistry* **266**: 22510–22514.

Porembski S, Barthlott W. **2000**. Granitic and gneissic outcrops (inselbergs) as centers of diversity for desiccation-tolerant vascular plants. *Plant Ecology* **151**: 19–28.

Prieto-Dapena P, Castano R, Almoguera C, Jordano J. **2006**. Improved resistance to controlled deterioration in transgenic seeds. *Plant physiology* **142**: 1102.

Pritchard HW, Daws MI, Fletcher BJ, Gamene CS, Msanga HP, Omondi W. **2004**. Ecological correlates of seed desiccation tolerance in tropical African dryland trees. *American journal of botany* **91**: 863.

Pritchard HW, Tompsett PB, Manger KR. **1996**. Development of a thermal time model for the quantification of dormancy loss in *Aesculus hippocastanum* seeds. *Seed Science Research* **6**: 127–135.

Proctor MCF, Tuba Z. **2002**. Poikilohydry and homoihydry: antithesis or spectrum of possibilities? *New Phytologist* **156**: 327–349.

Puhakainen T, Hess MW, Mäkelä P, Svensson J, Heino P, Palva ET. **2004**. Overexpression of multiple dehydrin genes enhances tolerance to freezing stress in Arabidopsis. *Plant Molecular Biology* **54**: 743–753.

Pukacka S, Ratajczak E. **2007**. Ascorbate and glutathione metabolism during development and desiccation of orthodox and recalcitrant seeds of the genus *Acer*. *Functional Plant Biology* **34**: 601–613.

Qin X, Zeevaart JAD. **1999**. The 9-cis-epoxycarotenoid cleavage reaction is the key regulatory step of abscisic acid biosynthesis in water-stressed bean. *Proceedings of the National Academy of Sciences* **96**: 15354–15361.

Raynal M, Guilleminot J, Gueguen C, Cooke R, Delseny M, Gruber V. **1999**. Structure, organization and expression of two closely related novel Lea (late-embryogenesis-abundant) genes in *Arabidopsis thaliana*. *Plant Molecular Biology* **40**: 153–165.

Raz V, Bergervoet JH, Koornneef M. **2001**. Sequential steps for developmental arrest in Arabidopsis seeds. *Development* **128**: 243.

Reeves WM, Lynch TJ, Mobin R, Finkelstein RR. **2011**. Direct targets of the transcription factors ABA-Insensitive(ABI)4 and ABI5 reveal synergistic action by ABI4 and several bZIP ABA response factors. *Plant Molecular Biology* **75**: 347–363.

Reidt W, Wohlfarth T, Ellerström M, Czihal A, Tewes A, Ezcurra I, Rask L, Bäumlein H. **2000**. Gene regulation during late embryogenesis: the RY motif of maturation-specific gene promoters is a direct target of the FUS3 gene product. *The Plant Journal* **21**: 401–408.

Reyes JL, Rodrigo M-J, Colmenero-Flores JM, Gil J-V, Garay-Arroyo A, Campos F, Salamini F, Bartels D, Covarrubias AA. **2005**. Hydrophilins from distant organisms can protect enzymatic activities from water limitation effects in vitro. *Plant, Cell & Environment* **28**: 709–718.

Ricker JV, Tsvetkova NM, Wolkers WF, Leidy C, Tablin F, Longo M, Crowe JH. **2003**. Trehalose maintains phase separation in an air-dried binary lipid mixture. *Biophysical Journal* **84**: 3045–3051.

Rinne PLH, Kaikuranta PLM, Plas LHW van der, Schoot C van der. **1999**. Dehydrins in cold-acclimated apices of birch (Betula pubescens Ehrh.): production, localization and potential role in rescuing enzyme function during dehydration. *Planta* **209**: 377–388.

Roberts EH. **1973**. Predicting the storage life of seeds. *Seed Science and technology* **1**: 499–514.

Rohde A, Rycke RD, Beeckman T, Engler G, Montagu MV, Boerjan W. **2000**. ABI3 affects plastid differentiation in dark-grown Arabidopsis seedlings. *The Plant Cell* **12**: 35–52.

Rohde A, Prinsen E, Rycke RD, Engler G, Montagu MV, Boerjan W. **2002**. PtABI3 impinges on the growth and differentiation of embryonic leaves during bud set in poplar. *The Plant Cell* **14**: 1885–1901.

Rohila JS, Jain RK, Wu R. **2002**. Genetic improvement of Basmati rice for salt and drought tolerance by regulated expression of a barley *Hva1* cDNA. *Plant Science* **163**: 525–532.

Röhrig H, Schmidt J, Colby T, Bräutigam A, Hufnagel P, Bartels D. **2006**. Desiccation of the resurrection plant *Craterostigma plantagineum* induces dynamic changes in protein phosphorylation. *Plant, Cell & Environment* **29**: 1606–1617.

Rojas A, Almoguera C, Jordano J. **1999**. Transcriptional activation of a heat shock gene promoter in sunflower embryos: synergism between ABI3 and heat shock factors. *The Plant Journal* **20**: 601–610.

Rosnoblet C, Aubry C, Leprince O, Vu BL, Rogniaux H, Buitink J. **2007**. The regulatory gamma subunit SNF4b of the sucrose non-fermenting-related kinase complex is involved in longevity and stachyose accumulation during maturation of *Medicago truncatula* seeds. *The Plant Journal* **51**: 47–59.

Roxas VP, Smith RK, Allen ER, Allen RD. **1997**. Overexpression of glutathione S-transferase/glutathioneperoxidase enhances the growth of transgenic tobacco seedlings during stress. *Nature Biotechnology* **15**: 988–991.

Rubinovich L, Weiss D. **2010**. The Arabidopsis cysteine-rich protein GASA4 promotes GA responses and exhibits redox activity in bacteria and in planta. *The Plant Journal* **64**: 1018–1027.

Saito S, Hirai N, Matsumoto C, Ohigashi H, Ohta D, Sakata K, Mizutani M. **2004**. Arabidopsis CYP707As encode (+)-abscisic acid 8′-hydroxylase, a key enzyme in the oxidative catabolism of abscisic acid. *Plant Physiology* **134**: 1439–1449.

Sakata Y, Nakamura I, Taji T, Tanaka S, Quatrano RS. **2010**. Regulation of the ABA-responsive Em promoter by ABI3 in the moss *Physcomitrella patens*: Role of the ABA response element and the RY element. *Plant Signaling & Behavior* **5**: 1061–1066.

Santos-Mendoza M, Dubreucq B, Miquel M, Caboche M, Lepiniec L. **2005**. LEAFY COTYLEDON 2 activation is sufficient to trigger the accumulation of oil and seed specific mRNAs in Arabidopsis leaves. *FEBS letters* **579**: 4666–4670.

Santos-Mendoza M, Dubreucq B, Baud S, Parcy F, Caboche M, Lepiniec L. **2008**. Deciphering gene regulatory networks that control seed development and maturation in Arabidopsis. *The Plant Journal* **54**: 608–620.

Sato Y, Yokoya S. **2008**. Enhanced tolerance to drought stress in transgenic rice plants overexpressing a small heat-shock protein, sHSP17.7. *Plant Cell Reports* **27**: 329–334.

Sattler SE, Gilliland LU, Magallanes-Lundback M, Pollard M, DellaPenna D. **2004**. Vitamin E Is Essential for seed longevity and for preventing lipid peroxidation during germination. *The Plant Cell* **16**: 1419–1432.

Scharf K-D, Berberich T, Ebersberger I, Nover L. The plant heat stress transcription factor (Hsf) family: Structure, function and evolution. *Biochimica et Biophysica Acta* **1819**: 104–119.

Schliesky S, Gowik U, Weber APM, Bräutigam A. **2012**. RNA-seq assembly – are we there yet? *Frontiers in Plant Systems Biology* **3**: 220.

Schwartz SH, Qin X, Zeevaart JAD. **2003**. Elucidation of the indirect pathway of abscisic acid biosynthesis by mutants, genes, and enzymes. *Plant Physiology* **131**: 1591–1601.

Seo JS, Sohn HB, Noh K, Jung C, An JH, Donovan CM, Somers DA, Kim DI, Jeong S-C, Kim C-G, *et al.* 2012. Expression of the Arabidopsis *AtMYB44* gene confers drought/salt-stress tolerance in transgenic soybean. *Molecular Breeding* **29**: 601–608.

Shen B, Allen WB, Zheng P, Li C, Glassman K, Ranch J, Nubel D, Tarczynski MC. **2010**. Expression of *ZmLEC1* and *ZmWRI1* increases seed oil production in maize. *Plant physiology* **153**: 980-987.

Shen B, Sinkevicius KW, Selinger DA, Tarczynski MC. **2006**. The Homeobox gene *GLABRA2* affects seed oil content in Arabidopsis. *Plant Molecular Biology* **60**: 377–387.

Shendure J, Ji H. **2008**. Next-generation DNA sequencing. *Nature biotechnology* **26**: 1135–1145.

Shih M, Lin S, Hsieh J, Tsou C, Chow T, Lin T, Hsing YC. **2004**. Gene cloning and characterization of a soybean (*Glycine max* L.) LEA protein, GmPM16. *Plant Molecular Biology* **56**: 689–703.

Shih M, Hsieh T, Lin T, Hsing YC, Hoekstra FA. **2010**. Characterization of two soybean (*Glycine max* L.) LEA IV proteins by circular dichroism and fourier transform infrared spectrometry. *Plant and Cell Physiology* **51**: 395–407.

Shih M-D, Hsieh T-Y, Jian W-T, Wu M-T, Yang S-J, Hoekstra FA, Hsing Y-IC. **2012**. Functional studies of soybean (*Glycine max* L.) seed LEA proteins GmPM6, GmPM11, and GmPM30 by CD and FTIR spectroscopy. *Plant Science* **196**: 152–159.

Shimada TL, Shimada T, Takahashi H, Fukao Y, Hara-Nishimura I. **2008**. A novel role for oleosins in freezing tolerance of oilseeds in *Arabidopsis thaliana*. *The Plant Journal* **55**: 798–809.

Shinozaki K, Yamaguchi-Shinozaki K, Seki M. **2003**. Regulatory network of gene expression in the drought and cold stress responses. *Current Opinion in Plant Biology* **6**: 410–417.

Singh S, Cornilescu CC, Tyler RC, Cornilescu G, Tonelli M, Lee MS, Markley JL. **2005**. Solution structure of a late embryogenesis abundant protein (LEA14) from *Arabidopsis thaliana*, a cellular stress-related protein. *Protein Science* **14**: 2601–2609.

Sinniah UR, Ellis RH, John P. **1998**. Irrigation and seed quality development in rapid-cycling Brassica: soluble carbohydrates and heat-stable proteins. *Annals of Botany* **82**: 647–655.

Sivamani E, Bahieldin A, Wraith JM, Al-Niemi T, Dyer WE, Ho T-HD, Qu R. **2000**. Improved biomass productivity and water use efficiency under water deficit conditions in transgenic wheat constitutively expressing the barley *HVA1* gene. *Plant Science* **155**: 1–9.

Smirnoff N. **1993**. The role of active oxygen in the response of plants to water deficit and desiccation. *New Phytologist* **125**: 27–58.

Smýkal P, Mašín J, Hrdý I, Konopásek I, Žárský V. **2000**. Chaperone activity of tobacco HSP18, a small heat-shock protein, is inhibited by ATP. *The Plant Journal* **23**: 703–713.

Söderman E, Mattsson J, Engström P. **1996**. The Arabidopsis homeobox gene *ATHB-7* is induced by water deficit and by abscisic acid. *The Plant Journal* **10**: 375–381.

Söderman EM, Brocard IM, Lynch TJ, Finkelstein RR. **2000**. Regulation and function of the Arabidopsis *ABA-insensitive4* gene in seed and abscisic acid response signaling networks. *Plant Physiology* **124**: 1752–1765.

Soulages JL, Kim K, Walters C, Cushman JC. **2002**. Temperature-induced extended helix/random coil transitions in a group 1 Late Embryogenesis-Abundant protein from soybean. *Plant Physiology* **128**: 822–832.

Soulages JL, Kim K, Arrese EL, Walters C, Cushman JC. **2003**. Conformation of a Group 2 Late Embryogenesis Abundant protein from soybean. evidence of Poly (l-Proline)-type II structure. *Plant Physiology* **131**: 963–975.

Sreenivasulu N, Radchuk V, Strickert M, Miersch O, Weschke W, Wobus U. **2006**. Gene expression patterns reveal tissue-specific signaling networks controlling programmed cell death and ABA- regulated maturation in developing barley seeds. *The Plant Journal* **47**: 310–327.

Stacy RAP, Munthe E, Steinum T, Sharma B, Aalen RB. **1996**. A peroxiredoxin antioxidant is encoded by a dormancy-related gene,*Per1*, expressed during late development in the aleurone and embryo of barley grains. *Plant Molecular Biology* **31**: 1205–1216.

Stacy RAP, Nordeng TW, Culiáñez-Macià FA, Aalen RB. **1999**. The dormancy-related peroxiredoxin anti-oxidant, PER1, is localized to the nucleus of barley embryo and aleurone cells. *The Plant Journal* **19**: 1–8.

Steadman KJ, Pritchard HW, Dey PM. **1996**. Tissue-specific soluble sugars in seeds as indicators of storage category. *Annals of Botany* **77**: 667–674.

Steponkus PL, Uemura M, Joseph RA, Gilmour SJ, Thomashow MF. **1998**. Mode of action of the COR15a gene on the freezing tolerance of *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences* **95**: 14570–14575.

Stone SL, Kwong LW, Yee KM, Pelletier J, Lepiniec L, Fischer RL, Goldberg RB, Harada JJ. **2001**. LEAFY COTYLEDON2 encodes a B3 domain transcription factor that induces embryo development. *Proceedings of the National Academy of Sciences of the United States of America* **98**: 11806-11811.

Stone SL, Williams LA, Farmer LM, Vierstra RD, Callis J. **2006**. KEEP ON GOING, a RING E3 ligase essential for Arabidopsis growth and development, is involved in abscisic acid signaling. *The Plant Cell* **18**: 3415–3428.

Stone SL, Braybrook SA, Paula SL, Kwong LW, Meuser J, Pelletier J, Hsieh TF, Fischer RL, Goldberg RB, Harada JJ. **2008**. Arabidopsis LEAFY COTYLEDON2 induces maturation traits and auxin activity: Implications for somatic embryogenesis. *Proceedings of the National Academy of Sciences* **105**: 3151-3156.

Stupnikova I, Benamar A, Tolleter D, Grelet J, Borovskii G, Dorne A-J, Macherel D. **2006**. Pea seed mitochondria are endowed with a remarkable tolerance to extreme physiological temperatures. *Plant Physiology* **140**: 326–335.

Suarez-Rodriguez MCS, Edsgärd D, Hussain SS, Alquezar D, Rasmussen M, Gilbert T, Nielsen BH, Bartels D, Mundy J. **2010**. Transcriptomes of the desiccation-tolerant resurrection plant *Craterostigma plantagineum*. *The Plant Journal* **63**: 212–228.

Sugliani M, Brambilla V, Clerkx EJM, Koornneef M, Soppe WJJ. **2010**. The conserved splicing factor SUA controls alternative splicing of the developmental regulator ABI3 in Arabidopsis. *The Plant Cell* **22**: 1936–1946.

Sun WQ, Leopold AC. **1993**. The glassy state and accelerated aging of soybeans. *Physiologia Plantarum* **89**: 767–774.

Sun WQ, Irving TC, Leopold AC. **1994**. The role of sugar, vitrification and membrane phase transition in seed desiccation tolerance. *Physiologia Plantarum* **90**: 621–628.

Sun WQ, Leopold AC. **1995**. The Maillard reaction and oxidative stress during aging of soybean seeds. *Physiologia Plantarum* **94**: 94–104.

Sun WQ. **1997**. Glassy state and seed storage stability: the wlf kinetics of seed viability loss at t>tgand the plasticization effect of water on storage stability. *Annals of Botany* **79**: 291–297.

Sun WQ, Davidson P. **1998**. Protein inactivation in amorphous sucrose and trehalose matrices: effects of phase separation and crystallization. *Biochimica et Biophysica Acta* **1425**: 235–244.

Sun WQ. **1999**. Desiccation sensitivity of recalcitrant seeds and germinated orthodox seeds: can germinated orthodox seeds serve as a model system for studies of recalcitrance? In: Marzalina M, Khoo KC, Jayanthi N, Tsan FY, Krishnapillay B, eds. *Recalcitrant seeds*. Forest Research Institute, Kuala Lumpur, Malaysia. 29-42

Sun W, Bernard C, Van De Cotte B, Van Montagu M, Verbruggen N. **2001**. At-HSP17.6A, encoding a small heat-shock protein in Arabidopsis, can enhance osmotolerance upon overexpression. *The Plant Journal* **27**: 407–415.

Sun W, Van Montagu M, Verbruggen N. **2002**. Small heat shock proteins and stress tolerance in plants. *Biochimica et Biophysica Acta* **1577**: 1–9.

Šunderlíková V, Salaj J, Kopecky D, Salaj T, Wilhem E, Matušíková I. **2009**. Dehydrin genes and their expression in recalcitrant oak (*Quercus robur*) embryos. *Plant Cell Reports* **28**: 1011–1021.

Suzuki M, Kao CY, McCarty DR. **1997**. The conserved B3 domain of VIVIPAROUS1 has a cooperative DNA binding activity. *The Plant Cell* **9**: 799–807.

Suzuki M, Kao C-Y, Cocciolone S, McCarty DR. **2001**. Maize *VP1* complements Arabidopsis *abi3* and confers a novel ABA/auxin interaction in roots. *The Plant Journal* **28**: 409–418.

Suzuki M, Ketterling MG, Li Q-B, McCarty DR. **2003**. Viviparous1 alters global gene expression patterns through regulation of abscisic acid signaling. *Plant Physiology* **132**: 1664–1677.

Suzuki N, Mittler R. **2006**. Reactive oxygen species and temperature stresses: A delicate balance between signaling and destruction. *Physiologia Plantarum* **126**: 45–51.

Suzuki M, Wang HH-Y, McCarty DR. **2007**. Repression of the LEAFY COTYLEDON 1/B3 Regulatory Network in plant embryo development by VP1/ABSCISIC ACID INSENSITIVE 3-LIKE B3 genes. *Plant Physiology* **143**: 902–911.

Suzuki M, McCarty DR. **2008**. Functional symmetry of the B3 network controlling seed development. *Current opinion in plant biology* **11**: 548–553.

Svensson J, Palva ET, Welin B. **2000**. Purification of recombinant *Arabidopsis thaliana* dehydrins by metal ion affinity chromatography. *Protein Expression and Purification* **20**: 169–178.

Swain SM, Reid JB, Kamiya Y. **1997**. Gibberellins are required for embryo growth and seed development in pea. *The Plant Journal* **12**: 1329–1338.

Tadege M, Ratet P, Mysore KS. **2005**. Insertional mutagenesis: a Swiss Army knife for functional genomics of *Medicago truncatula*. *Trends in Plant Science* **10**: 229–235.

Taji T, Ohsumi C, Iuchi S, Seki M, Kasuga M, Kobayashi M, Yamaguchi-Shinozaki K, Shinozaki K. **2002**. Important roles of drought- and cold-inducible genes for galactinol synthase in stress tolerance in Arabidopsis thaliana. *The Plant Journal* **29**: 417–426.

Tanaka M, Kikuchi A, Kamada H. **2008**. The Arabidopsis histone deacetylases HDA6 and HDA19 contribute to the repression of embryonic properties after germination. *Plant Physiology* **146**: 149–161.

Tejedor-Cano J, Prieto-Dapena P, Almoguera C, Carranco R, Hiratsu K, Ohme-Takagi M, Jordano J. **2010**. Loss of function of the HSFA9 seed longevity program. *Plant, Cell & Environment* **33**: 1408–1417.

Tezuka K, Taji T, Hayashi T, Sakata Y. **2013**. A novel *abi5* allele reveals the importance of the conserved Ala in the C3 domain for regulation of downstream genes and salt tolerance during germination in Arabidopsis. *Plant Signaling & Behavior* **8**: e23455.

Thomas SG, Phillips AL, Hedden P. **1999**. Molecular cloning and functional expression of gibberellin 2- oxidases, multifunctional enzymes involved in gibberellin deactivation. *Proceedings of the National Academy of Sciences* **96**: 4698–4703.

Thompson AJ, Jackson AC, Parker RA, Morpeth DR, Burbidge A, Taylor IB. **2000**. Abscisic acid biosynthesis in tomato: regulation of zeaxanthin epoxidase and 9-cis-epoxycarotenoid dioxygenase mRNAs by light/dark cycles, water stress and abscisic acid. *Plant Molecular Biology* **42**: 833–845.

Thorburn AW, Brand JC, Truswell AS. **1987**. Slowly digested and absorbed carbohydrate in traditional bushfoods: a protective factor against diabetes? *The American Journal of Clinical Nutrition* **45**: 98–106.

Tiedemann J, Rutten T, Mönke G, Vorwieger A, Rolletschek H, Meissner D, Milkowski C, Petereck S, Mock HP, Zank T, *et al.* 2008. Dissection of a complex seed phenotype: novel insights of FUSCA3 regulated developmental processes. *Developmental biology* **317**: 1–12.

To A, Valon C, Savino G, Guilleminot J, Devic M, Giraudat J, Parcy F. **2006**. A network of local and redundant gene regulation governs Arabidopsis seed maturation. *The Plant Cell* **18**: 1642.

Tolleter D, Hincha DK, Macherel D. **2010**. A mitochondrial late embryogenesis abundant protein stabilizes model membranes in the dry state. *Biochimica Biophysica* Acta **1798**: 1926-1933.

Tolleter D, Jaquinod M, Mangavel C, Passirani C, Saulnier P, Manon S, Teyssier E, Payet N, Avelange-Macherel M-H, Macherel D. **2007**. Structure and function of a mitochondrial late embryogenesis abundant protein are revealed by desiccation. *The Plant Cell* **9**: 1580–1589.

Tommasi F, Paciolla C, Arrigoni O. **1999**. The ascorbate system in recalcitrant and orthodox seeds. *Physiologia Plantarum* **105**: 193–198.

Tompsett PB, Pritchard HW. **1993**. Water status changes during development in relation to the germination and desiccation tolerance of *Aesculus hippocastanum* L. seeds. *Annals of Botany* **71**: 107 –116.

Tsukagoshi H, Saijo T, Shibata D, Morikami A, Nakamura K. **2005**. Analysis of a sugar response mutant of Arabidopsis identified a novel B3 domain protein that functions as an active transcriptional repressor. *Plant Physiology* **138**: 675–685.

Tsukagoshi H, Morikami A, Nakamura K. **2007**. Two B3 domain transcriptional repressors prevent sugar-inducible expression of seed maturation genes in Arabidopsis seedlings. *Proceedings of the National Academy of Sciences* **104**: 2543-2547.

De Tullio MC, Arrigoni O. **2003**. The ascorbic acid system in seeds: to protect and to serve. *Seed Science Research* **13**: 249–260.

Tunnacliffe A, Wise MJ. **2007**. The continuing conundrum of the LEA proteins. *Naturwissenschaften* **94**: 791–812.

Tweddle JC, Dickie JB, Baskin CC, Baskin JM. **2003**. Ecological aspects of seed desiccation sensitivity. *Journal of Ecology* **91**: 294–304.

Uno Y, Furihata T, Abe H, Yoshida R, Shinozaki K, Yamaguchi-Shinozaki K. **2000**. Arabidopsis basic leucine zipper transcription factors involved in an abscisic acid-dependent signal transduction pathway under drought and high-salinity conditions. *Proceedings of the National Academy of Sciences* **97**: 11632–11637.

Vandecasteele C, Teulat-Merah B, Morère-Le Paven M-C, Leprince O, Ly Vu B, Viau L, Ledroit L, Pelletier S, Payet N, Satour P, *et al.* 2011. Quantitative trait loci analysis reveals a correlation between the ratio of sucrose/raffinose family oligosaccharides and seed vigour in Medicago truncatula. *Plant, Cell & Environment* **34**: 1473–1487.

Vaucheret H. **2006**. Post-transcriptional small RNA pathways in plants: mechanisms and regulations. *Genes & Development* **20**: 759–771.

Veerappan V, Wang J, Kang M, Lee J, Tang Y, Jha AK, Shi H, Palanivelu R, Allen RD. **2012**. A novel HSI2 mutation in Arabidopsis affects the PHD-like domain and leads to derepression of seed-specific gene expression. *Planta* **236**: 1–17.

Velasco R, Salamini F, Bartels D. **1998**. Gene structure and expression analysis of the drought- and abscisic acid-responsive CDeT11-24 gene family from the resurrection plant *Craterostigma plantagineum* Hochst. *Planta* **204**: 459–471.

Verdier J, Lalanne D, Pelletier S, Torres-Jerez I, Righetti K, Bandyopadhyay K, Leprince O, Chatelain E, Vu BL, Gouzy J, *et al.* 2013. A regulatory network-based approach dissects late maturation processes related to the acquisition of desiccation tolerance and longevity of *Medicago truncatula* seeds. *Plant Physiology* doi.org/10.1104/pp.113.222380.

Vertucci CW. **1990**. Calorimetric studies of the state of water in seed tissues. *Biophysical Journal* **58**: 1463–1471.

Vicré M, Sherwin HW, Driouich A, Jaffer MA, Farrant JM. **1999**. Cell wall characteristics and structure of hydrated and dry leaves of the resurrection plant *Craterostigma wilmsii*, a microscopical study. *Journal of Plant Physiology* **155**: 719–726.

Vicré M, Lerouxel O, Farrant J, Lerouge P, Driouich A. **2004a**. Composition and desiccation-induced alterations of the cell wall in the resurrection plant *Craterostigma wilmsii*. *Physiologia Plantarum* **120**: 229–239.

Vicré M, Farrant JM, Driouich A. **2004b**. Insights into the cellular mechanisms of desiccation tolerance among angiosperm resurrection plant species. *Plant, Cell & Environment* **27**: 1329–1340.

Vierling E. **1991**. The roles of heat shock proteins in plants. *Annual Review of Plant Physiology and Plant Molecular Biology* **42**: 579–620.

Vierling E, Nagao RT, DeRocher AE, Harris LM. **1988**. A heat shock protein localized to chloroplasts is a member of a eukaryotic superfamily of heat shock proteins. *The EMBO Journal* **7**: 575–581.

Vijay N, Poelstra JW, Künstner A, Wolf JBW. **2012**. Challenges and strategies in transcriptome assembly and differential gene expression quantification. A comprehensive in silico assessment of RNA-seq experiments. *Molecular Ecology* **22**: 620-634.

Vilardell J, Goday A, Freire MA, Torrent M, Martínez MC, Torne JM, Pagès M. **1990**. Gene sequence, developmental expression, and protein phosphorylation of RAB-17 in maize. *Plant Molecular Biology* **14**: 423–432.

Walker BD, Kowalski M, Goh WC, Kozarsky K, Krieger M, Rosen C, Rohrschneider L, Haseltine WA, Sodroski J. **1987**. Inhibition of human immunodeficiency virus syncytium formation and virus replication by castanospermine. *Proceedings of the National Academy of Sciences* **84**: 8120–8124.

Wall PK, Leebens-Mack J, Chanderbali AS, Barakat A, Wolcott E, Liang H, Landherr L, Tomsho LP, Hu Y, Carlson JE, *et al.* 2009. Comparison of next generation sequencing technologies for transcriptome characterization. *BMC genomics* **10**: 347.

Walters C. **1998**. Understanding the mechanisms and kinetics of seed aging. *Seed Science Research* **8**: 223–244.

Walters C, Ried JL, Walker-Simmons MK. **1997**. Heat-soluble proteins extracted from wheat embryos have tightly bound sugars and unusual hydration properties. *Seed Science Research* **7**: 125–134.

Walters C, Farrant J, Pammenter N, Berjak P. **2002**. Desiccation Stress and damage. In: *Desiccation and survival in plants: Drying without dying*, Black M, Pritchard HW, eds. Wallingford, UK, CABI Publishing. 263-291.

Walters C, Wheeler LM, Grotenhuis JM. **2005a**. Longevity of seeds stored in a genebank: species characteristics. *Seed Science Research* **15**: 1–20.

Walters C, Hill LM, Wheeler LJ. **2005b**. Dying while dry: kinetics and mechanisms of deterioration in desiccated organisms. *Integrative and Comparative Biology* **45**: 751–758.

Walters C, Koster K. **2007**. Structural dynamics and desiccation damage in plant reproductive organs. In: Jenks MA and Wood AJ eds. *Plant Desiccation Tolerance*. Blackwell publishing, Iowa USA, 251-280.

Walters C, Berjak P, Pammenter N, Kennedy K, Raven P. **2013**. Preservation of recalcitrant seeds. *Science* **339**: 915–916.

Wang H, Caruso LV, Downie AB, Perry SE. **2004**. The embryo MADS domain protein AGAMOUS-Like 15 directly regulates expression of a gene encoding an enzyme involved in gibberellin metabolism. *The Plant Cell* **16**: 1206–1219.

Wang HL, Grusak MA. **2005**. Structure and development of *Medicago truncatula* pod wall and seed coat. *Annals of Botany* **95**: 737–747.

Wang F, Perry SE. **2013**. Identification of direct targets of FUSCA3, a key regulator of Arabidopsis seed development. *Plant Physiology* **161**: 1251–1264.

Waters ER, Lee GJ, Vierling E. **1996**. Evolution, structure and function of the small heat shock proteins in plants. *Journal of Experimental Botany* **47**: 325–338.

Webb MA, Arnott HJ. **1982**. Cell wall conformation in dry seeds in relation to the preservation of structural integrity during desiccation. *American Journal of Botany* **69**: 1657.

Wehmeyer N, Hernandez LD, Finkelstein RR, Vierling E. **1996**. Synthesis of small heat-shock proteins is part of the developmental program of late seed maturation. *Plant Physiology* **112**: 747 –757.

Wehmeyer N, Vierling E. **2000**. The expression of small heat shock proteins in seeds responds to discrete developmental signals and suggests a general protective role in desiccation tolerance. *Plant Physiology* **122**: 1099–1108.

Wen C-K, Chang C. **2002**. Arabidopsis RGL1 encodes a negative regulator of gibberellin responses. *The Plant Cell* **14**: 87–100.

Wenger Y, Galliot B. **2013**. RNAseq versus genome-predicted transcriptomes: a large population of novel transcripts identified in an Illumina-454 Hydra transcriptome. *BMC Genomics* **14**: 204.

White CN, Proebsting WM, Hedden P, Rivin CJ. **2000**. Gibberellins and seed development in maize. evidence that gibberellin/abscisic acid balance governs germination versus maturation pathways. *Plant Physiology* **122**: 1081–1088.

Wind JJ, Peviani A, Snel B, Hanson J, Smeekens SC. **2013**. ABI4: versatile activator and repressor. *Trends in Plant Science* **18**: 125–132.

Winter D, Vinegar B, Nahal H, Ammar R, Wilson GV, Provart NJ. **2007**. An 'electronic fluorescent pictograph' browser for exploring and analyzing large-scale biological data sets. *PLoS ONE* **2**: e718.

Wise MJ. **2002**. The POPPs: clustering and searching using peptide probability profiles. *Bioinformatics* **18**: S38–S45.

Wise MJ. **2003**. LEAping to conclusions: A computational reanalysis of late embryogenesis abundant proteins and their possible roles. *BMC Bioinformatics* **4**: 52.

Wise MJ, Tunnacliffe A. **2004**. POPP the question: what do LEA proteins do? *Trends in Plant Science* **9**: 13–17.

Wisniewski M, Webb R, Balsamo R, Close TJ, Yu X-M, Griffith M. **1999**. Purification, immunolocalization, cryoprotective, and antifreeze activity of PCA60: A dehydrin from peach (*Prunus persica*). *Physiologia Plantarum* **105**: 600–608.

Wolkers WF, McCready S, Brandt WF, Lindsey GG, Hoekstra FA. **2001**. Isolation and characterization of a D-7 LEA protein from pollen that stabilizes glasses in vitro. *Biochimica et Biophysica Acta (BBA)* **1544**: 196–206.

Wood AJ. **2007**. The nature and distribution of vegetative desiccation-tolerance in hornworts, liverworts and mosses. *The Bryologist* **110**: 163–177.

Wood AJ, Jenks MA. **2007**. Plant desiccation tolerance: diversity, distribution and real-world applications. In: Jenks MA and Wood AJ eds. *Plant Desiccation Tolerance*. Blackwell publishing, Iowa USA, 3-10.

Wright E, Dixon RA, Wang Z-Y. **2006**. *Medicago truncatula* transformation using cotyledon explants. In: Wang K, ed. *Methods in Molecular Biology. Agrobacterium Protocols*. Humana Press, 129–136.

Wu L, Zhou M, Shen C, Liang J, Lin J. **2012**. Transgenic tobacco plants over expressing cold regulated protein CbCOR15b from Capsella bursa-pastoris exhibit enhanced cold tolerance. *Journal of Plant Physiology* **169**: 1408–1416.

Xing X, Liu Y, Kong X, Liu Y, Li D. **2011**. Overexpression of a maize dehydrin gene, *ZmDHN2b*, in tobacco enhances tolerance to low temperature. *Plant Growth Regulation* **65**: 109–118.

Xiong L, Ishitani M, Lee H, Zhu J-K. **2001**. The Arabidopsis LOS5/ABA3 Locus Encodes a Molybdenum Cofactor Sulfurase and Modulates Cold Stress– and Osmotic Stress–Responsive Gene Expression. *The Plant Cell* **13**: 2063–2083.

Xu D, Duan X, Wang B, Hong B, Ho T, Wu R. **1996**. Expression of a Late Embryogenesis Abundant protein gene, *HVA1*, from barley confers tolerance to water deficit and salt stress in transgenic rice. *Plant Physiology* **110**: 249–257.

Yamagishi K, Nagata N, Yee KM, Braybrook SA, Pelletier J, Fujioka S, Yoshida S, Fischer RL, Goldberg RB, Harada JJ. **2005**. TANMEI/EMB2757 encodes a WD repeat protein required for embryo development in Arabidopsis. *Plant Physiology* **139**: 163–173.

Yamaguchi S. **2008**. Gibberellin Metabolism and its Regulation. *Annual Review of Plant Biology* **59**: 225–251.

Yamamoto A, Kagaya Y, Toyoshima R, Kagaya M, Takeda S, Hattori T. **2009**. Arabidopsis NF-YB subunits LEC1 and LEC1-LIKE activate transcription by interacting with seed-specific ABRE-binding factors. *The Plant Journal* **58**: 843–856.

Yang P, Li X, Shipp MJ, Shockey JM, Cahoon EB. **2010**. Mining the bitter melon (*Momordica charantia*) seed transcriptome by 454 analysis of non-normalized and normalized cDNA populations for conjugated fatty acid metabolism-related genes. *BMC Plant Biology* **10**: 250.

Yang C, Bratzel F, Hohmann N, Koch M, Turck F, Calonje M. **2013**. VAL- and AtBMI1-Mediated H2Aub initiate the switch from embryonic to postgerminative growth in Arabidopsis. *Current Biology* **23**: 1324–1329.

Yobi A, Wone BWM, Xu W, Alexander DC, Guo L, Ryals JA, Oliver MJ, Cushman JC. **2012**. Comparative metabolic profiling between desiccation-sensitive and desiccation-tolerant species of Selaginella reveals insights into the resurrection trait. *The Plant Journal* **72**: 983–999.

Yoshida T, Fujita Y, Sayama H, Kidokoro S, Maruyama K, Mizoi J, Shinozaki K, Yamaguchi-Shinozaki K. **2010**. AREB1, AREB2, and ABF3 are master transcription factors that cooperatively regulate ABRE-dependent ABA signaling involved in drought stress tolerance and require ABA for full activation. *The Plant Journal* **61**: 672–685.

Yotsui I, Saruhashi M, Kawato T, Taji T, Hayashi T, Quatrano RS, Sakata Y. **2013**. ABSCISIC ACID INSENSITIVE3 regulates abscisic acid-responsive gene expression with the nuclear factor Y complex through the ACTT-core element in Physcomitrella patens. *New Phytologist* **199**: 101–109.

Zeng Y, Kermode AR. **2004**. A gymnosperm *ABI3* gene functions in a severe abscisic acid-insensitive mutant of Arabidopsis (*abi3-6*) to restore the wild-type phenotype and demonstrates a strong synergistic effect with sugar in the inhibition of post-germinative growth. *Plant Molecular Biology* **56**: 731–746.

Zeng Y, Zhao T, Kermode AR. **2013**. A conifer ABI3-Interacting protein plays important roles during key transitions of the plant life cycle. *Plant Physiology* **161**: 179–195.

Zha H-G, Liu T, Zhou J-J, Sun H. **2013**. MS-desi, a desiccation-related protein in the floral nectar of the evergreen velvet bean (*Mucuna sempervirens* Hemsl): molecular identification and characterization. *Planta* **238**: 77–89.

Zhang X, Garreton V, Chua N-H. **2005**. The AIP2 E3 ligase acts as a novel negative regulator of ABA signaling by promoting ABI3 degradation. *Genes & Development* **19**: 1532–1543.

Zhang X, Clarenz O, Cokus S, Bernatavichute YV, Pellegrini M, Goodrich J, Jacobsen SE. **2007**. Whole-genome analysis of histone H3 lysine 27 trimethylation in Arabidopsis. *PLoS Biol* **5**: e129.

Zhang Y, Yang J, Lu S, Cai J, Guo Z. **2008**. Overexpressing *SgNCED1* in tobacco increases ABA level, antioxidant enzyme activities, and stress tolerance. *Journal of Plant Growth Regulation* **27**: 151–158.

Zhao T-Y, Corum III JW, Mullen J, Meeley RB, Helentjaris T, Martin D, Downie B. **2006**. An alkaline α-galactosidase transcript is present in maize seeds and cultured embryo cells, and accumulates during stress. *Seed Science Research* **16**: 107–121.

Zhou Y, Tan B, Luo M, Li Y, Liu C, Chen C, Yu C-W, Yang S, Dong S, Ruan J, *et al.* 2013. HISTONE DEACETYLASE19 interacts with HSL1 and participates in the repression of seed maturation genes in Arabidopsis seedlings. *The Plant Cell* **25**: 134–148.

Zhulidov PA, Bogdanova EA, Shcheglov AS, Vagner LL, Khaspekov GL, Kozhemyako VB, Matz MV, Meleshkevitch E, Moroz LL, Lukyanov SA, *et al.* 2004. Simple cDNA normalization using kamchatka crab duplex-specific nuclease. *Nucleic Acids Research* **32**: e37.

Zou M, Guan Y, Ren H, Zhang F, Chen F. **2007**. Characterization of alternative splicing products of bZIP transcription factors OsABI5. *Biochemical and Biophysical Research Communications* **360**: 307–313.

Annexes

Annexe 1. Séquences des amorces et températures d'hybridation utilisées pour le chapitre II.

Validation des mutants *Mtabi3* et clonage du gène *MtABI3* chez A17

Fonction	Nom amorce	Séquence 5' - 3'	Tm
RT-PCR + Tnt1 insertion	Mt-ABI3-F	GCAACAAGTTCAAGTGCAAG	60°C
RT-PCR	Mt-ABI3-R	TCATTTTTTTCTCTTTGGTGAAG	
screening mutant	Tnt-R	CACAGGTTCTGCTCGTTCACTG	60°C
	Tnt-F	ACAGTGCTACCTCCTCTGGATG	
PCR inverse	iPCRABI3_F	GATGAGACATAAACCGCCGT	60°C
	iPCRABI3_R	GAAAAACCAGCATGGGAAAA	
Amorces internes pour séquençage de contrôle	Mt-ABI3-seq1	CTCAGCAACAGCAACAAGAT	
	Mt-ABI3-seq2	TAACGCATATTCCAAACTCG	
	Mt-ABI3-seq3	TATTGATTTCCATAACCTCCAAAA	

Annexe 2. Contigs de *C. australe* homologues des gènes codant pour les protéines LEA identifiées chez *M. truncatula*.

L'identification des contigs est basée sur les resultants de blastn contre les bases de données de *M. truncatula* MtGI11 and Mt3.5 (e-value < 10^{-6}). Les contigs surlignés en gris sont associés à plusieurs gènes de *M. truncatula* ; ND = pas d'homologue identifié dans le transcriptome de *C. australe*.

Famille (classification PFAM)	Nom de la protéine	Base de données	*M. truncatula* ID	Contigs *C. australe*	E-value	Longueur alignement	% identité
dehydrin (PF00257)	DHN3	MtGI11	TC175037	Ca_11990	1,E-11	989	52,4
				Ca_12667	6,E-07	952	48,3
	PM12	Mt3.5	Medtr3g117190	Ca_10586	3,E-14	800	55,4
				Ca_5668	6,E-14	676	53,8
	dehydrin-cognate	Mt3.5	Medtr3g117290	Ca_9276	1,E-11	630	54,7
	BudCar5	Mt3.5	Medtr7g086340	Ca_16167	6,E-09	423	63,9
				Ca_21263	2,E-08	354	61,3
				Ca_31427	4,E-09	359	69
				Ca_41650	1,E-09	284	62,2
	SLTI66	MtGI11	TC182121	ND	ND	ND	ND
	RAB21	Mt3.5	Medtr8g106140	Ca_20634	2,E-14	279	62,5
LEA_5 (PF00477)	EM6	Mt3.5	Medtr4g016960	Ca_14307	2,E-83	305	83,1
	EM1	MtGI11	AJ498523	Ca_2036	3,E-41	464	71,2
LEA_4 (PF02987)	SBP65	Mt3.5	Medtr4g079690	Ca_7340	7,E-07	961	56,1
		MtGI11	TC185888	Ca_1960	1,E-08	1120	47,3
	PM10	MtGI11	TC174929	Ca_3035	2,E-11	1747	54,4
	PM18	MtGI11	TC183861	Ca_330	6,E-15	1206	67,7
				Ca_8462	8,E-22	502	60,3
	Maturation polypeptide2	Mt3.5	Medtr1g061730	Ca_330	8,E-11	1047	62,9
				Ca_8462	2,E-15	668	71
	LEAM	Mt3.5	Medtr2g014040	Ca_7604	5,E-22	1128	55,4
	CapLEA.I	MtGI11	TC175990	Ca_42722	3,E-12	789	62,5
				Ca_8841	3,E-20	275	66,8
	CapLEA.II	MtGI11	TC194635	Ca_8841	7,E-12	821	61,6
	CapLEA.III	MtGI11	TC178278	Ca_8841	1,E-10	739	62,9
LEA_1 (PF03760)	PM1	Mt3.5	Medtr7g093170	Ca_8304	1,E-08	557	57,4
	D-113.I	Mt3.5	Medtr7g093140	Ca_8304	1,E-08	328	62,5
	D113.II	Mt3.5	Medtr7g093160	Ca_8304	4,E-08	255	59,4
	PM29	MtGI11	TC183357	Ca_6614	3,E-38	744	55,1
SMP (PF04927)	D-34.I	Mt3.5	Medtr1g072090	ND	ND	ND	ND
	PM25	MtGI11	TC174777	Ca_6007	2,E-61	1030	69
	D-34.II	MtGI11	TC183570	Ca_25629	5,E-34	386	80,1
	D-34.III	Mt3.5	Medtr2g076230	Ca_23377	3,E-10	581	47,2
	ECP31	Mt3.5	Medtr3g034660	Ca_11459	6,E-13	806	69,3
	D-34.IV	MtGI11	TC196118	ND	ND	ND	ND
	LEA14-like c	MtGI11	BQ124186	Ca_11619	6,E-62	801	66,8
				Ca_15389	9,E-13	583	64,8

LEA_2 (PF03168)				Ca_21839	4,E-47	540	71,7
	LEA14	Mt3.5	Medtr7g079180	Ca_11619	6,E-22	459	73
				Ca_15389	5,E-11	495	64
				Ca_21839	6,E-18	456	74,3
	LEA-like a	Mt3.5	Medtr5g088850	Ca_13037	1,E-121	819	81,5
				Ca_13771	1,E-75	360	83,8
				Ca_6845	3,E-179	960	83,6
	LEA-like b	MtGI11	TC174481	Ca_13037	1,E-88	376	85,6
				Ca_13771	1,E-87	838	58,9
				Ca_6845	6,E-110	805	77,8
LEA_3 (PF03242)	LEA5	Mt3.5	Medtr4g120040	ND	ND	ND	ND
	LEA-like protein	MtGI11	NP7252631	ND	ND	ND	ND
	ARG-2	MtGI11	TC176432	Ca_46429	6,E-10	221	62,3
				Ca_15390	4,E-12	788	56,2
				Ca_16897	3,E-08	713	56,9
	ARG-2-like	MtGI11	TC187764	Ca_15390	1,E-14	723	55,1
				Ca_16897	1,E-10	635	58,7
				Ca_42980	1,E-11	272	58,8

Annexe 3. Description des spots identifiés dans le protéome stable à la chaleur de cotylédons de *C. australe*.
Les spots associés à des polypeptides de protéines LEA sont indiqués en gras. MW match = poids moléculaire de la protéine identifiée ; pI match = point isolélectrique de la protéine identifiée ; MW obs = poids moléculaire du spot visible sur le gel ; pI obs = point isolélectrique du spot visible sur le gel.

Spot	Match ID	Description du match	Nombre de peptides	MW match	pI match	MW obs	pI obs
1	**O04132**	**SRC1 (soybean gene regulated by cold) (BudCar)**	2	12	7,1	17	6
2	B2D2G4	small heat-shock protein	3	25	8,5	17	5
3	**Q70EM0**	**LEA D-11 (DHN3)**	2	24	6	25	6
4	**Q8GUS4**	**Dehydrin (DHN3)**	2	20	5,8	26	6
5	**TC174308**	**Dehydrin-cognate (DHN cogn)**	5	24	5,5	35	4
6	O04237	Transcription factor	5	40	6,4	45	6,5
7	Q2TUW1	Abscisic stress ripening-like protein	5	25	5,5	40	5
8	Q2TUW1	Abscisic stress ripening-like protein	1	25	5,5	40	5
9	**Q39871**	**Late embryogenesis abundant protein (MP2)**	4	51	6,3	65	5
10	B1Q2X4	Protein disulfide isomerase (PDIL-1)	15	59	5,1	65	4
13	Q9FQ13	Cystatin-like protein	4	13	9,2	<15	9
14	no match						
16	Q5PSJ1	Glutaredoxin (grxC1)	2	13	8,8	<15	8
17	no match						
18	**O49817**	**CAPLEA**	6	16	9,6	15,5	8
19	**O49817**	**CAPLEA**	4	16	9,6	15,5	7,5
	Q41649	Peptidyl-prolyl cis-trans isomerase	4	16	6,7	15,5	7,5
22	Q93WK3	Bowman-Birk inhibitor	2	11	6,9	<15	7,5
23	Q8RU22	Bowman-Birk type proteinase isoinhibitor C (WSTI-C)	1	12	5,2	<15	6,5
24	O65639	Cold shock protein 1	1	31	8,2	<15	6,5
25	O65639	Cold shock protein 1	1	31	8,2	17	7,5
26	Q41231	Pathogen and wound inducible antifungal protein CBP20	3	22	7,9	15,5	6,5
27	Q9M4Z9	Pathogenesis-related protein 4 (PR-4)	5	14	5,6	15,5	7
28	P02877	Pro-hevein (HEV1)	3	22	5,6	15,5	7
29	P29063	Pathogenesis-related protein PR-4B	1	17		<15	6,5
	A9NR43	Putative uncharacterized protein	1	16		<15	6,5
	Q41230	CBP20 Fragment	3	22	7,9	<15	6,5
31	B9SP82	Major latex protein putative	1	14	6,7	<15	7
	A4RXS9	Predicted protein	1	25	5,8	<15	7
	A4Q7L3	Putative uncharacterized protein	1	19	5,1	<15	7
	A9SF58	Predicted protein	1	90		<15	7
	Q9SWS4	Ripening-related protein	1	18	5,4	<15	7
32	B9SP82	Major latex protein putative	1	14	6,7	<15	7
33	Q9Z2B5	Eukaryotic translation initiation factor alpha kinase	2	123	5,1	16	6,5
	P09761	Wound-induced protein WIN1	1	21	8,6	16	6,5
35	P93501	Ubiquitin-like protein	4	45	8,2	<15	5
36	Q6S3E6	Polyubiquitin (Ubq1)	8			<15	5,5
37	**P93510**	**Em protein**	2	11	5,8	<15	5,5
38	Q9SQL5	Superoxide dismutase [Cu-Zn] (SOD1)	2	15	5,3	<15	5,5

39	Q40540	Pit2 protein	3	10	4,9	15	5,5
40	A9PB52	Endoribonuclease L-PSP domain containing protein	4	22	8,8	<15	5
41	no match						
42	Q8L6E7	Double-headed Trypsin inhibitor	1	11	6,4	<15	5
44	no match						
45	Q39840	Cysteine proteinase inhibitor	7	10	5,7	<15	4,5
46	Q06445	Cysteine proteinase inhibitor	1	11	6,1	<15	5
48	Q6F3C5	ML domain protein, putative, expressed	3	17	5,7	<15	4
49	A0A9R6	Histone H4	1	11,5	<15	4	1
50	Q6F3C5	ML domain protein, putative, expressed	2	17	5,7	<15	4,5
51	**P93510**	**Em protein**	2	11	5,8	<15	5
	Q94JQ4	Endoribonuclease L-PSP family protein	3	20	8,5	<15	5
52	no match						
54	O65198	Superoxide dismutase [Cu-Zn], chloroplastic	4	21	6	15	4,5
	Q6YNS1	Putative glycine rich RNA binding protein	3	17	7,8	15	4,5
55	Q8RW11	Putative glycine rich protein	4	17	5	15	4,5
	Q9ZST5	PII protein	3	0	0	15	4,5
56	B4UW51	Class II small heat shock protein Le-HSP17.6 (Fragment)	1	15		15	5
57	Q71F44	Eukaryotic translation initiation factor 5A isoform VIII	3	17	5,5	15,5	5
58	Q71F44	Eukaryotic translation initiation factor 5A isoform VIII	7	17	5,5	15,5	5,5
61	TC177423	Small heat shock protein	6	26	7,7	16	4,5
62	TC177423	Small heat shock protein	4	26	7,7	16	5
63	B9IGD8	Phosphoethanolamine binding protein	5	19	5,6	16	4,5
64	no match						
65	Q5WM51	Trypsin protein inhibitor 3	1	23	5	17	4
66	Q0GPA4	Cystatin (CPI)	6	28	5,6	19	6,5
67	Q84QX4	Superoxide dismutase (MnSOD)	11	0	0	17	6
68	B9RR63	GroES chaperonin, putative	6	27	8,9	18	5,5
69	A5BR41	Superoxide dismutase	2	25	6,8	20	5,5
70	Q41401	Chitinase homologue	12	36	5,2	22	6
71	Q39529	Agglutinin II	6	32	6,4	22	5
72	Q9ZP90	Ferritin (FER)	5	28	5,8	18	4,5
73	Q0PRR6	Calmodulin	5	17	3,9	15	3,5
74	**Q9SPJ6**	**Maturation protein pPM32 (LEAm)**	2	19	5,3	25	4,5
75	**B9SYF1**	**SBP65**	4	61	8,5	65	7,5
76	no match						
77	no match						
78	P85959	Unknown protein 3 Fragment	1				
79	Q8GT65	Serpin-like protein	14	42	0	30	5
80	B6TQB3	Ran-binding protein 1	3	23	4,7	29	3,5
82	Q75QC2	Glutamate-rich protein (LjGluRP1)	12	17	3,9	31	3,5
	Q8LAU8	Dessication-related protein, putative	2	34	8,8	31	3,5
83	Q75QC1	Glutamate-rich protein (LjGluRP2)	6	16	4	35	3,5
84	Q6TKQ6	putative ankyrin-repeat protein	2	38	4,5	33	4
85	Q93YR3	FAM10 family protein	8	47	4,8	40	4,5
86	A0A762	Calreticulin-1 (Gm crt-1)	12	48	4,4	50	4

87	Q9LZJ2	Hypothetical protein	6	35	5,3	31	4,5
88	B7FIV8	Putative uncharacterized protein	7	35	5,4	35	4,5
	B4G218	Putative uncharacterized protein	1	38		35	4,5
	A2Q669	Desiccation related protein PCC13 62 putative	2	32	6,1	35	4,5
	A2Q670	Desiccation related protein PCC13 62 putative	1	33	6,8	35	4,5
	Q9C7Z0	Dessication-related protein, putative	1	33		35	4,5
	A2Q670	Desiccation-related protein PCC13-62, putative	1	33	6,8	35	4,5
	P38661	Probable protein disulfide-isomerase A6	5	41		35	4,5
	B9RFH6	Protein disulfide isomerase putative	1	40	5,6	35	4,5
	B3TLY7	Disulfide-isomerase-like protein	1	41		35	4,5
89	A9SR07	Predicted protein (Heavy-metal-associated domain)	3	9	4,6	20	4
90	A2PYH3	Alpha chain of nascent polypeptide associated complex	4	22	4,1	22	4
91	A2PYH3	Alpha chain of nascent polypeptide associated complex	3	22	4,1	23	4
92	**Q2Q4X9**	**Seed maturation protein PM25**	4	26	4,6	25	4
	P46266	14-3-3-like protein	6	29	4,7	25	4
93	B4G0G4	HSP20-like domain containing protein	3	34	5,1	31	5
94	P25974	Beta-conglycinin, beta chain (CG-4)	3	51	5,8	26	4,5
95	B9GXX4	Predicted protein	8	73	4,4	>65	4
96	D7SKY2	protein disulphide isomerase	17	64	4,4	>65	4
97	A5BEE7	HSP70	20	74	5,4	>65	4
98	Q41401	Chitinase homologue	11	36	5,2	25	5,5
99	Q9ZW79	Clathrin interactor EPSIN 2	1	31	9,3	25	6
101	Q8FU38	Putative molybdopterin biosynthesis protein	1	37	5	40	6
102	A5B6Y2	Putative uncharacterized protein	5	52	4,9	30	4,5
	A7NU11	Chromosome chr18 scaffold 1 whole genome shotgun	2	28	4,7	30	4,5
	Q8LAU8	Dessication related protein putative	1	34	8,9	30	4,5
	B9R777	Cysteine protease putative	1	52	5,2	30	4,5
	B9SGK4	Cysteine protease putative	1	38	9,5	30	4,5
	B9IF23	Predicted protein	2	40	6,7	30	4,5
103	B6TQB3	Ran-binding protein 1	5	23	4,7	30	4,5
108	no match						
110	Q70I32	Transcription factor homolog BTF3-like protein	4	18	5,5	22	#N/A

Annexe 4. Intensité normalisée des spots associés aux LEA dans les protéomes stables à la chaleur de graines matures de *M. truncatula* et *C. australe* (moyenne de trois gels avec 150 µg de protéines de chaque espèce).

Identifiant SSP de chaque spot	Espèces	Numéro de spot sur gel référence *C. australe* (figure 2)	Numéro de spot sur gel référence *M. truncatula* (Chatelain *et al.*, 2012) (**Boudet *et al.*, 2006**)	Protéine LEA associée aux spots (rouge = *C. australe* ; bleu = *M. truncatula*)	Famille de LEA selon la classification PFAM	Intensité normalisée du spot	Ecart type
5502	*C. australe*	4		CaDHN3		7309	1225
6402	*C. australe*	3		CaDHN3		112883	31027
7605	*M. truncatula*		11	MtDHN3		20490	5013
7608	*M. truncatula*		12	MtDHN3		57272	8406
1601	*C. australe*	5		CaDHN cogn	DHN	11705	892
5305	*C. australe*	1		CaBudCar	PF00257	72374	29055
5307	*C. australe*			CaBudCar		3053	730
6309	*C. australe*			CaBudCar		16583	6204
6202	*M. truncatula*		**32**	MtBudCar		597	418
6308	*M. truncatula*		**33**	MtBudCar		563	146
3104	*C. australe*	51		CaEM1		7159	5647
2204	*M. truncatula*		71	CaEM1	LEA_5	22026	4526
5101	*C. australe*	37		CaEM6	PF00477	4792	1863
4103	*M. truncatula*		44	MtEM6		19606	2303
8706	*C. australe*	75		CaSBP65		364	51
5806	*M. truncatula*		1	MtSBP65		2737	2284
6801	*M. truncatula*		2	MtSBP65		6530	3035
6803	*M. truncatula*		3	MtSBP65		18181	4226
6805	*M. truncatula*		4	MtSBP65		22961	15983
6808	*M. truncatula*		5	MtSBP65		7079	4250
7808	*M. truncatula*		56	MtSBP65		408	207
8804	*M. truncatula*		57	MtSBP65		466	276
8805	*M. truncatula*		60	MtSBP65		338	123
3710	*M. truncatula*			MtPM10		3765	105
4704	*M. truncatula*		52	MtPM10		9701	5350
5701	*M. truncatula*		8	MtPM10		18395	11331
5705	*M. truncatula*			MtPM10		2904	1590
4604	*M. truncatula*		14	MtPM18	LEA_4	7374	3207
5602	*M. truncatula*		15	MtPM18	PF02987	35061	15381
6601	*M. truncatula*		16	MtPM18		2472	1550
2704	*C. australe*	9		CaMP2		790	237
3707	*C. australe*			CaMP2		3374	1408
3711	*C. australe*			CaMP2		2010	541
6501	*M. truncatula*		55	CaMP2		1299	713
7501	*M. truncatula*		19	CaMP2		5663	2177
7503	*M. truncatula*		62	CaMP2		22318	15483
2406	*C. australe*	74		CaLEAM		534	323
1505	*M. truncatula*		21	CaLEAM		9689	494
8202	*C. australe*	18		CaCapLea		14972	4317
8301	*C. australe*	19		CaCapLea		6621	3243
8201	*M. truncatula*		38	MtCapLea		3221	1388
8203	*M. truncatula*		39	MtCapLea		27423	3284
9101	*M. truncatula*		43	MtD113	LEA_1	12060	11188
9302	*M. truncatula*		32	MtPM1	PF03760	12250	8871
1304	*M. truncatula*		65	MtD34.I		6200	798
1403	*M. truncatula*		63	MtD34.II	SMP	8578	2262
405	*C. australe*	92		CaPM25	PF04927	4141	1704
404	*M. truncatula*		22	CaPM25		16039	2440

Annexe 5. Intensité normalisée des spots associés aux protéines LEA dans les protéomes stables à la chaleur de graines matures de *M. truncatula* R108 (type sauvage), *Mtabi3-1* et *Mtabi3-2*.

Les valeurs présentées sont les moyennes de 4 gels pour R108 et 3 gels pour *Mtabi3-1* et *Mtabi3-2* avec 150 µg de protéines.

			R108		***Mtabi3-1***		***Mtabi3-2***				
Numéro de spot sur le gel référence de *M. truncatula* R108 (Chatelain thèse, 2011)	Identification	SSP	Intensité normalisée du spot	Ecart type	Intensité normalisée du spot	Ecart type	Intensité normalisée du spot	Ecart type	ratio WT/*abi3-1*	ratio WT/*abi3-2*	P-value difference WT-*abi3* (paired t-test Statgraphics)
106	SBP65	2202	23381	14973	2981	782	1318	430	0,13	0,06	1,9E-03
58	SBP65	8804	7890	2056	916	51	258	146	0,12	0,03	2,2E-06
96	SBP65	8601	6588	3601	1246	453	241	192	0,19	0,04	8,0E-03
59	SBP65	7801	1205	355	157	272	100	173	0,13	0,08	6,1E-05
50	SBP65	1601	824	457	300	311	0	0	0,36	0,00	1,4E-02
57	SBP65	7902	802	552	169	293	275	476	0,21	0,34	1,7E-01
56	SBP65	7901	372	254	120	208	16	29	0,32	0,04	1,2E-01
100	PM25	401	7889	1025	4185	2307	2117	500	0,53	0,27	3,5E-04
91	PM18	6202	11886	7891	895	425	251	317	0,08	0,02	1,5E-02
89	PM18	6203	7245	4319	648	456	449	383	0,09	0,06	2,2E-02
84	PM18	3601	4981	2268	1961	527	1361	390	0,39	0,27	1,3E-02
87	PM18	6303	3361	2006	365	556	258	442	0,11	0,08	1,1E-01
8	PM10	5804	9639	2976	2276	1912	1350	828	0,24	0,14	8,9E-03
52	PM10	4803	1169	1006	296	512	201	349	0,25	0,17	2,1E-01
94	PM10	3505	888	977	5154	2946	3531	2492	5,80	3,98	2,5E-01
32	PM1	9301	2020	2634	18045	15326	20108	11217	8,93	9,95	1,0E-01
105	MP2	8301	30831	11121	3045	723	1447	245	0,10	0,05	4,3E-04
19	MP2	7704	1143	138	1877	664	1179	106	1,64	1,03	2,5E-01
55	MP2	6603	1075	838	232	402	303	525	0,22	0,28	1,8E-01
99	LEAM	1503	4144	5616	13987	13753	6833	8001	3,38	1,65	7,6E-01
44	EM6	4001	21817	6784	0	0	0	0	0,00	0,00	1,4E-03
71	EM prot	2201	13844	5064	0	0	0	0	0,00	0,00	8,9E-05
93	DHN3	7702	17060	5610	46237	33034	19162	13043	2,71	1,12	3,7E-01
92	DHN3	7701	2807	821	10729	10647	6899	8093	3,82	2,46	6,1E-01
97	D34.I	501	5515	1259	1976	1921	1976	275	0,36	0,36	2,9E-04
101	D34.II	1401	304	608	2452	1756	896	1050	8,07	2,95	4,7E-01
43	D113.I	9102	30341	6446	17694	2888	9929	4056	0,58	0,33	3,0E-03
39	CapLea	9204	33137	13756	70186	8081	47395	9440	2,12	1,43	5,4E-02
107	CapLea	8303	26653	13025	26115	2461	4414	1040	0,98	0,17	2,2E-02
38	CapLea	7202	3662	1282	3408	5904	0	0	0,93	0,00	1,1E-02

Annexe 6. Contigs Ca associés aux gènes Mt codant pour les régulateurs et répresseurs majeurs de la maturation des graines.

Les différents gènes codant pour les régulateurs majeurs de la maturation ainsi que l'identifiant associé chez *M. truncatula* sont indiqués dans la première colonne. Les contigs Ca reliés à chaque identifiant Mt ainsi que la e-value associée sont indiquées dans les deux colonnes suivantes. Les colonnes CaG, CaY et CaB indiquent les intensités mesurées (en log) pour chaque sonde associée au contig Ca au cours du développement de la graine de *C. australe*. Le ratio entre graine mature et immature est indiqué dans la colonne CaG/CaB. Pour les gènes associés à plusieurs contigs Ca, la valeur moyenne d'intensité et de ratio (en rouge dans le tableau) a été utilisée pour la figure 3.8.

Gene | Mt ID	Ca probe_id	E value Ca Mt	CaG	CaY	CaB	CaG/CaB
LEC1 | Mt3.5.1s042061	Ca_26921	1,0E-14	5,85	5,87	3,75	-2,10
FUS3 | Medtr7g083700.1	Ca_4589	2,0E-47	5,18	5,46	1,67	-3,51
	Ca_15566	1,0E-86	6,29	6,46	1,81	-4,48
			5,74	5,96	1,74	-4,00
ABI3 | Mt3.5.1s031063	Ca_3733	6,0E-115	6,64	5,86	3,13	-3,51
ABI3-like	Ca_2432	7,0E-109	5,82	5,62	4,63	-1,19
ABI5 | Medtr7g104480.1	Ca_1893	5,0E-56	7,76	7,67	6,48	-1,28
bZIP53 | Medtr1g087040.1	Ca_31118	7,0E-07	6,48	6,74	5,74	-0,74
bZIP67 | Medtr2g086340.1	Ca_3355	3,0E-49	5,58	5,55	3,24	-2,34
bZIP12 (EEL) | Medtr3g010660.1	Ca_22908	2,0E-35	2,01	2,40	1,06	-0,95
PIL5 | Medtr7g099540.1	Ca_37012	2,0E-32	3,69	2,90	2,73	-0,96
	Ca_6536	2,0E-67	5,15	5,63	5,26	0,11
	Ca_9849	3,0E-89	5,57	6,02	4,50	-1,07
			4,80	4,85	4,16	-0,64
SOMNUS | Medtr4g078660.1	Ca_4368	1,0E-21	7,47	7,76	8,12	0,65
ABI4 | Medtr5g082950.1	Ca_4350	1,0E-08	7,85	8,34	5,59	-2,26
	Ca_5926	3,0E-08	7,53	8,07	6,70	-0,83
			7,69	8,21	6,14	-1,55
PKL | Medtr3g106210.1	Ca_2641	0	1,30	1,15	1,46	0,17
	Ca_17133	4,0E-176	1,22	1,25	1,46	0,24
	Ca_19802	2,0E-109	1,75	1,77	2,23	0,48
			1,42	1,39	1,71	0,29
SUA | Mt3.5.1s030573	Ca_1906	0	5,19	5,30	5,61	0,42
	Ca_6035	0	2,22	2,34	2,71	0,50
			3,70	3,82	4,16	0,46
AIP2 | Mt3.5.1s012165	Ca_7028	1,0E-103	5,15	5,36	5,60	0,44
VAL2 | Medtr5g075410.1	Ca_10504	1,0E-63	5,33	5,51	4,58	-0,76
	Ca_35689	5,0E-33	1,41	1,95	1,04	-0,37
	Ca_557	1,0E-85	6,16	6,38	4,26	-1,90
			4,30	4,61	3,29	-1,01
VAL2-like | Medtr4g028130.1	Ca_29282	4,0E-61	1,99	2,95	3,68	1,69
	Ca_35469	5,0E-76	4,59	4,99	5,30	0,72
	Ca_36931	1,0E-109	2,65	3,46	4,05	1,40
	Ca_45067	1,0E-63	2,67	3,36	4,09	1,43
	Ca_48089	5,0E-46	4,48	4,50	4,74	0,27
	Ca_9272	0	4,57	5,07	5,75	1,18
			3,49	4,05	4,60	1,11
VAL3 | Medtr2g011940.1	Ca_11607	3,0E-100	4,15	4,47	5,53	1,39
	Ca_1433	0	3,03	3,54	5,24	2,21
	Ca_19164	1,0E-94	2,96	3,54	4,20	1,25
	Ca_25466	2,0E-37	1,57	1,95	3,95	2,38
	Ca_42719	9,0E-30	1,54	2,07	4,53	2,99
	Ca_7519	0	1,36	1,73	2,66	1,30
			2,43	2,88	4,35	1,92

Annexe 7. Alignement des séquences protéiques d'AtABI3, CaABI3 et CaABI3-like.

Les séquences génomiques et les ADNc des homologues *ABI3* chez *C. australe* ont été clonés par PCR dégénérée. Les séquences protéiques intégrales ont été déduites de ces clonages pour CaABI3 et CaABI3-like puis alignées par le logiciel MultAlin (matrice Blosum 62) à celle disponible sur la base de données TAIR10 pour AtABI3. Les domaines conservés A1, B1, B2 et B3 ont été positionnés. Un cinquième domaine conservé a été ajouté en pointillé. Ce domaine n'est pas décrit dans la littérature excepté chez le pois où un deuxième domaine acide (A2) a été mis en évidence (Gagete *et al.*, 2009). Chez le mutant *abi3-7*, l'alanine située au niveau de la flèche orange (position 518 sur cet alignement) est substituée par une thréonine. Chez le mutant *abi3-1*, l'aspartate situé au niveau de la flèche bleue (position 646 sur cet alignement) est substitué par une asparagine (Bies-Ethève *et al.*, 1999).

```
            1       10        20        30        40        50        60        70        80        90       100       110       120       130
            |--------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------|
CaABI3like   MKGQVELQVGDLH-AGMVTETDPVGFGAMENADAMAINEREIWLHS-----------DQEDFLGVNDASMFYAEFPPLPDFPCMSSSSSSPSTPAPAKTMTCSSSSSS-----SSAASWAVLKSDVEE
    CaABI3   MKYEVELQAGDLHVAGMVTDTDPVGFGSMEDTDTLAVNEREMWLQN-----------DQDDLLVVNDASMFYAEFPPLPDFPCMSSSSSSSSTPAPVKTVTCSSSSTSS-----SSAASWAVLKSNAEE
    AtABI3   MKSLHVAANAGDLAEDCGILGGDADDTVLMDGIDEVG---REIWLDDHGGDNNHVHGHQDDDLIVHHDPSIFYGDLPTLPDFPCMSSSSSSSTSPAPVNAIVSSASSSSAASSSTSSAASWAILRSDGED
                                                                                                                                  A1
            131     140       150       160       170       180       190       200       210       220       230       240       250       260
            |--------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------|
CaABI3like   DAEQKNR-HYYYKQNDPVDACAPPAVLSSTASMEISQPTDPSHHGVVGDCMDVMETFGYMDLLESNHFFDPASIFQNEENPLQEFPQEQRPQQEEHAVVAQQQPPPCAQQQEPEEGEVIVCATTNNNGEE
    CaABI3   DNVVKNNCHHYGMQHDPPVNAPPAALSSSTAPMEISQPPDPILD-CIGDCMDVMDTFGYMELLESDDFFDPASIFQNEANPLAEFPQEHVPPQEEHAMVAQQQQQS-QPQQEPEVDQLVACTTSNSEEEI
    AtABI3   PTPNQNQ---YASGNCDDSSGALQSTASMEIPLDSSQGFGCGEGG--GDCIDMMETFGYMDLLDSNEFFDTSAIFSQDDDTQNPNLMDQTLERQEDQVVVPMMEN-----------------NSGGDMQM
                                                                                  A2
            261     270       280       290       300       310       320       330       340       350       360       370       380       390
            |--------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------|
CaABI3like   IQGEGGPDGEMSTVFLEWLKSNKDSVSANDLRSVKLKKATIECAAKRLGGGKEAMKQLLKLILEWVQTSHLQNRRRKEN--APSNLLSQFQDPSQNQNTNPGSNLNPSSFGPESNTCFNQSHWIPPPPYG
    CaABI3   IQGEGGADDEMSTVFLEWLKSNKDSVSANDLRSVKLKKATIECAARRLGGGKEAMKQLLKLILQWVQTIHLHNKRRKENDNTASNPPTQIQDPCQNQNTNTGSNLNPSSFGPELNTCLNQSPWISQP-YW
    AtABI3   MNSSLEQDDDLAAVFLEWLKNNKETVSAEDLRKVKIKKATIESAARRLGGGKEAMKQLLKLILEWVQTNHLQRRRTTTTTNLSYQQSFQQDPFQNPNPNNNNLIPPSD-----QTCFSPSTWVPPPP--
                                                                  B1
            391     400       410       420       430       440       450       460       470       480       490       500       510       520
            |--------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------|
CaABI3like   ADPAALVTAPPPFPQPMNIGYVGDPYTNGA-SNNINGTNN--HFPYITEYHMLESAHSWPPS-QFTVASHYNQLFGDNNLHTAGPPLAAAAATFTGYGN-QYPYQFF---HVPGDRLMRLGPSATKEARK
    CaABI3   TDQAALMAATPPFPQPV-LGYVGDPYRNGAASNNINGTNHNPYPPASTEYHLLESAPSWPPS-QFTVASHYNQSFGDSNLYPA--------VAFGGYGNTQYPYQFF---HGPGDRVVRLGPSATKEARK
    AtABI3   --QQQAFVSDPGF------GYMPAPN----------------YPPQPEFLPLLESPPSWPPPPQSGPMPHQQFPMPPTSQYNQ----FGDPTGFNGYNMNPYQYPYVPAGQMRDQRLLRLCSSATKEARK
                                                                                                                                  B2
            521     530       540       550       560       570       580       590       600       610       620       630       640       650
            |--------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------|
CaABI3like   KRMARQRRSLSHHRHHNHHGQQNQLNQNQTATDPLARLGSDNCTAVLAAHANPANWVHWRAMAGGAASVAPPMIPVESPAGQPVMDQTAMQAQNHHQGRVSSDRRQVWKPEKNLRFLLQKVLKQSDVGNL
    CaABI3   KRMARQRRSLSHHRHHSQH-------QNQGGADPQARVGGGDNCIATAAHANPANWVYWHTMAGGVASMAP-VVPAKPPAGQPVMDRTAMQTQNYHQGRVSSDRRQGWKPEKNLRFLLQKVLKQSDVGNL
    AtABI3   KRMARQRRFLSHHHRHNNNNNNNNNNNQQNQTQIGETCAAVAPQLNPVATTATGGTWMYWPNVPA-----VPPQLPPVMETQLPTMDRAGSASAMPRQ-QVVPDRRQGWKPEKNLRFLLQKVLKQSDVGNL
                B2                                                                                                                  B3
            651     660       670       680       690       700       710       720       730       740       750       760       770       780
            |--------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------|
CaABI3like   GRIVLPKKEAETHLPELEARDGISIAMEDIGTSRVWNMRYRYWPNNKSRMYLLENTGDFVRANGLQEGDFIVIYSDVKCGKFMIRGVKVRQQ-GAKPETKKAGKSQKNQHGNNAATTAGTVANNGTPSSL
    CaABI3   GRIVLPKKEAETHLPELEARDGISITMEDIGTSRVWNMRYRYWPNNKSRMYLLENTGDFVRANGLQEGDFIVIYSDVKCGKQMIRGVKVRQP-GAKPETKKAGKSQKNQHVNNAATTPSAGAGAGTINCT
    AtABI3   GRIVLPKKEAETHLPELEARDGISLAMEDIGTSRVWNMRYRFWPNNKSRMYLLENTGDFVKTNGLQEGDFIVIYSDVKCGKYLIRGVKVRQPSGQKPEAPPSSAATKRQNKSQRNINNNSPSANVVVAS-
                                                                                                                                  B3
            781     790   796
            |--------+-----|
CaABI3like   P---QQKNEKAAK
    CaABI3   PCSPKQKNEKAVKYHI
    AtABI3   PTSQTVK
```

Annexe 8. Séquences des amorces et températures d'hybridation utilisées pour le chapitre IV.

Tm = température d'hybridation utilisée pour l'amplification PCR. Les nucléotides ajoutés pour les recombinaisons TOPO et Gateway sont indiqués en rouge.

Clonage du gène et du promoteur *MtABI3* chez R108

Fonction	Nom amorce	Séquence 5' - 3'	Tm
Clonage séquence codante pour sur-expression racines	MtABI3atg-TOPO-F	CCACATGGAGTGTGGTTTAGATCTGC	60°C
	MtABI3stop1-R	TCATTTTTTTCTCTTTGGTGAAG	
Amorces internes pour séquençage de contrôle	MtABI3int1-F	GTATGGTTTTCTTGGAGTGG	
	MtABI3int1-R	TAGTCCCAATGTATCATCATCAT	
	MtABI3int2-F	GCAACAAGTTCAAGTGCAAG	
	MtABI3int3-F	CTTTGTGGAGCAGAAGACCTT	
	MtABI3int4-F	CAGTACCATATTTCTTGCACCA	
	MtABI3int2-R	CGATTCAATTACGCCGGGAA	
	MtABI3int3-R	CCACTCCAAGAAAACCATAC	

Clonage du gène *CaABI3*

Fonction	Nom amorce	Séquence 5' - 3'	Tm
PCR dégénérée	ABI3degF	ACNAARGARGCNMGVMARAA	60°C
	ABI3degR	ACYTTHACBCCYCKKATMA	
PCR inverse	CaABI3_inv_fw	CCCCAGATTCAATTTCTGTTTT	60°C
	CaABI3_inv_rv	CAGTTCTGGCAGATGGGTTT	
	CaABI3_inv_fw2	ATCCTTTCTATCTGCTATTTTCTCTC	60°C
	CaABI3_inv_rv2	ATTTGGGTGGGGGGATTC	
Clonage séquence codante pour sur-expression racines	CaABI31atg1-TOPO-F	CACCATGAAGTATGAAGTGGAATTGCAAGC	67°C
	CaABI31stop1-R	TTATATATGATATTTTACTGCTTTTTCATTTTTCTGC	
Clonage séquence codante pour complémentation	CaABI31atg1-attB5-F	GGGGACAACTTTGTATACAAAAGTTGATGAAGTATGAAGTGGAATTGCAAGC	67°C
	CaABI31stop1-attB2-R	GGGGACCACTTTGTACAAGAAAGCTGGGTTTATATATGATATTTTACTGCTTTTTCATTTTTCTGC	
Amorces internes pour séquençage de contrôle	CaABI31int1_F	GAGAAATCCCAATTTGATTC	
	CaABI31int2_F	GGAAGAGATCATCCAAGGAGAG	
	CaABI31int1_R	TTGCCTCCAGTTCTGGCAGA	
	CaABI31int2_R	CGGGTACTGAGTGTTCCCAT	

Clonage du gène *CaABI3-like*

Fonction	Nom amorce	Séquence 5' - 3'	Tm
PCR inverse	DL-CaABI3Linv5'-3	CTATGGTGGCCTTCTTGAGC	60°C
	DL-CaABI3Linv3'-BamHI	CCTTCGAATCTTCTGAGCCA	
	DL-CaABI3Linv5'-1	TGAAAGTCGCTGCCGCT	60°C
	DL-CaABI3Linv3'-1	GACTCTATTTTTTCGTCTTCTGAGT	
Clonage séquence codante pour sur-expression racines	DL-CaABI3L-F	CACCATGAAGGGTCAAGTGGAACTGC	64°C
	DL-CaABI3L-R	TTATTTTGCTGCTTTTTCATTTTTCTG	

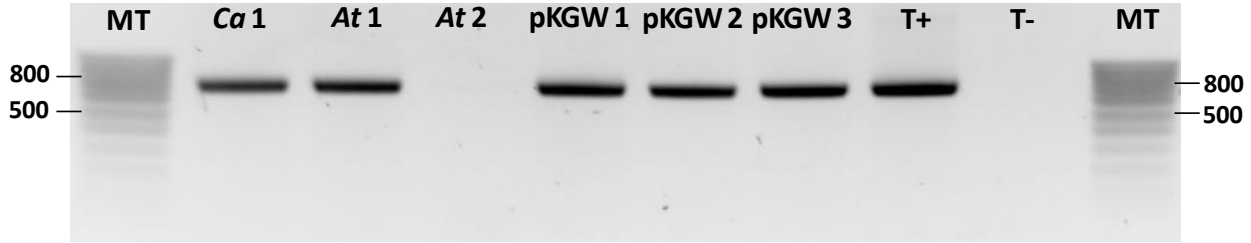
Clonage du gène *AtABI3*

Fonction	Nom amorce	Séquence 5' - 3'	Tm
Clonage promoteur pour complémentation	AtABI3pr2-attB1-F	GGGGACAAGTTTGTACAAAAAAGCAGGCTACATGCCATCTCGACGTACA	67°C
	AtABI3pr1-attB5r-R	GGGGACAACTTTTGTATACAAAGTTGCGTTGAAGTGGAAATGAAACAATA	
Clonage séquence codante pour complémentation	AtABI3atg1-attB5-F	GGGGACAACTTTGTATACAAAAGTTGATGAAAAGCTTGCATGTGG	64°C
	AtABI3stop1-attB2-R	GGGGACCACTTTGTACAAGAAAGCTGGGTTCATTTAACAGTTTGAGAAGTTGGTG	
Amorces internes pour séquençage de contrôle	AtABI3int1-F	GAAGATGAGGGGAGATTTACTATC	
	AtABI3int2-F	GGTTTATATGTATATGTATAGCCAGATAG	
	AtABI3int3-F	TACTTGAATCTCCACCGTCA	
	AtABI3int4-F	AGGTATCACATGATCTGCTCAA	
	AtABI3int5-F	TGTAGCCTTTGCTTTCCATT	
	AtABI3int6-F	ACATGCCTGCTCCAAACTAT	
	AtABI3int7-F	CAACAAGCTTTTGTCTCGGA	
	AtABI3int1-R	AATGGAAAGCAAAGGCTACA	
	AtABI3int2-R	ATGTGGCATGGGACCAGACT	
	AtABI3int3-R	GGTTTGGAGTCGGATCTTCT	

Vérification des transformants complémentés par *CaABI3*, *AtABI3* et pKGW

Fonction	Nom amorce	Séquence 5' - 3'	Tm
Amplification région codant pour résistance Kanamycine	Pk7WG2D_Kan_F	CGTCAAGAAGGCGATAGAAG	61°C
	Pk7WG2D_Kan_R	GATCGAGGCATGATTGAAC	

Gel de contrôle : taille attendue 793 pb
Ca = lignée *CaABI3*/*abi3-5* ; At = lignée *AtABI3*/*abi3-5* ; pKGW = lignée pKGW/*abi3-5* ; T^+ = plasmide pKGW purifié ; T^- = PCR sans matrice ; MT = marqueur de taille 100 bp



Thèse de Doctorat

Julien DELAHAIE

Comparaison moléculaire des graines orthodoxes de *Medicago truncatula* et récalcitrantes de *Castanospermum australe* : une nouvelle approche pour comprendre l'acquisition de la tolérance à la dessiccation

Molecular comparison of orthodox *Medicago truncatula* seeds and recalcitrant *Castanospermum australe* seeds: a new model to unravel desiccation tolerance acquisition

Résumé

La tolérance à la dessiccation (TD) est définie comme l'aptitude à survivre à l'état sec et à reprendre un métabolisme normal après le retour à des conditions hydriques favorables. Contrairement aux graines orthodoxes, qui acquièrent la TD au cours de leur maturation, les graines récalcitrantes ne survivent pas à la dessiccation. L'analyse comparative du développement de ces deux types de graines constitue donc un modèle intéressant pour mettre en évidence des mécanismes spécifiquement impliqués dans la TD. Par des approches protéomique et transcriptomique, ce travail a permis de caractériser le développement de graines récalcitrantes de *Castanospermum australe* et de le comparer à celui de graines orthodoxes de *Medicago truncatula*, espèce de la même sous-famille des Fabacées. Nos résultats montrent que certaines protéines de type LEA (Late Embryogenesis Abundant) sont absentes ou faiblement accumulées dans les graines matures de *C. australe* comparé à celles de *M. truncatula*. Le profil des LEA ressemble à celui du mutant *Mtabi3*, déficient pour le régulateur majeur de la maturation des graines orthodoxes ABI3 (ABscissic acid Insensitive 3). L'analyse transcriptomique révèle une forte répression de *CaABI3* et de ses gènes cibles en fin de développement chez les graines de *C. australe*, alors qu'ils restent fortement exprimés tout au long de la maturation des graines de *M. truncatula*. Deux gènes codant pour ABI3 ont été clonés chez *C. australe* : *CaABI3* et *CaABI3-like*. Par sur-expression ectopique dans des racines de *M. truncatula*, seul CaABI3-like est en mesure d'activer les mêmes cibles que MtABI3. De plus, *CaABI3* ne complémente pas le mutant *abi3-5* d'*Arabidopsis thaliana*. Cette analyse renforce l'importance d'ABI3 pour expliquer la sensibilité à la TD des graines récalcitrantes. Elle révèle par ailleurs que la graine mature de *C. australe* présente certaines caractéristiques d'une graine prête à germer tout en exprimant de nombreux gènes de réponse au stress.

Mots clés
ABI3 - *Castanospermum australe* - LEA - *Medicago truncatula* - protéomique - graine récalcitrante - tolérance à la dessiccation - transcriptomique

Abstract

Desiccation tolerance (DT) is defined as the ability to survive in the dry state and resume metabolic activity upon imbibition. Unlike orthodox seeds, that acquire DT during their maturation, recalcitrant seeds do not survive desiccation. A comparative analysis of orthodox and recalcitrant seeds development constitutes an interesting model to highlight mechanisms that are involved in DT. Using proteomic and transcriptomic analyses, the development of the recalcitrant *Castanospermum australe* seeds was characterized and compared to development of orthodox *Medicago truncatula* seeds, species from the same sub-family of Fabaceae. Our results show that most LEA (Late Embryogenesis Abundant) proteins are absent or faintly accumulated in mature seeds of *C. australe* compared to *M. truncatula*. The LEA protein profile is similar to seeds of the *Mtabi3* mutant that is deficient in the expression of the major regulator of orthodox seed maturation ABI3 (ABscissic acid Insensitive 3). Transcriptome analysis reveals a strong repression of *CaABI3* and its putative targets at the end of *C. australe* seed development while the transcript levels of these target genes remain high until the end of maturation of *M. truncatula* seeds. Two genes coding for ABI3 were cloned in *C. australe*: *CaABI3* and *CaABI3-like*. Ectopic expression in *M. truncatula* roots demonstrated that only *CaABI3-like* is able to activate the same targets as *MtABI3*. Moreover, *CaABI3* does not complement the *abi3-5* mutant of *Arabidopsis thaliana*. This analysis strengthens the implication of ABI3 in the desiccation sensitivity of recalcitrant seeds. In addition, mature *C. australe* seeds display certain characteristics of a seed that is preparing for germination and express many stress response genes.

Key Words
ABI3 - *Castanospermum australe* - desiccation tolerance - LEA - *Medicago truncatula* - proteomics - recalcitrant seed - transcriptomics

L'Université Nantes Angers Le Mans