



HAL
open science

Recherche de déterminants génétiques de la date de floraison chez la Légumineuse modèle, *Medicago truncatula*

Jean-Baptiste Pierre

► **To cite this version:**

Jean-Baptiste Pierre. Recherche de déterminants génétiques de la date de floraison chez la Légumineuse modèle, *Medicago truncatula*. Sciences agricoles. Agrocampus - Ecole nationale supérieure d'agronomie de rennes, 2008. Français. NNT : . tel-00653807

HAL Id: tel-00653807

<https://theses.hal.science/tel-00653807>

Submitted on 20 Dec 2011

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.



N° d'ordre :
N° de série :



THÈSE

Présentée devant

L'ÉCOLE NATIONALE SUPERIEURE AGRONOMIQUE DE RENNES

POUR OBTENIR LE TITRE DE DOCTEUR DE L'ENSAR

Mention : Biologie et Agronomie

par

Jean-Baptiste PIERRE

Recherche de déterminants génétiques de la date de floraison chez la Légumineuse modèle, *Medicago truncatula*

*Préparée à l'Unité de Génétique et d'Amélioration des Plantes Fourragères
INRA, 86600 LUSIGNAN
Bourse cofinancée par l'INRA et la région Poitou-Charentes*

Soutenue le 28 janvier 2008 devant le jury :

Mr. Grégoire THOMAS	Président du jury
Mme Evelyne COSTES	Rapporteur
Mr. Pascal RATET	Rapporteur
Mr. Philippe LONNET	Examinateur
Mr. Richard THOMPSON	Examinateur
Mr. Christian HUYGHE	Directeur de thèse
Mme Bernadette JULIER-KOUBAÏTI	Encadrant



N° d'ordre :
N° de série :



THÈSE

Présentée devant

L'ÉCOLE NATIONALE SUPERIEURE AGRONOMIQUE DE RENNES

POUR OBTENIR LE TITRE DE DOCTEUR DE L'ENSAR

Mention : Biologie et Agronomie

par

Jean-Baptiste PIERRE

Recherche de déterminants génétiques de la date de floraison chez la Légumineuse modèle, *Medicago truncatula*

*Préparée à l'Unité de Génétique et d'Amélioration des Plantes Fourragères
INRA, 86600 LUSIGNAN
Bourse cofinancée par l'INRA et la région Poitou-Charentes*

Soutenue le 28 janvier 2008 devant le jury :

Mr. Grégoire THOMAS	Président du jury
Mme Evelyne COSTES	Rapporteur
Mr. Pascal RATET	Rapporteur
Mr. Philippe LONNET	Examinateur
Mr. Richard THOMPSON	Examinateur
Mr. Christian HUYGHE	Directeur de thèse
Mme Bernadette JULIER-KOUBAÏTI	Encadrant

REMERCIEMENTS

Au-delà d'un véritable défi scientifique, la réalisation d'une thèse s'inscrit comme une réelle aventure humaine. Ce travail a été réalisé grâce au concours de nombreuses personnes que je tiens à remercier dans ces quelques lignes.

Mes remerciements s'adressent en premier lieu aux financeurs de ces travaux : le Département de Génétique et d'Amélioration des Plantes (DGAP) de l'INRA, et le Conseil Régional du Poitou-Charentes.

J'aimerais aussi remercier Christian Huyghe, mon directeur de thèse, d'une part pour m'avoir offert l'opportunité de réaliser ce projet de recherche au sein de l'unité de génétique et d'amélioration des plantes fourragères, et d'autre part pour ses relectures attentives de mon manuscrit. Sa disponibilité et son sens critique ne se sont jamais démentis et ont permis à cette thèse d'évoluer dans la compréhension de son sujet.

Je tiens à exprimer tous mes remerciements à Bernadette Julier-Koubaiti, ma co-directrice de thèse. Son expertise dans le domaine de la génétique quantitative, sa rigueur scientifique ainsi que sa grande disponibilité tout au long de cette thèse (même durant son congé maternité et ses week-ends) et particulièrement lors de la rédaction, ont grandement contribué à l'aboutissement de cette thèse. C'est toujours avec enthousiasme qu'elle m'a accueilli dans son bureau pour m'aider à progresser dans la réflexion scientifique. Je la remercie également de m'avoir laissé une grande autonomie dans la gestion de mon projet de thèse tout en restant très présente et pour m'avoir fait participer au congrès international de Tunis en 2007.

Je remercie les membres du comité de thèse : Thierry Huguet, Catherine Rameau, Joëlle Ronfort et Maria Manzanares pour leurs conseils avisés et pour l'intérêt qu'ils ont porté à ce travail. J'adresse un remerciement particulier à Evelyne Costes et Pascal Ratet pour avoir accepté d'être rapporteurs de cette thèse, et à Grégoire Thomas de présider le jury.

Je remercie également Gilles Boutet et Marie-Reine Perretant de la plate-forme de génotypage de l'INRA de Clermont-Ferrand.

Je tiens à remercier Anne-Marie Chèvre et l'ensemble des co-encadrants du groupe des doctorants et post-doctorants du DGAP de l'INRA. Les Journées Jeunes Chercheurs (JJC) ont toujours été des moments de formations et d'échanges très appréciés.

Mes remerciements vont aussi sincèrement à Philippe Barre, Isabelle Litrico, Yves Barrière, Sandrine Flajoulot, Jean-Paul Sampoux et Marc Ghesquière pour leurs précieuses connaissances en génétique, et à Ela Frak, Jean-Louis Durand et Claude Varlet-Grancher pour leurs avis éclairés d'écophysiologistes. Leurs disponibilités et leurs appréciations scientifiques et techniques ont toujours été pertinentes.

J'adresse ici toute ma gratitude à l'équipe « Bases génétiques », aux serristes et à tous les membres de la plate-forme de biologie moléculaire : Joël Jousse, Jean-François Bourcier (le tennis de table est leurs points forts), Frank Gelin, René Minault, Philippe Cormenier (spécialité : remise des vertèbres), Denise Cadier, Céline Talon et Chrystel Gibelin. Le reste de l'équipe ne m'en voudra certainement pas si je remercie tout particulièrement Françoise Durand qui s'est énormément investie dans ce travail. Elle a été les petites mains habiles du laboratoire et a su s'adapter aux difficultés des mises au point et du peu de routine dans les manipulations, tout en conservant un grand sens de l'organisation et des relations humaines excellentes. Ces deux ans passés à ses côtés ont été extraordinaires. Je profite de l'occasion qui m'est donnée pour remercier Matthieu Bogard qui a participé activement à ce travail de thèse durant son stage de master recherche. Il a réalisé un travail d'envergure en s'imprégnant d'un sujet complexe. J'ai apprécié de travailler avec lui pour son sérieux et son investissement.

Merci à Eric Marchoux pour sa précieuse aide et ses compétences dans la recherche d'articles et d'ouvrages écrits depuis la création du monde ou même en langues étrangères.

Je remercie les amis et collègues que j'ai connu pendant cette thèse et bien sûr Jérôme, Cédric, Cécile, Alban(e), ainsi que l'ensemble des stagiaires, CDD et permanents que j'ai pu côtoyer.

Enfin, ce travail, je le dois à mes parents qui m'ont toujours soutenu pendant ces trois ans, mais aussi tout au long de mes études ainsi qu'à Mélanie qui a dû me supporter, et ceci malgré la distance.

LISTE DES ABRÉVIATIONS

Sigles des organismes

CBBC : Center of biotechnology of Borj Cedria
CNRS : Centre national de la recherche scientifique
ENSAT : Ecole nationale supérieure d'agronomie de Toulouse
IMGAG : International medicago genome annotation group
INRA : Institut national de la recherche agronomique
NCBI : National center for biotechnology information
NSF : National science foundation
SARDI : South australian research and development institut
TIGR : The institute for genomic research

Abréviations générales

ABA : Abscissic acid	EAGLE : Ecological and association genomics in legumes
ADN : Acide désoxyribonucléique	EMS : Ethyl methyl sulfonate
ADNc : Acide désoxyribonucléique complémentaire	EST : Expressed sequence tag
ADNg : Acide désoxyribonucléique génomique	FP6 : Framework programme 6
ADN-t : Acide désoxyribonucléique de transfert	GLIP : Grain legumes integrated project
AG : Acide gibbérellique	HD : haploïdes doublées
ARN : Acide ribonucléique	JC : plantes de jours courts
ARNm : Acide ribonucléique messenger	JL : plantes de jours longs
BAC : Bacterial artificial chromosome	LeARN : Legume rcRNA
BC : Back-cross	LIS : Legume information system
BLAST : Basic local alignment search tool	LOD : Logarithm of the odd score
CIM : Composite Interval Mapping	LR : lignée recombinante
DBF : date of begining of flowering	MAC : méristème apical caulinaire
DV : durée de la vernalisation	MCQTL : Multi-cross QTL
	MsC27 : Medicago sativa cDNA 27
	NPK : azote, phosphore, potassium
	ORF : Open reading frame

Pb : Paire de bases	TBE : Tris Borate EDTA
PCR : Polymerase chain reaction	TC : Tentative Consensus
QTL : Quantitative trait locus	TILLING : Targeting induced local lesions in genomes
RIL : Recombinant inbred line	Tnt1 : Transposon tagged 1
RKIP : Raf Kinase Inhibitor Protein	SNP : Single nucleotide polymorphism
RN : nœud reproducteur	SSH : Suppression subtractive hybridation
RNAi : Ribonucléic acid interference	SIM : Simple Interval Mapping
RT-PCR : Reverse-transcriptase polymerase chain reaction	
SSR : Single sequence repeat / marqueur microsatellite	

Gènes

AG : Agamous	LD : Lumini dependens
AP : Apetala	LFY : Leafy
CCA : Circadian clock associated	LHY : Late elongated hypocotyls
CO : Constans	MTAP : Medicago truncatula apetala
CRY : Cryptochrome	NORK : Nodulation receptor kinase K
DMI2 : Doesn't make infection 2	PHYA : Phytochrome A
ELF : Eucalyptus LEAFY	PHYB : Phytochrome B
FCA : Flowering time controle in Arabidopsis	PKS : Phytochrome kinase substrate
FD : Flowering locus D	PIM : Proliferating inflorescence meristem
FLC : Flowering locus C	RGA : Regulator of gibberellic acid
FRI : Frigida	SOC : Suppressor of overexpression of CO
FT : Flowering locus T	TOC : Timing of CAB expression
GAI : Gibberellic acid insensitive	VIP : Vernalization independent plant
GI : Gigantea	VRN : Vernalization

SOMMAIRE

1	<u>INTRODUCTION GÉNÉRALE.....</u>	<u>1</u>
2	<u>SYNTHÈSE BIBLIOGRAPHIQUE</u>	<u>5</u>
2.1	MEDICAGO TRUNCATULA.....	5
2.1.1	TAXONOMIE ET BIOLOGIE.....	5
2.1.1.1	Genre Medicago.....	5
2.1.1.2	Morphologie et cycle de développement.....	8
2.1.2	DIVERSITE DISPONIBLE	11
2.1.2.1	Diversité naturelle.....	11
2.1.2.2	Mutants induits	13
2.1.3	GENOMIQUE FONCTIONNELLE, GENETIQUE ET TRANSCRIPTOMIQUE.....	15
2.1.3.1	Coordination internationale des initiatives de génomique.....	15
2.1.3.2	ADNc, EST, et analyses transcriptomiques.....	16
2.1.3.3	Cartes physiques, génétiques et cytogénétiques	18
2.2	LE PROCESSUS DE FLORAISON ET LA DATE DE FLORAISON CHEZ LES DICOTYLEDONES	
	ANNUELLES.....	20
2.2.1	PHYSIOLOGIE	20
2.2.1.1	Photopériodisme : induction et rôle adaptatif	20
2.2.1.2	Horloge circadienne	22
2.2.1.3	Influence de la température	23
2.2.1.4	Modèles physiologiques de la réponse aux facteurs du milieu.....	25
2.2.1.5	Relation entre date de floraison et processus de développement.....	26
2.2.2	CONTROLE GENETIQUE DE LA DATE DE FLORAISON	28
2.2.2.1	Gènes connus et leurs interactions chez <i>A. thaliana</i>	28
2.2.2.2	La variabilité allélique naturelle connue	32
2.2.3	ROLE DE LA FLORAISON DANS L'ADAPTATION ET LA VALEUR AGRONOMIQUE.....	33
2.3	RECHERCHE DE ZONES DU GENOME ET DES GENES IMPLIQUES DANS LE DETERMINISME D'UN	
	CARACTERE QUANTITATIF	35
2.3.1	CONSTRUCTION DE CARTES GENETIQUES.....	35
2.3.1.1	Critères de construction.....	35
2.3.1.2	Cartographie fine	36
2.3.2	CARTOGRAPHIE DE LOCUS A EFFETS QUANTITATIFS	36
2.3.2.1	Avantages et inconvénients des descendances utilisées.....	37
2.3.2.2	Analyses QTL sur des populations multiples	38
2.3.3	APPROCHE DE GENES CANDIDATS POSITIONNELS	40
2.3.3.1	Choix des gènes candidats.....	41
2.3.3.2	Recherche de polymorphisme et cartographie fine.....	41
2.3.3.3	Validation du gène candidat	42
2.4	CONCLUSION	44
3	<u>MATÉRIEL ET MÉTHODES.....</u>	<u>47</u>
3.1	EFFETS PHOTOPERIODE, LIGNEE ET LEUR INTERACTION POUR LA DATE DE FLORAISON	47
3.1.1	MATERIEL VEGETAL.....	47
3.1.2	NOTATIONS PHENOTYPIQUES	48
3.1.3	ANALYSES STATISTIQUES.....	48

3.2	DETECTION DE QTL DE LA DATE DE FLORAISON DANS TROIS POPULATIONS DE CARTOGRAPHIE	49
3.2.1	MATERIEL VEGETAL.....	49
3.2.2	GENOTYPAGE DE MARQUEURS SUR LA POPULATION LR4	50
3.2.3	STATISTIQUES ET ANALYSES QTL SUR LES POPULATIONS LR1, LR4 ET LR5.....	52
3.3	MODELISATION DE LA DATE DE FLORAISON	53
3.4	CARTOGRAPHIE FINE D'UN QTL DE LA DATE DE FLORAISON	54
3.4.1	MATERIEL VEGETAL.....	54
3.4.2	GENOTYPAGE.....	56
3.4.3	CARTOGRAPHIE FINE D'UN QTL SUR LE CHROMOSOME 7	58
3.4.4	BIO-INFORMATIQUE	59
3.5	EXPRESSION DE QUELQUES GENES CANDIDATS DE LA DATE DE FLORAISON	61
3.5.1	MATERIEL VEGETAL.....	61
3.5.2	RT-PCR SEMI-QUANTITATIVE.....	62
4	<u>RÉSULTATS</u>	<u>67</u>
4.1	EFFETS PHOTOPERIODE, LIGNEE ET LEUR INTERACTION POUR LA DATE DE FLORAISON	67
4.1.1	VARIATION DE LA DATE DE FLORAISON EN REPONSE A LA PHOTOPERIODE	67
4.1.2	MODELISATION DE LA DATE DE FLORAISON.....	71
4.2	DETECTION DE QTL POUR LA DATE DE FLORAISON DANS TROIS POPULATIONS DE	
	CARTOGRAPHIE	73
4.2.1	VARIATIONS QUANTITATIVES DE LA DATE DE FLORAISON DANS TROIS POPULATIONS DE	
	CARTOGRAPHIE.....	73
4.2.2	IDENTIFICATION DE QTL.....	77
4.2.3	PROJECTION DE CARTES ET ANALYSE MULTI-CROISEMENTS.....	80
4.2.4	CARTOGRAPHIE FINE D'UN QTL SUR LE CHROMOSOME 7	82
4.2.5	SELECTION DE GENES CANDIDATS	89
4.2.6	SEQUENÇAGE DES GENES CANDIDATS.....	93
4.3	EXPRESSION DE QUELQUES GENES CANDIDATS DE LA DATE DE FLORAISON	103
5	<u>DISCUSSION GÉNÉRALE ET PERSPECTIVES</u>	<u>108</u>
5.1	EFFET DE LA PHOTOPERIODE SUR LA DATE DE FLORAISON.....	108
5.2	LA MODELISATION DE LA DATE DE FLORAISON	109
5.3	ANALYSE QTL MULTI-CROISEMENTS	111
5.4	CARTOGRAPHIE FINE ET CARTOGRAPHIE PHYSIQUE.....	112
5.5	LE CHOIX DES GENES CANDIDATS	115
5.6	EXPRESSION DES GENES CANDIDATS	118
5.7	VERS UNE VALIDATION DU GENE CONSTANS	119
5.8	QUELLE STRATEGIE GENETIQUE POUR LA DATE DE FLORAISON ?	121
6	<u>CONCLUSION</u>	<u>124</u>
	<u>REFERENCES BIBLIOGRAPHIQUES</u>	<u>126</u>

ANNEXES

INTRODUCTION GÉNÉRALE

1 INTRODUCTION GÉNÉRALE

Les Légumineuses (Fabacées) constituent un taxon végétal très important d'un point de vue biologique et écologique mais également agronomique et environnemental. En effet, leur culture fournit depuis les débuts de l'agriculture une source essentielle de protéines pour l'alimentation humaine et animale. Leur capacité à fixer l'azote atmosphérique génère un double intérêt, économique par une moindre utilisation d'engrais azotés de synthèse, et écologique par une limitation des fuites azotées. En Europe, le déficit structurel en protéines végétales constitue un argument favorable à l'accroissement des surfaces cultivées en Légumineuses (soja, pois, lupin, luzerne,...) où en dépit de leurs nombreux avantages elles occupent moins de 5 % de la surface allouée aux grandes cultures. La situation apparaît toutefois un peu plus favorable si on prend en compte les prairies semées reposant sur des associations entre Graminées et Légumineuses. En effet, près de 60 % des trois millions d'hectares de prairies temporaires reposent aujourd'hui sur des associations (Huyghe 2005). Cependant, malgré une politique volontariste qui a favorisé de 1973 à 1992 le développement de ces cultures riches en protéines, le taux de couverture en protéines utilisées (en complément) dans l'alimentation animale de l'Union Européenne n'a jamais excédé 38 %. En France, le déficit était de 54 % en 2003 (Crepon 2005). Les Fabacées suscitent donc un regain d'intérêt avec le souhait que l'agriculture moderne évolue vers des pratiques durables plus respectueuses de l'environnement.

Les travaux sur la plante modèle *Arabidopsis thaliana* ont permis des avancées considérables dans la compréhension des bases génétiques et moléculaires du fonctionnement des plantes (Bouchez 2004). Cependant, l'utilisation de cette espèce n'est pas possible pour la compréhension de mécanismes tels que la symbiose nodulaire avec des bactéries pour la fixation d'azote atmosphérique ou encore la mycorhization (Gianinazzi-Pearson 1996; Harrison 1997). De plus, la distance phylogénétique entre les familles botaniques représentées parmi les plantes cultivées est trop importante et rend donc nécessaire de disposer dans chaque famille d'une espèce modèle, comme *A. thaliana* pour les Crucifères ou encore *Oryza sativa* pour les céréales. Dans ce contexte, *Medicago truncatula* et *Lotus japonicus* ont été proposées comme espèces modèles des Légumineuses et ont été choisies depuis quelques années par la communauté scientifique internationale (Barker et al. 1990; Cook 1999; Udvardi 2001).

Ces deux espèces présentent des caractéristiques génétiques similaires, mais développent des nodosités de types différents (croissance indéterminée/déterminée). Cependant, *M. truncatula* est phylogénétiquement plus proche de la plupart des Légumineuses cultivées en Europe que *Lotus japonicus* (Choi et al. 2004). Les Légumineuses ont souvent des génomes complexes. C'est le cas de la luzerne (*Medicago sativa*), espèce fourragère pérenne, autotétraploïde et allogame ainsi que du pois (*Pisum sativum*), espèce protéagineuse qui présente un génome de grande taille. Des études précises peuvent être menées sur *M. truncatula*, espèce diploïde, annuelle, à cycle court et autogame. De nombreuses ressources génétiques sont disponibles (banques BAC, EST/ADNc, cartes génétiques et physiques, mutants, variabilité génétique naturelle, séquences) et il existe un fort degré de synténie avec les génomes de la luzerne et du pois. On peut ainsi envisager de réaliser des études génétiques sur cette espèce, en espérant que les gènes expliquant la variabilité génétique chez *M. truncatula* expliquent aussi les variations détectées chez les Légumineuses cultivées.

La morphogenèse aérienne inclut des caractères de croissance, de développement et de phénologie, et conditionne fortement la valeur d'usage des Légumineuses. Parmi ces caractères, la floraison est un événement majeur du cycle de vie. Cette période est déterminante pour le succès reproductif. Elle correspond à la transition généralement non réversible d'un méristème végétatif produisant des feuilles et tiges, en méristème floral à l'origine des différentes structures verticillées composant la fleur. La régulation de ce phénomène morphogénétique est le fait d'un réseau complexe de signalisations, qui permet à la plante de réaliser sa reproduction dans des conditions pédoclimatiques optimales (Putterill et al. 2004).

L'étude de différents écotypes et de mutants d'*A. thaliana* a permis de distinguer quatre voies principales contrôlant la date de floraison, dont deux sont induites par des signaux environnementaux (photopériode et vernalisation) et deux autres par des signaux endogènes (voie hormonale et voie autonome) (Koornneef et al. 1998; Blazquez 2005). La date de floraison, qui est un caractère quantitatif, serait contrôlée par plus de 80 locus chez *A. thaliana* (Levy and Dean 1998).

L'identification précise, le clonage et la caractérisation de ces gènes constituent un défi majeur d'autant que la présence de gènes homologues chez différentes espèces suggère une bonne conservation des voies de contrôle de la date de floraison chez les angiospermes (Tasma and Shoemaker 2003).

La recherche des gènes de la floraison chez *M. truncatula* permet d'une part d'établir la conservation ou non des gènes avec ceux décrits principalement chez *A. thaliana*, et d'autre part de proposer les gènes qui expliquent la variabilité observée pour ce caractère. Dans ce contexte, l'objectif de ma thèse est d'identifier des gènes expliquant la variabilité génétique pour la date de floraison chez *M. truncatula*. La démarche développée consiste à :

- (i) décrire la variabilité génétique existant pour la date de floraison en réponse à la photopériode ;
- (ii) rechercher des QTL (*Quantitative Trait Locus*) de la date de floraison dans trois populations de lignées recombinantes connectées ;
- (iii) déterminer par une approche gènes candidats positionnels, les gènes qui pourraient correspondre au QTL ayant le plus fort effet et préciser par cartographie fine et étude d'expression le gène candidat le plus vraisemblable à ce QTL.

La synthèse bibliographique rappellera tout d'abord les caractéristiques biologiques et les ressources génétiques disponibles de *M. truncatula*, puis définira le processus de floraison et la date de floraison d'un point de vue physiologique et génétique, et enfin exposera la recherche de zones du génome et des gènes impliquées dans le déterminisme d'un caractère quantitatif. Puis, le matériel végétal et les méthodes seront décrits. Les résultats seront ensuite exposés et discutés. Une réflexion sera proposée sur la stratégie et les méthodes utilisées, et les perspectives possibles seront dégagées.

SYNTHÈSE BIBLIOGRAPHIQUE

2 SYNTHÈSE BIBLIOGRAPHIQUE

2.1 *Medicago truncatula*

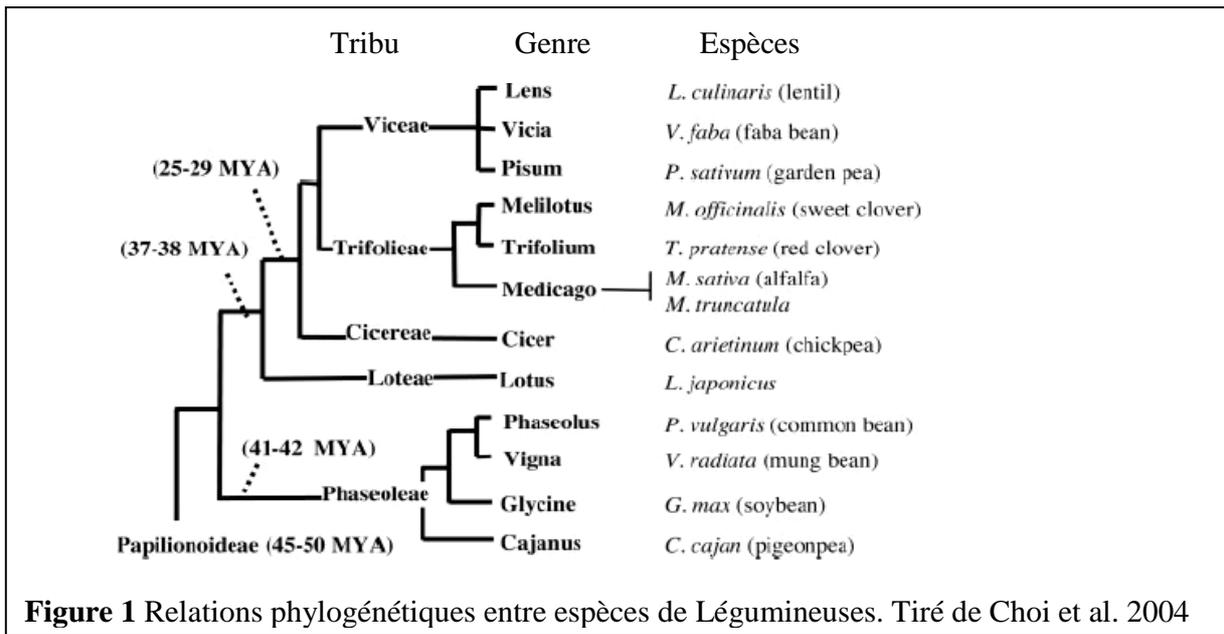
2.1.1 Taxonomie et biologie

2.1.1.1 Genre *Medicago*

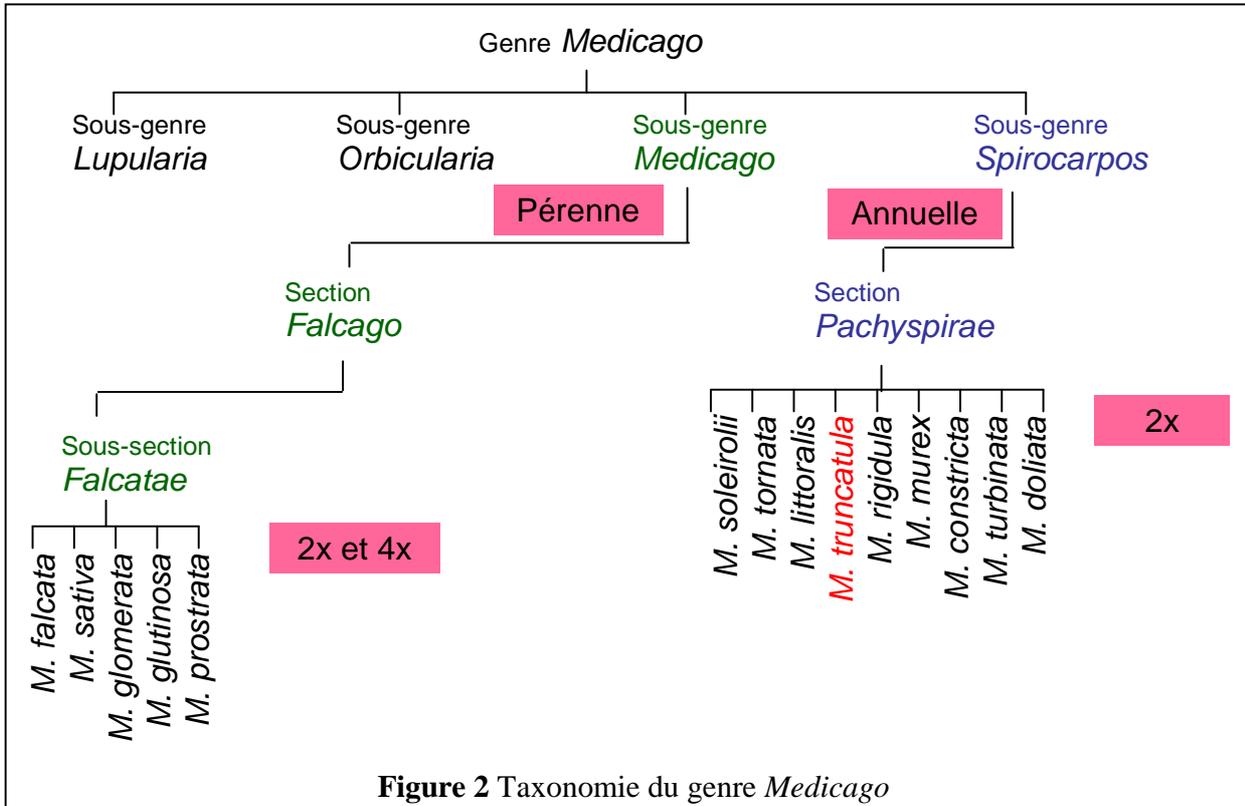
Le genre *Medicago* (les luzernes) regroupe de nombreuses espèces annuelles ou vivaces, le plus souvent herbacées, à feuilles trifoliolées. Le nom scientifique du genre n'est pas lié à ses propriétés médicales, mais au fait que, selon Théophraste (372 - 287 av. J.-C.), la luzerne serait originaire de Médie (région allant du Nord-Ouest de l'Iran à l'Azerbaïdjan) (Théophraste traduit par Amigues 1989). Le nom anglais et espagnol alfalfa proviendrait du nom arabe *al-fac-façah*, qui signifie « le père des aliments » car il servait de fourrage aux chevaux. Quant au nom vernaculaire, il est emprunté à l'occitan *luserna*, en raison de l'aspect brillant des graines de la plante.

A la suite de Vavilov (1887 - 1943) il est admis que le centre d'origine du genre *Medicago* est le Croissant fertile qui correspond à la région géographique occupée par l'Israël, la Cisjordanie, le Liban et des parties de la Jordanie, la Syrie, l'Irak, l'Égypte et sud-ouest de la Turquie d'où il aurait été apporté en Europe par le roi perse Darius (550 - 486 av. J.-C.). Son extension européenne ne commence réellement qu'avec l'Empire romain (Genier et al. 1992). La plupart des espèces constituant le genre *Medicago* sont décrites depuis le XVI^{ème} siècle ; ainsi dans *Species Plantarum* de Carl von Linné en 1753, neuf espèces sont décrites dont certaines avec plusieurs variétés botaniques (Heyn 1963). Plusieurs travaux importants réalisés au cours du XIX^{ème} siècle avaient abouti à une description complète du genre *Medicago*, mais mal reliés entre eux, ils comportaient un nombre important de synonymies. Il faut attendre les ouvrages de Heyn (1963) et (Lesins and Lesins 1979) pour obtenir une réelle synthèse.

Le genre *Medicago* appartient à l'ordre des Fabales, superfamille des Légumineuses, famille des *Fabaceae* et sous-famille des *Faboideae*. Le genre *Medicago* est placé dans la tribu des *Trifolieae* qui comprend la majorité des Légumineuses fourragères tempérées comme la luzerne cultivée et le trèfle, et qui est phylogénétiquement proche de la tribu des *Vicieae* incluant les Légumineuses à graines comme le pois (*P. sativum*) et la féverole (*Vicia faba*) (Figure 1) (Doyle and Luckow 2003; Choi et al. 2004).



Le genre *Medicago* est particulièrement proche des genres *Melilotus* et *Trigonella* et a été classé en quatre sous-genres sur la base de la morphologie des gousses et des graines : *Lupularia*, *Medicago*, *Orbicularia* et *Spirocarpos* (Figure 2) (Lesins and Lesins 1979).



Ces sous-genres comprennent au total 55 espèces dont 21 sont pérennes, 33 sont annuelles et une est de pérennité non définie. Le sous-genre *Spirocarpos* est composé de 29 espèces annuelles (médics) et comporte quatre sections : *Intertextae*, *Leptospirae*, *Pachyspirae* et *Rotatae*. *M. truncatula* (appelée communément luzerne tronquée) appartient à la section *Pachyspirae* aux côtés de huit autres espèces : *constricta*, *doliata*, *littoralis*, *murex*, *rigidula*, *soleirolii*, *tornata* et *turbinata*.

Le nombre chromosomique de base du genre *Medicago* est $x = 8$, excepté chez quatre espèces annuelles *M. constricta* Dur., *M. praecox* DC., *M. polymorpha* L., et *M. rigidula* L. qui ont un nombre chromosomique de base $x = 7$. Chez les différentes espèces du genre, trois niveaux de ploïdie existent : diploïdes ($2n = 2x = 14$ et $2n = 2x = 16$) qui représentent la majorité des espèces, tétraploïdes ($2n = 4x = 32$) et hexaploïdes ($2n = 6x = 48$) (Quiros and Bauchan 1988). La luzerne cultivée (*M. sativa*) est une espèce tétraploïde pérenne alors que *M. truncatula* est diploïde et annuelle. Tout comme les espèces annuelles *M. polymorpha* et à un niveau moindre *M. orbicularis*, *M. truncatula* est une espèce colonisatrice et ubiquiste (Olivieri 1991). *M. truncatula* produit des graines en abondance avec un temps de génération autour de 10 à 12 semaines de graines à graines en conditions contrôlées. La taille de son génome est d'environ 470 Mb/1C (trois à quatre fois supérieure à celle d'*A. thaliana*, équivalente à celle d'*O. sativa* et environ neuf fois inférieure à celle de *P. sativum*) (Young et al. 2005). Son mode de reproduction est principalement autogame, si bien que les individus sont généralement homozygotes.

L'aire de répartition actuelle de *M. truncatula* se situe sur l'ensemble du pourtour méditerranéen et également pour une partie en Europe centrale et Europe du Nord (Heyn 1963). Les populations spontanées de *M. truncatula* sont fréquemment exploitées dans les zones de parcours, souvent par des ovins au pâturage. La culture de cette espèce est très marginale dans le pourtour méditerranéen. A l'opposé, dans les systèmes prairiaux australiens *M. truncatula* occupe une place prépondérante avec plus de 4.5 millions d'hectares cultivés et représente ainsi l'espèce la plus cultivée parmi les médics (Nair et al. 2006). La principale particularité des cultivars sélectionnés en Australie réside dans leurs précocités et leurs adaptations aux conditions pédoclimatiques.

Les plus couramment utilisés sont Jemalong, Cyprus, Paraggio et Sephi (Olivieri 1991). En France, une seule variété de *M. truncatula* est inscrite au catalogue officiel (Salernes).

2.1.1.2 Morphologie et cycle de développement

M. truncatula (Figure 3) est une espèce de taille intermédiaire, de 15 à 80 cm, velue, à port variable, souvent prostré. Elle est présente sur des sols lourds, marneux ou argileux. Le pédoncule est plus court que le pétiole et l'inflorescence porte de une à cinq fleurs jaunes groupées en racèmes (Lesins and Lesins 1979). Les gousses sont cylindriques (7 - 12 mm de diamètre), glabres, très dures, à spires jointives et serrées (3.5 - 6), aux épines recourbées, souvent perpendiculaires au plan des spires. Elles contiennent de trois à douze graines. Le poids de 1 000 graines varie de 3.3 à 6 g.

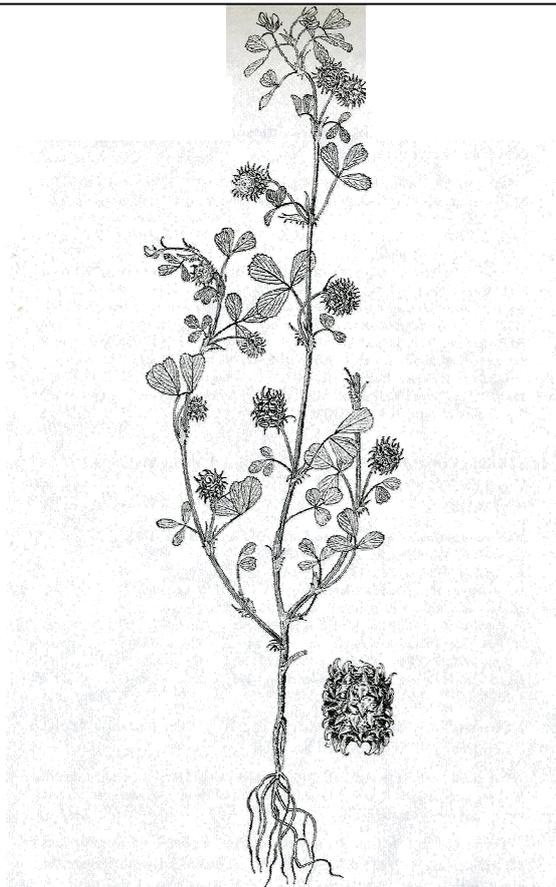


Figure 3 Plante et gousse de *M. truncatula*. Tiré de Clarkson (1977) et Prospero et al. (1995)

Au-delà de ces particularités, *M. truncatula* possède les spécificités du genre *Medicago*, c'est-à-dire une corolle papilionacée constituée d'un étendard, de pétales formant deux ailes libres, d'une carène formée par les deux pétales inférieurs soudés ; neuf étamines soudées forment la colonne staminale, une dixième étamine est libre. La corolle et la colonne staminale constituent le dispositif de déclenchement de la fleur, caractéristique du genre, retrouvé aussi chez les espèces autogames. Le calice est formé de cinq sépales soudés. Les feuilles sont trifoliolées, terminées par un mucron. Les stipules sont collées au pétiole. L'absence chez *Medicago* d'une articulation entre les cotylédons et l'hypocotyle au stade plantule constitue un caractère distinctif important avec les genres *Melilotus* et *Trigonella* (Prosperi et al. 1995b).

La description de *M. truncatula* par les écophysiologistes a permis de mieux appréhender son développement et sa croissance (Djemel et al. 2005; Wang and Grusak 2005; Moreau et al. 2007). Pour le développement, les critères phénotypiques tels que nombre de feuilles, de tiges, d'organes reproducteurs, de gousses ou de graines ont été caractérisés. Ils permettent la désignation des modifications qualitatives issues de l'initiation et de l'apparition de nouveaux organes végétatifs et reproducteurs. La croissance de la plante par l'augmentation du poids et de la taille résultant de l'accumulation d'éléments chimiques et d'eau a été analysée. Peu d'outils de caractérisation de l'état des plantes de *M. truncatula* ont ainsi été décrits et ces données étaient très partielles car bien souvent limitées au début de la floraison. La mise en place d'un cadre d'analyse du développement de *M. truncatula* avec une terminologie précise d'identification des différents organes et d'une échelle de notations des stades de développement est très récente (Moreau et al. 2006; Bucciarelli et al. 2006).

La plante produit tout d'abord deux cotylédons puis la première feuille unifoliolée appelée feuille 0. Viennent ensuite les feuilles trifoliolées (1 à 5 suivant l'ordre d'apparition) sur l'axe principal, organisées en rosette selon une phyllotaxie qui peut être dextre ou sénestre. A partir de la feuille 6, la disposition devient alterne sur l'axe principal. Le phyllochrone qui représente l'intervalle de temps entre l'apparition de deux feuilles étalées sur un axe, est de 2.5 à 3 jours en croissance rapide de la plante à une température moyenne de 15.3°C, soit 37 à 45 °C.J pour une feuille (Chabaud et al. 2006).

Les ramifications primaires (B0, B1 et B2) sont nommées d'après leur position sur l'axe principal. Généralement des ramifications secondaires et tertiaires apparaissent à la base axillaire des feuilles de ramifications primaires et secondaires respectivement (Figure 4). La croissance indéterminée de *M. truncatula* se caractérise par l'aptitude de l'apex à rester végétatif. La progression séquentielle du développement reproductif s'effectue le long des axes depuis le premier nœud jusqu'à l'apex. La position du premier nœud reproducteur varie selon l'axe mais il n'existe pas de modèle qui explique son emplacement exact. Un nœud reproducteur porte de deux à trois fleurs ou inflorescences, souvent en plus d'un axe d'ordre supérieur.

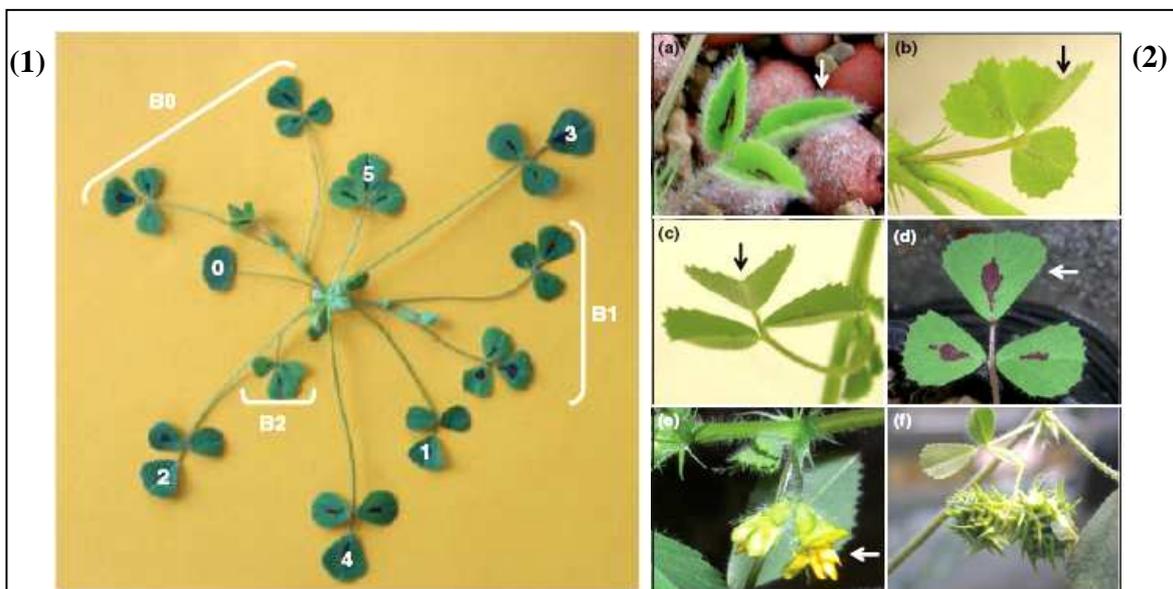


Figure 4

- (1) Phyllotaxie de l'axe principal et dispersion des ramifications primaires (B0 à B2).
 (2) Code décimal pour noter le développement végétatif et définir le développement des stades reproducteurs
- (a) 0.25 : les trois feuilles trifoliolées sont séparées et les bords de la feuille trifoliolée centrale forment un angle aigu
 - (b) 0.50 : les bords de la feuille trifoliolée centrale forment un angle de 90°
 - (c) 0.75 : les bords de la feuille trifoliolée centrale forment un angle obtus
 - (d) 0.00 : le limbe de la feuille trifoliolée centrale est plat
 - (e) un nœud est considéré reproducteur quand au moins une fleur est ouverte
 - (f) des graines sur un nœud reproducteur

Dans de bonnes conditions de culture, les plantes peuvent alors produire jusqu'à trois cycles de graines à graines par an (Chabaud et al. 2006). En conditions naturelles, la croissance végétative a lieu principalement en automne, hiver et au début du printemps, c'est-à-dire sous des températures fraîches et des photopériodes courtes.

La floraison se produit au printemps sous des photopériodes encore courtes et les graines mûturent à la fin du printemps ou au début de l'été (Prosperi et al. 1995a).

Les stades végétatifs sont définis par le nombre de feuilles trifoliolées déployées par axe ainsi que par le nombre d'entre-nœuds. Les stades reproducteurs sont quant à eux déterminés par (i) l'identification de la position du premier nœud reproducteur pour chaque axe et (ii) le nombre de nœuds reproducteurs par axe (Benlloch et al. 2003; Moreau et al. 2006). Cette méthode de notation a l'avantage de permettre une caractérisation phénotypique plus complète qu'une simple description à l'échelle d'un organe individuel ou de la nature des organes présents sur la plante.

2.1.2 Diversité disponible

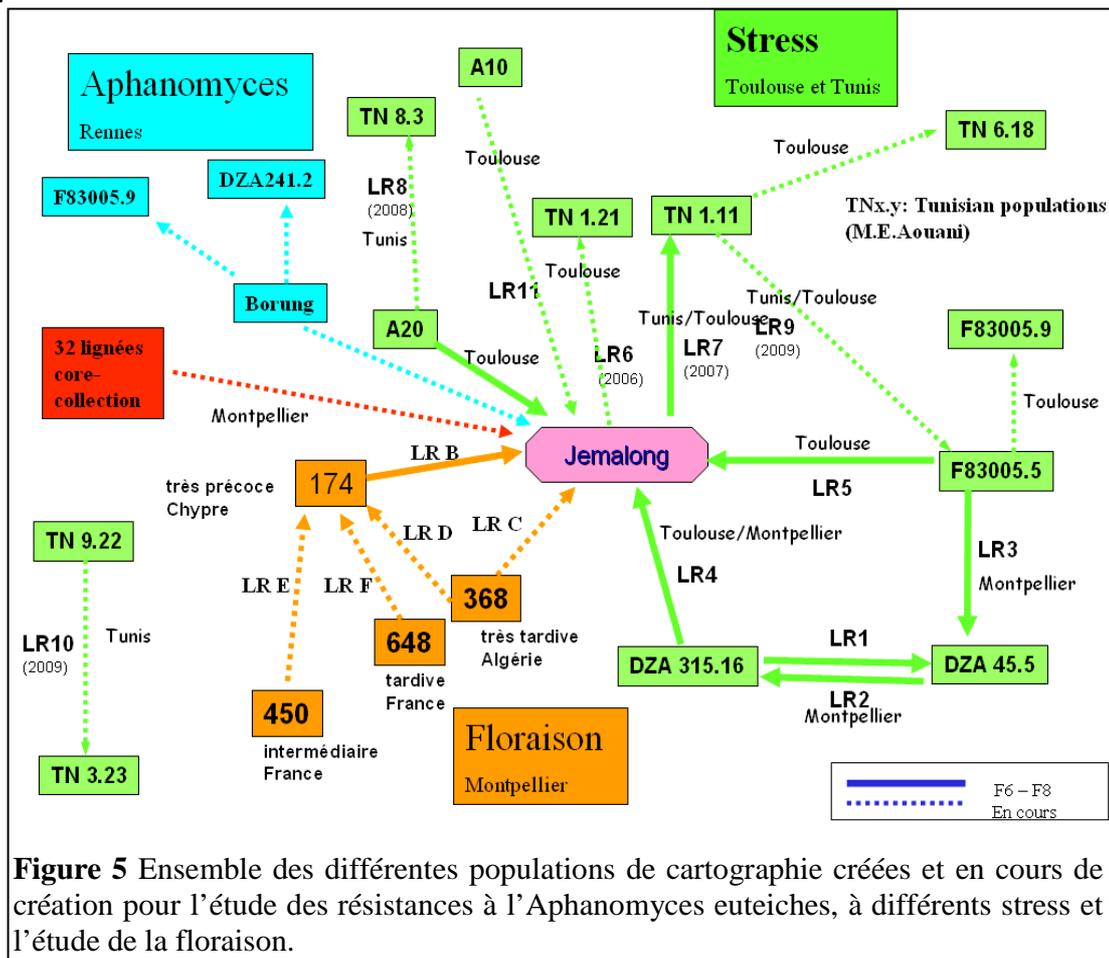
2.1.2.1 *Diversité naturelle*

Chez *M. truncatula*, il existe une forte diversité génétique rendue accessible pour l'expérimentation grâce, principalement, à une collection de plusieurs centaines de populations naturelles (principalement issues de France, Espagne, Portugal, Grèce, Maroc et Algérie) constituée à l'INRA de Montpellier (Ronfort et al. 2006). En 2005, la collection de l'INRA (<http://bioweb.ensam.inra.fr/multicrop>) comprenait 833 populations et 944 lignées fixées. Sur le long terme, la collection s'enrichira de 1 500 populations et de plus de 3 000 lignées fixées (Nair et al. 2006). Une autre collection de 5 120 accessions de *M. truncatula* est disponible en Australie, au centre de ressources génétiques SARDI (*South Australian Research and Development Institut*) (<http://www.sardi.sa.gov.au>) (Ellwood et al. 2006). De telles collections d'accessions permettent de conserver et mettre à la disposition de la communauté scientifique une ressource génétique considérable bien que le nombre de collectes pour certains pays méditerranéens soit faible (Ronfort et al. 2006).

Le criblage de ces ressources génétiques a révélé une large diversité génétique et une grande variation allélique au sein de cette espèce autogame entre populations mais aussi à l'intérieur de certaines populations (Bonnin et al. 1996; Bonnin et al. 2001). L'étude des flux génétiques permet de comprendre les mécanismes évolutifs qui expliquent le maintien de la diversité au sein de populations très autogames, et de déterminer comment cette diversité génétique est organisée à l'intérieur d'une population spécifique, mais aussi entre populations (Jenczewski et al. 2003).

Certaines de ces populations ont été analysées pour des caractères phénotypiques liés à des composantes agronomiques chez les Légumineuses cultivées. Des variants naturels ont été mis en évidence pour la reconnaissance symbiotique plante-*Sinorhizobium* ainsi que pour des lignées présentant divers degrés de résistance ou de sensibilité à des champignons pathogènes (*Phytophthora medicaginis*, *Colletotrichum trifolii*) et des nématodes (*Meloidogyne incognita*, *Ditylenchus dipsaci*) (Tirichine et al. 2000; Moussart et al. 2007).

Une grande variation a été mise en évidence chez cinq accessions de *M. littoralis*, 24 accessions de *M. truncatula* et dans une population de lignées recombinantes pour les caractères de date de floraison, de morphologie (poids sec, nombre de tiges primaires et secondaires) et de dynamique d'élongation des tiges (longueur, diamètre, vitesse d'élongation,...). Il est également intéressant de noter que lors de cette étude, la date de floraison fut trouvée négativement corrélée au diamètre et à la longueur des tiges (Julier et al. 2007). Une autre étude portant sur 700 lignées issues de 560 populations a montré une grande diversité pour la date de floraison (Delalande et al. 2004).



Enfin, du matériel scientifique a été créé à l'INRA de Montpellier, l'INRA de Rennes, l'ENSAT et le CBBC de Tunis à partir de la diversité naturelle : plus de 500 descendances F2, F3 et générations suivantes, et trois populations de 200 lignées recombinantes chacune, en génération F6 à F8. Trois grands thèmes de recherche sont abordés : la résistance à l'aphanomyces, la résistance à différents stress et l'étude de la floraison (Figure 5).

2.1.2.2 *Mutants induits*

La génétique inverse définit les techniques permettant à partir d'un gène ou fragment d'ADN l'étude des fonctions de ce gène et de ses produits, par opposition à la génétique classique dont le but est de localiser le gène responsable de l'altération d'une fonction ou d'un caractère. La génétique inverse (bas débit) s'inscrit dans la génomique fonctionnelle (haut débit) qui est la compréhension du fonctionnement des gènes et des autres composants du génome. La génération de mutants est indispensable à des études de génétique inverse ou de génomique fonctionnelle.

Le développement de la génomique fonctionnelle de *M. truncatula* passe par la réalisation de mutagenèse à haut-débit sur l'ensemble du génome, par les techniques d'ADN-t ou de transposons. Bien que le nombre de lignées d'insertion pour atteindre la saturation du génome soit estimé autour de 300 000, cette stratégie a déjà fait les preuves de son efficacité pour la génétique directe ou inverse chez *A. thaliana* (Parinov and Sundaresan 2000).

Plusieurs programmes de mutagenèse (induction de mutations dans un génome) par des agents physico-chimiques classiques ont été réalisés ou sont en cours : par rayons γ , par Ethyl-Methyl-Sulfonate (EMS) et par neutrons rapides (Penmetsa and Cook 2000).

Des programmes de mutagenèse (rayons γ , ADN-t et transposon de tabac *Tnt1*) ont été développés sur *M. truncatula* par l'INRA, l'Institut des sciences du végétal (CNRS à Gif, France), chez Plant Research International (Wageningen en Hollande) ou par la Noble Foundation (USA). La majorité de ces programmes a été réalisée sur l'accession de référence A17 (lignée sélectionnée à partir du cultivar australien Jemalong) et sur la lignée R108 spécifiquement pour le transposon *Tnt1* (d'Erfurth et al. 2003).

En 2005, 1 750 lignées ont été produites : 200 lignées depuis les transposons *Tnt1*, 50 depuis les rayons γ et 1 500 lignées depuis les ADN-t de transfert (Nair et al. 2006). Des collections (familles M2 mutagénisées à l'EMS) de mutants de TILLING (*Targeting Induced Local Lesions IN Genomes*) ont été construites à l'université Davis de Californie (USA) et à l'INRA de Dijon (France) (Colbert et al. 2001). Des programmes de production de mutants sont développés notamment depuis le programme européen FP6 « Grain Legumes » (2004 - 2008) qui a permis la production d'une population de TILLING de 9 000 individus M2 dérivée de 5000 M1, et prévoit la création d'environ 8 000 lignées mutantes *Tnt1*. L'étude de mutants offre la possibilité d'analyser des gènes impliqués dans divers processus biologiques comme la perception/transduction de signaux biotiques, le développement (morphogénèse et organogénèse), la régulation de la nutrition minérale,... C'est ainsi que les mutants créés ont permis le criblage à haut débit de divers phénotypes, particulièrement les mutants affectés dans les interactions symbiotiques et mycorhiziennes. Chez *M. truncatula*, ce travail a été particulièrement approfondi dans le cas des mutants de nodulation grâce à leur identification et leur caractérisation phénotypique et génétique. Les mécanismes de reconnaissance bactérie-plante ont pu ainsi être disséqués, ce qui a permis le clonage en 2002 du gène DMI2 codant pour un récepteur kinase et impliqué dans l'établissement de la mycorhization chez *M. truncatula*, la luzerne, le pois et le lotier (Catoira et al. 2000; Catoira et al. 2001; Endre et al. 2002; Stracke et al. 2002; Kistner and Parniske 2002).

Chez *M. truncatula* il existe encore peu de mutants décrits dans le développement floral (Domoney et al. 2006). Cependant le mutant MEDICAGO TRUNCATULA APETALA (*mtap*) obtenu par mutagenèse EMS se distingue par une conversion homéotique de certains verticilles floraux (corolle et androcée) en sépale et carpelle. Ce genre de mutation est similaire à celle observée chez *A. thaliana* pour les gènes APETALA3 / PISTALLATA (Bowman et al. 1989; Penmetsa and Cook 2000). L'insertion du rétrotransposon *Tnt1* dans un homologue des gènes PROLIFERATING INFLORESCENCE MERISTEM (PIM) / APETALA1 / SQUAMOSA induit une conversion de la fleur en inflorescence et altère les fleurs en transformant les sépales en feuilles. Il provoque ainsi un phénotype qui ressemble aux mutants *pim* du pois (Benlloch et al. 2006).

2.1.3 Génomique fonctionnelle, génétique et transcriptomique

2.1.3.1 *Coordination internationale des initiatives de génomique*

L'étude de la génétique et de la génomique chez *M. truncatula* est aujourd'hui au centre de vastes programmes de recherche initiés à l'échelle mondiale impliquant plusieurs instituts (<http://www.medicago.org>). Ainsi un programme international de recherche sur *M. truncatula* est soutenu par le *Plant Genome Program* de la *National Science Foundation* (NSF). Ce programme vise au séquençage des régions riches en gènes avec l'objectif d'atteindre 80 % du génome de *M. truncatula* fin 2006 avec un système d'annotation automatique coordonnée par IMGAG (*International Medicago Genome Annotation Group*). Ainsi, près de 20 000 gènes ont été prédits, avec une longueur moyenne de 2 172 pb (Tableau 1).

Tableau 1 Prédications des gènes et découpages de gènes effectuées par le logiciel IMGAG

	Nombre total	Longueur moyenne (pb)	Longueur Minimum (pb)	Longueur Maximum (pb)	Densité (kb/gène ou exon)
Gène	20 118	2 172	91	49 909	4.32
Exon	61 605	448	4	42 848	1.41

Aux USA, quatre chromosomes sont pris en charge par l'université d'Oklahoma, deux par le TIGR (*The Institute for Genomic Research*). En Europe, un chromosome est étudié par l'institut Sanger (UK) et un par le Génoscope d'Evry en France. Au total, 1 950 BAC représentant 188.55 Mpb ont été séquencés (Tableau 2).

Tableau 2 Répartition du séquençage des huit chromosomes de *M. truncatula*

	Centre	Taille (Mpb)	Pourcentage de redondance	BAC
Chromosome 1	OU	19.36	22.73	213
Chromosome 2	TIGR	20.88	10.73	211
Chromosome 3	Sanger	26.55	0	272
Chromosome 4	OU	26.74	13.28	275
Chromosome 5	Genoscope	27.32	7.65	264
Chromosome 6	OU	13.83	4.48	139
Chromosome 7	TIGR	24.3	2.3	242
Chromosome 8	OU	25.22	19.11	284
Non ancré		4.35		50
Total		188.55	10.64	1950

L'institut américain privé *Noble Foundation* finance également un important programme de génomique fonctionnelle, protéomique, métabolomique et bioinformatique (<http://www.noble.org/medicago>). L'autre grand programme multinational s'intitule « *Grain Legumes Integrated Project* » (GLIP), soutenu financièrement par le 6^e *Framework Programme* (FP6) de la Commission Européenne (2004 - 2008). Organisé autour de huit modules, il s'efforce de développer de nouvelles stratégies afin d'augmenter l'utilisation des Légumineuses à graines pour la consommation humaine et animale (<http://www.eu grainlegumes.org>). *M. truncatula* y trouve sa place principalement dans la recherche de mutants dans des gènes ciblés.

A l'heure actuelle, de nombreux outils bio-informatiques sont disponibles chez *M. truncatula*. Ils concernent les prédictions et annotations de gènes, d'EST (*Expressed Sequence Tag*) et de protéines, mais permettent aussi la réalisation d'alignements de séquences sur les bases de données spécifiques de *M. truncatula* ou des Légumineuses (Tableau 3).

Tableau 3 Quelques bases de données bio-informatiques disponibles pour *M. truncatula*

Données	Nom	Site
Gènes (prédictions et annotations) BAC et BAC-end Marqueurs	IMGAG (International Medicago Genome Annotation Group).	http://medicago.org/genome/
EST	TIGR (The Institute for Genomic Research)	http://compbio.dfc.harvard.edu/tgi/cgi-bin/tgi/Blast/index.cgi
Alignements nucléotidique et protéique	Université d'Oklahoma	http://www.genome.ou.edu/medicago.html
EST, ORF, cartes génétiques <i>M. truncatula</i> , <i>L. japonicus</i> , <i>G. max</i>	LIS (Legume Information System)	http://www.comparative-legumes.org

2.1.3.2 ADNc, EST, et analyses transcriptomiques

L'investissement consenti sur *M. truncatula* a permis l'obtention d'un grand nombre de séquences EST, ou de marqueurs de séquences exprimées utilisés comme marqueurs dans l'identification de gènes exprimés. C'est le résultat le plus visible du progrès sur cette plante modèle. Historiquement les deux premiers lots d'EST ont été produits à partir d'une banque d'ADNc issus de pointes racinaires enrichies en poils absorbants et d'une banque issue de jeunes nodosités (Covitz et al. 1998; Györgyey et al. 2000).

Depuis, les programmes de séquençage systématiques de la NSF, la *Noble Foundation* et du Génoscope ont permis le dépôt de 180 000 EST en 2002 et 227 000 fin 2006 dont 95 % ont été générés à partir de banques standards d'ADNc (http://medicago.org/genome/genome_stats.php) (Frugoli and Harris 2001; VandenBosch and Stacey 2003) (Tableau 4).

Tableau 4 Ensemble des EST et TC disponibles au 11/01/2006 sur <http://medicago.org>

	Nombre	Alignement génomique
EST	226 923	
TC	18 612	52.4 %
Dont TC de haute qualité	9 396	56.0 %
Singletons	18 266	38.9 %

Ces EST sont issus de plusieurs dizaines de banques d'ADNc qui représentent différents organes, stades de développement et traitements par des stimuli biotiques ou abiotiques (http://www.ncbi.nlm.nih.gov/dbEST/dbEST_summary.html). Plusieurs bases de données regroupent désormais tous les EST de *M. truncatula*, bien souvent répertoriées en clusters, chacun représentant un gène. En 2006, le TIGR a fourni 18 000 consensus provisoires TC (*Tentative Consensus*) assorties d'une annotation automatique (Schiex et al. 2003). Les TC sont créés par l'assemblage d'EST et dans certains cas de séquences entières ou partielles d'ADNc. Ils contiennent également des informations concernant l'origine et l'abondance des EST qui le composent.

La banque de données internationale LeARN (*Legume rcRNA*) référence les ADNc non redondants. Au total, 41 banques d'ADNc non-soustractives ont été construites sur la lignée A17 et deux sur la lignée R108. La génération de clones ADNc pleine longueur a été entreprise car cela présente de nombreux avantages dans l'identification de gènes et pour les études fonctionnelles des protéines correspondantes. Le programme du TIGR est actuellement le seul à référencer des ADNc issus de pools de tissus.

L'une des applications du séquençage massif des EST et ADNc est le développement de macro- et micro-arrays pour l'analyse du transcriptome (Schena et al. 1995). En effet, le clustering des EST a permis de distinguer certains gènes spécifiques et ainsi choisir, pour représenter physiquement chaque gène, le meilleur clone d'ADNc. Les macro- et micro-arrays développés ont rassemblé près de 6 000 gènes principalement à partir d'ADNc de racines, nodules et mycorhizes en 2003-04 (El Yahyaoui et al. 2004; Manthey et al. 2004). Il faut attendre 2005 pour voir l'incorporation d'ADNc de fleurs et de gousses issus de différents stades de développement (Firnhaber et al. 2005). Deux puces ont pu être produites, la première de 16 500 oligonucléotides (70-mer) a été fabriquée par l'université de Bielefeld ; la seconde est une puce Affymetrix de 51 000 oligonucléotides (dont 19 000 dérivés du génome et 32 000 dérivés des EST) couvrant 70 % des gènes ou séquences codantes du génome. L'autre utilisation des EST et des ADNc pleine longueur concerne l'annotation des gènes, ainsi que la prédiction de protéines. Ils représentent donc une clef essentielle dans les approches protéomiques (Watson et al. 2003; Imin et al. 2005).

2.1.3.3 *Cartes physiques, génétiques et cytogénétiques*

Des banques génomiques en vecteur BAC ont été construites et couvrent physiquement les huit chromosomes de *M. truncatula* (Nam et al. 1999). En effet, le séquençage systématique des extrémités des BAC a permis le recouvrement d'une grande partie du génome et la production d'une carte physique (<http://www.genome.ou.edu/genome.html>). De nombreux établissements de recherche participent au développement de ces cartes physiques dans l'objectif de réaliser du clonage positionnel de gènes, de la génomique comparative des Légumineuses et le séquençage massif du génome de *M. truncatula* (Gualtieri and Bisseling 2002; Choi et al. 2004; Sato et al. 2007).

La création et l'étude de cartes génétiques permettent d'une part d'identifier des zones du génome intervenant dans les variations phénotypiques de caractères quantitatifs et ainsi colocaliser ces zones du génome avec des gènes candidats, et d'autre part, de tester la synténie entre les espèces. Les croisements sont réalisés entre lignées identifiées pour leurs caractéristiques phénotypiques contrastées et les marqueurs moléculaires développés permettent

d'ancrer les cartes génétiques entre elles et avec la carte physique (développée grâce au contigage de BAC).

Les premières cartes génétiques ont été réalisées sur la base de la carte élaborée sur *M. sativa* diploïde (Kiss et al. 1993). Elles ont été établies à partir d'une population F2 de 124 individus puis d'une population de 195 lignées recombinantes (RILs) F6 obtenues après le croisement entre les lignées Jemalong6 et DZA315.16 (Thoquet et al. 2002). Au total, 530 marqueurs ont été placés sur cette carte et définissent huit groupes de liaison ($2n = 16$) pour un total de 1 106 centiMorgans (Huguet et al. 2004). Une autre carte F2 fut produite par D. Cook (université de Davis, Californie) à partir d'un croisement entre les lignées Jemalong-A17 et A20 avec 300 marqueurs co-dominants cartographiés dans des EST (<http://mtgenome.ucdavis.edu>). A l'heure actuelle, cinq autres cartes génétiques sur des populations de RILs en F6 ont été produites à l'INRA pour lesquelles il existe des liens par le biais de microsatellites montrant du polymorphisme entre les lignées parentales (Nair et al. 2006). De plus, *M. truncatula* a une forte synténie avec la luzerne tétraploïde et diploïde et le pois (Julier et al. 2003; Choi et al. 2004; Aubert et al. 2006).

La notion de nouveaux marqueurs moléculaires « universels » (MTE) a été développée sur *M. truncatula* pour ancrer directement toute carte avec la carte physique (Huguet et al. 2007). Chaque locus correspond à un BAC dans lequel on définit différents marqueurs microsatellites permettant sa cartographie dans toute population. L'avantage que procurent ces marqueurs espacés régulièrement porte sur la localisation identique quelque soit le croisement et facilite ainsi la recherche de caractères quantitatifs et de gènes candidats entre différents croisements (Huguet et al. 2007). Ainsi, grâce aux données de séquençage des BAC de *M. truncatula*, les cartes ne sont pas seulement liées mais aussi ancrées avec la carte physique.

L'élaboration de cartes cytogénétiques donne des informations sur l'organisation du génome, la répartition des régions d'euchromatine et d'hétérochromatine ainsi que la position des régions centromériques. Elles facilitent le positionnement et l'ordonnancement des BAC sur les chromosomes (Kulikova et al. 2001). L'existence de cartes intégrées (génétique, physique et cytogénétique) chez *M. truncatula* accélère donc la localisation précise de gènes d'intérêt et peut faciliter le clonage positionnel.

2.2 Le processus de floraison et la date de floraison chez les dicotylédones annuelles

2.2.1 Physiologie

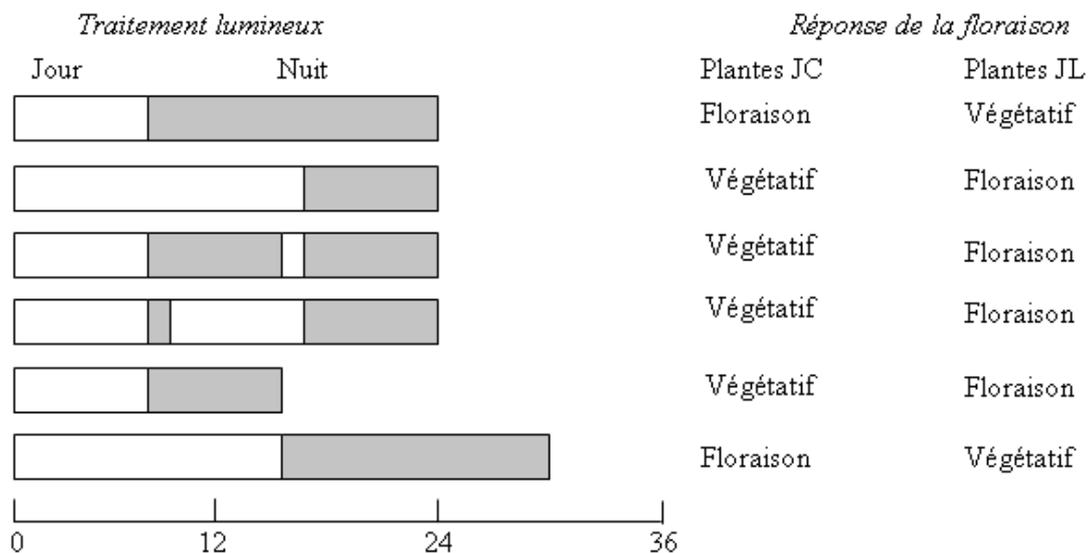
Une plante ne fleurit que si elle atteint un certain développement végétatif (maturité de floraison), avant lequel elle est dans une phase de juvénilité. Au delà de la phase de juvénilité, chez beaucoup d'espèces, l'aptitude à fleurir ne sera acquise qu'après l'intervention de facteurs externes comme la lumière et la température.

2.2.1.1 Photopériodisme : induction et rôle adaptatif

L'influence de la longueur du jour sur l'induction de la floraison fut découverte dans les années 1920 et fut qualifiée de photopériodisme (Garner 1933). L'étude du photopériodisme s'est surtout focalisée sur le passage de l'état végétatif à l'état floral, mais la photopériode influence aussi de nombreux autres aspects du développement des plantes, comme la tubérisation, la chute des feuilles et la dormance (Yakir et al. 2006).

Le photopériodisme est une réponse à la durée du jour et à l'alternance de périodes d'éclaircissement et d'obscurité. Les espèces sont classées en trois catégories selon leur réponse à la photopériode : les plantes de jours courts (JC) (*e.g. O. sativa*), les plantes de jours longs (JL) (*e.g. A. thaliana*) et les plantes photo-indifférentes (*e.g. Z. mays*) (Tableau 5) (Thomas and Vince-Prue 1997).

Tableau 5 Effet de la durée de la période obscure sur la floraison pour des plantes de jours courts (JC) ou longs (JL). Tiré de Thomas and Vince-Prue (1997).



Le besoin d'une photopériode peut être qualitatif auquel cas elle est absolument nécessaire, ou quantitatif, la photopériode favorable permet alors d'avancer la floraison. La photopériode critique détermine les limites supérieures (plantes de jours courts) ou inférieures (plantes de jours longs). La différence entre les espèces de JC et de JL est basée sur leur réponse à des jours plus courts ou plus longs que la période diurne critique. Cette période critique absolue varie d'une espèce à l'autre, elle peut être plus courte chez un génotype de JL que chez un génotype de JC. Les *Medicago* annuelles sont connues pour être des plantes de jours longs (Clarkson and Russell 1975). Cependant, chez *M. truncatula* la réponse de la floraison à la photopériode est mal connue.

Le passage réel d'une croissance végétative à une croissance reproductrice se déroule dans les zones méristématiques. Cependant, le signal photopériodique n'est pas perçu par les bourgeons méristématiques mais par les feuilles (Zeevaart 2006). Cette séparation des lieux de perception et de réponse suggère qu'un stimulus transmissible est nécessaire. Julius von Sachs (1832-1897) est le premier à concevoir l'hypothèse d'un stimulus hormonal qui a été ensuite appelé florigène en 1936 par Chailakhyan et dont la fonction est conservée entre les espèces (Lang 1965; Zeevaart 1976; Zeevaart 2006). Cependant, des études proposent que le florigène ne soit pas un constituant unique mais multifactoriel, la floraison étant induite par un ratio spécifique d'hormones connues et de métabolites (Bernier et al. 1993). Ainsi, le florigène est présenté comme un concept physiologique et non plus comme une simple entité chimique. Si la feuille perçoit le stimulus photopériodique alors elle est capable de mesurer le temps. Ainsi, le photopériodisme implique une famille de chromoprotéines : les phytochromes (Samach and Coupland 2000). Ceux-ci sont localisés dans le hyaloplasme des cellules végétales et sont capables de percevoir la lumière rouge clair et rouge sombre, et d'engendrer une réponse caractéristique en fonction de la longueur d'onde reçue.

Dans la nature, le photopériodisme procure aux plantes la garantie d'une floraison dans leur niche temporelle ou d'une réduction de la compétition avec d'autres espèces ayant un photopériodisme différent, ou que l'accomplissement de leur cycle de développement aille à terme, avant que ne se produisent les attaques hivernales ou les sécheresses estivales (Jeong and Clark 2005). Excepté aux tropiques, la longueur du jour est le témoin le plus sûr du changement de saison.

Ainsi, chez les plantes photosensibles, la réponse à la photopériode est plus marquée sous les latitudes subtropicales ou tempérées, là où les variations saisonnières de la durée du jour sont plus prononcées que dans les zones tropicales où l'insensibilité à la photopériode prédomine (Thomson et al. 2006). De plus, dans la mesure où la pollinisation croisée est indispensable ou avantageuse, une floraison synchronisée par la photopériode servirait à assurer un maximum de pollinisations ou à coordonner la floraison avec l'arrivée d'un insecte pollinisateur particulier (Yakir et al. 2006). La floraison ainsi que d'autres réponses peuvent se dérouler dans des limites assez larges de photopériode, et il a été prouvé que des populations étaient capables de s'adapter génétiquement aux latitudes, donnant naissance à des écotypes physiologiques. Chez de nombreuses espèces comme le sorgho (*Sorghum bicolor*), le chénopode (*Chenopodium*), l'oseille glaciaire (*Oxyria digyna*), la lampourde (*Xanthium*), le fraisier (*Fragaria vesca*), il existe des écotypes ou des variants photopériodiques qui se caractérisent par des photopériodes critiques différentes (Shitaka and Hirose 1998; Craufurd et al. 1999; Bertero et al. 1999; Heide 2005; Heide and Sonstebly 2007). Plus les individus sont récoltés sous des latitudes septentrionales, plus la longueur de la photopériode critique s'accroît.

2.2.1.2 Horloge circadienne

De nombreux aspects du comportement des plantes dont les nyctinasties (repliements diurnes des feuilles) et le photopériodisme sont soumis à des cycles fluctuants sur une période de 24 heures. Ces cycles, appelés rythmes circadiens (*circa*, environ- *dies*, jour), sont sous le contrôle d'un système interne de mesure du temps, appelé horloge biologique et mise en évidence dès 1729 par De Mairan (Hopkins 2003). Le rythme se manifeste spontanément en conditions naturelles ; il peut se maintenir pendant un certain temps en conditions constantes de lumière ou d'obscurité mais s'atténue et disparaît à long terme. En conditions naturelles, le rythme est initié toutes les 24 heures par la transition jour / nuit. (Bünning 1964) suggéra l'existence d'une alternance chez les végétaux de phase photophile et de phase scotophile. Dans la phase photophile (plante sensible), la lumière a un effet positif, alors que dans la phase scotophile, la lumière a un effet négatif ou neutre sur la floraison. Il proposa également que la réponse des plantes à la lumière dans le cadre du photopériodisme était reliée au moment de son application par rapport à ces phases d'un rythme endogène. Une expérience chez le soja (plante de JC) démontre effectivement l'existence d'un rythme de sensibilité à la lumière dans le cas de la floraison.

Il y a, en effet, perception du stimulus lumineux et naissance d'un signal lequel n'est pas lu pour la réponse florale à certaines heures du cycle photopériodique (Brest et al. 1971).

2.2.1.3 *Influence de la température*

La température (son énergie thermique) a des effets essentiels sur la croissance et sur le développement des végétaux. La croissance constitue l'aspect quantitatif : augmentation de masse, de volume, de taille de la plante alors que le développement représente l'aspect qualitatif : la mise en place d'organes nouveaux. La croissance des plantes ne peut avoir lieu que si elles reçoivent un minimum d'énergie thermique. L'élévation de la température accélère la croissance jusqu'à un optimum, puis la ralentit jusqu'à un maximum au-delà duquel les plantes périssent. Par ailleurs, l'optimum thermique varie au cours du développement.

Si les basses températures ralentissent les phénomènes physiologiques, elles n'en sont pas moins parfois nécessaires à la levée d'inhibitions comme les dormances végétatives. Certaines plantes nécessitent donc de basses températures au début de leur développement ou lorsqu'elles ont dépassé leur stade juvénile, pour élaborer leur aptitude à fleurir.

Un traitement par le froid peut rendre possible ou accélérer la floraison. Par exemple, le fait que le pois (*P. sativum*) soit vernalisé à l'état de semences pendant quelques jours (3 – 5 °C) réduit le nombre de nœuds nécessaires avant l'apparition des premières fleurs (Highkin 1956). Les plantes dépendantes de la vernalisation peuvent accomplir celle-ci, pour certaines déjà à l'état de semences à peine germées (*e.g.* les céréales d'hiver), voire parfois immatures (*e.g.* le seigle « Petkus »), pour d'autres seulement à l'état de pousses plus ou moins avancées (*e.g.* les bisannuelles) (Purvis and Gregory 1937; Wellensiek 1963).

Le zéro de végétation (variable selon les espèces) est la température au dessus de laquelle la croissance est positive (au-dessous, la croissance est nulle). Au-delà du zéro de végétation, la température explique la croissance de la plante. Ainsi, la durée de la période de croissance est essentiellement fixée par un certain nombre de degrés-jour. Plus la température est élevée et plus le nombre de degrés-jour nécessaire pour un stade physiologique (*e.g.* la floraison) sera rapidement atteint. Le calcul de la somme de degré-jours est donc le meilleur moyen pour tenir compte de l'effet de la température (Durand et al. 2007).

L'action accélératrice de la température sur les processus biologiques peut être conçue à la base de deux hypothèses (Belehradek 1932; Johnson and Thornley 1985) :

- (i) chimique : la température accélère les réactions biochimiques (enzymatiques) ;
- (ii) physique : la température abaisse la viscosité du protoplasma, favorise les transports moléculaires (diffusion protoplasmique) et accélère alors les réactions biochimiques.

Chez les Légumineuses annuelles (*Trifolium subterraneum*, *M. truncatula*, *M. scutella*, *M. polymorpha*, *M. littoralis*, *M. tornata*, *M. rugosa*), la date de floraison est décrite comme étant contrôlée par trois facteurs (Aitken 1955; Clarkson and Russell 1975) : 1- la longueur du jour, 2- la vernalisation (inférieure à 5 °C) et 3- la température post-vernalisation (supérieure à 10 °C). Des plantes non vernalisées de *M. truncatula* ont un développement abondant de ramifications et une croissance procombante des tiges jusqu'à la floraison (5 à 6 semaines après la germination). En chambre de culture, les observations ont permis de montrer qu'une longue vernalisation suivie d'une haute température post-vernalisation (16 - 30 °C) combinée à une longue photopériode (18 - 24 h) réduisent le délai d'apparition de la première fleur (Aitken 1955; Clarkson and Russell 1975; van Heerden 1984; Hochman 1987).

Les différents types de réponses à la photopériode peuvent être modifiés par les conditions de l'environnement comme la température (Salisbury 1963; Thomas and Vince-Prue 1997). Ainsi, chez la plupart des plantes de jour long, l'abaissement de la température au cours de la période obscure réduit l'effet inhibiteur de cette dernière et permet ainsi son extension, tandis que l'élévation de température augmente l'inhibition. Ainsi, la durée maximale critique de la période obscure peut être augmentée d'une heure chez *A. thaliana* en abaissant la température nocturne aux environs de 0 °C (Laibach 1951). Chez d'autres espèces, au contraire, comme le trèfle (*Trifolium pratense*), l'élévation de température nocturne favorise l'induction florale (Wellensiek 1966). Cependant, chez *M. truncatula*, dans des conditions de culture en serre avec des changements de saison en termes de durée de jour et de température, l'effet de la vernalisation sur la date de floraison est moindre que l'effet de la photopériode. D'autre part, alors que l'effet de la vernalisation sur la date de floraison d'un génotype est important en jours longs, il peut être supplanté par l'effet de la photopériode chez d'autres génotypes (Chabaud et al. 2006). Il existe donc probablement un effet génotype en réponse à la vernalisation et la photopériode.

2.2.1.4 Modèles physiologiques de la réponse aux facteurs du milieu

La modélisation permet d'améliorer la compréhension du processus de floraison mais aussi de prédire le comportement de la plante en intégrant les effets du milieu. Les paramètres de modèles sont aussi des caractères dont on peut analyser le déterminisme génétique (Tardieu 2003). La possibilité de créer des modèles mathématiques de la réponse de la croissance et de la floraison à divers facteurs environnementaux (température, photopériode) a été exploitée dès 1985 sur des Légumineuses annuelles (Roberts et al. 1985; Roberts et al. 1996). Un modèle empirique de régression linéaire multiple a été ainsi proposé pour relier le taux de développement à la température et la photopériode :

$$1/f = a + bT + cP \quad (1)$$

où $1/f$ correspond au taux de développement vers la floraison (inverse du temps nécessaire exprimé en jour depuis l'émergence jusqu'à la floraison) et a , b , c sont des coefficients empiriques de régression, T la température moyenne et P la photopériode moyenne depuis la levée jusqu'à la floraison. Ce type de modèle additif a été utilisé pour la modélisation de la floraison chez le quinoa (*Chenopodium quinoa*) et le pois chiche (*Cicer arietinum*) mais n'a jamais été appliqué chez *M. truncatula* (Roberts et al. 1985; Roberts et al. 1996; Bertero et al. 1999; Soltani et al. 2006).

D'autres modèles plus complexes ont également vu le jour dont un modèle empirique chez *M. truncatula*, sur le cultivar Jemalong (van Heerden 1984) :

$$y = -0.00496 - 0.00691Tm + 0.00036Tm^2 - 0.02023\log P \quad (2)$$

où y est l'inverse du nombre de jours de la levée à la floraison, ce qui équivaut au taux de développement du modèle précédent, Tm est la température minimale moyenne et $-\log P$ est l'opposé du logarithme de la photopériode moyenne.

Récemment, des modèles décrivant l'apparition du premier nœud reproducteur (1RN) (1) et la date de floraison de la première tige primaire (DBF) (2) chez le cultivar Jemalong en réponse à la durée de vernalisation (DV) ont été développés (Moreau et al. 2007) :

$$1RN = a (1 + b * r^{DV}) \quad \text{où } a = 2.80 ; b = 5.40 ; r = 0.77 \quad (3)$$

$$DBF (°C.J) = DA + [(1RN + 2) * RLa^{-1}] \quad (4)$$

où DA représente la date d'apparition de la première tige primaire (en °C.J avec une température de base de 5°C) ; (1RN + 2) comprend le nombre de feuilles et les deux cotylédons ; RLa^{-1} est le temps nécessaire pour l'apparition de chaque feuille (°C.J) (phyllochrone). Ce modèle ne prend pas en compte l'impact de la photopériode et nécessite la notation de la date d'apparition de la première tige primaire pour prédire le début de la floraison.

Ces différents modèles ont chacun leurs avantages (réalité physiologique des résultats, applications futures, ...) et leurs limites (possibilité et qualité de la simulation, intervalle de confiance des paramètres, la prise en compte de la variabilité inter-plantes, ...) dans les applications. Aucun n'a été généralisé à plusieurs génotypes, alors même qu'on attend que certains paramètres soient génétiquement déterminés, comme cela a été montré chez *A. thaliana* (Nielson et al. 1976).

2.2.1.5 *Relation entre date de floraison et processus de développement*

L'environnement influe sur la croissance et l'architecture des plantes. L'organogenèse continue, spécifique du règne végétal, permet de répondre à ces variations de l'environnement (Huber et al. 1999). Typiquement, chez les plantes à fleurs, le méristème apical racinaire donne naissance à l'appareil racinaire, tandis que le méristème apical caulinaire (MAC) donne naissance à toute la partie aérienne. Le MAC est caractérisé par un développement de type itératif, au cours duquel sont produites des unités séquentielles, appelées phytomères. Un phytomère comporte une feuille, portant éventuellement à son aisselle un méristème axillaire, l'ensemble constituant un nœud, soutenu par une portion de tige, appelée l'entre-nœud. Au cours du développement de la plante des changements progressifs de l'identité du méristème interviennent.

Lors de la transition florale, le MAC végétatif formé au cours de l'embryogenèse devient un méristème d'inflorescence produisant des feuilles modifiées, ou feuilles caulines, et des méristèmes floraux latéraux. Ces méristèmes floraux (au contraire des méristèmes végétatifs et d'inflorescence) ont une croissance définie ou déterminée, puisque leurs activités cessent après la production d'un nombre défini d'organes. L'activité des méristèmes aériens détermine ainsi la base de l'architecture de la plante :

- (i) lorsque seul le méristème apical primaire est actif, l'architecture générale de la plante est linéaire ;
- (ii) si des méristèmes axillaires sont activés, des hampes florales secondaires sont formées, donnant naissance à une architecture branchée plus complexe.

Le développement des ramifications peut être considéré comme le résultat d'un recrutement au sein d'une banque de bourgeons produite et portée par la plante. Pour une plante isolée, ce recrutement est le résultat de compromis entre les facteurs génétiques, les relations à plus ou moins courte distance au sein de la plante et l'action des facteurs abiotiques. Le fonctionnement de cette banque de bourgeons a des conséquences directes sur l'aptitude de la plante à produire fleurs, fruits et graines (Harper 1977; Jongejans et al. 2005).

Les modèles décrivant la date de floraison vus précédemment relient le taux de développement, équivalant à $1/f$ (où f représente la date de floraison) à une température et une photopériode moyennes. Dans un modèle de type multiplicatif, ce taux de développement peut être expliqué comme une fonction de température (T) et de photopériode (PP) (Hammer et al. 1989):

$$R_t = R_{\max} f(T) f(PP) \quad (5)$$

Où R_t est le taux de développement au jour t ; R_{\max} le maximum de développement aux températures et photopériodes optimales et le produit $f(T) f(PP)$ constitue un « jour physiologique » (PD). La floraison est donc prédite pour se produire lorsque :

$$\sum_t R_t = 1 \text{ soit } 1/R_{\max} = \sum_t (PD) = \sum_t f(T) f(PP) \quad (6)$$

Expérimentalement, une relation entre développement végétatif et date de floraison a été constatée. Chez le pois, il existe une relation entre la date de floraison, le numéro du premier nœud florifère et la hauteur de la plante. Les génotypes ayant une vitesse d'apparition des feuilles élevée fleurissent plus rapidement que les génotypes dont le taux d'apparition des feuilles est plus faible. De plus, quand ce taux ne diffère pas, la date de floraison est en relation avec la position du premier nœud reproducteur (Dumoulin et al. 1994).

D'un point de vue génétique, la date de floraison chez le pois semble être déterminée par le gène majeur LEAFY contrôlant le nœud de la première fleur chez huit génotypes parentaux (Bourion et al. 2002). Chez *M. truncatula*, la colocalisation sur les cartes génétiques a été observée entre des locus des caractères quantitatifs de la date de floraison et de la croissance (vitesse d'élongation et longueur des tiges, nombre de tiges primaires, ratio entre poids de feuilles et de tiges) suggérant implicitement un contrôle par des gènes communs ou étroitement liés (Julier et al. 2007).

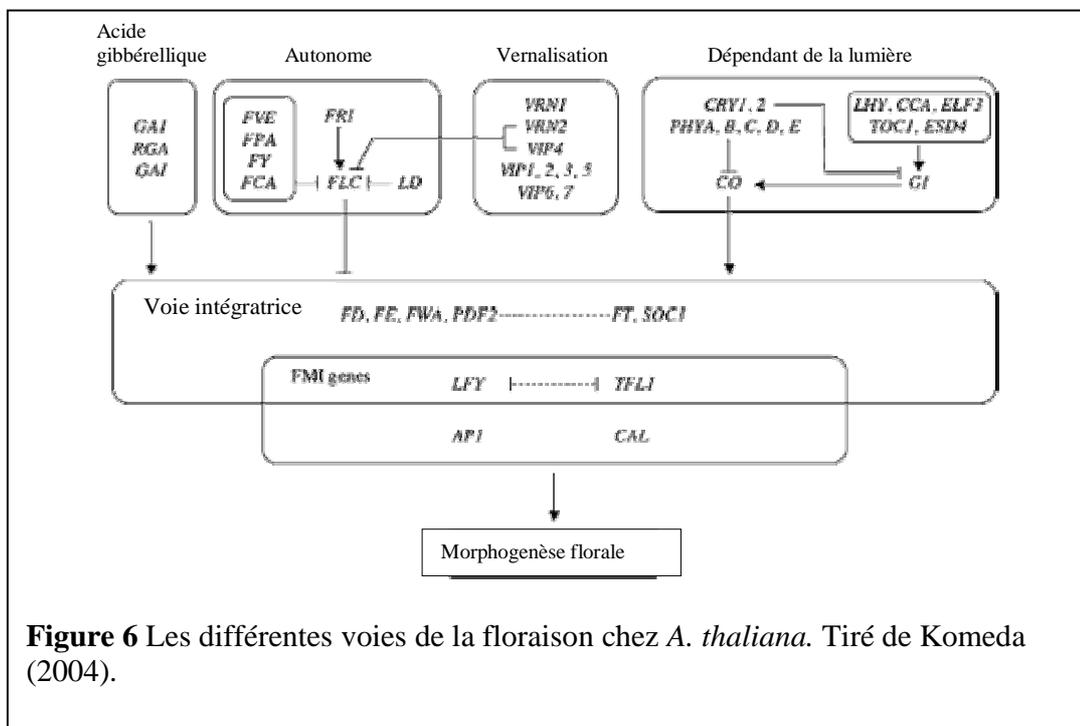
2.2.2 Contrôle génétique de la date de floraison

2.2.2.1 Gènes connus et leurs interactions chez *A. thaliana*

Quand *A. thaliana* a atteint un certain développement (phase juvénile), des signaux de l'environnement déclenchent la transition vers le développement floral. Le méristème caulinaire arrête de produire des primordia foliaires et se met à produire des primordia d'organes floraux : pétales, sépales, étamines, carpelles (Vernoux et al. 2000). Les nombreuses études de génétique sur le développement de la fleur menées principalement sur *A. thaliana* ont conduit à conclure que ce processus de développement est déterminé par un réseau très complexe et conservé de gènes opérant en cascade et qui s'expriment dans différents territoires spécifiques (Komeda 2004; Kramer and Hall 2005). La différenciation est, en effet, une expression spatio-temporelle particulière de gènes spécifiques impliqués dans l'élaboration de la floraison, de l'identité du méristème floral et de son développement, et enfin dans la construction des organes floraux.

Pour la plante, le passage à l'état reproductif requiert de nombreuses ressources. Komeda (2004) a émis l'hypothèse que la plante prépare ses ressources pour la floraison et que cette floraison est un processus développemental répressible. Ainsi, les gènes de floraison tardive sont représentatifs de cette fonction développementale répressible (Koornneef et al. 2004).

Les études de relations épistatiques entre les gènes de floraison ont permis l'élaboration d'un modèle physiologique reposant sur quatre voies de réponse contrôlant la floraison chez *A. thaliana* (Figure 6) (Mouradov et al. 2002; Corbesier and Coupland 2005; Imaizumi and Kay 2006).



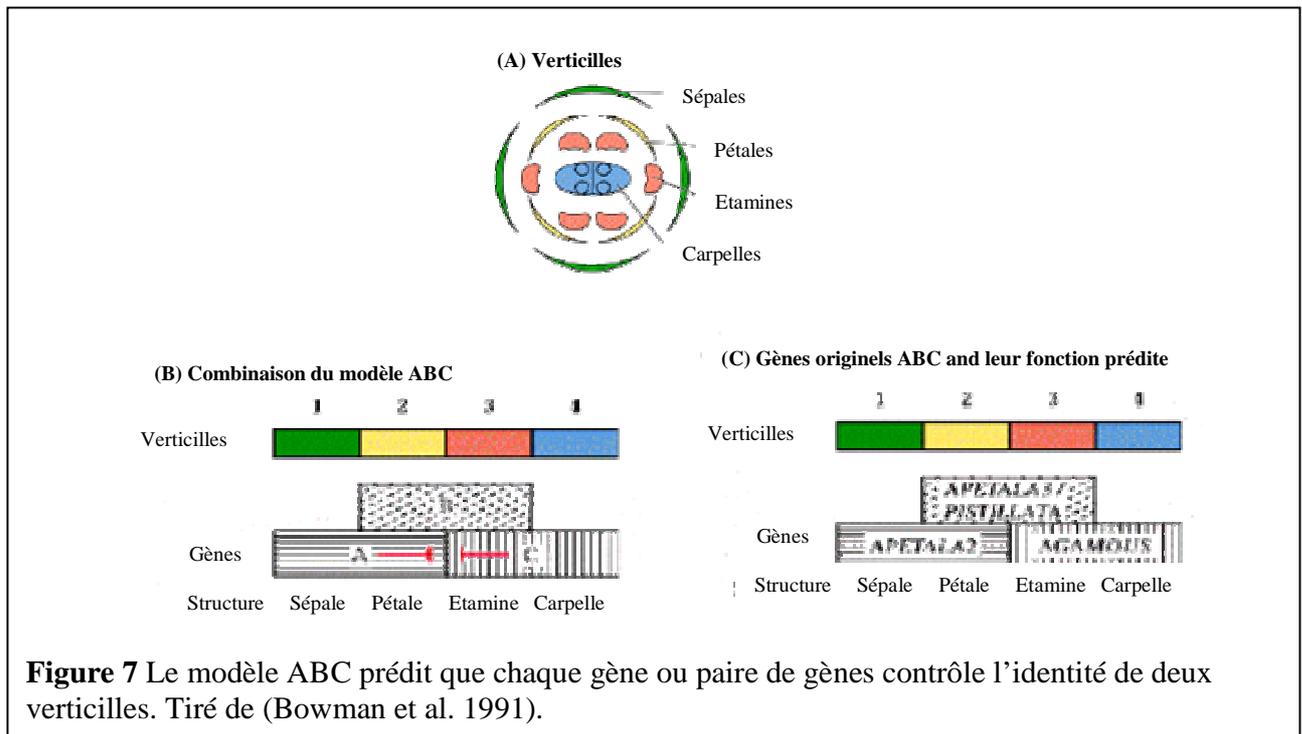
- (i) la voie dépendante de l'acide gibbérellique (AG), où AG est un activateur de la floraison chez les plantes de jours longs d'*A. thaliana*. Des mutants *gal* d'*A. thaliana* ne fleurissent jamais en conditions de jours courts et ont perdu l'activité du gène intégrateur LEAFY (LFY) (Blazquez et al. 1998). Il existe de nombreux gènes pour la synthèse et la perception de GsA. Ainsi GA active les gènes REGULATOR OF GIBBERELLIC ACID (RGA) et GIBBERELLIC ACID INSENSITIVE (GAI) qui restituent le phénotype des mutants *gal*, impliquant une autorégulation de cette voie (Dill et al. 2004) ;

- (ii) la voie autonome car les plantes requièrent non seulement des facteurs externes (environnementaux) mais aussi des facteurs internes (développementaux) pour activer la floraison. Le gène FRIGIDA (FRI) qui est un régulateur positif du répresseur de floraison FLOWERING LOCUS C (FLC) est responsable d'une partie des différences de date de floraison entre les écotypes d'*A. thaliana* (Johanson et al. 2000). Les gènes FLOWERING TIME CONTROLE IN ARABIDOPSIS (FCA) et LUMINI DEPENDENS (LD) répriment l'expression du gène FLC. La présence de nombreux gènes homologues (FVE, FPA et FY qui ont une fonction similaire au gène FCA) permet d'assurer le bon déroulement de la floraison (Page et al. 1999; Schomburg et al. 2001) ;
- (iii) la voie dépendante de la vernalisation où deux catégories de gènes ont été identifiées : VERNALIZATION (VRN1 à VRN5) et VERNALIZATION INDEPENDENT PLANT (VIP1 à VIP7) : VRN2 et VIP4 sont des répresseurs de FLC, induisant des liens entre voie autonome et voie dépendante de la vernalisation (Sheldon et al. 2000) ;
- (iv) la voie dépendante de la lumière rouge est captée par les protéines phytochromes codées par le gène PHYTOCHROME A (PHYA) ; les récepteurs de la lumière bleue sont des protéines cryptochromes codées par les gènes CRYPTOCHROME (CRY) (Lin et al. 1998; Briggs et al. 2001). PHY et CRY répriment les gènes inducteurs de la floraison CONSTANS (CO) et GIGANTEA (GI) (Mockler et al. 1999). Les gènes LATE ELONGATED HYPOCOTYL (LHY), CIRCADIAN CLOCK ASSOCIATED (CCA1), EUCALYPTUS LEAFY 3 (ELF3) et TIMING OF CAB EXPRESSION 1 (TOC1) transforment le signal physique pour le transmettre au gène GI qui activera CO (Doyle et al. 2002). CO est le gène intégrateur de la voie dépendante de la lumière.

Les quatre voies décrites convergent vers quelques gènes intégrateurs où le gène LFY joue un rôle de premier plan (Borner et al. 2000). Il régule la transcription de APETALA 1 (AP1), AP3, AGAMOUS (AG). LFY cause un blocage de la transition du méristème végétatif en méristème floral, les fleurs étant remplacées par des pousses feuillées. La protéine codée par le gène LFY serait un facteur de transcription car elle contient une région riche en proline et des régions acides et basiques caractéristiques des facteurs de transcription.

L'expression du gène est détectable avant apparition des primordia floraux et s'arrête quand les quatre types d'organes sont en place (sépalés, pétales, étamines et carpelles). LFY est donc le déclencheur du développement floral mais c'est le gène FT qui détermine le moment de l'initiation florale. Ce gène FT requiert la présence de CO en jours longs pour une régulation positive (Samach et al. 2000). CO interagit directement avec FT. L'expression de FT et SOC1 est contrôlée positivement par la voie dépendante de la lumière et la voie autonome (après inhibition de FLC) (Lee et al. 2000). Le signal de la vernalisation augmente l'expression de SOC1 qui est alors autorégulée par la voie dépendante des acides gibbérelliques (Borner et al. 2000). Ainsi FT, LFY et SOC1 sont les gènes intégrateurs de la floraison chez *A. thaliana* (Lee et al. 2000; Blazquez 2000).

Pour la morphogenèse florale, un modèle physiologique appelé « ABC » explique l'identité et le positionnement des organes floraux qui sont issus du méristème floral (Bowman et al. 1991; Komeda 2004). Ce modèle est basé sur trois classes de gènes A, B et C, qui sont exprimés dans le méristème floral. La fonction A est liée aux gènes de l'identité méristématique florale (LFY, AP1, CAL) alors que la fonction C conditionne la croissance déterminée ou indéterminée du bourgeon. L'expression de A et C uniquement engendre une cascade d'activation de gènes, ce qui mène respectivement à la formation de sépalés et de carpelles. L'expression de B se réalise toujours de façon combinée avec A ou C. Ainsi, l'expression de ces gènes en combinaison avec B résulte en la formation de pétales et d'étamines. Les fonctions des gènes A+B et B+C sont conservées dans la plupart des espèces mais les interactions incluant spécifiquement les fonctions A ou C sont plus complexes. Par conséquent, les mutations des gènes d'identité florale causent le développement des organes floraux dans des zones inappropriées. L'expression spécifique en temps opportun de ces régulateurs est donc critique au développement normal de la fleur (Figure 7).



2.2.2.2 La variabilité allélique naturelle connue

Chez *M. truncatula*, il existe peu d'études portant sur la variabilité allélique. Cependant, le programme EAGLE (*Ecological and Association Genomics in Legumes*) coordonné par l'INRA de Montpellier en 2005 a initié le séquençage de la variabilité naturelle disponible chez *M. truncatula* et chez le pois. Sur un ensemble établi de génotypes, l'objectif est de séquencer un grand nombre de fragments génomiques (environ 200 fragments de 600 - 800 nucléotides). Ces fragments sont choisis dans des exons ou introns de gènes candidats pour des phénotypes d'intérêts agronomiques tels que la floraison, permettant des études d'associations génotypes / phénotypes.

Différentes études ont été menées chez *A. thaliana* sur la variabilité allélique de gènes impliqués dans la floraison. Ainsi, la variation allélique naturelle a été analysée sur le gène *CRY2*^{EDI} (Olsen et al. 2004). Six allèles, dont l'allèle cloné *CRY2*^{EDI} de précocité de floraison en réponse à des jours courts, ont été étudiés sur 100 accessions. *CRY2*^{EDI} s'est avéré présent à une faible fréquence, spécifique d'un génotype, suggérant que cet allèle est géographiquement très localisé (El-Assal et al. 2001).

Le gène FRI détermine une floraison tardive chez *A. thaliana*, qui ne peut être rendue possible qu'après l'action antagoniste du froid hivernal via la voie de la vernalisation. Une étude menée sur 25 écotypes pour la diversité allélique du gène FRI montre un grand nombre d'haplotypes (20) et une faible diversité nucléotidique. Les auteurs émettent l'hypothèse d'une expansion démographique rapide et d'une forte pression de sélection qui a permis de maintenir ces mutations (Le Corre et al. 2002). Près de 40 % des écotypes d'*A. thaliana* possèdent un gène FRI non fonctionnel, ce qui les rend capables de fleurir rapidement à n'importe quelle période de l'année (Johanson et al. 2000). Cette perte de fonction est causée par différentes mutations spontanées, causant un codon STOP.

L'étude des relations entre la diversité moléculaire et la variation de caractères quantitatifs cartographiés peut être utilisée en vue de la validation génétique d'un gène candidat (Buckler IV and Thornsberry 2002).

2.2.3 Rôle de la floraison dans l'adaptation et la valeur agronomique

L'apparition des angiospermes (*aggeion*, réceptacle – *sperma*, graines) ou plantes à fleurs dès le début du Jurassique (environ 150 millions d'années) a représenté un succès évolutif qui, par radiation adaptative, a permis aux angiospermes de coloniser tous les biotopes. Bien que le processus de floraison et son contrôle soient encore plus anciens (apparition des gymnospermes au Carbonifère, vers 300 millions d'années), différents types d'innovations ont en effet émergé, comme la formation d'un fruit entraînant une amplification des possibilités de dissémination des graines par les animaux, les phénomènes d'auto-incompatibilité entraînant un brassage génétique, et enfin l'accumulation des réserves dans la graine pour la germination plus efficace que chez les gymnospermes (Bateman et al. 2006; Feild and Arens 2007). C'est pourquoi les angiospermes représentent environ 250 000 espèces contre 600 chez les gymnospermes. Après la germination et le développement végétatif, la floraison représente une transition particulièrement complexe dans le cycle de développement.

Les caractères sexuels sont importants au regard des processus d'adaptation et d'évolution car les gènes influençant la reproduction gouvernent à la fois leur propre transmission mais aussi ceux de l'organisme tout entier. Les différentes observations réalisées révèlent une grande diversité de systèmes de pollinisation (allogamie associée ou non avec auto-incompatibilité ou autogamie) coexistant la plupart du temps dans des conditions écologiques similaires. Les variations dans l'architecture florale impliquent en effet des décisions hiérarchiques sur l'allocation optimale des ressources aux développements des organes mâles et femelles (Delph 1990). De plus, du fait de leur immobilité, les plantes requièrent des vecteurs de transfert de pollen provoquant des adaptations florales diverses pour s'associer à l'agent pollinisateur. Les plantes ont donc mis en place pour la floraison différents types de réponses aux conditions environnementales (Delph 2003). Malgré un système de reproduction très autogame de l'espèce *M. truncatula* (environ 95 %), certaines populations maintiennent des diversités importantes soulevant l'hypothèse d'événements d'allogamie impliquant des différents génotypes, suivie d'autofécondations (Bonnin et al. 2001).

L'époque de floraison joue un rôle important sur la régularité de la production car les conditions climatiques influencent directement ou indirectement, à ce stade, les diverses composantes de l'élaboration de la production en fruits et du remplissage des graines (sensibilité au gel, concordance de floraison, pollinisation et fécondation). Une floraison précoce peut exposer au gel les stades sensibles des organes de fructification alors qu'une floraison tardive génèrera un risque de stress hydrique pour la fructification (Totland 1999; Sparks et al. 2000; Williams and Abberton 2004). La date de floraison est également d'une importance majeure pour l'adaptation des cultures aux conditions environnementales. La précocité de floraison conditionne en effet directement les zones de culture possibles de nombreuses espèces et constitue un critère majeur de sélection en amélioration végétale. Un bon exemple est celui du maïs, une plante d'origine tropicale pour laquelle la sélection de variétés précoces, moins exigeantes en sommes de températures élevées pour leur développement, a rendu possible la culture dans les régions tempérées. Ainsi, sur le seul territoire français, les variétés de maïs sont réparties en sept groupes de précocité qui déterminent les zones où ces variétés peuvent être cultivées. Pour le pois protéagineux, les variétés tardives permettent une meilleure adaptation de la culture aux climats froids par évitement des gelées de fin de printemps.

2.3 Recherche de zones du génome et des gènes impliqués dans le déterminisme d'un caractère quantitatif

2.3.1 Construction de cartes génétiques

2.3.1.1 Critères de construction

L'élaboration d'une carte génétique consiste à ordonnancer des marqueurs définissant chacun un locus unique. Ces cartes indiquent les positions relatives des marqueurs les uns par rapport aux autres grâce à l'analyse statistique de ségrégation des marqueurs au cours des générations. La construction d'une carte de liaisons génétiques s'appuie sur l'estimation du taux de recombinaison entre tous les locus pris deux à deux. Deux marqueurs sont d'autant plus difficilement séparables qu'ils sont proches l'un de l'autre. Ces marqueurs doivent être polymorphes (présenter différents allèles identifiables) et peuvent être localisés dans un exon (variations phénotypiques parfois détectables) ou (plus couramment) dans un intron (variations détectables au niveau de l'ADN). Les distances se mesurent en centiMorgans (cM).

La première étape de ce jeu de construction nécessite donc de déterminer le contenu allélique des marqueurs (*i.e.* le génotype) pour chacun des individus de la population de cartographie. Ainsi les principaux critères de l'élaboration d'une carte génétique sont :

- (i) développer la population de cartographie appropriée (descendance simple en ségrégation souvent issue d'un croisement biparental entre deux lignées pures : F₂, lignées recombinantes, back-cross, ...) et décider de sa taille (50 à plusieurs milliers d'individus) (Mohan et al. 1997) ;
- (ii) choisir le type de marqueurs moléculaires pour génotyper la population de cartographie ;
- (iii) tester les parents avec les marqueurs moléculaires pour rechercher du polymorphisme et passer ces marqueurs sur l'ensemble de la population de cartographie ;
- (iv) exécuter des analyses de liaisons (calculer les fréquences de recombinaison entre marqueurs, établir des groupes de liaison, estimer les distances sur la carte et déterminer l'ordre de la carte).

2.3.1.2 *Cartographie fine*

Les effectifs des descendance classiquement utilisées ne permettent pas d'espérer une précision de la position des QTL d'effet modéré inférieure à 15 ou 20 cM (De Vienne 1998). De tels segments peuvent comprendre plusieurs centaines de gènes, ce qui est rédhibitoire pour toute tentative de caractérisation ou de clonage positionnel de QTL.

L'utilisation d'une carte très dense de marqueurs établie sur un très grand nombre d'individus et l'étude des relations entre le polymorphisme de ces marqueurs et le caractère pourront permettre de cartographier finement le QTL. Ainsi, la cartographie fine consiste à rechercher et ordonner des marqueurs localisés dans un très petit intervalle autour du locus d'intérêt. La cartographie fine se fera à partir de lignées quasi-isogéniques, populations de lignées recombinantes ou de plantes F2 (Michelmore et al. 1991). Le nombre d'individus de la population à analyser pour construire une carte génétique fine dépend du taux de recombinaison génétique autour du locus cible. Une fois identifiées, les plantes recombinantes autour du gène cible sont génotypées pour tous les marqueurs de la région, qui seront alors ordonnés précisément. Cette analyse permettra le développement d'une carte génétique fine.

2.3.2 Cartographie de locus à effets quantitatifs

La génétique quantitative constitue une alternative à la génétique inverse pour établir des liens entre une fonction et un ou des gènes. En effet, la génétique inverse n'est applicable qu'à un gène unique. Or la plupart des caractères importants au niveau agronomique sont soumis à des variations continues et sont contrôlés par plusieurs locus. La cartographie de QTL responsables de la variation de ces caractères permet de disséquer leur architecture génétique. Les principes de base de détection des QTL ne seront pas détaillés ici (se reporter par exemple à de Vienne 1995). La caractérisation phénotypique de plantes utilisées pour la cartographie fine permet de localiser précisément le gène cible.

2.3.2.1 *Avantages et inconvénients des descendances utilisées*

La puissance des tests dans la recherche de QTL sera fortement déterminée par la taille de la population (Charcosset and Gallais 1996). Il sera préférable, à moyens d'évaluation phénotypique constants, de privilégier le nombre d'individus étudiés, relativement au nombre de répétitions de ces individus. Les populations en ségrégation les plus efficaces pour cartographier les locus responsables des variations sont composées de lignées fixées. L'établissement de ces populations permanentes ou « immortelles » est obtenu par autofécondations successives d'individus F2. A chaque génération, un individu est choisi, qui sera à l'origine de la génération suivante. Ces populations de lignées offrent plusieurs avantages :

- (i) le taux d'hétérozygotie diminue à chaque génération (il n'est plus que de 3% en génération F6) ;
- (ii) après plusieurs autofécondations, tous les individus d'une lignée n sont génétiquement identiques entre eux et identiques à la lignée $n-1$ (à l'hétérozygotie près). Par le jeu des méioses successives, chaque lignée correspond à une combinaison particulière d'allèles (lignées recombinantes ou Recombinant Inbred Lines, RILs).

Dans les descendances back-cross (BC), haploïdes doublées (HD) ou RILs, contrairement aux F2, il n'y a que deux moyennes à comparer (homozygote A contre homozygote B) pour effectuer une recherche QTL. Mais ces descendances ne permettent pas d'estimer tous les effets. Ainsi, dans le cas des HD et des RILs, la valeur de l'effet additif peut être estimée mais pas la dominance. Chez les BC, où l'on compare un génotype homozygote avec un hétérozygote, l'effet estimé ne sera égal à l'effet additif que sous l'hypothèse d'absence de dominance. Chez les RILs, la recombinaison entre le marqueur et le QTL fait diminuer l'effet additif plus rapidement qu'avec les autres types de descendance. Pour une population F2, bien que les individus hétérozygotes au locus étudié (50 % de la population totale) ne soient pas informatifs pour estimer l'effet additif, ils permettent la création de population de taille importante assez rapidement (utilisable en cartographie fine). Cependant, à taille de population égale, la puissance des tests effectués à partir de descendances de HD ou de RILs devrait donc être nettement supérieure, d'autant que plusieurs individus peuvent être analysés pour chaque génotype (De Vienne 1998).

2.3.2.2 *Analyses QTL sur des populations multiples*

Ayant obtenu des QTL dans plusieurs populations et/ou avec des conditions variables, il convient de chercher à intégrer l'ensemble de ces résultats. Tout d'abord proposée pour intégrer et résumer les résultats obtenus de différentes études animales, ce type de méta-analyse multi-croisements a été adapté avec succès à la génétique et aux études QTL (Goffinet and Gerber 2000; Khatkar et al. 2004; Jourjon et al. 2005), tant dans le domaine animal que végétal (Belknap and Atkins 2001; Hanocq et al. 2007). Plusieurs outils statistiques (logiciels) ont été développés et sont disponibles pour tester si des QTL sont communs et, le cas échéant, de préciser leur localisation : Biomecator, MCQTL et Meta-QTL (Arcade et al. 2004; Jourjon et al. 2005; Veyrieras 2005; Veyrieras et al. 2007).

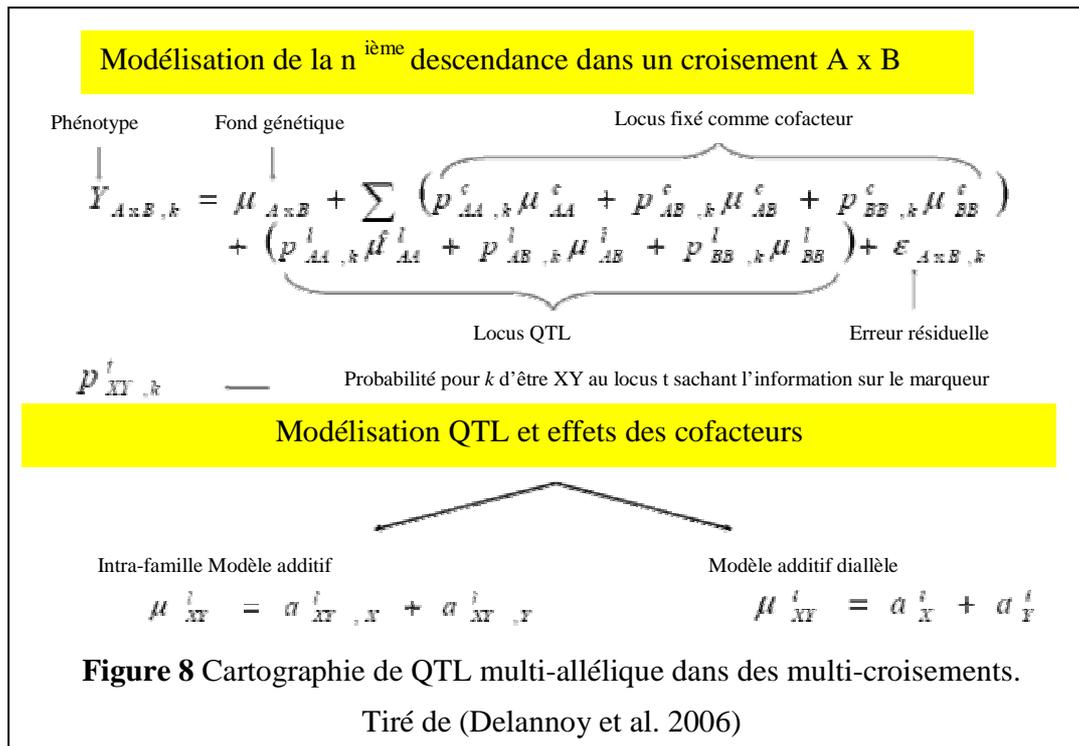
L'utilisation de Biomecator nécessite des études indépendantes de populations pour réaliser les projections des cartes et QTL avant de procéder au calcul de QTL communs. Ce logiciel permet de réaliser une méta-analyse compilant les données de cartes génétiques et de détection QTL. La méthode se base sur le critère Akaike qui estime combien de QTL différents sont représentés par le QTL détecté dans différentes expériences (Goffinet and Gerber 2000).

MCQTL et Meta-QTL regroupent les QTL détectés dans une même région du génome. MCQTL prend en compte des connexions possibles entre populations et permet le calcul de l'effet de chaque allèle. Cette méta-analyse multi-croisements lie les populations en supposant que la localisation des QTL est la même dans toutes les populations. Ceci permet :

- (i) de trouver des QTL communs à plusieurs populations d'une même espèce ou de différentes espèces sur la base de marqueurs communs avec pour objectif la réduction de l'intervalle de confiance du QTL étudié ;
- (ii) d'analyser la conservation de la synténie entre différentes espèces pour la prédiction de gènes candidats.

Le modèle mis en œuvre par MCQTL est un modèle de régression linéaire dans lequel les probabilités des QTL sont calculées à partir des données multiples sur les marqueurs (Figure 8) (Delannoy et al. 2006).

La première étape d'une méta-analyse multi-croisements consiste à réaliser une carte consensus à partir des marqueurs communs tout en fournissant les possibles connections des parents de populations de cartographie ainsi que le type de populations (F2, F3, RILs). Dans un second temps, il faut charger les données de génotypes et phénotypes pour enfin lancer la procédure d'analyse selon des seuils de détection choisis.



Meta-QTL établit grâce à un modèle consensus, la position des marqueurs et des QTL sur l'ensemble du génome. Ce logiciel fusionne tout d'abord différentes cartes génétiques en une carte consensus qui est optimale en terme des moindres carrés pondérés et peut être utilisée pour rechercher l'hétérogénéité du taux de recombinaison entre études. Cependant, la méthode manque de connaissances sur la précision effective des distances entre marqueurs dans chaque expérience de cartographie, dû aux données manquantes et aux données propres des marqueurs (codominant / dominant). Pour contrer ce problème, une carte génétique peut être fixée comme référence. Puis, supposant que les QTL peuvent être projetés sur la carte consensus, Meta-QTL permet un regroupement selon un modèle de mélange Gaussien et calcule combien de QTL sont reliés aux QTL observés (Veyrieras et al. 2007).

2.3.3 Approche de gènes candidats positionnels

Une approche de clonage positionnel est utilisée lorsqu'il n'existe aucun gène candidat évident ou aucun défaut biochimique spécifique. Une telle approche qui tente d'identifier un gène d'intérêt à partir de la seule connaissance de sa localisation chromosomique, est lourde et longue à mettre en œuvre. En effet, la localisation primaire, qui s'effectue par une analyse de liaison classique, est suivie d'une cartographie fine de la région d'intérêt, permettant de réduire progressivement l'intervalle critique contenant le gène. Il faut ensuite identifier des clones chevauchants et couvrant la région avant de choisir le gène responsable du caractère. Cependant, le choix du gène est difficile car sa localisation implique un taux d'erreur statistique. Il est donc important de pouvoir établir une liste de gènes candidats pour ce caractère. A l'inverse, une approche purement gène candidat est difficile à envisager. Il faut disposer d'aide sur la localisation du gène ayant un rôle dans l'élaboration du caractère et donc de le confronter à une analyse QTL. Ainsi, la stratégie de clonage positionnel est souvent associée à une approche de cartographie comparée et de gènes candidats (Tableau 6) (Pflieger et al. 2001).

Tableau 6 Principe de l'approche « gènes candidats positionnels »

<p>1- Le choix des gènes repose sur leurs localisations génétiques Hypothèse : tous les gènes situés à proximité du locus à caractériser deviennent candidats</p> <p>2- La recherche de polymorphisme moléculaire pour des gènes dont la fonction correspondrait au QTL de date de floraison Hypothèse : polymorphisme dans le(s) gène(s) pourrait expliquer la variation de date de floraison</p> <p>3- La cartographie fine Hypothèse : tous les gènes qui ne coségrègent pas avec le locus d'intérêt sont rejetés</p> <p>4- La validation définitive du rôle définitif du gène candidat Hypothèse : transformants, mutants,... modifient le phénotype pour le gène transformé, muté</p>
--

2.3.3.1 *Choix des gènes candidats*

Lorsqu'un caractère a été décrit de façon détaillée, il devient possible d'utiliser des banques de données pour identifier les gènes candidats. De plus en plus de gènes issus de plantes modèles comme *A. thaliana*, ou de plantes cultivées comme *O. sativa*, *P. sativum*, *Zea mays* ayant été cartographiés dans des régions sub-chromosomiques spécifiques, les approches positionnelles de gènes candidats, *via* les connaissances de la cartographie comparée, sont souvent privilégiées. C'est par cette approche que le gène de résistance au phoma a été identifié chez le colza (Mayerhofe et al. 2005).

Le choix de gènes candidats peut reposer sur leurs localisations génétiques et il s'agit alors de gènes candidats positionnels car tous les gènes (connus pour intervenir sur le caractère et éventuellement ceux dont on ne connaît pas la fonction) situés à proximité du locus à caractériser deviennent des candidats.

2.3.3.2 *Recherche de polymorphisme et cartographie fine*

La principale hypothèse sous-jacente au choix des gènes candidats, dont la fonction pourrait correspondre au gène majeur ou au QTL recherché, est qu'il existe un polymorphisme dans le gène candidat qui peut être responsable de la variation phénotypique observée.

Lorsqu'un ou plusieurs gènes candidats ont été choisis, des expériences doivent être conduites pour sélectionner le plus vraisemblable. Chez les plantes, deux stratégies sont utilisées :

- (i) la confrontation entre la localisation génétique du gène candidat et celui du locus d'intérêt (Wei et al. 1999) ;
- (ii) des associations entre la variation phénotypique du caractère et des polymorphismes moléculaires dans le gène candidat sont recherchées dans une série d'individus non apparentés (*core-collection*) (Morot-Gaudry and Briat 2004).

La cartographie est essentiellement utile pour rejeter tous les gènes candidats qui ne coségrègent pas avec le locus d'intérêt. Dans le cas des gènes majeurs, l'association entre un gène candidat et le locus doit être totale.

Ainsi, le passage à la cartographie fine de la région d'intérêt, en étudiant plusieurs centaines d'individus, conduit à rejeter certains gènes préalablement jugés intéressants sur de plus petits effectifs. En parallèle, le polymorphisme dans le gène candidat peut être étudié pour rechercher d'éventuelles mutations expliquant la variation de son expression.

L'association entre un gène candidat et un QTL est moins évidente à démontrer du fait de la position imprécise des QTL, ce qui implique un intervalle de confiance couvrant plusieurs mégabases (Dupuis and Siegmund 1999). La cartographie fine est donc nécessaire pour réduire cet intervalle de confiance autour du QTL. Plusieurs exemples de colocalisation de QTL avec des gènes candidats pour des enzymes impliquées dans le métabolisme carboné ont été publiés (Prioul et al. 1999).

Une corrélation statistique ou une liaison génétique étroite entre un polymorphisme d'un gène candidat et une variation d'un caractère peut être étudiée (Thornsberry et al. 2001). Cependant, elle ne suffit pas à démontrer de manière irréfutable une relation de cause à effet entre le gène et le phénotype car le polymorphisme du gène candidat peut être simplement en déséquilibre de liaison avec le gène responsable du phénotype étudié.

2.3.3.3 *Validation du gène candidat*

La validation d'un gène candidat doit être réalisée chez les parents des populations de cartographie. La transformation génétique est le moyen ultime pour valider l'intervention d'un gène candidat dans la variation d'un caractère monogénique. Dans le cas des caractères complexes, la validation par transformation n'est pas immédiate, à cause de l'effet partiel des gènes impliqués. Les expériences de transformations consistent généralement à compléter un phénotype déficient avec une construction « sens » afin de restaurer un phénotype fonctionnel. Des constructions « anti-sens » sont également utilisées, notamment dans le cas de gènes appartenant à des familles multigéniques. Dans ce cas, le transgène inhibe l'expression du gène fonctionnel et la preuve est apportée si le transformant montre un phénotype déficient.

Une autre stratégie consiste à compléter des lignées pour le caractère étudié. Lorsqu'un clone supposé porter le gène d'intérêt est identifié, soit celui-ci est sous-cloné et intégralement séquencé afin de rechercher les cadres ouverts de lecture ORF (*Open Reading Frame*), soit une banque de cosmides est triée de façon à séquencer un fragment plus petit. On étudie alors le polymorphisme des ORF identifiés et l'on recherche d'éventuelles relations entre polymorphisme et variation du caractère. Des constructions plasmidiques porteuses de ces ORF sont produites et utilisées pour transformer des plantes afin de valider leur rôle dans la variation du caractère par complémentation.

L'inactivation d'un gène ciblé par des mutants EMS TILLING ou des mutants d'insertion d'un ADN-t ou d'un transposon, chez les espèces où la technique est maîtrisée, constitue une stratégie alternative de validation mais est réservée à quelques espèces modèles car elle impose de disposer d'une collection de mutants d'insertion et de réaliser son criblage systématique pour le phénotype étudié.

La génétique d'association, très intéressante pour la validation de fonctions de gènes, doit être considérée comme l'un des éléments d'une stratégie de validation de fonction, en parallèle à la transgénèse et à la recherche de mutants. Chez les espèces autogames comme *M. truncatula*, les éléments disponibles actuellement sur le déséquilibre de liaison conduisent à s'interroger sur l'intérêt de l'approche comme outil de validation fonctionnelle.

Des niveaux de déséquilibre de liaison (DL) relativement faibles et la décroissance assez rapide du DL avec la distance physique (environ 25 kb) suggèrent que l'espèce présente une importante taille efficace et/ou que l'autogamie est assez récente (bien que les études menées montraient que ce DL était variable selon les populations considérées) (Ronfort et al. 2006). La faisabilité de cette approche devra tenir compte du fait que si plusieurs gènes proches physiquement présentent de forts déséquilibres de liaison et sont liés au caractère, il sera difficile de déterminer celui qui est en cause.

2.4 Conclusion

M. truncatula est phylogénétiquement proche de la plupart des Légumineuses cultivées en Europe. Elle fait partie des Galégoïdes qui contient les tribus des Trifoliées (luzernes, trèfles), Viciées (pois, féveroles, lentilles, vesces) et Cicérées (pois chiche) (Choi et al. 2004). De plus, elle offre maintenant de nombreuses ressources génomiques (lignées recombinantes, banques BAC, ADNc, EST, cartes génétiques et physiques, mutants,...) qui permettent une meilleure appréhension génétique des caractères étudiés avant le transfert des connaissances à des plantes cultivées (Bell et al. 2001).

La floraison est une étape importante dans la vie d'une plante car elle conditionne le succès reproductif et donc la survie de la plante. *A. thaliana* fournit un grand nombre d'informations concernant les gènes et les différentes voies génétiques contrôlant le processus de floraison. Beaucoup d'études physiologiques de la floraison ont été menées chez le pois et ce n'est que récemment qu'elles ont émergées chez *M. truncatula*. La possibilité d'étudier ce processus mettant en œuvre différentes cascades de gènes, donne lieu à une meilleure compréhension de la floraison chez *M. truncatula* ainsi que des différents gènes la contrôlant.

Ainsi, les connaissances et outils couplés à une approche originale de gènes candidats positionnels permettent de rechercher les gènes responsables des variations génétiques pour la date de floraison chez *M. truncatula*. Cette démarche est présentée dans la suite du document.

MATÉRIEL ET MÉTHODES

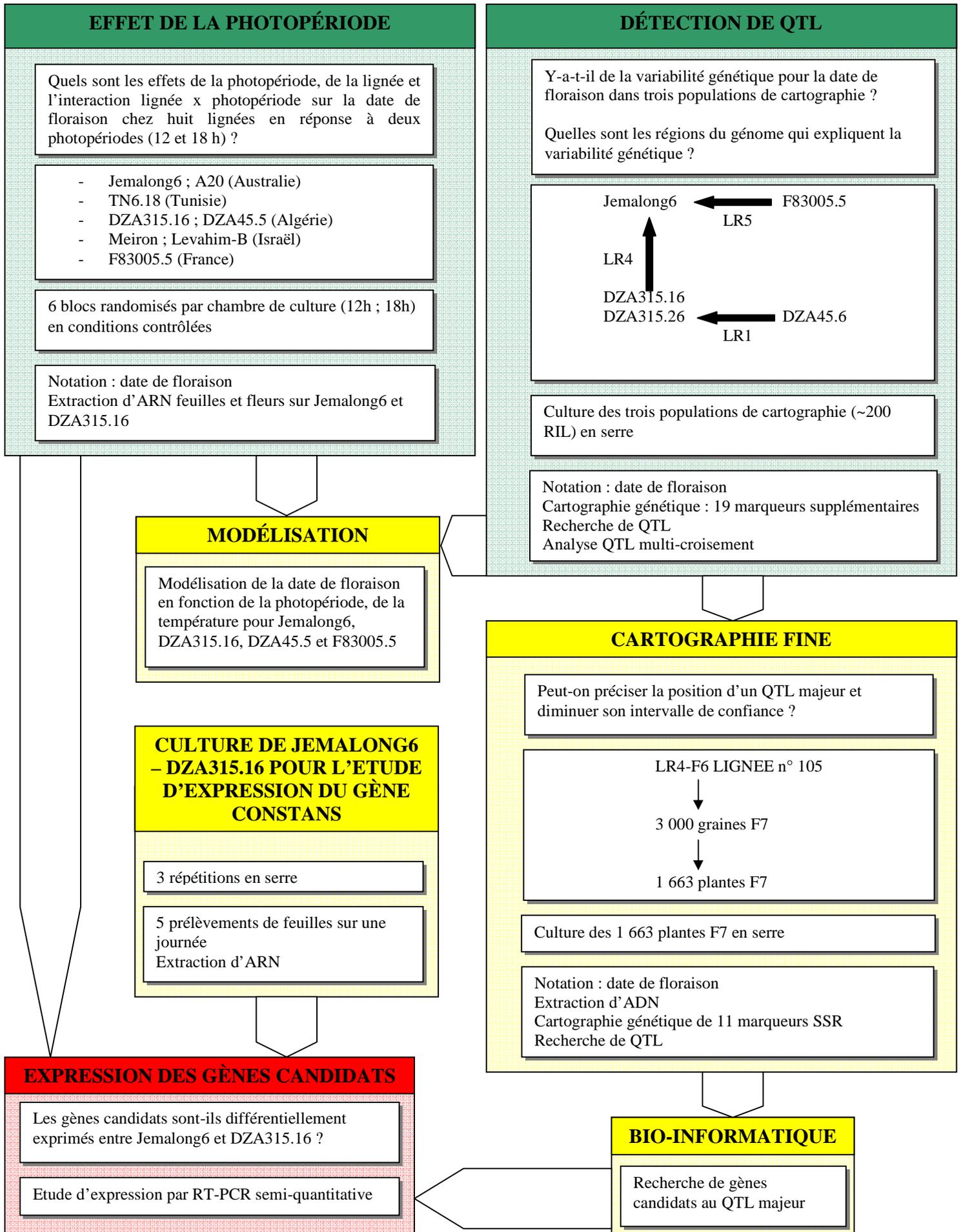


Figure 9 Schéma global résumant les différentes parties de la thèse, le matériel végétal et les dispositifs. Les connexions entre parties sont représentées par des flèches.

3 MATÉRIEL ET MÉTHODES

La Figure 9 présente le schéma global des différentes parties de la thèse, incluant le matériel végétal ainsi que les dispositifs expérimentaux.

3.1 Effets photopériode, lignée et leur interaction pour la date de floraison

3.1.1 Matériel végétal

Afin d'étudier les effets de la durée de photopériode, de la variabilité génétique et de leurs interactions sur la date de floraison de *M. truncatula*, huit lignées provenant de différentes origines ont été sélectionnées : DZA315.16 et DZA45.6 (dérivées de populations algériennes), F83005.5 (dérivée d'une population française), TN6.18 (dérivée d'une population tunisienne), Meiron et Levahim-B (dérivées de populations israéliennes), Jemalong6 (dérivée du cultivar australien Jemalong) et A20 (dérivée d'une population australienne et provenant probablement du Maroc) (Thoquet et al. 2002; Torregrosa et al. 2004; Aydi et al. 2004). DZA315.16, DZA45.6, F83005.5 et Jemalong6 ont été fournis par le centre de ressources biologiques de l'INRA de Montpellier (<http://www.montpellier.inra.fr/BRC-MTR/>), Meiron et Levahim-B par A. Samach de l'université israélienne de Jérusalem et TN6.18, A20 par T. Huguet de l'ENSA de Toulouse.

Les graines ont été scarifiées manuellement avec du papier de verre et mises dans des boîtes de Pétri le 12 octobre 2005. Imbibées pendant 24 heures à température ambiante, les graines ont ensuite été vernalisées à 4 °C pendant sept jours. Les graines germées ont alors été transférées dans des pots individuels le 20 octobre 2005 et réparties dans deux chambres de culture. Le plan d'expérimentation comprenait deux durées de photopériode (12 et 18 heures) et le dispositif était organisé en blocs randomisés avec six répétitions dans chaque chambre, chaque répétition étant représentée par une plante.

Les pots étaient remplis à 70 % d'un mélange de tourbe et d'argile, 20 % de compost et 10 % de sable. La différence de quantité d'énergie lumineuse apportée sous les deux photopériodes a été minimisée en utilisant d'une part des lampes à sodium pendant 12 heures dans les deux chambres de culture, et des lampes à incandescence pour les six heures supplémentaires de la chambre de culture 18 heures, car elles produisent une faible quantité d'énergie.

Ainsi, les plantes en jours courts (12 heures) recevaient 19 085 mmol/m²/jour de flux de photons tandis qu'en jours longs (12 + 6 heures), les plantes recevaient 19 177 mmol/m²/jour soit une différence d'énergie lumineuse entre les traitements inférieure à 0.5 %. L'humidité relative était de 80 % et les températures étaient similaires entre les deux chambres de culture : 20.0 °C (jour) et 17.9 °C (nuit) dans la chambre de culture « photopériode longue », et 19.9 °C (jour), 18.0 °C (nuit) dans la chambre de culture « photopériode courte ». Quatre apports de solutions NPK (3-2-6) à 3 % ont été réalisés, correspondant à 20 ml par plante, le 25 novembre, 05, 12 et 26 décembre 2005. Les plantes ont été arrosées (120 ml) trois fois par jour.

3.1.2 Notations phénotypiques

Une plante a été notée fleurie quand elle avait une fleur ouverte sur la première ramification primaire (Moreau et al. 2006). Tout au long de la période de floraison, la date de floraison était notée individuellement à raison de trois fois par semaine avant d'être transformée en somme de degrés-jours depuis l'implantation en pot avec une température de base (*T_b*) de 5.0 °C. Cette transformation en somme de températures permet de comparer les résultats de l'expérimentation en chambre de culture avec ceux des essais menés en serre. La somme des degrés-jours nécessaires depuis le semis jusqu'à la floraison est calculée ainsi :

$$\text{Nombre de degrés-jours} = \sum [(T_{\min} + T_{\max}) - T_b] / 2$$

Avec *T_{min}* la température minimale, *T_{max}* la température maximale journalières et *T_b* la température de base.

3.1.3 Analyses statistiques

Une analyse de variance a été réalisée pour tester l'effet du génotype sur la date de floraison en utilisant la procédure GLM de type III (SAS 2001) selon le modèle suivant :

$$Y_{ijk} = + L_i + \text{Photopériode}_j + \text{Bloc}_k + L_i \times \text{Photopériode}_j + e_{ijk},$$

Avec Y_{ijk} la valeur du génotype i pris dans la photopériode j et le bloc k , L_i l'effet de la lignée i , Photopériode $_j$ l'effet de la photopériode j , Bloc $_k$ l'effet du bloc k , $L_i \times$ Photopériode $_j$ l'interaction lignée \times photopériode et e_{ijk} l'erreur résiduelle. La normalité des résidus a été testée en utilisant la procédure UNIVARIATE de SAS (SAS 2001).

Afin de mieux comprendre l'interaction lignée \times photopériode, un calcul d'écovalence a été mené ainsi (Wricke and Weber 1986) :

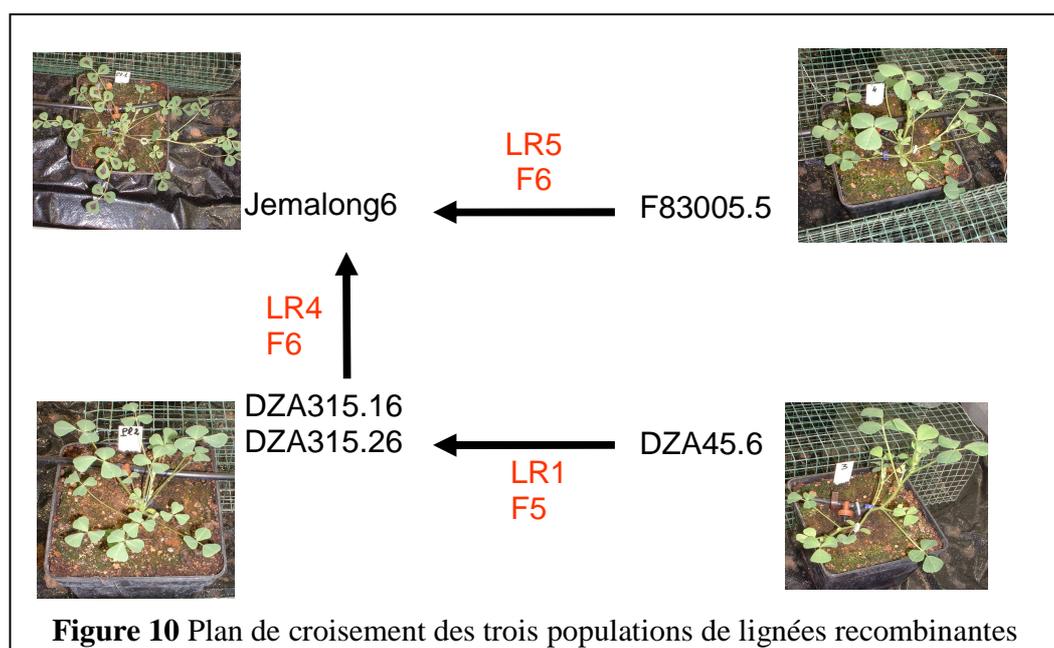
$$W_i = \sum_j (\bar{y}_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})^2$$

Avec W_i qui représente l'écovalence de l'individu i ; \bar{y}_{ij} est la valeur moyenne de la lignée i dans l'environnement j ; $\bar{y}_{i.}$ est la valeur moyenne de la lignée i pour tous les environnements ; $\bar{y}_{.j}$ est la valeur moyenne dans l'environnement j de tous les individus i ; et enfin $\bar{y}_{..}$ est la valeur moyenne de tous les individus dans tous les environnements. Plus l'écovalence est forte et plus la lignée contribue à l'interaction lignée \times photopériode

3.2 Détection de QTL de la date de floraison dans trois populations de cartographie

3.2.1 Matériel végétal

Cette expérimentation a été menée en serre au printemps et à l'automne 2005 dans l'objectif de détecter des QTL de la date de floraison. Deux populations de lignées recombinantes (RILs) ont été utilisées : LR1 (196 RILs en génération F5) et LR5 (173 RILs en génération F7). Les lignées LR1 ont été fournies par le BRC de l'INRA de Montpellier (<http://www.montpellier.inra.fr/BRC-MTR>) et les lignées LR5 par Thierry Huguet (ENSAT de Toulouse). Elles ont été obtenues respectivement par les croisements de DZA315.26 \times DZA45.6 et Jemalong6 \times F83005.5 (Figure 10). Une autre population de RILs, LR4 (199 RILs en génération F7) issue du croisement entre Jemalong6 \times DZA315.16, a été étudiée en 2000, 2002, 2003 et 2004 (Julier et al. 2007). Les deux lignées DZA315.16 et DZA315.26 sont supposées identiques car elles ont été extraites de la même population et sont phénotypiquement identiques en particulier pour la date de floraison (JM Prospero, comm. pers.).



Les plans des essais de la population LR4 ont déjà été décrits (Julier et al. 2007). Pour les populations LR1 et LR5, il n'y a pas eu de répétitions sauf pour 15 lignées par population et les lignées parentales qui ont été répétées trois fois. Les mêmes conditions de scarification et de vernalisation que décrites en 2-1-1 ont été utilisées. Avant la transplantation des graines germées, la terre a été inoculée avec *Sinorhizobium meliloti*. Pour l'essai de printemps, la température moyenne journalière a varié entre 15.5 et 35.8 °C. La photopériode était de 13h57 de jour au début de l'expérimentation (04/04/2005) et de 17h03 à la fin (13/06/2005). En automne, la température moyenne journalière a varié entre 12.7 et 20.3 °C. La photopériode était de 13h21 au début de l'expérimentation (20/09/2005) et de 9h45 à la fin (20/12/2005). Un complément d'éclairage (41 lampes de 250 W couvrant environ 40 m²) a été fourni à partir du 04/10/2005 pour une durée de photopériode journalière totale de 16 heures. La date de floraison a été notée comme précédemment (3.1.2).

3.2.2 Génotypage de marqueurs sur la population LR4

Un recherche de QTL avait été effectuée par Julier et al., 2007 sur la population LR4 et avait permis de mettre en évidence un QTL majeur sur le chromosome 7. Des marqueurs ont été ajoutés aux marqueurs cadres du chromosome 7. Au total, 19 marqueurs ont complété l'intervalle de confiance du QTL majeur de la date de floraison, entre les positions 53 et 65 cM (Tableau 7).

Ces 19 marqueurs comprenaient 16 marqueurs SSR et trois marqueurs-gènes auparavant séquencés (CONSTANS, ABA et Sucrose synthase aussi appelé mtic686). Le gène ABA-responsive comprenait un polymorphisme de 60 paires de bases dans le premier intron et le gène CONSTANS un polymorphisme au niveau de la boîte TATA. Les PCR ont été faites dans un volume final de 25 µl contenant 50 ng d'ADN génomique, 0.2 µM de chaque amorce, 200 µM de dNTP, 1X de tampon 10X, 2 mM de MgCl₂, et 0.625 unité de Taq polymérase. Le thermocycleur utilisé était une PTC100 (MJ Research) avec les conditions d'amplification suivantes : 94 °C pendant 4 minutes, suivi de 35 cycles composés de 94 °C pendant 1 minute, la température d'hybridation optimale pendant 1 minute, 72 °C pendant 1.5 minute, et enfin une extension à 72 °C pendant 5 minutes. Les produits PCR ont été analysés sur gel d'agarose (2 %) dans du tampon TBE 0.5X.

Tableau 7 Description des nouveaux marqueurs cartographiés sur le chromosome 7 de la population LR4

Marqueur						
Nom	Type*	BAC	Amorce sens (5'-3')	Amorce anti-sens (5'-3')	Ta (°C)	Longueur du produit PCR (pb) chez Jemalong6
MTIC724	SSR	AC135311	GCTTTTGAGGTCAAAGTACAA	CAGGGAGTAGCTTGAAATCA	55	143
MTIC246	SSR	AC139601	TTTTCTTCTCAATGGCTTC	TGTTGTCATCAATCTCCTCA	55	117
MTIC470	SSR	AC162276	GGTTCGTGTATTTGTTTCGAT	CCCTTCACAGAATGATTGAT	55	144
MTIC175	SSR	AC133780	TTCCTCATTGACTCTCCCTCTC	AAATGCGATCAGAGGTTTGG	60	112
MTIC40	SSR	AC149129	CATCATTAAACAACAACGGCAAT	TGCAAACACAGAACCGAAGA	64.4	147
MTIC705	SSR	AC137822	ATGCATGTCGGAAAGTAAAT	CGGTGCGAGTATAGTGATG	59.3	139
MTIC719	SSR	AC133779	CTAGCTAGCATTGCTCCAT	TCATCCTAATTGGAGTCCTG	59.1	141
MTIC705b	SSR	AC137822	ATGGAAGGTGGAACCTATCT	GGTGTCGACTGATCCTAGC	55	139
MTIC783	SSR	AC136505	GCAAATCTTCATTTCATCTCA	ATGCTGTAGTTTCTGGACCT	57.8	147
MTIC720	SSR	AC133780	TTGTGGTGACTAGTGATTGG	ATGTGAAGTAAATCCCTTGC	58.1	138
MTIC502	SSR	AC123593	GGATGAAATTGAAAGGAAAACAA	CAACAATCAACTAAGCATACTATTCG	55	202
MTIC760	SSR	AC137822	CGAAGCAATGTAATTTTGTG	GGCCTTCAGATGTTAAAAGTT	55	502
Constans	gene	AC133780	GAGTTTCGCGGAGTCTG	TAGTCCTACCTTCCTCCGA	55.2	272
MTIC252	SSR	AC165429	TCTGTCCTTGTTTTTGTTC	AATTTTGTGAAAGGGGATTT	55	143
MTIC701	SSR	AC122169	ACACCTTCATTGGTGGCTTC	AGGGAAAATCGCCATGAGTT	57.3	186
ABA	gene	AC133780	CAGAGGTAGGCACAATT	GGAACCTTTACTTATCAGA	50.4	544
MTIC714	SSR	AC122169	TAGAAAAGCACACAAGCTG	TGCTACGTATCAAATCAACAA	58.2	143
MTIC241	SSR	AC157894	TCATGAATCTTTGACTTTTTG	AACTGTAGATAGAAATTCCTTGA	55	123
MTIC686	gene	-	GGCATGATGTTCAAATAATGC	CTCGGCAGCAATTTCACTGGC	64	-

* SSR = « Simple Sequence Repeat »

3.2.3 Statistiques et analyses QTL sur les populations LR1, LR4 et LR5

En analyse de variance, l'effet de la lignée et de la saison sur la date de floraison a été recherché par la procédure GLM de type III selon le modèle suivant :

$$Y_{ijk} = L_i + Saison_j + Bloc_k + L_i \times Saison_j + e_{ijk},$$

Avec Y_{ijk} la valeur de la lignée i pris dans la saison j et le bloc k , L_i l'effet de la lignée i , $Saison_j$ l'effet de la saison j , $Bloc_k$ l'effet du bloc k , $L_i \times Saison_j$ l'interaction lignée x saison et e_{ij} l'erreur résiduelle. La normalité des résidus a été testée en utilisant la procédure UNIVARIATE de SAS.

Des analyses de variance ont ainsi été réalisées sur les populations de cartographie LR1 et LR5 (automne et printemps) pour tester la présence d'effets lignées, saisons et interactions lignée x saison sur la date de floraison des lignées parentales et des 15 lignées répétées trois fois.

La procédure PROC VARCOMP a permis d'estimer la variance génotypique (σ^2_G) et la variance résiduelle (σ^2_R). L'héritabilité au sens large (h^2) qui représente la part de la variance génétique dans la variance phénotypique totale a été calculée ainsi pour chaque population et chaque saison :

$$h^2 = \frac{\sigma^2_G}{\sigma^2_G + \sigma^2_R / b} \text{ avec } b = \text{nombre de répétitions.}$$

La cartographie de QTL a été effectuée avec le logiciel QTLcartographer en utilisant la procédure de *Composite Interval Mapping* (CIM) (Basten et al. 1994; Basten et al. 2002). Le seuil d'acceptation d'un QTL détecté après 1 000 permutations était un $LOD \geq 3$ (Churchill and Doerge 1994). Les positions des QTL ont été estimées à l'emplacement du pic maximum du LODscore dans la région considérée. Un intervalle de confiance de part et d'autre du pic de chaque QTL a été estimé à $LOD - 1$ (Lander and Botstein 1989).

Des cartes cadres réalisées avec des marqueurs SSR étaient disponibles pour les populations LR1, LR5 (T. Huguet, non publié) et LR4 (Julier et al. 2007).

Elles incluaient 91, 84 et 71 marqueurs et couvraient respectivement 562.3, 551.5, et 608.4 cM. Par ailleurs, une carte LR4 de référence incluant 133 marqueurs SSR a été employée. Les marqueurs des cartes LR1 et LR5 ont été projetés sur cette carte de référence LR4 en utilisant les marqueurs communs avec le logiciel Biomercator (Arcade et al. 2004). La carte LR1 a été projetée en premier sur la carte LR4 et a ainsi produit une nouvelle carte. LR5 projetée à son tour sur cette nouvelle carte a produit une carte consensus LR1-LR5-LR4. Cet ordre de projection a été choisi car LR1 possède plus de marqueurs que LR5. Chaque QTL présent sur les cartes LR1 et LR5 avec les positions les plus probables ainsi que leur intervalle de confiance ont été projetés sur la carte de LR4 sur la base de leur distance relative aux marqueurs flanquants.

Une cartographie multi-croisements a été menée avec le logiciel MCQTL qui prend en compte le fait que les trois croisements utilisés comportent des parents communs. La possibilité de combiner les données de différents croisements permet d'augmenter la puissance de détection de QTL (Jourjon et al. 2005). La méthode dite de cartographie itérative de QTL (iQTLm) utilisant des cofacteurs génétiques a été choisie. Ces cofacteurs ont été sélectionnés automatiquement et les tests d'effets des QTL réalisés grâce au calcul du F . Le seuil de F était déterminé par un test de 1 000 permutations afin de correspondre à un risque global de 10 % pour la détection des QTL. Le seuil d'acceptation d'un QTL était $\text{LOD} \geq 3$.

3.3 Modélisation de la date de floraison

La date de floraison a été modélisée à partir des données disponibles en 3.1 (effet de la photopériode testée en chambre de culture) et 3.2 (détection de QTL de la date de floraison dans trois populations de cartographie) sur les lignées parentales, en utilisant un modèle de régression linéaire multiple (Summerfield et al. 1985; Erskine et al. 1990). Le taux de développement jusqu'à la floraison ($1/f$ avec f la date de floraison en jours) était décrite comme une fonction de la température moyenne (T en °C) et de la photopériode moyenne (P en h) :

$$1/f = a + bT + cP$$

avec a , b et c des constantes spécifiques des lignées.

Le modèle a été appliqué à chacune des quatre lignées parentales des populations de cartographie (Jemalong6, DZA315.16, DZA45.6, F83005.5). La procédure PROC REG de SAS a été utilisée pour estimer les paramètres a , b et c du modèle, avec l'option FREQ pour indiquer la fréquence de chaque observation.

3.4 Cartographie fine d'un QTL de la date de floraison

Afin de diminuer l'intervalle de confiance du QTL majeur détecté sur la population LR4, une cartographie fine a été mise en place. Elle vise à déceler des recombinaisons entre marqueurs proches, dans une population de grande taille.

3.4.1 Matériel végétal

Parmi les 199 RILs de la population LR4 (Julier et al. 2007), trois lignées F6 ont été identifiées pour être hétérozygotes dans la région du QTL majeur détecté sur le chromosome 7 et homozygotes sur le reste du génome (lignées 48, 100 et 105). Des plantes sœurs de la lignée 105 ont été testées pour leur hétérozygotie, et l'une d'elle a été isolée, autofécondée, et a fourni une descendance de 3 000 graines F7 (Figure 11). Cette génération a été appelée « pseudo-F2 » car elle ségrége sur la portion hétérozygote comme une génération F2. Les graines ont été scarifiées manuellement avec du papier de verre et placées dans des boîtes de Pétri pour une imbibition de 24 heures à température ambiante. Les graines imbibées ont été vernalisées à 4 °C pendant sept jours et les 2 400 graines qui avaient germé ont été transférées en serre le 3 mai 2006. Du fait d'une forte hétérogénéité au niveau de la germination, les repiquages ont été effectués à cinq dates différentes : la majorité le 03 mai et le reste les 04, 05, 09 et 10 mai 2006. Les plantes se sont développées dans des conditions naturelles de longueur de jour croissante de 14h23 à 15h46 durant l'expérimentation. La température atteignait en moyenne entre 18 et 35 °C, le minimum était compris entre 9 et 25 °C, et le maximum entre 26 et 45 °C.

Soit G_p la taille physique du génome de *Medicago truncatula* en paires de bases : 550.10^6

Soit G_g la taille génétique du génome de *Medicago truncatula* en centiMorgans : 608.4

La distance génétique (D) de 100 000 pb correspond à :

$$D = (G_g / G_p) \times 100\,000$$

$$D = (608.4 / 550.10^6) \times 100\,000 = 0.11 \text{ cM}$$

Il y a donc en moyenne 0.11 % de chances d'observer un crossing-over dans 100 000 pb chez un individu donné. Autrement dit, la probabilité qu'un individu ne présente pas de crossing-over dans 100 000 pb est de :

$$P_{100} = 1 - (D / 100)$$

$$P_{100} = 1 - (0.11 / 100) = 0.9989$$

La probabilité qu'aucun individu parmi N ne présentent de crossing-over est : $(P_{100})^N$.

On veut que cette probabilité n'excède pas 5% et on peut donc déduire N de l'équation :

$$\ln 0.05 = N \ln (P_{100})$$

$$N = \ln (0.05) / \ln P_{100}$$

$$N = \ln (0.05) / \ln 0.9989 = \mathbf{2\,722 \text{ individus.}}$$

Figure 11 Taille utile de la population en ségrégation pour avoir 95 % de chances d'observer au moins un crossing-over dans 100 000 pb

La date de floraison de chacune des 1663 plantes F7 obtenues dans le cadre de la cartographie fine a été enregistrée en somme de degrés-jours en utilisant une température de base de 5 °C. L'analyse de variance des données phénotypiques de la population de cartographie avait révélé un effet « date de repiquage » ($P < 0.001$), c'est-à-dire que plus les plantes étaient implantées tardivement et plus le temps nécessaire à la floraison était faible. Pour corriger ce biais, les dates de floraison ont été ajustées comme suit :

$$Y'_{ij} = Y_{ij} + (Y_1 - Y_j)$$

Avec Y'_{ij} la date de floraison ajustée de l'individu i implanté à la date j, Y_{ij} la date de floraison de l'individu i implanté à la date j, Y_1 la date de floraison moyenne des individus implantés le 03 mai 2006 (la majorité des plantes, soit 1149 plantes) et Y_j la date de floraison moyenne des individus implantés à la date j.

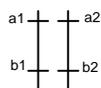
3.4.2 Génotypage

L'ADN des 2 400 individus a été extrait à partir de 50 mg de jeunes feuilles trifoliées (conservées à -80 °C) par une technique classique d'extraction rapide sur plaque 96 puits (Cheung et al. 1993) (Annexe 1). Les quantités d'ADN extraits sont dosées pour 15 individus par plaque (extraction considérée homogène sur l'ensemble d'une plaque) sur gel d'agarose (1 %) à partir d'une solution d'ADN dilué au 1/5^{ième} qui est comparée à des solutions témoins. La quantité d'ADN extrait est comprise entre 0.0013 mg et 0.003 mg.

La cartographie fine a nécessité des moyens à haut débit disponibles sur la plateforme de génotypage de l'INRA Clermont-Ferrand. Sur les 2400 plantes repiquées, 1663 plantes F7 ont été génotypées sur cette plateforme avec les deux marqueurs MTIC40 et MTIC714 bornant le QTL majeur de 7.5 cM (Julier et al. 2007). Ils ont été passés sur un séquenceur capillaire ABI PRISM 3100 (Applied Biosystems). Les conditions PCR sont les suivantes : 1X de tampon 10X, 2.3 mM de MgCl₂, 830 μM de dNTP, 500 nM d'amorces sens, 50 nM d'amorces anti-sens, 520 nM d'amorces M13 marquée, 0.2 U de Taq et 25 ng d'ADN, le volume final est ajusté à 6.5 μl. Le programme PCR se compose : 30 secondes à 95 °C, 30 secondes à 62 °C, 30 secondes à 72 °C (7 cycles avec 1 °C en moins par cycle), puis 30 secondes à 95 °C, 30 secondes à 55 °C, 30 secondes à 72 °C (28 cycles), et 5 minutes d'élongation à 72 °C. Ceci a permis d'identifier dans l'intervalle du QTL détecté en F6, 216 plantes ayant subi un événement de recombinaison et 1424 plantes homozygotes ou hétérozygotes (Figure 12).

Les contigs compris à l'intérieur de l'intervalle de confiance du QTL ont été listés grâce aux données disponibles sur le site <http://www.medicago.org>. Onze marqueurs SSR ont été développés (dont certains sont identiques à ceux ajoutés précédemment sur LR4) sur ces contigs afin de les ordonner. Ces marqueurs, ainsi que la séquence des amorces utilisées pour les amplifier ont été choisis parmi une liste disponible en ligne (<http://www.medicago.org/genome/downloads.php>). Les principales caractéristiques des marqueurs utilisés pour cette cartographie fine sont résumées dans le Tableau 8.

➤ La plante **LR4-105** hétérozygote a=MTIC40 b=MTIC714 (7.5 cM)



➤ Les gamètes sont :



➤ Les 3 000 plantes sont :

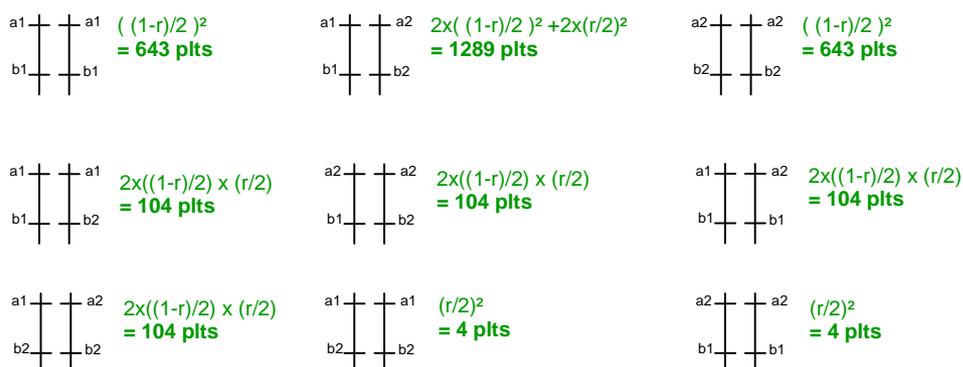


Figure 12 Calcul théorique du nombre de plantes recombinantes, homozygotes, hétérozygotes et doubles recombinants pour deux marqueurs dans une descendance de 3 000 graines F2

Tableau 8 Caractéristiques des marqueurs utilisés pour réaliser la cartographie fine

Nom du marqueur	Type de marqueur	Nom du BAC	Amorces Forward (5'-3')	Amorces reverse (3'-5')	Tm (°C)	Taille des produits PCR (pb)	
						Jemalong6	DZA315.16
MTIC40	SSR	AC149129	CATCATTAACAACAACGGCAAT	TGCAAACACAGAACCGAAGA	64,4	147	144
MTIC719	SSR	AC133779	CTAGCTAGCATTTGCTCCAT	TCATCCTAATTGGAGTCCTG	59,1	141	143
MTGSP4e09	SSR	AC136505	GCAAATCTTCATTCATCTCA	ATGCTGTAGTTTCTGGACCT	57,8	147	145
MTIC720	SSR	AC133780	TTGTGGTGACTAGTGATTGG	ATGTGAAGTAAATCCCTTGC	58,1	138	136
CONSTANS	gène	AC133780	GAGTTTCGCGGAGTCTG	TAGTCCTACCTTCCTCCGA	55,2	272	277
ABA	gène	AC133780	CAGAGGTAGGCACAATT	GGAACCCTTTACTTATCAGA	50,4	544	481
MTGSP2e05	SSR	AC137822	ATGGAAGGTGGAACCTATCT	GGTGTGCGACTGATCCTAGC	55	139	133
MTIC705	SSR	AC137822	ATGCATGTCGAAAGTAAAT	CGGTGCGAGTATAGTGATG	59,3	139	133
19 15a	SSR	AC149134	AGCAAAGCCCAATACCACAT	AACGAGTCCATCTCCATGCT	63,6	220	186
MTGSP1h04	SSR	AC122169	ACACCTTCATTGGTGGCTTC	AGGGAAAATCGCCATGAGTT	57,3	186	178
MTIC714	SSR	AC122169	TAGAAAAGCACAACAAGCTG	TGCTACGTATCAAATCAACAA	58,2	143	149
H4_42i13a	SSR	AC161750	TGTATGCCAAGCATCGGTTA	TGCCAAGAGGAAACTTGGTT	63,5	286	295
7a11b	SSR	AC143341	CACGTGGGATGTCACCACTA	GCCTTGCTGCAGAAGCTATT	63,4	323	285

Tm : « melting temperature » ou température à laquelle statistiquement 50% des amorces ne sont plus hybridées aux brins d'ADN complémentaires
 SSR = « Simple Sequence Repeat » : marqueurs microsatellites constitués de répétitions en tandem de motifs mono-, di-, tri- ou tétranucléotidiques

Les amplicons des différents marqueurs ont été obtenus dans les conditions PCR suivantes : 1X de tampon 10X, 1.5 ou 3.0 mM de MgCl₂, 200.0 µM de dNTP, 0.1 µM d'amorces marquées (M13) sens, 0.2 µM d'amorces anti-sens et 0.1 µM de queue M13 (800 sens), 0.625 U de Taq, 50 ng d'ADN, le volume est ajusté à 25 µl final avec de l'eau. Les amplicons ont ensuite été séparés par électrophorèse sur gel d'acrylamide à l'aide d'un séquenceur LI-COR Gene ImagIR2 DNA analyser system.

3.4.3 Cartographie fine d'un QTL sur le chromosome 7

Le génotypage des 1663 individus de la population de cartographie fine a été utilisé pour la cartographie génétique (logiciel Carthagène v1.0) de cette portion de chromosome 7. Les données ont été analysées comme une population F2 avec un ratio de ségrégation 1 : 2 : 1 à chaque locus marqueur. Un test de Chi-deux a été réalisé pour s'assurer que la distribution observée était conforme à celle attendue à chaque locus marqueur. Les marqueurs présentant une distorsion de ségrégation ont été écartés. La carte a été visualisée à l'aide du logiciel Biomeator.

Les génotypes ont été divisés en quatre classes :

- (i) homozygotes pour les allèles Jemalong6 aux deux marqueurs bornes MTIC40 et MTIC714 ;
- (ii) homozygotes pour les allèles DZA315.16 ;
- (iii) hétérozygotes ;
- (iv) les génotypes ayant une recombinaison entre les marqueurs MTIC40 et MTIC714.

Afin de réaliser une analyse QTL sur les 1663 individus de la cartographie fine, une analyse de variance a été faite sur les trois premières classes pour démontrer statistiquement les différences pour la date de floraison entre les classes et tester.

L'analyse QTL par *Simple Interval Mapping* (SIM) avec un intervalle de 0.1 cM a été menée sur les 1663 individus avec le package r-qt1 sur le logiciel R v2.4.1 (Lander and Botstein 1989; Broman et al. 2003).

3.4.4 Bio-informatique

Le chromosome 7 a (09/07/2007) une taille de 24.3 Mpb (<http://www.medicago.org/genome/stats.php>) et se compose de 242 BAC. La liste des BAC compris dans l'intervalle de confiance du QTL majeur de la date de floraison a été établie sur la base des positions des marqueurs bornes, grâce au site de l'université d'Oklahoma (<http://www.genome.ou.edu/medicago.html>). Les prédictions des gènes candidats sur ces BAC ont été analysées afin d'identifier les gènes déjà connus pour intervenir dans la date de floraison chez d'autres espèces.

Dans l'intervalle de confiance du QTL trouvé sur la population pseudo-F2, les gènes candidats ont été sélectionnés par une approche *in silico*. Les prédictions protéiques des 573 gènes localisés sous l'intervalle de confiance du QTL cartographié sur la population pseudo-F2 ont été obtenues par IMGAG (<http://www.medicago.org/genome/downloads.php>) ou Fgenesh (http://www.tigr.org/tigr-scripts/medicago/gbrowse_bacs.pl). Ces séquences protéiques ont été ensuite analysées par BLASTp sur la base de données Uniprot-Swissprot (Evalue = 1e-6). Seules les séquences ayant une forte similarité avec des gènes connus et décrits dans la bibliographie pour avoir un impact sur la date de floraison chez des espèces modèles ont été sélectionnées (Putterill et al. 2004; Hecht et al. 2005). Un BLASTn a été réalisé sur la base de données des EST des Légumineuses (<http://www.comparative-legumes.org/lis/blasttool.html>) pour ne retenir que des gènes prédits qui sont également exprimés. Les séquences génomiques et codantes des gènes gardés proviennent de la base de données IMGAG (http://www.tigr.org/tigr-scripts/medicago/IMGAG/imgag_annotator.pl?browse=1).

La recherche de cis-éléments et/ou de facteurs de transcription (CAT-box, TATA-box,...) a été réalisée en ligne sur le site <http://www.dna.affrc.go.jp/sigscan/signal1.pl> à partir de la base de données nommée PLACE. Les structures 3D proviennent de la base de données NCBI (en ligne) et les représentations graphiques sont effectuées grâce au logiciel Cn3D.

Différents couples d'amorces ont été testés sur ADN génomique et ADNc pour l'ensemble des six gènes candidats (Tableau 9).

Tableau 9 Ensemble des couples d'amorces testés par gène candidat, sur ADNg et ADNc

Gènes	BAC	Amorces sens		Amorces anti-sens		Ta testées	[MgCl ₂] testées (mM)	ADNc (*)	Longueur (pb)		
		noms	séquences	noms	séquences						
MitPKS	AC149134	PKS_B_1F	GCAGGTGCTATCTCAGTTAA	PKS_B_2R	TGTTGAAGATGAATTAGGGT	50 - 56	1.5 / 3		1296		
		PKS1FM13	CTTTGATTGTAACCTGCCTC	PKS2R	ACAGCTCACTAACAAACCGT	53 - 57.3	1.5 / 3		1300		
MitCO	AC133780	PKS_B_1F	GCAGGTGCTATCTCAGTTAA	PKS_B_2R	TGTTGAAGATGAATTAGGGT	56	1.5 / 3	*	369		
		Co9F	CTCTCACTCTCTTTCTCG	Co12R	CAGACTCCGCGAAACTC	53.5	1.5		3400		
		Co12F	GAGTTTCGCGGAGTCTG	Co8R	TAGTCTACCTTCTTCCGA	53 - 60	1.5		212		
		Co12F	GAGTTTCGCGGAGTCTG	Co31R	GCTAATCTTGGNATHYAYGT	50 - 54	1.5 - 3		3000		
		Co1F	TAAGTTGGCAATAATTTTCAT	Co2R	CCACGCAACGCACAC	45 - 53.9	3		1596		
		Co3F	TTTGCTCGTCGGTCATATC	Co4R	GGCATTATTTTACCATTTTTC	45 - 56	3		889		
		Co1F	TAAGTTGGCAATAATTTTCAT	Co4R	GGCATTATTTTACCATTTTTC	45 - 53.9	3		1733		
		Co7F	AATTCAGACGCCAAA	Co8R	TAGTCTACCTTCTTCCGA	43.9 - 55	1.5		2461		
		Co5F	ATGGGATGAGTTACTTGG	Co6R	GAGCTGAAATTTCTGAGGGA	50.2 - 55	1.5		1445		
		Co33F	TCTGGGATTTTTYAYTAYAG	Co31R	GCTAATCTTGGNATHYAYGT	55	3	*	1199		
		Co9F	CTCTCACTCTCTTTCTCG	Co12R	CAGACTCCGCGAAACTC	54	1.5	*	1268		
		Co40F	CCACACACTTCAACTTCTTA	Co41R	GTGTATGAAGATGCATAAGC	50 - 65	1.5 / 3		1350		
		Co40F	CCACACACTTCAACTTCTTA	Co43R	GCATAAGCATAGGCATGGCA	50 - 65	1.5 / 3		1338		
		Co42F	TTAGTATGGACTTTACTTCT	Co41R	GTGTATGAAGATGCATAAGC	50 - 65	1.5 / 3		1290		
		Co42F	TTAGTATGGACTTTACTTCT	Co43R	GCATAAGCATAGGCATGGCA	50 - 65	1.5 / 3		1278		
		MitFTa	AC123593	FT120F	GTTGGTGACTGRTAYTCNAG	FT110R	GGAAATAAGACARACNCGNAA	60 - 54	3	*	100
				FT1F	CATATCCGAGGAACACTTGA	FT2R	AGCAATGAGTAGTGGGCG	56.7			2790
				FT120F	GTTGGTGACTGRTAYTCNAG	FT110R	GGAAATAAGACARACNCGNAA	55	1.5	*	100
				FT7F	ATGGCTGGTAGCAGTAGGAA	FT8R	GCCACTCTCCCTCTGACA	63 - 52	3	*	507
				FT7F	ATGGCTGGTAGCAGTAGGAA	FT8R	GCCACTCTCCCTCTGACA	55 / 56.4 / 58.2	1.5	*	507
FT120F	GTTGGTGACTGRTAYTCNAG			FT110R	GGAAATAAGACARACNCGNAA	52.4 - 55	1.5 / 2		1100		
FT120F	GTTGGTGACTGRTAYTCNAG			FT130R	CCCTTTGTCTAYACNGTNGA	52.4 - 55	1.5 / 2		1100		
FT7F	ATGGCTGGTAGCAGTAGGAA			FT8R	GCCACTCTCCCTCTGACA	63 - 52	3		1700		
FT7F	ATGGCTGGTAGCAGTAGGAA			FT8R	GCCACTCTCCCTCTGACA	53 - 65	1.5 / 3		1700		
FT7F	ATGGCTGGTAGCAGTAGGAA			FT8R	GCCACTCTCCCTCTGACA	55 / 56.4 / 58.2	1.5		1700		
MitFTb	AC123593			FT160F	TGTATTAGACCNCNTNGRAA	FT170R	GTTTACTCTTTRGCNTAYGA	55	3	*	340
				FT160F	TGTATTAGACCNCNTNGRAA	FT170R	GTTTACTCTTTRGCNTAYGA	60 - 54	3	*	340
				FT160F	TGTATTAGACCNCNTNGRAA	FT170R	GTTTACTCTTTRGCNTAYGA	54 / 55	1.5 / 2 / 3	*	340
				FT160F	TGTATTAGACCNCNTNGRAA	FT170R	GTTTACTCTTTRGCNTAYGA	54.2 / 55.2 / 56 / 57	2 / 3	*	340
				FT3F	AATACATATCCTTACTCGC	FT4R	GACTGTGTGATAGACGA	51.4			1125
				FT9F	ATGGCAAGTGGTAGCAGACC	FT11R	TGAACCTCTAATTTCTGGTTT	63 - 52	3	*	393
				FT9F	ATGGCAAGTGGTAGCAGACC	FT11R	TGAACCTCTAATTTCTGGTTT	64 - 53	3	*	393
				FT9F	ATGGCAAGTGGTAGCAGACC	FT11R	TGAACCTCTAATTTCTGGTTT	53 / 53.9 / 55	1.5	*	393
				FT9F	ATGGCAAGTGGTAGCAGACC	FT11R	TGAACCTCTAATTTCTGGTTT	53.3 / 53.9 / 55	3	*	393
				FT9F	ATGGCAAGTGGTAGCAGACC	FT10R	TTGGCCTATGAAATGGATGG	55 / 56.4 / 58.2	2	*	349
		FT9F	ATGGCAAGTGGTAGCAGACC	FT10R	TTGGCCTATGAAATGGATGG	58.2 / 55 / 61.9 / 64.1	1.5	*	349		
		FT9F	ATGGCAAGTGGTAGCAGACC	FT10R	TTGGCCTATGAAATGGATGG	61.9 / 64.1	2	*	349		
		FT9F	ATGGCAAGTGGTAGCAGACC	FT10R	TTGGCCTATGAAATGGATGG	63.2 / 64.1 / 65	1.5	*	349		
		FT9F	ATGGCAAGTGGTAGCAGACC	FT10R	TTGGCCTATGAAATGGATGG	58.2 / 64.1	1 / 1.25 / 1.5	*	349		
		FT160F	TGTATTAGACCNCNTNGRAA	FT170R	GTTTACTCTTTRGCNTAYGA	54.3	3		502		
		FT9F	ATGGCAAGTGGTAGCAGACC	FT11R	TGAACCTCTAATTTCTGGTTT	63 - 52	3		590		
		FT9F	ATGGCAAGTGGTAGCAGACC	FT11R	TGAACCTCTAATTTCTGGTTT	53 - 65	1.5 / 3		590		
		FT9F	ATGGCAAGTGGTAGCAGACC	FT10R	TTGGCCTATGAAATGGATGG	53 - 65	1.5 / 3		530		
		FT9F	ATGGCAAGTGGTAGCAGACC	FT10R	TTGGCCTATGAAATGGATGG	55 / 56.4 / 58.2	2		530		
		MitFTc	AC123593	FT200F	TTTGGTATGTAYAYTAYTG	FT210R	TAAAACACTACNTCNTTYCC	52.4 - 55	1.5 / 2		1232
FT200F	TTTGGTATGTAYAYTAYTG			FT210R	TAAAACACTACNTCNTTYCC	54.3 / 55	2 / 2.5 / 2.75 / 3		1232		
FT200F	TTTGGTATGTAYAYTAYTG			FT210R	TAAAACACTACNTCNTTYCC	54.3 / 55	1.5 / 1.25 / 1.5 / 2		1232		
FT5F	TTGCTCCTTCATGACTGG			FT6R	AATTGATCGACACTGCCTAA	64	2.5		2961		
FT5F	TTGCTCCTTCATGACTGG			FT6R	AATTGATCGACACTGCCTAA	58 / 60	3		2961		
FT5F	TTGCTCCTTCATGACTGG			FT6R	AATTGATCGACACTGCCTAA	55 - 65	1.5 / 2 / 3		2961		
FT5F	TTGCTCCTTCATGACTGG			FT6R	AATTGATCGACACTGCCTAA	48 - 56	1.5 / 3		2961		
FT5F	TTGCTCCTTCATGACTGG			FT6R	AATTGATCGACACTGCCTAA	62 - 54	3		2961		
FT5F	TTGCTCCTTCATGACTGG			FT6R	AATTGATCGACACTGCCTAA	60 - 67	3		2961		
FT5F	TTGCTCCTTCATGACTGG			FT6R	AATTGATCGACACTGCCTAA	55.4	3		2961		
FT5F	TTGCTCCTTCATGACTGG			FT6R	AATTGATCGACACTGCCTAA	56 - 67	3		2961		
FT5F	TTGCTCCTTCATGACTGG			FT6R	AATTGATCGACACTGCCTAA	54 - 56	1.5 / 3		2961		
FT5F	TTGCTCCTTCATGACTGG			FT6R	AATTGATCGACACTGCCTAA	55.4	3		2961		
FT180F	CCTCTTGGTGNATHGGNGA			FT200R	CAATAGTTGTAYATHCCNAA	54.3 / 55	1 / 1.25 / 1.5 / 1.75 / 2 / 2.5		700		
FT12F	AAATTTGGTCGACCCTCTTG			FT13R	ATCTCCTTCCGCCGC	53 - 65	1.5 / 3		2123		
FT12F	AAATTTGGTCGACCCTCTTG			FT13R	ATCTCCTTCCGCCGC	60.2 / 61.9 / 63.2	1.5		2123		
FT12F	AAATTTGGTCGACCCTCTTG			FT13R	ATCTCCTTCCGCCGC	58.2 / 60.2 / 62.9	1.5		2123		
FT180F	CCTCTTGGTGNATHGGNGA			FT200R	CAATAGTTGTAYATHCCNAA	60 - 54	3	*	300		
FT180F	CCTCTTGGTGNATHGGNGA			FT200R	CAATAGTTGTAYATHCCNAA	55 / 56	1.5 / 2 / 3	*	300		
FT12F	AAATTTGGTCGACCCTCTTG			FT13R	ATCTCCTTCCGCCGC	63 - 52	3	*	503		
FT12F	AAATTTGGTCGACCCTCTTG	FT13R	ATCTCCTTCCGCCGC	66 - 54	3		2123				
FT12F	AAATTTGGTCGACCCTCTTG	FT13R	ATCTCCTTCCGCCGC	60.2 / 61.9 / 63.2	1.5	*	503				
FT12F	AAATTTGGTCGACCCTCTTG	FT13R	ATCTCCTTCCGCCGC	60.2 / 63.2	2	*	503				
FT180F	CCTCTTGGTGNATHGGNGA	FT200R	CAATAGTTGTAYATHCCNAA	55	1.5 / 3	*	300				
MitFD	AC149134	FD1F	GAGGCAGAAGAGGATGATTA	FD2R	GAACCTTTGATCTCCAGCTC	53 - 58	1.5 / 3		500		
		Fdg_1F	TTTTTGTCTTAGTTCGTTGC	Fdg_2R	TAATCACTCCTTCTCCTC	53 - 63	1.5 / 3		2812		
		FDg3F	CCCCAAGAATGTCACCTACT	FDg4R	ACATGGCACAGACGTACA	53 - 63	1.5 / 3		2386		
		FD1F	GAGGCAGAAGAGGATGATTA	FD2R	GAACCTTTGATCTCCAGCTC	58	3	*	191		

Ta : température « annealing » ou température d'hybridation des amorces aux brins d'ADN

* tests réalisés sur ADNc

3.5 Expression de quelques gènes candidats de la date de floraison

L'objectif était d'évaluer l'expression de six gènes candidats présents dans l'intervalle de confiance du QTL majeur.

3.5.1 Matériel végétal

Les plantes utilisées (Jemalong6 et DZA315.16) provenaient de l'essai sur les effets photopériode, lignée et leur interaction pour la date de floraison, et seules les plantes en conditions de jours courts (12 heures) ont été utilisées. Des échantillons de feuilles et fleurs ont été prélevés avant et après la floraison et placés dans l'azote liquide avant d'être conservés à -80 °C (Tableau 10).

Tableau 10 Dates des prélèvements effectués sur Jemalong6 et DZA315.16 en chambre de culture (photopériode courte : 12 heures)

Stades	Jemalong6			DZA315.16		
	Bloc	Prélèvements	Floraison	Bloc	Prélèvements	Floraison
Feuilles avant floraison	1	29/11/2005	23/12/2005	1	29/11/2005	09/01/2006
	2	29/11/2005	30/12/2005	2	29/11/2005	02/01/2006
	3	29/11/2005	03/01/2006	3	29/11/2005	20/01/2006
	4	29/11/2005	23/12/2005	4	29/11/2005	12/01/2006
	5	29/11/2005	06/01/2006	5	29/11/2005	13/01/2006
	6	29/11/2005	24/12/2005	6	29/11/2005	02/01/2006
Feuilles à floraison	1	23/12/2005	23/12/2005	1	11/01/2006	09/01/2006
	2	11/01/2006	30/12/2005	2	11/01/2006	02/01/2006
	2	23/12/2005	30/12/2005	2	23/12/2005	02/01/2006
	3	11/01/2006	03/01/2006	3	27/01/2006	20/01/2006
	4	23/12/2005	23/12/2005	4	11/01/2006	12/01/2006
	5	11/01/2006	06/01/2006	5	11/01/2006	13/01/2006
Fleurs à floraison	6	23/12/2005	24/12/2005	6	11/01/2006	02/01/2006
	6	23/12/2005	24/12/2005	6	23/12/2005	02/01/2006
	2	11/01/2006	30/12/2005	1	11/01/2006	09/01/2006
	3	11/01/2006	03/01/2006	2	11/01/2006	02/01/2006
	5	11/01/2006	06/01/2006	3	27/01/2006	20/01/2006
				4	11/01/2006	12/01/2006
Feuilles après floraison				6	11/01/2006	02/01/2006
	1	11/01/2006	23/12/2005	1	27/01/2006	09/01/2006
	4	11/01/2006	23/12/2005	2	27/01/2006	02/01/2006
	6	11/01/2006	24/12/2005	4	27/01/2006	12/01/2006
				5	27/01/2006	13/01/2006
				6	27/01/2006	02/01/2006
Fleurs après floraison	1	11/01/2006	23/12/2005	1	27/01/2006	09/01/2006
	4	11/01/2005	23/12/2005	2	27/01/2006	02/01/2006
				4	27/01/2006	12/01/2006
				5	27/01/2006	13/01/2006
				6	27/01/2006	02/01/2006
				6	27/01/2006	02/01/2006

Parmi les six gènes identifiés sous le QTL majeur du chromosome 7, l'un d'entre eux, CONSTANS (MtCO) était décrit dans la bibliographie pour s'exprimer différemment durant la journée avec une augmentation de son expression durant la phase lumineuse et une diminution en phase obscure. L'analyse de l'expression de MtCO a donc été menée pendant une journée (avant floraison) sur deux lignées parentales (Jemalong6 et DZA315.16).

Les graines ont été scarifiées manuellement avec du papier de verre et disposées dans des boîtes de Pétri le 27 juin 2007 pour une imbibition de 24 heures. Elles ont ensuite été vernalisées à 4 °C pendant sept jours puis transférées en serre le 2 juillet 2007. Le dispositif expérimental comprenait 15 plantes de chaque lignée randomisées en trois blocs. Le 24 juillet 2007, des échantillons de jeunes feuilles trifoliolées (sur la première ramification primaire) ont été collectés, à raison de trois plantes par lignée prélevée, toutes les quatre heures entre 6h30 et 22h30. Ce 24 juillet 2007, le soleil s'était levé à 6h30 pour se coucher à 21h42.

3.5.2 RT-PCR semi-quantitative

Les ARN totaux ont été extraits à partir de 50 mg de jeunes feuilles trifoliolées et/ou de fleurs ouvertes après broyage dans l'azote liquide grâce au kit *RNeasy* sur colonne (Qiagen). Les ARN totaux ont alors été traités à la DNase et rétro-transcrits respectivement par les kits *DNeasy* et *Superscript first strand synthesis* (Invitrogen). Le dosage des ARN totaux s'est effectué par spectrométrie et la présence (quantité et qualité) d'ADNc contrôlée par l'amplification d'un gène constitutif MsC27 (Crespi et al. 1994).

Le principe de la RT-PCR semi-quantitative consiste à quantifier le niveau d'expression d'un gène donné par rapport à celui d'un gène constitutif exprimé tout au long du développement et dans l'ensemble des organes de la plante (Marone et al. 2001). En pratique, il s'agit de réaliser, au sein de la même réaction PCR, l'amplification du gène ciblé ainsi que celle du gène témoin. Les amplicons sont ensuite séparés par électrophorèse et révélés après un bain de BET. L'intensité des bandes obtenues est alors quantifiée à l'aide d'un logiciel ou bien d'un analyseur. De nombreuses mises au point ont été nécessaires, telles que les compatibilités de (i) tailles, (ii) températures d'hybridation et concentrations en MgCl₂, (iii) nombres de cycles PCR et (iv) concentrations en amorces du gène ciblé et du gène constitutif.

- (i) la taille des amplicons : dans le cadre de cette étude, la séparation des amplicons du gène témoin et des gènes candidats a été réalisée sur gel d'agarose 2 % ce qui implique que les tailles des amplicons soient suffisamment différentes pour être distinguées sur ce type de gel (≥ 10 pb) ;
- (ii) la température d'hybridation et la concentration en $MgCl_2$: elles doivent être compatibles pour amplifier à la fois le gène cible et le gène constitutif de façon spécifique. Elles ont été définies après un test sur ADN génomique comportant un gradient de températures et deux concentrations en $MgCl_2$ (amplification visible et spécifique) ;
- (iii) le nombre de cycles d'amplification : il a été défini de manière à ce que la réaction d'amplification des gènes ciblés soit en phase exponentielle afin de visualiser les différences d'expression entre les deux lignées. En effet, lors de cette phase, la quantité de fragments amplifiés est proportionnelle à la quantité de transcrits présents dans l'échantillon et reflète donc le niveau d'expression du gène dans l'échantillon analysé (Chelly et al. 1990). Pour optimiser ce paramètre, une cinétique de la réaction a été réalisée (24 / 28 / 32 / 36 cycles testés sur ADN génomique) ;
- (iv) la concentration en amorces du gène constitutif : dans cette étude, les gènes candidats ciblés se sont avérés beaucoup moins exprimés que le gène constitutif MsC27, ce qui entraînait un effet de compétition. Pour pallier ce problème, différentes concentrations en amorces MsC27 ont été testées (0.2, 0.1, 0.05 μM) pour une concentration en amorces du gène ciblé de 0.2 μM .

Les amorces utilisées pour amplifier les gènes candidats dans l'essai de suivi d'expression ont été dessinées dans les séquences exoniques grâce au logiciel Oligo6 v6.71 (Molecular Biology Insights, Inc. Cascade, CO, USA). Les amorces ont été choisies avec une longueur moyenne de 20 nucléotides, un pourcentage en bases GC compris entre 30 et 60 % pour les amorces et le produit PCR, la présence de bases GC du côté 3', l'absence de formation en épingle à cheveu ou de simple boucle, une température de fusion comprise entre 57 et 62 °C et une taille de produit PCR différente de celle du gène constitutif MsC27 et de l'ADN génomique. Les produits PCR ont été séquencés, après purification, par l'entreprise Millegen (Labège) et les séquences ont été alignées avec les prédictions de séquences codantes pour s'assurer que les gènes amplifiés étaient les gènes ciblés.

Les PCR ont été réalisées sur des thermocycleurs PTC100 et PTC200 (MJ research Inc.), avec le kit *Taq Platinum* (Invitrogen) comme suit : ADNc (2 ng/ μ l), 1X de tampon 10X, 200 μ M dNTP, 0.05 μ M de chaque amorce sens et anti-sens MsC27, 0.2 μ M pour chaque amorce du gène cible (identique pour tous les gènes), 0.625 unités de la polymérase, une concentration optimale en MgCl₂, et enfin un volume d'eau ultra-pure pour ajuster le volume total à 25 μ l final. Les conditions PCR pour chacun des gènes candidats sont décrites dans le Tableau 11.

Les produits PCR ont été chargés sur un gel d'agarose (2 %) contenant du bromure d'éthidium dans du tampon TBE 0.5X. Un marqueur de taille 100 pb a également migré pour confirmer la taille du produit PCR. Chaque réaction d'amplification contenait un contrôle négatif et un contrôle positif ADN génomique. Les bandes révélées par fluorescence ont été acquises par une caméra ST-50 inclus dans un compartiment Fisher Bioblock Scientific et la quantification a été effectuée avec le logiciel ImageQuant (Molecular dynamics ©, 1998). L'intensité des bandes a été exprimée en pixels.

Le ratio d'expression entre le gène ciblé et le gène constitutif a été calculé pour normaliser les variations initiales de concentration des échantillons. La moyenne et la déviation standard du ratio de chacune des six répétitions (pour chaque organe et stade de développement) ont été calculées après la normalisation. Un test de Student a été réalisé pour tester la différence entre les lignées sur le ratio d'expression.

Tableau 11 Tests de concentrations en amorces gène témoin / gène cible ; de concentrations en MgCl₂, du nombre de cycles PCR

Gène	Concentrations finales				Ta	Nombre de cycle PCR	ADNc (*)
	ADNg ou ADNc	MgCl ₂ (mM)	amorces (μM)				
			MsC27	Gène cible			
MsC27	50	1.5 / 3	0.2	-	53 - 58	35	
MsC27/PKS	50	1.5 / 3	0.05 ; 0.2	0.2	55.2	24 / 28 / 32 / 36	
MsC27/CO	50	1.5 / 3	0.05 ; 0.2	0.2	55.2	24 / 28 / 32 / 36	
MsC27/FD	50	1.5 / 3	0.05 ; 0.2	0.2	55.2	24 / 28 / 32 / 36	
MsC27	25 / 37.5 / 50 / 100	3	0.2	-	56.6	35	*
MsC27/PKS	50	1.5 / 3	0.05	0.2	55.2	28 / 32 / 36	*
MsC27/CO	50	1.5 / 3	0.05	0.2	55.2	28 / 32 / 36	*
MsC27/FD	50	1.5 / 3	0.05	0.2	55.2	28 / 32 / 36	*
MsC27/FTa	50	1.5 / 3	0.05	0.2	55.2	28 / 32 / 36	*
MsC27/FTb	50	1.5 / 3	0.05	0.2	55.2	28 / 32 / 36	*
MsC27/FTc	50	1.5 / 3	0.05	0.2	55.2	28 / 32 / 36	*
MsC27/PKS	50	1.5 / 3	0.05 ; 0.1	0.2	55.2	31 / 35	*
MsC27/CO	50	1.5 / 3	0.05 ; 0.1	0.2	55.2	31 / 35	*
MsC27/FD	50	1.5 / 3	0.05 ; 0.1	0.2	55.2	31 / 35	*
MsC27/FTa	50	1.5 / 3	0.05 ; 0.1	0.2	55.2	31 / 35	*
MsC27/FTb	50	1.5 / 3	0.05 ; 0.1	0.2	55.2	31 / 35	*
MsC27/FTc	50	1.5 / 3	0.05 ; 0.1	0.2	55.2	31 / 35	*

Ta : « Annealing temperature » ou température d'hybridation des amorces aux brins d'ADN
 (*) tests réalisés sur ADNc

RÉSULTATS

4 RÉSULTATS

Les données présentées dans les paragraphes 4.1, 4.2.1, 4.2.2, et 4.2.3 ont fait l'objet de deux publications (Annexes 2 et 3).

4.1 Effets photopériode, lignée et leur interaction pour la date de floraison

4.1.1 Variation de la date de floraison en réponse à la photopériode

Des études ont déjà été menées chez de nombreuses espèces afin de caractériser la réponse des plantes vis-à-vis de la photopériode. Elles démontrent clairement une diversité génétique expliquant les comportements contrastés qui ont pu être sélectionnés au cours de l'histoire des plantes (Searle and Coupland 2004; Hayama and Coupland 2004; Imaizumi and Kay 2006). Chez *M. truncatula*, la diversité génétique en réponse à la photopériode n'a jamais été étudiée. Ce paragraphe met en avant la recherche de cette diversité génétique en réponse à la photopériode chez des lignées de différentes origines.

Huit lignées provenant de populations ou de cultivars originaires d'Algérie, Tunisie, Israël, France et Australie ont été étudiées. Elles ont été choisies pour couvrir une large gamme d'origine et des adaptations à des contraintes environnementales différentes. Deux photopériodes (12 et 18 heures de jour) ont été obtenues sans différences d'apports d'énergie lumineuse. Ces deux traitements ont été appliqués dans un dispositif expérimental composé de six blocs et installé en chambres de culture. Une analyse de variance pour la date de floraison a montré un effet significatif de la photopériode, de la lignée et de l'interaction lignée x photopériode (Tableau 12).

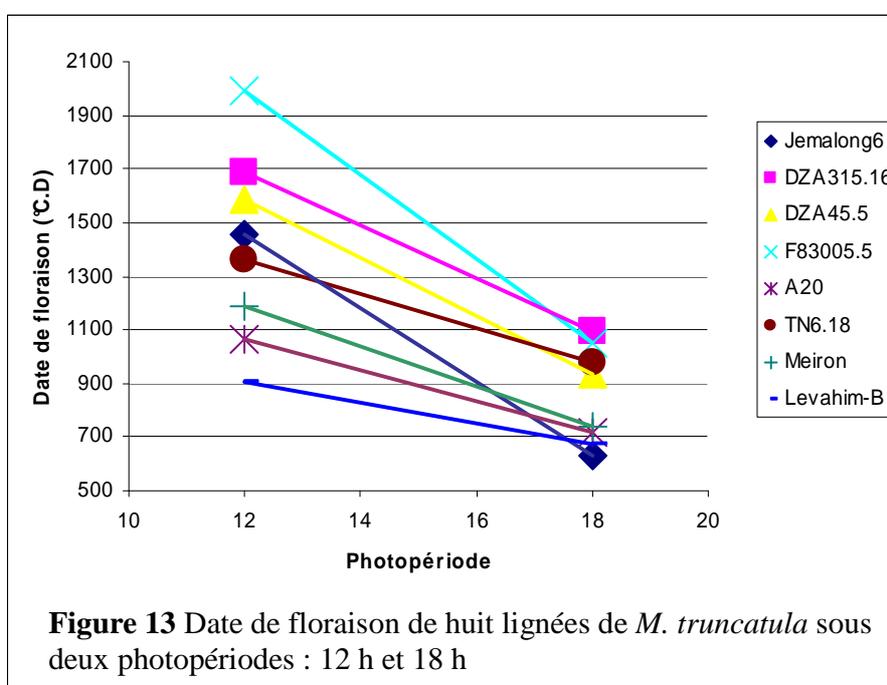
Tableau 12 Analyse de variance (carrés moyens) pour la date de floraison de huit lignées de *M. truncatula* sous deux photopériodes

Effet	Carrés moyens
Photopériode	19543***
Bloc (Photopériode)	9 NS
Lignée	2118***
Lignée x Photopériode	498***
Erreur	24

NS: non significatif à $P = 0.05$, ***: significatif à $P < 0.001$

L'effet bloc hiérarchisé à photopériode n'est pas significatif ($P > 0.05$). Ceci souligne l'homogénéité du dispositif expérimental. L'erreur représente la variabilité due à tous les autres effets non contrôlés, sa valeur faible dans ce dispositif indique que le modèle est précis.

Toutes les lignées ont une floraison plus précoce en photopériode longue (18 h) qu'en photopériode courte (12 h), avec en moyenne, respectivement 1404 et 853 °C.J, soit une différence de 551 °C.J (Figure 13).



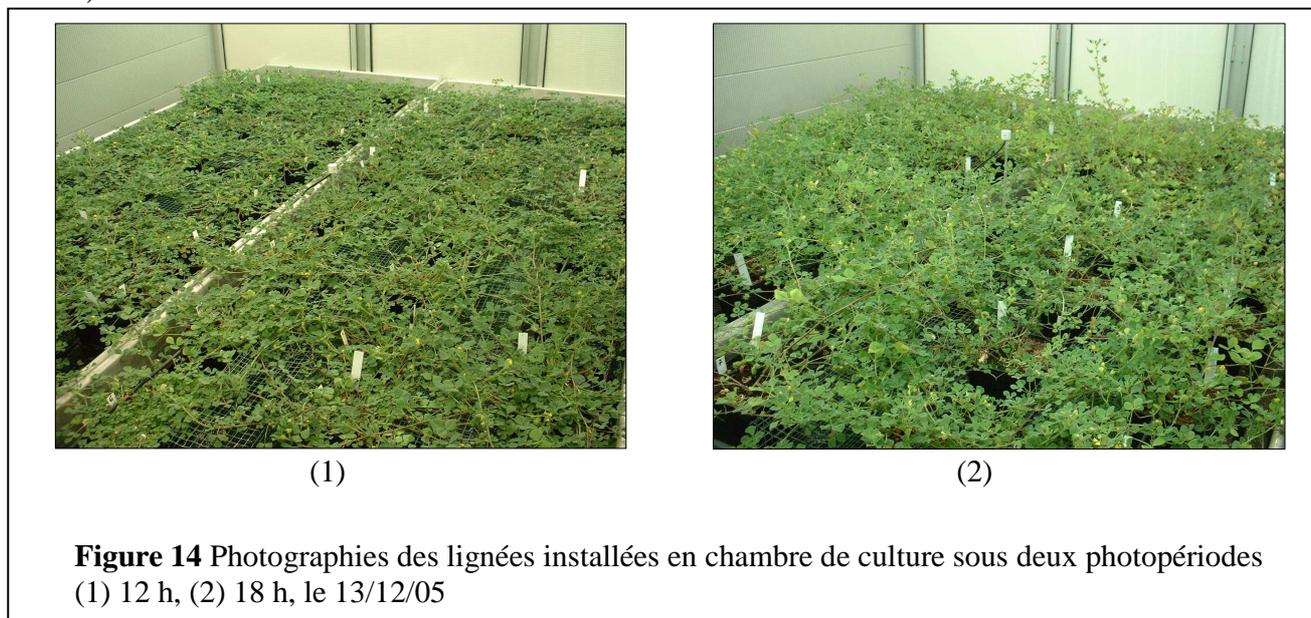
Ainsi, sous 12 heures de lumière, Levahim-B est la lignée la plus précoce tandis que F83005.5 est la plus tardive. En photopériode longue (18 h), Jemalong6, Levahim-B, A20 et Meiron fleurissent précocement alors que DZA45.6, TN6.18, F83005.5 et DZA315.16 fleurissent tardivement. La contribution de chaque lignée dans l'interaction lignée x photopériode a été mesurée grâce à l'écovalence (Tableau 13).

Tableau 13 Date de floraison en degrés-jours pour huit lignées sous deux photopériodes (12 et 18 h). Evaluation de la contribution de chaque lignée à l'interaction lignée x photopériode (l'écovalence de chaque lignée est exprimée en pourcentage de l'écovalence)

Lignée	Date de floraison				Ecovalence (%)
	12 h		18 h		
	°C.J	Jours	°C.J	Jours	
F83005.5	1995	105.0	1049	55.2	37.0
Jemalong6	1454	76.5	633	33.3	17.2
DZA45.6	1583	83.3	937	49.3	2.1
DZA315.16	1685	88.7	1096	57.7	0.3
Meiron	1186	62.4	741	39.0	2.7
TN6.18	1359	71.5	979	51.5	6.9
A20	1064	56.0	716	37.7	9.7
Levahim-B	903	47.5	671	35.3	24.1

Ce calcul montre que DZA315.16 a la plus faible écovalence (0.3 %) ce qui correspond à la lignée qui contribue le moins à l'interaction lignée x photopériode. En effet, DZA315.16 fleurit après 1 685 °C.J à 12 heures de traitement et 1 096 °C.J à 18 heures, soit une différence de 589 °C.J, proche de la différence de moyenne entre les deux traitements. Au contraire, F83005.5, Levahim-B et Jemalong6 ont une forte valeur d'écovalence avec respectivement 37.0, 24.1 et 17.2 %. En photopériode courte (12 h), F83005.5, Levahim-B et Jemalong6 fleurissent respectivement à 1995, 903 et 1454 °C.J alors qu'en photopériode longue (18 h), elles fleurissent à 1049, 671 et 633 °C.J, soit une différence de précocité de 946, 232 et 821 °C.J. F83005.5 et Jemalong6 ont une floraison plus accélérée que la moyenne des lignées. Au contraire, Levahim-B est relativement insensible à la différence de photopériode. Ces trois lignées ont donc une réponse à la photopériode différente de celle des autres lignées et elles contribuent fortement à l'interaction lignée x photopériode. Les autres lignées ont un pourcentage d'écovalence compris entre 2.1 % (DZA45.6) et 9.7 % (A20).

D'un point de vue morphologique, les plantes exposées à une photopériode courte (12 heures) présentent une tige principale courte ou inexistante et un port rampant. A l'inverse, les plantes subissant une photopériode longue ont une tige principale longue et un port plus érigé (Figure 14).



Les différences entre lignées peuvent être liées à leur origine géographique. F83005.5 dont la floraison est très accélérée par une photopériode courte provient du Var (France), une région au nord de la zone de distribution de l'espèce dans laquelle des risques de froid existent. La population F83005 a une floraison tardive au printemps (ou floraison tardive en photopériode courte 12 h), mais cette floraison est rapidement initiée en photopériode plus longue, qui correspond à un moment de l'année où les risques de froid sont minimales. Les plantes issues d'Afrique de Nord (DZA45.6, DZA315.16, TN6.18 et A20), ont un comportement semblable en réponse à la photopériode. Ce comportement est donc compatible avec la survie des populations (réalisation du cycle de vie de graine à graine) dans leurs conditions climatiques. Pourtant, la population DZA315 a été collectée en altitude (1 070 m) alors que les trois autres populations proviennent de plaines. Cependant, même si cette population subit certainement des périodes de froid hivernal, les conditions au moment de la floraison ne sont pas forcément très différentes dans ce site en comparaison des autres sites. Les deux lignées israéliennes Levahim-B et Meiron n'ont pas la même réponse à la photopériode. Leur origine géographique n'est pas précisée, mais pourrait provenir d'un site où les contraintes liées à la saison sont minimales, mais où la population a nécessité d'un cycle très court.

4.1.2 Modélisation de la date de floraison

La modélisation de la date de floraison en réponse à la température et à la photopériode chez *M. truncatula* doit permettre : (i) d'améliorer la compréhension du fonctionnement de la date de floraison en réponse à ces facteurs et (ii) de comparer les quatre lignées parentales des populations de cartographie.

Le modèle de régression linéaire multiple choisi pour modéliser le taux de développement depuis le semis jusqu'à la floraison ($1/f$) est une fonction de la température moyenne (T) et de la photopériode moyenne (P) :

$$1/f = a + bT + cP$$

Le modèle a été appliqué à chacune des quatre lignées parentales des populations de cartographie (Jemalong6, DZA315.16, DZA45.6, F83005.5) à partir des données de l'essai réalisé en chambre de culture en 2006 sur la variation de la date de floraison en réponse à la photopériode et des essais menés en serre au printemps et à l'automne 2005 pour étudier les populations de cartographie (paragraphe 4.2). On observe que le temps entre le semis et la floraison (f) diffère fortement entre les quatre lignées parentales et selon la photopériode et la température appliquées (Tableau 14).

Ainsi, selon les conditions photo-thermales, f varie de 33.3 à 76.5 jours pour Jemalong6, 56.7 à 88.7 jours pour DZA315.16, 48.2 à 83.3 jours pour DZA45.6, et de 48.2 à 105 jours pour F83005.5. Les dates de floraison observées sont les plus précoces dans les conditions de printemps 2005 pour toutes les lignées excepté pour Jemalong6 qui fleurit le plus tôt en condition de photopériode longue en chambre de culture 2006. A l'inverse, les floraisons de l'ensemble des quatre lignées sont tardives à 19 °C (température moyenne) et 12 heures d'éclairement (photopériode courte).

Tableau 14 Modélisation de la date de floraison observée (f) chez quatre lignées parentales à partir des températures et photopériodes moyennes issues de trois essais.

Expérimentations	Lignée	Température	Photopériode	Date de floraison	
				Observée (f)	Prédite
Chambre de culture	Jemalong6	19.0	18.0	33.3	34.9
Chambre de culture		19.0	12.0	76.5	83.3
Printemps 2005		22.8	15.4	37.0	35.6
Automne 2005		17.7	15.4	70.5	52.2
Chambre de culture	DZA315.16	19.0	18.0	57.7	59.6
Chambre de culture		19.0	12.0	88.7	91.7
Printemps 2005		24.0	15.7	56.7	55.5
Automne 2005		17.8	15.7	83.0	73.8
Chambre de culture	DZA45.6	19.0	18.0	49.3	51.9
Chambre de culture		19.0	12.0	83.3	88.6
Printemps 2005		23.1	15.6	48.2	47.4
Automne 2005		17.7	15.5	76.2	69.8
Chambre de culture	F83005.5	19.0	18.0	55.2	55.3
Chambre de culture		19.0	12.0	105.0	105.4
Printemps 2005		23.2	15.6	48.2	48.4
Automne 2005		17.7	15.5	78.0	75.0

Les paramètres a , b et c du modèle sont hautement significatifs ($P < 0.001$). Le pourcentage de variance expliqué par le modèle (R^2) varie entre 93 % pour Jemalong6 et 99 % pour F83005.5 (Tableau 15).

Tableau 15 Modélisation du taux de croissance jusqu'à la floraison ($1/f$) pour quatre lignées de *M. truncatula*

Lignée	Paramètres du modèle $1/f = a + bT + cP$ *						R^2
	a	SE de a	b	SE de b	c	SE de c	
Jemalong6	-0.05467	0.00744	0.00176	0.000347	0.00277	0.0002162	0.93
DZA315.16	-0.01394	0.00214	0.00069	0.000095	0.00098	0.0000775	0.94
DZA45.6	-0.02823	0.00226	0.00124	0.000095	0.00133	0.0000896	0.95
F83005.5	-0.03503	0.00012	0.00144	0.000005	0.00143	0.0000050	0.99

* Tous les paramètres étaient significatifs ($P < 0.001$)

SE = Ecart-type

Les paramètres b et c sont positifs ce qui indique que les jours longs et les températures chaudes accélèrent la floraison. Le paramètre b dénote que Jemalong6, F83005.5 et DZA45.6 ont une floraison plus accélérée par des températures chaudes que DZA315.16.

Le paramètre c est plus grand pour Jemalong6 que pour les trois autres lignées, Jemalong6 est donc plus affectée par la durée de photopériode. La date de floraison prédite par les modèles est indiquée dans le Tableau 14. La date de floraison prédite est toujours supérieure à la floraison observée dans les conditions de croissance des chambres de cultures alors qu'elle est inférieure dans les autres conditions (excepté pour F83005.5 au printemps 2005).

Ainsi, la date de floraison des quatre lignées étudiées peut être prédite dans des conditions de température et de photopériode proches de celles utilisées dans le modèle.

4.2 Détection de QTL pour la date de floraison dans trois populations de cartographie

4.2.1 Variations quantitatives de la date de floraison dans trois populations de cartographie

L'analyse de la date de floraison dans trois populations de cartographie doit permettre l'identification de variations quantitatives pour ce caractère. Trois populations de lignées recombinantes ont été étudiées : LR1 et LR5, au printemps et à l'automne 2005, et LR4 aux printemps 2000 (à Montpellier), 2002, 2003 et à l'automne 2004 à Lusignan. Une analyse de variance a été réalisée sur les lignées parentales installées en 2005 (Tableau 16).

Tableau 16 Analyses de variance (carrés moyens) pour la date de floraison de quatre lignées parentales et de 15 RILs répétées trois fois pour chacune des deux populations de cartographie (LR1, LR5)

	RILs					
	Parents		LR1		LR5	
	Carrés moyens	DL	Carrés moyens	DL	Carrés moyens	DL
Lignée	112262***	3	555389***	13	69873***	14
Saison	495022***	1	311926***	1	2005316***	1
Répétition (saison)	11159 NS	4	10855 NS	4	1964 NS	4
Lignée x saison	22777 NS	3	9493 NS	13	20501**	14
Erreur	11580	23	5130	49	6549	55

NS : non significatif à $P = 0.05$, ** et ***: significatif à $P < 0.01$ et $P < 0.001$, respectivement
DL : nombre de degrés de liberté

Les effets saisons et lignées parentales sont hautement significatifs ($P < 0.001$) tandis que l'effet répétition dans saison, ainsi que l'interaction lignée parentale x saison ne sont pas significatifs.

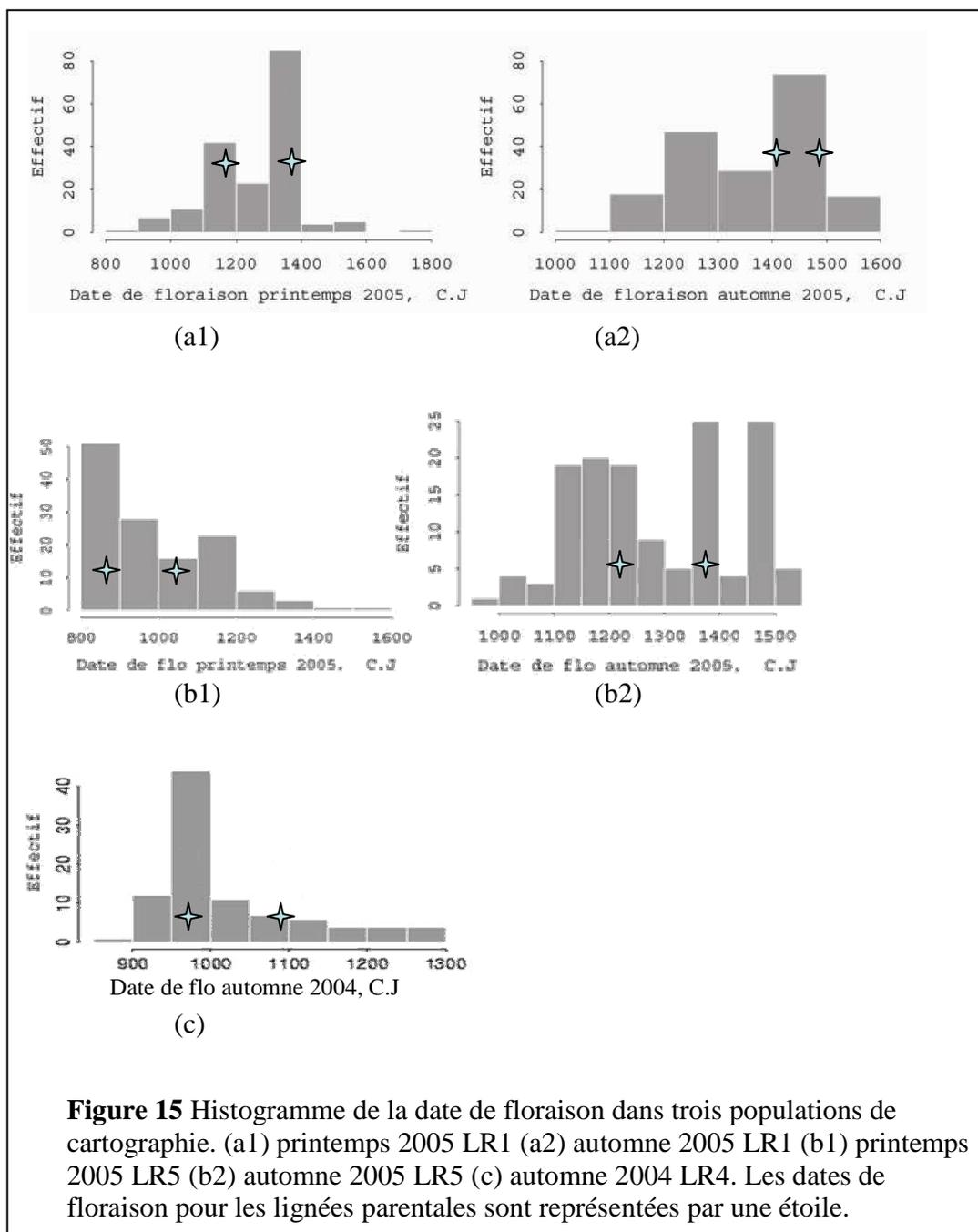
Les analyses de variance menées sur chacune des populations LR1 et LR5 (15 lignées répétées trois fois pour chaque population) sont présentées dans le Tableau 16. Pour la population LR1, les effets lignées et saisons (printemps ou automne) sont significatifs et les carrés moyens sont du même ordre de grandeur. L'interaction lignée x saison n'est pas significative. Pour la population LR5, les effets lignées et saisons sont significatifs, cependant le carré moyen de l'effet saison est plus important que celui de l'effet lignée. Il existe une interaction lignée x saison. Pour ces deux populations, l'effet répétition dans saison n'est pas significatif. Les analyses de variance réalisées sur LR4 en 2002 et 2003 montrent un effet lignée et une interaction lignée x année (Julier et al. 2007).

Pour la population LR1, DZA45.6 fleurit plus tôt que DZA315.26. Pour les populations LR4 et LR5, Jemalong6 fleurit toujours plus précocement que DZA315.16 et F83005.5 (Tableau 17). La moyenne de la date de floraison des RILs est assez similaire au printemps et à l'automne pour la population LR1 (1288 et 1401 °C.J) tandis qu'elle diffère de façon conséquente pour la population LR5 (1012 et 1315 °C.J). La date de floraison moyenne pour la population LR4 est comprise entre 790 et 1147 °C.J aux printemps 2000, 2002 et 2003, mais atteint 1402 °C.J à l'automne 2004 (Tableau 17).

Tableau 17 Moyenne et gamme de variation pour la date de floraison, exprimée en degrés-jours depuis le semis, pour trois populations de RILs et leurs parents

Population	Lieu/Année	Parents		RILs	
		Nom	Date de floraison (°C.J)	Moyenne (°C.J)	Variation (°C.J)
LR1	Lusignan 2005 printemps	DZA45.6	1132	1288	913 - 1784
		DZA315.26	1389		
	Lusignan 2005 automne	DZA45.6	1440	1401	1040 - 1613
		DZA315.26	1502		
LR4	Montpellier 2000 printemps	Jemalong6	740	790	639 - 1181
		DZA315.16	848		
	Lusignan 2002 printemps	Jemalong6	968	1023	878 - 1282
		DZA315.16	1093		
	Lusignan 2003 printemps	Jemalong6	1069	1147	899 - 1500
		DZA315.16	1123		
	Lusignan 2004 automne	Jemalong6	1050	1402	867 - 2120
		DZA315.16	1364		
LR5	Lusignan 2005 printemps	F83005.5	1104	1012	831 - 1568
		Jemalong6	872		
	Lusignan 2005 automne	F83005.5	1379	1315	1022 - 1554
		Jemalong6	1277		

Les quatre lignées utilisées comme parents des populations de cartographie diffèrent donc pour leur date de floraison. Les RILs provenant des croisements DZA315.26 x DZA45.6 (LR1), Jemalong6 x DZA315.16 (LR4) et Jemalong6 x F83005.5 (LR5) montrent une grande variation pour la date de floraison. Les courbes de distributions des dates de floraison pour les populations LR1 et LR5 sont présentées en Figure 15.



Le fait que les RILs de LR4 aient une distribution non normale pourrait être dû à une floraison tardive de certaines lignées durant les premières générations de création de la population, floraison tardive qui a conduit à la perte de ces lignées (Julier et al. 2007). Chez les trois populations, des lignées transgressives pour la date de floraison (RILs montrant une floraison plus précoce ou plus tardive que celle des parents) ont été observées, comme pour des caractères morphologiques dans la population LR4 (Julier et al. 2007).

Un effet saison (printemps vs. automne) a été observé. En effet, la somme de degrés-jours nécessaire pour atteindre la floraison est plus élevée à l'automne qu'au printemps. Il est important de noter que la longueur du jour augmente au printemps et diminue à l'automne avec une amplitude comprise entre trois à quatre heures (voir Matériel et Méthodes). La date de floraison de Jemalong6 et des populations qui lui sont apparentées, est plus accélérée en photopériode longue au printemps 2005 que celles des autres lignées, comme observé dans l'essai photopériode. En automne 2005, la photopériode était plus courte et n'accélère donc pas la date de floraison d'aucune lignée.

L'héritabilité au sens large de la date de floraison a été calculée pour les 15 lignées répétées trois fois (populations LR1 et LR5). Cette héritabilité représente la part de la variance génétique dans la variance phénotypique totale. Elle est de 82.3 % pour LR1 et 75.5 % pour LR5. Dans la population LR4, l'héritabilité était de 82.2 % (Julier et al. 2007).

4.2.2 Identification de QTL

Afin de détecter les zones du génome contrôlant les variations quantitatives observées pour la date de floraison, une recherche QTL est réalisée dans trois populations de cartographie.

Des cartes cadres composées de marqueurs SSR étaient disponibles pour les populations LR1, LR5 et LR4 (voir Matériel et Méthodes). La détection de QTL a été réalisée avec une méthode de CIM sur chacune des populations de cartographie. Au total, trois QTL ont été identifiés sur LR1, onze sur LR4 et quatre sur LR5. Les pourcentages de variance expliquée par les QTL s'échelonnent entre 5 et 60 % (Tableau 18).

Pour LR1, des QTL ont été détectés à la même position (41.8 cM) sur le chromosome 7 au printemps et à l'automne 2005, expliquant respectivement 31 et 34 % de la variation de la date de floraison. Un autre QTL est également présent à l'automne 2005 sur le chromosome 4, le pourcentage de variation observé est de 16 %. Les allèles DZA45.6 ont un effet positif pour l'ensemble des QTL (entre 49.5 et 91.4 °C.J), c'est-à-dire qu'ils induisent une floraison tardive. L'intervalle de confiance des QTL est en moyenne de 7.5 cM.

Tableau 18 QTL de la date de floraison détectés par Composite Interval Mapping dans trois populations de cartographie : LR1 (DZA315.26 x DZA45.6), LR4 (Jemalong6 x DZA315.16) et LR5 (Jemalong6 x F83005.5)

Population	Lieu, année, saison	Chromosome	Position et intervalle de confiance sur chaque carte (cM)	Position et intervalle de confiance projeté sur la carte consensus	LODscore	Effet*	R ²
LR1	Lusignan 2005 printemps	7	41.8 (34 – 44)	58.3 (42 - 62)	9.4	91.4	0.31
		4	45.0 (42 – 50)	62.1 (59 – 67)	3.6	49.5	0.16
	Lusignan 2005 automne	7	41.8 (38 – 43)	58.3 (51 – 59)	9.9	73.5	0.34
LR4	Montpellier 2000 printemps	1	64.0 (58 – 69)	64.0 (58 – 69)	3.6	27.1	0.05
		7	57.5 (55 – 58)	57.5 (55 – 58)	18.6	-74.1	0.42
		8	18.0 (12 – 23)	18.0 (12 – 23)	4.3	-31.4	0.08
	Lusignan 2002 printemps	1	68.0 (64 – 77)	68.0 (64 – 77)	3.1	28.1	0.08
		7	56.7 (55 – 61)	56.7 (55 – 61)	9.4	-58.6	0.37
		8	62.0 (53 – 68)	62.0 (53 – 68)	3.7	-31.8	0.10
	Lusignan 2003 printemps	5	0.0 (0 – 4)	0.0 (0 – 4)	3.8	33.0	0.06
		7	56.3 (56 – 57)	56.3 (56 – 57)	18.5	-99.9	0.60
		8	58.0 (52 – 62)	58.0 (52 – 62)	5.3	-43.6	0.10
	Lusignan 2004 printemps	7	57.8 (56 – 59)	57.8 (56 – 59)	14.1	-213.0	0.32
		8	4.0 (0 – 12)	4.0 (0 – 12)	4.3	-115.0	0.10
	LR5	Lusignan 2005 printemps	7	62.9 (47 – 65)	61.3 (41 – 64)	4.2	-56.5
8			2.0 (0 – 11)	1.6 (0 – 8)	5.7	-67.3	0.17
Lusignan 2005 automne		4	58.3 (55 – 60)	70.4 (65 – 72)	3.8	-45.2	0.10
		7	63.8 (58 – 67)	62.3 (55 – 65)	5.5	-66.8	0.20

* Effet de l'allèle Jemalong6 pour LR4 et LR5, l'allèle DZA315.26 pour LR1

Pour la population LR4, un QTL a été localisé sur le chromosome 7 pendant les quatre années d'étude, sur une position similaire comprise entre 56.3 et 57.8 cM. L'intervalle de confiance moyen est de l'ordre de 3.5 cM. L'allèle Jemalong6 sur le chromosome 7 conduit à une floraison plus précoce de 58.6 à 213.0 °C.J. Ces QTL expliquent entre 32 et 60 % de la variation observée pour la date de floraison. Sur le chromosome 8, plusieurs QTL sont également présents : en position 18.0 cM (printemps 2000), 62.0 cM (printemps 2002), 58.0 cM (printemps 2003) et 4.0 cM (automne 2004). Le pourcentage de variation expliquée est en moyenne de 9.5 %. Pour ces QTL du chromosome 8, l'allèle Jemalong6 a un effet négatif (floraison précoce) compris entre 31.4 et 115.0 °C.J. D'autres QTL ont été trouvés, tout d'abord sur le chromosome 1, au printemps 2000 en position 64.0 cM et au printemps 2002 en position 68.0 cM, et également sur le chromosome 5 au printemps 2003 en position 0.0 cM. Pour les QTL présents sur les chromosomes 1 et 5, le pourcentage de variation expliquée est compris entre 5 et 8 % et les allèles Jemalong6 montrent un effet positif (27.1 à 33.0 °C.J) sur la date de floraison.

La détection de QTL de la morphogenèse aérienne menée entre 2002 et 2004 a montré la colocalisation entre QTL de la date de floraison et longueur de tiges sur le chromosome 7 (Julier et al., 2007).

Pour LR5, un QTL a été trouvé au printemps et à l'automne 2005 sur le chromosome 7, en positions 62.8 et 70.4 cM. Il explique respectivement 11 et 27 % de la variation et couvre un intervalle de confiance de 18 et 9 cM. Au printemps 2005, un QTL a aussi été détecté sur le chromosome 8 en position 2.0 cM (variation expliquée de 17 %) et à l'automne 2005 un autre QTL sur le chromosome 4 en position 58.3 cM explique 10 % de la variation. Pour tous les QTL de cette population, les allèles Jemalong6 ont un effet négatif (entre 45.2 et 67.3 °C.J).

L'utilisation de plusieurs populations maximise la diversité allélique et donc le nombre de QTL détectés. Ainsi, des QTL ont été trouvés sur les chromosomes 1, 4, 5, 7 et 8 en fonction de la population et de l'année ou saison. Un QTL majeur présent dans tous les environnements sur le chromosome 7 des trois populations explique de 11 à 60 % de la variation totale. Cette région génomique a donc un effet fort sur la variation génétique de la date de floraison, bien que cet effet soit plus important chez LR1 et LR4 (R^2 entre 0.30 et 0.60) que chez LR5 ($R^2 < 0.20$). D'autres QTL dépendants de la saison ont été trouvés sur d'autres chromosomes. Ces QTL de moindre importance comparés à celui présent sur le chromosome 7 pourraient être reliés aux effets des conditions environnementales (rayonnement, photopériode, températures) ou de cultures (sol, irrigation, ...) qui varient selon les expériences. Il existe également une autre région génomique commune à deux populations (entre 0 - 16.1 cM sur le chromosome 8 pour les populations LR4 et LR5) où l'allèle Jemalong6 a le même effet négatif.

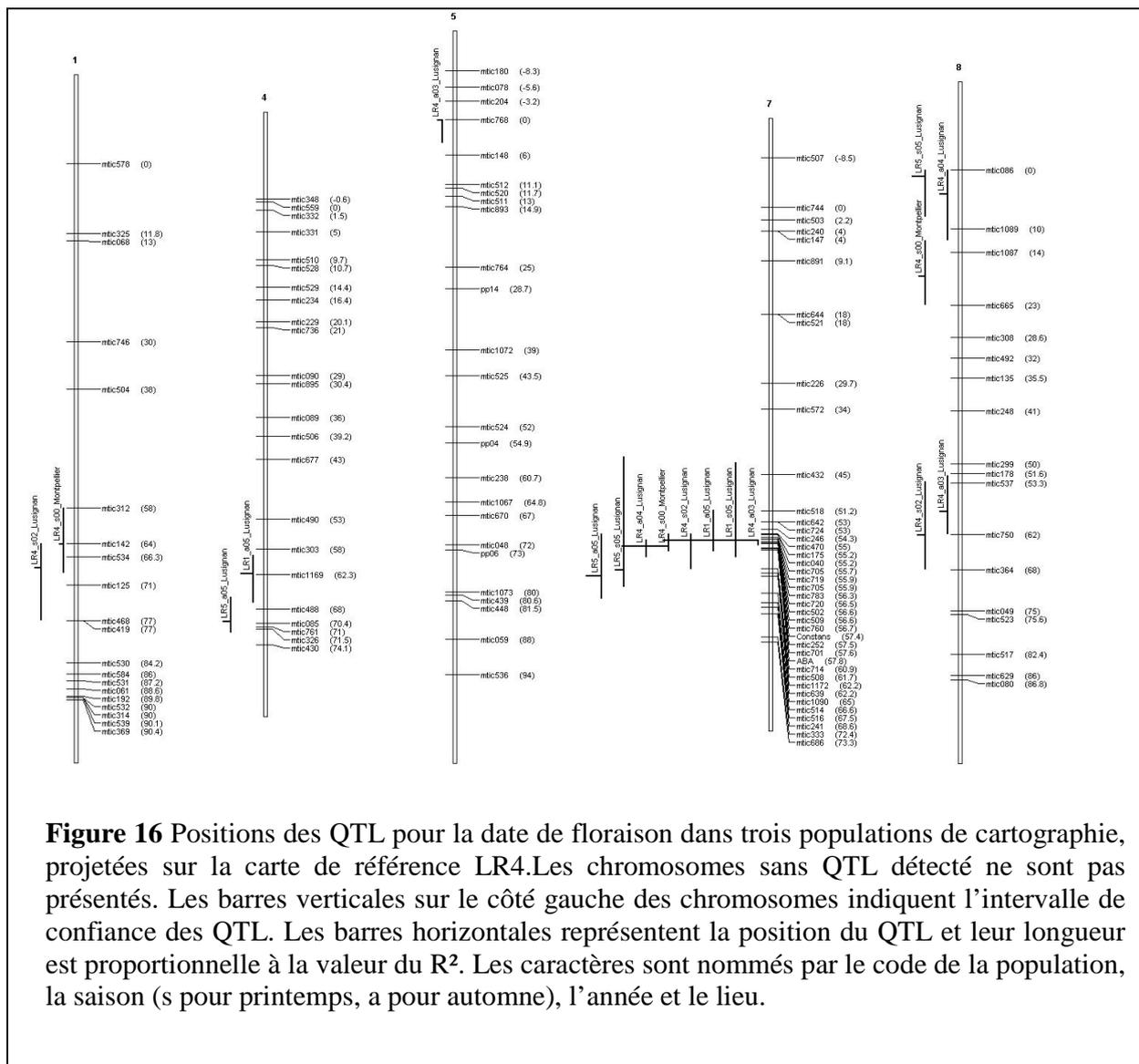
Les allèles à chaque QTL peuvent avoir des effets contrastés sur la date de floraison. Les allèles Jemalong6 induisent soit un effet négatif (floraison précoce) comme sur les chromosomes 4, 7 et 8 (populations LR4 et LR5) ou bien un effet positif (floraison tardive) comme sur les chromosomes 1 et 5 de la population LR4. A l'inverse, les allèles DZA45.6 sur les chromosomes 7 et 4 ont un effet positif sur la date de floraison (population LR1). Les allèles DZA315.16 / DZA315.26 ont un effet négatif dans la population LR1 mais négatif ou positif sur LR4. Les allèles F83005.5 ont un effet positif dans la population LR5.

Sur le chromosome 7, l'allèle DZA315.16 / DZA315.26 a un effet positif dans LR1 mais un effet positif ou négatif dans LR4. Les effets des allèles au QTL du chromosome 7 sont en adéquation avec la gamme de date de floraison observée pour les lignées : Jemalong6 fleurit plus tôt que DZA315.16 et F83005.5, et DZA45.6 fleurit plus tôt que DZA315.26. La présence de lignées transgressives peut être expliquée par une accumulation d'allèles positifs ou négatifs provenant des parents. Une situation similaire a été trouvée chez *A. thaliana* (El Lithy et al. 2004; Loudet et al. 2005).

La recherche de QTL spécifiques des paramètres du modèle (4.1.2) aurait pu être envisagée. Cependant, cela nécessitait la mise en place de conditions de culture permettant l'indépendance des paramètres de photopériode et température sur l'ensemble de la population étudiée.

4.2.3 Projection de cartes et analyse multi-croisements

L'objectif est d'identifier des QTL communs aux populations pour la date de floraison. Les différentes cartes et les QTL détectés chez LR1, LR4 et LR5 ont été compilés. En effet, à partir de la carte de référence LR4 et des marqueurs communs aux trois cartes, une projection a permis de réaliser une carte consensus comprenant l'ensemble des marqueurs. La carte de référence LR4 a été utilisée pour calculer la position des QTL. La position de chaque QTL ainsi que son intervalle de confiance ont été projetés des cartes LR1 et LR5 sur la carte LR4 grâce au logiciel Biomeqan (Figure 16). A cette étape, des colocalisations de QTL sont visibles, sur les chromosomes 1, 2, 7 et 8, mais il convient de tester si réellement ces QTL sont communs aux trois populations et quelle part de la variation ils expliquent. Pour cela, le logiciel MCQTL permet de réaliser la cartographie de QTL tout en considérant les connexions qui existent entre les populations (plan multi-croisements) (Jourjon et al. 2005).



Deux QTL ont été trouvés (Tableau 19) :

- (i) sur le chromosome 7, un QTL a été détecté en position 57.8 cM avec un intervalle de confiance de 0.9 cM. Il explique 25 % de la variation totale dans les trois populations. Les allèles de Jemalong6 et F83005.5 ont un effet négatif (195.8 et 72.8 °C.J respectivement) mais les allèles de DZA315.16 et DZA45.6 ont un effet positif (54.7 et 214.0 °C.J respectivement) ;

- (ii) sur le chromosome 8, un autre QTL a été trouvé en position 10.0 cM et couvre un intervalle de confiance de 16.1 cM. Ce QTL explique 3 % de la variation totale et l'effet des allèles est négatif pour Jemalong6 (64.3 °C.J) alors qu'il est positif pour les autres lignées (entre 8.8 et 40.2 °C.J).

Tableau 19 QTL pour la date de floraison obtenus après l'analyse multi-croisements effectuée par MCQTL dans trois populations de cartographies

Chromosome	Position (cM)	Intervalle de confiance (cM)	LOD score	Effet des allèles (°C.J)				R ²
				Jemalong6	DZA315.16	DZA45.6	F83005.5	
7	57.8	57.2 - 58.1	82.7	-195.8	54.7	214.0	-72.8	0.25
8	10.0	0.0 - 16.1	7.9	-64.3	8.8	15.3	40.2	0.03

L'effet du fond génétique varie selon les populations : il est de 1 229.0 °C.J pour LR1, 1 399.5 °C.J pour LR4 et 1 446.8 °C.J, pour LR5. Le modèle global explique 80 % de la variation du caractère. Les deux QTL détectés sur le chromosome 1 de la population LR4 n'ont pas pu être analysés en multi-croisements car ils sont spécifiques d'une seule population. Il en va de même sur le chromosome 8 entre les positions 52 et 68 cM. Le fait qu'aucun QTL consensus n'ait été trouvé sur le chromosome 4, malgré la présence de deux QTL issus de deux populations (LR1 et LR5), est certainement dû au mauvais chevauchement des intervalles de confiances.

L'intervalle de confiance très réduit de 0.9 cM du QTL présent sur le chromosome 7 est borné par les marqueurs MTIC760 et MTIC714 qui correspondent à mtgsp004g09 et mtgsp002g05 sur la carte intégrée de l'université du Minnesota (<http://www.medicago.org/genome/map.php>). Cet intervalle de confiance couvre 26 BAC (http://www.medicago.org/genome/show_BAC_chr2.php?chr=7).

4.2.4 Cartographie fine d'un QTL sur le chromosome 7

Du fait de l'important effet du QTL détecté sur le chromosome 7 sur l'ensemble des trois populations de cartographie et en particulier sur LR4, nous avons choisi de cartographier finement cette zone du génome. L'intervalle de confiance considéré est celui calculé sur LR4 en génération F6, il représente 7.5 cM. Le but de cette étude consiste à réduire l'intervalle de confiance du QTL afin de diminuer le nombre de gènes candidats.

Pour cela, une pseudo-population F2 de 1663 plantes issues d'une plante F6 hétérozygote aux deux marqueurs bornant le QTL majeur (mais fixée dans tous les autres QTL détectés) a été créée et étudiée (Figure 17).

Sur 3 000 graines, 1 663 plantes ont pu être utilisées.

Le nombre d'individus N nécessaire pour avoir 95 % de chances d'observer un crossing-over est égal à :

$$N = \ln 0.05 / \ln P100$$

$$\ln P100 = \ln 0.05 / N$$

$$\ln P100 = \ln 0.05 / 1663 = - 0.001801$$

$$P100 = 1 - 0.001801 = 1 - (0.1801 \text{ cM} / 100)$$

Sachant que $1 \text{ cM} = 550.10^6 / 608.4 = 904 \text{ 010.52 pb}$

Alors 0.18 cM représente **162 812 pb**.

Avec 1 663 individus, il y a 95 % de chances d'observer un crossing-over dans 16 2812 pb soit 163 kpb.

Figure 17 Calcul de la taille du fragment étudié dans lequel il y a 95 % de chance d'observer un crossing-over avec 1 663 plantes chez une population F2

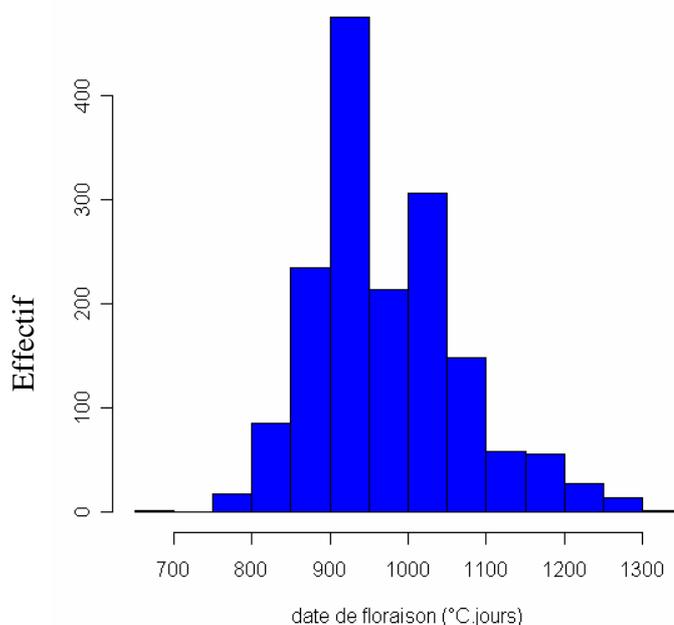
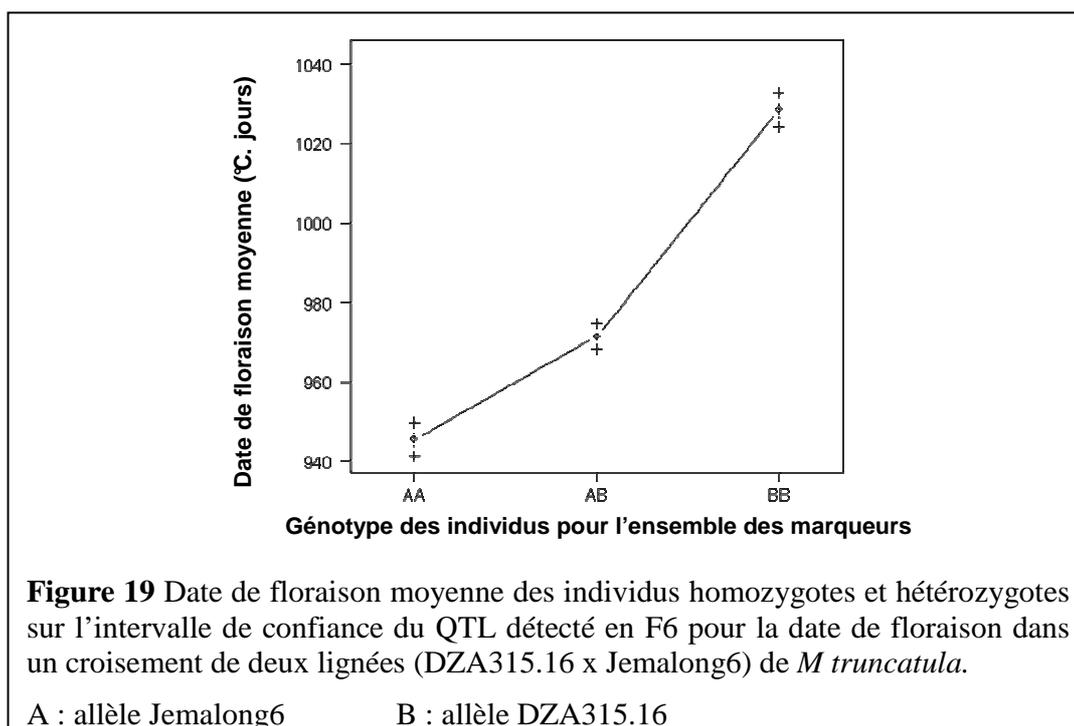


Figure 18 Distribution de la date de floraison des individus de la pseudo-population F2

D'importantes variations pour la date de floraison ont pu être observées dans la pseudo-population F2 (Figure 18). En effet, la floraison s'échelonne entre 700 et 1302 °C.J avec une moyenne de 979 °C.J et une valeur médiane de 950 °C.J. L'examen des dates moyennes de floraison des plantes génétiquement fixées et des hétérozygotes pour l'ensemble des marqueurs révèle que :

- (i) les individus homozygotes pour l'allèle Jemalong6 fleurissent en moyenne à 946 °C.J, ceux homozygotes pour l'allèle DZA315.16 à 1026 °C.J, et les hétérozygotes autour de 972 °C.J ;
- (ii) il existe bien une différence significative entre les trois classes (homozygotes pour l'allèle Jemalong6, homozygotes pour l'allèle DZA315.16 et hétérozygotes) ($P < 0.001$) ;
- (iii) il existe une dominance partielle de l'allèle Jemalong6 à ce locus. En effet, la Figure 19 qui présente la date de floraison moyenne des individus fixés à l'intérieur de l'intervalle de confiance du QTL détecté en F6, illustre un cas intermédiaire entre un effet additif des allèles et un effet de dominance. Ainsi, les hétérozygotes, bien qu'ils soient nettement plus précoces que les homozygotes pour l'allèle DZA315.16, sont faiblement plus tardifs que les homozygotes pour l'allèle Jemalong6. Statistiquement, cela se traduit par une date de floraison moyenne du groupe d'homozygotes (allèles DZA315.16 et Jemalong6 compris) significativement différente de celle des hétérozygotes ($P < 0.001$). Cette dominance partielle de l'allèle Jemalong6 explique que la distribution des individus pseudo-F2 suit une courbe en cloche dissymétrique avec une fréquence plus élevée des individus précoces.



Les proportions mendéliennes caractéristiques d'une population F2 ($\frac{1}{4}$, $\frac{1}{2}$, $\frac{1}{4}$) sont respectées (Tableau 20). La distribution des individus homozygotes ou hétérozygotes aux marqueurs bornes MTIC40 et MTIC714 (*i.e* homozygotes ou hétérozygotes sur l'ensemble du QTL) suggère un contrôle monogénique de la date de floraison dans cette population. Le QTL détecté pourrait donc dissimuler un gène majeur voire plusieurs gènes suffisamment proches pour paraître liés dans la population étudiée.

Tableau 20 Test du χ^2 évaluant la ségrégation des allèles chez 1423 individus non recombinants (homozygotes ou hétérozygotes) pour l'ensemble des marqueurs à l'intérieur du QTL

Allèle	Effectifs observés	Effectifs théoriques	χ^2
AA	401	355.75	0.007
AB	658	711.5	
BB	364	355.75	

A = allèle du parent Jemalong6 ; B = allèle du parent DZA315.16
hypothèse nulle : les allèles ségrégent dans les proportions $\frac{1}{4}$, $\frac{1}{2}$, $\frac{1}{4}$

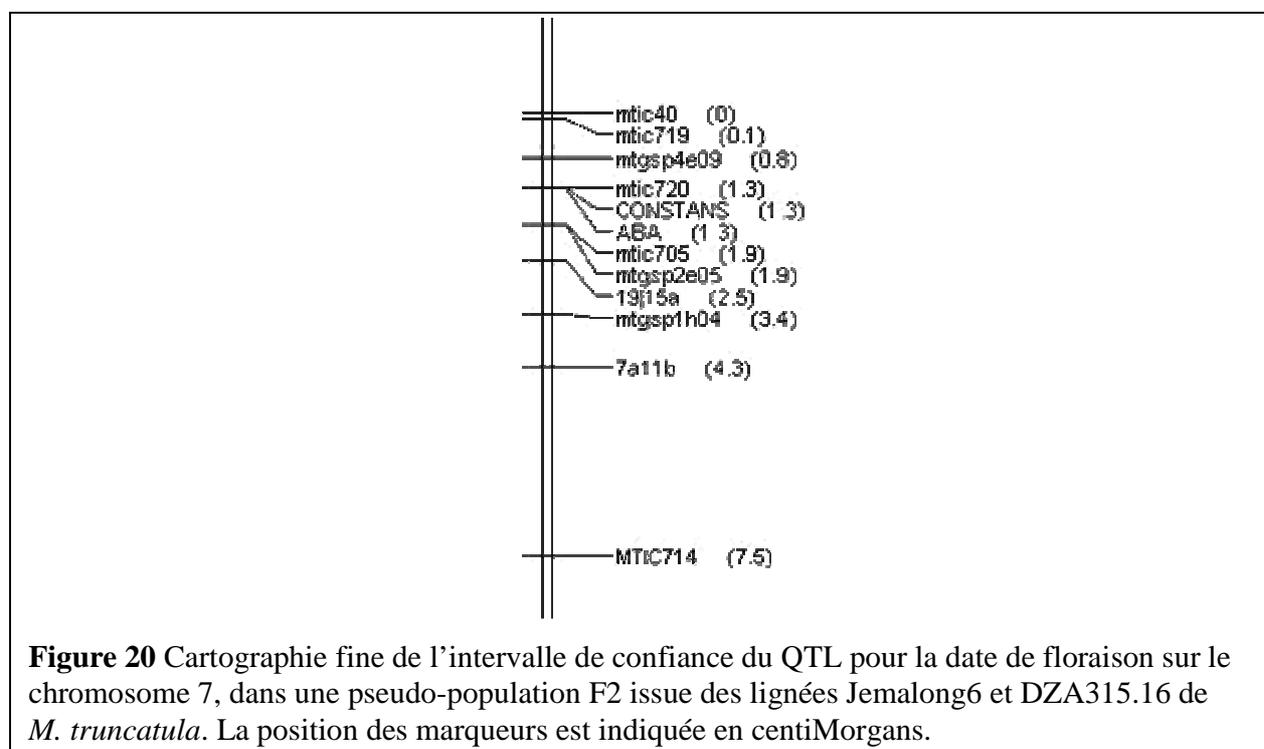
Un test de χ^2 a permis de comparer les effectifs observés et théoriques aux différents marqueurs (Tableau 21). Seul le marqueur h4_42i13a ne suit pas la ségrégation attendue. En effet, celui-ci présente une absence totale de ségrégation, il a donc été abandonné.

Tableau 21 Test du χ^2 évaluant la ségrégation des allèles à chaque marqueur

Nom des marqueurs	Effectifs observés			Effectifs théoriques			χ^2
	AA	AB	BB	AA	AB	BB	
MTIC40	449	762	429	410	820	410	0.013
MTIC719	448	756	426	410	820	410	0.01
MTGSP4e09	448	755	427	410	820	410	0.009
MTIC720	451	755	427	410	820	410	0.007
CONSTANS	451	756	427	410	820	410	0.007
ABA	450	753	429	410	820	410	0.006
MTIC705	456	756	426	410	820	410	0.004
MTGSP2e05	456	755	427	410	820	410	0.004
19 15a	454	758	422	410	820	410	0.007
h4_42i13a	0	0	1660	410	820	410	1680
7a11b	451	753	416	410	820	410	0.017
MTGSP1h04	456	755	414	410	820	410	0.005
MTIC714	458	771	411	410	820	410	0.014

A = allèle du parent Jemalong6 ; B = allèle du parent DZA315.16

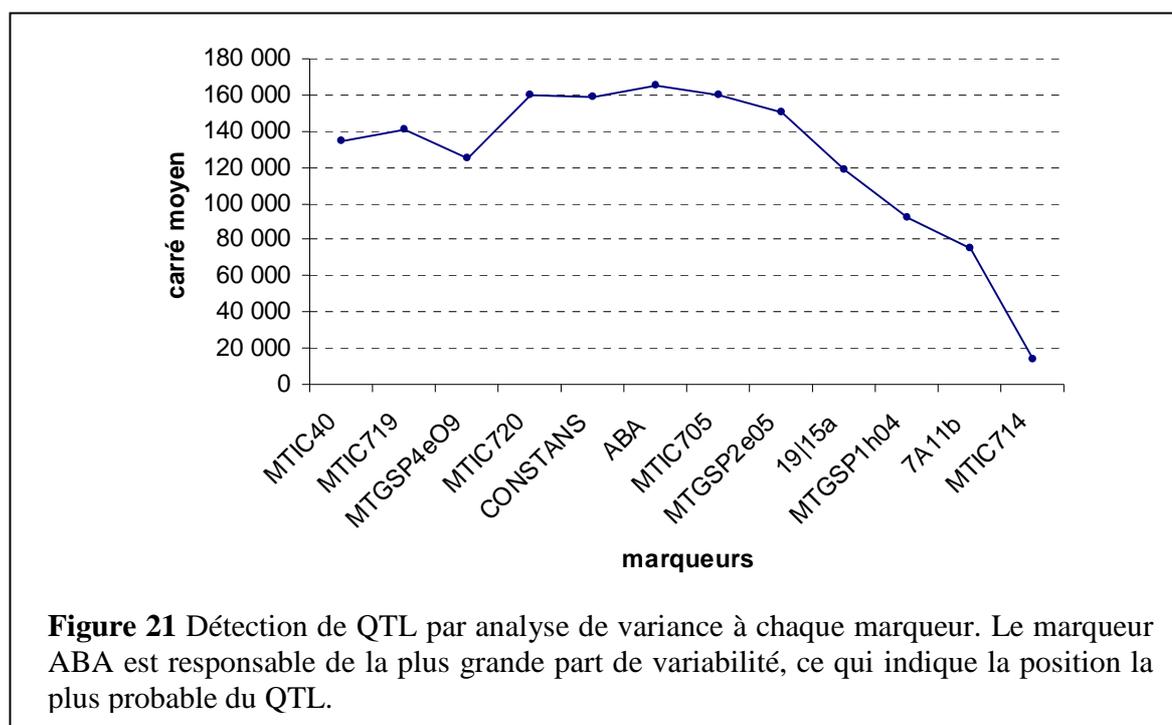
hypothèse nulle : les allèles ségrégent dans les proportions $\frac{1}{4}$, $\frac{1}{2}$, $\frac{1}{4}$



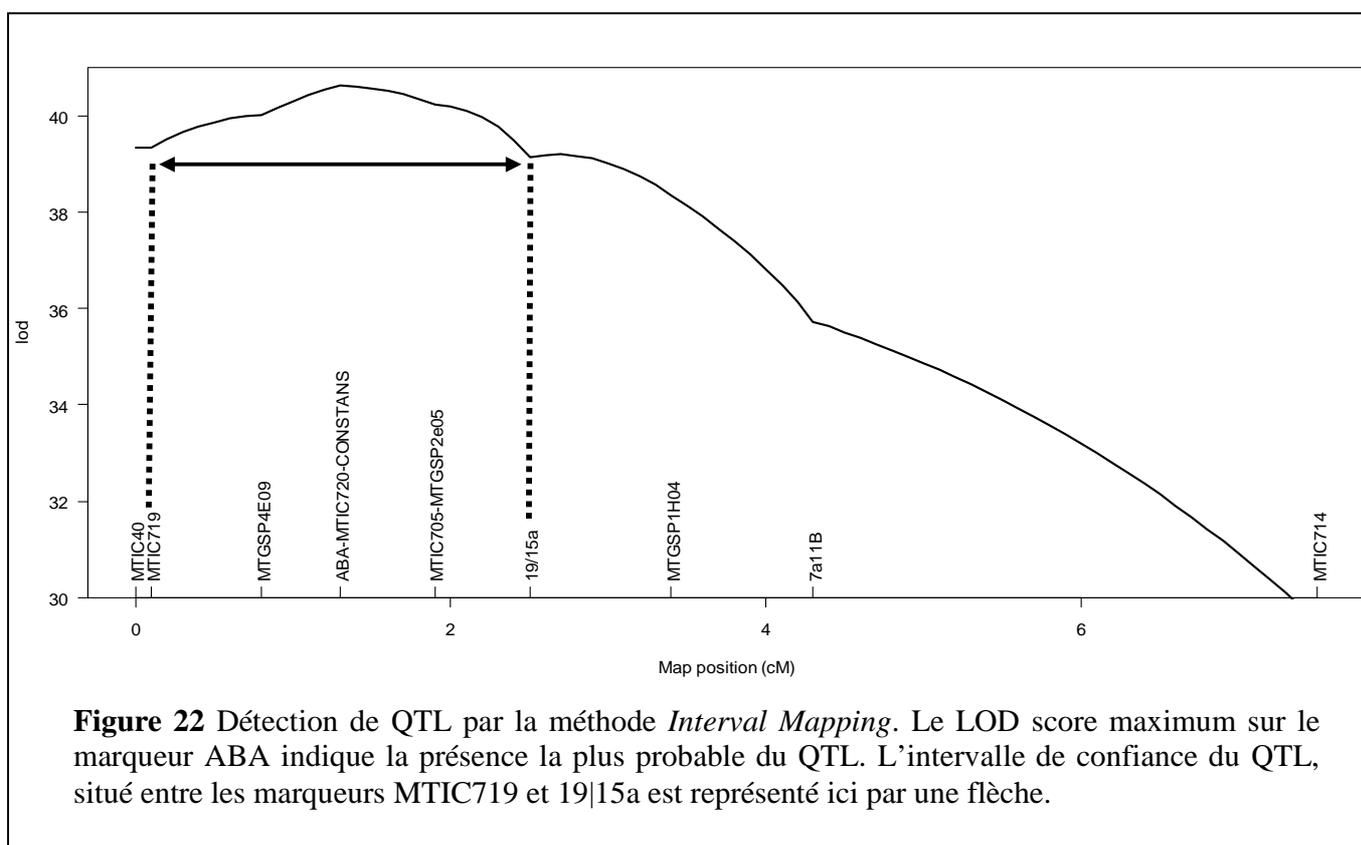
La carte établie dans le cadre de cette étude donne une longueur de 7.5 cM entre les deux marqueurs bornes MTIC40 et MTIC714 (Figure 20). Elle comprend 12 marqueurs inégalement répartis le long de cette portion du chromosome 7. Les marqueurs CONSTANS, ABA, MTIC720 d'une part et mtgsp002e05 et MTIC705 d'autre part, sont situés à une distance génétique de 0.0 cM, c'est-à-dire qu'il n'y a pas de recombinaison observée entre ces marqueurs. Ce résultat était attendu puisque ces derniers avaient été définis respectivement sur les BAC AC133780 et AC137822. La densité de marqueurs est importante entre 0.0 et 4.3 cM, mais il n'y a pas de marqueur entre 4.3 et 7.5 cM.

Deux méthodes de détection de QTL ont été employées : (i) l'analyse de variance par marqueur et (ii) l'analyse par *Interval Mapping*.

- (i) l'analyse de variance par marqueur indique un effet du locus ABA plus fort que l'effet des autres marqueurs, ce qui signifie donc que le QTL se trouve proche de la position du marqueur ABA (Figure 21). Cette analyse permet d'écarter le marqueur MTIC714 car son effet n'est pas significatif sur la date de floraison. Cette méthode n'admet pas la création d'un intervalle de confiance du QTL puisque, dans le cas présent, chaque marqueur, à l'exception de MTIC714, a un effet significatif sur la date de floraison ;



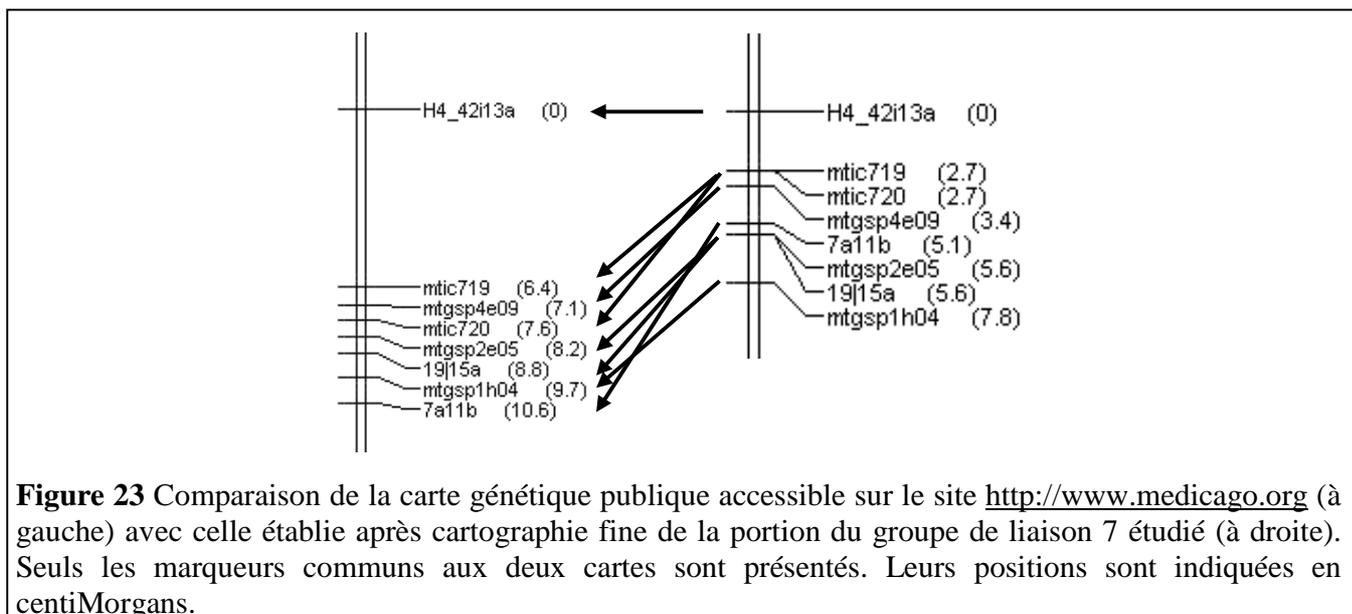
- (ii) la méthode d'*Interval Mapping* présente l'avantage de définir des limites probables du QTL. Le résultat obtenu avec un pas de 0.1 cM est tout à fait comparable à celui de l'analyse de variance : un seul QTL, dont la position la plus vraisemblable se trouve autour du marqueur ABA (LODscore = 40.62). L'intervalle de confiance autour du marqueur ABA couvre une région de 2.4 cM (Figure 22) bornée par les marqueurs MTIC719 (équivalent au marqueur mtgsp003a06) et 19|15a. Au total, 19 BAC sont référencés dans cet intervalle (http://www.medicago.org/genome/show_BAC_chr2.php?chr=7).



Une comparaison de la carte obtenue après cartographie fine avec cette carte publique permet de constater, sur cette faible portion du génome de *M. truncatula*, des divergences (Figure 23) :

- (i) l'ordre des marqueurs est différent (les couples MTIC720 / mtgsp4e09 et 7a11b / mtgsp1h04 voient leurs positions inversées) ;

- (ii) certains marqueurs cartographiés à une distance de 0.0 cM sur la carte publique sont à une distance de 1.2 cM pour le couple MTIC719 / MTIC720 et 0.6 cM pour le couple MTGSP2e05 / 19|15a, ce qui indique que la résolution de la carte présentée ici est de meilleure qualité, sans doute du fait de la taille de la population étudiée (1 663 individus).

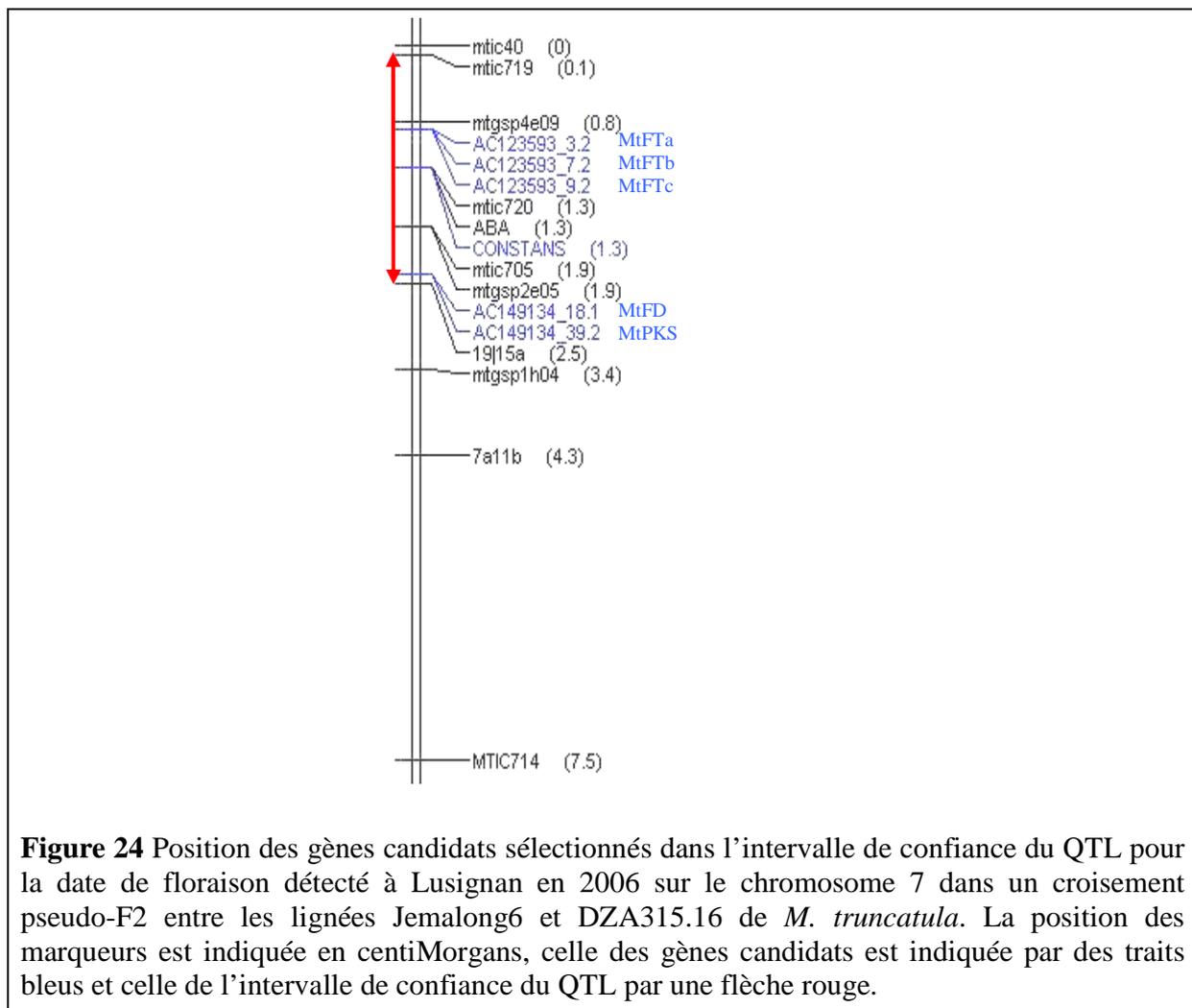


4.2.5 Sélection de gènes candidats

La stratégie d'approche des gènes se base sur une démarche de gènes candidats positionnels (voir synthèse bibliographique). Cette stratégie devrait permettre d'identifier, dans la zone couverte par le QTL majeur détecté sur le chromosome 7, plusieurs gènes candidats au contrôle de la date de floraison.

L'intervalle de confiance considéré est celui détecté après la cartographie fine, soit 2.4 cM entre les marqueurs MTIC719 et 19|15a. Cet intervalle est plus grand que celui considéré par l'analyse multi-croisements (0.9 cM) mais il permet de ne pas écarter les gènes qui seraient statistiquement aux bornes du QTL. A partir des données bioinformatiques et des ressources bibliographiques disponibles chez *M. truncatula* et chez d'autres espèces modèles, six gènes candidats ont été sélectionnés à l'intérieur de l'intervalle de confiance du QTL parmi 573 gènes prédits par IMGAG.

Ils ont été nommés, pour plus de facilité, « Mt » pour *Medicago truncatula*, suivi du nom du gène. Quatre d'entre eux (MtCO, MtFTa, MtFTb et MtFTc) sont au centre de l'intervalle de confiance (1.3 cM) alors que MtPKS et MtFD sont sur le BAC AC149134 où est ancré le marqueur borne 19|15a (2.5 cM) (Figure 24).



- 1- MtPKS (AC149134_39.2) : présente une forte similarité avec le gène PHYTOCHROME KINASE 1 (PKS), codant pour un substrat de phosphorylation pour le phytochrome B (PHYB) et jouant un rôle négatif sur l'activité de ce phytochrome (Fankhauser et al. 1999). L'EST de *M. truncatula* présentant le plus de similarité est BF644398. La protéine module l'activité kinase de PHYB et pourrait ainsi influencer sa localisation sub-cellulaire. Des plantes transgéniques sur-exprimant PKS1 montrent en effet un phénotype similaire à celui du mutant *PHYB*.

Or, il a été démontré que PHYB joue un rôle inhibiteur sur l'initiation florale des plantes de jours courts ou de jours longs avec néanmoins une inhibition plus prononcée sous des conditions de photopériode inhibitrice. Chez le pois, un mutant *PHYB (iv-1)* présente un phénotype précoce et une moindre sensibilité à la photopériode (Weller et al. 1995).

Parmi les six gènes sélectionnés, les cinq locus suivants, regroupés sur une distance d'environ 1.7 cM présentent une forte similarité avec des gènes connus comme étant impliqués dans une cascade d'interactions géniques promouvant la floraison :

- 2- MtCO (AC133780_7.1) : présente une forte similarité avec le gène CONSTANS d'*A. thaliana* qui peut être considéré comme un des gènes clés contrôlant la transition florale chez cette plante (Putterill et al. 1995) et dont l'expression est contrôlée par l'horloge circadienne (Hayama and Coupland 2004). Il code pour une protéine à domaine B-BOX en « doigt de zinc » s'exprimant principalement dans les feuilles avant floraison et induisant l'expression du gène FT chez *A. thaliana* (Wigge et al. 2005; Imaizumi and Kay 2006). L'EST TC86982 de *M. truncatula* présente le plus de similarité avec la séquence codante.
- 3- , 4- ,5- Trois séquences génomiques en tandem sur le BAC AC123593 de *M. truncatula* ont déjà été nommées respectivement MtFTLa (AC123593_3.2), MtFTLb (AC123593_7.2) et MtFTLc (AC123593_9.2) (Hecht et al. 2005). Ils présentent tous trois une forte similarité avec le gène FT d'*A. thaliana* dont le transfert de l'ARNm des feuilles à l'apex constituerait, du moins en partie, le signal « florigène » et accélère ainsi la transition florale en condition de photopériode inductive (Huang et al. 2005). Ce gène semble très conservé chez les Angiospermes (Imaizumi and Kay 2006; Lifschitz and Eshed 2006). Il a été démontré que ce gène s'exprime d'abord dans les feuilles et qu'il est impliqué dans une boucle de rétroaction positive provoquant une augmentation de sa propre expression (Huang et al. 2005). Il n'a pas été trouvé d'EST qui corresponde spécifiquement à ces trois gènes (MtFTLa, MtFTLb et MtFTLc) dans les bases de données ce qui laisse douter de leur expression (Hecht et al. 2005).

6- MtFD (AC149134_18.1) présente une forte similarité de séquence avec le gène FD d'*A. thaliana*, un facteur de transcription interagissant avec la protéine FT et initiant l'expression séquentielle des gènes de l'identité florale du méristème au niveau des cellules méristématiques de l'apex (Abe et al. 2005; Blazquez 2005). L'EST TC99455 présente une forte similarité avec la séquence codante MtFD.

Le résultat de l'alignement effectué sur les bases de données protéiques (BLASTp) et les principales caractéristiques des séquences similaires aux gènes candidats sélectionnés sont présentés Tableau 22, Tableau 23 et Tableau 25.

Tableau 22 Résultats du BLASTp pour les gènes candidats sélectionnés (matrice BLOSUM62, banque de données uniprot_swissprot)

Gène	Numéro d'accèsion	Taille de la séquence protéique (acides aminés)	Séquence similaire chez <i>A. thaliana</i>	E value
MtFTLa	AC123593_3.2	177	FLOWERING LOCUS T	$4 e^{-72}$
MtFTLb	AC123593_7.2	131	FLOWERING LOCUS T	$6 e^{-32}$
MtFTLc	AC123593_9.2	172	FLOWERING LOCUS T	$2 e^{-53}$
MtCO	AC133780_7.1	416	Protéine en doigt de zinc CONSTANS-LIKE 15	$9.3 e^{-73}$
MtPKS	AC149134_39.2	498	PKS1_ARATH Phytochrome kinase substrate 1	$7 e^{-04}$
MtFD	AC149134_18.1	320	Protéine FD (Facteur de transcription BZIP14) AtbZIP14	$8 e^{-15}$

MtFTLa, MtFTLb et MtFTLc présentent de fortes similarité avec la séquence protéique du gène FT d'*A. thaliana*. Cependant, MtFTLb contient une séquence plus courte d'environ 40 acides aminées. MtCO montre une forte ressemblance avec CONSTANS-LIKE alors que MtPKS et MtFD ont une ressemblance plus faible avec les séquences protéiques prédites, respectivement, PKS1 et AtbZIP14.

Tableau 23 Caractéristiques des séquences similaires aux gènes candidats positionnels étudiés identifiées par BLASTp. Toutes les séquences de gènes ont été obtenues chez *A. thaliana*

Gène	Type de protéine codée	Fonction	Références
FT (Flowering locus T)	Raf-kinase	Signal florigène Induction des gènes FD et SOC1	Huang et al. 2005
CO (Constans)	Protéine en “doigts de zinc”	Accélération de la floraison en conditions de photopériode inductive	Putterill et al. 1995
PKS (Phytochrome Kinase Substrate)	Substrat de phosphorylation	Régulation de les voies de signalisation impliquant les phytochrome A et B	Lin, 2000
FD (Flowering locus D)	Facteur de transcription “b- ZIP”	Induction des gènes conduisant à la formation des différentes structures florales	Blazquez, 2005

4.2.6 Séquençage des gènes candidats

L'objectif est d'étudier le polymorphisme de séquence (ADNg et ADNc) chez deux lignées pour éventuellement le relier aux différences de précocité observées.

Le séquençage est réalisé chez les deux lignées parentales contrastées pour leur date de floraison (Jemalong6 précoce, DZA315.16 tardif) pour les six gènes candidats, excepté le gène MtCO pour lequel le séquençage a été effectué sur les quatre lignées parentales. Une étape de mise au point des amorces a été effectuée. Plusieurs couples d'amorces ont été testés pour leur spécificité et leur niveau d'amplification, et deux ont été choisis par gène (ADNg et ADNc) (Tableau 24).

- 1- L'ADNc de MtPKS a été séquençé en partie (294 pb entre les positions 529 et 823 pb du gène) et ne révèle aucun polymorphisme de séquence.
- 2- MtCO présente du polymorphisme exonique de changement de paire de bases (SNP) en position +389, +1 368, +2 413 et +2 511 pour respectivement DZA315.16 et DZA45.5, F83005.5, DZA45.5, DZA315.16. Il y a une délétion de séquence GAA chez F83005.5 en position +407 (
- 3-

4-

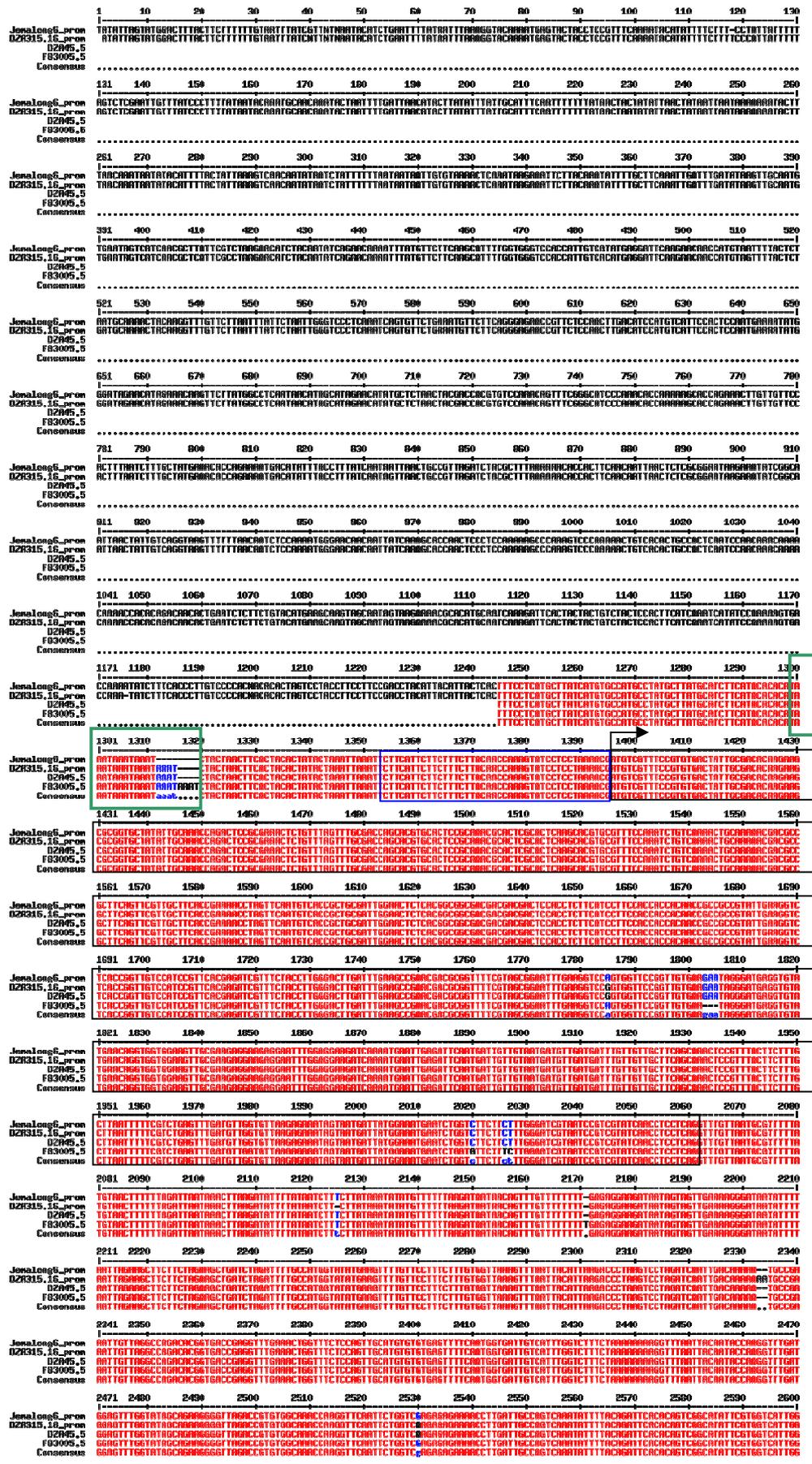
5- Figure 25). Le séquençage de l'ARNm pour Jemalong6 et DZA315.16 ne montre pas de polymorphisme.

Le séquençage du promoteur a été effectué 1400 pb en amont du codon ATG. La recherche de cis-élément révèle entre autre la présence d'une boîte TATA dont la structure est proche de celle trouvé dans un gène du pois (Tjaden et al., 1995). L'alignement présenté en

Figure 25 montre la présence de polymorphisme de répétitions du motif AAT dans la boîte TATA chez DZA315.16, DZA45.5 et F83005.5 par rapport à Jemalong6. Il y a une répétition de motif pour DZA315.15 et DZA45.5, et deux répétitions pour F83005.5. Cette différence de nombre de répétitions pourrait entraîner une modification de la fixation de l'ARN polymérase II lors de la transcription de l'ADN en ARNm (Figure 26). En effet, des délétions ou mutations de la boîte TATA n'abolissent pas totalement la transcription. Les deux effets observés sont : une diminution du taux de transcription et une perte de la fidélité du site de transcription (O'Shea-Greenfield and Smale 1992). La traduction protéique effectuée à partir de ces données ne permet pas d'envisager de modifications induisant une protéine tronquée ou non-fonctionnelle chez Jemalong6 ou DZA315.16.

Tableau 24 Couples d'amorces permettant d'amplifier spécifiquement les six gènes candidats (ADN génomique et ADNc *)

Gène	BAC	Amorce sens		Amorce anti-sens		T _a	[MgCl ₂] (mM)	ADNc (*)	Longueur (pb)
		Nom	Séquence	Nom	Séquence				
MtPKS	AC149134	PKS_B_1F	GCAGGTGCTATCTCAGTTAA	PKS_B_2R	TGTTGAAGATGAATTAGGGT	56	1.5		1296
		PKS_B_1F	GCAGGTGCTATCTCAGTTAA	PKS_B_2R	TGTTGAAGATGAATTAGGGT	56	1.5	*	369
MtCO	AC133780	Co7F	AATCCAGACGCCAAA	Co8R	TAGTCCTACCTTCCTCCGA	57.3	1.5		2461
		Co5F	ATGGGATGAGGTTACTTGG	Co6R	GAGCTGAAATTTCTGAGGGA	55.3	1.5		1445
		Co33F	TCTGGGATTTTAAAYTAYAG	Co31R	GCTAATCTTGGNATHAYGT	55	3	*	1199
		Co9F	CTCTCACTCTTTCTCG	Co12R	CAGACTCCGCGAAACTC	54	1.5	*	1268
		Co40F	CCACACACTTCAACTTCTTA	Co41R	GTGTATGAAGATGCATAAGC	54.2	3		1350
		FT1F	CATATCCGAGGAACACTTGA	FT2R	AGCAATGAGTAGTGGGCAG	56.7	3		2790
MtFTa	AC123593	FT7F	ATGGCTGGTAGCAGTAGGAA	FT8R	GCCACTCTCCCTCTGACA	61.7	3	*	507
		FT3F	AATACATATCCTTTACTCGC	FT4R	GACTGTGTTGATAGACGA	51.4			1125
MtFTb	AC123593	FT9F	ATGGCAAGTGGTAGCAGACC	FT11R	TGAACTTCTAATTCTGTTT	52.9	3	*	393
		FT5F	TTGCTCCTTCATGACTGG	FT6R	AATTGATCGACACTGCCTAA	55.4	3		2961
MtFTc	AC123593	FT12F	AAATTTGGTCCGCCCTCTTG	FT13R	ATCTCCTTCCGCCCGC	52.9	2	*	503
		FD1F	GAGGCAGAAGAGGATGATTA	FD2R	GAAACTTTGATCTCCAGCTC	58	1.5		500
MtFD	AC149134	FD1F	GAGGCAGAAGAGGATGATTA	FD2R	GAAACTTTGATCTCCAGCTC	58	1.5		500
		FD1F	GAGGCAGAAGAGGATGATTA	FD2R	GAAACTTTGATCTCCAGCTC	58	3	*	191



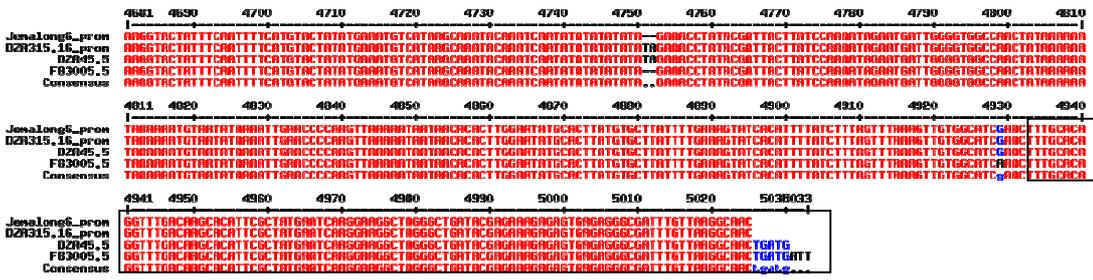
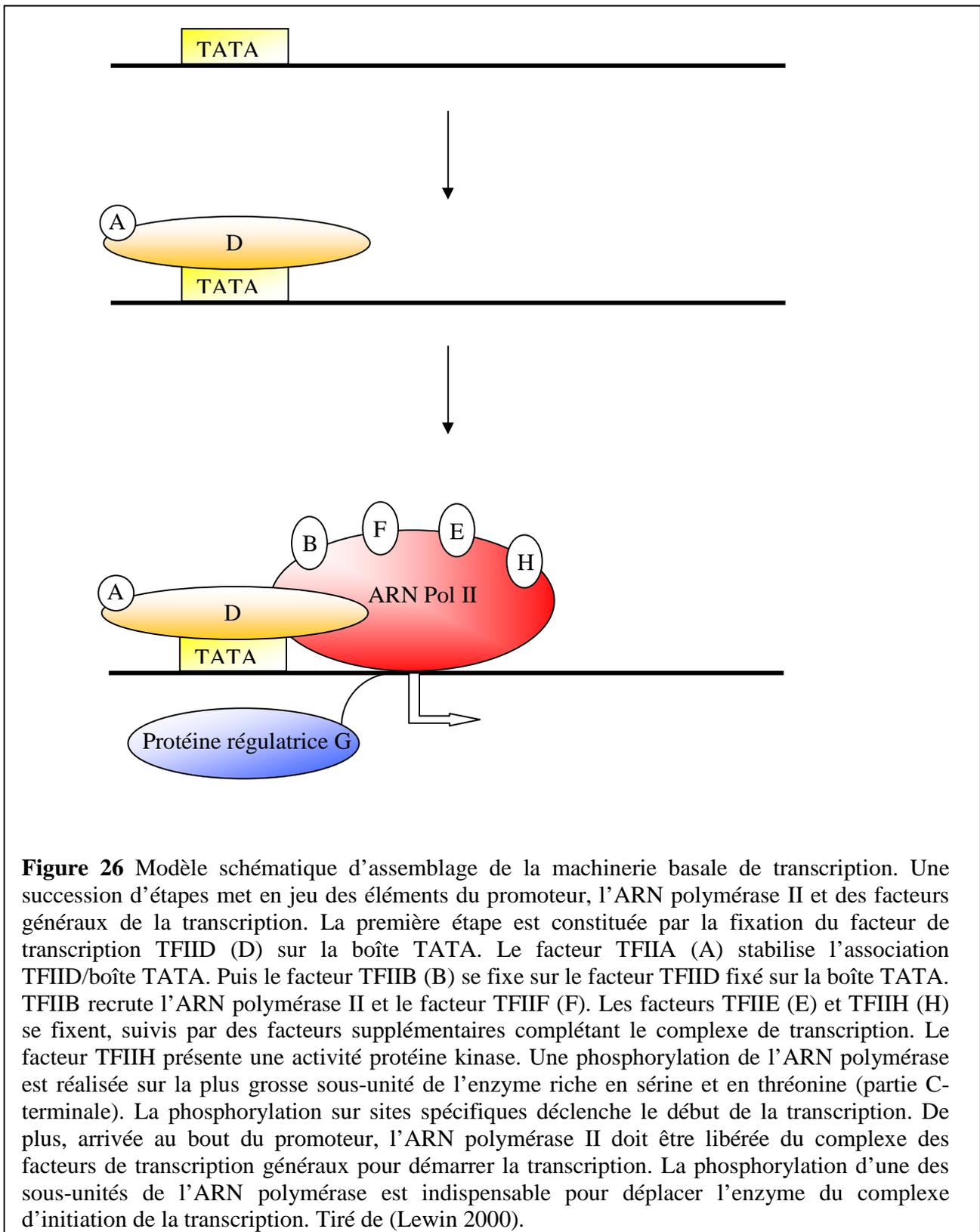
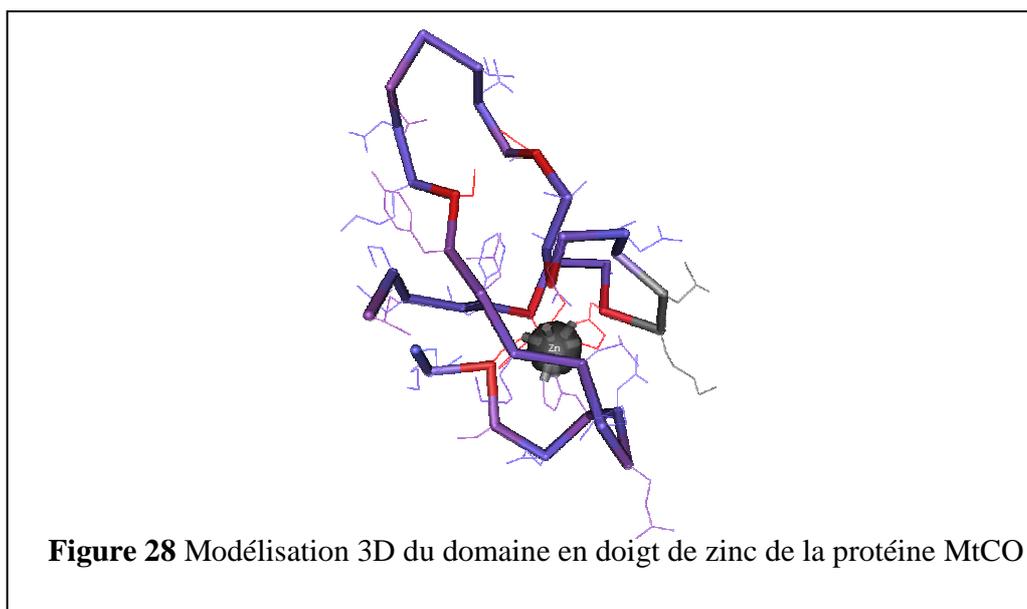
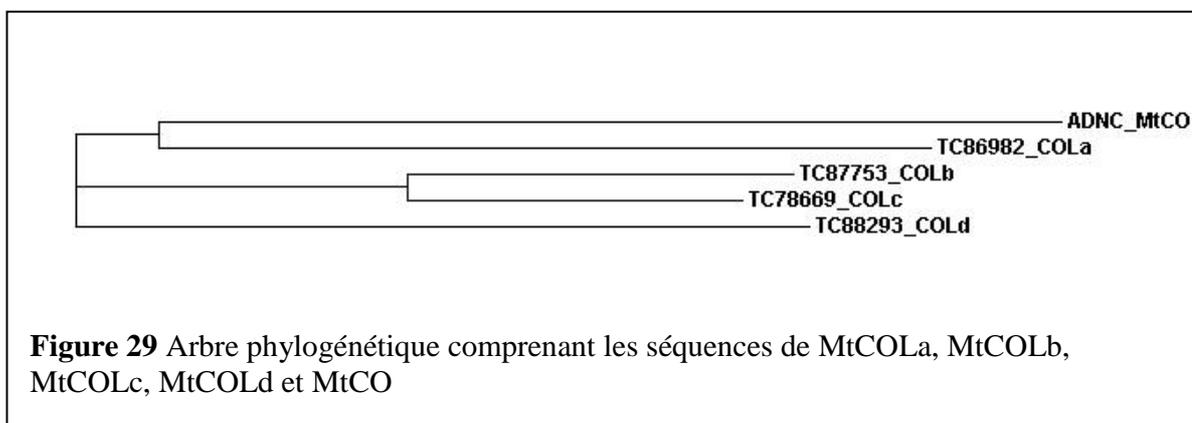


Figure 25 Alignement (logiciel Multalign) des séquences génomiques pour Jemalong6, DZA315.16, DZA45.5 et F83005.5 et les séquences promotrices (-1400pb du codon ATG) de Jemalong6 et DZA315.16 pour le gène MtCO. La boîte TATA est représentée par un cadre vert, les exons par un cadre noir et le codon ATG par une flèche. La boîte 3'UTR s'étend de la position 5 042 à 5196pb, soit, sur le BAC de la position 77 651 à 77 805pb. La boîte 5'UTR est représentée par un carré bleu.





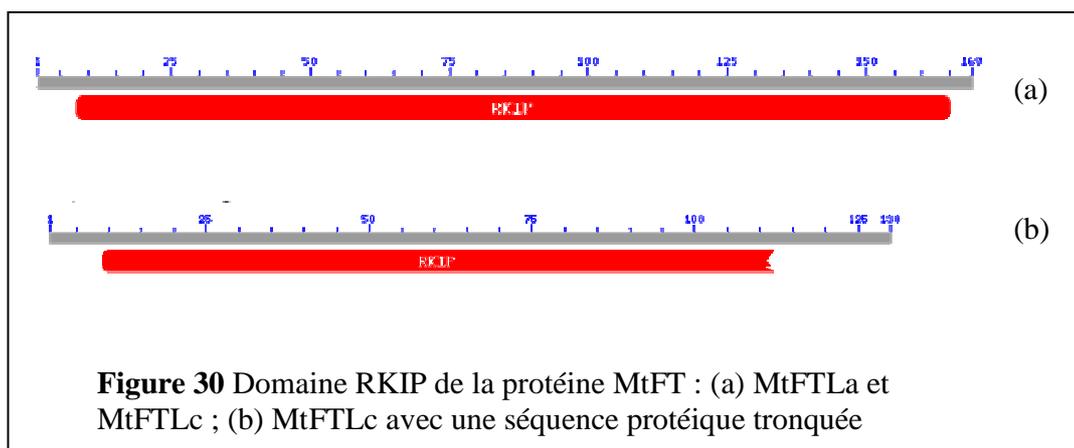
Quatre groupes de CONSTANS (COLa à COLd) ont été décrits dans la bibliographie chez *M. truncatula* (Hecht et al. 2005). La séquence d'ADNc de Jemalong6 a été comparée à quatre séquences caractéristiques de ces différents groupes dans un arbre phylogénétique (Figure 29).



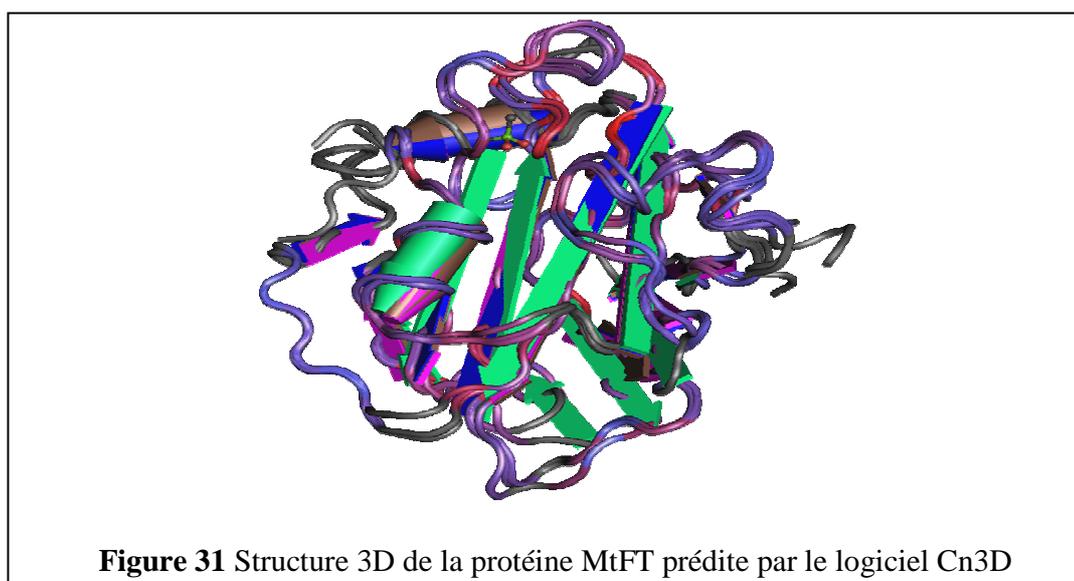
MtCO sur le chromosome 7 appartient au groupe COLa, qui contient MtCOLa (TC86982) et PsCOLa. Ces séquences sont proches des gènes CO / COL1 / COL2 chez *A. thaliana*.

- 6- ,4- ,5- MtFTa et MtFTc ne présentent aucun polymorphisme de séquences, que ce soit dans l'ADNg ou l'ARN. Le séquençage de MtFTb révèle une insertion exonique de 2 pb chez DZA315.16 qui entraîne un codon STOP et donc une protéine tronquée à 123 acides aminés au lieu de 131. MtFTLa, MtFTLb et MtFTLc codent pour une protéine RKIP (*Raf Kinase Inhibitor Protein*).

Les protéines RKIP inhibent la phosphorylation et l'activation d'autres protéines en interrompant la formation de certains complexes. Il est également possible que les protéines RKIP aient d'autres fonctions comme celles d'une molécule d'échafaudage qui soutient l'activation de cascade de gènes dans certaines circonstances ou localise cette cascade à une organelle spécifique (Chetoui 2005). Le codon STOP présent chez MtFTLb pourrait rendre sa protéine RKIP non-fonctionnelle (Figure 30).



La structure secondaire se compose de sept hélices alpha, douze feuillets bêta. Cette protéine MtFT est visualisable (prédiction) dans son intégralité en 3D dans la Figure 31.



Les promoteurs de ces trois gènes n'ont pas été séquencés. L'absence d'EST propres à ces trois gènes dans les bases de données internationales ne permet pas de prédire l'expression spécifique de MtFTLa, MtFTLb ou MtFTLc.

- 6- L'ADNc de MtFD a été séquencé en partie (130 pb entre les positions 808 et 938 pb du gène) et ne révèle aucun polymorphisme de séquence entre Jemalong6 et DZA315.16.

Le bilan des séquences est présenté dans le Tableau 25. La position des exons dans les gènes MtPKS et MtFD sont prédits par le logiciel IMGAG et n'ont pas pu être vérifiés. MtCO, MtFTLa, MtFTLb et MtFTLc ont un découpage introns - exons effectué à partir du séquençage de l'ADN génomique et de l'ARNm pleine longueur. MtPKS ne possède pas d'intron, la longueur de la séquence génomique est équivalente à la longueur de l'ARNm. Les trois MtFT présentent des différences importantes de longueur d'ADN génomique (588, 1704 et 2139 pb). Ces différences sont moins importantes pour l'ARNm entre MtFTa et MtFTc. MtFTb a une longueur d'ARNm inférieure aux deux autres copies.

Tableau 25 Positions des gènes candidats sur les BAC, longueur des gènes avec le découpage en exons et la longueur de l'ARNm

Gène	BAC		Longueur ADNg (pb)	Exons		Longueur ARNm (pb)
	Numéros	Position (pb)		Nombre	Position (pb)	
MtPKS	AC149134	161253 - 162749	1496	1	1 - 1496	1496
MtCO	AC133780	50126 - 53771	3645	4	1 - 577 1325 - 1468 2321 - 2582 2832 - 2854	1003
MtFTLa	AC123593	10132 - 11836	1704	4	1 - 203 247 - 310 415 - 457 961 - 1186	533
MtFTLb	AC123593	30118 - 30706	588	3	1 - 206 291 - 354 456 - 583	396
MtFTLc	AC123593	41468 - 43607	2139	4	1 - 194 456 - 519 647 - 689 1867 - 2088	520
MtFD	AC149134	109454 - 112188	2734	2	1 - 979 2654 - 2734	1059

4.3 Expression de quelques gènes candidats de la date de floraison

Les six gènes décrits ont été étudiés pour leur expression sur les lignées Jemalong6 et DZA315.16, cultivés en photopériode courte (12 h). Le but est de relier la différence de floraison observée entre Jemalong6 (précoce) et DZA315.16 (tardif) à celle du taux d'ARNm transcrits pour chaque gène candidat.

Le niveau d'expression des gènes candidats sélectionnés par étude bio-informatique et analyse de séquence a été déterminé par RT-PCR semi-quantitative. Tous les prélèvements ont été effectués entre 9 et 11 heures du matin, soit après 1 à 3 heures d'éclairage. Les conditions optimales de RT-PCR semi-quantitative sont présentées dans la Tableau 26.

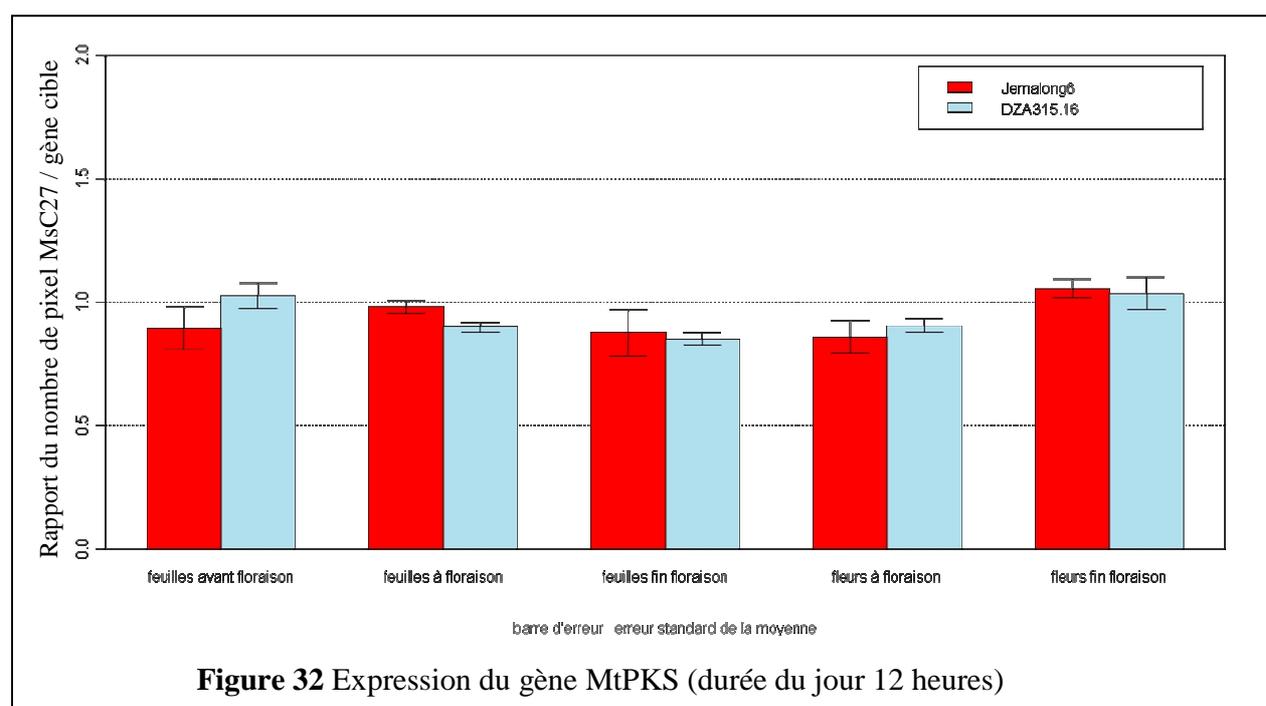
Tableau 26 Paramètres optimum pour l'amplification par RT-PCR semi-quantitative des gènes candidats sélectionnés à l'intérieur de l'intervalle de confiance du QTL

Numéro d'accension	Nom du gène	Ta (°C)	[MgCl ₂] (mM)	Nombre de cycles	Taille de l'amplicon (pb)
AC123593_3.2	MtFTLa	55.2	1.5	-	507
AC123593_7.2	MtFTLb	57.3	3.0	35	255
AC123593_9.2	MtFTLc	63.2	2.0	-	503
AC133780_7.1	MtCO	53.5	3.0	31	1268
AC149134_39.2	MtPKS	56.0	3.0	35	369
AC149134_18.1	MtFD	57.3	3.0	35	103
MsC27	MsC27	53.0 à 58.0	1.5 à 3.0	32	311

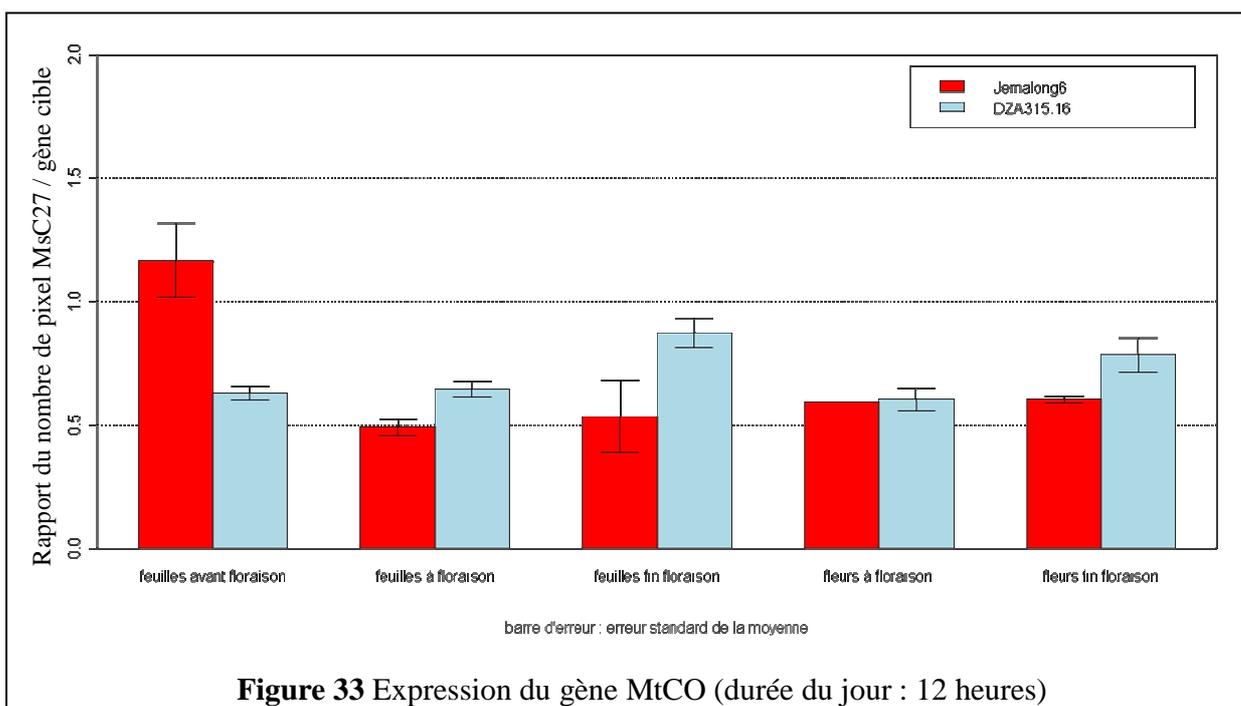
La technique met en œuvre deux couples d'amorces : le gène candidat et le gène témoin MsC27. L'utilisation du gène MsC27 en tant que témoin du niveau d'expression s'est avérée techniquement très avantageuse. En effet, ce gène présente des conditions d'amplification très flexibles avec une température d'hybridation des amorces comprise entre 53 °C et 58 °C et la possibilité de l'amplifier dans une gamme de concentrations en MgCl₂ comprises entre 1.5 et 3 mM avec un bon niveau d'amplification et de spécificité. Ces caractéristiques ont permis d'amplifier MsC27 avec l'ensemble des gènes sélectionnés et ainsi de réaliser un suivi d'expression semi-quantitatif des gènes sélectionnés.

A l'aide d'un outil informatique, un rapport d'expression est établi sur photographie d'un gel d'agarose par comparaison du nombre de pixels entre le gène cible et le gène témoin pour Jemalong6 et DZA315.16. Cette opération permet de confronter le niveau d'expression de chacun des deux parents. La mise au point des paramètres est primordiale pour l'analyse informatique.

- (i) étude transcriptomique des gènes MtFTLa et MtFTLc : ces deux gènes n'ont pas pu être amplifiés sur ADNc pour une raison inconnue et ce, malgré l'utilisation de différents couples d'amorces qui étaient opérationnels sur ADNg. Soit ces gènes ne s'expriment pas aux stades et dans les organes analysés, soit leurs ARNm ont été dégradés avant la rétro-transcription ;
- (ii) étude transcriptomique des gènes MtPKS, MtFD et MtFTLb : aucune variation significative de l'expression de MtFD ou de MtFTLb n'a pu être mise en évidence dans les différents types d'échantillons étudiés. MtPKS montre une légère diminution de son niveau d'expression dans les feuilles pour les stades avant floraison et après floraison chez DZA315.16 ($P < 0.05$). Cette légère variation n'a pas été observée chez Jemalong6. Excepté dans le cas des feuilles prélevées à floraison où une expression significativement plus forte de MtPKS chez Jemalong6 a été détectée ($P < 0.05$), le niveau d'expression de ces trois gènes s'est avéré similaire entre les deux lignées dans tous les organes et à tous les stades de prélèvement (Figure 32) ;

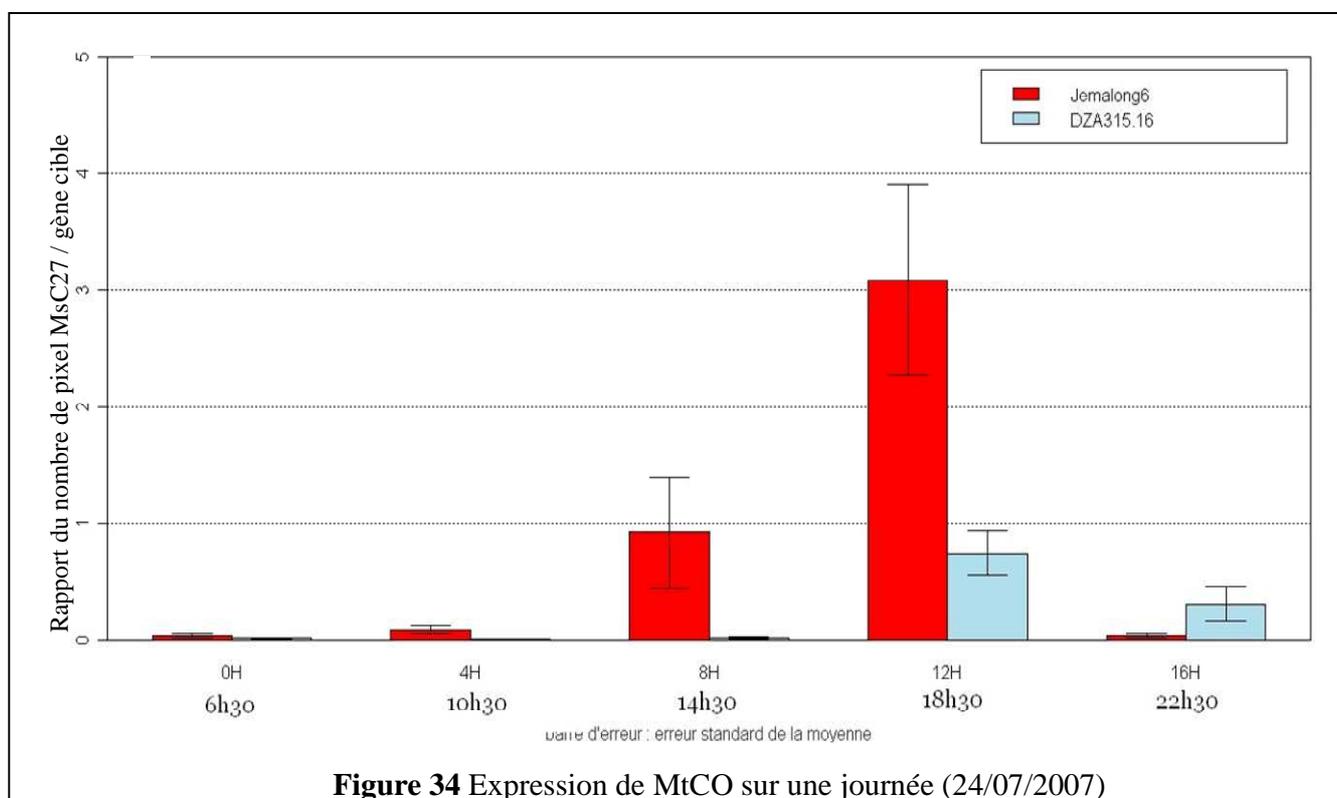


- (iii) étude transcriptomique du gène MtCO : chez Jemalong6, une diminution nette de l'expression de ce gène a été observée entre les échantillons prélevés avant floraison et ceux prélevés à floraison ($P < 0.05$), son expression ne variant plus dans les feuilles à la fin de la floraison. Au contraire, chez DZA315.16, alors que l'expression de MtCO reste stable jusqu'à la floraison, celle-ci augmente significativement en fin de floraison ($P < 0.05$). Dans les échantillons de fleurs, aucune variation significativement différente de l'expression de MtCO n'a pu être enregistrée quelle que soit la lignée étudiée. L'expression de ce gène est apparue similaire entre Jemalong6 et DZA315.16 dans les échantillons de fleurs. La comparaison de l'expression de ce gène chez les deux lignées étudiées a permis de mettre en évidence des niveaux d'expression significativement différents dans les échantillons de feuilles prélevés avant floraison ($P < 0.05$) (Figure 33).



En effet, dans ce type d'échantillon, il a pu être observé une expression deux fois plus importante de MtCO chez Jemalong6 par rapport à DZA315.16. Cependant, le gène CO est décrit dans la bibliographie pour avoir une expression cyclique durant la journée (dégradation de son ARN pendant la nuit). Afin de confirmer la piste de ce gène candidat, l'étude de son expression sur une journée a été réalisée.

L'étude transcriptomique de MtCO sur une journée (24/07/2007) à partir de cinq points de prélèvement permet d'observer les différences d'expression entre les deux lignées Jemalong6 et DZA315.16. Les prélèvements ont été effectués en même temps chez Jemalong6 et DZA315.16 avant la réalisation de la RT-PCR semi-quantitative. Le profil d'expression suit une courbe chez les deux lignées avec un pic d'expression 12 heures après le lever du soleil, c'est-à-dire à 18 heures 30. Malgré une forte variabilité d'expression, des différences du taux d'expression de MtCO sont visibles entre Jemalong6 et DZA315.16 après 4 et 8 heures de jour ($P = 0.06$). Douze heures après le lever du soleil, une différence d'expression significativement plus importante de MtCO est observée chez Jemalong6 que chez DZA315.16 ($P < 0.05$) (Figure 34). En fin de journée (22 h 30), le niveau d'expression a diminué chez les deux lignées mais reste plus élevé pour DZA315.16. Il semble donc que la différence de précocité entre Jemalong6 (précoce) et DZA315.16 (tardif) soit imputable au gène candidat MtCO, inducteur de la floraison. En effet, son expression plus élevée pour Jemalong6 que DZA315.16 pourrait induire une floraison plus précoce chez Jemalong6.



DISCUSSION GÉNÉRALE ET PERSPECTIVES

5 DISCUSSION GÉNÉRALE ET PERSPECTIVES

5.1 Effet de la photopériode sur la date de floraison

L'altération de la date de floraison en réponse à la photopériode correspond aux interactions entre les signaux environnementaux et l'horloge circadienne (rythmes endogènes ayant une période de 24 heures) (Hayama and Coupland 2004). Les *Medicago* annuelles sont connues pour être des plantes vernalisables de jours longs (Clarkson and Russell 1975). En chambre de culture avec des photopériodes variant de 12.2 à 16 heures et une température entre 15 et 21.3 °C, (Moreau et al. 2007) ont montré que la photopériode a un impact mineur sur la date de floraison de Jemalong6 comparée à l'influence de la vernalisation (4 °C appliqués de 1 à 21 jours selon les expériences). D'autres auteurs ont observé que l'effet de la vernalisation sur la date de floraison de Jemalong est plus important en conditions de jours longs, mais qu'il peut devenir secondaire par rapport à l'effet de la photopériode chez d'autres génotypes (Chabaud et al. 2006).

La date de floraison influence la possibilité de réaliser un cycle complet de reproduction. La photopériode varie selon la saison et la latitude géographique. Elle génère des changements dans la morphogénèse des plantes car la réponse à la durée du jour et de la nuit représente implicitement l'énergie lumineuse totale (Thomas and Vince-Prue 1997). Peu de connaissances étaient disponibles sur l'effet de la durée de la photopériode sur la date de floraison chez *M. truncatula*, ainsi que la part de la variation génétique pour cette réponse à la photopériode. Dans cette étude, la date de floraison des huit lignées de *M. truncatula* (provenant du bassin Méditerranéen et d'Australie) est plus précoce en jours longs qu'en jours courts. De plus, la morphologie des plantes est fortement altérée. Ceci montre nettement que la durée de photopériode, qui correspond ici uniquement à la perception de la longueur du jour (changement de cycle jour / nuit), est d'une grande influence sur le développement phénologique et morphologique des plantes. Une interaction significative entre lignée et photopériode a été observée : la date de floraison est plus accélérée par une durée longue de photopériode pour F83005.5 et Jemalong6 que pour les autres lignées alors que Levahim-B est peu influencé.

Les différences entre les lignées pour la date de floraison en réponse à la photopériode pourraient indiquer une adaptation à l'environnement. En effet, chez le riz (*O. sativa*), les variétés tropicales sont des plantes de jours longs tandis que les variétés non-tropicales sont insensibles à la photopériode (Thomson et al. 2006). Chez le maïs (*Zea mays*), l'exposition à des jours longs retarde le moment de l'initiation et du développement de l'inflorescence terminale, bien qu'il existe de nombreuses variations entre les hybrides (Krüger 1984). On peut ainsi concevoir que des populations de *M. truncatula* se sont naturellement adaptées pour leur forte / faible sensibilité à la photopériode, dans l'objectif d'assurer la floraison à un moment optimal dépendant des contraintes climatiques. Les populations avec une floraison précoce en photopériode courte (floraison précoce au printemps) seraient adaptées à des climats à printemps tièdes et étés chauds, le cycle de vie étant court. Les lignées avec une floraison tardive en photopériode courte (floraison tardive au printemps) échapperaient au froid dans les climats à printemps frais mais pourraient être affectées par un stress hydrique en fin de printemps. Dans une collection d'accessions méditerranéennes de *M. truncatula*, une corrélation a été établie entre la date de floraison et la pluviométrie annuelle des sites de collecte, elle-même reliée à la latitude, les accessions les plus tardives à floraison étant issues des lieux les plus arrosés et les plus septentrionaux (Julier et al. 2007). La réponse à la photopériode contribue à cette adaptation.

5.2 La modélisation de la date de floraison

L'intérêt majeur de la modélisation réside dans son utilisation pour la prédiction de la date de floraison. Dans le modèle ($1/f = a + bT + cP$), les paramètres a , b , et c ont été spécifiquement estimés pour chacune des quatre lignées étudiées. Les valeurs de la date de floraison prédite sont proches de celles observées avec un R^2 compris entre 0.93 et 0.99. Les modèles établis permettent une bonne prédiction de la date de floraison dans les conditions de culture. Ils sont cependant calculés sur un nombre limité de données, provenant d'essais réalisés en chambre de culture et en serre. Ainsi, les valeurs intégrées de température et de photopériode des essais menés en serre sont des moyennes depuis le semis jusqu'à la floraison de la lignée considérée. Le modèle ne prend pas en compte les fluctuations de température ni de photopériode qui ont pu avoir lieu. La généralisation de ce modèle avec les valeurs spécifiques des lignées et son application à des conditions différentes de culture demanderait l'ajout des données expérimentales.

Enfin, il serait utile de prendre en compte la durée de la vernalisation dans le modèle afin d'étudier son impact sur la floraison. Cependant, il n'existe pas de modèle décrit intégrant photopériode, température et vernalisation.

(Moreau et al. 2007) ont étudié l'impact de la durée de vernalisation sur la date de floraison. Pour une durée de vernalisation identique à celle de notre étude (7 jours à 4 °C), les valeurs de la date de floraison de la lignée Jemalong6 ont été prédites par le modèle de Jemalong6 (Tableau 27).

Tableau 27 Modélisation de la date de floraison observée chez Jemalong6 dans les conditions d'expérience de (Moreau et al. 2007)

Expérimentation	Température	Photopériode	Date de floraison	
			Observée	Prédite
Chambre de culture	19.0	18.0	33.3	34.9
Chambre de culture	19.0	12.0	76.5	83.3
Printemps 2005	22.8	15.4	37.0	35.6
Automne 2005	17.7	15.4	70.5	52.2
(Moreau et al. 2007)	15.3	12.2	45.8	165.2

La différence importante entre la date de floraison observée et prédite (119.4 jours) peut être expliquée par les températures basses subies en serre, qui sont autant de jours de vernalisation accélérant la floraison, non pris en compte par le modèle. De plus, dans le dispositif de (Moreau et al. 2007) les graines pré-germées ont subi une photopériode de 16 h pendant un jour. Il a été constaté que des graines pré-germées qui ont passé trois jours à une photopériode de 18 h avant d'être transférées à 12 h avaient une date de floraison et une morphologie semblable à celles restant à 18 h (observation personnelle). On peut donc envisager que le traitement de 16 h subi par les plantes pré-germées ait influencé le développement de Jemalong6 et donc de sa date de floraison.

5.3 Analyse QTL multi-croisements

Cette analyse QTL est la première réalisée chez *M. truncatula* à partir de plusieurs populations. Une analyse similaire menée sur le maïs, et plus spécifiquement sur l'architecture génétique de la date de floraison, a permis de préciser les QTL contrôlant le caractère (Chardon et al. 2004). Notre étude fournit des indications fortes sur le contrôle génétique de la date de floraison chez *M. truncatula*.

La précision de la cartographie QTL dépend de nombreux facteurs comme l'héritabilité du caractère, le nombre de QTL impliqués, la distribution des QTL le long du génome, le nombre de marqueurs ainsi que leur distribution sur le génome, et enfin la taille de la population (Knapp and Teuber 1990). La précision de la position du QTL dépend plus de la taille de la population que du nombre de marqueurs. Pour cette raison, des populations relativement grandes ont été utilisées, avec des cartes cadres. Un plan d'étude multi-croisements offre la possibilité d'augmenter la taille de la population pour la détection de QTL. Ce type de dispositif augmente la probabilité qu'un polymorphisme allélique soit présent dans les lignées parentales, permettant le repérage de QTL qui seraient indétectables dans un croisement simple où les deux parents seraient fixés pour le même allèle. Elle permet également l'estimation des effets du QTL pour les différentes lignées parentales, et augmente la puissance de détection du QTL, donnant une position du QTL plus exacte. L'accumulation d'informations provenant des différents croisements est un atout considérable dans la détection de QTL (Rebai and Goffinet 1993; Muranty 1996). Ainsi, l'obtention d'un QTL commun aux trois populations étudiées a permis d'affiner sa localisation sur cette région du chromosome 7. L'ajout de nouveaux marqueurs dans la région du QTL majeur sur le chromosome 7 a permis de préciser la position du QTL. Un résultat frappant de notre étude sur les trois populations de cartographie est l'obtention d'un seul QTL expliquant 25 % de la variation dans les trois populations. Sa position a été estimée avec un intervalle de confiance de seulement 0.9 cM, car la population globale était d'environ 600 RILs. Un autre QTL sur le chromosome 8 explique 3 % de la variation. Ces deux QTL et le fond génétique dans le modèle utilisé par MCQTL qui prend en compte l'ensemble des populations et le rôle de chacune d'entre-elles (date de floraison = population + QTL1 + QTL2) expliquent 80 % de la variation globale.

Le fait que certains QTL spécifiques de populations ne soient pas détectés dans cette approche multi-croisements est dû à leur faible contribution sur la variation dans l'ensemble des trois populations. Le QTL du chromosome 7 était présent dans les trois populations et avait un effet fort sur la date de floraison quelle que soit la saison. Les QTL avec de forts effets sont souvent contrôlés par des gènes majeurs qui peuvent alors être clonés. Le gène Hd1 chez le riz explique un QTL majeur contrôlant la réponse de la date de floraison à la photopériode, EDI chez *A. thaliana* explique un QTL majeur déclenchant la floraison en photopériode de jours longs (Yano et al. 2000; El-Assal et al. 2001). Ainsi l'hypothèse d'un gène majeur impliqué dans le QTL du chromosome 7 observé dans les trois populations est vraisemblable.

5.4 Cartographie fine et cartographie physique

La carte génétique réalisée pour la cartographie fine du QTL majeur du chromosome 7 présente quelques divergences avec la carte physique de *M. truncatula* proposée en ligne (<http://www.medicago.org/genome/>). Cette carte physique publique, disponible pour la communauté scientifique, a été réalisée à partir du séquençage partiel de la lignée Jemalong6 de *M. truncatula* (assemblage des BACs en contigs), puis les contigs ont été ordonnés en suivant la ségrégation de marqueurs dans une population F2 de 69 individus. Cette carte présente néanmoins des points faibles importants : le séquençage n'étant pas complètement terminé, de nombreux contigs ne sont pas ancrés les uns aux autres et de nombreuses coupures jalonnent la carte. De plus, l'effectif faible (69 individus) et le type de population étudiée (F2) entraînent une mauvaise estimation des distances génétiques et un mauvais ordonnancement des contigs.

Cependant, la carte physique constitue un solide appui. En effet, couplée à une carte génétique robuste établie par cartographie fine, elle fournit la liste des contigs de BAC contenant les marqueurs préalablement ordonnés, et établit ainsi une liste de gènes candidats. Une telle liste de gènes candidats ne peut être considérée comme exhaustive pour trois raisons majeures :

- (i) il est évident que la présence de coupures entre les contigs limite cette liste ;

- (ii) étant donné que seuls les contigs cartographiés dans l'intervalle de confiance du QTL détecté en F6 ont été utilisés pour établir cette liste, certains contigs mal cartographiés et donc certaines prédictions de gènes ont pu être omises ;
- (iii) certains contigs de BAC ne sont pas pris en compte car leurs marqueurs ne ségrègent pas comme h4_42i13a dans cette étude. Ils n'ont donc pas été cartographiés dans l'intervalle.

Pour la mise en oeuvre de cette méthode de gènes candidats positionnels, le choix des gènes candidats repose sur leur localisation et donc leur co-ségrégation avec le locus d'intérêt (Morot-Gaudry and Briat 2004). La cartographie fine du QTL réalisée sur 1663 plantes pseudo-F2 a permis de réduire l'intervalle de confiance du QTL à 2.4 cM et de sélectionner six gènes candidats, alors que le QTL détecté en F6 présentait un intervalle de confiance de 7.5 cM.

L'intervalle de confiance fourni par la cartographie fine (2.4 cM) n'atteint pas la précision de l'analyse multi-croisements (0.9 cM). La population de cartographie fine s'apparente à une population F2 quant à sa puissance de détection et sa structure aux locus considérés. Ainsi, dans cette pseudo-population F2, environ la moitié des individus, les hétérozygotes au locus étudié, ne sont pas informatifs pour estimer l'effet additif. A effectif égal, la puissance des tests effectués à partir des descendances de lignées recombinantes est supérieure aux populations F2 (De Vienne 1998). Ainsi, l'effectif total d'environ 600 RILs lors de l'analyse multi-croisements apporte un gain de précision face aux 1 663 plantes F2 utilisées en cartographie fine. Il aurait pu être envisagé de réduire l'intervalle de confiance du QTL détecté en F7 (cartographie fine) en augmentant la taille de la population. En effet, la perte d'individus (1 663 au lieu d'environ 3 000 initialement prévus) a diminué la puissance de détection de la population F2. Le passage de 2 722 individus théoriques (pour avoir 95 % de chance d'observer un crossing-over dans moins de 100 000 pb, soit 0.11 cM) à 1 663 plantes a augmenté de 63 % la longueur étudiée, allant à l'encontre de la stratégie. En effet, à condition que le nombre de recombinants entre les deux nouvelles bornes du QTL (MTIC719 et 19|15a) soit suffisant, la résolution de la détection de QTL aurait permis de limiter encore le nombre de gènes candidats en éliminant par exemple PKS et FD qui sont situés près de 19|15a, c'est à dire à une extrémité du QTL.

L'effectif de la population de cartographie fine aurait pu être calculé pour observer un crossing-over dans 50 000 pb, ce qui aurait représenté l'étude de 5 400 plantes. Le coût financier élevé du génotypage et du phénotypage d'une telle population dans la perspective d'éliminer certains gènes candidats justifie difficilement le déploiement de tels moyens.

L'intervalle de confiance trouvé par l'analyse multi-croisements (0.9 cM) est inférieur à celui de la cartographie fine (2.4 cM). Le choix de l'intervalle détecté lors de la cartographie fine pour la recherche de gènes candidats se base sur :

- (i) le fait que les cartes génétiques des populations LR1, LR3 et LR4 ne sont pas assez robustes car réalisées sur un faible effectif de RILs (environ 200 par population). Il existe donc des erreurs de cartographie de certains marqueurs. Ainsi, malgré la probabilité d'observer un QTL commun à plusieurs populations et la finesse du résultat, il existe un risque d'une mauvaise estimation des contigs de BAC et donc des gènes candidats ;
- (ii) la robustesse de la carte détectée lors de la cartographie fine. Le nombre élevé d'individus permet une meilleure cartographie des marqueurs sur cette région du chromosome 7. Bien que la distance couverte (2.4 cM) par le QTL soit relativement grande, le nombre de gènes candidats pour la date de floraison est assez faible. Dans le cas présent, l'intervalle de confiance n'est donc pas un frein à la recherche du ou des gènes contrôlant le caractère. A l'inverse, (Yang et al. 2007) ont montré la présence de nombreux gènes candidats à la résistance à l'antracnose chez *M. truncatula* sur 200 000 pb, ce qui rend difficile l'analyse des différents gènes.

Notre analyse réalisée par la cartographie fine a permis de réduire nettement le nombre de gènes à étudier et d'envisager ainsi une étude transcriptomique de ces gènes par RT-PCR semi quantitative.

5.5 Le choix des gènes candidats

Parmi les six gènes candidats (MtPKS, MtCO, MtFTLa, MtFTLb, MtFTLc, MtFD), cinq sont décrits dans la bibliographie pour jouer un rôle majeur dans le contrôle de la date de floraison. En effet, les étapes activatrices de la floraison en relation avec la date de floraison ont été particulièrement étudiées chez *A. thaliana* et servent maintenant de modèle aux études menées sur la date de floraison chez d'autres espèces (Zeevaart 2006; Zeevaart 2007). Ainsi les gènes CONSTANS (CO) et FLOWERING LOCUS T (FT) jouent un rôle central dans l'induction de la floraison en jours longs chez *A. thaliana*. On peut envisager que leurs homologues MtCO, MtFTLa, MtFTLb et MtFTLc aient un rôle clé dans la date de floraison chez *M. truncatula*.

Le groupe I des séquences COL d'*A. thaliana* comprend quatre sous-groupes chez *M. truncatula* (répartis sur les chromosomes 1, 7, 8). Le gène MtCO, qui appartient au groupe Ia (MtCOLa), est connu pour son rôle de promoteur de l'initiation florale (Hayama and Coupland 2004). CO code pour une protéine nucléaire ayant un motif en doigt de zinc qui s'exprime de façon plus ou moins abondante selon la phase de jour (l'expression augmente) ou de nuit (l'expression diminue) et qui, en réponse aux jours longs, induit la transcription de *FT* dans le phloème des feuilles. En effet, ce gène est impliqué dans la perception de la photopériode selon le modèle des coïncidences externes (Imaizumi and Kay 2006) ; son expression, qui oscille journalièrement, est contrôlé par l'horloge circadienne. Au niveau protéique, l'activité de CONSTANS est régulée de manière à n'être efficace qu'à la lumière. Ainsi, la coïncidence entre un niveau d'expression élevé de CONSTANS et la lumière conduit à l'induction de gènes intégrateurs de la floraison. Or, l'étude transcriptomique réalisée ici a permis de mettre en évidence une expression significativement supérieure de MtCO chez Jemalong6 par rapport à DZA315.16 dans les feuilles avant floraison mais également lors de la cinétique d'expression sur 24 heures, ce qui peut être relié à l'induction plus précoce de la floraison chez Jemalong6.

Chez *A. thaliana*, FT agit dans l'apex en formant un complexe avec un facteur de transcription bZIP FD (Abe et al. 2005; Wigge et al. 2005). Bien que FT et FD soient produits dans différents sites, l'hétérodimère FT/FD formé agit dans l'apex et active les gènes APETALA1 (AP1) et SUPPRESSOR OF OVEREXPRESSION OF CO1 (SOC1), ce dernier active alors LEAFY (LFY) (Figure 35). Longtemps désavouée, l'hypothèse d'un stimulus commun à toutes plantes activant la floraison a été récemment reprise. Le candidat le plus évident de ce stimulus serait la protéine du gène FT (Corbesier et al. 2007). La mise en évidence physiologique de l'universalité du florigène s'est basée sur la différence de réponse aux jours courts ou longs selon les espèces. Récemment, il a été montré par transgénèse que le gène SINGLE FLOWER TRUSS (orthologue de FT chez la tomate) d'indifférence à la longueur du jour chez la tomate peut se substituer à la nécessité des jours longs chez *A. thaliana* (Lifschitz et al. 2006). Le gène FT exprimé sous conditions de jours courts chez *Cucurbita moschata* induit la floraison d'*A. thaliana* en conditions de jours longs (Lin et al. 2007). La fonction des orthologues de FT apparaît fortement conservée chez les plantes à fleurs aussi bien les monocotylédones que les dicotylédones (Zeevaart 2007). De plus, la protéine de FT en tant que florigène chez les plantes herbacées a été également démontrée chez les arbres angiospermes (Bohlenius et al. 2006; Hsu et al. 2006). Bien que le contrôle de l'expression de FT varie selon le type de réponse des plantes à la photopériode, la protéine apparaît toujours identique, elle est mobile des feuilles à l'apex et d'un poids approximatif de 20 kDA (Giavalisco et al. 2006). Chez *M. truncatula*, trois homologues du gène FT d'*A. thaliana* sont sur le même BAC (AC123593) du chromosome 7 (Hecht et al. 2005). Ces trois gènes FT sont reliés par synténie aux gènes GIGAS du pois (*P. sativum*) (chromosome 5). Les plantes mutantes *Gigas* ont une floraison tardive et ont une réponse plus forte à la photopériode et la vernalisation (Beveridge and Murfet 1996).

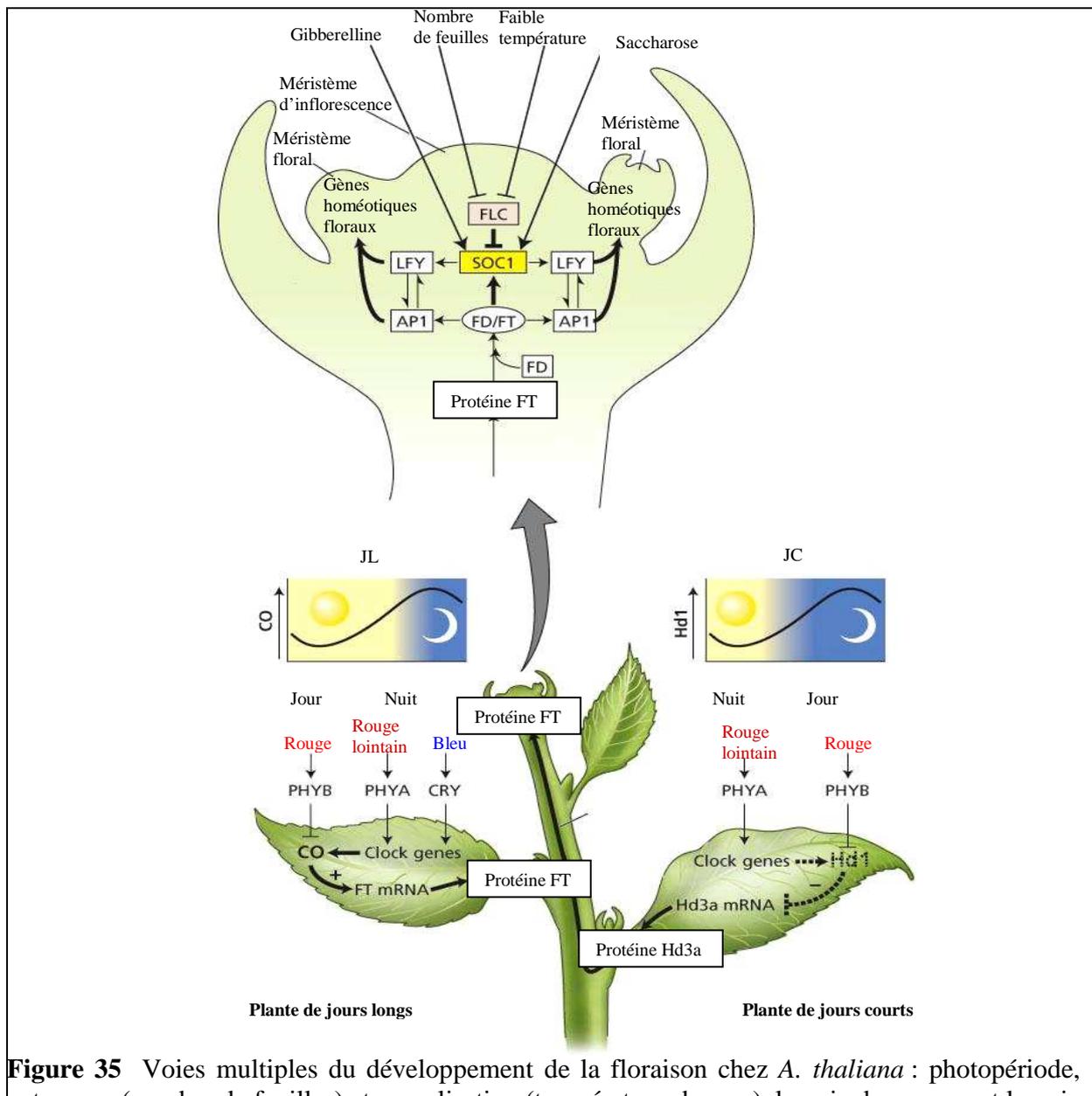


Figure 35 Voies multiples du développement de la floraison chez *A. thaliana* : photopériode, autonome (nombre de feuilles) et vernalisation (températures basses), la voie du sucrose et la voie des gibbérellines. Pour la photopériode, le parcours est situé dans les feuilles et comporte la transmission d'un stimulus floral, la protéine florale. Chez les plantes de jours longs comme *A. thaliana*, la protéine FT est produite dans le phloème en réponse à l'accumulation de la protéine CO sous des jours longs. Elle est ensuite transportée par le phloème jusqu'au méristème apical. Chez les plantes de jours courts comme le riz, le stimulus floral Hd3a est accumulé lorsque la protéine répresseur, Hd1, n'est pas produite en jours courts. Hd3a est transporté par le phloème jusqu'au méristème apical. Chez *A. thaliana*, FT se lie à FD et le complexe protéique FT / FD active les gènes AP1 et SOC1 qui provoquent l'expression du gène LFY. LFY et AP1 déclenchent ensuite des gènes homéotiques floraux. La voie autonome (nombre de feuilles) et la voie de la vernalisation (température basse) agissent sur le méristème apical pour réguler négativement FLC, un répresseur de SOC1. Le saccharose et la voie des gibbérellines, également localisée dans le méristème, active l'expression de SOC1. Tiré de (Zeevaart 2007).

5.6 Expression des gènes candidats

Six gènes décrits dans la bibliographie pour intervenir dans la date de floraison et détectés dans la région du QTL majeur ont été choisis pour la RT-PCR semi-quantitative. L'avantage majeur de l'utilisation de la RT-PCR semi quantitative a été la mise en évidence des différences d'expression entre Jemalong6 et DZA315.16 alors que la RT-PCR classique n'aurait permis de comparaisons qu'en terme de présence / absence d'amplification. Or, dans le cas présent, tous les gènes étudiés se sont avérés exprimés ou non-exprimés chez les deux lignées dans tous les échantillons étudiés, ce qui aurait conduit, en RT-PCR classique, à l'amplification de l'ensemble de ces gènes sans permettre de comparaison entre Jemalong6 et DZA315.16.

Les niveaux d'expression de MtFTLb, MtFD et MtPKS n'étant pas significativement différents entre Jemalong6 et DZA315.16 dans les échantillons de feuilles prélevés avant floraison ou à floraison, ces gènes ne permettent pas d'expliquer les différences de précocités observées entre les deux lignées étudiées.

L'étude transcriptomique des gènes candidats sélectionnés à l'intérieur de l'intervalle de confiance du QTL suggère que les différences d'expression du MtCO, dans les feuilles avant floraison et durant la cinétique d'expression sur 24 heures, expliquent, du moins en partie, les différences de précocité de Jemalong6 et DZA315.16.

Il aurait également pu être envisagé de réaliser cette étude transcriptomique par RT-PCR quantitative. Cependant, cette technique requiert des moyens importants (environ 9 000 euros pour les six gènes étudiés) et les résultats obtenus en PCR semi quantitative apparaissent satisfaisants pour un coût relativement modéré (environ 900 euros pour les six gènes étudiés).

Néanmoins, certains aspects du mode opératoire pourraient être améliorés. En effet, les échantillons ont été prélevés en chambre de culture à des dates précises, ce qui génère des différences de stades physiologiques entre Jemalong6 et DZA315.16. Ce type de regroupement ne tient pas compte du délai séparant le prélèvement de la date de floraison.

Ceci peut éventuellement constituer un élément expliquant la variabilité observée dans les échantillons étudiés. L'une des solutions pourrait être de prendre en compte les résultats des modèles de prédictions de la date de floraison afin d'échelonner les prélèvements d'échantillons sur une gamme similaire entre les lignées depuis le semis jusqu'à la floraison.

5.7 Vers une validation du gène CONSTANS

Parmi l'ensemble des gènes candidats étudiés par RT-PCR semi-quantitative, le gène MtCO s'avère différentiellement exprimé dans les feuilles entre Jemalong6 et DZA315.16. MtCO pourrait donc expliquer les différences de précocité observées entre ces deux lignées. D'un point de vue génomique, l'analyse de séquence a révélé un polymorphisme du nombre de répétitions dans la boîte TATA du promoteur, c'est-à-dire une insertion d'un motif TATA supplémentaire chez DZA315.16 et DZA45.5, et de deux motifs chez F83005.5 par rapport à Jemalong6. La difficulté de relier ce polymorphisme de séquence au phénotype réside particulièrement dans le manque de connaissances sur les actions des parties régulatrices des génomes, particulièrement chez les végétaux. Les mutations observées dans la boîte TATA pourraient expliquer les différences de date de floraison et les QTL trouvés dans les populations LR5 (Jemalong6 x F83005.5) et LR4 (Jemalong6 x DZA315.16). Cependant, elles ne permettent pas d'expliquer les QTL observés dans la population LR1 (DZA315.26 x DZA45.6) pour laquelle le nombre de motifs est identique entre les lignées parentales. Il pourrait donc s'agir d'un autre polymorphisme dans la zone promotrice. La réalisation du séquençage du promoteur dans une *core-collection* permettrait d'établir une série allélique et de relier la séquence génique au phénotype.

Le résultat de l'analyse transcriptomique coïncide avec la détection de QTL effectuée auparavant puisque le gène MtCO est cartographié sous le pic du QTL. De plus, le rôle prépondérant de gènes orthologues de CONSTANS dans le déterminisme de la date de floraison chez d'autres espèces a déjà pu être démontré. Lors d'une étude de cartographie comparée, la présence d'un orthologue du gène CONSTANS a été révélée dans l'intervalle de confiance d'un QTL pour la date de floraison chez *Brassica nigra* (Lagercrantz et al. 1996).

Chez l'avoine (*Avena sativa*), une étude de ségréants groupés effectuée sur trois populations de lignées recombinantes développées à partir de trois variétés brésiliennes d'avoine a permis de mettre en évidence un QTL majeur orthologue du locus HD1 chez le riz et de CONSTANS chez *A. thaliana* (Locatelli et al. 2006).

Différents types de validation fonctionnelle du gène candidat MtCO peuvent être envisagés :

- (i) la génétique d'association permet d'associer la variabilité phénotypique pour la date de floraison à la variabilité génétique. Cette approche est réalisée sur un grand nombre de plantes (*core-collection*) représentant le maximum de la variabilité pour ce caractère ;
- (ii) la génétique inverse par l'étude des banques de mutants par EMS TILLING ou *Tnt1* qui existent chez *M. truncatula* permettrait de rechercher des mutants de MtCO et d'observer leurs phénotypes. Cependant, les mutations observées par EMS TILLING et *Tnt1* s'insèrent dans plusieurs régions du génome. Il y a donc un travail important de croisements à réaliser pour obtenir un génotype ne portant que le gène CO muté ;
- (iii) la génétique sens qui permet à partir du phénotype altéré d'une plante, de remonter au gène muté, dans ce cas aussi, la multitude des mutations dans le génome limite l'observation d'un phénotype typique ;
- (iv) la transformation génétique (par *Agrobacterium tumefaciens*) convertirait une plante précoce en tardive et inversement chez *M. truncatula* (*i.e.* DZA315.16 précoce et Jemalong6 tardif). Il existe un protocole de transformation optimisé pour Jemalong6 mais pas pour DZA315.16 (Boisson-Dernier et al. 2001) ;
- (v) chez *A. thaliana*, il est possible de compléter des mutants KO (*Knock-out*) qui n'expriment plus le gène CO par les gènes MtCO de Jemalong6 et DZA315.16. Le phénotype de ces mutants complétés montrera s'ils retrouvent l'expression du gène muté ;
- (vi) l'inactivation du gène MtCO par RNAi : la séquence du siRNA est très spécifique de sa cible l'ARNm, ce qui permet la dégradation de l'ARNm ciblé qui n'est pas traduit en protéine. Le phénotype observé correspond à une inactivation de ce gène.

Bien que l'approche de génétique d'association ou l'étude de mutants semblent autant de possibilités vers la validation du gène MtCO, il n'en reste pas moins que la transformation de *M. truncatula* (iv) représenterait la validation ultime. L'accumulation d'arguments indépendants permet de renforcer le choix du gène isolé.

5.8 Quelle stratégie génétique pour la date de floraison ?

Il est à noter que les déterminants des variations naturelles de la date de floraison semblent très différents d'une espèce à l'autre malgré une conservation du signal florigène. En effet, chez *A. thaliana*, ce sont les gènes FRIGIDA (FRI) et FLOWERING LOCUS C (FLC) impliqués respectivement dans la voie de la vernalisation et dans la voie autonome qui jouent un rôle prépondérant (Salathia et al. 2006). L'action majeure d'orthologues de FT a également été mise au jour puisque, chez les Fabacées, la cartographie fine d'un QTL expliquant 73 % de la variation observée de la date de floraison chez le soja (*Glycine max*) a permis d'identifier le gène FT1 comme étant impliqué dans le contrôle de la date de floraison dans un croisement de cette espèce (Yamanaka et al. 2005). Chez *M. truncatula*, aucun gène orthologue de FRI ou de FLC n'a pu être identifié (Hecht et al. 2005). De plus, l'étude présentée ici indique que les différences observées pour la date de floraison entre les lignées Jemalong6 et DZA315.16 seraient imputables au gène MtCO, impliqué dans la voie de régulation de la date de floraison dépendante de la lumière. Il semble donc que l'évolution des génomes ait conduit différents gènes à contrôler spécifiquement la date de floraison (car les voies générales sont identiques), ce qui reflète l'adaptation des plantes et la diversité de leurs stratégies de développement. Chez *A. thaliana*, il a été montré récemment le rôle de micro-ARN dans la régulation de l'expression de CO (Jung et al. 2007). A l'heure actuelle, aucune étude ne révèle la quantité, la répartition et la position de ces micro-ARN. Il est donc possible d'en découvrir sur le chromosome 7, proche du gène MtCO, d'autant que plusieurs gènes trouvés sur cette portion du génome (2.4 cM) ont un rôle essentiel dans l'élaboration de la date de floraison.

L'étude de ces résultats sur le déterminisme génétique de la date de floraison et le rôle majeur que semble jouer le gène candidat MtCO mènent à envisager d'étudier le rôle de ses homologues sur la date de floraison chez d'autres Légumineuses cultivées. En effet, des travaux ont montré un fort degré de conservation macrosyntenique entre *M. truncatula* (marqueurs microsatellites dans un ordre similaire) et *M. sativa*, mais aussi *Pisum sativum*, *Lotus japonicus* et dans une moindre mesure *Glycine max* (Gualtieri and Bisseling 2002; Julier et al. 2003; Choi et al. 2004; Aubert et al. 2006).

Endre en 2002 a publié le premier exemple d'utilisation de la conservation de la synténie avec *M. truncatula* pour le clonage positionnel du gène NODULATION RECEPTOR KINASE (NORK) de la luzerne, orthologue de DMI2. Pour la floraison, qui est un processus clé dans le développement de la plante, il pourrait être envisagé l'étude du polymorphisme de MtCO chez plusieurs Légumineuses et des études d'association.

CONCLUSION

6 CONCLUSION

La floraison est une période critique au cours du développement des Angiospermes. L'adaptation à l'environnement et le succès reproductif au sein d'une saison sont dépendants de la date de floraison. En conséquence, l'identification et la caractérisation des gènes impliqués dans le contrôle de la date de floraison revêtent une importance majeure dans la compréhension de ce processus.

L'objectif de la thèse était d'identifier des gènes expliquant la variabilité génétique pour la date de floraison chez *M. truncatula*.

Dans un premier temps, et de façon à compléter les données existantes sur l'effet des paramètres du milieu sur la date de floraison, l'effet de la photopériode (12 h et 18 h) a été analysé pour huit lignées. Les données montrent qu'il existe de la variabilité génétique pour la date de floraison entre ces huit lignées, que la floraison est plus précoce en jours longs qu'en jours courts et qu'il existe une interaction lignée x photopériode.

Le développement de la plante jusqu'à la floraison a été modélisé selon un modèle additif prenant en compte les facteurs température et photopériode moyennes, avec des paramètres spécifiques de chaque lignée qui indiquent des réponses différentes de chaque lignée aux températures et photopériodes.

Les variations quantitatives de la date de floraison dans trois populations de lignées recombinantes impliquant quatre lignées parentales ont été évaluées et une détection de QTL a été menée. Un QTL majeur de date de floraison a été repéré sur le groupe de liaison 7 dans trois populations de lignées recombinantes expliquant de 10 à 60 % de la variabilité observée. Grâce à l'analyse multi-croisement, un QTL consensus a été mis en évidence ayant un intervalle de confiance de seulement 0.9 cM. La cartographie fine de ce QTL majeur, puis l'analyse bioinformatique des BAC inclus dans l'intervalle de confiance du QTL ont permis de sélectionner six gènes potentiellement impliqués dans le contrôle de la date de floraison.

L'étude de leur expression par RT-PCR semi-quantitative, sur des échantillons de feuilles et de fleurs prélevés à différents stades de développement chez deux lignées contrastées pour leur date de floraison, suggère qu'un des gènes candidats, MtCO, un homologue du gène CONSTANS d'*A. thaliana*, pourrait expliquer les différences de précocité observées chez les deux lignées. Chez *A. thaliana*, CONSTANS est décrit comme un des gènes clés impliqués dans la perception de la longueur du jour et dont l'expression est contrôlée par l'horloge circadienne. Son séquençage pleine longueur a révélé du polymorphisme entre les deux parents avec une TATA-box modifiée et une insertion intronique de 2 pb.

Une stratégie de validation fonctionnelle est à envisager mais l'étude de la diversité allélique du gène MtCO dans une core-collection comme celle en cours de réalisation à l'INRA de Montpellier fournirait également de réelles informations quant à l'adaptation des plantes à leur milieu d'origine et l'impact de MtCO.

Cette approche des gènes candidats positionnels met en évidence la cohérence des résultats QTL avec l'ensemble des gènes candidats présents sur le chromosome 7. Il semblerait donc que le gène contrôlant le caractère de la date de floraison soit MtCO et non un gène d'architecture.

Le fait que le QTL observé sur le chromosome 7 soit détecté sous toutes les photopériodes testées est cohérent avec un effet quantitatif et non qualitatif de la photopériode. Ainsi, la recherche de QTL de la date de floraison ne doit pas se restreindre à l'effet de la photopériode mais également de la vernalisation, qui joue également un rôle important sur la date de floraison (Delalande et al. 2004; Moreau et al. 2007). Il serait donc judicieux d'approfondir cette piste afin de détecter des QTL spécifiques à l'effet de la vernalisation sur la date de floraison. Ceci permettrait de trouver le(s) gène(s) contrôlant la réponse à la vernalisation chez *M. truncatula* et de mieux appréhender l'impact respectif de la vernalisation et de la photopériode.

Enfin, la génomique comparative des Légumineuses devrait permettre le transfert des connaissances de *M. truncatula* aux espèces cultivées, en particulier la luzerne et le pois. Ainsi, la cartographie du gène MtCO chez la luzerne ainsi que l'étude de sa diversité allélique dans une collection apporterait une indication sur le rôle de MtCO sur la date de floraison mais peut être aussi sur la morphogenèse aérienne.

RÉFÉRENCES BIBLIOGRAPHIQUES

- Abe M, Kobayashi Y, Yamamoto S, Daimon Y, Yamaguchi A, Ikeda Y, Ichinoki H, Notaguchi M, Goto K, Araki T (2005) FD, a bZIP protein mediating signals from the floral pathway integrator FT at the shoot apex. *Science* 309:1052-1056
- Aitken Y (1955) Flower initiation in pasture legumes. I. factors affecting flower initiation in *Trifolium subterraneum*. *Aust J Agric Res* 26:212-244
- Aubert G, Morin J, Jacquin F, Loridon K, Quillet MC, Petit A, Rameau C, Lejeune-Henaut I, Huguet T, Burstin J (2006) Functional mapping in pea, as an aid to the candidate gene selection and for investigating synteny with the model legume *Medicago truncatula*. *Theor Appl Genet* 112:1024-1041
- Aydi S, Drevon JJ, Abdelly C (2004) Effect of salinity on root-nodule conductance to the oxygen diffusion in the *Medicago truncatula*-*Sinorhizobium meliloti* symbiosis. *Plant Physiology and Biochemistry* 42:833-840
- Barker D, Bianchi S, Blondon F, Dattée Y, Duc G, Essad S, Flament P, Gallusci P, Génier G, Guy P, Muel X, Tourneur J, Denarie J, Huguet T (1990) *Medicago truncatula*, a model plant for studying the molecular genetics of the *Rhizobium*-legume symbiosis. *Plant Mol Biol Rep* 8:40-49
- Basten CJ, Weir BS, Zeng ZB (1994) Zmap-a QTL cartographer. Guelph, Ontario, Canada, pp 65-66
- Bateman RM, Hilton J, Rudall PJ (2006) Morphological and molecular phylogenetic context of the angiosperms: contrasting the 'top-down' and 'bottom-up' approaches used to infer the likely characteristics of the first flowers. *J Exp Bot* 57:3471-3503
- Belehradek J (1932) A propos de la base théorique des coefficients de température des processus protoplasmiques. *Protoplasma* 16:102-131
- Belknap JK, Atkins AL (2001) The replicability of QTLs for murine alcohol preference drinking behavior across eight independent studies. *Mammalian Genome* 12:893-899
- Bell CJ, Dixon RA, Farmer AD, Flores R, Inman J, Gonzales RA, Harrison MJ, Paiva NL, Scott AD, Weller JW, May GD (2001) The *Medicago* Genome initiative: a model legume database. *Nucleic Acids Res* 29:114-117
- Benlloch R, d'Erfurth I, Ferrandiz C, Cosson V, Beltran JP, Canas LA, Kondorosi A, Madueno F, Ratet P (2006) Isolation of mtpim proves Tnt1 a useful reverse genetics tool in *Medicago truncatula* and uncovers new aspects of AP1-like functions in legumes. *Plant Physiol* 142:972-983
- Benlloch R, Navarro C, Beltran JP, Canas LA (2003) Floral development of the model legume *Medicago truncatula*: ontogeny studies as a tool to better characterize homeotic mutations. *Sexual Plant Reproduction* 15:231-241

- Bernier G, Havelange A, Houssa C, Petitjean A, Lejeune P (1993) Physiological Signals That Induce Flowering. *Plant Cell* 5:1147-1155
- Bertero HD, King RW, Hall AJ (1999) Modelling photoperiod and temperature responses of flowering in quinoa (*Chenopodium quinoa* Willd.). *Field Crops Research* 63:19-34
- Beveridge CA, Murfet IC (1996) The gigas mutant in pea is deficient in the floral stimulus. *Physiol Plant* 96:637-645
- Blazquez MA (2000) Flower development pathways. *Journal of Cell Science* 113:3547-3548
- Blazquez MA (2005) The right time and place for making flowers. *Science* 309:1024-1025
- Blazquez MA, Green R, Nilsson O, Sussman MR, Weigel D (1998) Gibberellins promote flowering of *Arabidopsis* by activating the LEAFY promoter. *Plant Cell* 10:791-800
- Bohlenius H, Huang T, Charbonnel-Campaa L, Brunner AM, Jansson S, Strauss SH, Nilsson O (2006) CO/FT regulatory module controls timing of flowering and seasonal growth cessation in trees. *Science* 312:1040-1043
- 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. *Mol Plant Microbe Interact* 14:695-700
- Bonnin I, Prospero JM, Olivieri I (1996) Genetic markers and quantitative genetic variation in *Medicago truncatula* (Leguminosae): A comparative analysis of population structure. *Genetics* 143:1795-1805
- Bonnin I, Ronfort J, Wozniak F, Olivieri I (2001) Spatial effects and rare outcrossing events in *Medicago truncatula* (Fabaceae). *Mol Ecol* 10:1371-1383
- Borner R, Kampmann G, Gleissner R, Wisman E, Apel K, Melzer S (2000) A MADS domain gene involved in the transition to flowering in *Arabidopsis*. *The Plant Journal* 24:591-599
- Bouchez D (2004) L'espèce modèle *Arabidopsis thaliana*. In: Morot-Gaudry J, Briat JF (eds) *La génomique en biologie végétale*. pp 289-301
- Bourion V, Fouilloux G, Le Signor C, Lejeune-Henaut I (2002) Genetic studies of selection criteria for productive and stable peas. *Euphytica* 127:261-273
- Bowman JL, Smyth DR, Meyerowitz EM (1989) Genes directing flower development in *Arabidopsis*. *The Plant Cell* 1:37-52
- Bowman JL, Smyth DR, Meyerowitz EM (1991) Genetic interactions among floral homeotic genes of *Arabidopsis*. *Development* 112:1-20
- Brest DE, Hoshizaki T, Hamner KC (1971) Rhythmic leaf movements in Biloxi soybean and their relation to flowering. *Plant Physiol* 47:676-681

- Briggs WR, Beck CF, Cashmore AR, Christie JM, Hughes J, Jarillo JA, Kagawa T, Kanegae H, Liscum E, Nagatani A, Okada K, Salomon M, Rudiger W, Sakai T, Takano M, Wada M, Watson JC (2001) The phototropin family of photoreceptors. *Plant Cell* 13:993-997
- Broman KW, Wu H, Sen S, Churchill GA (2003) R/qtl: QTL mapping in experimental crosses. *Bioinformatics* 19:889-890
- Bucciarelli B, Hanan J, Palmquist D, Vance CP (2006) A standardized method for analysis of *Medicago truncatula* phenotypic development. *Plant Physiol* 142:207-219
- Buckler IV ES, Thornsberry JN (2002) Plant molecular diversity and applications to genomics. *Curr Opin Plant Biol* 5:107-111
- Bünning E (1964) *The physiological clock*. Heidelberg
- Catoira R, Galera C, de Billy F, Penmetsa RV, Journet EP, Maillet F, Rosenberg C, Cook D, Gough C, Denarie J (2000) Four genes of *Medicago truncatula* controlling components of a nod factor transduction pathway. *Plant Cell* 12:1647-1665
- Catoira R, Timmers ACJ, Maillet F, Galera C, Penmetsa RV, Cook D, Denarie J, Gough C (2001) The HCL gene of *Medicago truncatula* controls *Rhizobium*-induced root hair curling. *Development* 128:1507-1518
- Chabaud M, Lichtenzveig J, Ellwood S, Pfaff T, Journet EP (2006) Vernalization, crossings and testing for pollen variability. <http://www.noble.org/MedicagoHandbook/pdf/Vernalization.pdf> .
- Charcosset A, Gallais A (1996) Estimation of the contribution of quantitative trait loci (QTL) to the variance of quantitative trait by means of genetic markers. *TAG Theoretical and Applied Genetics* 93:1193-1201
- Chardon F, Virlon B, Moreau L, Falque M, Joets J, Decousset L, Murigneux A, Charcosset A (2004) Genetic architecture of flowering time in maize as inferred from quantitative trait loci meta-analysis and synteny conservation with the rice genome. *Genetics* 168:2169-2185
- Chelly J, Montarras D, Pinset C, Berwald-Netter Y, Kaplan JC, Kahn A (1990) Quantitative estimation of minor mRNAs by cDNA-polymerase chain reaction - Application to dystrophin mRNA in cultured myogenic and brain cells. *Eur J Biochem* 187:691-698
- Chetoui N (2005) Caractérisation du rôle de la protéine kinase MEK1 dans les voies de transduction des MAP Kinases. *Faculté de médecine; Université Laval; Québec*,
- Cheung W, Hubert N, Landry B (1993) A simple and rapid DNA microextraction method for plant, animal, and insect suitable for RAPD and other PCR analyses. *PCR Methods Appl* 3:69-70

- Choi HK, Mun JH, Kim DJ, Zhu H, Baek JM, Mudge J, Roe B, Ellis N, Doyle J, Kiss GB, Young ND, Cook DR (2004) Estimating genome conservation between crop and model legume species. *Proc Natl Acad Sci USA* 101:15289-15294
- Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. *Genetics* 138:963-971
- Clarkson NM, Russell JS (1975) Flowering responses to vernalization and photoperiod in annual medics (*Medicago ssp.*). *Aust J Agric Res* 26:831-838
- Colbert T, Till BJ, Tompa R, Reynolds S, Steine MN, Yeung AT, McCallum CM, Comai L, Henikoff S (2001) High-throughput screening for induced point mutations. *Plant Physiol* 126:480-484
- Cook DR (1999) *Medicago truncatula* - a model in the making! *Curr Opin Plant Biol* 2:301-304
- Corbesier L, Coupland G (2005) Photoperiodic flowering of Arabidopsis: integrating genetic and physiological approaches to characterization of the floral stimulus. *Plant Cell and Environment* 28:54-66
- Corbesier L, Vincent C, Jang SH, Fornara F, Fan QZ, Searle I, Giakountis A, Farrona S, Gissot L, Turnbull C, Coupland G (2007) FT protein movement contributes to long-distance signaling in floral induction of Arabidopsis. *Science* 316:1030-1033
- Covitz KMY, Amidon GL, Sadee W (1998) Membrane topology of the human dipeptide transporter, hPEPT1, determined by epitope insertions. *Biochemistry (Mosc)* 37:15214-15221
- Craufurd PQ, Mahalakshmi V, Bidinger FR, Mukuru SZ, Chantreau J, Omanga PA, Qi A, Roberts EH, Ellis RH, Summerfield RJ, Hammer GL (1999) Adaptation of sorghum: characterisation of genotypic flowering responses to temperature and photoperiod. *Theor Appl Genet* 99:900-911
- Crepon K (2005) Diversité des productions animales et des systèmes d'élevage en Europe. Impacts induits sur les stratégies d'approvisionnement en protéines, éléments de réflexion. *OCL - Oleagineux, Corps Gras, Lipides* 12:211-216
- Crespi MD, Jurkevitch E, Poiret M, Daubentoncarafa Y, Petrovics G, Kondorosi E, Kondorosi A (1994) Enod40, A Gene Expressed During Nodule Organogenesis, Codes for A Nontranslatable Rna Involved in Plant-Growth. *EMBO J* 13:5099-5112
- d'Erfurth I, Cosson V, Eschstruth A, Lucas H, Kondorosi A, Ratet P (2003) Efficient transposition of the Tnt1 tobacco retrotransposon in the model legume *Medicago truncatula*. *Plant J* 34:95-106
- De Vienne D (1998) Les marqueurs moléculaires en génétique et biotechnologies végétales. INRA, Paris

- Delalande M, Ronfort J, Prosperi JM (2004) Diversity for flowering time in a large collection of *Medicago truncatula* Gaertn. Québec, pp
- Delannoy D, Cathelin R, Jourjon MF, Sylvain J, Marcel J, Ngom B, Mangin B (2006) MCQTL Multi-allelic QTL mapping in multi-cross designs.
- Delph LF (1990) Sex-differential resource allocation patterns in the subdioecious shrub *HEBE SUBALPINA*. *Ecology* 71:1342-1351
- Delph LF (2003) Sexual dimorphism in gender plasticity and its consequences for breeding system evolution. *Evolution & Development* 5:34-39
- Dill A, Thomas SG, Hu JH, Steber CM, Sun TP (2004) The Arabidopsis F-box protein SLEEPY1 targets gibberellin signaling repressors for gibberellin-induced degradation. *Plant Cell* 16:1392-1405
- Djemel N, Guedon D, Lechevalier A, Salon C, Miquel M, Prosperi JM, Rochat C, Boutin JP (2005) Development and composition of the seeds of nine genotypes of the *Medicago truncatula* species complex. *Plant Physiology and Biochemistry* 43:557-566
- Domoney C, Duc G, Ellis THN, Ferrandiz C, Firnhaber C, Gallardo K, Hofer J, Kopka J, Kuster H, Madueno F, Munier-Jolain NG, Mayer K, Thompson R, Udvardi M, Salon C (2006) Genetic and genomic analysis of legume flowers and seeds. *Curr Opin Plant Biol* 9:133-141
- Doyle JJ, Luckow MA (2003) The rest of the iceberg. Legume diversity and evolution in a phylogenetic context. *Plant Physiol* 131:900-910
- Doyle M, Davis S, Bastow R, McWatters H, Kozma-Bognar L, Nagy F, Millar A, Amasino R (2002) The ELF4 gene controls circadian rhythms and flowering time in *Arabidopsis thaliana*. *Nature (London)* 419:74-77
- Dumoulin V, Ney B, Eteve G (1994) Variability of Seed and Plant Development in Pea. *Crop Sci* 34:992-998
- Dupuis J, Siegmund D (1999) Statistical methods for mapping quantitative trait loci from a dense set of markers. *Genetics* 151:373-386
- Durand R, Bonhomme R, Derieux M (2007) Seuil optimal des sommes de températures Application au maïs (*Zea mays* L.). *Agronomie* 2:589-597
- El Lithy ME, Clercx EJM, Ruys GJ, Koornneef M, Vreugdenhil D (2004) Quantitative trait locus analysis of growth-related traits in a new *Arabidopsis* recombinant. *Plant Physiol* 135:444-458
- El Yahyaoui F, Kuster H, Ben Amor B, Hohnjec N, Puhler A, Becker A, Gouzy J, Vernie T, Gough C, Niebel A, Godiard L, Gamas P (2004) Expression profiling in *Medicago truncatula* identifies more than 750 genes differentially expressed during nodulation,

- including many potential regulators of the symbiotic program(1[w]). *Plant Physiol* 136:3159-3176
- El-Assal SED,onso-Blanco C, Peeters AJM, Raz V, Koornneef M (2001) A QTL for flowering time in *Arabidopsis* reveals a novel allele of CRY2. *Nat Genet* 29:435-440
- Ellwood SR, D'Souza NK, Kamphuis LG, Burgess TI, Nair RM, Oliver RP (2006) SSR analysis of the *Medicago truncatula* SARDI core collection reveals substantial diversity and unusual genotype dispersal throughout the Mediterranean basin. *Theor Appl Genet* 112:977-983
- Endre G, Kereszt A, Kevei Z, Mihacea S, Kaló P, Kiss GB (2002) A receptor kinase gene regulating symbiotic nodule development. *Nature* 417:962-966
- Erskine W, Ellis RH, Summerfield RJ, Roberts EH, Hussain A (1990) Characterization of responses to temperature and photoperiod for time to flowering in a world lentil collection. *TAG Theoretical and Applied Genetics* 80:193-199
- Fankhauser C, Yeh KC, Lagarias JC, Zhang H, Elich TD, Chory J (1999) PKS1, a substrate phosphorylated by phytochrome that modulates light signaling in *Arabidopsis*. *Science* 284:1539-1541
- Feild TS, Arens NC (2007) The ecophysiology of early angiosperms. *Plant Cell and Environment* 30:291-309
- Firnhaber C, Puhler A, Kuster H (2005) EST sequencing and time course microarray hybridizations identify more than 700 *Medicago truncatula* genes with developmental expression regulation in flowers and pods. *Planta* 222:269-283
- Frugoli J, Harris J (2001) A brief tour of the cell cycle. *The Plant Cell* 13:458-463
- Garner WW (1933) Comparative responses of long-day and short-day plants to relative length of day and night. *Plant Physiol* 8:347-356
- Genier G, Guy P, Prosperi J (1992) Lucernes. Amelioration des especes vegetales cultivees: objectifs et criteres de selection
- Gianinazzi-Pearson V (1996) Plant cell responses to arbuscular mycorrhizal fungi : getting to the roots of the symbiosis. *The Plant Cell* 8:1871-1883
- Giavalisco P, Kapitzka K, Kolasa A, Buhtz A, Kehr J (2006) Towards the proteome of *Brassica napus* phloem sap. *Proteomics* 6:896-909
- Goffinet B, Gerber S (2000) Quantitative trait loci: a meta-analysis. *Genetics* 155:463-473
- Gualtieri G, Bisseling T (2002) Microsynteny between the *Medicago truncatula* SYM2-orthologous genomic region and another region located on the same chromosome arm. *Theor Appl Genet* 105:771-779

- Györgyey J, Vaubert D, Jiménez-Zurdo JI, Charon C, Troussard L, Kondorosi A, Kondorosi E (2000) Analysis of *Medicago truncatula* nodule expressed sequence tags. *Mol Plant Microbe Interact* 13:62-71
- Hammer GL, Vanderlip RL, Gibson G, Wade LJ, Henzell RG, Younger DR, Warren J, Dale AB (1989) Genotype-by-environment interaction in grain sorghum. II. Effects of temperature and Photoperiod on ontogeny. *Crop Sci* 29:376-384
- Hanocq E, Laperche A, Jaminon O, Laine AL, Le Gouis J (2007) Most significant genome regions involved in the control of earliness traits in bread wheat, as revealed by QTL meta-analysis. *Theor Appl Genet* 114:569-584
- Harper JL (1977) Population biology of plants. London;New York;San Francisco
- Harrison MJ (1997) The arbuscular mycorrhizal symbiosis: An underground association. *Trends in Plant Science* 2:54-60
- Hayama R, Coupland G (2004) The molecular basis of diversity in the photoperiodic flowering responses of *Arabidopsis* and rice. *Plant Physiol* 135:677-684
- Hecht V, Foucher F, Ferrandiz C, Macknight R, Navarro C, Morin J, Vardy ME, Ellis N, Beltran JP, Rameau C, Weller JL (2005) Conservation of *Arabidopsis* flowering genes in model legumes. *Plant Physiol* 137:1420-1434
- Heide OM (2005) Ecotypic variation among European arctic and alpine populations of *Oxydactylus*. *Arctic Antarctic and Alpine Research* 37:233-238
- Heide OM, Sonstebj A (2007) Interactions of temperature and photoperiod in the control of flowering of latitudinal and altitudinal populations of wild strawberry (*Fragaria vesca*). *Plant Physiol* 130:280-289
- Heyn CC (1963) The annual species of MEDICAGO. Jerusalem
- Hochman Z (1987) Quantifying vernalization and temperature promotion effects on time of flowering of three cultivars of *Medicago truncatula* Gaertn. *Aust J Agric Res* 38:279-286
- Hopkins WG (2003) Physiologie des stress et biotechnologie. *Physiologie végétale*. Bruxelles, pp 449-725
- Hsu CY, Liu YX, Luthe DS, Yuceer C (2006) Poplar FT2 shortens the juvenile phase and promotes seasonal flowering. *Plant Cell* 18:1846-1861
- Huang T, Bohlenius H, Eriksson S, Parcy F, Nilsson O (2005) The mRNA of the *Arabidopsis* gene FT moves from leaf to shoot apex and induces flowering. *Science* 309:1694-1696
- Huber H, Lukacs S, Watson MA (1999) Spatial structure of stoloniferous herbs: an interplay between structural blue-print, ontogeny and phenotypic plasticity. *Plant Ecology* 141:107-115

- Huguet T, Gherardi M, Chardon F, Sartorel E, Prosperi JM, Chennaoui-Kourda H, Aouani ME (2007) Creation of a consensus genetic-physical (CGPM) for the identification of *Medicago truncatula* genes involved in natural variation. Tunis, pp 12-
- Huguet T, Thoquet P, Gherardi M, Cardinet G, Prioul S, Lazrek F, Aouani ME, Laouar M, Abdelguerfi A, Kurchak O, Jacquet C, Torregrosa C, Julier B, Kiss E, Batut J, Prosperi JM (2004) 5th European Conference on grain legumes. 169-170
- Huyghe C (2005) Prairies et cultures fourragères en France. Entre logiques de production et enjeux territoriaux. Paris
- Imaizumi T, Kay SA (2006) Photoperiodic control of flowering: not only by coincidence. Trends in Plant Science 11:550-558
- Imin N, Nizamidin M, Daniher D, Nolan KE, Rose RJ, Rolfe BG (2005) Proteomic analysis of somatic embryogenesis in *Medicago truncatula*. Explant cultures grown under 6-benzylaminopurine and 1-naphthaleneacetic acid treatments. Plant Physiol 137:1250-1260
- Jenczewski E, Ronfort J, Chèvre AM (2003) Crop-to-wild gene flow, introgression and possible fitness effects of transgenes. Environmental Biosafety Resources 2:9-24
- Jeong S, Clark SE (2005) Photoperiod regulates flower meristem development in *Arabidopsis thaliana*. Genetics 907-915
- Johanson U, West J, Lister C, Michaels S, Amasino R, Dean C (2000) Molecular analysis of FRIGIDA, a major determinant of natural variation in *Arabidopsis* flowering time. Science 290:344-347
- Johnson IR, Thornley JHM (1985) Temperature dependence of plant and crop processes. Ann Bot 55:1-24
- Jongejans E, de Kroon H, Berendse F (2005) The interplay between shifts in biomass allocation and costs of reproduction in four grassland perennials under simulated successional change. Oecologia 147:369-378
- Jourjon MF, Jasson S, Marcel J, Ngom B, Mangin B (2005) MCQTL: multi-allelic QTL mapping in multi-cross design. Bioinformatics 21:128-130
- Julier B, Flajoulot S, Barre P, Cardinet G, Santoni S, Huguet T, Huyghe C (2003) Construction of two genetic linkage maps in cultivated tetraploid alfalfa (*Medicago sativa*) using microsatellite and AFLP markers. BMC Plant Biol 3:9-
- Julier B, Huguet T, Chardon F, Ayadi R, Pierre JB, Prosperi JM, Barre P, Huyghe C (2007) Identification of quantitative trait loci influencing aerial morphogenesis in the model legume *Medicago truncatula*. Theor Appl Genet 114:1391-1406

- Jung JH, Seo YH, Seo PJ, Reyes JL, Yun Y, Chua NH, Park CM (2007) The GIGANTEA-regulated microRNA172 mediates photoperiodic flowering independent of CONSTANS in *Arabidopsis*. *The Plant Cell* 19:2736-2748
- Khatkar MS, Thomson PC, Tammen I, Raadsma HW (2004) Quantitative trait loci mapping in dairy cattle: review and meta-analysis. *Genetics Selection Evolution* 36:163-190
- Kiss B, Csanadi G, Kalman K, Kalo P, Okresz L (1993) Construction of a basic genetic map for alfalfa using RFLP, RAPD, isozyme and morphological markers. *Mol Gen Genet* 238:129-137
- Kistner C, Parniske M (2002) Evolution of signal transduction in intracellular symbiosis. *Trends in Plant Science* 7:511-518
- Knapp E, Teuber L (1990) Environmental factors and plant phenotype affect alfalfa floret tripping. *Crop Sci* 30:270-275
- Komeda Y (2004) Genetic regulation of time to flower in *Arabidopsis thaliana*. *Annu Rev Plant Biol* 55:521-535
- Koornneef M, Alonso-Blanco C, Blankestijn-de Vries H, Hanhart CJ, Peeters AJM (1998) Genetic interactions among late-flowering mutants of *Arabidopsis*. *Genetics* 148:885-892
- Koornneef M, Alonso-Blanco C, Vreugdenhil D (2004) Naturally occurring genetic variation in *Arabidopsis thaliana*. *Annu Rev Plant Biol* 55:141-172
- Kramer EM, Hall JC (2005) Evolutionary dynamics of genes controlling floral development. *Curr Opin Plant Biol* 8:13-18
- Krüger GHJ (1984) Effect of photoperiod on initiation of staminate inflorescence in different hybrids of *Zea mays*. *South African Journal of Botany* 3:81-82
- Kulikova O, Gualtieri G, Geurts R, Kim DJ, Cook D, Huguet T, de Jong JH, Fransz PF, Bisseling T (2001) Integration of the FISH pachytene and genetic maps of *Medicago truncatula*. *Plant J* 27:49-58
- Lagercrantz U, Putterill J, Coupland G, Lydiat D (1996) Comparative mapping in *Arabidopsis* and *Brassica*, fine scale genome collinearity and congruence of genes controlling flowering time. *Plant J* 9:13-20
- Laibach F (1951) Über sommer- und winteranuelle Rassen von *Arabidopsis thaliana* (L.) Heynh. Ein Beitrag zur Atiologie der Blütenbildung. *Beiträge zur Biologie der Pflanzen*. pp 173-210
- Lander ES, Botstein D (1989) Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 121:185-199
- Lang A (1965) Physiology of flower initiation. *Encyclopedia of plant physiology* XV:1380-1536

- Le Corre V, Roux F, Reboud X (2002) DNA polymorphism at the FRIGIDA gene in *Arabidopsis thaliana*: extensive nonsynonymous variation is consistent with local selection for flowering time. *Mol Biol Evol* 19:1261-1271
- Lee H, Suh SS, Park E, Cho E, Ahn JH, Kim SG, Lee JS, Kwon YM, Lee I (2000) The AGAMOUS-LIKE 20 MADS domain protein integrates floral inductive pathways in *Arabidopsis*. *Genes & Development* 14:2366-2376
- Lesins KA, Lesins I (1979) *Genus Medicago (Leguminosae). A Taxogenetic study.* W. Junk, The Hague, Boston, London.
- Levy Y, Dean C (1998) The transition to flowering. *Plant Cell* 10:
- Lewin B (2000) *Genes VII.* Oxford
- Lifschitz E, Eshed Y (2006) Universal florigenic signals triggered by FT homologues regulate growth and flowering cycles in perennial day-neutral tomato. *J Exp Bot* 57:3405-3414
- Lifschitz E, Eviatar T, Rozman A, Shalit A, Goldshmidt A, Amsellem Z, Alvarez JP, Eshed Y (2006) The tomato FT ortholog triggers systemic signals that regulate growth and flowering and substitute for diverse environmental stimuli. *Proc Natl Acad Sci USA* 103:6398-6403
- Lin CT, Yang HY, Guo HW, Mockler T, Chen J, Cashmore AR (1998) Enhancement of blue-light sensitivity of *Arabidopsis* seedlings by a blue light receptor cryptochrome 2. *Proc Natl Acad Sci USA* 95:2686-2690
- Lin MK, Belanger H, Lee YJ, Varkonyi-Gasic E, Taoka KI, Miura E, Xoconostle-Cazares B, Gendler K, Jorgensene RA, Phinney B, Lough TJ, Lucas WJ (2007) FLOWERING LOCUS T protein may act as the long-distance florigenic signal in the cucurbits. *Plant Cell* 19:1488-1506
- Locatelli AB, Federizzi LC, Milach SCK, Wight CP, Molnar SJ, Chapados JT, Tinker NA (2006) Loci affecting flowering time in oat under short-day conditions. *Genome* 49:1528-1538
- Loudet O, Gaudon V, Trubuil A, Daniel-Vedele F (2005) Quantitative trait loci controlling root growth and architecture in *Arabidopsis thaliana* confirmed by heterogeneous inbred family. *Theor Appl Genet* 110:742-753
- Manthey K, Krajinski F, Hohnjec N, Firnhaber C, Puhler A, Perlick AM, Kuster H (2004) Transcriptome profiling in root nodules and arbuscular mycorrhiza identifies a collection of novel genes induced during *Medicago truncatula* root endosymbioses. *Mol Plant Microbe Interact* 17:1063-1077
- Marone M, Mozzetti S, De Ritis D, Pierelli L, Scambia G (2001) Semiquantitative RT-PCR analysis to assess the expression levels of multiple transcripts from the same sample. *Biol Proced Online* 3:19-25

- Michelmore RW, Paran I, Kesseli RV (1991) Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. *Proc Natl Acad Sci USA* 88:9828-9832
- Mockler TC, Guo HW, Yang HY, Duong H, Lin CT (1999) Antagonistic actions of Arabidopsis cryptochromes and phytochrome B in the regulation of floral induction. *Development* 126:2073-2082
- Mohan M, Nair S, Bhagwat A, Krishna TG, Yano M, Bhatia CR, Sasaki T (1997) Genome mapping, molecular markers and marker-assisted selection in crop plants. *Mol Breeding* 3:87-103
- Moreau D, Salon C, Munier-Jolain N (2006) Using a standard framework for the phenotypic analysis of *Medicago truncatula*: an effective method for characterizing the plant material used for functional genomics approaches. *Plant Cell Environ* 29:1087-1098
- Moreau D, Salon C, Munier-Jolain N (2007) A model-based framework for the phenotypic characterisation of the flowering of *Medicago truncatula*. *Plant Cell Environ* 30:213-224
- Morot-Gaudry J, Briat JF (2004) *La génomique en biologie végétale*. Paris
- Mouradov A, Cremer F, Coupland G (2002) Control of flowering time: Interacting pathways as a basis for diversity. *Plant Cell* 14:S111-S130
- Moussart A, Onfroy C, Lesne A, Esquibet M, Grenier E, Tivoli B (2007) Host status and reaction of *Medicago truncatula* accessions to infection by three major pathogens of pea (*Pisum sativum*) and alfalfa (*Medicago sativa*). *European Journal of Plant Pathology* 117:57-69
- Muranty H (1996) Power of tests for quantitative trait loci detection using full-sib families in different schemes. *Heredity* 76:156-165
- Nair RM, Hughes SJ, Ellwood S, Olivier R, Greene SL, Delalande M, Wen J, Oldroyd GE (2006) *Medicago truncatula* Stock centres. *Medicago truncatula Handbook*
- Nam YW, Penmetsa RV, Endre G, Uribe P, Kim D, Cook DR (1999) Construction of a bacterial artificial chromosome library of *Medicago truncatula* and identification of clones containing ethylene-response genes. *Theor Appl Genet* 98:638-646
- Nielson M, Lehman W, Kodet R (1976) Resistance in alfalfa to *Acyrtosiphon kondoi*. *Journal of Economic Entomology* 69:471-472
- O'Shea-Greenfield A, Smale ST (1992) Roles of TATA and initiator elements in determining the start site location and direction of RNA polymerase II transcription. *The journal of biological chemistry* 267:1391-1402
- Olivieri I (1991) Apports des marqueurs moléculaires à la prospection et à la gestion des ressources génétiques.

- Olsen KM, Halldorsdottir SS, Stinchcombe JR, Weinig C, Schmitt J, Purugganan MD (2004) Linkage disequilibrium mapping of Arabidopsis CRY2 flowering time alleles. *Genetics* 167:1361-1369
- Page T, Macknight R, Yang CH, Dean C (1999) Genetic interactions of the Arabidopsis flowering time gene FCA, with genes regulating floral initiation. *Plant J* 17:231-239
- Parinov S, Sundaresan V (2000) Functional genomics in Arabidopsis: large-scale insertional mutagenesis complements the genome sequencing project. *Curr Opin Biotechnol* 11:157-161
- Penmetsa RV, Cook DR (2000) Production and characterization of diverse development mutants of *Medicago truncatula*. *Plant Physiol* 123:1387-1397
- Pflieger S, Lefebvre V, Causse M (2001) The candidate gene approach in plant genetics: a review. *Mol Breeding* 7:275-291
- Prioul J, Pelleschi S, Sene M, Thevenot C, Causse M, Vienne D, Leonardi A (1999) From QTLs for enzyme activity to candidate genes in maize. *J Exp Bot* 50:1281-1288
- Prosperi JM, Guy P, Balfourier F (1995a) Ressources génétiques des plantes fourragères et à gazon. Paris
- Prosperi JM, Guy P, Grenier G, Angevain M (1995b) Les luzernes ou le genre *Medicago*. In: Prospero JM, Guy P, Balfourier F (eds) Ressources génétiques des plantes fourragères et à gazon. pp 131-140
- Purvis ON, Gregory FG (1937) Studies in vernalisation of cereals. *Ann Bot* 1:569-591
- Putterill J, Laurie R, Macknight R (2004) It's time to flower: the genetic control of flowering time. *Bioessays* 26:363-373
- Putterill J, Robson F, Lee K, Simon R, Coupland G (1995) The Constans gene of *Arabidopsis* promotes flowering and encodes a protein showing similarities to zinc-finger transcription factors. *Cell* 80:847-857
- Quiros CF, Bauchan GR (1988) The genus *Medicago* and the origin of the *Medicago sativa* complex. In: Hanson AA, Barnes DK, Hill RR (eds) Alfalfa and Alfalfa Improvement. pp 93-124
- Rebai A, Goffinet B (1993) Power of tests for QTL detection using replicated progenies derived from a diallel cross. *Theor Appl Genet* 86:1014-1022
- Roberts EH, Hadley P, Summerfield RJ (1985) Effects to temperature and Photoperiod flowering in chickpeas (*Cicer arietinum* L.). *Ann Bot* 55:881-892

- Roberts EH, Qi A, Ellis RH, Summerfield RJ, Lawn RJ, Shanmugasundaram S (1996) Use of field observations to characterise genotypic flowering responses to photoperiod and temperature: A soyabean exemplar. *Theor Appl Genet* 93:519-533
- Ronfort J, Bataillon T, Santoni S, Delalande M, David JL, Prospero JM (2006) Microsatellite diversity and broad scale geographic structure in a model legume: building a set of nested core collection for studying naturally occurring variation in *Medicago truncatula*. *BMC Plant Biol* 6:28-
- Salathia N, Davis SJ, Lynn JR, Michaels SD, Amasino RM, Millar AJ (2006) FLOWERING LOCUS C-dependent and -independent regulation of the circadian clock by the autonomous and vernalization pathways. *BMC Plant Biol* 6:
- Salisbury FB (1963) *The Flowering process*. New York
- Samach A, Coupland G (2000) Time measurement and the control of flowering in plants. *Bioessays* 22:38-47
- Samach A, Onouchi H, Gold SE, Ditta GS, Schwarz-Sommer Z, Yanofsky MF, Coupland G (2000) Distinct roles of CONSTANS target genes in reproductive development of *Arabidopsis*. *Science* 288:1613-1616
- Sato S, Nakamura Y, Asamizu E, Isobe S, Tabata S (2007) Genome sequencing and genome resources in model legumes. *Plant Physiol* 144:588-593
- Schena M, Shalon D, Davis RW, Brown PO (1995) Quantitative Monitoring of Gene-Expression Patterns with A Complementary-Dna Microarray. *Science* 270:467-470
- Schiex T, Gouzy J, Moisan A, de Oliveira Y (2003) FramedD: a flexible program for quality check and gene prediction in prokaryotic genomes and noisy matured eukaryotic sequences. *Nucleic Acids Res* 31:3738-3741
- Schomburg FM, Patton DA, Meinke DW, Amasino RM (2001) FPA, a gene involved in floral induction in *Arabidopsis*, encodes a protein containing RNA-recognition motifs. *Plant Cell* 13:1427-1436
- Searle I, Coupland G (2004) Induction of flowering by seasonal changes in photoperiod. *EMBO J* 23:1217-1222
- Sheldon CC, Rouse DT, Finnegan EJ, Peacock WJ, Dennis ES (2000) The molecular basis of vernalization: The central role of FLOWERING LOCUS C (FLC). *Proc Natl Acad Sci USA* 97:3753-3758
- Shitaka Y, Hirose T (1998) Effects of shift in flowering time on the reproductive output of *Xanthium canadense* in a seasonal environment. *Oecologia* 114:361-367
- Soltani A, Hammer GL, Torabi B, Robertson MJ, Zeinali E (2006) Modeling chickpea growth and development: Phenological development. *Field Crops Research* 99:1-13

- Sparks TH, Jeffree EP, Jeffree CE (2000) An examination of the relationship between flowering times and temperature at the national scale using long-term phenological records from the UK. *International Journal of Biometeorology* 44:82-87
- Stracke S, Kistner C, Yoshida S, Mulder L, Sato S, Kaneko T, Tabata S, Sandal N, Stougaard J, Szczyglowski K, Parniske M (2002) A plant receptor-like kinase required for both bacterial and fungal symbiosis. *Nature* 417:959-962
- Summerfield RJ, Muehlbauer F, Roberts EH (1985) *Lens culinaris* Medic. In: Havelly A (ed) A handbook of flowering. CRC Press, Boca Raton, Florida, pp 118-124
- Tardieu F (2003) Virtual plants: modelling as a tool for the genomics of tolerance to water deficit. *Trends in Plant Science* 8:9-14
- Tasma IM, Shoemaker RC (2003) Mapping flowering time gene homologs in soybean and their association with maturity (*E*) loci. *Crop Sci* 43:319-328
- Tjaden G, Edwards JW, Coruzzi GM (1995) Cis-Elements and Trans-Acting Factors Affecting Regulation of A Nonphotosynthetic Light-Regulated Gene for Chloroplast Glutamine-Synthetase. *Plant Physiol* 108: 1109-1117
- Théophraste traduit par Amigues S (1989) Recherche sur les plantes. Paris
- Thomas B, Vince-Prue D (1997) Photoperiodism in plants. Academic Press, London
- Thomson MJ, Edwards JD, Septiningsih EM, Harrington SE, McCouch SR (2006) Substitution mapping of *dth1.1*, a flowering-time quantitative trait locus (QTL) associated with transgressive variation in rice, reveals multiple sub-QTL. *Genetics* 172:2501-2514
- Thoquet P, Ghérardi M, Journet EP, Kereszt A, Ané JM, Prosperi JM, Huguet T (2002) The molecular genetic linkage map of the model legume *Medicago truncatula*: an essential tool for comparative legume genomics and the isolation of agronomically important genes. *BMC Plant Biol* 2:1-
- Thornsberry JM, Goodman MM, Doebley J, Kresovich S, Nielsen D, Buckler IV ES (2001) Dwarf8 polymorphisms associate with variation in flowering time. *Nat Genet* 28:286-289
- Tirichine L, de Billy F, Huguet T (2000) *Mtsym6*, a gene conditioning *Sinorhizobium* strain-specific nitrogen fixation in *Medicago truncatula*. *Plant Physiol* 123:845-851
- Torregrosa C, Cluzet S, Fournier J, Huguet T, Gamas P, Prosperi JM, Esquerre-Tugaye MT, Dumas B, Jacquet C (2004) Cytological, genetic, and molecular analysis to characterize compatible and incompatible interactions between *Medicago truncatula* and *Colletotrichum trifolii*. *Mol Plant Microbe Interact* 17:909-920
- Totland O (1999) Effects of temperature on performance and phenotypic selection on plant traits in alpine *Ranunculus acris*. *Oecologia* 120:242-251

- Udvardi MK (2001) Legume models strut their stuff. *Mol Plant Microbe Interact* 14:6-9
- van Heerden JM (1984) Influence of temperature and daylength on the phenological development of annual *Medicago* species with particular reference to *M.truncatula*, cv. Jemalong. *South African Journal of Plant Soil* 1:73-78
- VandenBosch KA, Stacey G (2003) Summaries of legume genomics projects from around the globe. *Community resources for crops and models. Plant Physiol* 131:840-865
- Vernoux T, Autran D, Traas J (2000) Developmental control of cell division patterns in the shoot apex. *Plant Mol Biol* 43:569-581
- Veyrieras JB, Goffinet B, Charcosset A (2007) MetaQTL: a package of new computational methods for the meta-analysis of QTL mapping experiments. *Bmc Bioinformatics* 8:
- Wang HL, Grusak MA (2005) Structure and development of *Medicago truncatula* pod wall and seed coat. *Ann Bot* 95:737-747
- Watson BS, Asirvatham VS, Wang LJ, Sumner LW (2003) Mapping the proteome of barrel medic (*Medicago truncatula*). *Plant Physiol* 131:1104-1123
- Wei F, Gobelman-Werner K, Morroll SM, Kurth J, Mao L, Wing R, Leister D, Schulze-Lefert P, Wise RP (1999) The Mla (powdery mildew) resistance cluster is associated with three NBS-LRR gene families and suppressed recombination within a 240-kb DNA interval on chromosome 5S (1HS) of barley. *Genetics* 153:1929-1948
- Wellensiek SJ (1963) Dividing cells as the prerequisite for vernalization. *Plant Physiol* 832-835
- Wellensiek SJ (1966) Photoperiod and temperature in the long-days plants. *Zeitschrift fur Pflanzenphysiologie* 54:377-385
- Weller JL, Nagatani A, Kendrick RE, Murfet IC, Reid JB (1995) New Iv Mutants of Pea Are Deficient in Phytochrome-B. *Plant Physiol* 108:525-532
- Wigge PA, Kim MC, Jaeger KE, Busch W, Schmid M, Lohmann JU, Weigel D (2005) Integration of spatial and temporal information during floral induction in *Arabidopsis*. *Science* 309:1056-1059
- Williams TA, Abberton MT (2004) Earlier flowering between 1962 and 2002 in agriculture varieties of white clover. *Oecologia* 138:122-126
- Wricke G, Weber WE (1986) *Quantitative genetics and selection in plant breeding*. Walter de Gruyter, Berlin
- Yakir E, Hilman D, Green RM (2006) Regulation of output from the plant circadian clock. *the FEBS Journal* 274:335-345

- Yamanaka N, Watanabe S, Toda K, Hayashi M, Fuchigami H, Takahashi R, Harada K (2005) Fine mapping of the FT1 locus for soybean flowering time using a residual heterozygous line derived from a recombinant inbred line. *Theor Appl Genet* 110:634-639
- Yang S, Gao M, Deshpande S, Lin S, Roe BA, Zhu H (2007) Genetic and physical localization of an anthracnose resistance gene in *Medicago truncatula*. TAG Theoretical and Applied Genetics in press:
- Yano M, Katayose Y, Ashikari M, Yamanouchi U, Monna L, Fuse T, Baba T, Yamamoto K, Umehara Y, Nagamura Y, Sasaki T (2000) Hd1, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the arabidopsis flowering time gene CONSTANS. *Plant Cell* 12:2473-2483
- Young ND, Cannon SB, Sato S, Kim D, Cook DR, Town CD, Roe BA, Tabata S (2005) Sequencing the genespaces of *Medicago truncatula* and *Lotus japonicus*. *Plant Physiol* 137:1174-1181
- Zeevaart JAD (1976) Physiology of flower formation. *Ann Rev Plant Physiol* 27:321-348
- Zeevaart JAD (2006) Florigen coming of age after 70 years. *Plant Cell* 18:1783-1789
- Zeevaart JAD (2007) FT protein, not mRNA, is the phloem-mobile signal for flowering. *Plant physiology online*

ANNEXES

Extraction rapide d'ADN en plaque 96 puits (méthode Cheung) (BM-MO-MAR-24)

Ce protocole est utilisé pour extraire rapidement de l'ADN mais il ne permet pas d'éliminer les protéines (Applications : génotypage, amplification par PCR).

Quelques points importants avant de commencer :

Préparation du matériel (travailler sur glace)

- ~40-60 mg de feuilles fraîches par puit.
 - 1 bille d'acier de 5 mm par puit (ray-grass, luzerne) ou 4 billes d'acier de 3 mm par puit (*Festuca eskia*).
 - Mettre un couvercle en caoutchouc, essuyer la plaque.
 - Congeler dans de l'azote liquide et conserver à -80 °C.
-
- ✓ Broyer 2 à 3 min à la mélangeuse à peinture.
 - ✓ Déposer à la pipette multicanaux distributrice 500 µl de tampon d'extraction (*ajouter extemporanément, le sarcosyl (Festuca eskia), le bisulfite de sodium et la RNase à 10 µg/ml final dans le tampon*). Thermosouder 6 sec un thermoseal (face brillante vers le haut, face mate vers la plaque). Agiter au vortex vitesse maximum.
 - ✓ Incuber 1h à 60°C et agiter toutes les 15 min par retournement.
 - ✓ Centrifuger 30 min à 4 000 rpm à 20°C.
 - ✓ Pendant la centrifugation, mettre 90 µl d'acétate d'ammonium 10M et 200 µl d'isopropanol dans une plaque 96 puits de 650 µl.
 - ✓ A la fin de la centrifugation, pipeter 200 µl de surnageant à la multicanaux, et le transférer dans la plaque contenant l'isopropanol et l'acétate. Thermosouder un easypeal. Mélanger délicatement par renversement (6 fois minimum).
 - ✓ Laisser 15 min minimum à température ambiante (ray-grass, luzerne) ou au congélateur à -20°C dans le cas de *Festuca eskia* (**Précipitation**)
 - ✓ Centrifuger 30 min à 4 000 rpm à 20°C.
 - ✓ Vider le surnageant par retournement de la plaque.
 - ✓ Bien égoutter sur papier absorbant.
 - ✓ Rincer les culots avec 100 µl d'éthanol 70% (**1^{er} rinçage**)
 - ✓ Centrifuger 10 min à 4 000 rpm à 4°C.
 - ✓ Vider l'éthanol par retournement de la plaque et égoutter sur papier absorbant.
 - ✓ Rincer les culots avec 100 µl d'éthanol 70% (**2^{ème} rinçage**)
 - ✓ Centrifuger 10 min à 4 000 rpm à 4°C.
 - ✓ Vider l'éthanol par retournement de la plaque et égoutter sur papier absorbant.
 - ✓ Sécher les culots 20 min maximum à 60°C à l'étuve.

- ✓ Sur glace, reprendre les culots dans 50 μ l de TE 10.1 pH8. Thermosouder un easypeal.
- ✓ Laisser une nuit à 4°C.

Tampon d'extraction :

	<u>pour 4 plaques</u>	<u>pour 2 plaques</u>
NaCl (2,2M final)	25.38 g	12,69 g
Tris (0,2M final)	4.85 g	2,425 g
EDTA (0,7M final)	5.2 g	2,6 g
	qsp 200 ml	qsp 100 ml

ajuster à pH 8 avec une solution d'HCl.

Autoclaver et stocker à 4°C.

Ajouter juste avant utilisation :

bisulfite de sodium	1 g	0,5 g
sarcosyl (1% final)	2 g	1 g
(chauffer au bain-marie pour dissoudre si nécessaire)		
RNase (10 μ g/ml final)	200 μ l	100 μ l

NB1 : N-Lauroylsarcosine sodium = Sarcosyl

Le sarcosyl est utilisé pour certaines extractions difficiles comme Festuca eskia. C'est un détergent tesioactif qui permet de casser les noyaux des cellules pour libérer l'ADN. Il n'est pas indispensable d'en rajouter pour l'extraction d'ADN de ray-grass ou de luzerne.

NB2 : Le bisulfite de sodium est un agent anti-oxydant.

Acétate d'ammonium 10 M

38.5 g d'acétate d'ammonium

qsp 50 ml

- Autoclaver et stocker à 4°C.

L'acétate d'ammonium se dissout totalement lors de l'autoclavage. A la sortie de l'autoclave, bien remuer le flacon pour homogénéiser et éviter la cristallisation.

TE 10-1

- 121,1 mg Tris
- 37,3 mg EDTA
- qsp 100 ml
- Ajuster à pH8 avec une solution d'HCl
- Autoclaver et stocker à 4°C.

ANNEXE 2

ANNEXE 3

Annexe

Jean-Baptiste PIERRE, 2008. Recherche de déterminants génétiques de la date de floraison chez la Légumineuse modèle, *Medicago truncatula*

RÉSUMÉ :

La morphogenèse aérienne inclut des caractères de croissance, de développement et de phénologie, et conditionne fortement la valeur d'usage des Légumineuses. Parmi ces caractères, la floraison est un événement majeur du cycle de vie car elle est déterminante pour le succès reproductif. Elle correspond à la transition généralement non réversible d'un méristème végétatif produisant des feuilles et tiges, en un méristème floral. La régulation de ce phénomène morphogénétique est le fait d'un réseau complexe de signalisations. Les légumineuses cultivées ont souvent des génomes complexes. C'est le cas de la luzerne (*Medicago sativa*), espèce fourragère pérenne, tétraploïde et allogame ainsi que du pois (*P. sativum*) qui présente un génome de grande taille. Des études précises peuvent être menées sur la légumineuse modèle *Medicago truncatula*, espèce diploïde, annuelle, à cycle court et autogame. De nombreuses ressources génétiques et génomiques sont disponibles chez cette espèce qui possède un fort degré de synténie avec la luzerne et le pois. De plus des gènes intervenant dans le déterminisme de la date de floraison ont été décrits chez *A. thaliana* et chez le pois.

L'objectif de la thèse est d'identifier des zones du génome et des gènes en utilisant les connaissances et outils développés chez *M. truncatula*, *A. thaliana* et *P. sativum* dans le déterminisme génétique de la date de floraison chez *M. truncatula*. Après une analyse de l'effet de la photopériode sur la date de floraison d'une gamme de lignées, une approche « gènes candidats positionnels » a été mise en oeuvre. La méthodologie employée consiste à montrer la variabilité génétique de la date de floraison en réponse à la photopériode, rechercher des QTL (*Quantitative Trait Locus*) de date de floraison dans trois populations connectées de lignées recombinantes, réaliser une méta-analyse QTL afin de détecter les régions conservées dans le contrôle du caractère entre populations, cartographier finement un QTL majeur et repérer les gènes candidats présents dans son intervalle de confiance. L'expression de ces gènes a été comparée pour deux lignées afin d'associer au caractère les gènes différenciellement exprimés.

En chambre de culture, la date de floraison de huit lignées a été mesurée sous deux traitements : 12 heures et 18 heures d'éclairement. Les données montrent qu'il existe de la variabilité génétique pour la date de floraison entre ces huit lignées, que la floraison est plus précoce en jours longs qu'en jours courts et qu'il existe une interaction lignée x photopériode. Un QTL majeur de date de floraison a pu être repéré sur le groupe de liaison 7 dans les trois populations de lignées recombinantes expliquant de 10 à 60 % de la variabilité observée. En méta-analyse sur les trois populations, un QTL consensus a été mis en évidence ayant un intervalle de confiance de seulement 0.9 cM. La cartographie fine de ce QTL a été réalisée sur la descendance (1663 plantes) d'une plante F6 hétérozygote au QTL détecté dans la population LR4. L'intervalle du QTL ainsi détecté couvre 2.4 cM. Six gènes homologues de gènes de floraison décrits chez *A. thaliana* ont été identifiés dans l'intervalle de ce QTL établi par cartographie fine. Leur séquençage pleine longueur a révélé du polymorphisme entre les deux parents : pour MtCO, homologue de CONSTANS et pour MtFTLc, homologue de FT. Par contre, aucun polymorphisme n'a été détecté pour deux autres homologues de FT (MtFTLa et MtFTLb) ni pour PKS. Une analyse de l'expression différentielle par RT-PCR semi quantitative des six gènes candidats a été réalisée chez deux lignées parentales contrastées pour leur date de floraison. Seul le gène MtCO est différenciellement exprimé entre ces deux lignées ; ce gène est donc actuellement le principal candidat pour expliquer la variation du caractère révélée à ce QTL sur le chromosome 7, dans ces populations.

Mots clés : *Medicago truncatula*, date de floraison, QTL, méta-analyse, cartographie fine, gènes candidats positionnels, étude transcriptomique, CONSTANS.

INRA, Unité de Génétique et d'Amélioration des Plantes Fourragères, 86600 Lusignan, France

ABSTRACT

The aerial morphogenesis includes traits such as growth, development or phenology, and has an important effect on yield and quality of legume crops. Among these traits, flowering is a major event of the life cycle and therefore crucial for the reproductive success. It corresponds to the irreversible transition of the meristem that generates plant leaves and stems into a floral meristem. The morphogenetic regulation of this phenomenon is a complex network of signals. Legume crops often have complex genomes such as the perennial forage species of alfalfa (*Medicago sativa*), that is tetraploid and allogamous and pea (*P. sativum*), with a large genome. Specific studies can be conducted on the model legume *Medicago truncatula* which is an annual autogamous and diploid species, and has a short reproductive cycle. Many genetic and genomic resources are available in this species that has a high degree of synteny with alfalfa and pea. In addition, genes involved in the determinism on the flowering date were described in *A. thaliana* and in pea.

The aim of the thesis was to identify genomic regions and genes in *M. truncatula* using the knowledge and tools developed in *M. truncatula*, *A. thaliana* and *P. sativum* for genetic determinism of the flowering date. After analysing the photoperiod effect on flowering date of a set of lines, a "positional candidate genes" strategy has been implemented. After the analysis of the genetic variability of the date of flowering in response to photoperiod, the methodology consisted into searching QTL (*Quantitative Trait Locus*) of flowering date in three connected populations of recombinant lines, conducting a meta-analysis QTL to detect common regions involved in the control of this trait between populations, realising a fine mapping of a major QTL and to identify major candidate genes in its confidence interval. The expression of these genes was compared between two parental lines to relate traits with genes that are differentially expressed.

The flowering date of eight lines was measured in growth chambers under two photoperiods: 12 hours and 18 hours of light. The data showed that there was genetic variability for the flowering date among the eight lines, that flowering was earlier under long days than under short days and that there was an interaction between line and photoperiod. On chromosome 7, a major QTL for flowering date was detected in three populations of recombinant lines explaining between 10 and 60% of the variation observed. A meta-analysis on the three populations revealed a consensus QTL with a confidence interval of only 0.9 cM. Fine mapping of this genome region was conducted in a pseudo-F₂ population (1663 plants) derived from a heterozygous F₆ plant at the QTL from the population LR4. Six genes homologous to flowering genes described in *A. thaliana* were identified in the confidence interval (2.4 cM) of the QTL detected in the region of fine mapping. Their full-length sequences revealed polymorphism between the two parents : for MtCO homologous to CONSTANS, and for MtFTLc homologous to FT. In contrast, there was no polymorphism observed for two other FT homologues (MtFTLa and MtFTLb) nor PKS. An analysis of the differential expression by semi-quantitative RT-PCR of the six candidate genes was performed in two parental lines contrasting for flowering date. Only MtCO gene was differentially expressed between these two lines. This gene is currently the best candidate to explain the variation of the trait detected at the QTL on chromosome 7 in these populations.....

Key-words : *Medicago truncatula*, flowering date, QTL, meta-analysis, fine mapping, positional candidate gene, transcriptomic study, CONSTANS.

INRA, Unité de Génétique et d'Amélioration des Plantes Fourragères, 86600 Lusignan, France

Jean-Baptiste PIERRE, 2008. Recherche de déterminants génétiques de la date de floraison chez la Légumineuse modèle, *Medicago truncatula*
