



Université
de Toulouse

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :

Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)

Présentée et soutenue par :
Audrey COURTIAL

Le mardi 30 octobre 2012

Titre :

Vers l'identification de gènes contrôlant la dégradabilité de la paroi secondaire lignifiée chez le maïs à travers l'élucidation de QTLs à effets forts

ED SEVAB : Développement des plantes

Unité de recherche :

UMR 5546 CNRS-UPS, Laboratoire de Recherche en Sciences Végétales (LRSV)

Directeur(s) de Thèse :

Yves BARRIÈRE, Directeur de recherche INRA, Lusignan
Jacqueline GRIMA-PETTENATI, Directrice de recherche CNRS, Toulouse

Rapporteurs :

David CAPARRÓS RUIZ, Directeur de recherche CSIC, Barcelone
Béatrice DENOYES-ROTHAN, Directrice de recherche INRA, Bordeaux

Autre(s) membre(s) du jury :

Fabienne GUILLON, Directrice de recherche INRA, Nantes
Clémentine VITTE, Chargée de recherche CNRS, Le Moulon
Philippe MOREAU, Directeur de recherche Caussade Semences, Caussade
Christophe JACQUET, Professeur des Universités Toulouse III, Toulouse

RESUME

La dégradabilité des parois des plantes fourragères est un facteur limitant à la fois leur valeur énergétique pour l'alimentation des ruminants et le rendement d'hydrolyse des polysaccharides pariétaux pour la production de biocarburants de seconde génération. L'identification des gènes contrôlant la formation des parois lignifiées chez les graminées est donc un enjeu majeur, tant au niveau fondamental en raison des caractéristiques spécifiques des parois des monocotylédones, qu'au niveau appliqué dans l'optique d'améliorer la dégradabilité de ces dernières. Les recherches conduites ont donc visé à identifier les gènes contrôlant des propriétés de ces parois lignocellulosiques chez le maïs en prenant comme modèle un cluster de QTLs à effets forts localisé dans le bin 6.06 des lignées recombinantes (RILs) de maïs précoces issues d'un croisement entre les lignées parentales F288 x F271.

De façon inattendue, nos résultats ont montré que les positions estimées de ces QTLs se situaient dans une région monomorphe entre les deux parents. Une densification ciblée de la carte génétique a permis de révéler que ces QTLs "fantômes" correspondaient en fait à des QTLs localisés sur deux positions proches, mais distinctes (bins 6.05 et 6.07). La densification de la carte a aussi permis de détecter des QTLs majeurs au bin 4.09. La détection de nouveaux QTLs lors de nouvelles expérimentations au champ a également permis de conforter l'implication des acides hydroxycinnamiques (acides *p*-coumarique, férulique et diféruliques) et de la composition monomériques des lignines dans la dégradabilité des parois.

Afin d'identifier les gènes sous-jacents à ces QTLs, plusieurs démarches ont été entreprises après avoir remise à jour la liste de gènes candidats putatifs à partir de la bibliographie. Cette liste de gènes comprend 287 gènes impliqués i) dans la biosynthèse des composés phénoliques des parois, ii) dans la régulation des gènes de biosynthèse, et iii) dans la mise en place des tissus lignifiés. Une étude d'expression différentielle entre le parent F271 porteur des allèles défavorables à la dégradabilité des parois et quatre RILs porteuses des allèles favorables aux QTLs majeurs du bin 6.06 a permis de mettre en évidence 360 gènes différentiellement exprimés, dont peu appartiennent à la liste des gènes candidats pré-établie. En parallèle, une approche de séquençage ciblé de BACs porteurs de la région d'intérêt a été initiée chez les deux lignées parentales F271 et F288, confirmant le très grand polymorphisme entre lignées de maïs. Enfin, une synthèse des données QTLs publiées dans six familles RILs, fondée sur une carte consensus construite à partir de la séquence physique de la lignée maïs de référence B73, a permis de mettre en évidence 35 QTLs consensus de dégradabilité des parois, avec 79 % des gènes de la liste pré-établie sous-jacents à ces QTLs.

Au vu de l'ensemble de ces résultats, plusieurs types de gènes semblent sous-tendre la variabilité génétique de la dégradabilité d'une paroi secondaire de maïs. De plus, plusieurs gènes pourraient simultanément expliquer la présence d'un QTL ou d'un cluster de QTLs. Ces déterminants génétiques putatifs sont souvent des gènes MYB et NAC, ainsi que des gènes COV ou zinc finger. Toutefois, l'implication de gènes de fonctions inconnues ou non encore reliées à la mise en place des parois est probable, avec aussi celle, possible, de miRNA. La poursuite du séquençage ciblé du bin 6.05 devrait permettre d'identifier le ou les gènes responsables de ces QTL majeurs chez la famille RIL F288 x F271. Les gènes candidats devront également être validés grâce à une analyse fonctionnelle.

Mots clés - *Zea mays*, dégradabilité, paroi, lignines, acides *p*-hydroxycinnamiques, QTL, gènes candidats, transcriptome, BAC, séquençage.

ABSTRACT

Highlighting the genetic determinants of the lignified cell wall assembly in grasses is a major challenge for both basic research, because of the specific parietal traits of these species, and for plant breeding based on marker-assisted selection. Cell wall degradability is a limiting factor of plant energy value for cattle feeding, as well as for the production of second-generation biofuel. The research conducted thus aims to identify genes involved in cell wall related traits, taking as a model a cluster of strong effect QTLs located in the bin 6.06 of the maize recombinant inbred line (RIL) progeny derived in the F288 x F271 cross.

Unexpectedly, it was shown that the estimated positions of these QTLs were located in an area that was monomorphic between the two parental lines. Targeted densification of the genetic map enabled us to show that these "ghost" QTLs correspond in fact to QTLs located on two close but distinct positions, in bins 6.05 and 6.07. In addition, map densification also allowed the identification of previously undetected major QTLs in bin 4.09. QTL detection from new field experiments have also allowed to consolidate the involvement of *p*-hydroxycinnamic acids (*p*-coumaric, ferulic and diferulic acids), and of the lignin monomeric composition, in the variation of cell wall degradability.

In order to identify the genes underlying these QTLs, several strategies were developed, after updating the putative candidate gene list. This list now includes 287 genes involved i) in the biosynthesis of cell wall phenolic compounds, ii) in the regulation of the former genes, and iii) in the lignified tissue patterning. A study of differential expression between the parent F271 exhibiting unfavorable alleles for cell wall degradability, and four RILs, with favorable alleles for the major QTLs in bin 6.06, highlighted 360 differentially expressed genes. However, only a few of them belong to the pre-established list of candidate genes. In addition, a sequencing approach targeted on BACs carrying the QTL region of interest, was initiated for both F271 and F288 parental lines, confirming the great polymorphism between these parental maize lines. Finally, a synthesis of QTL data available from six RIL progenies was performed based on a consensus physical map drawn from marker physical positions of the maize B73 reference genome. This analysis allowed to highlight 35 consensus QTLs for cell wall degradability, and 79 % of genes of the pre-established list underlie these QTLs.

As a tentative conclusion, several types of genes appeared to be involved in the genetic variation of the maize secondary cell wall degradability. Moreover, several genes could explain simultaneously the presence of a QTL or a QTL cluster. These putative genetic determinants are often transcription factors such as MYB and NAC, COV and zinc finger genes. However, the involvement of genes of unknown function or with a function not yet related to the cell wall metabolism, as well as the implication of miRNA can not be discarded. Further sequencing targeted to bin 6.05 should contribute to identify the gene or genes responsible for these major QTLs in the RIL progeny F288 x F271. Candidate genes should be latter validated through functional analyses.

Key words - *Zea mays*, degradability, cell wall, lignins, *p*-hydroxycinnamic acids, QTL, candidate genes, transcriptomic, BAC, 454 sequencing.

REMERCIEMENTS

Je souhaite, tout d'abord, remercier Yves et Jacqueline pour m'avoir fait confiance tout au long de ma thèse. Yves, sans vous cette thèse n'aurait pas pu être possible, merci d'être venu me chercher à la fin de mon master et d'avoir toujours été présent pour répondre à mes interrogations, même à distance. Un grand merci pour votre aide permanente, pour votre réactivité et pour tous vos allers-retours entre Lusignan et Toulouse. Jacqueline, merci, d'avoir délaissé un peu l'*Eucalyptus* pour du maïs. Merci pour l'attention que tu m'as accordée tout au long de ma thèse, pour ton aide précieuse et la rigueur dont tu as fait preuve lors de la rédaction du manuscrit.

Je remercie l'association Promaïs, qui a financé mes travaux de thèse, et leur président Claude Tabel pour m'avoir intégrée au sein du programme ZeaWall. Merci à Philippe Moreau de m'avoir si gentiment accueillie au sein de l'entreprise Caussade Semences et d'avoir fortement contribué à tous les aspects administratifs et financiers. Merci à tous les membres du projet ZeaWall qui sont venus échanger autour de mes résultats de thèse, au cours de rendez vous annuels, et qui ont fait avancer ce projet.

Je souhaite également remercier les directeurs successifs du Laboratoire de Recherche en Sciences Végétales, Guillaume Bécard d'abord, Elisabeth Jamet et Christophe Roux maintenant, pour m'avoir accueillie dans leurs locaux et m'avoir ainsi permis de réaliser cette thèse fort enrichissante, soutenue par l'ensemble des collègues et des compétences des différentes équipes. Merci à l'ensemble du personnel du laboratoire pour votre accueil et votre gentillesse, avec une pensée toute particulière pour les services communs dont notamment le noyau administratif Michèle, Catherine et Nicole, mais aussi les bioinformaticiens, Hélène et Bruno. Hélène, j'espère que je ne t'aurai pas trop dégoûté du maïs. Merci également à tous les non permanents du laboratoire avec qui j'ai eu le plaisir de partager de bons moments : Sophie, Elsa, Thomas, Coline, Marion, Alexandra, Olivier, Damien...

Je remercie toute mon équipe ex "Régulation Transcriptionnelle et Formation du Xylème" qui depuis un peu moins d'un an s'est agrandie en devenant l'équipe "Génomique fonctionnelle de l'*Eucalyptus*". Merci à vous tous pour votre gentillesse, votre bonne humeur et vos bons conseils. La liste est longue : Jacqueline, Hua, Isabelle, Nathalie, Annabelle, Fabien, Marçal, Anna, Hong, Bang, Sahar, Chantal et Christiane, sans oublier les membres qui n'ont été que de passage durant mes trois années de thèse : Edouardo, Najib, Victor, Malha, Nadia, Alan...

Merci aux personnes de l'INRA de Lusignan pour votre accueil très chaleureux. Même si mes visites ont été rares, ce fut un plaisir à chaque fois. Merci à Nathalie, Annelise, Corinne, Sabrina, Catherine pour votre sympathie et vos aides diverses (administratifs, commandes, bibliographie, analyse aux laboratoires...). Un merci tout particulier à l'équipe maïs, Christiane, Dominique et

Pascal, pour le temps qu'ils ont passé aux petits soins des maïs et sans qui les essais aux champs et les manipulations, qui en découlaient, n'auraient pas été possibles.

Merci également à tous les membres du CNRGV, qui fut mon second bureau durant une bonne partie de ma troisième année de thèse. Ce fut un grand plaisir de travailler avec vous, d'ailleurs au grand désespoir des gens de mon laboratoire qui ne me voyaient plus. Merci pour votre grande gentillesse et votre contribution à mes travaux.

Merci aux Versaillais, Valérie, Matthieu, Yu, Sylvain, pour tous les échanges que l'on aura eus sur le sujet de thèse. Valérie et Matthieu, un grand merci, également pour votre implication dans ce projet et votre contribution à la rédaction des différents projets de publication.

Merci également aux autres personnes qui sont intervenues à des moments clefs de la thèse : Brigitte Mangin et Jean-Marc Gion pour leur expertise dans les analyses QTLs, Marie-Françoise Devaux et Laurent Helary pour leur recommandation de la lame saphir, Yves Martinez et Alain Jauneau pour toute l'aide que vous m'aurez apportée en histologie, la plateforme génomique de Toulouse pour votre sympathie et votre accueil dans ma grande période génotypage.

Je tiens à remercier les membres du comité de thèse, qui ont évalué mon travail durant ces trois années de thèse. Merci donc à Valérie Méchin, Charles-Eric Durel, Richard Sibout, Philippe Moreau, pour tous vos conseils.

Un grand merci à Béatrice Denoyes-Rothan et à David Caparrós Ruiz d'avoir accepté d'être mes rapporteurs et ainsi d'évaluer mes travaux de thèse. Merci également à Fabienne Guillon, Clémentine Vitte, Philippe Moreau, Christophe Jacquet, pour leur participation au jury.

Ma reconnaissance va aussi aux différentes personnes que j'ai rencontrées pendant mes stages durant mes études. A savoir l'équipe fraise, qui m'a donnée le goût à la recherche et l'équipe du laboratoire de R2n (RAGT).

Enfin, merci à toute ma famille pour son soutien et sa présence tout au long de ma thèse. Un grand merci en particulier à mes parents, ma sœur, mon frère, qui ont toujours été là dans les bons et les mauvais moments. Et bien sur je n'oublie pas mon chéri, Florian qui a su me supporter et me soutenir durant ces trois années, ainsi que son père pour ses talents de bricoleur.

SOMMAIRE

INTRODUCTION GENERALE.....	1
1. Le maïs	1
1.1. Rappels généraux sur l'espèce et la plante "maïs"	1
1.1.1. Taxonomie.....	1
1.1.2. Morphologie.....	1
1.1.3. Cycle de développement	2
1.2. De l'anatomie d'une tige de maïs à la structure des parois	4
1.2.1. Anatomie	4
1.2.2. Organisation structurale et composition des parois cellulaires	5
1.3. Génomique structurale du maïs	6
1.3.1. Caryotype du maïs et origine	7
1.3.2. Taille du génome et complexité	8
1.4. Quelques rappels généraux sur l'importance du maïs.....	9
1.4.1. Expansion de la culture du maïs.....	9
1.4.2. Production	9
1.4.3. Utilisations du maïs et enjeux	11
1.4.3.1. Utilisations	11
1.4.3.2. Valorisation alimentaire et énergétique du grain.....	11
1.4.3.3. Valorisation alimentaire et énergétique du maïs plante entière.....	12
2. Vers l'amélioration de la valeur alimentaire et énergétique de la biomasse lignocellulosique chez le maïs.....	13
2.1. La valeur alimentaire et énergétique de la biomasse lignocellulosique	13
2.2. Amélioration et prédiction de la valeur énergétique et de la dégradabilité de la biomasse lignocellulosique	14
2.3. Les parois lignocellulosiques, facteurs limitant la dégradabilité.....	16
2.3.1. Facteurs biochimiques.....	16
2.3.2. Facteurs histologiques	18
2.4. Les gènes impliqués dans la formation des parois secondaires lignifiées	19
2.4.1. La voie de biosynthèse des lignines	20
2.4.1.1. La voie de biosynthèse des monolignols.....	20
2.4.1.2. Le transport et la polymérisation des monolignols en lignines	21
2.4.2. La voie de biosynthèse des acides hydroxycinnamiques	22
2.4.3. Les facteurs de transcription régulant les voies de biosynthèse et/ou la mise en place des tissus lignifiés.....	23

2.5. Régions génomiques impliquées dans la dégradabilité des parois et dans les caractères associés et gènes candidats.....	25
3. Travaux de thèse.....	27
<i>Chapitre I - Approche transcriptomique de la variabilité génétique entre quatre RILs de la descendance recombinante de maïs F288 x F271 ayant une dégradabilité de parois élevée et le parent faible F271</i>	29
1.1. Avant propos	29
1.2. Résultats	30
1.3. Conclusions et perspectives.....	32
1.4. Article 1 - Comparative expression of cell wall related genes in four maize RILs and one parental line of variable lignin content and cell wall degradability (2012, Maydica).....	33
<i>Chapitre II - QTLs fantômes et QTLs en positions multiples dans le bin 6.06 de la famille RIL F288 x F271. Approche par marquage ciblé de la descendance en s'appuyant sur le génome séquencé B73. Recherche de gènes candidats</i>	52
2.1. Avant propos	52
2.2. Résultats	53
2.3. Conclusions et perspectives.....	55
2.4. Article 2 - Targeted linkage map densification improved positions of maize cell wall related qtls, puzzle out ghost QTLs, and highlighted significant epistatic QTL interactions (submitted to TAG)	57
<i>Chapitre III - Recherche de QTLs de dégradabilité des parois, de teneur en lignines, de structure des lignines, et de teneur en acides p-hydroxycinnamiques dans la descendance recombinante F288 x F271 (Essais maïs en plantes sans épis réalisés en 2010 et 2011, avec récolte au stade ensilage)</i>	71
3.1. Avant propos	71
3.2. Résultats	71
3.3. Conclusions et perspectives.....	72
3.4. Article 3 - QTLs for cell wall degradability, lignin content and structure, and p-hydroxycinnamic acid contents in the F288 x F271 early maize RIL progeny (in progress)	74
<i>Chapitre IV - Séquençage ciblé de BACS situés dans le bin 6.05 sous des QTLs liés à la dégradabilité des parois détectés dans la famille recombinante F288 x F271</i>	89
4.1. Avant propos	89
4.2. Résultats	89
4.3. Conclusions et perspectives.....	90
4.4. Article 4 -Targeted BAC sequencing during search for candidate genes underlying a major cell wall related QTL in an early maize progeny (in progress).....	91

<i>Chapitre V - Recherche de gènes candidats putatifs impliqués dans la mise en place et la formation des parois lignifiées du maïs et colocalisations avec des QTLs de dégradabilité de parois issus d'une synthèse des données sur six familles de lignées recombinantes de maïs précoces et demi-précoces</i>	103
5.1. Avant propos	103
5.2. Résultats	103
5.3. Conclusions et perspectives.....	104
5.4. Article 5 - Search for cell wall degradability consensus QTLs based on a synthesis of data from six RIL progenies, and search for underlying candidate genes (in progress).....	105
<i>Chapitre VI - Approche de la variabilité histologique de la répartition des tissus lignifiés dans une tige de maïs</i>	128
6.1 Avant propos	128
6.2. Résumé	129
6.3. Introduction	130
6.4. Matériel et méthodes	132
6.5. Résultats	135
6.6. Discussion	137
6.7. Conclusions et perspectives.....	139
<i>Conclusion générale et perspectives</i>	142
<i>Références bibliographiques</i>	151

Abréviations et définitions

Termes généraux

AGPM.	Association Générale des Producteurs de Maïs
bm	brown-midrib (coloration brune de la nervure centrale des feuilles)
GNIS	Groupement National Interprofessionnel des Semences
ha	hectare
ma	millions d'années
mya	million years ago

Termes liés à la génétique et à la génomique

BAC	Bacterial Artificial Chromosome
Bin	région du génome d'environ 20 centiMorgans (cM) délimitée par deux core markers et désignée par le numéro du chromosome suivi par deux chiffres consécutifs désignant les régions successives sur le chromosome (01, 02, 03...)
bp	paires de bases
cM	centiMorgans (unité de recombinaison)
CIM	Composite Interval Mapping [recherche de QTLs en prenant des marqueurs en cofacteurs, par opposition au SIM (Simple Interval Mapping) qui teste l'effet de marqueurs pris isolément]
CNV	Copy Number Variation (nombre de copie d'un gène à une position donnée)
EST	Expressed Sequences Tag
F2	première génération d'autofécondation dans un hybride (F1) de deux lignées [chaque génération d'autofécondation est ainsi notée F2, F3, ..., les lignées sont considérées comme fixées à la génération F7 (l'habitude anglosaxonne est de nommer S0 la F1, puis S1 la F2, etc ..., S1 signifiant Self1 pour self-pollination 1)]
HMW DNA	High molecular weight DNA (ADN de haut poids moléculaire extrait lors des préparations des banques BAC)
INDEL	INsertion/DEletion (insertion ou délétions d'un nombre limité de paires de base)
LTR	Long Terminal Repeat
PAV	Presence / Absence Variation (absence d'un fragment de génome pouvant inclure des gènes)
PFGE	Pulsed-Field Gel-Electrophoresis (électrophorèse sur gel en champ pulsé utilisée pour la sélection des fragments d'ADN au cours de la création des banques BAC ainsi que pour l'évaluation de la taille des inserts d'ADN)
PP	Plate Pool (pool des 384 clones contenus dans une plaque de stockage 384 puits)
QTL	Quantitative Trait Loci (locus génomique associé à un caractère quantitatif)
eQTL	QTL d'expression (QTL après étude d'expression)
RIL	Recombinant Inbred Lines (famille de plein-frères dérivée d'un hybride simple par SSD sans sélection)

HRIL	RIL hautement recombinante (après une ou plusieurs génération d'intercroisement des F2 ou obtenues dans une descendance de deux RIL recombinantes)
NIL	Near Isogenic Lines [lignées obtenues par rétrocroisements (backcrosses) successifs permettant l'introgession d'un (ou d'un faible nombre de) petit fragment d'un génotype B dans un génotype receveur A (et donc de mesurer isolément l'effet de chaque région de B dans le génome de A)]
SSD	Single Seed Descent (production d'une famille où chaque plante F2 d'un croisement donne une descendance unique à chaque génération)
SNP	Single Nucleotide Polymorphism (changement au niveau d'une paire de base)

Termes liés à l'histologie

Ab	aire bleue
Ar	aire rouge
LCM	Laser Capture Microdissection
Ncs	nombre de couches cellulaires dans le sclérenchyme cortical

Caractéristiques et dosages parois et lignines

C-C	liaison carbone-carbone
NDF	Neutral Detergent Fiber (parois sans les pectines selon Goering et van Soest)
KL	lignines mesurées par la méthode Klason
ADF	Acid Detergent Fiber (lignines + cellulose selon Goering et van Soest)
ADL	Acid Detergent Lignin (lignines acido-résistante selon Goering et van Soest)
H	unité <i>p</i> -hydroxyphényle des lignines
G	unité gaïacyle des lignines
S	unité syringyle des lignines
pCA	acide <i>p</i> -coumarique
EsterpCA	acide <i>p</i> -coumarique estérifié aux unités S des lignines
FA	acide férulique
esterFA	acide férulique lié en ester aux arabinoxylanes
etherFA	acide férulique lié en éther aux unités G des lignines
diFA	acides diféruliques
5-5diFA	acides diféruliques liés en 5-5
8-O-4diFA	acides diféruliques liées en 8-O-4
NIRS	spectrométrie de réflectance dans le proche infrarouge

Caractéristiques liés aux échantillons de plantes et leur dégradabilité

MS	matière sèche
MO	matière organique

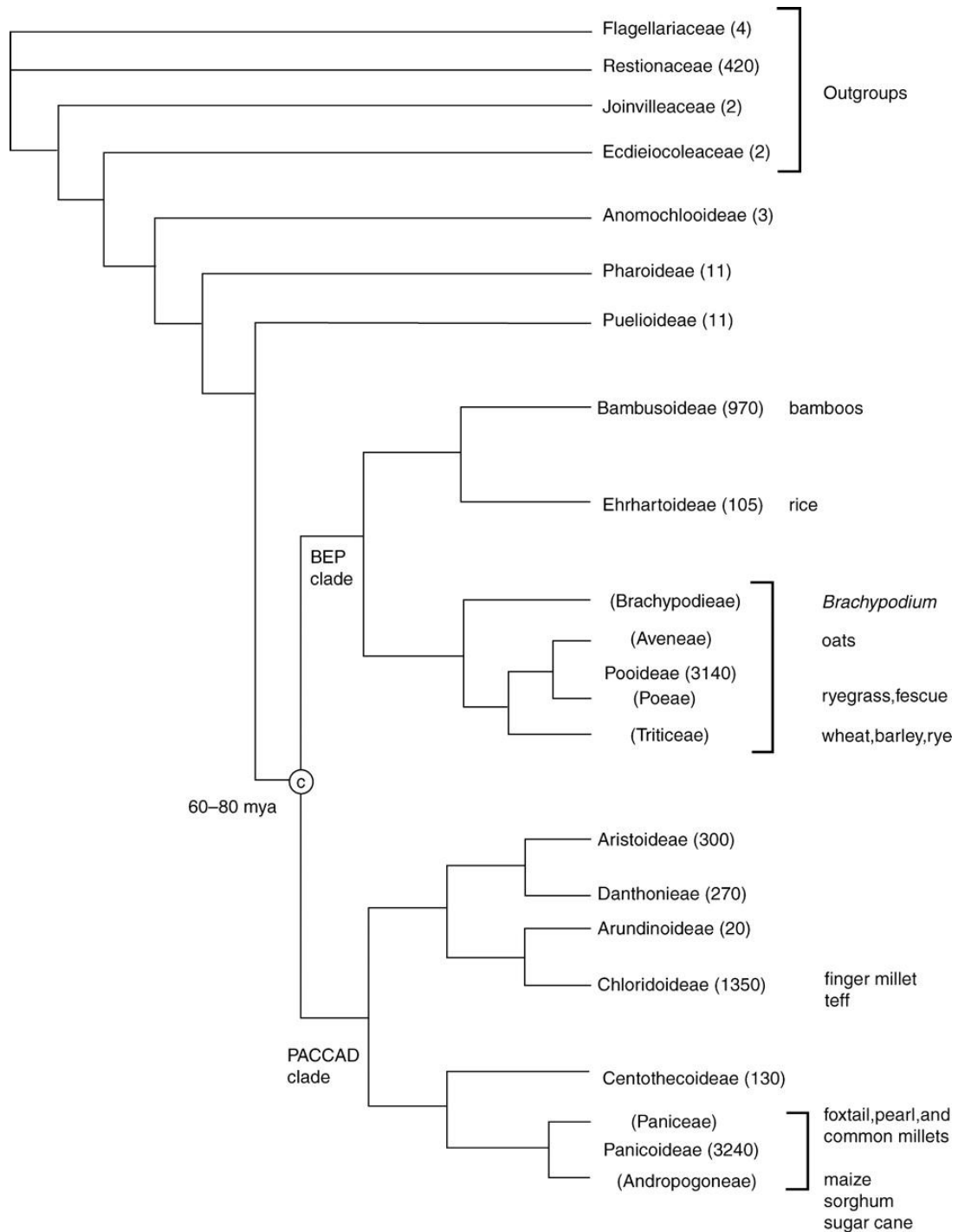
dMS	digestibilité de la matière sèche
dMO	digestibilité de la matière organique
IVNDFD	<i>In Vitro</i> Neutral Detergent Fiber Digestibility
DINAGZ	DIgestibilité <i>in vitro</i> du Non-(Amidon+Glucides solubles+matières aZotées)
PDI	Protéines Digestibles dans l'Intestin
UF	Unité Fourragère (équivalent à la valeur énergétique de 1kg d'orge)
UFL	Unité Fourragères Lait (besoins UF pour la production de lait)
UE	Unité d'Encombrement (estimation grossière de l'ingestibilité d'un fourrage)

Gènes et Enzymes

4CL	4-hydroxycinnamate:CoA ligase
ABC	ATP binding cassette
Act	acyltransférase
ALDH	aldéhyde déshydrogénase
β -Glu	β -glucosidases
C3H	<i>p</i> -coumarate 3-hydroxylase
C4H	cinnamate 4-hydrolase
CAD	cinnamyl alcool déshydrogénase
CCR	cinnamoyl-CoA réductase
CCoAOMT	caffeoyle-CoA O-méthyltransferase
COMT	caféique O-méthyltransferase
COV	COntinuous Vascular ring (mutant d' <i>Arabidopsis</i> ayant plus de régions lignifiées)
ERF	ethylene response factor
F5H	ferulate-5-hydroxylase
Feruloyl-AXX	feruloyl-arabinoxylanes
GST	glutathione S-transferases
HCT	hydroxycinnamoyl transférase
Lac	laccase
OMT	O-méthyltransferase
PAL	phénylalanine ammonia-lyase
Pox	peroxydase
REF1	reduced epidermal fluorescence1 (ALDH)
SAD	sinapyl alcool déshydrogénase
SHP	shatterproof (mutant parois d' <i>Arabidopsis</i> à siliques indéhiscentes)
TAL	tyrosine ammonia-lyase
UGT	uridine-diphosphate-glucosyltransferase
ZRP4-OMT	ZRP4-like O-méthyltransferase (ZRP4 = Zea root preferential)

INTRODUCTION GENERALE

Figure 1 - Arbre phylogénétique de la famille des graminées. "©" dénote l'embranchement des graminées divisé en 12 sous-familles, divisées parfois en tribus (noms entre parenthèses). Le groupe « outgroup » correspond à des familles monocotylédones proches des graminées. Le nombre entre parenthèses indique le nombre approximatif d'espèces connues dans la sous-famille considérée (Doust et al, 2007).



1. Le maïs

1.1. Rappels généraux sur l'espèce et la plante "maïs"

1.1.1. Taxonomie

Le maïs, *Zea mays*, est une plante herbacée annuelle appartenant à la classe des monocotylédones (ou Liliopsida). Cette classe, comprenant les plantes à fleurs dont l'embryon possède un seul cotylédon, constitue une évolution relativement récente du règne végétal (200 millions d'années ou ma selon l'acronyme français), comparée à celle des dicotylédones qui ont très longtemps été seules représentantes des plantes à fleurs (Wolfe et al, 1989; Herendeen et Crane, 1995). On relève parmi les plantes monocotylédones une vingtaine de familles, avec notamment les graminées, les orchidées, les palmiers.

Le maïs fait partie de la famille des graminées (ou poacées) qui est divisée en 12 sous-familles (Figure 1 ; Doust, 2007). La quasi-totalité des céréales cultivées dont le blé et le riz et les deux espèces de graminées modèle pour la génomique que sont *Brachypodium distachyon* (Garvin, 2007) et *Setaria italica* (Bennetzen et al, 2012; Zhang et al, 2012) font également partie de cette famille, qui est apparue il y a environ 65 ma (Jacobs et al, 1999; Kellogg, 2001), probablement peu de temps après l'extinction des dinosaures. Au sein des graminées qui sont essentiellement divisées en deux grands clades [BEP (Bambusoïdés/Ehrhartoïdés/Poïdés) et PACCAD (Panicoïdeae/Arundinoïdeae/Chloridoïdeae/Centothecoïdeae/Aristidoïdeae/Danthonioïdeae)] le maïs appartient à la tribu des Andropogoneae, dans la grande sous-famille des Panicoïdeae.

1.1.2. Morphologie

Le maïs est une plante de grande culture, dont la taille peut varier de 40 cm à cinq mètres, et se situe généralement entre un et trois mètres pour les variétés couramment cultivées. La tige est le plus souvent unique (Figure 2), le tallage (tiges secondaires qui émergent à la base de la plante) est rare chez les variétés modernes. Il a été contre-sélectionné au cours des générations de sélection massale, d'abord semble-t-il pour faciliter la récolte manuelle des épis. Il était en effet plus facile de récolter un gros épi sur une tige unique, plutôt que de nombreux petits épis répartis sur plusieurs tiges comme chez les maïs primitifs, et *a fortiori* chez l'ancêtre téosinte. L'amélioration moderne des plantes s'est poursuivie sur cet idéotype, en augmentant le nombre de rangs et le nombre de grains par rang de cet épi unique sur une tige unique. De gros diamètre, la tige du maïs est remplie d'une moelle sucrée et présente une succession de nœuds et d'entrenœuds. Au niveau de chaque nœud est insérée une feuille

Figure 2 - Morphologie du maïs.

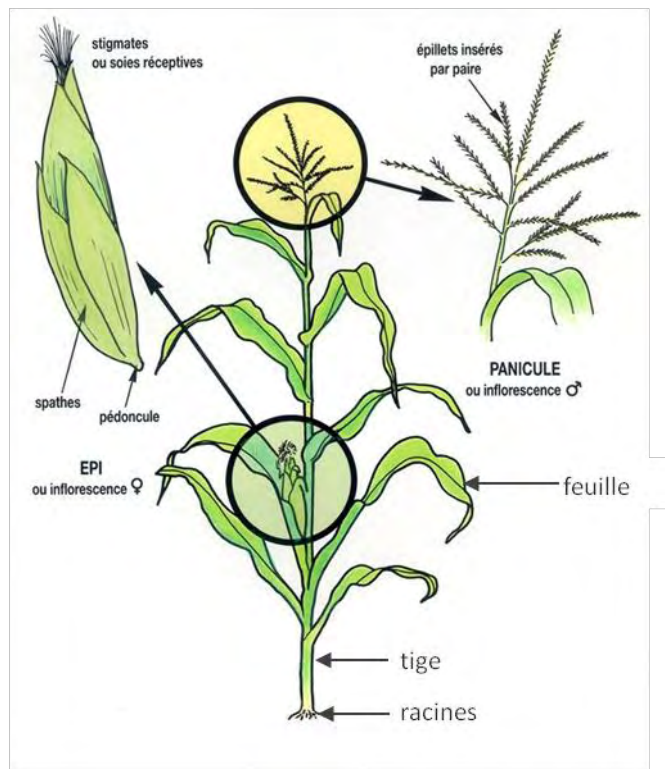
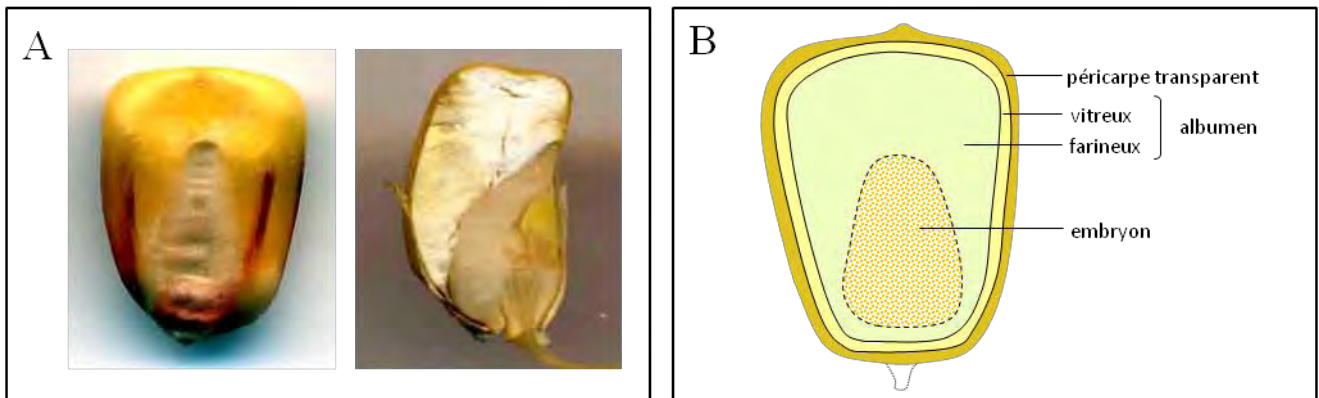


Figure 3 - Grain de maïs. A - Photos d'un grain de maïs vu de face et en coupe. B - Composition.



alternativement d'un côté et de l'autre de la tige. Les feuilles sont typiques des graminées avec gaines et ligules, des nervures parallèles, et de grande taille (jusqu'à 10 cm de large et 1 mètre de long).

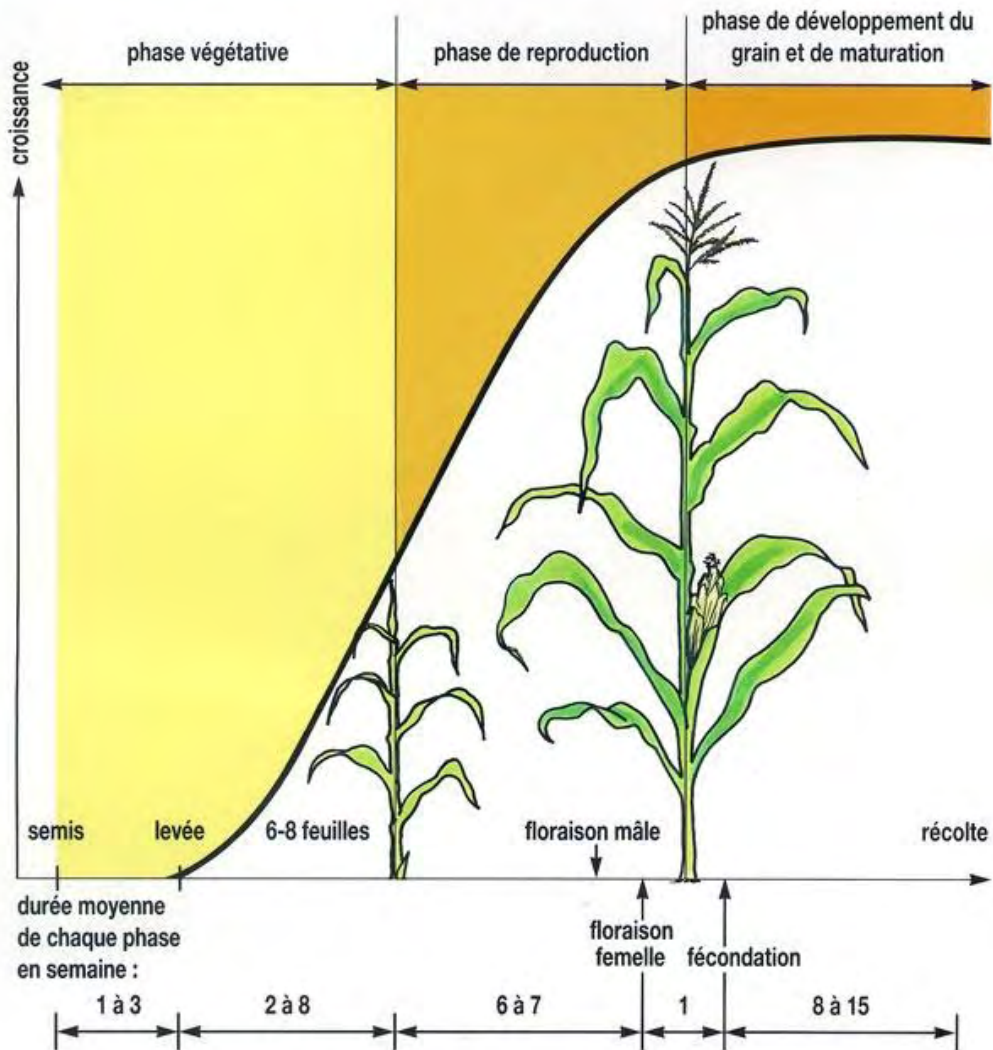
Le système racinaire comprend un très grand nombre de racines qui naissent sur les nœuds situés à la base de la tige. Ces racines peuvent atteindre une profondeur supérieure à un mètre, elles explorent ainsi un volume significatif de terre, mais une irrigation trop précoce, comme une sélection trop intense sur la productivité en biomasse aérienne, peut conduire à un développement racinaire limité et contribuer au phénomène de verse en végétation. Un développement d'abord horizontal des racines participe aussi de façon importante à la résistance à la verse en végétation.

Le maïs présente un appareil reproducteur à inflorescences séparées sur le même pied : c'est une plante monoïque. Les fleurs mâles sont regroupées en une inflorescence terminale appelée panicule, qui apparaît après la dernière feuille. Cette panicule plus ou moins ramifiée est constituée d'épillets regroupant chacun deux fleurs à trois étamines. Les fleurs femelles sont regroupées sur des inflorescences latérales appelées épis. Un pied peut donner naissance à deux (éventuellement trois ou quatre épis chez des génotypes prolifiques en épis, et/ou en cas de très faible densité), mais un seul atteint généralement un développement complet. L'axe de l'épi, appelé rafle, porte de 10 (lignées) à 35-40 (hybrides) rangs d'ovules, avec de 8 à 18-20 ovules par rang, en nombre pair, les fleurs étant comme sur la panicule émises par paire (une fleur sessile et une fleur pédicellée). Il est entouré de feuilles modifiées : les spathes, qui se dessèchent à maturité du grain. À l'extrémité supérieure et au moment de la floraison femelle, les spathes laissent dépasser les stigmates filiformes ou soies. Un épi de plante hybride peut porter, après fécondation, jusqu'à 500 grains à maturité. Le grain de maïs est formé d'un embryon et d'un tissu de réserve l'albumen (Figures 3A et 3B). Il est enveloppé par le péricarpe, fine membrane translucide ou parfois colorée. L'albumen est constitué essentiellement de grains d'amidon sauf la couche périphérique située sous le péricarpe qui contient des grains d'aleurone riches en protéines. Le grain de type amylicé est constitué essentiellement d'amidon (environ 70 %) mais aussi de protéines (10 %) de lipides (5 %) ainsi que de minéraux (calcium, phosphore) et de vitamines. Il peut être corné, denté, ou "pop corn", selon son état plus ou moins farineux ou vitreux et sucré, chez les maïs doux chez lesquels une étape de polymérisation de l'amidon est manquante (mutation sh2 de l'ADP glucose pyrophosphorylase, bin 3.09).

1.1.3. Cycle de développement

Le maïs, qui est une plante ayant une photosynthèse de type C4 Kranz, présente un cycle de développement (Figure 4) de longueur très variable selon la précocité de floraison des génotypes. Il y a des maïs de type "quarantain" faisant leur cycle en 40 jours en conditions chaudes, des maïs à cycle long, mais à production de biomasse limitée, en région fraîche (Bretagne, Normandie, Europe du Nord) avec alors souvent une appellation "précoce", et des maïs à cycle long également en région plus chaude, avec une forte production de biomasse, (Sud-Est, Sud-Ouest, Europe centrale, Corn Belt) avec

Figure 4 - Cycle de développement d'une plante de maïs (source Arvalis reprise par le GNIS, Groupement National Interprofessionnel des Semences).



Source : ARVALIS - Institut du Végétal

une appellation "tardif". Grâce à son système de photosynthèse spécifique, le maïs valorise très bien la lumière, la chaleur et l'eau, avec une production de 40 kg/MS.ha par mm d'eau évaporée (en comparaison de 30 kg/MS.ha.mm pour le blé). Le développement végétatif de la plante est ainsi spectaculaire, permettant la fabrication d'une quantité importante de matière sèche en peu de temps.

Dans les zones tempérées de l'hémisphère nord, le maïs possède pour la plupart des génotypes cultivés actuellement un cycle de vie d'environ six mois : semé en avril-mai, le maïs fleurit en juillet-août et ses grains sont récoltés en octobre-novembre. La plante entière peut aussi être récoltée et ensilée avant la maturité complète du grain, on parle alors de maïs fourrage. La distinction du maïs fourrage et du maïs grain est assez récente, le maïs étant autrefois utilisé en fourrage vert immature, récolté bien avant le remplissage du grain.

La phase végétative correspond à la formation de l'ensemble de la tige et des feuilles, et se termine lorsque 7 à 15 feuilles ont leur ligule visible (feuille étalée). La durée de cette phase dépend de la précocité de la plante et des conditions climatiques.

La phase de reproduction correspond à la formation et au développement des organes reproducteurs. La formation de la panicule et de l'épi a lieu entre mi-juin et mi- à fin juillet selon les conditions de milieu et met un terme à la production de nœuds foliaires. Un (éventuellement plusieurs) bourgeon axillaire se transforme en épi 1 mois avant la floraison, au stade 7-8 feuilles étalées, et les soies commencent à s'allonger. Le bourgeon terminal s'allonge dès la fin de la phase végétative pour se transformer en panicule, qui émerge du cornet de feuilles. La formation du pollen débute 2 à 3 semaines avant la floraison qui a lieu mi-juillet/mi-août. L'émission de pollen (floraison mâle) précède de 2 à 4 jours l'apparition des soies hors de l'épi et leur réceptivité (floraison femelle), favorisant la fécondation croisée. Le maïs est donc une plante allogame, qui peut être auto fertile avec une forte dépression de consanguinité. Dans une parcelle avec une diversité génétique, l'allo-pollen est plus compétitif que l'auto-pollen. La sélection a toutefois favorisé la protogynie, c'est-à-dire la sortie des stigmates avant la déhiscence des anthères. Le pollen mature tombe sur les soies, y germe et le noyau migre jusqu'à l'ovule pour le féconder.

Durant la phase de développement du grain et de maturation, les grains se gonflent d'abord d'eau, puis se remplissent en stockant sucres (amidon), protéines et matières grasses. Au cours de la perte d'eau et de l'accumulation des sucres, ils passent par trois stades : le stade laiteux (80 % d'eau et 20 % de matière sèche), le stade pâteux (50 % d'eau) et le stade vitreux (35 à 38 % d'eau). La plante arrive à maturité courant septembre à novembre selon le type de maïs (fourrage ou grain). Le maïs fourrage se récolte lorsque la teneur en matière sèche de la plante entière est comprise entre 30 et 35 %. Les plantes entières sont alors broyées et stockées à l'abri de l'air et de la lumière. Le maïs grain est, quant à lui, récolté à un taux d'humidité compris entre 15 % et 30 %. Le grain est ensuite séché pour atteindre 15 % d'humidité avant d'être conservé et commercialisé.

Figure 5 - Développement de la tige de maïs.

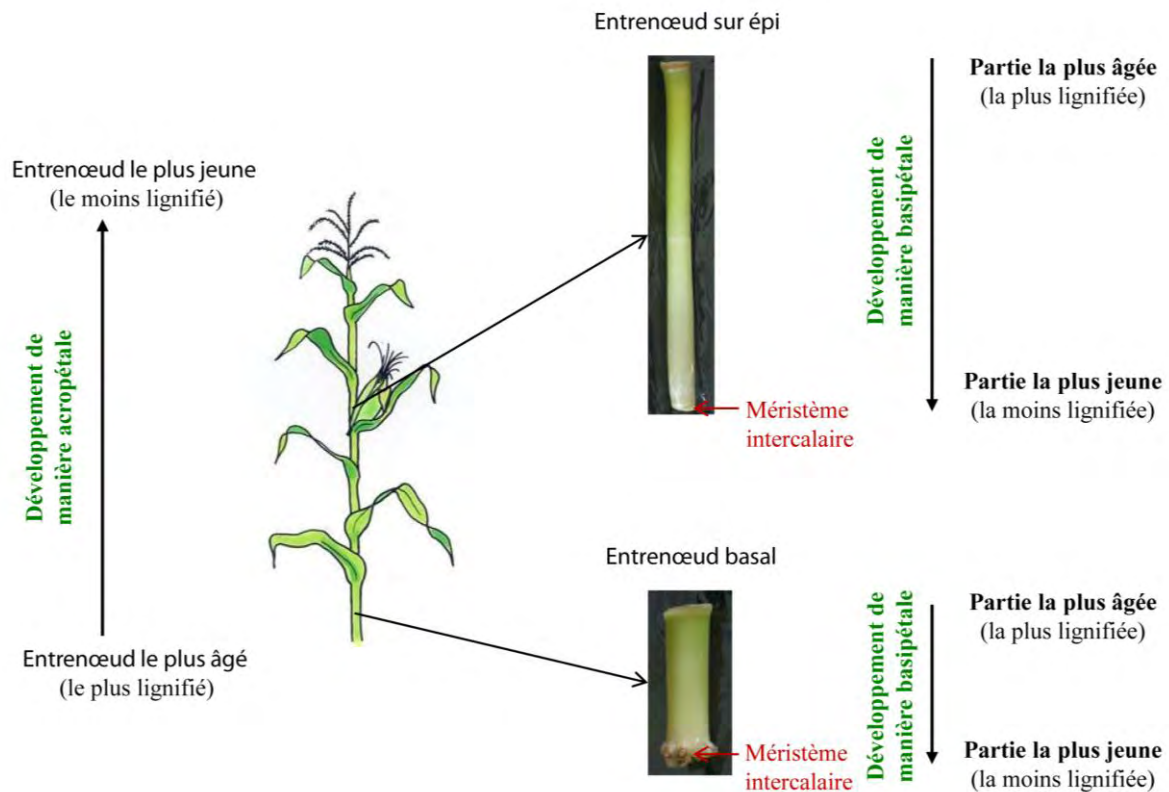
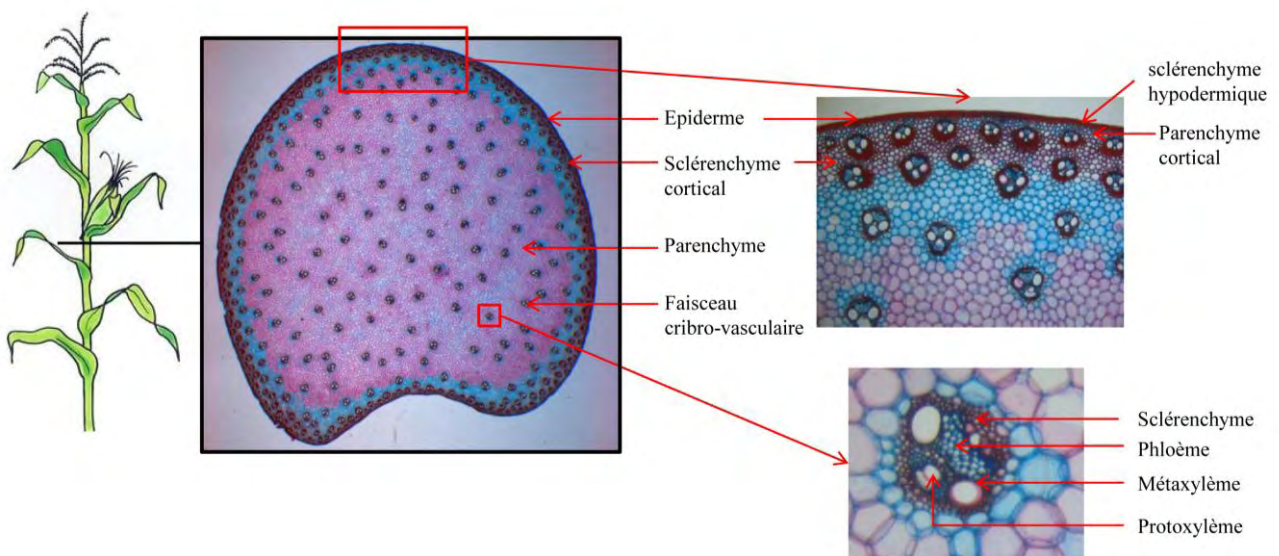


Figure 6 - Anatomie d'une tige de maïs. Coupe transversale du haut de l'entrenœud sous épi au stade panicule visible, observée après coloration Fasga (tissus rouges : tissus lignifiés, tissus bleus : tissus cellulósiques).



1.2. De l'anatomie d'une tige de maïs à la structure des parois

1.2.1. Anatomie

La tige de maïs présente différents tissus constitués de différentes cellules, à différents stades de différenciation selon leur âge mais également selon leur type. Cette différenciation s'accompagne de modification de la paroi pecto-cellulosique (paroi composée de pectines et de cellulose), dont les qualités sont déterminantes pour les fonctions assurées par la cellule au sein de son tissu. Les modifications majeures pouvant intervenir correspondent à l'épaississement de la paroi secondaire, à sa rigidification à la suite d'un dépôt de lignines, pouvant aboutir à la mort des cellules destinées à être des vaisseaux conducteurs.

Tout d'abord, la tige de maïs se développe de manière acropétale (Figure 5), l'entrenœud de base étant ainsi le plus ancien et l'entrenœud apical le plus récent (Morrison et al, 1994). A l'inverse, les entrenœuds se développent de manière basipétale, les méristèmes intercalaires, donnant naissance à de nouvelles cellules, étant situés à la base de chaque entrenœud. La partie la plus basale de l'entrenœud est ainsi la plus jeune et la moins différenciée et la partie la plus haute la plus âgée et la plus différenciée (Scobbie et al, 1993; Migné et al, 1996).

En plus des différences longitudinales dues à l'âge, plusieurs types de cellules, à paroi plus ou moins rigide et lignifiée, sont observés au sein d'une coupe transversale (Figure 6). L'organisation du système vasculaire d'une plante monocotylédone est très différente de celle des dicotylédones et sont dites "sans homologie" (Terashima et Fukushima, 1993; Tomlinson, 1995). Ainsi, le maïs, qui fait partie des monocotylédones, ne présente pas de cambium bifacial (méristème secondaire) et donc pas de croissance secondaire. Les différents tissus, présents dans la tige de maïs, sont décrits ci-après de l'extérieur vers l'intérieur de la tige. L'épiderme, qui assure un rôle de protection, est constitué d'une seule assise de cellules vivantes dont la paroi peut être épaissie vers l'extérieur par une cuticule. Le parenchyme cortical, peu présent au niveau de la tige, correspond à des cellules peu différenciées, toujours vivantes et à paroi mince et souple. Il s'agit de parenchyme chlorophyllien qui réalise la photosynthèse. Un anneau de sclérenchyme, formé de cellules mortes à paroi épaisse lignifiée et présentes sous forme de fibres, constitue un tissu épais à rôle de soutien. Ce tissu peut englober le premier cercle de faisceaux cribro-vasculaires (ou libéro-ligneux). Ces faisceaux, en forme de V et pénétrant radialement, sont dispersés, chez les monocotylédones, sur toute la surface, permettant une légère croissance en épaisseur. Chaque faisceau est constitué de plusieurs types de tissus, plus ou moins lignifiés : le xylème et le phloème. Du protoxylème, xylème peu différencié constitué de parenchyme de réserve et de trachéide, est présent à la pointe. Les trachéides sont formées de cellules mortes, allongées en tube, qui possèdent une paroi secondaire lignifiée de façon discontinue en spirale ou en anneaux. Du métaxylème, xylème formé après le protoxylème et plus différencié, est situé juste au dessus. Il est constitué de vaisseaux de type ponctué, avec une paroi entièrement lignifiée sauf au

Figure 7 - Paroi végétale.

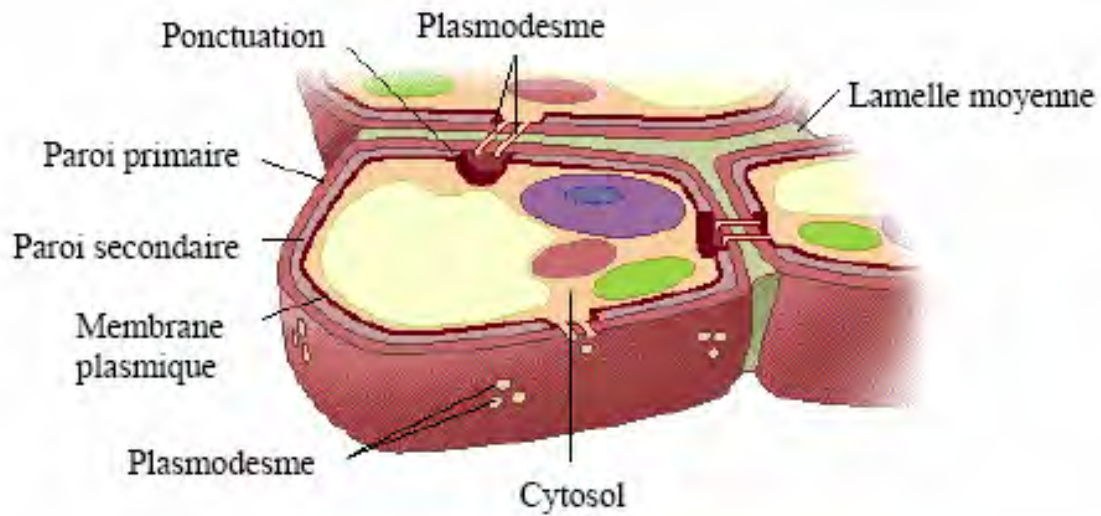


Tableau 1 - Comparaison des compositions pariétales des parois primaires et secondaires des monocotylédones et dicotylédones (tableau issu de Vogel, 2008 et corrigé d'après la source Ishii, 1997). MLG: Mixed linkage glucans, XyG : xyloglucan. Les différentes références sont ^b : Mitchell et al, 2007; ^c : Ishii , 2001; ^d : O'Neil et Work, 2003; ^e : Zablackis et al, 1995; ^f : Hatfield et al, 1999; ^g : Ebringerova et al, 2005.

	Primary wall		Secondary wall	
	Grass	Dicot	Grass	Dicot
Cellulose	20-30 ^{b,c}	15-30 ^{c,d,e}	35-45 ^{c,f}	45-50 ^c
Hemicelluloses				
Xylans	20-40 ^d	5 ^c	40-50 ^{c,g}	20-30 ^{c,g}
MLG	10-30 ^d	Absent	Minor	Absent
XyG	1-5 ^{c,d,g}	20-25 ^g	Minor	Minor
Mannans and glucomannans	Minor	5-10 ^d	Minor	3-5 ^g
Pectins	5 ^c	20-35 ^d	0.1 ^c	0.1 ^c
Structural proteins	1 ^d	10 ^{d,e}	Minor	Minor
Phenolics				
Ferulic acid and p-coumaric acid	0.5-1.5 ^c	Minor (except order Caryophyllales)	0.5-1.5 ^c	Minor (except order Caryophyllales)
Lignin	Minor	Minor	7-10 ^c	20 ^c
Silica			5-15 ^c	Variable

^a Numbers in this table were taken from several sources to provide rough approximations of generalized cell wall composition from typical dicots and

niveau de pores appelés ponctuation (Shane et al, 2000). Des fibres de sclérenchyme mais aussi du parenchyme, sont également présents dans le métaxylème. Trachéides et vaisseaux constituent les éléments conducteurs de la sève brute. Finalement, au dessus du xylème, est présent le phloème, qui comprend les éléments conduisant la sève élaborée. Ces éléments conducteurs sont formés par des cellules vivantes anuclées et à paroi non lignifiée. Ces cellules sont accompagnées de rares fibres, de parenchyme et de cellules compagnes, cellules à paroi mince qui maintiennent en vie les éléments conducteurs. Finalement, les faisceaux sont entourés d'une gaine de sclérenchyme et séparés par du parenchyme dit intrafasciculaire. Ce dernier tissu présente des parois qui peuvent se lignifier avec l'âge.

1.2.2. Organisation structurale et composition des parois cellulaires

Les parois cellulaires des végétaux sont dites primaires ou secondaires selon leurs compositions (Figure 7). Les parois de cellules contigües sont séparées par la lamelle moyenne qui se forme entre deux cellules voisines dès les premiers stades de croissance cellulaire. Elle permet la cohésion entre les cellules et les tissus grâce à son ciment pectique.

Des différences significatives existent entre les parois des plantes dicotylédones et celles des monocotylédones, elles sont récapitulées dans le Tableau 1 (Vogel, 2008). Le maïs est aujourd'hui la plante modèle pour l'étude de la paroi des monocotylédones (Lawrence et al, 2007), avec depuis peu *Brachypodium* (Garvin, 2007), et maintenant des sétaires (Bennetzen et al, 2012; Zhang et al, 2012) pour les plantes en C4.

La paroi primaire se dépose autour du protoplaste durant la période d'élongation cellulaire et est ainsi associée aux jeunes tissus en croissance et en cours de différenciation. Elle reste donc plastique et élastique pour ne pas limiter la croissance. De nature pecto-cellulosique, elle est constituée de cellulose, d'hémicelluloses, de pectines et de protéines pariétales. La charpente de la paroi primaire est surtout assurée par les microfibrilles de cellulose qui sont déposées lors de la division et de l'expansion cellulaire et qui forment un maillage lâche. La paroi primaire du maïs est une paroi de type II, paroi commune aux graminées et à la majorité des monocotylédones de la clade des commélinidées (The Angiosperm Phylogeny Group, 2009), alors que la paroi de type I est commune aux dicotylédones et à quelques monocotylédones (non commélinidées). Les différences de composition entre ces deux types de parois sont illustrées par Yokoyama et Nishitani en 2004 (Figure 8). La paroi de type II est constituée majoritairement de microfibrilles de cellulose associées aux hémicelluloses qui sont en plus faible quantité, alors que celle de type I est constituée d'autant de cellulose que d'hémicelluloses. Les hémicelluloses de la paroi de type II sont constituées principalement de glucurono-arabinoxylanes (Nishitani et Nevins, 1991; Carpita et Gibeaut, 1993) et de β -glucanes (Kato et al, 1982), alors que celle de la paroi de type I sont constituées principalement de xyloglucanes. La paroi de type II est également constituée de beaucoup moins de pectines

Figure 8 - Représentation schématique des parois primaires de type I et II (Yokoyama et Nishitani, 2004).

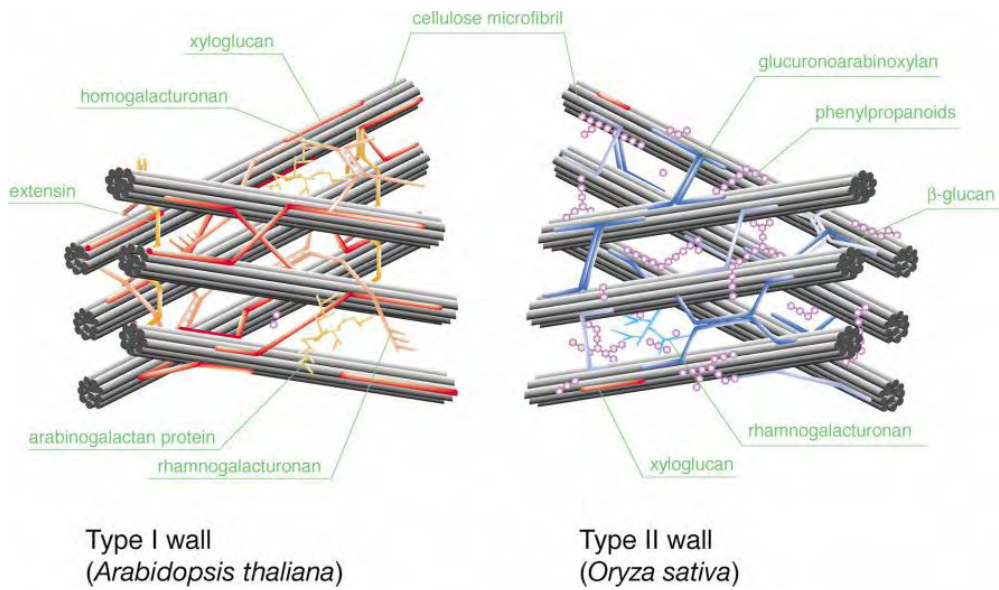
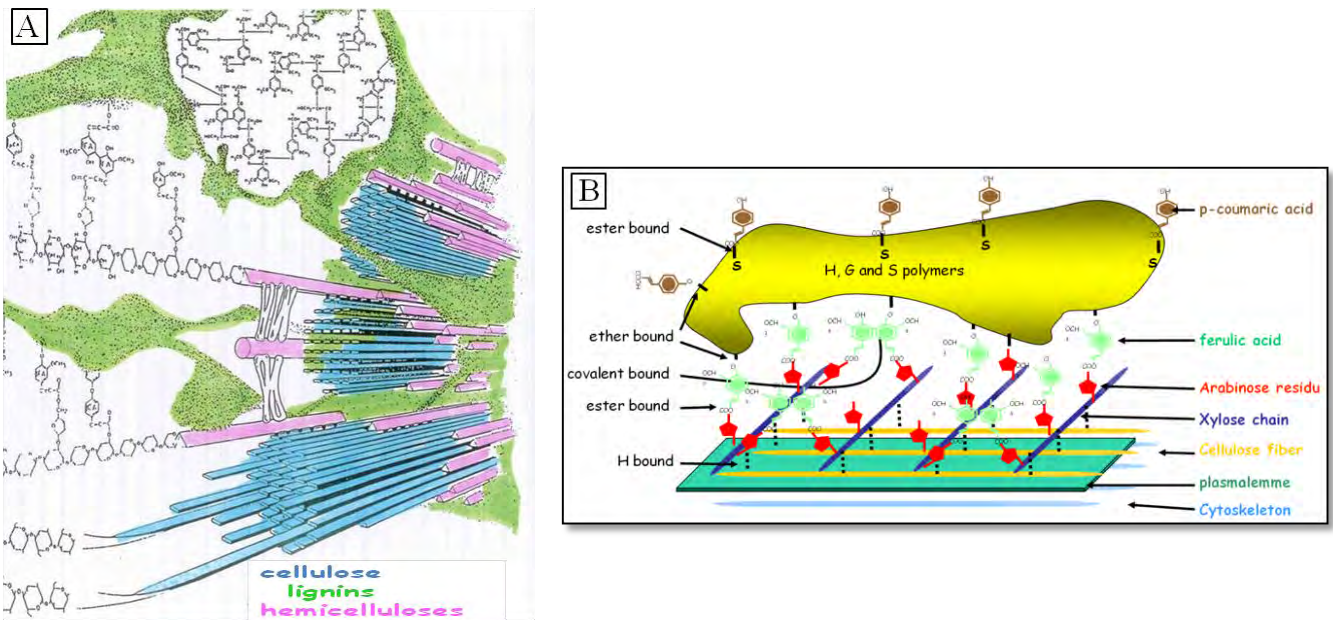


Figure 9 - Structure et constituants de la paroi secondaire du maïs. A - modèle structural (Bidlack et al, 1992). B - liaisons des acides hydroxycinnamiques aux différents constituants de la paroi.



(rhamnogalacturane) que celles de type I (rhamnogalacturane et homogalacturonane). Finalement, la paroi de type II présente des composés phénylpropanoïdes, acide férulique et dans une moindre mesure acide *p*-coumarique (Iiyama et al, 1990), que ne présente pas la paroi de type I.

La paroi secondaire est liée à la phase de développement terminal durant laquelle la cellule perd son élasticité et atteint sa taille limite. Cette paroi se met en place du côté interne de la paroi primaire. La paroi secondaire du maïs (Figures 9A et 9B) est constituée essentiellement de polysaccharides : cellulose et hémicelluloses majoritairement constituées de chaînes d'arabinoxylanes, ainsi qu'une faible quantité de pectines. Comme pour la paroi primaire de type I, la cellulose est retrouvée en plus grande quantité que les hémicelluloses. La paroi contient également une faible quantité de composés phénoliques: lignines et acides hydroxycinnamiques. Les lignines sont constituées, après polymérisation de trois monolignols, de trois sous-unités : *p*-hydroxyphényle (H), gâiacycle (G) et syringyle (S). Chez le maïs, les sous-unités H représentent en moyenne 4 % du polymère de lignines, les sous-unités G 35 % et les sous-unités S 61 % (Lapierre, 1993; Barrière et al, 2007). Chez les dicotylédones, le polymère de lignines est composé essentiellement de sous-unités G (entre 14 et 66 %) et S (entre 34 et 86 %) avec quelques traces de sous-unités H. Les monomères de lignines sont interconnectés par plusieurs types de liaisons. Ces liaisons peuvent être soit condensées et résistantes aux attaques chimiques comme les liaisons carbone-carbone (C-C) de type 5-5 et phényl-*O*-phényl éthers, soit non condensées regroupant les liaisons éther labiles β -*O*-4 (Terashima et Fukushima, 1993). Les principaux acides hydroxycinnamiques trouvés au niveau de la paroi sont les acides férulique (FA), diféruliques (diFA) et *p*-coumarique (pCA). Avec 5 % d'acide férulique et diféruliques et 3 % d'acide *p*-coumarique, les parois du maïs, et plus largement des graminées, sont particulièrement riches en acides hydroxycinnamiques (Grabber et al, 2004). Les acides féruliques relient les hémicelluloses aux lignines (principalement aux sous-unités G) *via* des liaisons esters (hémicelluloses-acides) et éthers (acides-lignines). Les acides diféruliques relient aussi les chaînes d'hémicelluloses entres-elles et au polymère de lignines. Les acides férulique et diféruliques sont ainsi à la base de réticulations pariétales et marquent le site de dépôt des lignines. L'acide *p*-coumarique est lié essentiellement aux sous-unités S des lignines *via* des liaisons esters et ce caractère est un bon indicateur du dépôt des lignines.

1.3. Génomique structurale du maïs

La lignée B73 représente à ce jour le génome de référence du maïs (Schnable et al, 2009, maizegenome.org). Son génome est presque entièrement séquencé et assemblé. La version actuelle de cet assemblage (B73 RefGen_v2, release 5b60) comprend plus de 110 000 gènes prédits avec Fgenesh et GeneBuilder. Cet ensemble de gènes est restreint à un sous-ensemble d'environ 39 000 gènes, par exclusion des éléments transposables, des pseudogènes, des contaminants, et des gènes "low-

confidence annotations". La première version du génome de référence (RefGen_v1) de B73 comprenait 32000 gènes filtrés (filtered set) et il semblait ainsi manquer 5% des gènes, et au plus 15%.

Les données de séquençage sont également disponibles pour deux autres lignées, Mo17 (Rokhsar et al, 2009; <http://www.phytozome.net/maize.php>) et Palomero (Vielle-Calzada et al, 2009; <http://www.ncbi.nlm.nih.gov/Traces/wgs/?val=AECO01>).

Les connaissances acquises sur ces génomes constituent une ressource de données importantes pouvant être valorisée dans différentes études de génomique structurelle et fonctionnelle de maïs. Des études sur l'origine du génome du maïs, et sur sa complexité, ont ainsi bénéficié de ces avancées de séquençage.

1.3.1. Caryotype du maïs et origine

Le maïs actuel a une méiose diploïde et possède 10 paires de chromosomes ($2n = 20$). Chaque chromosome est divisé en bins, région d'environ 20 centiMorgans (cM) délimitée par deux core markers (Gardiner et al, 1993) et désignée par le numéro du chromosome suivi par deux chiffres consécutifs désignant les régions successives sur le chromosome (01, 02, 03...). Le génome des plantes de la famille des graminées est bien conservé (Devos, 2005). Ainsi la synténie (regroupement des gènes sur un même chromosome) et la colinéarité (ordre des gènes sur les chromosomes) sont bien conservées entre le maïs, le blé et le riz. Les graminées ont en plus une origine monophylétique (Kellogg, 2001). Le progéniteur ancestral à 5 chromosomes aurait donné naissance à l'ancêtre intermédiaire des graminées à 12 chromosomes. Cet ancêtre intermédiaire, après différentes duplications, translocations réciproques et des événements de fusion de chromosome, aurait à son tour permis la formation de deux ancêtres dont dérivent toutes les graminées actuelles : un ancêtre "Pooideae" à 12 chromosomes et un ancêtre "Panicoideae" à 10 chromosomes (Murat et al, 2010). Le génome de maïs, descendant de l'ancêtre Panicoideae, a ensuite subi une duplication (plus exactement une amphipolyploïdisation) de l'ensemble de son génome il y a 5 à 12 ma, le distinguant des Panicoideae proches (Schnable et al, 2009). Selon Salse (2012), l'histoire du caryotype de maïs pourrait être résumée par l'équation $(((12 + 2 - 4) \times 2) + 7 - 17)$. Cette équation reflète tout d'abord que l'ancêtre intermédiaire des monocotylédones à 12 paires de chromosomes a subi 2 fissions et 4 fusions aboutissant à des ancêtres Panicoideae à 10 paires de chromosomes. Cet ancêtre à 10 chromosomes a ensuite subi une fusion/duplication entière du génome, formant un intermédiaire de 20 paires de chromosomes. Ce génome à 20 chromosomes a finalement subi 7 fissions de chromosome et 17 fusions, atteignant la structure de génome de maïs moderne de type allotétraploïde segmental avec 10 paires de chromosomes.

1.3.2. Taille du génome et complexité

En plus de sa structure d'allotétraploïde segmental, la taille du génome de maïs a énormément augmenté pendant les 3 derniers millions d'années, atteignant ainsi une longueur physique voisine de 2300 Mb, soit 15 fois la taille du génome d'*Arabidopsis thaliana* et 5 fois celle du riz. En plus de la grande taille de son génome, il possède un très fort contenu en séquences répétées (80 % à 85 % du génome) impliquant massivement des éléments mobiles. Le maïs est considéré par certains comme le plus grand "dépositaire" de rétrotransposons qui constituent entre 65 et 70 % de son génome (Meyers et al, 2001). Soixante-quinze pour cent de la séquence génomique de la lignée B73 est composée de rétrotransposons de type LTR (long terminal repeat) (Schnable et al, 2009) dont la forte prolifération a été responsable notamment de l'augmentation de la taille de son génome (San Miguel et al, 1998). Ces deux caractéristiques défavorables de son génome jouent *a priori* contre son séquençage. Par ailleurs, les transposons et les séquences d'ADN répétées jouent un rôle significatif dans la génération de la diversité allélique, soit avec la création de mutations annulant la fonctionnalité de la protéine, soit de façon plus intéressante en modifiant la régulation du gène.

Le génome de maïs est ainsi caractérisé par une forte diversité génétique, qui se traduit par du polymorphisme de séquence et aussi de la variation structurale (Buckler et al, 2006; Messing et Dooner, 2006). Tout d'abord, le maïs présente un taux élevé de polymorphisme de séquence, SNP (Single Nucleotide Polymorphism) et INDEL (INsertion/DEletion), par rapport à beaucoup d'espèces (Buckler et al, 2006). Ainsi, la fréquence de SNP entre lignées de maïs est plus élevée que celle observée entre les humains et les chimpanzés (Buckler et al, 2006). En moyenne, B73 et Mo17 présentent, au niveau de leur génome, un INDEL toutes les 300 paires de bases (pb) et un SNP toutes les 80 pb (Vroh et al, 2006; Fu et al, 2006), et au niveau des transcrits un SNP toutes les 300 pb (Barbazuk et al, 2007). En comparant 2 à 2 différentes lignées de maïs, on trouve, en moyenne, un polymorphisme, tout type confondu, toutes les 100 pb (Tenailon et al, 2001; Ching et al, 2002). De plus on observe de nombreuses variations de structure du génome entre différentes lignées de maïs. Ainsi entre deux lignées on peut observer plusieurs centaines de variation de copies d'un gène (Copy Number Variation, CNV), correspondant à des duplications, mais aussi plusieurs milliers de fragments de génome portant un ou plusieurs gènes et/ou éléments régulateurs, qui sont soit présents soit absents chez certains génotypes (Presence-Absence Variation, PAV) (Springer et al, 2009; Beló et al, 2010; Lai et al, 2010; Swanson-Wagner et al, 2010).

Toutes ces variations alléliques, qui incluent des variations du contenu en fragments géniques, des variations des éléments répétés environnant des gènes, et des variations entraînant des différentiels d'expression de gènes contribuent très probablement à la diversité phénotypique et à l'effet d'hétérosis (vigueur de l'hybride par rapport aux lignées parentales), résultant de la combinaison de ces variations (Springer et Stupar, 2007). Le phénomène d'hétérosis est particulièrement important chez le maïs.

Figure 10 - Evolution des données économiques sur le maïs dans le monde (source FAO).

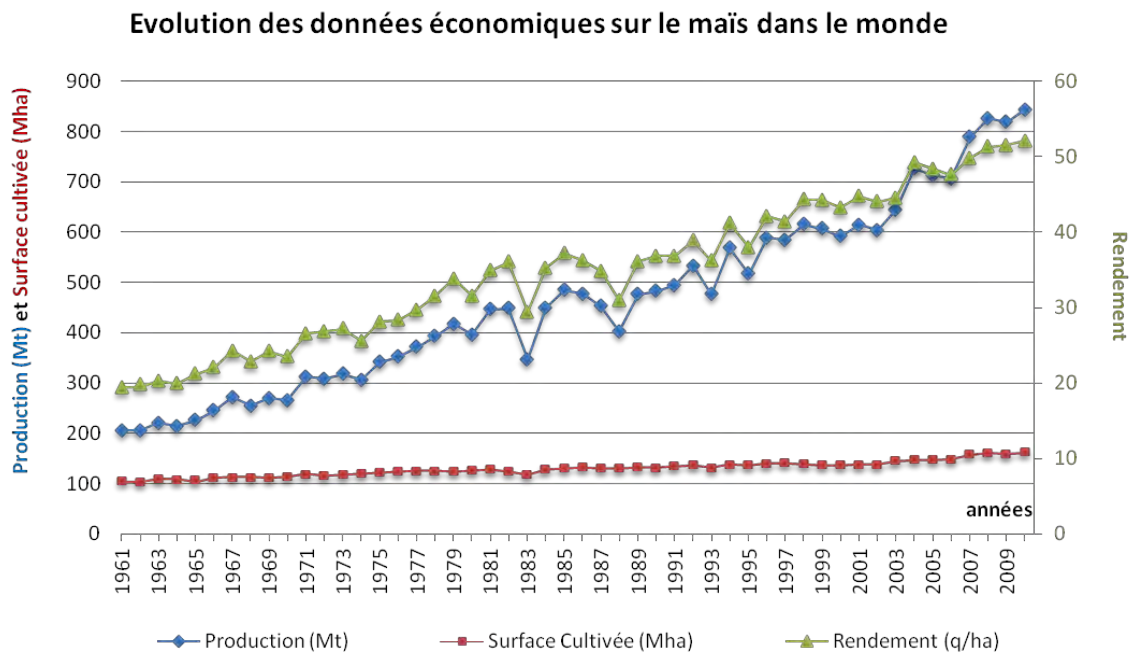


Figure 11 - Répartition de la production mondiale de maïs en 2010 par continent (source FAO).

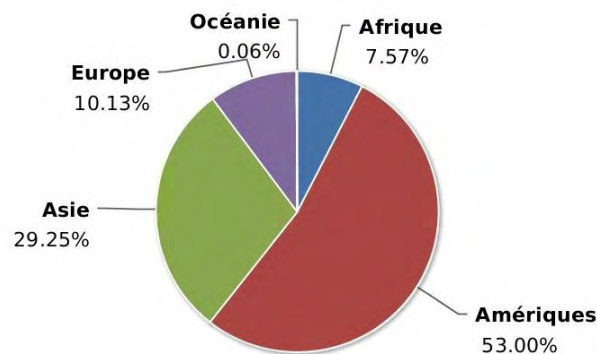
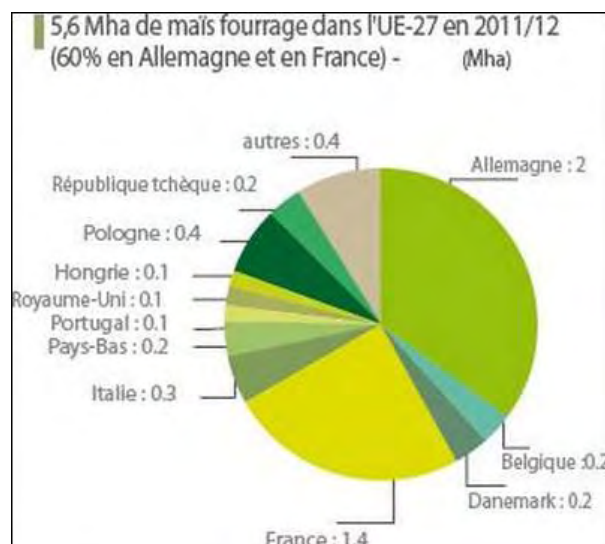


Figure 12 - Surface de maïs fourrage cultivé dans l'Union Européenne des 27 en 2011 (source A.G.P.M., Association Générale des Producteurs de Maïs).



1.4. Quelques rappels généraux sur l'importance du maïs

1.4.1. Expansion de la culture du maïs

Le maïs a été domestiqué, au Sud-Ouest du Mexique, il y a environ 7000 ans, à partir d'une descendance mutante de téosinte de la sous-espèce *Zea mays ssp parviglumis* (Matsuoka et al, 2002; Vigouroux et al, 2002). Il s'est ensuite répandu sur le continent américain avant d'être introduit en Europe par plusieurs explorateurs. Christophe Colomb fut le premier à introduire du maïs des Caraïbes en Espagne, en 1493 (Rebourg et al, 2003). De là, il s'est diffusé dans toutes les régions d'Europe méridionale au climat suffisamment chaud et humide. Le maïs fut également introduit, depuis le Canada et le Nord des USA, en Normandie par Giovanni da Verrazano (1524) et Jacques Cartier (1535), avant de gagner l'Europe vers l'est et le nord. Les régions centrales furent progressivement le lieu d'hybridation entre les écotypes de maïs entrés par le Sud de l'Europe et ceux entrés par le Nord-Ouest. L'expansion de la culture du maïs toucha aussi largement l'Afrique, vers 1540 en Egypte, et vers 1550 dans le golfe de Guinée où il est introduit par les Portugais. Il a également été introduit en Asie, et présent en Chine depuis cette même époque. Il a aussi été envisagé que ce soit le navigateur Zheng He sous la dynastie Ming qui aurait été le premier à découvrir l'Amérique en 1421 (sur la base d'une carte antérieure à Christophe Colomb sur laquelle figure clairement l'Amérique), avec une première introduction très précoce du maïs. Le maïs est ainsi aujourd'hui présent sur les cinq continents, et s'est progressivement adapté aux différents climats et latitudes, mais demeure néanmoins sensible aux basses températures, à la fois au printemps et en fin de végétation.

1.4.2. Production

La production mondiale de maïs grain est actuellement la production la plus importante au monde (844 millions de tonnes (Mt) récoltés en 2010) devant le riz (672 Mt) et le blé (651 Mt), d'après les données publiées par la FAO. Son importance économique n'a cessé d'augmenter depuis 50 ans. Sa production a été multipliée par 4,1, reflet de l'évolution des surfaces et des rendements (Figure 10). En effet, les surfaces ont progressé de 65 % et les rendements moyens en grain sont passés de 19 à 52 quintaux à l'hectare (q/ha) mais avec de grandes disparités selon les pays. En fonction des demandes intérieures, les productions sont plus ou moins exportées, et cette situation peut engendrer des fluctuations importantes des cours. La consommation mondiale ne cesse de se développer, en particulier en raison de l'augmentation de la consommation humaine de viandes animales, et se situe même depuis quelques années au dessus de l'offre ce qui oblige les principaux consommateurs à puiser dans leurs stocks. En 2010, 162 millions d'hectare (Mha) à travers le monde ont été cultivés en

Figure 13 - Répartition des surfaces de maïs grain et fourrage cultivé en Europe en 2011 (source A.G.P.M.).

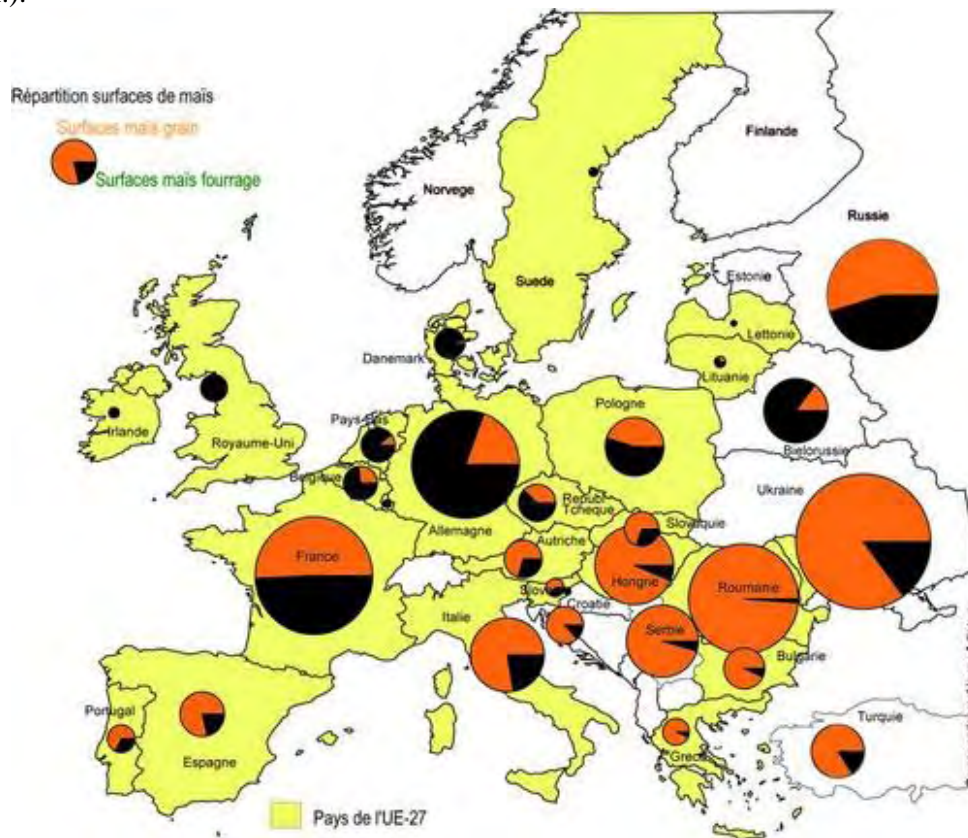


Figure 14 - Répartition de la surface cultivée des différentes cultures de maïs en France en 2011 (source A.G.P.M.).

Le maïs en France, un total de 3 Mha en 2011/12

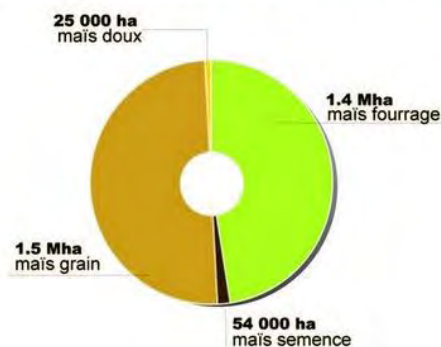
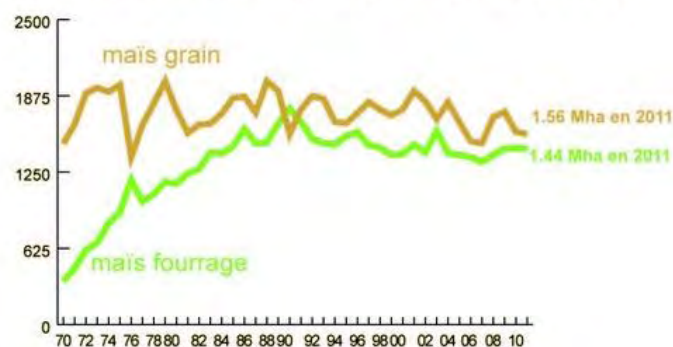


Figure 15 - Evolution des surfaces de maïs grain et fourrage cultivées en France entre 1970 et 2011 (source A.G.P.M.).

Evolution des surfaces de maïs grain et de maïs fourrage en France



maïs (4900 Mha de terre agricole dont 1300 Mha de terres arables), avec des zones à haut rendement (supérieur à 50 q/ha) limitées principalement aux Etats-Unis (rendements de 96 q/ha en 2010) et l'Europe (61 q/ha) avec notamment la France (environ 90 q/ha de moyenne sur les 10 dernières années). Près de 30 % des superficies de maïs grain sont localisées dans ces zones dont 33 Mha aux Etats-Unis et 14 Mha en Europe dont 8,5 Mha dans l'Union Européenne des 27 (UE-27). Le reste des surfaces se répartit entre l'Asie (54 Mha), le continent Africain (33 Mha) et l'Amérique du Sud (19 Mha). Ainsi au niveau de la production, quatre continents se partagent le marché (Figure 11) avec une dominance du continent américain (53 %) et asiatique (29 %). Cette production est dominée par quatre grands centres qui assurent 75 % des disponibilités mondiales. Les Etats-Unis avec une production en 2010 de plus de 283 Mt, soit près de 34 % des ressources mondiales, sont les leaders. La Chine est le deuxième producteur avec plus de 178 Mt. L'Amérique du Sud avec plus de 92 Mt produites principalement par le Brésil et l'Argentine est troisième. L'Europe suit, juste derrière, avec 86 Mt dont 14 Mt produites en France, premier pays producteur européen et huitième au niveau mondial.

La production de maïs fourrage est importante surtout au niveau de l'UE-27 et son niveau élevé de production semble être une spécificité de l'Europe. Sa surface en Europe ne cesse de se développer. Elle a augmenté pour atteindre 5,6 Mha dans l'UE-27 en 2011 (Figure 12). L'Allemagne et la France concentrent 60 % de la surface totale de l'UE-27. La surface de maïs fourrage est plus importante que la surface de maïs grain dans certains pays dont notamment l'Allemagne, le Pays-Bas, la Belgique (Figure 13), ce qui s'explique aussi par des raisons climatiques avec une maturité du grain difficile à atteindre dans les régions les plus septentrionales.

En France, le maïs est la deuxième production végétale derrière le blé en terme de surface. En 2011, un total de 3 Mha était consacrée à la culture des différents types de maïs (Figure 14) soit environ 10 % de la surface agricole utile totale (source A.G.P.M., Association Générale des Producteurs de Maïs). Le maïs grain et fourrage, représentent à eux deux presque l'intégralité des surfaces dédiées au maïs qui se répartissent dorénavant de manière à peu près équivalente. Les surfaces consacrées à la culture de maïs (grain et ensilage) se sont rapidement développées entre 1955 et 1975, passant de 450 000 à 2 837 000 ha. L'évolution de la culture du maïs fourrage s'est poursuivie jusqu'en 1986, passant de 300 000 ha en 1970 à 1 600 000 ha (Figure 15). En 1988, les surfaces totales en maïs atteignaient 3 481 000 hectares, dont 1 973 000 hectares pour le grain et 1 508 000 hectares pour le fourrage. Depuis, les surfaces ont peu évolué, oscillant ces dernières années autour de 3 Mha. Cette évolution correspond à l'extension géographique de la culture. Presque uniquement localisée dans le Sud-Ouest dans les années 50, elle est désormais présente sur l'ensemble du territoire français. Cette extension est liée à la création de variétés de plus en plus précoces pouvant s'adapter à des régions plus fraîches du pays.

Figure 16 - Les filières de valorisation énergétique de la biomasse en biocarburants (source <http://www.plateforme-biocarburants.ch>).

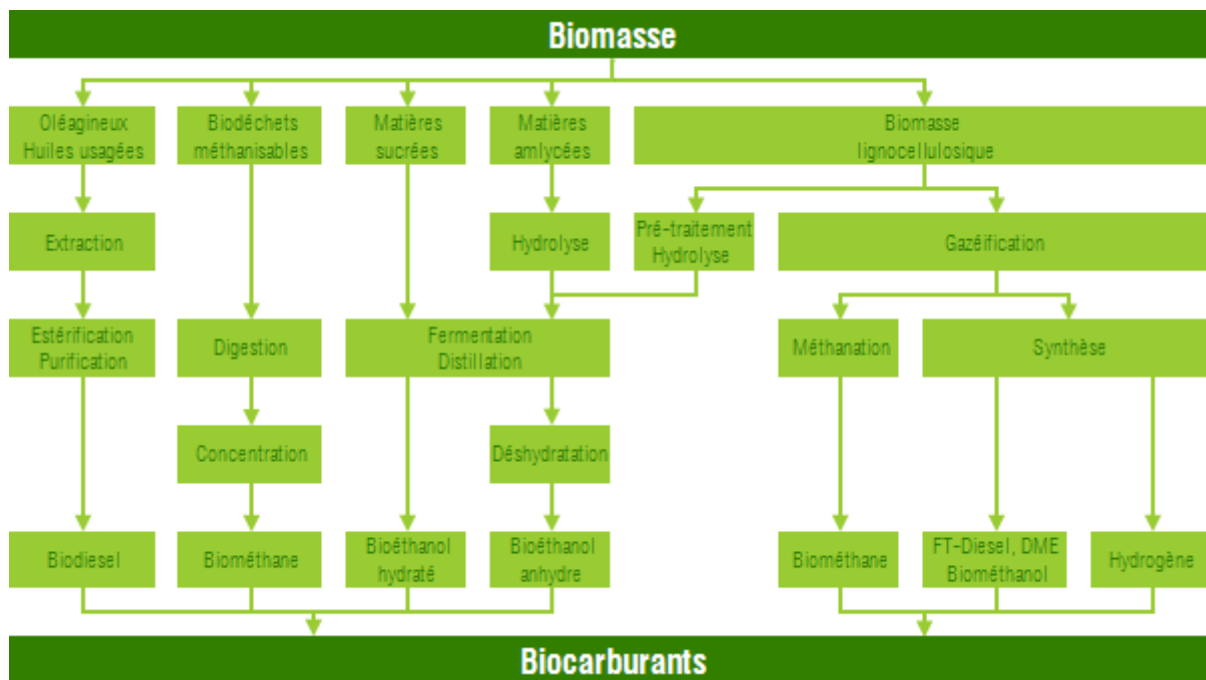
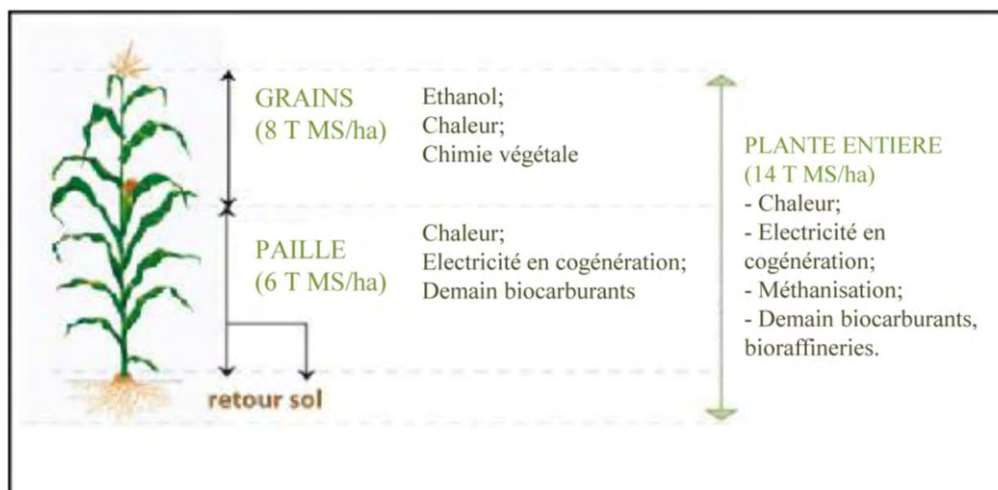


Figure 17 - Les filières de valorisation énergétique de la biomasse maïs (source A.G.P.M.).



1.4.3. Utilisations du maïs et enjeux

1.4.3.1. Utilisations

Simple à produire, à récolter et à stocker, le maïs possède une grande aptitude à produire de la biomasse lignocellulosique et amylacée qui présente une forte valeur alimentaire et énergétique. Le maïs peut être valorisé aussi bien par son appareil végétatif que par ses grains. L'ensemble de ses atouts, économiques et nutritionnels, en grain comme en fourrage, est responsable de son succès planétaire, en premier lieu en matière d'alimentation animale et humaine, mais aussi dans de nombreuses applications énergétiques ou non-énergétiques (biomatériaux et chimie du végétal).

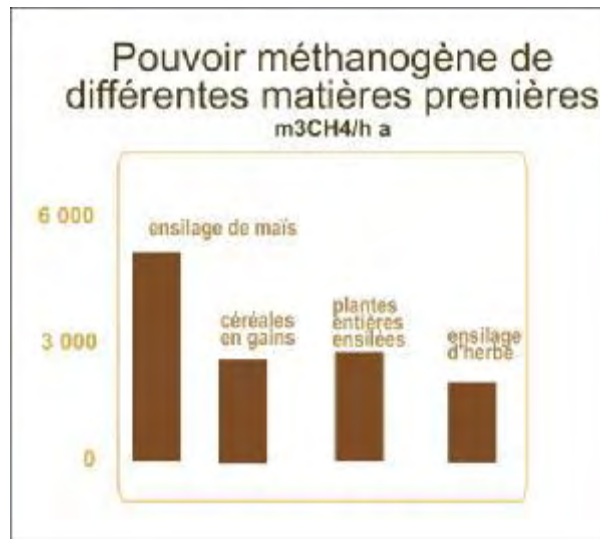
De nos jours, 70 % de la consommation mondiale de maïs, que ce soit sous forme grain ou plante entière ensilée (maïs fourrage), correspond à l'alimentation animale. Le maïs, qui offre une source d'alimentation animale sûre et riche, est un aliment de base pour les animaux (volailles, porc et bovins). Il permet d'augmenter la productivité en lait, en viande et en œufs et d'obtenir des produits laitiers et de la viande de très bonne qualité.

L'utilisation du maïs en tant que source d'énergie renouvelable tend à se développer et s'inscrit dans une politique globale "d'énergies renouvelables". La politique européenne énergétique vise, pour 2020, 20 % d'énergies renouvelables dans la note énergétique globale, dont 10 % dans la consommation des transports. Les principales filières de production des biocarburants sont indiquées sur la Figure 16. Le maïs, grain ou plante entière, est une source d'énergie renouvelable permettant principalement la production de bioéthanol incorporé à l'essence, de biogaz ou encore, de chaleur et d'électricité (Figure 17). La conversion biologique de la biomasse amylacée ou lignocellulosique en bioéthanol consiste en deux étapes : la conversion enzymatique des polysaccharides pariétaux complexes en sucres simples fermentescibles (la saccharification/hydrolyse) et la fermentation. Le biogaz est produit par méthanisation, qui correspond à une fermentation anaérobie appelée aussi digestion. Le biogaz est généralement utilisé comme combustible pour la production locale de chaleur ou la cogénération d'électricité et de chaleur. Toutefois, l'utilisation du biogaz comme carburant dans des véhicules à gaz semblerait se développer de plus en plus.

1.4.3.2. Valorisation alimentaire et énergétique du grain

Le maïs grain est un aliment complet (sucres, lipides, minéraux et vitamines) à forte valeur énergétique grâce à sa richesse en amidon et à l'huile de son germe. Il possède la valeur énergétique la plus élevée de toutes les céréales (1,27 UFL/kg, Unité Fourragère Lait, 110 % par rapport à l'orge). Il est en revanche pauvre en protéines (10 %), et les rations animales doivent être complétées avec des tourteaux de soja. D'un point de vue énergétique, 0,44 UFL sont nécessaire pour produire 1 kg de lait, ainsi 1 kg de maïs grain (1,27 UFL/kg) permet de produire 2,9 kg de lait (sous réserve de

Figure 18 - Pouvoir méthanogène de différentes matières premières dont le maïs fourrage (source A.G.P.M.).



complémenter en matières azotées, et en prenant en compte par ailleurs les besoins métaboliques d'entretien, et éventuellement de gestation, des vaches laitières). Le grain, destiné à l'usage alimentaire des animaux, peut être utilisé sous plusieurs formes. Les grains récoltés humides peuvent être broyés puis ensilés pour être utilisés essentiellement pour l'alimentation des porcs, parfois mélangés à des rafles broyées (corn-cob-feed), tout au long de l'année. Les grains sont également consommés sous forme sèche directement incorporés dans les rations des animaux (volailles) ou sous forme d'aliments composés (farines et granulés).

Le grain est utilisé dans la production de bioéthanol (fermentation de l'amidon). Ce biocarburant est dit de première génération. Un hectare de maïs permet de produire 34 hl de bioéthanol, contre 24 pour les céréales comme le blé et 38 pour le sorgho.

1.4.3.3. Valorisation alimentaire et énergétique du maïs plante entière

Le maïs plante entière, récoltée avant la maturité du grain et ensilé (maïs fourrage), est une source importante de biomasse lignocellulosique et d'énergie (0,90 UFL/kg, au moins 130 % par rapport à un foin de graminées). Le maïs fourrage est devenu, aux côtés des graminées prairiales, un des piliers des systèmes d'alimentation des bovins où il représente jusqu'à 60 % de la ration des animaux. Le maïs fourrage est ainsi, en France, le fourrage annuel le plus cultivé (84 % de la surface totale couverte en fourrages annuels). Bien conservé, il est apprécié des vaches et garde une bonne valeur énergétique. Il peut apporter à la vache jusqu'à 80 % de l'énergie et 40 % de l'azote, qui lui sont nécessaires. D'un point de vue énergétique, 1 kg de maïs ensilage (0,90 UFL/kg) permet de produire 2,0 kg de lait (sous réserve de compléter en matières azotées, et en prenant en compte par ailleurs les besoins métaboliques d'entretien, et éventuellement de gestation, des vaches laitières).

La plante entière de maïs est également une source d'énergies renouvelables diverses : chaleur, électricité, biogaz et bioéthanol. Tout d'abord, la matière sèche peut être valorisée dans la production de chaleur et d'électricité par combustion des rafles maïs aussi de la tige et des feuilles. Depuis peu, l'ensemble de la biomasse du maïs et les résidus de cultures sont considérés, comme source majeure de polysaccharides pariétaux, pour la production de biocarburant de deuxième génération : bioéthanol et biogaz. D'une part, le maïs fourrage est un substrat fort intéressant pour la production de biogaz, du fait de son pouvoir méthanogène supérieur à celui observé pour d'autres matières premières (Figure 18). D'autre part, l'importante biomasse lignocellulosique du maïs riches en sucres pariétaux permet la production de bioéthanol. La production de biocarburant de deuxième génération représente une alternative aux carburants d'origine fossile (pétrole, gaz) et aux carburants de première génération qui entrent en compétition avec l'alimentation. Le développement de biocarburants de deuxième génération à partir de la biomasse lignocellulosique est fondamental pour le futur et représente une solution qui devrait permettre de répondre aux objectifs fixés en matière d'énergies renouvelables dans la consommation finale de carburants par les transports (10 % en 2020) sans impacter l'alimentation.

2. Vers l'amélioration de la valeur alimentaire et énergétique de la biomasse lignocellulosique chez le maïs

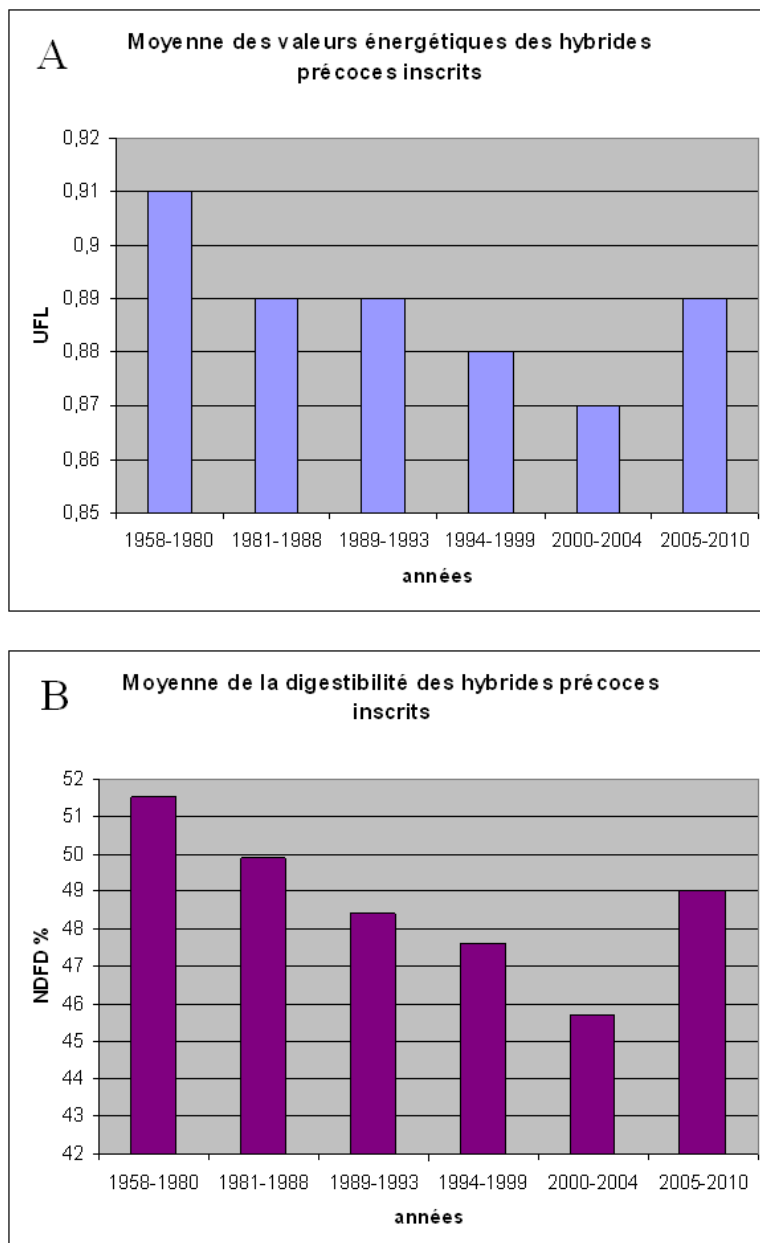
Parmi toutes les parties utilisables du maïs, la plante entière ainsi que les résidus de culture du maïs grain représentent une biomasse importante qui est une source de polysaccharides (amidon, cellulose, sucres solubles, hémicelluloses) valorisables en énergie. Le maïs est, de ce fait, une plante d'intérêt économique à la fois dans l'alimentation animale mais aussi depuis peu dans la production de biocarburants de deuxième génération (bioéthanol et biogaz). Au vu de ces deux utilisations, il est important de connaître les facteurs qui peuvent limiter l'apport alimentaire et énergétique du maïs plante entière.

2.1. La valeur alimentaire et énergétique de la biomasse lignocellulosique

La valeur alimentaire d'une plante fourragère est la quantité d'éléments nutritifs apportée par la plante lorsqu'elle est distribuée à volonté aux animaux. Cette valeur alimentaire intègre deux notions : la valeur nutritive et l'ingestibilité (Baumont et al, 1999). L'ingestibilité permet d'évaluer la quantité de l'aliment que l'animal peut ingérer volontairement en l'absence de choix. Elle est exprimée par sa valeur d'encombrement en UE (Unité d'Encombrement). La valeur nutritive d'un aliment permet, quant à elle, d'évaluer la contribution de l'aliment à la couverture des besoins nutritionnels de l'animal. Elle recouvre sa valeur énergétique exprimée en UF (Unité Fourragère), sa valeur azotée exprimée en PDI (Protéines Digestibles dans l'Intestin) et sa teneur en minéraux. La teneur en matière azotée du maïs n'apparaît pas comme une composante limitante de la valeur nutritive et donc alimentaire car elle est facile à corriger par un apport complémentaire d'azote dans les rations. Par ailleurs, la valeur azotée de la plante entière présente peu de variabilité chez le maïs, et une héritabilité faible. En revanche, la valeur énergétique des aliments est une composante clef de la valeur nutritive et alimentaire. Le principal facteur de variation de la teneur en énergie nette des aliments est la digestibilité de l'énergie brute qu'ils contiennent et qui est très étroitement liée à la digestibilité de la matière organique (dMO). La digestibilité, exprimée en %, correspond à la proportion d'un aliment qui disparaît dans le tube digestif et donc à la part de la nourriture valorisée par l'animal. Ainsi, avec l'ingestibilité, la digestibilité est un facteur clef de la valeur alimentaire. Par ailleurs, il existe des relations significatives entre digestibilité et ingestibilité, la digestibilité expliquant 50 % de la variabilité observée pour l'ingestibilité (Barrière et al, 2003a).

L'ingestibilité est liée pour l'animal à la vitesse de réduction en petites particules de l'aliment dans la bouche, à sa vitesse de dégradation par les micro-organismes du rumen, et finalement à la vitesse de sortie des particules indigestibles (et/ou non digérées) hors du rumen vers le feuillet par l'orifice réticulo-omasal (Minson et Wilson 1994; Jung et Allen 1995; Fernandez et al, 2004). L'ingestibilité

Figure 19 - Valeurs phénotypiques de deux caractères mesurés chez des hybrides maïs inscrits entre 1958 et 2010 (essai moutons effectués à l'INRA de Lusignan entre 1960 et 2004, les valeurs 2005-2010 sont estimées à partir de régressions sur des données in vitro). A - Valeur énergétique UFL (Unité Fourragère Lait). B - digestibilité (NDFD : Digestibilité des parois mesurée par des moutons en cage).



conditionne autant la variabilité des nutriments énergétiques apportés par l'ensilage de maïs au bovin que la digestibilité de cet ensilage. Les comparaisons d'hybrides conduites à Lusignan ont permis de mettre en évidence l'ingestibilité exceptionnelle de l'hybride Dk265 (obtention RAGT-semences, 1986), qui est depuis un modèle pour la recherche. Il existe clairement des déterminants propres de l'ingestibilité, pour un niveau de digestibilité donné. Ces déterminants ne sont pas complètement identifiés ou compris, mais ils sont très certainement liés à la friabilité des tissus, qui serait elle-même liée à la réticulation entre les composants des parois. Il existe ainsi des relations significatives entre digestibilité, ingestibilité et vitesse de dégradation dans le rumen (Barrière et al, 2003a). Les facteurs génétiques spécifiques correspondants semblent rares dans le matériel maïs actuel, et leur existence n'a été clairement prouvée que chez l'hybride Dk265, et sans doute dans l'hybride Bahia de Syngenta, à un moindre niveau (Barrière et al, 1995; Barrière et al, 2004b).

La production d'un biocarburant de 2^{ème} génération est basé sur la valorisation de la biomasse cellulosique et dépend de la teneur en sucres qui peuvent être transformés en énergie. Les sucres complexes (cellulose, hémicelluloses) doivent dans un premier temps être convertis en sucres simples fermentescibles par hydrolyse (réaction consistant à rompre les longues chaînes de polymères). Cette étape dite de saccharification dépend de la dégradabilité de la biomasse lignocellulosique par les enzymes hydrolytiques tout comme c'est le cas pour l'alimentation animale. Le terme digestibilité, utilisé en premier lieu avec l'utilisation du maïs dans l'alimentation animale, peut ainsi aujourd'hui, avec l'essor des biocarburants, être remplacé par le terme dégradabilité.

2.2. Amélioration et prédiction de la valeur énergétique et de la dégradabilité de la biomasse lignocellulosique

Il a longtemps été considéré que la productivité en terme de biomasse et la qualité d'une plante entière de maïs pouvaient être déterminées par l'abondance des grains. La distinction du maïs fourrage et du maïs grain est d'ailleurs assez récente. Après son introduction en Europe, le maïs "grain" a d'abord été utilisé en fourrage vert immature, récolté bien avant le remplissage du grain, si ce n'est avant la floraison. L'utilisation en ensilage a progressivement remplacé l'utilisation en fourrage au cours de la première moitié du XX^{ème} siècle. La sélection des variétés a longtemps été faite sans prise en compte de caractères plante entière (valeur énergétique et digestibilité) et a ciblé la productivité en grain, la résistance à la verse, à la casse, ainsi qu'à la pyrale et à la pourriture des tiges, ce qui a conduit à une perte de la valeur énergétique UFL et de la digestibilité chez les hybrides inscrits entre 1958 et 2010 (Figures 19A et 19B). Une contre sélection non prévue de ces derniers caractères a sans doute eu lieu car une valeur alimentaire élevée est souvent antagoniste d'une bonne tenue de tige ou d'une bonne résistance à la casse ou aux ravageurs. Mais, il ne faut pas exclure également, avec un poids sans doute au moins égal, une dérive génétique, corrélative à la non prise en compte des

caractères de valeur alimentaire, et au choix, sur d'autres critères, de ressources génétiques comme le groupe "Iodent" ayant, pour partie "par hasard", une plus faible valeur énergétique en plante entière. Ainsi seulement 32 % des hybrides inscrits avant 1980 avaient une digestibilité inférieure à celle de l'hybride bien connu LG11, inscrit en France en 1970. Cette proportion monte à 93 % avec les hybrides inscrits à partir de 1999. La diminution moyenne de valeur énergétique observée par rapport à LG11 correspond environ, dans la ration journalière d'une vache laitière, à l'énergie nécessaire à la production de 2 à 4 kg de lait par animal et par jour.

Cette situation a conduit à la prise en compte du critère valeur énergétique (UFL) pour l'inscription des variétés au catalogue officiel à partir de 1999 (Barrière et al, 2004a; Murigneux et al, 2006). L'amélioration de la valeur énergétique est liée à une amélioration de la dégradabilité des parois, qui est alors devenue un nouveau critère clef dans la sélection des variétés dont la biomasse lignocellulosique est utilisée à des fins alimentaires et de biocarburants. La valeur énergétique et la digestibilité peuvent être mesurées *in vitro*, en utilisant notamment des solubilités enzymatiques (Aufrère et Michalet-Doreau, 1983). Ainsi, la mesure de la digestibilité *in vitro* de la matière sèche IVDMD (*in vitro* dry matter digestibility) est basée sur la mesure de la quantité de la matière sèche solubilisée après dégradation par des enzymes (amylase, pepsines et un mélange de cellulases et d'hémicellulases), d'après Aufrère et Michalet-Doreau (1983). Cette mesure, associée à la teneur en matières azotées totales, est une composante des équations de prévision actuellement utilisées pour prédire la valeur énergétique (Andrieu et Aufrère, 1996, Andrieu et al, 1999). Par ailleurs, la solubilité enzymatique permet d'obtenir un classement des génotypes sur leur valeur de digestibilité, reflétant assez bien celui obtenu à partir des mesures effectuées sur moutons, au moins sur de larges gammes de variations. A partir de cette mesure de solubilité et de celle de la teneur en paroi NDF (Neutral Detergent Fibre), il est possible d'estimer la digestibilité des parois cellulaires selon deux méthodes. La première méthode d'estimation est celle de Struik (1983) et Dolstra et Medema (1990), qui permet de calculer la valeur appelée IVNDFD (*In vitro* NDF digestibility) en supposant que la partie non-NDF de la plante est complètement digestible $[IVNDFD = 100 \times (IVDMD - (100 - NDF)) / NDF]$. La deuxième méthode, selon Argillier et al, (1995) et Barrière et al, (2003b), permet d'estimer la valeur DINAGZ correspondant à la digestibilité *in vitro* de la matière sèche (IVDMD) à laquelle est soustrait la teneur en amidon ST (starch), en glucides solubles SC (soluble carbohydrate) et en protéine brutes CP (crude protein), considérant que ces trois constituants sont complètement digestibles $[DINAGZ = 100 \times (IVDMD - ST - SC - CP) / (100 - ST - SC - CP)]$. L'ensemble de ces valeurs peut également être prédit en spectrométrie dans le proche infrarouge, NIRS (Near Infrared spectroscopy, Roussel et al, 2002; Barrière et al, 2007). Contrairement à ces caractères, l'ingestibilité est un caractère difficile à évaluer en routine, du fait qu'il n'en existe pas de prédicteurs en laboratoire actuellement. L'estimation de l'ingestibilité d'un hybride ne peut se faire que par une mesure sur des bovins en production en dépit du test de plusieurs approches, dont des tests de résistance mécanique. Ceci conduit à des coûts expérimentaux extrêmement élevés, et à la possibilité d'étudier seulement un

Tableau 2 - Proportion et digestibilité moyenne des principaux constituants du contenu cellulaire du maïs (en % du total contenu cellulaire + parois, teneur en matière sèche de la plante entière supérieure à 25%).

Composition du contenu cellulaire	Amidon	Glucides solubles	Protéines	Cendres	Lipides
% de la matière sèche	30%	10%	8%	4%	4%
Digestibilité moyenne	100%	100%	90-99%	10-95%	90-99%

Tableau 3 - Proportion et digestibilité moyenne des principaux constituants de la paroi cellulaire du maïs (en % du total contenu cellulaire + parois, teneur en matière sèche de la plante entière supérieure à 25%).

Composition de la paroi cellulaire	Hémicelluloses	Cellulose	Lignines
% de la matière sèche	20%	20%	4%
Digestibilité moyenne	20-100%	20-100%	0%

nombre très limité de génotypes. Ainsi, pour l'instant, l'ingestibilité, facteur limitant l'alimentation animale, ne fait pas partie des critères de sélection, mais fait l'objet de travaux de recherche.

2.3. Les parois lignocellulosiques, facteurs limitant la dégradabilité

Une plante de maïs est constituée en moyenne de 56 % de contenu cellulaire (grain compris) dont les constituants sont quasiment entièrement digestibles/dégradables (Tableau 2). Le contenu cellulaire est donc un facteur non limitant de la valeur énergétique du maïs. Les 44 % restants correspondent à la paroi cellulaire dont les différents constituants présentent une variabilité importante de la dégradabilité (Tableau 3). Ainsi des variations génétiques très significatives de la dégradabilité *in vivo* des parois ont été observées par Barrière et al (2003b et 2004a) avec des valeurs allant de 35 % à 60 %. Pour un génotype donné, la dégradation des tissus diminue en même temps que la plante vieillit, les parois s'épaississant et subissant des changements de composition chimique. La dégradabilité de la plante entière serait un reflet de sa teneur en parois végétales et de leur dégradabilité (Baumont et al, 2008) et dépendrait ainsi à la fois de caractères biochimiques et histologiques (Barrière et al, 2009b; Jung et al, 2012).

2.3.1. Facteurs biochimiques

Les parois sont constituées de différents polymères organisés et liés entre-eux. La composition biochimique des parois et les structures engendrées par les liaisons entre ses polymères sont largement responsables des variations observées pour la dégradabilité des parois.

Tout d'abord, les différents composés majeurs présentent une variabilité importante de leur dégradabilité. Alors que la cellulose et les hémicelluloses sont *a priori* totalement dégradables à l'état pur, les lignines ne le sont pas et créent des barrières physiques aux enzymes de dégradation. La cellulose et les hémicelluloses, au sein de la paroi, ont alors une dégradabilité qui varie entre 90 % et 40 % selon qu'elles sont plus ou moins "incrustées" de lignines. La teneur en lignines est considérée comme le principal facteur limitant de la digestibilité des fourrages au sein du rumen des bovins (Jung et Deetz, 1993; Jung et Allen, 1995; Argillier et al, 1998; Méchin et al, 2000; Zhang et al, 2011) et de l'hydrolyse des polysaccharides dans les bioréacteurs pour la production de biocarburant (Carroll et Somerville, 2009; Grabber et al, 2009; Harris et DeBolt, 2010; Li et al, 2010; Pauly et Keegstra, 2010). Toutefois, les variations de teneurs en lignines expliqueraient seulement 50 % de la variation observée pour la dégradabilité (Barrière et al, 2009b). La teneur en lignines peut être dosée selon deux méthodes. La première méthode appelé ADL/ NDF, permet de mesurer la partie acido-insoluble des lignines, ADL (Acid Detergent Lignin) comprise dans les parois (NDF) (Goering et Van Soest, 1970), la partie acido-soluble des lignines étant éliminée lors de la première étape chimique de cette

procédure. Les lignines dosées par l'ADL sont négativement liées à la digestibilité des fourrages par les animaux. La deuxième méthode, qui est la méthode Klason, (KL), permet de mesurer la teneur de l'ensemble des lignines présentes dans les tissus (Dence et Lin, 1992). Les teneurs en lignines Klason ne sont généralement que peu liées à la digestibilité des fourrages par les animaux.

Des liaisons labiles éther (liaison β -O-4), en plus des liaisons résistantes C-C, relient les différentes sous-unités des lignines entre elles. Il a été montré récemment que la teneur en liaisons β -O-4 limite la dégradabilité (Zhang et al, 2011) bien que ces liaisons soient les plus faciles à hydrolyser que les liaisons C-C. Ces liaisons conduiraient en fait, à des polymères de lignines non condensés, linéaires et "adsorbés" à plat sur les chaînes de polysaccharides (Besombes et Mazeau, 2005), ce qui rendrait plus difficile l'accès des enzymes hydrolytiques aux polysaccharides.

La teneur en acide *p*-coumarique estérifié (pCAest) est également corrélée négativement à la dégradabilité des parois (Méchin et al, 2000; Fontaine et al, 2003; Méchin et al 2005; Riboulet et al, 2008b; Taboada et al, 2010; Zhang et al, 2011). Le pCA, qui est lié essentiellement aux sous-unités S des lignines, sous-unités déposées en dernier, est de ce fait un signe de dépôt tardif des lignines, expliquant probablement en partie cette corrélation négative (Grabber et al, 2004). De plus, le pCAest semble être présent préférentiellement dans les polymères de lignines non condensés, probablement dû au fait que les unités S-pCA qui arrivent au niveau de la paroi sont plus propices à des liaisons β -O-4 (Zhang et al, 2011). Cette acylation influencerait donc la structure géométrique des lignines et ainsi l'accessibilité des polysaccharides des parois cellulaires aux enzymes hydrolytiques.

Les acides férulique et diféruliques ont également un effet négatif sur la dégradabilité des polysaccharides et leur fermentation (Grabber et al, 1998a; Grabber et al, 1998b; Jung et al, 1998; Grabber et al, 2004; Grabber, 2005; Grabber et al, 2009). En effet, ces acides sont à la base de réticulations pariétales qui, tout comme les lignines, agissent comme des barrières empêchant les enzymes hydrolytiques d'atteindre les polysaccharides pariétaux dans le rumen des animaux ou dans les fermenteurs industriels limitant ainsi l'hydrolyse enzymatique. Des corrélations négatives entre présence d'acide férulique étherifié (etherfA) et digestibilité ont été observées (Méchin et al, 2000; Méchin et al, 2005; Riboulet et al, 2008b). Fontaine et al (2003) considèrent que la teneur en acide férulique étherifié est un critère à prendre en compte dans l'amélioration de la valeur alimentaire du maïs fourrage. De manière plus générale, les réticulations pariétales, qui participent probablement à la rigidité de la tige (Schopfer, 1996; MacAdam et Grabber, 2002), à la structure et aux propriétés mécaniques de la paroi végétale (Ranocha et al, 2002; Fernandez et al, 2004) pourraient être des cibles intéressantes pour l'amélioration de la valeur alimentaire du maïs fourrage et plus particulièrement de l'ingestibilité (dont les déterminants sont en grande partie liés aux qualités mécaniques des tissus lors de la mastication), mais aussi pour la production de biocarburant.

Finalement, la composition monomérique (H, G et S) des lignines semble influencer la dégradabilité mais reste encore un sujet à controverse. Les résultats peuvent être différents d'une part selon l'espèce étudiée, mais aussi selon le fond génétique. Ainsi chez le maïs, l'impact des trois sous-

unités des lignines sur la dégradabilité de la plante peut être variable. Dans certains cas, la composition du polymère de lignines aurait peu d'influence sur la dégradabilité pariétale (Grabber et al, 1997; Zhang et al, 2011) alors que dans d'autres cas la dégradabilité des parois semble être corrélée positivement au ratio S/G, dû à un impact à la fois négatif de la teneur en sous-unité G et à la fois nul de la teneur en sous-unités S (Méchin et al, 2000; Méchin et al, 2005). Elle peut être aussi corrélée négativement à la teneur en sous-unité H et/ou S (Riboulet et al, 2008b; Barrière et al, 2009b), résultat contradictoire avec celui de Méchin et al (2000 et 2005). Des études avec des cellules de maïs dont les parois sont lignifiées artificiellement ont montré également que la composition en monomère de lignines n'affectait pas la fermentation de la paroi cellulaire (Grabber et al, 2009). Des études sur des plantes d'*Arabidopsis* présentant des teneurs en sous unités G ou S modifiées ont également montré que la composition en monomère n'affectait pas la digestibilité enzymatique (sauf en cas de prétraitement (Li et al, 2010). D'après Grabber et al (1997), il est possible que les variations de dégradabilité associées à la composition des lignines soient en fait dues à d'autres changements associés de structure et de chimie de la paroi. Une substitution des monolignols par divers composés phénoliques atypiques permet d'améliorer la dégradabilité (Grabber et al, 2010; Ralph, 2010) soit en limitant la lignification, soit en réduisant l'hydrophobicité de la lignine ou les liaisons aux polysaccharides, soit en créant des liaisons au sein du polymère de lignines plus facilement clivables.

Deux modèles de régression à deux variables prenant en compte les facteurs biochimiques, (KL, pCAest, etherFA, teneur en sous-unité S ou G, ratio S/G) ont été établis (Méchin et al, 2000; Méchin et al, 2005). Ces deux modèles font intervenir les deux variables pCAest et ratio S/G et permettraient d'expliquer seulement une partie des variations observées pour la digestibilité (58 % et 57 %).

2.3.2. Facteurs histologiques

Les feuilles et les tiges de maïs sont constituées de différents tissus composés de types cellulaires différents dont le degré de différenciation dépend de l'âge et du type. Certains tissus contiennent des types cellulaires caractérisés par la présence de parois secondaires lignifiées, parois qui présentent une épaisseur, une structure et une composition particulière. Les différents tissus des tiges de maïs : parenchyme, fibres, vaisseaux conducteurs présentent notamment une grande variabilité du degré de lignification de leur paroi secondaire (Méchin et al, 2000; Méchin et al, 2005). Les variations architecturales, dues à des différences de proportion et de répartition spatiale des tissus, au sein du système aérien de la plante sont ainsi également responsables des variations observées de la dégradabilité des tiges.

En 1995, Wilson et Mertens ont émis l'hypothèse que l'architecture des tissus et des cellules des graminées influençait la dégradabilité de la plante, et ce probablement de façon plus importante que la composition chimique des parois, la structure anatomique influençant d'abord l'accessibilité aux parois par les enzymes de dégradation. L'épaisseur des parois limiterait la surface à l'intérieur des

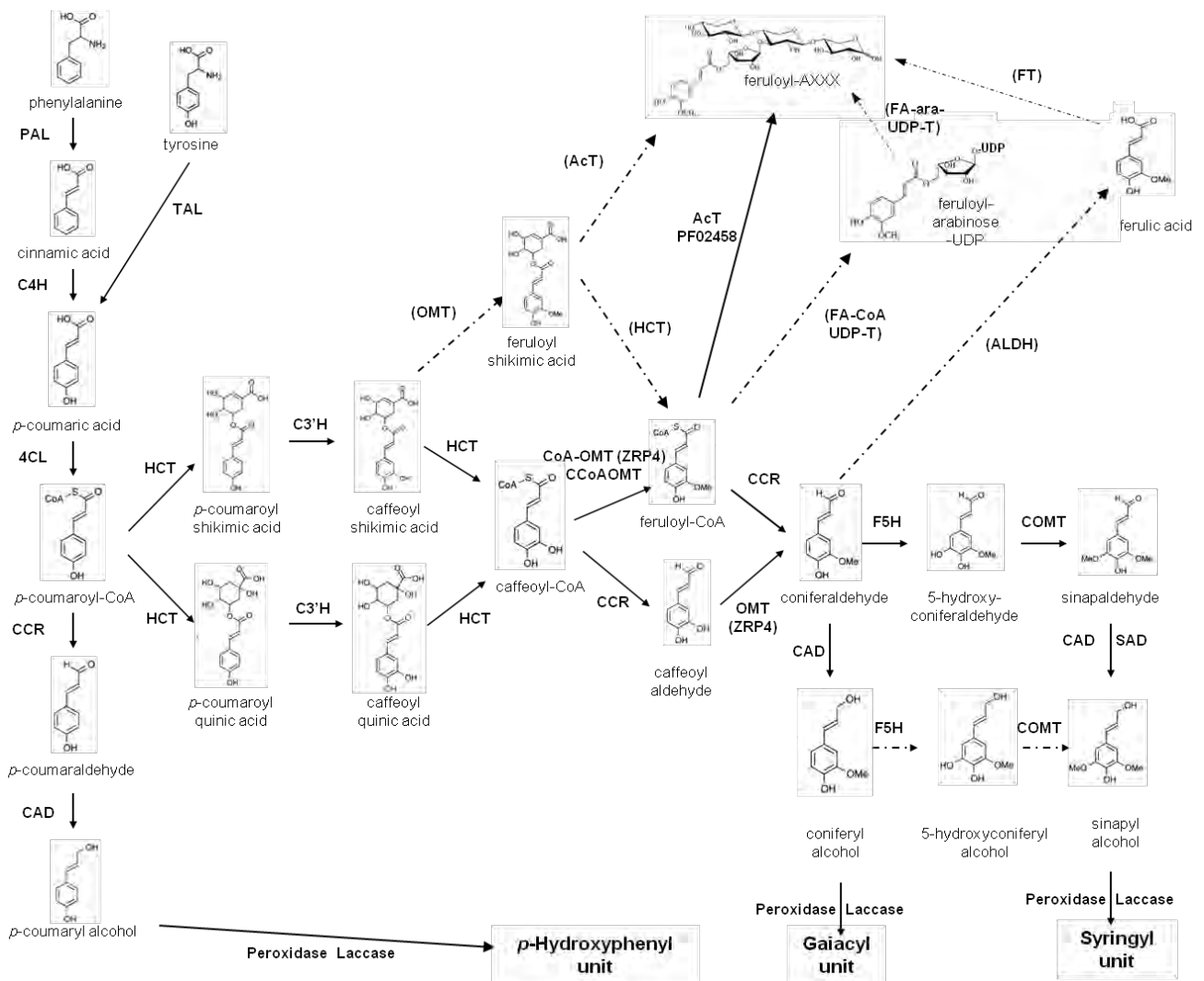
cellules et donc le déplacement des enzymes, d'où une dégradation limitée du sclérenchyme (Wilson et Hatfield, 1997). Le mutant *brown-midrib bm3*, plus digestible que le maïs normal, a des parois du sclérenchyme plus fines et en partie dégradées. Dans le rumen, le phloème et le parenchyme sont dégradés et disparaissent les premiers, tandis que les tissus lignifiés (sclérenchyme, fibres périfasciculaires et xylème) sont peu ou pas dégradés (Grenet, 1997).

Chez le maïs, des variations de surface des tissus lignifiés, de surface de cortex, d'épaisseur et de nombre de couches cellulaires dans le sclérenchyme cortical ont été observées au sein de coupes transversales d'entre-nœud sous épi au stade maturité ensilage (Méchin et al, 2005; Riboulet et al, 2008b). Certaines de ces variations ont montré un impact significatif sur la dégradabilité (Méchin et al, 2005). La surface lignifiée correspondant à la surface rouge (Ar) révélée par la coloration différentielle Fasga (Tolivia et Tolivia, 1987 et Figure 6) a été corrélée positivement à la teneur en lignine Klason (KL) et négativement à la digestibilité IVNDFD (*In Vitro* Neutral Detergent Fiber Digestibility). L'importance relative des tissus lignifiés, $Ar/(Ar+Ab)$, où Ab correspond à la surface non lignifiée colorée en bleue, est aussi corrélée positivement à la teneur en lignine Klason (KL) et négativement avec la digestibilité IVNDFD. Le nombre de couches cellulaires dans le sclérenchyme cortical (Ncs) a également été corrélé positivement à la teneur en lignine Klason (KL). Parmi les caractères biochimiques (NDF, KL, S, G, pCAest, etherFA estérifié) et histologiques (Ar, Ab, Ncs) étudiés par Méchin et al (2005), la surface de tissu lignifié Ar, à elle seule, expliquerait 40 % des variations observées pour la digestibilité de la paroi cellulaire et serait ainsi le caractère le plus représentatif de la digestibilité de la paroi cellulaire. La prise en compte des facteurs histologiques Ar et Ncs, en plus des facteurs biochimiques (estpCA et S/G), dans l'équation de régression permettant d'expliquer les variations de digestibilité IVNDFD permet d'améliorer très nettement cette explication. L'équation de régression expliquerait alors 89 % des variations de digestibilité contre 57 % quand seuls les facteurs biochimiques sont pris en compte (Méchin et al, 2005).

2.4. Les gènes impliqués dans la formation des parois secondaires lignifiées

L'impact des lignines, des acides hydroxycinnamiques, a conduit à de nombreuses études visant à identifier les gènes impliqués dans la formation des parois secondaires lignifiées chez le maïs mais aussi chez d'autres espèces (*Arabidopsis*, *Eucalyptus*, peuplier, riz...). Ces gènes sont (1) les gènes de la voie de biosynthèse des lignines, (2) les gènes de la voie de biosynthèse des acides hydroxycinnamiques, (3) les facteurs de transcription régulant ces voies de biosynthèse et/ou la mise en place des tissus lignifiés. Les gènes potentiellement impliqués dans la lignification appartiennent à d'importantes familles multigéniques et la fonction de chaque homologue est encore inconnue, en particulier à cause de redondance fonctionnelle soupçonnée. Une synthèse de l'ensemble de ces gènes chez le maïs a été faite par Barrière et al (2009a) et a été remise à jour depuis au vu de l'avancée des

Figure 20 - Voies putatives de biosynthèse des monolignols et de l'acide férulique chez le maïs et les graminées [d'après Riboulet (2007) et Barrière et al (2009)].



études sur la lignification et de la disponibilité du génome maïs de référence (maizesequence.org, release v2). Cette nouvelle liste a été valorisée dans l'ensemble des travaux "génomique" de cette thèse et est présentée dans le chapitre 5. Ainsi, seules les familles de gènes impliqués dans la formation des parois secondaires lignifiées sont présentées dans cette partie "introduction générale".

2.4.1. La voie de biosynthèse des lignines

2.4.1.1. La voie de biosynthèse des monolignols

La biosynthèse des monolignols a lieu en aval de la voie de biosynthèse de l'acide shikimique à partir de la phénylalanine (Riboulet, 2007, Barrière et al, 2007; Barrière et al, 2009b). La plupart des gènes de la voie de biosynthèse des monomères de lignines sont identifiés (Figure 20). La première partie de cette voie de biosynthèse est commune aux phénylpropanoïdes alors que la dernière partie est spécifique aux monolignols. La première enzyme de la voie de biosynthèse des lignines est une phénylalanine ammonia-lyase (PAL) qui permet de dériver la phénylalanine, produite par la voie de biosynthèse de l'acide shikimique, du métabolisme primaire (synthèse des protéines) vers le métabolisme secondaire avec la synthèse des monolignols. La PAL catalyse ainsi la désamination de la phénylalanine en acide cinnamique. La cinnamate 4-hydrolase (C4H) assure ensuite l'hydroxylation de l'acide cinnamique en acide *p*-coumarique. Une activité tyrosine ammonia-lyase (TAL) permet également la synthèse de cet acide *p*-coumarique directement à partir de tyrosine (deuxième départ possible de la voie des monolignols chez le maïs). Une 4-hydroxycinnamate:CoA ligase (4CL) permet le couplage de l'acide *p*-coumarique avec le coenzyme A pour donner le *p*-coumaroyl-CoA. Le *p*-coumaroyl-CoA est hydroxylé en caffeoyl-CoA via l'action de deux enzymes : une hydroxycinnamoyl transférase (HCT) et une *p*-coumarate 3-hydroxylase (C3H). Une caffeoyl-CoA O-méthylTransferase (CCoAOMT) permet la méthylation du caffeoyl-CoA en feruloyl-CoA. Une cinnamoyl-CoA réductase (CCR) réduit ensuite les trois esters cinnamoyl-CoA (*p*-coumaroyl-CoA, caffeoyl-CoA et feruloyl-CoA) en trois aldéhydes (respectivement en *p*-coumaraldehyde, caffeoyl aldehyde et coniferaldehyde). Une O-méthyltransferase (OMT) pourrait également méthyler le caffeoyl-aldehyde en coniferaldehyde, mais cette voie est actuellement inconnue chez le maïs. Le coniferaldehyde est hydroxylé par l'action d'une ferulate-5-hydroxylase (F5H) pour donner du 5-hydroxyconiferaldehyde qui est à son tour méthyler en sinapaldehyde par une acide caféique O-méthyltransferase (COMT). Le *p*-coumaraldehyde, le coniferaldehyde et le sinapaldehyde sont enfin réduits par l'action d'une cinnamyl alcool déshydrogénase (CAD) et/ou d'une sinapyl alcool déshydrogénase (SAD) en trois monolignols : l'alcool *p*-coumarylique, l'alcool coniférylique et l'alcool sinapylique. Le passage de l'alcool coniférylique à l'alcool sinapylique semble possible par l'action d'une F5H et d'une COMT.

Les voies aboutissant aux différentes sous-unités monomériques, H, G et S peuvent être partiellement indépendantes (Dixon et al, 2001; Barrière et al, 2009b). En effet, la réduction de

l'activité PAL, C4H, 4CL, CCoAOMT ou CCR peut aboutir à une variation du ratio S/G dans le polymère de lignines alors que ces familles de gènes sont théoriquement communes à la synthèse des deux sous-unités (Kajita et al, 1997; Sewalt et al, 1997; Piquemal et al, 1998; Ralph et al, 1998; Do et al, 2007). Certains gènes de ces familles multigéniques seraient ainsi spécifiques à la formation d'une sous-unité mais la fonction réelle de chacun de ces homologues est bien souvent obscure. Les voies de biosynthèse des sous-unités G et S sembleraient diverger dès la troisième enzyme de la voie de biosynthèse, l'enzyme 4CL1 chez *Arabidopsis* étant spécifique de la sous-unité G (Lee et al, 1997). Ces deux sous-unités seraient de plus synthétisées par deux familles d'enzymes différentes dans les dernières étapes de synthèse. La biosynthèse des sous-unités G résulterait probablement de la méthylation du cafféoyl-CoA par une CCoAOMT, alors que la biosynthèse des sous-unité S semblerait résulter préférentiellement de la méthylation du cafféoyl-aldéhyde par une OMT inconnue de type COMT ou ZRP4-OMT (Guo et al, 2001b; Parvathi et al, 2001; Chen et al, 2006; Barrière et al, 2007; Do et al, 2007).

2.4.1.2. Le transport et la polymérisation des monolignols en lignines

Les trois monolignols produits sont transportés du cytosol vers la paroi par des transporteurs ABC (Miao et Liu, 2010; Kaneda et al, 2011; Liu, 2012) très probablement sous forme de glucosides (Lanot et al, 2006). La glycosylation des monolignols est effectuée par des enzymes uridine-diphosphate-glucosyltransferases (UGT) spécifiques. Les glucosides sont déglucosylés en monolignol aglycone au sein de la paroi cellulaire par des β -glucosidases spécifiques pour permettre la polymérisation des lignines. Ainsi, deux gènes codant pour des β -glucosidases, β -Glu45 et β -Glu46 sont fortement exprimés dans des organes en cours de lignification chez *Arabidopsis* (Escamilla-Trevino et al, 2006). La polymérisation est ensuite initiée par un mécanisme d'oxydation enzymatique faisant intervenir des peroxydases et/ou des laccases qui génèrent des radicaux se couplant spontanément (Boudet, 2000; Onnerud et al, 2002; Boerjan et al, 2003; Liu, 2012). Il existe de nombreuses peroxydases sécrétables de classe III et laccases chez les végétaux supérieurs (73 peroxydases et 17 laccases chez *Arabidopsis*) et il est difficile d'assigner une fonction spécifique à ces enzymes, en particulier à cause de redondance fonctionnelle soupçonnée. Certaines enzymes spécifiquement impliquées dans la lignification ont été identifiées chez *Arabidopsis* (Ostergaard et al, 2000; Nielsen et al, 2001; Tokunaga et al, 2009; Berthet et al, 2011), chez le peuplier (Christensen et al, 1998; Ranocha et al, 2002; Sasaki et al, 2006), chez le tabac (Blee et al, 2003; Kavousi et al, 2010), chez l'épicéa (Fagerstedt et al, 2010). Chez le maïs, au moins cinq peroxydases et cinq laccases semblent impliquées dans la polymérisation des monolignols (De Obeso et al, 2003; Caparrós-Ruiz et al, 2006; Guillaumie et al, 2007b; Andersen et al, 2009; Barrière et al, 2009a), dont notamment *ZmPox2* et *ZmPox3* exprimées dans les tissus en cours de lignification (De Obeso et al, 2003). De façon complémentaire, les ARNm de *ZmPox2* ont été localisés plus précisément au niveau des tissus vasculaires et épidermiques (De Obeso et al, 2003) et une

insertion d'un élément transposable MITE dans le gène *ZmPox3* a été associée à la digestibilité de la paroi chez les lignées cornées précoces (Guillet-Claude et al, 2004).

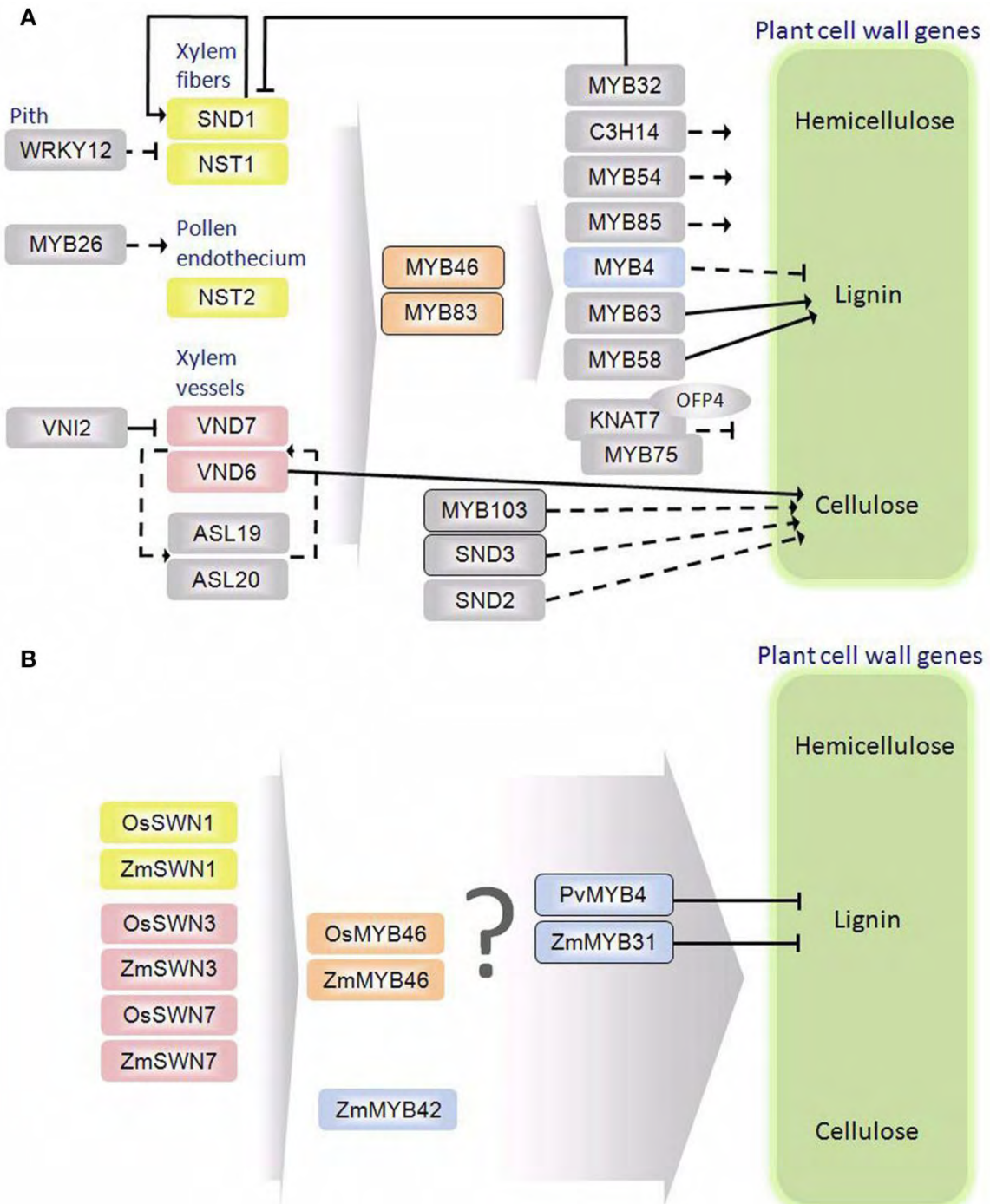
2.4.2. La voie de biosynthèse des acides hydroxycinnamiques

La biosynthèse de l'acide *p*-coumarique est bien connue, cet acide est en effet le deuxième composé produit à partir de la phénylalanine, par l'action d'une enzyme C4H. Par contre, les enzymes impliquées dans sa liaison à la sous-unité S et la formation du S-pCA avant son incorporation au polymère de lignines (Grabber et Lu, 2007) restent inconnues.

La voie de biosynthèse de l'acide férulique (Riboulet, 2007; Barrière et al, 2007; Barrière et al, 2009b), quant à elle, reste mal comprise chez les monocotylédones (Figure 20). D'après différentes études de mutants de Medicago et tabac, la voie de synthèse du FA divergerait avant l'étape CCoAOMT (Chen et al, 2006). Ceci étant, il n'a pas été vérifié que toutes les CCoAOMT aient été effectivement dérégulées dans le mutant CCoAOMT. La synthèse du FA dépend de l'enzyme C3H. L'hypothèse la plus probable actuellement admise chez les dicotylédones est celle mise en évidence chez *Arabidopsis* où une aldéhyde déshydrogénase (ALDH) a été clonée dans le mutant *reduced epidermal fluorescence1* (REF1) qui accumule moins d'ester d'acide sinapique que la plante sauvage. Cette ALDH permettrait ainsi l'oxydation du sinapaldéhyde en acide sinapique (Nair et al, 2004). Des tests *in vitro* ont montré que cette ALDH était également active sur le coniféraldéhyde, permettant alors d'envisager une synthèse d'acide férulique. Finalement, l'acide férulique pourrait être converti en feruloyl-glucose par une glucosyl-transférase (GT). Toutefois, la voie ALDH n'a pas été mise en évidence chez les monocotylédones. Il existe au moins six ALDH chez le maïs (Skibbe et al, 2002; Guillaumie et al, 2007b). Trois sont mitochondriales dont *ZmRF2A* lié à la stérilité mâle cytoplasmique (Skibbe et al, 2002; Liu et Schnable, 2002) et trois sont cytosoliques *ZmREF2C*, *ZmREF2D1* et *ZmREF2D2* et orthologues à REF1. L'expression des ALDH étudiées par Guillaumie et al (2007b) est faible dans des tissus jeunes et matures, à l'exception de l'ALDH AY109842 (GRMZM2G135341), orthologue de l'ALDH22A1 d'*Arabidopsis*, qui montre un niveau d'expression plus élevé. D'autres voies de biosynthèse menant aux feruloyl-arabinoxylanes (feruloyl-AXXX) ont été envisagées soit *via* l'acide cafféoyl-shikimique par l'action successive d'OMT et d'acyltransférase (AcT), enzymes encore inconnues. Il est probable que le feruloyl-glucose parfois mis en évidence soit une forme de stockage et non pas une forme de transfert (Buanafina, 2009).

A ce jour, la voie la plus probable conduisant aux arabinoxylanes féruloylés chez les monocotylédones est celle qui passerait par un féruloyl-CoA. Une activité féruloyl-CoA-arabinoxylane-O-hydroxycinnamoyl trisaccharide transférase (AcT) a ainsi été mise en évidence (Yoshida-Shimokawa et al, 2001) dans des cultures de cellules de riz alimentées en féruloyl-CoA et arabinoxylanes-trisaccharide (AXX) avec la formation de féruloyl-arabinoxylan-trisaccharides

Figure 21 - Régulation transcriptionnelle de la formation de la paroi secondaire chez *Arabidopsis* (A) et chez les monocotylédones (B). Les rectangles représentent des facteurs de transcription. L'ovale indique une protéine interagissant. Les flèches continues et les rectangles entourés indiquent des interactions directes bien établies. Les flèches discontinues indiquent les interactions directes supposées. Les rectangles contenant les gènes orthologues entre *Arabidopsis* et les graminées sont de la même couleur. (Handakumbura et Hazen, 2012).



(FAXX). En outre, des gènes impliqués dans la féruloylation des arabinoxyanes ont été identifiés comme des gènes codant pour des AcT exprimées spécifiquement chez les graminées et non pas chez les dicotylédones, groupe chez lequel cette fonction particulière est censée être absente (Mitchell et al, 2007). L'un des groupes de gènes d'AcT les plus différentiellement exprimés comprenait ainsi des membres de la famille Pfam PF02458 qui codent pour des AcT CoA-dépendantes. Ces gènes sont donc les meilleurs candidats encodant des enzymes catalysant le transfert d'un groupe féruloyl sur une chaîne d'arabinoxyanes dans une paroi de graminées. Cette voie enzymatique n'est peut-être pas la seule. Une réaction entre un féruloyl-CoA et de l'UDP-arabinose a également été envisagée, ce qui donne un intermédiaire Feruloyl-Arabinose-UDP qui pourrait être transféré à la chaîne d'(arabino)xylanes au cours d'une réaction catalysée par une féruloyl-arabinose-UDP transférase (Buanafina, 2009). Des gènes des familles GT43 et GT47 permettraient finalement l'allongement des chaînes de xylanes.

2.4.3. Les facteurs de transcription régulant les voies de biosynthèse et/ou la mise en place des tissus lignifiés

La régulation transcriptionnelle des gènes impliqués dans les voies de biosynthèse des différents composés pariétaux et dans leur mise en place est un champ de recherche encore peu exploré en particulier chez les monocotylédones en dépit des travaux de plus en plus nombreux qui montrent l'importance de son rôle dans le contrôle de la formation la paroi secondaire chez les dicotylédones (Goicoechea et al, 2005; Ramsay et Glover, 2005; Demura et Fukuda, 2007 ; Legay et al, 2007; Zhong et al, 2007; Zhong et al, 2008; Rengel et al, 2009; Zhou et al, 2009; Bhargava et al, 2010; Zhao et al, 2010; Zhong et Ye, 2010; Zhong et al, 2010a; Zhong et al, 2010b; Zhong et Ye, 2012). En effet, tous les tissus présentent des niveaux de lignification différents, ne se lignifient pas en même temps, ni à la même vitesse, indiquant que les voies de biosynthèse des composés phénoliques et l'assemblage des surfaces lignifiées sont finement régulés dans le temps et dans l'espace.

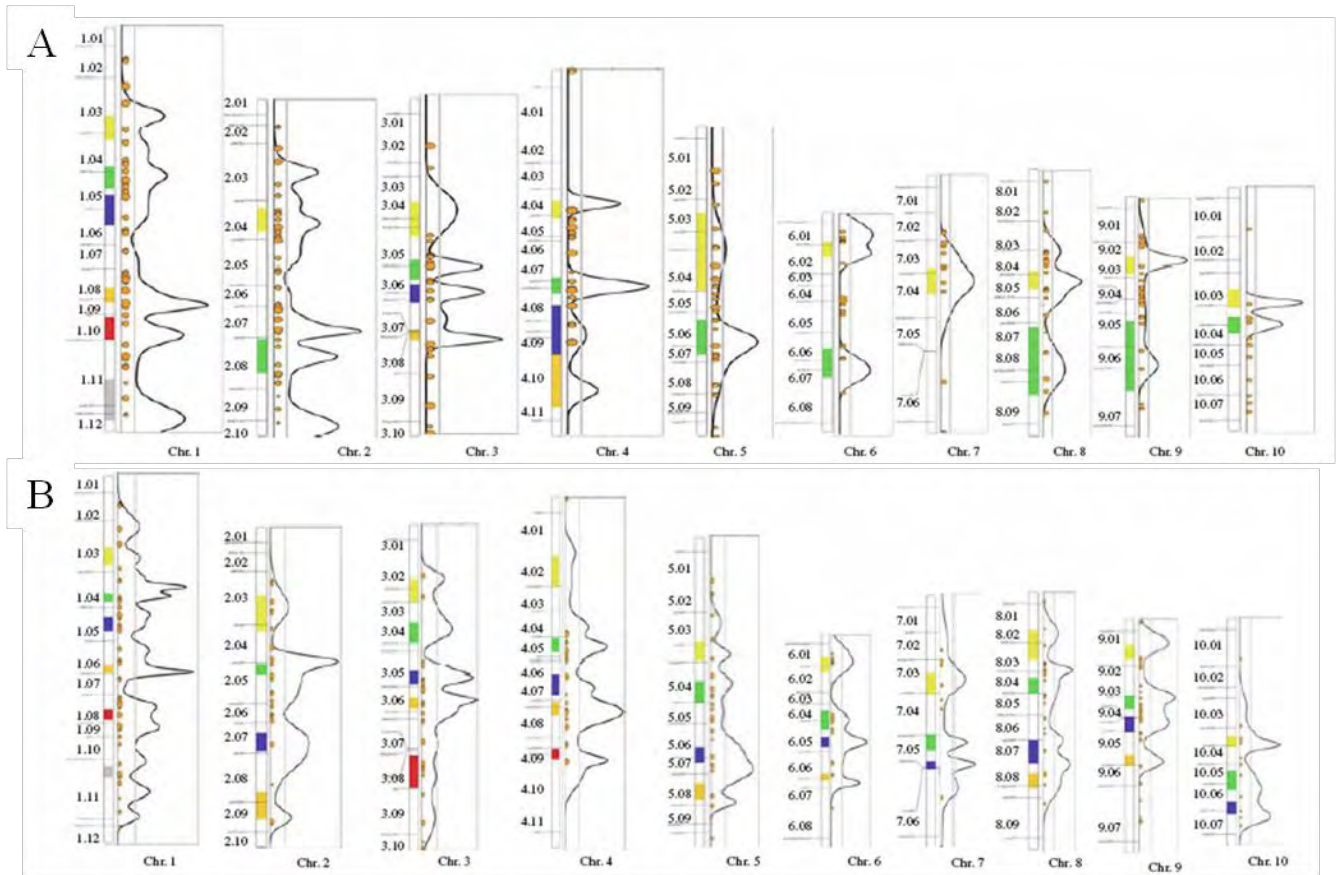
Les premiers gènes régulateurs de la voie de biosynthèse des phénylpropanoïdes sont les facteurs de transcriptions R2R3 MYB (Paz-Ares et al, 1987; Tamagnone et al, 1998). Par la suite, les différentes études ont mis en évidence de nombreux gènes impliqués dans la régulation de la voie de biosynthèse des phénylpropanoïdes et des monolignols et une hiérarchisation de ces derniers. Des revues bibliographiques font le bilan de ces facteurs de transcription (Zhong et Ye, 2009; Zhao et Dixon, 2011, Zhong et al, 2010c; Grima-Pettenati et al, 2012; Handakumbura et Hazen, 2012; Wang et Dixon, 2012) et présentent leur hiérarchisation (Figure 21). Les master gènes correspondraient à des gènes NAC (NST, VND), les facteurs de transcription MYB étant alors des gènes intermédiaires. Chez le maïs, deux gènes MYB, *ZmMYB31* et *ZmMYB42* ont tout d'abord été identifiés comme impliqués dans la régulation de certains gènes de la voie de biosynthèse des monolignols et/ou la mise en place des tissus lignifiés chez le maïs (Fornalé et al, 2006; Sonbol et al, 2009; Fornalé et al, 2010). Quatre

orthologues fonctionnels du master gène *AtSND1* (*ZmSWN2*, 3, 6, 7) et un orthologue fonctionnel d'*AtMYB46* (*ZmMYB46*) ont ensuite été identifiés et sont capables d'activer l'ensemble des programmes de biosynthèse des parois secondaires du maïs (Zhong et al, 2011). De plus, un inventaire complet de la famille des gènes MYB de type R2R3 chez le maïs a été récemment proposé, avec une classification de tous les membres de la famille selon leur séquence et leur classe fonctionnelle (Du et al, 2012). D'autres gènes régulant la voie de biosynthèse des lignines ont été identifiés souvent interagissant en complexe comme le gène *KNAT7* responsable du phénotype irregular xylem IRX (Brown et al, 2005; Zhong et al, 2008; Ko et al, 2009; Li et al, 2012).

De nombreuses familles de gènes sont également impliquées dans la mise en place des tissus lignifiés et dans leur différenciation. La famille des zinc finger est une des plus large de ces familles, avec des zinc-finger C2H2 (Rengel et al, 2009), des zinc finger CCCH appelée aussi C3H14 (Ko et al, 2009; Kim et al, 2012), des zinc finger C3HC4 (Ma et al, 2009), des zinc finger WRKY (Wang et al, 2010), des gènes LIM (Kawaoka et Ebinuma, 2001; Matthews et al, 2009), les gènes *KANADI* (Ilegems et al, 2010) et pour finir les gènes *DOF* (Yanagisawa, 2004; Rogers et al, 2005; Guo et al, 2009; Kushwaha et al, 2011). Des membres de la famille homeodomain-leucine zipper classe III (*HD-ZIP III*) ont également un rôle dans la différenciation des vaisseaux, dont *IFL1* (*Interfascicular Fiber Less 1/ Revoluta*) chez *Arabidopsis* (Zhong et Ye, 1999; Ratcliffe et al, 2000) mais aussi *RLD1* et *RLD2* chez le maïs (Juarez et al, 2004; Ilegems et al, 2010). D'autres gènes sont impliqués dans la mise en place des tissus lignifiés comme *COV1* (*continuous vascular ring*) (Parker et al, 2003), les gènes *bZIP GRAS SCARECROW* (DiLaurenzio et al, 1996; Lee et al, 2008), le gène *EgROP1* (Foucart et al, 2009; Rengel et al, 2009; Kwon et al, 2010), les gènes *ERF/AP2* (*ethylene responsive factor/APETALA2*) (Ambavaram et al, 2011), *SHATTERPROOF MADS-box* (Liljegren et al, 2000; Ko et Han, 2004), *kinesine* (Zhong et al, 2002, Zhang et al, 2010), *katanine* (Burk et Ye, 2002), *COBRA-like* (Li et al, 2003c; Ching et al, 2006; Sindhu et al, 2007; Dai et al, 2011) et les protéines associées aux microtubules dont *AtMAP70* (Pesquet et al, 2010; Pesquet et al, 2011).

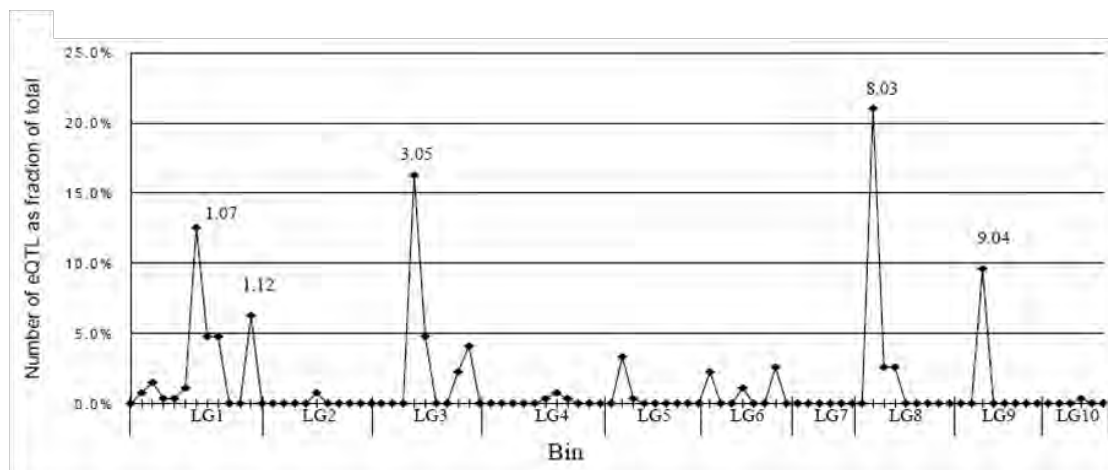
La complexité de la régulation de la formation de la paroi secondaire et de la mise en place des tissus lignifiés est confirmée par le nombre important de gènes candidats énumérés ci-dessus. De plus, de nombreux gènes agissent en complexe comme cela été montré pour la protéine *KNAT7* qui interagit avec les protéines de la famille des *Ovate*, *OFP1* et *OFP4* (Li et al, 2011), pour les protéines *MYB* qui interagissent avec les protéines *basic helix-loop-helix* (*bHLH*) et *WD40* (Ramsay et Glover, 2005), mais aussi pour les protéines *SCARECROW* qui interagissent avec le facteur de transcription *SHORT ROOT* et les protéines *HDZIP III* (Carlsbecker et al, 2010). A cette large liste de gènes s'ajoute les *miRNA*, *ARNs* non codant qui contrôle l'expression des gènes cibles au niveau post-transcriptionnel. Cent cinquante *miRNA* ont été identifiés chez le maïs (Zhang et al, 2006; Zhang et al, 2009). Les travaux sur les *miRNA* sont récents (Yang et al, 2007) et l'implication des *miRNA* dans la formation de la paroi secondaire lignifiée n'est prouvée que pour les familles *miR166*, *miR165* qui ciblent des *HD-ZIP* dont *RLD1* (Juarez et al, 2004; Ko et al, 2006). D'autres familles de *miRNA*

Figure 22 - Répartition des méta-QTL de dégradabilité des parois (A) et de composition des parois (B) sur le génome de maïs (Truntzler et al, 2010).



Legend (Truntzler et al, 2010) : Meta-analysis results for the ten chromosomes: a for digestibility traits and b for cell wall components traits. The bins are displayed on the left side of each chromosome, metaQTL confidence intervals are displayed as thick coloured bars on the chromosomes; the candidate genes are represented as yellow dots with different sizes depending on the class of candidate genes they belong to, class 1 genes being plotted with the bigger-sized dots). QTL overview is shown on the right of each chromosome; the horizontal line corresponds to the average overview U0.5 value (see text).

Figure 23 - Répartition des eQTL (expression Quantitative Trait Loci) de gènes impliqués dans la dégradabilité des parois sur le génome de maïs segmenté en bins (Shi et al, 2007). LG correspond à linkage group ou chromosome.



pourraient être impliquées comme la famille miR171 qui cible des scarecrow et la famille miR164, qui cible des NAC et des laccases, cette famille de gènes étant également la cible des miR397, miR408, miR528.

2.5. Régions génomiques impliquées dans la dégradabilité des parois et dans les caractères associés et gènes candidats

Au vu de la complexité de la dégradabilité des parois et de la formation des parois lignifiées, la génétique quantitative représente un outil adapté pour déchiffrer ces caractères et faciliter l'identification des gènes régulateurs ou structuraux impliqués dans ces caractères. Plusieurs types de familles peuvent être utilisées tel que des descendances F1, F2, mais aussi des familles de lignées recombinantes (RILs, Recombinant Inbred Lines). Les RILs sont produites par autofécondation des descendants du croisement entre deux lignées, ce qui conduit à une descendance de plein-frères. Les RILs sont des outils qui ont été privilégiés pour des études QTLs chez le maïs, en particulier en choisissant des lignées parentales nettement variables pour le caractère ciblé. L'étude de la valeur phénotypique chez les descendants, simultanément à la création d'une carte génétique après génotypage de l'ensemble des RILs, permet la mise en évidence de QTLs (quantitative trait locus, locus génomique associé à un caractère quantitatif).

Chez le maïs, de nombreux QTLs (Quantitative Trait Loci) relatifs aux propriétés des parois (dégradabilité, teneur en lignines, composition monomérique des lignines, teneur en acides hydroxycinnamiques) ont été identifiés dans différents fonds génétiques (Lübberstedt et al, 1997; Ni et al, 1998; Méchin et al, 2001; Roussel et al, 2002; Cardinal et al, 2003; Fontaine et al, 2003; Krakowsky et al, 2003; Krakowsky et al, 2005; Riboulet et al, 2008a; Barrière et al, 2008; Barrière et al, 2010b; Thomas et al, 2010; Barrière et al, 2012). Certains de ces travaux ont fait l'objet de synthèse QTLs (Barrière et al, 2007; Barrière et al, 2009a; Barrière et al, 2009b) et même d'une méta-analyse (Truntzler et al, 2010) révélant que de nombreuses régions génomiques étaient impliquées dans le caractère dégradabilité des parois et les caractères associés de façon indépendante au fond génétique. Selon la méta-analyse, environ 30 régions sont impliquées dans la dégradabilité des parois et 40 régions impliquées dans la composition chimique des parois (Figure 22). Une étude d'eQTLs (QTL d'expression) de gènes impliqués dans la dégradabilité de la paroi cellulaire (Shi et al, 2007) a également permis de mettre en évidence 271 eQTLs, localisés sur 24 bins (Figure 23) pour 89 Expressed Sequences Tag (EST), les variations transcriptionnelles étant à 74 % oligogéniques. Parmi l'ensemble des régions impliquées dans la dégradabilité des parois et dans la composition des parois, certaines régions (bin 1.07, 1.11, 2.08, 3.05, 5.04, 6.05/06, 8.03, 9.04 et 9.06) présentent une colocalisation de nombreux caractères phénotypiques mesurés en valeur propre (évaluation

phénotypique des RILs) ou/et topcross (évaluation phénotypique de la descendance issue des croisements de l'ensemble des RILs avec un testeur). Le bin 6.06 se distingue de ces régions du fait de la colocalisation de QTLs de teneur en lignine chez les deux familles B73 x De811 (Krakowsky et al, 2003; Krakowsky et al, 2005) avec des QTLs majeurs de dégradabilité des parois et de teneur en lignine chez la famille F288 x F271 (Roussel et al, 2002; Thomas et al, 2010), ces derniers QTLs expliquant 20 % à 40 % des variations observées pour ces caractères. La colocalisation des différents caractères pariétaux confirme les liens entre dégradabilité et lignification, dégradabilité et réticulation des parois ainsi que dégradabilité et teneur en acide *p*-coumarique. Par ailleurs, l'absence de QTL de teneur en lignine au niveau de région impliquée dans la dégradabilité permet de valider l'importance des autres caractères pariétaux. Quarante-deux pour cent des metaQTLs de dégradabilité des parois sont chevauchants avec des metaQTLs de composition des parois cellulaires, ce qui est en accord avec les équations de régression permettant de prédire la dégradabilité à partir seulement des caractères biochimiques. Par contre, aucune étude sur les régions génomiques impliquées dans la répartition des tissus lignifiés n'a été faite à ce jour.

Plusieurs régions impliquées dans la dégradabilité des parois et de teneur en lignine ont été mises en évidence chez *Arabidopsis*, espèce modèle des dicotylédones (Chavigneau et al, 2012). Des études QTL sur la qualité du bois et notamment de la teneur en lignines, de la composition en monomère de lignines ont également été menées chez le pin (Sewell et al, 2002; Markussen et al, 2003; Pot et al, 2006), le peuplier (Yin et al, 2010; Ranjan et al, 2010), et l'*Eucalyptus* (Freeman et al, 2009; Thumma et al, 2010; Gion et al, 2011). L'ensemble de ces études conforte l'existence d'un nombre relativement important de loci, et donc de gènes, impliqués dans les variations de dégradabilité des parois du maïs, et donc la nécessité d'identifier pour chaque ressource génétique d'intérêt, les déterminants majeurs à introgresser dans le matériel élite.

Pour l'instant, aucun gène candidat responsable des QTL de dégradabilité des parois et de caractères associés n'a été validé même si des colocalisations entre gènes candidats et QTL ont été recherchées chez le maïs, *Arabidopsis*, le peuplier, l'*Eucalyptus* (Riboulet et al, 2008a; Barrière et al, 2010b; Ranjan et al, 2010; Thomas et al, 2010; Gion et al, 2011; Barrière et al, 2012; Chavigneau et al, 2012).

Ces gènes semblent être majoritairement des gènes régulateurs de la biosynthèse des composés pariétaux ou encore de l'assemblage des tissus lignifiés plutôt que des gènes des voies de biosynthèse. Plusieurs raisons appuient cette hypothèse : (1) de nombreux gènes de la voie des phénylpropanoïdes sont différentiellement exprimés entre des lignées présentant des dégradabilités de paroi différente, un facteur de transcription pouvant réguler l'expression de l'ensemble de ces gènes (2) peu de gènes de la voie de biosynthèse des monolignols colocalisent avec les QTL des caractères pariétaux associés à la dégradabilité et dans l'intervalle support de certains QTL aucun gène de cette voie de biosynthèse n'est présent, (3) les eQTL (QTL d'expression) de gènes impliqués dans la dégradabilité de la paroi

cellulaire détectés par Shi et al (2007), correspondent à des trans eQTL, pouvant être dus à des facteurs de régulation contrôlant l'expression des gènes différentiellement exprimés, (4) enfin, certains QTL de dégradabilité des parois pourraient être liés à des gènes impliqués dans la répartition anatomique et la proportion des tissus lignifiés.

3. Travaux de thèse

Les facteurs majeurs limitant la dégradabilité des parois et donc la valeur énergétique du maïs sont les lignines, essentielles pour la croissance et le développement de la plante, l'acide *p*-coumarique, les réticulations pariétales impliquant l'acide férulique, les liaisons β -O-4 et la répartition anatomique des tissus lignifiés *versus* non lignifiés. L'amélioration de la qualité du fourrage et de la production de biocarburant peut donc être appréhendée d'un point de vue fondamental par une recherche des déterminants génétiques et moléculaires impliqués dans la biosynthèse des constituants de la paroi secondaire, dans leur mise en place et dans la régulation de ces processus. L'essentiel des travaux de thèse sont dédiés à cet objectif et constitue cinq chapitres dont les quatre premiers correspondent à des travaux effectués sur la famille RIL F288 x F271, famille développée à l'INRA de Lusignan pour la recherche de déterminants de la qualité du fourrage et pour laquelle l'identification de QTLs a déjà mis en évidence l'importance de la région du bin 6.06 (Roussel et al, 2002; Thomas et al, 2010). Pour mener à bien la recherche de gènes déterminant la variation de la dégradabilité des parois, la liste de gènes candidats potentiellement impliqués dans la formation des parois secondaires lignifiées (synthèse des constituants pariétaux et mise en place) proposée par Barrière et al (2009a) a été mise à jour tout au long de la thèse en se basant sur le génome de référence maïs B73 (maizesequence.org, release v2) et sur les connaissances bibliographiques chez le maïs mais aussi chez d'autres espèces (*Arabidopsis*, *Eucalyptus*, peuplier, riz...). Cette liste "*a priori*" a été utilisée dans l'ensemble des travaux "génomique" et est présentée dans le chapitre 5. Les cinq chapitres dédiés à la recherche des déterminants génétiques et moléculaires impliqués dans la dégradabilité de parois sont rédigés sous forme de publications et sont présentés ci-dessous avec le plan de la thèse:

Chapitre 1 : Une première étude transcriptomique comparée a été menée avec pour but la mise en évidence des gènes associés à la formation, régulation et/ou assemblage des parois différentiellement exprimés notamment au niveau des QTLs majeurs du bin 6.06. Ce travail s'appuie sur la comparaison du parent défavorable F271 (peu digestible) à quatre RIL ayant d'une part une dégradabilité de parois élevée et une faible teneur en lignine, d'autre part les allèles favorables F288 aux QTLs majeurs. Deux approches ont été suivies (i) une approche "*a priori*" fondée sur la comparaison des expressions des gènes identifiés comme putativement impliqués dans la formation des parois secondaires, (ii) une approche "*sans a priori*" visant à identifier l'ensemble des gènes différentiellement exprimés simultanément entre les quatre RIL et F271 (pour amortir l'effet du fond génétique autre que celui de la région du bin 6.06) et à repérer ceux présents sous les QTLs majeurs.

Chapitre 2 : L'objectif des travaux présentés dans ce chapitre était, grâce à une étape de densification ciblée de la carte génomique, d'affiner la position et l'intervalle support des QTLs détectés chez la famille de 131 RIL F288 x F271, avec les valeurs phénotypiques 1998-2000 afin de pouvoir identifier le ou les gènes candidats responsables de ces QTLs.

Chapitre 3 : Une deuxième détection de QTLs a été menée chez la famille RIL F288 x F271 avec cette fois-ci des données phénotypiques "valeur propre" obtenues au cours des essais 2010-2011. Cette étude avait deux objectifs (i) vérifier et/ou valider les résultats obtenus avec les données phénotypiques 1998-2000 (dégradabilité des parois, teneur en lignines), (ii) rechercher les QTLs pour des caractères de parois influençant la dégradabilité des parois: structure des lignines, teneur en acides *p*-hydroxycinnamiques. Une liste de gènes candidats sous-jacents les principales régions génomiques impliquées dans la dégradabilité et dans les caractères pariétaux associés a été établie à partir de la liste "*a priori*".

Chapitre 4 : Ce chapitre est une suite des chapitres 2 et 3 et a pour objectif final d'identifier le ou les gènes responsables des différents QTLs majeurs colocalisant au bin 6.05 en comparant la séquence génomique des deux lignées parentales F288 et F271 grâce à une approche de séquençage ciblé de BAC (Bacterial Artificial Chromosome) ayant inséré la région d'intérêt.

Chapitre 5 : Ce chapitre fait d'une part, comme cela été dit au dessus, le bilan des gènes potentiellement impliqués dans la formation des parois secondaires lignifiées et d'autre part la synthèse de QTLs de dégradabilité des parois et des caractères de parois associés détectés dans différents fonds génétiques (six familles RILs); l'objectif final étant de rechercher les gènes candidats pouvant expliquer les variations de dégradabilité des parois sous-jacent aux différents QTLs.

Chapitre 6 : Ce chapitre de thèse avait pour but d'évaluer l'impact de la répartition spatiale et quantitative des tissus à parois lignifiées *versus* non lignifiées dans la tige. Pour cela, un protocole pour l'étude histologique des tissus lignifiés a tout d'abord été mis au point afin d'évaluer un grand nombre d'échantillons. La cinétique de mise en place des tissus lignifiés a finalement été étudiée chez les deux lignées parentales F288 et F271.

Enfin, une conclusion générale et des perspectives sont proposées, faisant la synthèse des résultats obtenus et proposant des pistes pour poursuivre ce travail.

CHAPITRE I - APPROCHE TRANSCRIPTOMIQUE DE LA
VARIABILITÉ GÉNÉTIQUE ENTRE QUATRE RILs DE LA
DESCENDANCE RECOMBINANTE DE MAÏS F288 x F271
AYANT UNE DÉGRADABILITÉ DE PAROIS ÉLEVÉE ET LE
PARENT FAIBLE F271

1.1. Avant propos

Les essais conduits à la fin des années 1990 sur les 131 RILs développées à partir du croisement F288 x F271 avaient montré une variabilité de la dégradabilité des parois mesurée par le critère IVNDFD (*In Vitro* Neutral Detergent Fiber Digestibility) comprise entre 24,5 et 39,0 %, les parents F271 et F288 se situant respectivement à 23,3 et 36,9 %. Des résultats similaires avaient été mis en évidence pour le critère DINAGZ [*in vitro* digestibility of the “non-(Starch, soluble carbohydrates and crude protein)” part, selon Argillier et al, 1995], la variabilité des RILs étant comprise entre 45,7 et 55,7 %, et les parents ayant des valeurs de 43,6 et 53,8 %, confirmant des transgressions uniquement pour les valeurs élevées. En plus de ces deux caractères de dégradabilité, les teneurs en lignines ADL/NDF et KL/NDF, [avec NDF (Neutral Detergent Fiber) correspondant à la teneur en parois et ADL (Acid Detergent Lignin) et KL (Klason lignin) correspondant à deux estimations de la teneur en lignine] avaient également été mesurées.

La synthèse de ces essais avait ainsi permis la mise en évidence de QTLs de dégradabilité des parois (IVNDFD et DINAGZ) et de teneurs en lignines (ADL/NDF et KL/NDF), avec notamment des QTLs à effets majeurs dans la région du bin 6.06, à la fois dans les essais en valeur propre, et en combinaison, en croisement avec la lignée F286 qui a une dégradabilité de parois élevée. La variabilité phénotypique expliquée par les QTLs était d'au moins 20 % et atteignait 40 % pour la dégradabilité des parois IVNDFD en valeur propre (Roussel et al, 2002). La mise en évidence du ou des gènes explicatifs de ces QTLs majeurs constitue l'objectif final du travail de thèse. Pour cela, une approche transcriptomique a été entreprise et fait l'objet de ce premier chapitre.

Sur la base du critère IVNDFD, une vingtaine de RILs avaient une valeur de dégradabilité des parois au moins égale ou supérieure à F288. Nous en avons choisi quatre sur la base de leur marquage moléculaire (RIL39, RIL99, RIL54, RIL118, avec des valeurs IVNDFD de 38,4, 38,4, 36,7, et 35,2 %). Ces quatre lignées portaient toutes l'allèle favorable de F288 à la position des QTLs du bin 6.06, et étaient aussi variables que possible aux autres positions du génome non impliquées dans les QTLs. Des études d'expression des gènes ont été réalisées sur les quatre RILs et F271, dans l'entre-noeud sous l'épi aux deux stades "panicule émergeant du cornet" (S1) et "floraison femelle" (S2), qui sont les deux stades ayant une expression élevée des gènes liés à la formation des parois (Riboulet et al, 2009). Ces deux stades correspondent à la fin de l'élongation de cet entre-noeud et à la mise en place de la paroi secondaire lignifiée.

La décision de mise en place de cette expérimentation ayant été prise en 2008 avec l'INRA d'Evry, c'est la puce Affymetrix maïs ne comportant "que" 17.555 sondes qui a été utilisée, cette puce et cette technologie étant alors disponibles avec une grande robustesse d'utilisation. Il n'y a donc qu'un tiers environ du génome maïs représenté, avec de plus un très probable biais favorisant les gènes impliqués

dans la croissance des plantes et le remplissage du grain, correspondant aux EST les plus représentées à cette époque dans les bases de données, la séquence du génome de la lignée B73 n'étant pas encore disponible.

Différentes stratégies de valorisation des données obtenues ont été conduites, avec (1) une approche "*a priori*" fondée sur la comparaison des expressions des gènes identifiés comme putativement impliqués dans la formation des parois secondaires (liste de gènes confortée à partir de Barrière et al, 2009, et aussi chapitre 5), (2) une approche sans *a priori* permettant d'identifier l'ensemble des gènes différentiellement exprimés simultanément dans les quatre RILs par rapport au parent défavorable F271 afin d'amortir l'effet du fond génétique autre que celui du bin 6.06. L'identification des gènes correspondant aux sondes différentiellement exprimées a ensuite permis de distinguer ceux localisés sous l'intervalle support des QTLs majeurs du bin 6.06 de ceux situés ailleurs dans le génome.

Ce travail est détaillé dans la publication « Comparative expression of cell wall related genes in four maize RILs and one parental line of variable lignin content and cell wall degradability » publiée dans la revue *Maydica* (Courtial et al, 2012a).

1.2. Résultats

Une expression différentielle (Bonferroni P-value < 0,05) entre la lignée F271 et simultanément (et dans le même sens) les quatre RILs, a été mise en évidence pour 360 gènes à partir des 17.555 sondes disponibles sur la puce Affymetrix. Alors que 25 % de ces gènes ont une fonction inconnue, 3 % semblent impliqués dans le métabolisme secondaire et/ou celui de la paroi cellulaire et 9% sont liés à la régulation de la transcription. Toutefois, il est difficile de conclure sur l'implication de l'un de ces gènes dans la variation de la dégradabilité et de la teneur en lignines.

Parmi les 360 gènes, neuf gènes sont sous-jacents aux QTLs à effets majeurs du bin 6.06 et 36 aux QTLs du bin 1.01 et 3.05, régions pour lesquels les quatre RILs présentent l'allèle de F288. Parmi les neuf gènes sous-jacents aux QTLs à effets majeurs du bin 6.06, les plus proches des positions des QTLs sont trois gènes qui codent des protéines de fonction inconnue et le gène expansine EXPA5. Le gène FK506 (codant pour une peptidyl-prolyl cis-trans isomérase) est le plus proche de la nouvelle position des QTLs détectés avec la nouvelle carte génétique densifiée en marqueurs (voir chapitre 2), cette nouvelle position étant plus amont. Parmi les gènes ayant une annotation fonctionnelle, les gènes EXPA5 et FK506 sont les candidats les plus probables (ou les moins improbables) au vu de leur fonction. Chez *Arabidopsis*, un orthologue de ZmEXPA5, AtEXPA8 (At2g40610) est aussi situé sous un QTL de dégradabilité de parois (Chavigneau et al, 2012). Pour ce qui est des gènes présents sous

les QTLs des bins 1.01 et 3.05, il est également difficile de conclure sur leur implication dans la variation de la dégradabilité et de la teneur en lignines.

L'approche avec *a priori* a permis de mettre en évidence des gènes différentiellement exprimés non retenus dans l'approche sans *a priori*, en raison des critères statistiques de stringence appliqués. Ainsi l'expression de ces gènes dans une des RILs par rapport à F271 présente une valeur de Bonferroni P-value supérieure à 0,05.

Seuls quatre gènes potentiellement impliqués dans la voie de biosynthèse des monolignols et leur transport sont différentiellement exprimés entre les quatre RILs et la lignée F271. Ces gènes sont : la PAL2b (GRMZM2G118345, bin 2.03), la C3H1 (GRMZM2G138074, bin 3.06), l'OMT ZRP4-like2b SBP1 (GRMZM2G085924, bin 2.05) et une O-glycosyl hydrolase (GRMZM2G119941, bin 2.01). Quatre gènes putativement impliqués dans la polymérisation des monolignols sont eux différentiellement exprimés entre les quatre RILs et la lignée F271. Il s'agit de ZmPox2 (GRMZM2G040638, bin 1.04), ZmPox3 (GRMZM2G135108, bin 6.05), ZmLac Atlac12-like (GRMZM2G132169, bin 3.06) et ZmLac Atlac12-like (GRMZM2G336337, bin 8.08). Il semblerait donc logique de considérer que dans ce fond génétique, les gènes impliqués dans la biosynthèse des monolignols, leur transport et dans une moindre mesure dans leur polymérisation soient que rarement des cibles des facteurs impliqués dans les différences étudiées, au moins pour les stades considérés et pour les sondes présentes sur la puce.

En revanche, cette étude transcriptomique a permis de mettre en évidence trois nouveaux facteurs de transcription MYB (GRMZM2G037650, bin 1.07; GRMZM2G017520, bin 3.04; GRMZM2G041415, bin 3.05) et un NAC (GRMZM2G176677, bin 2.03) très probablement impliqués dans la mise en place des parois secondaires lignifiées du maïs, en plus de ceux précédemment décrits. Ces nouveaux gènes colocalisent d'ailleurs avec des QTLs liés aux parois dans d'autres familles recombinantes. De même, l'expression différentielle du gène codant pour une protéine zinc finger de type C3HC4 (GRMZM2G062724, bin 1.01) et sa colocalisation sous des QTLs conduisent à le considérer comme impliqué dans les variations des caractéristiques pariétales observées, et conforterait sa place sur la liste des gènes impliqués dans la mise en place des parois du maïs. En supposant que ces cinq gènes soient effectivement co-responsables des variations observées, le déterminant génétique premier serait par conséquent un (ou des) gène(s) impliqué(s) dans un mécanisme en amont de la biosynthèse de la lignine et l'assemblage des tissus lignifiés, et qui serai(en)t sans doute localisé(s) dans le bin 6.06.

1.3. Conclusions et perspectives

Cette étude aura permis de mettre en évidence de nombreux gènes différentiellement exprimés entre le parent F271 de mauvaise dégradabilité et quatre RILs de bonne dégradabilité possédant les allèles favorables pour trois régions avec des QTLs impliqués dans ce caractère dont les QTLs majeurs du bin 6.06. Toutefois, sans étude supplémentaire, il est difficile de conclure sur l'implication de l'un de ces gènes dans la variation de la dégradabilité et de la teneur en lignine, même pour les gènes localisés sous les QTLs. La synthèse présentée dans le chapitre 5 montre que cette situation existe également pour d'autres familles ou positions, avec plusieurs QTLs sans aucun gène candidat, ou sans candidats réellement clairs de part leur annotation fonctionnelle. Ceci conduirait à faire l'hypothèse que les gènes candidats recherchés pour la formation des parois, en dehors des MYB, NAC et quelques autres gènes impliqués dans le "tissue-patterning", puissent en fait être des gènes de fonction inconnue, ou bien des gènes ayant une annotation non encore reliée à la lignification. Par ailleurs, cette étude aura permis de révéler trois nouveaux gènes MYB, un gène NAC et un gène zinc finger C3HC4 comme très probablement impliqués dans la mise en place des parois secondaires lignifiées du maïs.

En plus de cette étude transcriptomique, d'autres travaux doivent être entrepris pour identifier le ou les gènes explicatifs des QTL de dégradabilité des parois. En effet, cette étude n'est pas exhaustive du fait de l'utilisation de la puce Affymetrix ne portant des sondes que pour le tiers du génome du maïs, ce qui conduit à ignorer les deux tiers des gènes ayant un intérêt possible, avec une sous représentation des gènes MYB et NAC liés à la régulation de la lignification des parois. De plus, les différences de dégradabilité et de lignification des parois observées entre génotypes ne sont pas forcément liées à des différences d'expression de gènes, mais peuvent être aussi en relation avec des différences de fonctionnalité des protéines par exemple. Le séquençage ciblé de certains gènes, voire de régions du génome, doit ainsi être considéré afin d'identifier d'éventuels polymorphismes responsables de ces variations phénotypiques. De tels travaux ont été entrepris avec d'une part le séquençage ciblé de gènes candidats sous-jacents les QTLs du bin 6.06 pour les deux lignées parentales F288 et F271 (Chapitre 2) et d'autre part la création d'une banque BAC pour ces deux lignées suivie du séquençage ciblé des BACs ayant inséré une partie de la région d'intérêt (Chapitre 4).

Comparative expression of cell wall related genes in four maize RILs and one parental line of variable lignin content and cell wall degradability

Audrey Courtial^{1,2,3}, Cyril Jourda^{1,2,3}, Sandrine Arribat⁴, Sandrine Balzergue⁴, Stéphanie Huguet⁴, Matthieu Reymond⁵, Jacqueline Grima-Pettenati^{2,3}, Yves Barrière^{1*}

¹INRA, Unité de Génétique et d'Amélioration des Plantes Fourragères, 86600 Lusignan, France

²Université de Toulouse, UPS, UMR 5546, Laboratoire de Recherche en Sciences Végétales, BP 42617 Auzeville, F-31326 Castanet-Tolosan, France

³CNRS, UMR 5546, BP 42617, F-31326 Castanet-Tolosan, France

⁴INRA, Unité de Recherche en Génétique Végétale (URGV), UMR 1165 - Université d'Evry Val d'Essonne - ERL CNRS 8196, 2 rue G. Crémieux, CP 5708, F-91057 Evry Cedex, France

⁵INRA, Institut Jean-Pierre Bourgin, 78026 Versailles, France

*Corresponding author: E-mail: yves.barriere@lusignan.inra.fr

Abstract

A comparison of gene expression in maize between the parental line F271 and four RILs derived from the cross F288 x F271 was investigated based on hybridization on the 17,555 probes Affymetrix micro-array, targeting nearly one third of the genes present in maize genomes. The parental line had unfavorable alleles for cell wall degradability traits at the major QTL position in bin 6.06, while the set of RILs had both the favorable allele and high cell wall degradability. 360 genes were differentially expressed in the four RIL in comparison to F271, including nine genes underlying the major QTL position and 36 underlying two other QTL positions. However, their proposed function (whenever is described) do not allow us to firmly consider their involvement in the observed variation of cell wall related traits. Only a few genes involved in monolignol biosynthesis or polymerization located elsewhere in the genome were differentially expressed between the four RILs and F271, corroborating with the fact that these genes are probably not involved in major determinants of cell wall degradability in the studied set of lines. Among the investigated regulation factors, three *ZmMYB*, one *NAC* and one *C3HC4* zinc finger were differentially expressed between the four RILs and F271, but they were not located in bin 6.06. Notwithstanding, the obtained results especially strengthened the probable involvement of these genes in maize secondary wall assembly and/or lignification.

Keywords: maize, degradability, lignin, cell wall, expression

Introduction

Most plant species dedicated to ruminant nutrition and/or proposed as feedstock for renewable energy production belong to the grass clade. Even if a variable part of soluble carbohydrates is still present in harvested grasses, most plant biomass is made up of secondary cell walls which are mostly comprised of cellulose and hemicelluloses embedded in a phenolic component matrix. This matrix is constituted of lignins and *p*-hydroxycinnamic acids which contribute to the plant standability and confer hydrophobicity to the vascular elements allowing transport of water and nutrients. Constitutive and neo-synthesized stress-lignins are involved in defense responses against diseases and pests. On the other hand, lignification is also the basic cause of reduced digestibility of forages in animal digestive tracks and of reduced conversion of plant biomass into bioethanol. Because of the negative impact of lignins for the pulp and paper industry, a sector of considerable economic importance, most of the efforts in genetics, genomics,

and biochemistry have been directed towards a better understanding of lignin biosynthesis in perennial and annual dicotyledonous including the Arabidopsis model system as well as in gymnosperm species. Conversely, fewer studies have been carried on in grasses. Deciphering the genetic determinants of cell wall biosynthesis and assembly in grasses is, thus, a strategic issue in breeding plants for both cattle feeding and second generation biofuel production.

Due to the large genetic resources characterized for cell wall related traits, the ease of self-pollination and crossing, the availability of genomic tools (ie. EST, transgenesis, transposon-tagging), and the recent release of its genomic sequence (Schnable et al, 2009), maize is both a plant of direct high economical interest and a model system for all grasses. QTL analyses for lignin and cell wall related traits have been investigated in several maize RIL progenies (Méchin et al, 2001; Roussel et al, 2002; Krakowsky et al, 2003; Cardinal et al, 2003; Krakowsky et al, 2005; Barrière et al, 2008; Riboulet et al, 2008a; Barrière et al, 2010;

Barrière et al, 2012). Even if co-localizations between candidate genes and QTLs have been pointed out in maize, but also in arabidopsis, poplar, and eucalyptus (Barrière et al, 2010; Ranjan et al, 2010; Thomas et al, 2010; Gion et al, 2011; Chavigneau et al, 2012), no gene has yet been shown to be responsible for the effect of any cell wall-related QTL. The latter situation is due in part to our insufficient knowledge of the major genetic determinants of cell wall biosynthesis, assembly and regulation, even if mutants, including brown-midrib mutants in maize, and genetically modified plants in the lignin pathway have been studied, and candidate genes listed in both maize and Arabidopsis (He et al, 2003; Barrière et al, 2004; Ko and Han, 2004; Minic et al, 2009; Pichon et al, 2006; Barrière et al, 2009; Tamasloukht et al, 2011; Chavigneau et al, 2012). The size of the QTL support intervals, which have an average length of nearly 20 cM corresponding in maize to 15 to 50 Mbp, and up to 500 genes, depending on their genomic location and recombination rate, is another reason for making candidate gene identification and validation difficult.

In the F288 x F271 early maize RIL progeny, several QTLs for cell wall related traits, including lignin content and cell wall degradability, were already mapped (Roussel et al, 2002; Thomas et al, 2010). Among the latter QTLs, those located in bin 6.06 explained a highly significant part of investigated trait variations, with R^2 values ranging from 20 to 39% for investigated traits. In order to contribute to the identification of candidate genes underlying QTLs, a large-scale gene expression study was carried out, comparing the parental line (F271) with unfavorable alleles for cell wall degradability traits in bin 6.06 and a set of four RILs with both favorable phenotypic and allelic values. Because the objective of the expression study was to get insights into the discovery of genes involved in cell wall variation for degradability and lignin content, mostly in relation with QTLs in bin 6.06, two strategies of gene expression study were simultaneously considered. Gene expressions between the F271 parental line and the set of RILs were investigated with a targeted hypothesis-driven strategy, considering genes putatively involved in cell wall biosynthesis and assembly, and regulation of these processes according to the list proposed by Barrière et al (2009). In addition, a screen for genes giving a significant differential expression between F271 and each of the RILs was considered, first focusing on those located in bin 6.06 QTL support intervals, and then on those elsewhere with a genome-wise approach.

Materials and Methods

RIL progeny development and experiments

A set of 131 RILs was initially developed by single seed descent from the cross between the two early dent inbred lines F288 and F271, which are of medium-high and low cell wall degradability, respec-

tively (Barrière et al, 2001). RIL progenies were then evaluated in field experiments for both their per se and topcross values, with F286 as flint tester of high cell wall degradability (Roussel et al, 2002). QTLs were detected for cell wall related traits, including lignin content (ADL/NDF and KL/NDF) and cell wall digestibility (IVNDFD and DINAGZ). NDF (Neutral Detergent Fiber) is an estimate of cell wall content and ADL (Acid Detergent Lignin) is an estimate of lignin content according to Goering and van Soest (1970). KL (Klason lignin) is gravimetric method used to assess the lignin content according to Dence and Lin (1992) including the acido-soluble part of lignin which is lost during the first step of the ADL procedure (Hatfield et al, 1994; Jung et al, 1997; Hatfield and Fukushima, 2005). IVNDFD (in vitro NDF digestibility) is an estimate of cell wall digestibility, based on the enzymatic solubility of Aufrère and Michalet-Doreau (1983), which is calculated according to Struik (1983) and Dolstra and Medema (1990), assuming that the non-NDF part is fully digestible. DINAGZ is another estimate of cell wall digestibility, similarly based on the same enzymatic solubility, which is estimated according to Argillier et al (1995) and Barrière et al (2003), assuming that starch, crude proteins, and soluble carbohydrates are fully digestible. Major QTLs were, thus, shown for these four cell wall traits in bin 6.06 (Table 1).

Line choice

The parental line F271, with alleles giving high lignin content and low cell wall degradability values at all QTL positions except in bin 9.02, was chosen as the "negative" control. Four RILs (RIL39, RIL54, RIL99, and RIL118) were chosen, firstly, for their parental F288 genotypes in bin 6.06, based on three markers, from positions 147.0 to 166.3 Mbp which overlapped QTL support intervals (Figure 1). The RIL39 differed from the three other RILs because it has F271 parental allele downstream QTL positions, while RIL99, RIL54, and RIL118 have F288 allele up until phi089 marker (position 166.3 Mbp). RILs were also chosen for their phenotypic values with high cell wall degradability and low lignin contents (Table 2). The four RILs had also the favorable alleles (increasing cell wall degradability) at QTL positions in bins

Table 1 - Position (cM), Lod value, explained phenotypic variation (R^2 , %), and additive value (add) of putative QTLs identified in bin 6.06 for ADL/NDF and KL/NDF lignin content, and DINAGZ and IVNDFD cell wall digestibility [131 RIL of the F288 x F271 progeny in per se experiments, from Roussel et al (2002) and Thomas et al (2010), traits as percentage].

QTL bin 6.06	Position	Lod	R^2	add value
ADL/NDF	142	6.2	19.5	0.30
KL/NDF	142	9.0	27.0	0.66
DINAGZ	142	8.4	25.6	1.26
IVNDFD	142	14.3	39.4	2.50

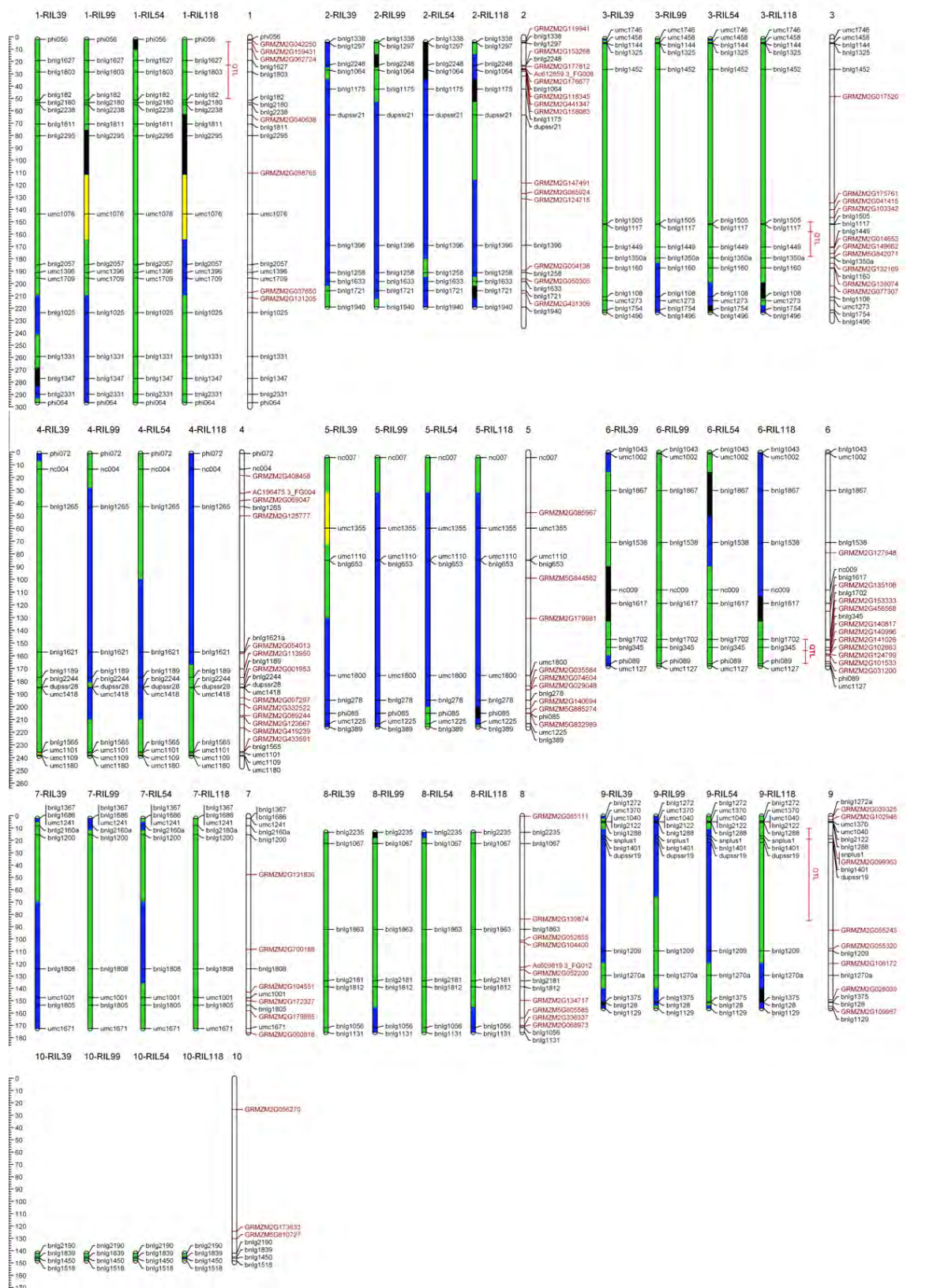


Figure 1 - Physical map of the genome of the four RILs (colour bars) and of the QTLs for cell wall related trait in the RIL F288 x F271 progeny. Candidate genes are positioned on the right bars (white bars). In green= F288 allele, blue= F271 allele, black= unknown allele and yellow= heterozygote.

Table 2 - Phenotypic values in the parental lines and the set of selected RILs for ADL/NDF and KL/NDF lignin content, and DINAGZ and IVNDFD cell wall digestibility [131 RIL of the F288 x F271 progeny in per se experiments, from Roussel et al (2002)].

	F288	F271	RIL39	RIL99	RIL54	RIL118
ADL/NDF	5.4	7.6	5.2	5.6	5.9	5.8
KL/NDF	13.0	16.1	12.3	12.8	13.1	13.9
DINAGZ	53.8	43.6	54.7	54.3	54.0	53.8
IVNDFD	36.9	23.3	38.4	38.4	36.7	35.2

1.02 and 3.05, while favorable alleles were not present in RIL118 in bin 9.02.

Field experiments and sampling for gene expression studies

Selected RILs and the parental line F271 were cropped at INRA Lusignan (Vienne, France) in 2009 in a block design with two four-row replicates. Row spacing was 0.75 m and the density was thinned to 90,000 plants/ha. Irrigation was applied during summer to prevent water stress.

Because [Guillaumie et al \(2007a, 2008\)](#) and [Riboulet et al \(2009\)](#) have shown that the genes of the lignin pathway were highly expressed in the below-ear internode in a period from tassel emergence (S1) to a few days after silking (S2), corresponding to the end of the elongation phase and the deposition of the secondary cell wall in this internode, below-ear internodes (without nodes) of five representative plants in each of the two replicates were, thus, harvested at this two stages S1 and S2. Sampling was done five hours after sunrise, for each of the four investigated RILs and the parental line F271. To limit experiment costs, biological repetitions of each RIL were pooled after cutting the internode in fragments of nearly 1 cm long. All samples were immediately frozen in liquid nitrogen and stored at -80°C.

RNA extraction

The internodes were crushed in liquid nitrogen using an Ika Mill crusher (IKA, Staufen, Germany) prior to RNA extraction. Total RNAs were extracted using the Nucleospin® RNA II kit (Macherey-Nagel). Some modifications of the protocol were added, including i) 600 µl of buffer RA1 were used, instead of 450 µl, and 6 µl of β-mercaptoethanol, instead of 3.5 µl, for the cell lyse, ii) 600 µl of 70% ethanol were used for the DNA binding, instead of 350 µl, and iii) finally two washes at 600 µl of buffer RA3, instead of one. RNA samples were checked for their integrity with the Agilent 2100 bioanalyzer according to the Agilent technologies (Waldbronn, Germany).

Affymetrix micro-array hybridization, statistical analysis, and gene expression investigations

One µg of total RNA was used to synthesize biotin-labeled cRNAs with the One-cycle cDNA synthesis kit (Affymetrix, Santa Clara, CA). Superscript II reverse transcriptase and T7-oligo (dT) primers were used to synthesize the single strand of cDNA at 42°C during 1 hour, followed by the synthesis of the double stranded cDNA by using DNA ligase, DNA polymerase I and RNaseH during 2 hours at 16°C.

Clean up of the double-stranded cDNA was performed with Sample Cleanup Module (Affymetrix) followed by in vitro transcription (IVT) in the presence of biotin-labeled UTP using GeneChip® IVT labelling Kit (Affymetrix). Quantity of the labelled-cRNA with RiboGreen® RNA Quantification Reagent (Turner Biosystems, Sunnyvale, CA) was determined after cleanup by the Sample Cleanup Module (Affymetrix). Fragmentation of 10µg of labelled-cRNA was carried out for 35 minutes at 94°C, followed by hybridization during 16 hours at 45°C to Affymetrix GeneChip® Maize Genome Array 18K representing approximately 14,850 maize transcripts, corresponding to 13,339 genes. After hybridization, the arrays were washed with 2 different buffers (stringent: 6x SSPE, 0.01% Tween-20 and non-stringent: 100mM MES, 0.1M [Na⁺], 0.01% Tween-20) and stained with a complex solution including Streptavidin R-Phycoerythrin conjugate (Invitrogen/molecular probes, Carlsbad, CA) and anti-Streptavidin biotinylated antibody (Vectors laboratories, Burlingame, CA). The washing and staining steps were performed in a GeneChip® Fluidics Station 450 (Affymetrix). The Affymetrix GeneChip® Maize Genome Arrays 18K were finally scanned with the GeneChip® Scanner 3000 7G piloted by the Command Console Launcher Tool.

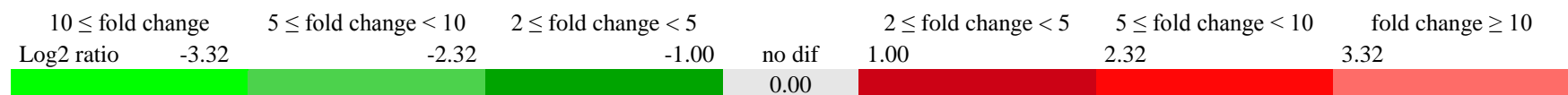
The data were normalized (Log2 values) with the gcrma algorithm ([Irizarry et al, 2003](#)), available in the Bioconductor package ([Gentleman et al, 2002](#)). To determine differentially expressed genes, a usual two group t-test, that assumes equal variance between groups, was performed. The variance of the gene expression per group is a homoscedastic variance, where genes displaying extremes of variance (too small or too large) were excluded. The raw P values were adjusted by the Bonferroni method, which controls the Family Wise Error Rate (FWER, [Ge et al, 2003](#)). A gene was declared differentially expressed if the Bonferroni P-value was less than 0.05. All this steps were performed on Affymetrix platform at IN-RA-URGV, Evry.

The physical positions of differentially expressed genes were considered according to the B73 genomic sequence ([Schnable et al, 2009](#), www.maize-esequence.org, release v2 5b60) allowing testing of their colocalizations with QTLs, especially with QTLs in bin 6.06.

Data Deposition

The raw.CEL files were imported in R software for data analysis. All raw and normalized data are avail-

S_tables 1 - Gene expression, in below-ear internodes at two growing stages, S1 = emerging tassel stage and S2 = silking stage. GRMZM numbers are given according to the release 5b.60 of the www.maizesequence.org database. Range of colors according to the level of differential expression between F271 and RILs, and similarly for S_tables 1A to 1F, a log₂ sum intensity in green or in red indicates that the gene is less expressed or more expressed in RILs, respectively, other log₂ sum intensities (not in green or red) were not found to be statistically significant after Bonferroni correction (P < 0.05).



S_table 1A - Expression of genes involved in monolignol biosynthesis in F271 and four RILs, in below-ear internodes, at two stages.

Gene	GRMZM gene	Bin	Emerging tassel stage (S1)					Silking stage (S2)				
			F271	RIL39	RIL99	RIL54	RIL118	F271	RIL39	RIL99	RIL54	RIL118
PAL2a	GRMZM2G441347	2.03	14.37	13.92	14.32	14.18	14.24	13.64	13.22	13.61	13.58	14.18
PAL2b	GRMZM2G118345	2.03	7.54	10.64	12.95	10.77	8.20	5.16	7.37	10.20	8.10	6.56
PAL3a (ZmPAL)	GRMZM2G074604	5.05	13.84	13.62	13.79	13.59	13.80	13.27	13.54	13.67	13.40	13.73
PAL3b	GRMZM2G029048	5.05	12.54	12.18	12.94	12.49	12.24	11.65	11.31	11.76	11.64	11.95
4CL4	GRMZM5G805585	8.06	4.33	4.78	4.57	5.11	5.09	4.59	4.71	4.69	4.44	4.58
4CL5	GRMZM2G055320	9.04	10.61	10.35	11.18	10.65	10.28	8.74	8.59	8.03	9.20	9.20
4CL6	GRMZM2G054013	4.06	2.47	2.63	3.58	2.90	2.91	2.45	2.44	2.49	2.43	2.40
C4H1	GRMZM2G139874	8.03	11.82	10.33	11.23	11.47	11.09	10.68	10.35	10.51	10.48	10.82
C3'H1	GRMZM2G138074	3.06	8.83	5.99	8.32	4.46	5.00	7.25	5.26	8.15	3.86	4.48
C3'H2	GRMZM2G140817	6.06	11.51	10.02	11.61	11.00	10.85	9.90	8.50	9.59	9.53	10.43
HCT1	GRMZM2G035584	5.05	13.30	12.38	13.36	13.05	12.85	11.54	11.26	11.40	11.70	12.13
HCT2	GRMZM2G158083	2.04	10.86	10.37	10.84	10.59	10.58	10.34	9.77	10.88	10.30	10.80
CCoAOMT1	GRMZM2G127948	6.01	13.35	12.92	13.85	13.17	12.95	12.48	11.99	12.69	12.32	12.63
CCoAOMT2	GRMZM2G099363	9.02	14.22	13.36	14.32	13.67	13.45	13.76	13.06	13.61	13.33	13.07
CCoAOMT3	GRMZM2G004138	2.07	9.18	8.40	8.12	6.89	8.22	10.14	9.72	9.74	8.18	8.89
CCoAOMT5	GRMZM2G332522	4.08	12.31	11.68	12.83	12.26	11.68	10.10	9.56	9.71	9.72	11.11
ZmCCR1	GRMZM2G131205	1.07	13.67	13.38	13.83	13.67	13.70	13.19	13.02	13.28	13.49	13.78
ZmCCR2	GRMZM2G131836	7.02	4.99	4.54	7.04	4.87	5.30	3.24	3.04	2.87	3.01	2.82
COMT	AC196475.3_FG004	4.04	14.50	14.20	14.32	14.30	14.45	14.36	14.03	14.15	13.95	14.27
ZmCAD1	GRMZM2G179981	5.04	11.35	11.03	11.28	11.33	10.85	12.01	11.12	11.45	11.41	11.04
ZmCAD2	GRMZM5G844562	5.04	13.84	13.63	14.02	13.51	13.88	13.14	12.75	13.46	12.96	13.49
ZmCAD10	GRMZM2G700188	7.02	8.74	10.47	8.71	10.44	9.43	7.58	8.99	7.50	8.45	9.03

able through the CATdb database (AFFY_ZeaWall-RIL_Maize & AFFY_ZeaWall-Ril2_MAIZE, [Gagnot et al, 2008](#)) and from the Gene Expression Omnibus (GEO) repository at the National Center for Biotechnology Information (NCBI), accession number GSE22439 & GSE 27070 ([Barrett et al, 2007](#)).

Results

Global comparative analysis of gene expression between F271 and four RILs

A differential expression between F271 and all of the four RILs was only shown for 338 probes out of the 17,555 probes available on the Affymetrix microarray ([Table 3](#)); these probes corresponding to 360 genes (including probes hybridizing two close paralogs). Most of the genes differentially expressed at stage S1 were also differentially expressed at stage S2. Differentially expressed genes were classified according to the MapMan BIN classification ([Thimm et al, 2004](#)). Twenty five percent of the genes were of unknown function, while nearly 2% had a function related to cell wall metabolism, 1% belongs to the secondary metabolism functional class, and 9% had a RNA related function including regulation of transcription ([Figure 2](#)).

Comparative expression of genes related to monolignol biosynthesis and polymerization

Monolignol biosynthesis genes

Expression data were available for 22 genes belonging to nine multigene families involved in monolignol biosynthesis ([Supplementary table 1A](#)). No probes or data were available neither for F5H (ferulate-5-hydroxylase), and nor for several members in different families. Only a few differences in transcript levels were observed between the four RILs and the parental line F271. The *PAL2b* (phenylalanine-ammonia-lyase) gene (GRMZM2G118345) was over-expressed in all RILs at stage S2 and in three RILs out of four at stage S1 as compared to the parental line.

Even if an under-expression of PAL genes should have been expected in RILs that are less lignified than F271, a higher involvement of the *PAL2b* gene could induce favorable differences in lignification patterns between the two plant types. However PAL enzymes are not specific of the lignin pathway since they catalyze the first step of the phenylpropanoid pathway which leads to many phenylpropanoids compounds other than lignins. The *C3'H1* (cinnamate-3-hydroxylase) gene (GRMZM2G138074) was under-expressed in RILs except RIL99, which had the F271 allele in this position ([Figure 1](#)). There was also a tendency towards an over-expression of the *CAD10* (cinnamyl-alcohol-dehydrogenase). This gene is a *ZmCAD2-type* gene, but its involvement in constitutive lignification has not yet been established. Finally, the *CCoAOMT3* showed a little lower expression in RIL54.

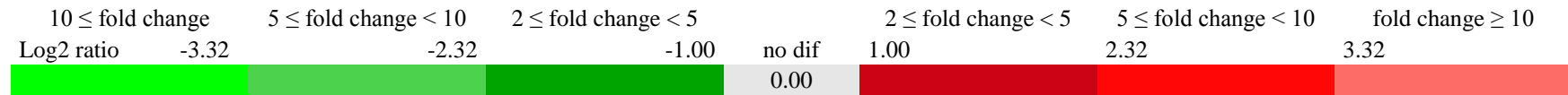
Monolignol related genes

Based on the B73 sequence ([Schnable et al, 2009](#)), 12 *ZRP4-like OMT* (*Zea Root Preferential O-methyl-transferase*) are present in the maize genome, of which its role is likely not limited to the methylation of suberin subunit precursors in plant roots as it has been initially described by [Held et al \(1993\)](#). It was hypothesized that member(s) of the family could be involved in methylation of caffeoyl aldehyde in a route preferentially oriented towards the biosynthesis of sinapyl alcohol ([Barrière et al, 2009](#)). The *ZRP4-like2b OMT* (GRMZM2G085924, *SBP1*, [Scott-Craig et al, 1998](#)) was over-expressed in the four RILs at stage S1 and in two RILs at stage S2, while it was under-expressed in one RIL at this stage ([Supplementary table 1B](#)). Simultaneously, a lower expression was shown, but only at stage S2, for the *ZRP4-like5c/5d OMT* gene(s), the corresponding probe targeting the two close and closely located genes GRMZM2G102863 (position 158.76 Mbp) and GRMZM2G124799 (position 158.84 Mbp). Monolignols are transported from the cytosol to the cell wall, very likely as monolignol glucosides ([Lanot et al, 2006](#)), after their glucosylation by specific uridine-diphosphate-glucosyltransferases

Table 3 - Number of probes giving significant differential expression between F271 and all of the four RILs [Expression data from below-ear internodes, at emerging tassel stage (S1) and silking stage (S2)].

	Stage S1 (only)	Stage S2 (only)	Stages S1 and S2
F271 > all 4 RIL	8	133	46
fold change \geq 20	1	7	17
10 \leq fold change < 20	2	16	13
5 \leq fold change < 10	4	41	14
2 \leq fold change < 5	1	69	2
All 4 RIL > F271	24	77	50
fold change \geq 20	1	1	25
10 \leq fold change < 20	6	10	11
5 \leq fold change < 10	14	26	14
2 \leq fold change < 5	3	40	0
Total	32	210	96

S_tables 1 - Gene expression, in below-ear internodes at two growing stages, S1 = emerging tassel stage and S2 = silking stage. GRMZM numbers are given according to the release 5b.60 of the www.maizesequence.org database. Range of colors according to the level of differential expression between F271 and RILs, and similarly for S_tables 1A to 1F, a log2 sum intensity in green or in red indicates that the gene is less expressed or more expressed in RILs, respectively, other log2 sum intensities (not in green or red) were not found to be statistically significant after Bonferroni correction ($P < 0.05$).



S_table 1B - Expression of monolignol related genes and monolignol polymerization genes in F271 and four RILs, in below-ear internodes, at two stages.

Gene	GRMZM gene	bin	Emerging tassel stage (S1)					Silking stage (S2)				
			F271	RIL39	RIL99	RIL54	RIL118	F271	RIL39	RIL99	RIL54	RIL118
OMT ZRP4-like1	GRMZM2G408458	4.03	6.12	6.13	6.21	6.23	6.12	6.13	5.96	6.14	6.11	6.13
OMT ZRP4-like2b SBP1	GRMZM2G085924	2.05	9.56	11.85	11.11	12.68	11.34	8.79	7.79	10.86	12.81	8.71
OMT ZRP4-like2c	GRMZM2G147491	2.05	2.21	2.25	2.27	2.23	2.19	2.22	2.28	2.24	2.26	2.23
OMT ZRP4-like3b	GRMZM2G106172	9.04	2.10	2.13	2.14	2.12	2.09	2.10	2.15	2.12	2.14	2.12
OMT ZRP4-like4a	GRMZM2G097297	4.08	14.10	11.51	13.65	13.30	12.48	14.26	11.52	13.33	13.90	11.97
OMT ZRP4-like5a	GRMZM2G140996	6.06	6.81	8.86	7.50	8.24	5.93	6.39	7.34	6.01	7.05	5.74
OMT ZRP4-like5b	GRMZM2G141026	6.06	2.12	2.98	2.20	2.18	2.10	2.12	2.32	2.13	2.15	2.13
OMT ZRP4-like5c/5d	GRMZM2G102863*	6.06	3.07	2.80	3.40	2.77	2.11	4.57	2.83	2.74	3.14	2.35
β-Glucosidase45/46-like	GRMZM5G810727	10.05	9.36	9.24	9.19	9.22	9.22	9.82	9.47	9.56	9.75	8.87
O-glycosyl hydrolase	GRMZM2G119941	2.01	6.93	2.13	2.13	2.12	2.43	4.54	2.14	2.12	2.13	2.12
ABC transporter	GRMZM2G177812	2.03	4.07	4.13	4.00	4.63	4.01	4.76	3.87	4.44	3.82	6.04
ABC transporter	GRMZM2G085111	8.00	9.64	9.08	9.28	9.75	9.81	10.32	9.50	10.29	10.42	9.34
ZmPox2	GRMZM2G040638	1.04	2.54	2.54	5.75	3.91	5.81	2.53	2.61	2.58	2.58	3.88
ZmPox3	GRMZM2G135108	6.05	5.72	8.13	8.70	7.78	6.74	4.88	4.88	3.85	4.46	5.60
ZmPox39	GRMZM2G085967	5.03	12.90	13.06	12.76	13.29	13.18	10.76	11.90	10.31	11.85	11.62
ZmPox54	GRMZM2G088765	1.05	3.24	5.77	2.91	6.70	3.31	2.19	2.79	2.11	2.23	2.11
ZmLac1	GRMZM5G842071	3.06	8.16	7.55	7.14	7.07	7.99	8.69	9.11	8.86	8.18	8.89
ZmLac Atlac12-like	GRMZM2G132169	3.06	5.43	8.03	8.55	8.40	8.18	2.86	2.85	2.70	2.78	2.83
ZmLac Atlac12-like	GRMZM2G336337	8.08	2.54	4.42	5.74	4.60	5.41	2.07	2.08	2.05	2.07	2.06

* The corresponding probe also targeted the closely located GRMZM2G124799 gene.

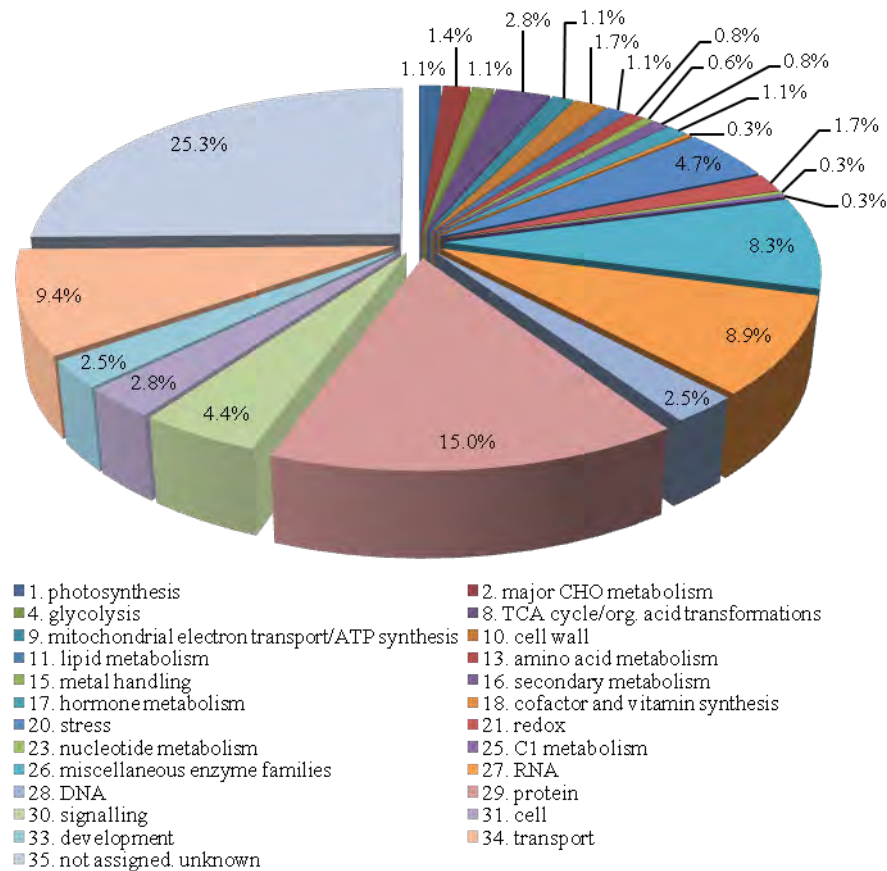


Figure 2 - Functional classification of genes differentially expressed between F271 and the four RILs (classification according to MapMan BIN, percentage of gene in each class).

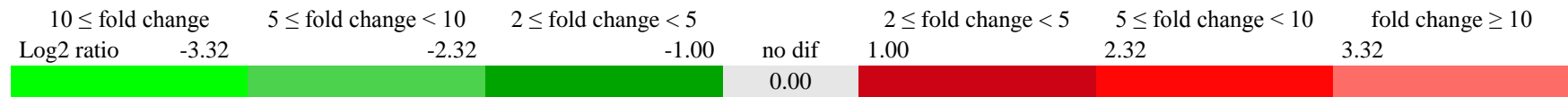
(UGT, no available probes). The release of monolignol aglycone from its glucosidic form at the cell wall for subsequent lignin polymerization is thought to be mediated by specific glucosidases, such as the two *Arabidopsis* β -Glu45 and β -Glu46 (β -glucosidases), of which encoding genes are strongly expressed in lignifying organs (Escamilla-Trevino et al, 2006). The only investigated maize ortholog was not differentially expressed between RILs and F271. However, an *Arabidopsis* ortholog annotated *O-glycosyl hydrolase* (GRMZM2G119941) was dramatically less expressed in all RILs than in F271. In addition, the two investigated maize ABC transporters, encoding proteins putatively involved in the transport of monolignols across membranes (Miao et al, 2010; Kaneda et al, 2011), were not differentially expressed between RILs and F271.

Monolignol polymerization

While peroxidases have long been considered as the unique class of oxidases involved in lignin polymerization, dehydrogenative polymerization of monolignols is, in fact, driven by both peroxidases and laccases (Boudet, 2000; Onnerud et al, 2002; Boerjan et al, 2003; Tokunaga et al, 2009; Fagerstedt et al, 2010; Berthet et al, 2011). This has also been shown in maize

(de Obeso et al, 2003; Caparros-Ruiz et al, 2006; Andersen et al, 2009). Peroxidases and laccases belong to large multigene families and it has been difficult to assign a specific function to any particular oxidase, especially because functional redundancy was often suspected. In *Arabidopsis*, only laccase double mutants clearly exhibited a modified lignin phenotype (Berthet et al, 2011). Nevertheless, the importance of oxidase redundancy could be partly reduced by the fact that many peroxidases or laccases have tissue specific expression patterns. When considering genes expressed in maize vascular and lignifying tissues (de Obeso et al, 2003; Caparros-Ruiz et al, 2006; Guillaumie et al, 2007b; Riboulet et al, 2009), at least five peroxidases and five laccases could be involved in monolignol polymerization. However, these latter numbers of members are likely under-estimated (Barrière et al, 2009). Among the four peroxidases for which expression data were available (Supplementary table 1B), the *ZmPox3* gene (GRMZM2G135108) was more expressed in all four RILs than in F271 at stage S1. In a partly contradictory way, a MITE insertion disrupting this gene was associated with higher cell wall degradability in a set of related European flint lines (Guillet-Claude et al, 2004). A similar tendency

S_tables 1 - Gene expression, in below-ear internodes at two growing stages, S1 = emerging tassel stage and S2 = silking stage. GRMZM numbers are given according to the release 5b.60 of the www.maizesequence.org database. Range of colors according to the level of differential expression between F271 and RILs, and similarly for S_tables 1A to 1F, a log2 sum intensity in green or in red indicates that the gene is less expressed or more expressed in RILs, respectively, other log2 sum intensities (not in green or red) were not found to be statistically significant after Bonferroni correction ($P < 0.05$).



S_table 1C - Expression of genes involved in the regulation of lignin biosynthesis in F271 and four RILs, in below-ear internodes, at two stages.

Gene	GRMZM gene	bin	Emerging tassel stage (S1)					Silking stage (S2)				
			F271	RIL39	RIL99	RIL54	RIL118	F271	RIL39	RIL99	RIL54	RIL118
ZmMYB AtMYB85-like	GRMZM2G037650	1.07	3.86	1.93	2.32	1.92	1.92	1.99	1.95	1.93	1.93	1.98
ZmMYB AtMYB85-like	GRMZM2G104551	7.03	11.79	10.97	11.98	11.62	11.45	10.38	10.88	11.25	11.28	12.24
ZmMYB AtMYB46-like	GRMZM2G172327	7.03	1.81	1.83	1.84	1.82	1.80	1.82	1.85	1.82	1.84	1.82
ZmMYB AtMYB103-like	GRMZM2G173633	10.04	7.47	6.23	4.41	6.82	6.86	8.00	7.28	5.50	7.52	7.45
ZmMYB AtMYB4-like	GRMZM2G124715	2.05	6.28	6.58	7.29	4.01	5.05	3.26	3.04	3.47	2.32	4.62
ZmMYB8 AtMYB4-like	GRMZM2G041415	3.05	5.96	2.41	5.08	2.36	3.17	4.14	2.27	2.42	2.22	2.47
ZmMYB AtMYB4-like	GRMZM2G089244	4.09	3.10	2.14	2.91	2.12	2.10	2.37	2.15	2.12	2.13	2.14
ZmMYB AtMYB4-like hv5	GRMZM2G000818	7.04	7.42	7.62	8.99	7.87	7.56	6.31	5.80	6.96	6.14	7.11
ZmMYB31 AtMYB4-like hv5	GRMZM2G050305	2.07	4.98	6.29	5.09	5.53	5.27	4.88	6.37	5.70	5.51	6.84
ZmMYB42 AtMYB4-like hv5	GRMZM2G419239	4.09	5.03	6.41	4.07	6.20	4.01	7.88	7.27	8.89	7.82	6.61
ZmMYB2 AtMYB86-like	GRMZM2G017520	3.04	4.17	10.65	10.15	9.74	9.66	3.51	8.77	8.44	8.19	8.21
ZmNAC ANAC082/VNI1-like	GRMZM2G176677	2.03	5.71	6.52	5.38	6.83	6.35	5.46	7.73	7.03	6.97	6.57
ZmNAC ANAC082/VNI1-like	GRMZM2G125777	4.04	5.31	3.68	5.17	3.66	5.21	5.90	4.30	5.32	4.27	5.08
ZmNAC ANAC103-like	GRMZM2G456568	6.05	9.47	9.38	10.98	11.14	11.37	9.74	9.48	10.87	10.86	11.31
ZmNAC ANAC103-like	GRMZM2G104400	8.03	11.86	11.70	11.21	11.24	11.79	11.85	11.61	11.51	11.43	11.63
ZmNAC ANAC025-like	GRMZM2G179885	7.04	2.45	2.50	2.54	2.49	2.41	6.97	7.70	3.03	5.89	3.84
ZmNAC	GRMZM2G068973	8.08	10.05	10.31	8.97	11.59	10.25	10.32	10.28	10.73	10.59	10.59
ZmNAC ANAC032-like	GRMZM2G123667	4.09	9.39	8.91	8.00	8.59	9.02	9.43	8.09	9.07	9.71	8.12
ZmNAC48 ANAC032-like	GRMZM2G014653	3.06	9.25	9.33	8.16	9.18	9.75	10.68	10.42	10.79	10.67	9.91
ZmNAC ANAC053-like	GRMZM2G113950	4.06	9.95	9.16	9.70	9.80	9.72	9.70	9.16	9.55	9.59	9.64
ZmNAC019 ANAC073/SND2-like	GRMZM2G134717	8.05	2.97	2.45	2.60	2.44	2.42	2.43	2.47	2.43	2.45	2.43
ZmNAC123 ANAC073/SND2-like	GRMZM2G031200	6.07	5.16	5.29	6.25	4.69	5.58	3.08	3.14	3.34	2.55	3.53
ZmNAC143 = ZmSWN2	GRMZM2G069047	4.05	2.17	2.19	2.20	4.96	2.71	2.16	2.19	2.18	2.18	2.19
ZmNAC168/186 = ZmSWN4	AC212859.3_FG008	2.03	2.81	2.22	5.29	2.42	2.20	2.19	2.23	2.30	2.17	2.28
KNAT7-At1g62990-like	GRMZM2G159431	1.01	10.70	10.64	11.70	11.10	11.12	9.31	9.33	9.91	9.96	10.35
KNAT7-At1g62990-like	GRMZM2G433591	4.09	8.15	9.02	7.93	8.45	8.96	8.93	8.93	8.86	8.90	9.02
KNAT7-At1g62990-like	GRMZM2G055243	9.03	10.31	10.30	10.00	9.98	10.47	10.70	10.92	11.08	10.86	10.64

towards a higher expression was also observed, except in RIL39, for *ZmPox2* (GRMZM2G040638), which was considered as involved in maize vascular vessels and epidermis lignification, at a greater extent than *ZmPox3* (de Obeso et al, 2003). None of the four maize laccases described by Caparros-Ruiz et al (2006) were present on the Affymetrix array. However, two of three genes encoding for other laccases were more expressed in RILs than in F271 at stage S1. These latter correspond to two closely related genes (GRMZM2G132169 and GRMZM2G336337), with 93% sequence identity, which are orthologous to *AtLac12* (At5g05390). *AtLac12* is a gene expressed in Arabidopsis stems, mostly in the medium part, at a lower extent than *AtLac4* (At2g38080) and *AtLac17* (At5g60020) and (Berthet et al, 2011). Higher transcript levels of some peroxidases and laccases were found in lines exhibiting lower lignin content, mostly at emerging tassel S1 stage, possibly indicating a variable way of lignin polymerization between the selected RILs and F271. In addition, other peroxidases or laccases not present on the array could also have a determinant effect on lignin polymerization and structure.

Comparative expression of genes involved in regulation of lignin biosynthesis

MYB transcription factors

The regulation of phenylpropanoid and lignin biosynthesis was the first role identified for a plant R2R3-MYB transcription factor (Paz-Ares et al, 1987; Tamagnone et al, 1998). Later, comprehensive investigations have used the Arabidopsis model system, but also woody plants, documenting the hierarchical network including MYB regulating the secondary wall biosynthetic program (Goicoechea et al, 2005; Legay et al, 2007; Zhong et al, 2008; Wilkins et al, 2009; Zhong and Ye, 2009; Legay et al, 2010). In addition, MYB genes can be either transcriptional activators of lignin gene expression as was shown for *EgMYB2* (Goicoechea et al, 2005), *AtMYB46*, *AtMYB58*, *AtMYB85*, and *AtMYB103*, or transcriptional repressors as was shown for *EgMYB1* (Legay et al, 2007; Legay et al, 2010), *AtMYB4*, and *AtMYB86*. In maize, the two *ZmMYB40* and *ZmMYB95* duplicated genes (or *ZmMYB-IF35* and *ZmMYB-IF25*) were shown to be involved in phenylpropanoid metabolism (Dias and Grotewold, 2003; Heine et al, 2007), but their involvement in regulation of monolignol biosynthesis is not yet established. Conversely, *ZmMYB31* and *ZmMYB42*, orthologous to barley *MYB-hv5* genes associated with vascular bundles (Wissenbach et al, 1993), were proven to be involved in the regulation of lignin pathway genes, with a repressive effect (Fornalé et al, 2006; Sonbol et al, 2009; Fornalé et al, 2010; Gray et al, 2012). Finally, the *ZmSWMYB46* (GRMZM2G052606) gene orthologous to *AtMYB46* was recently identified (Zhong et al, 2011). Several other MYB factors are surely involved in the regulation of maize lignification (Barrière et al, 2009, and

unpublished data). Expression data were available for only a few of the *ZmMYB* genes putatively involved in maize cell wall lignification (Supplementary table 1C). Corroborating the probable involvement of other *ZmMYB* genes in secondary wall assembly in maize, the *ZmMYB* (GRMZM2G037650) located in bin 1.07, which is a *EgMYB2*-type and thus a putative positive regulator of transcription, was less expressed in RILs than in F271 at stage S1. Complementarily, the *ZmMYB* (GRMZM2G017520) located in bin 3.04, which is a *EgMYB1*-type and thus a putative negative regulator of transcription, was more expressed in all four RILs at the two stages. However, and conversely, the *ZmMYB* (GRMZM2G041415) located in bin 3.05, ortholog of negative regulators of gene transcription, was less expressed in three RIL at stage S1, and in all RILs at stage S2.

NAC transcription factors

Lignin pathway genes are regulated upstream of the MYB transcription factor level by NAC transcription factors. A set of NAC domain transcription factors (SND1, NST1, VND6, VND7) has, thus, been shown to act as a positive master regulator of secondary cell wall biosynthesis in Arabidopsis (Yamaguchi et al, 2008; Zhong et al, 2008; Zhong and Ye, 2009; Zhou et al, 2009; Zhong et al, 2010). Orthologs of these NAC are present in other plant species, but no exhaustive information is available in maize. Several maize orthologs are, nevertheless, available from Barrière et al (2009), Shen et al (2009), and Zhong et al (2011) data, even if some of these NAC are possibly involved in other pathways than cell wall assembly. Based on data available for 13 NAC genes putatively involved in regulation of cell wall lignification, no differences in expression were clearly shown between RILs and F271 (Supplementary table 1C). There was, nevertheless, a tendency to a higher expression in the four RILs for one NAC (GRMZM2G176677) located in bin 2.03 at stage S2, and also in three RILs at the two stages for one NAC (GRMZM2G456568) located in bin 6.05. A lower expression was also shown for one NAC located in bin 7.04 in three RILs at stage S2.

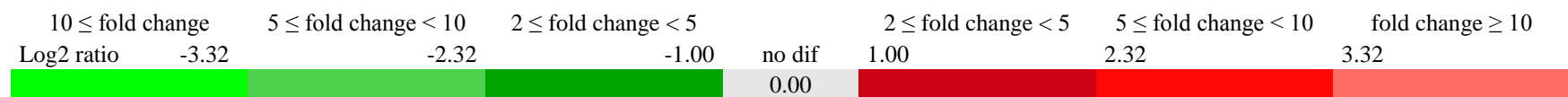
KNAT7 orthologous genes

KNAT7 was shown to be a transcriptional repressor of secondary wall biosynthesis, involved in "a negative feedback loop that maintains cellular homeostasis during developmental commitment to secondary cell wall deposition" (Li et al, 2012). However, none of the *KNAT7* orthologs for which data were available were differentially expressed between RILs and F271 (Supplementary table 1C).

Comparative expression of genes involved in regulation of lignified tissue assembly

Genes involved in the regulation of lignified tissue assembly are located upstream in pathways potentially regulating the plant final lignin content or cell wall degradability. It is, thus, questionable whether or not these genes could be good candidate genes explaining the effect of the detected QTLs. In fact, their

S_tables 1 - Gene expression, in below-ear internodes at two growing stages, S1 = emerging tassel stage and S2 = silking stage. GRMZM numbers are given according to the release 5b.60 of the www.maizesequence.org database. Range of colors according to the level of differential expression between F271 and RILs, and similarly for S_tables 1A to 1F, a log₂ sum intensity in green or in red indicates that the gene is less expressed or more expressed in RILs, respectively, other log₂ sum intensities (not in green or red) were not found to be statistically significant after Bonferroni correction (P < 0.05).



S_table 1D - Expression of genes involved in regulation of lignified tissue assembly in F271 and four RILs, in below-ear internodes, at two stages.

Gene	GRMZM gene	bin	Emerging tassel stage (S1)					Silking stage (S2)				
			F271	RIL39	RIL99	RIL54	RIL118	F271	RIL39	RIL99	RIL54	RIL118
Zinc finger HDZIP IFL1-RLD1	GRMZM2G109987	9.07	9.28	9.45	9.70	9.11	8.97	9.89	9.71	10.32	10.07	9.42
Zinc finger HDZIP IFL1-RLD2	GRMZM2G042250	1.01	11.14	10.33	10.94	10.76	10.72	11.10	10.36	10.70	10.98	10.24
Zinc finger C3HC4-type	GRMZM2G062724	1.01	4.63	6.17	6.55	8.25	7.39	4.36	5.21	5.89	5.57	8.40
Zinc finger C3HC4-type	GRMZM2G077307	3.07	11.38	6.80	6.47	11.13	8.33	11.07	7.80	7.37	11.44	8.75
Zinc finger C3HC4-type	GRMZM2G056270	10.03	9.34	9.30	8.90	8.95	9.73	10.95	10.39	11.07	10.72	10.15
Zinc finger DOF-type HCA2-like	GRMZM2G140694	5.06	3.34	3.22	3.24	3.18	3.16	3.79	3.20	3.41	2.82	3.19
Zinc finger Lim-like	GRMZM2G153268	2.03	2.01	2.04	2.05	2.03	2.01	2.02	2.06	2.03	2.05	2.03
Zinc finger Lim-like	GRMZM2G175761	3.05	7.16	7.23	6.54	6.16	6.19	6.18	6.25	5.07	5.52	6.63
bZIP GRAS SCARECROW-like	GRMZM2G431309	2.07	10.93	11.31	10.81	10.00	11.27	9.23	10.72	10.00	10.54	11.63
bZIP GRAS SCARECROW-like	GRMZM5G885274	5.06	11.33	10.66	11.05	11.41	11.26	10.86	10.79	11.02	11.09	10.62
bZIP GRAS SCARECROW-like	GRMZM2G153333	6.05	8.52	8.69	8.74	8.95	9.05	7.94	8.32	9.01	8.72	8.94
bZIP GRAS SCARECROW-like	GRMZM2G028039	9.06	11.33	10.66	11.05	11.41	11.26	10.86	10.79	11.02	11.09	10.62
COV1-like	GRMZM2G101533	6.06	9.98	9.43	9.44	9.57	9.97	10.48	9.89	10.02	10.07	9.86
COV1-like	GRMZM2G052200	8.05	5.21	5.00	4.60	5.48	4.86	5.25	5.97	5.99	6.20	6.24
COV LCV2-like	GRMZM2G149662	3.06	9.42	9.51	10.67	9.58	9.64	8.40	8.99	8.85	9.13	8.71
COV LCV2-like	GRMZM2G052855	8.03	7.54	8.40	7.84	7.59	7.26	5.51	6.66	5.09	5.36	6.13
ROP AtROP3 EgROP1	GRMZM2G001953	4.07	6.29	7.01	7.19	7.76	7.12	5.76	6.75	6.61	6.53	6.91
ROP AtROP3 EgROP1	AC209819.3_FG012	8.04	2.02	2.32	2.06	2.13	2.34	2.02	2.06	2.04	2.05	2.04
ROP AtROP3 EgROP1	GRMZM2G102946	9.00	7.84	7.69	8.99	6.99	7.56	6.05	6.89	6.34	5.82	7.07
AtMAP70-1-like	GRMZM5G832989	5.06	7.58	8.51	7.63	8.82	9.09	5.72	7.34	6.01	7.40	8.79
AtMAP70-1-like	GRMZM2G039325	9.00	7.79	8.78	8.13	5.97	7.80	7.07	7.69	6.94	6.63	7.51

involvement could be considered if observed differences in lignin content or cell wall degradability are related to variation in areas and patterning of lignified tissues. Such large differences were shown between lines, with lines of which parenchyma is not lignified, or with large differences in lignified tissues surrounding vascular bundles, or large variation in width of the lignified cortex (Méchin et al, 2005; Riboulet et al, 2008b).

Zinc finger encoding genes

Zinc finger proteins, which constitute one of the largest families of transcription factor regulatory proteins, are involved in the regulation of numerous elementary processes of plant development, including lignified tissue assembly. Among zinc-finger, members of the small class III homeodomain leucine-zipper (HD-ZIP) proteins have been shown to play a regulatory role in vascular differentiation, as illustrated by the investigations on the *IFL1* (*Interfascicular Fiber Less1*, synonymous *Revoluta*) genes of Arabidopsis or *Rolled Leaf* (*RLD1* and *RLD2*) genes in maize (Zhong and Ye, 1999; Ratcliffe et al, 2000; Juarez et al, 2004). The *RLD1* gene is regulated by the miR166 (Juarez et al, 2004), and the expression of the aspen *PtaHB1* gene, orthologous to *IFL1* and *RLD1*, which was closely associated with wood formation, was inversely correlated with the level of miR166 miRNA (Ko et al, 2006). In addition, interactions between *HDZIPIII* and *KANADI* gene family members were shown to be involved in the establishment of the spatial arrangement of phloem, cambium and xylem. It was considered that *HDZIPIII* and *KANADI* transcription factors control cambium activity, with *KANADI* proteins acting on auxin transport, and *HDZIPIII* proteins promoting axial cell elongation and xylem differentiation (Ilegems et al, 2010). Because there is no secondary cambium in grasses, the latter *HDZIPIII* and *KANADI* genes, obviously, have different functions than in Arabidopsis or aspen. Notwithstanding, the rolled-leaf mutations have consequences in lignified tissue patterning, even if no difference in expression of these genes was observed between F271 and any of the four RILs (Supplementary table 1D).

While the CCCH zinc finger protein (AtC3H14) may function as master regulator of secondary wall biosynthesis and activate all of the secondary wall phenolics and carbohydrate related genes tested (Ko et al, 2009), no clear involvement of C3HC4 RING finger proteins in regulation of lignin-related metabolism has been reported so far. RING finger proteins are involved in numerous cellular processes including transcription regulation, signal transduction, protein-protein interaction and ubiquitination (Ma et al, 2009). However, the *C3HC4 zinc finger* (GRMZM2G062724), located in bin 1.01, was more highly expressed in the four RILs than in F271, especially at the earlier S1 stage. In addition, the *C3HC4 zinc finger* GRMZM2G077307, located in bin 3.07, was under-expressed in three RIL.

Members of the LIM protein family contain zinc-binding domains. Unlike the classical zinc fingers, these domains do not bind DNA, but mediate interactions with other proteins (Matthews et al, 2009). The tobacco *NtLIM1* gene was shown to be a positive regulator of the lignin pathway (Kawaoka and Ebinuma, 2001). The closest maize ortholog of *NtLIM1* (GRMZM2G004959) has no probe on the Affymetrix array. Expression data were only available for two paralogs. The most expressed LIM gene is annotated in the maize sequence database as encoding the “pollen specific protein SF3” which was, in fact, clearly expressed in stalks. No differences in LIM gene expression were shown between RILs and F271.

bZIP GRAS SCARECROW-like genes

The GRAS SCARECROW and SCARECROW-like proteins belong to a plant specific transcription factor family which contains basic leucine zipper regions (Di Lorenzo et al, 1996; Lee et al, 2008). These proteins are involved in complex regulatory pathways regulating tissue patterning and differentiation. The SCARECROW protein is involved in bidirectional cell signaling mediated by miRNA165/6 and interfering with the transcription factor SHORT ROOT (SHR, equally expressed in stem and root) and class III homeodomain leucine zipper proteins towards the control of xylem patterning (Carlsbecker et al, 2010). Out of the four maize orthologs with available expression, the one located in bin 2.07 (GRMZM2G431309) was more expressed in three RILs, while the one located in bin 6.05 was more expressed in two RIL, at stage S2 for the two genes.

COV-like genes

The *COV1* (*CONTINUOUS VASCULAR RING*) recessive mutant of Arabidopsis has a great increase in stem vascular tissue at the inter-fascicular regions (Parker et al, 2003). In addition to the *COV1* gene (At2g20120), two *COV1-like* genes (*LCV2*, *Like-COV-2*, At1g43130, and *LCV3*, *Like-COV-3*, At2g18460) are present in the Arabidopsis genome. Ten genes encoding potentially for COV orthologs were found in the maize genome, but probes and/or data were available for only four genes, without differences in expression between RILs and F271.

EgROP1-like genes

A member of the plant ROP family (*EgROP1*) was shown to be preferentially expressed in the cambial zone and differentiating xylem of eucalyptus. Its over-expression in Arabidopsis altered vessel formation and fiber growth in secondary xylem (Rengel et al, 2009; Foucart et al, 2009). None of the three maize orthologs of *EgROP1* with available probe data was differentially expressed between RILs and F271.

Microtubule-Associated Proteins

Two microtubule-associated proteins (AtMAP70-5 and AtMAP70) were shown in Arabidopsis to be essential for defining where secondary cell wall polymers are applied at the cell cortex in wood-forming cells (Pesquet et al, 2010). There was higher expres-

sion of one maize ortholog (GRMZM2G832989), located in bin 5.06, in three RILs compared with F271, especially at the later S2 stage.

Differentially expressed genes underlying the major QTL position (bin 6.06)

As was previously quoted, only two (or three) genes, with a possible role in secondary wall biosynthesis, were differentially expressed between RILs and F271. The NAC transcription factor located in bin 6.05 could be an upstream regulator of cell wall assembly, but its higher expression concerned only three RILs out of four, a fact that did not convincingly designate it as a good candidate gene for the studied QTLs. A higher expression was also shown, but only at stage S2, for the *OMT ZRP4-like5c/5d* gene(s). Even if the *OMT-ZRP4* genes have been considered with a possible involvement in methylation of aromatic ring during cell wall phenolic compounds biosynthesis (Barrière et al, 2009), a differential expression only at stage S2 is not the most probable situation for a candidate gene with major effect. In addition, nine genes located in bin 6.06 QTL support intervals were differentially expressed between RILs and F271, including four genes of unknown function (Table 4).

Alpha-expansin gene

The α -expansin 5 (*EXPA5*), has a functional annotation which is also related to cell wall metabolism. Expansins appear to be involved in the disruption of hydrogen bonds between cellulose microfibrils and cross-linking hemicelluloses in the wall, restoring the long-term extension to cell walls (Li et al, 2003). While the rice genome contains at least 28 α -expansin genes (Shin et al, 2005), the whole number of α -expansin genes is not yet known in maize. Ten genes are annotated alpha-expansin in the maize sequence database, nine of them belonging to a group of 38 paralogs, allowing a probable redundancy in α -expansin activities. However, the

EXPA5 gene (GRMZM2G361064) located in bin 6.06 (position 154.73 Mbp) appears different from the others as it is the only one without any paralogs. This *ZmEXPA5* gene was expressed in region of leaves where elongation has ceased and where secondary cell wall deposition occurred (Muller et al, 2007). In addition, roles of expansins that do not involve wall expansion have already been shown. Expansins have been associated with the growth of protoxylem elements in Zinnia stems (Im et al, 2000) and expansins similarly appeared to be expressed during the differentiation of the tracheary elements, suggesting their possible involvement in secondary wall formation or/and primary wall disassembly (Milionni et al, 2001). The *EXPA5* gene was, at least, expressed 20 times more in RILs than in F271 at stage S1 and S2, except in two RILs at stage S2 with an only three times greater expression of this gene.

FKBP gene

FK506-binding proteins (FKBPs) belong to the large family of peptidyl-prolyl *cis-trans* isomerases, which are known to be involved in many cellular processes, such as cell signaling, protein trafficking and transcription (Harrar et al, 2001). A putative ortholog, GRMZM2G035922, in maize genome located in position 151.5 Mbp was less expressed in RILs than in F271, especially at stage S2. One gene of this family, *Pasticcino1 (PAS1)*, was shown to play an important role in the control of plant development, controlling cell division or differentiation. The PAS mutants of *Arabidopsis thaliana* show ectopic cell proliferation in cotyledons, extra layers of cells in the hypocotyl, and an abnormal apical meristem (Vittorioso et al, 1998).

Seed maturation protein reticulon (RTN) family

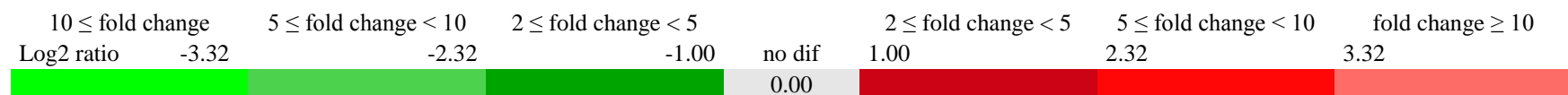
The reticulon family gathers a large group of membrane-associated proteins found throughout the eukaryotic kingdom. Reticulons principally localize to the endoplasmic reticulum, and they are involved in

Table 4 - Differentially expressed genes between F271 and four RILs underlying the major QTLs located in bin 6.06, in below-ear internodes, at two growing stages, S1 = emerging tassel stage and S2 = silking stage. GRMZM numbers are given according to the release 5b.60 of the www.maizesequence.org database. Range of colors according to the level of differential expression between F271 and RILs, a log2 sum intensity in green or in red indicates that the gene is less expressed or more expressed in RILs, respectively, other log2 sum intensities (not in green or red) were not found to be statistically significant after Bonferroni correction (P < 0.05).

Gene	GRMZM gene	bin	Pos. Mpb	Emerging tassel stage (S1)				Silking stage (S2)					
				F271	RIL39	RIL99	RIL54	RIL118	F271	RIL39	RIL99	RIL54	RIL118
FK506 binding protein	GRMZM2G035922	6.05	151.49	9.95	9.02	8.91	8.44	8.92	10.52	8.94	9.13	8.86	9.03
Unknown function	GRMZM2G125037	6.05	153.76	3.13	6.87	5.16	5.49	5.35	3.08	3.96	3.77	4.23	3.70
Alpha-expansin 5	GRMZM2G361064	6.06	154.13	3.49	7.07	8.58	8.97	7.92	2.61	7.18	8.46	7.21	8.84
Unknown function	GRMZM2G037111	6.06	154.99	6.36	3.41	3.12	3.22	3.12	6.59	4.70	3.22	4.59	3.20
Unknown function	GRMZM2G129166	6.06	155.07	4.52	8.71	9.41	9.11	10.00	3.73	8.44	8.98	8.85	9.16
Unknown function	GRMZM2G315199	6.06	156.12	7.33	7.26	7.50	8.34	8.33	5.86	8.13	7.64	7.66	7.73
Heat shock 60kDa protein	GRMZM2G074790	6.07	161.73	2.86	2.37	2.46	2.36	2.28	4.07	3.39	3.35	2.56	2.14
Seed maturation protein	GRMZM2G004828	6.07	162.64	7.49	2.25	4.03	4.38	3.21	7.04	2.17	2.17	2.16	2.16
Non-specific lipid-transfer protein	GRMZM2G320373	6.07	162.66	9.18	9.33	9.68	9.24	9.36	8.96	7.29	5.81	5.36	7.27



S_tables 1 - Gene expression, in below-ear internodes at two growing stages, S1 = emerging tassel stage and S2 = silking stage. GRMZM numbers are given according to the release 5b.60 of the www.maizesequence.org database. Range of colors according to the level of differential expression between F271 and RILs, and similarly for S_tables 1A to 1F, a log₂ sum intensity in green or in red indicates that the gene is less expressed or more expressed in RILs, respectively, other log₂ sum intensities (not in green or red) were not found to be statistically significant after Bonferroni correction (P < 0.05).



S_table 1E - Differentially expressed genes between F271 and four RILs underlying QTLs located in bin 1.02, in below-ear internodes, at two stages.

Gene	GRMZM gene	bin	Pos. Mpb	Emerging tassel stage (S1)					Silking stage (S2)				
				F271	RIL39	RIL99	RIL54	RIL118	F271	RIL39	RIL99	RIL54	RIL118
Hydroxymethylglutaryl-CoA synthase	GRMZM2G087207	1.01	4.02	3.25	5.51	7.44	6.21	6.94	2.35	3.48	3.84	4.69	4.80
Serine/threonine-protein kinase	GRMZM2G176604	1.01	5.41	5.55	3.17	4.39	3.67	3.47	4.31	2.53	2.62	2.64	2.71
Serine/threonine-protein kinase	GRMZM2G014366	1.01	6.37	5.55	3.17	4.39	3.67	3.47	4.31	2.53	2.62	2.64	2.71
Flower-specific γ -thionin	GRMZM2G392863	1.01	6.81	9.72	7.26	9.32	9.05	8.10	10.85	8.21	8.81	8.12	8.62
Endochitinase	GRMZM2G099454	1.01	7.41	10.58	7.72	7.66	7.21	9.41	11.50	8.45	9.46	7.27	9.48
Glutathione S-transferase 6	GRMZM2G096247	1.01	7.86	7.87	7.52	8.54	8.46	7.49	6.92	2.41	4.60	2.62	2.60
Unknown function	GRMZM2G047002	1.01	8.98	7.41	8.91	7.05	8.72	8.23	4.86	6.45	6.91	7.85	8.51
60S ribosomal protein L21	GRMZM2G149257	1.01	9.05	8.71	10.09	9.50	10.07	9.41	8.43	10.49	10.12	10.27	9.99
PLAC8 family protein	GRMZM2G027821	1.01	11.98	3.77	6.40	6.72	6.09	6.26	4.17	5.77	5.20	5.74	5.99
tRNA methyltransferase	GRMZM2G154087*	1.02	12.44	4.24	7.07	6.23	5.01	6.88	5.34	8.40	8.05	6.93	7.74
β -galactosidase	GRMZM2G417455	1.02	13.69	6.55	8.31	7.86	8.76	8.14	5.09	6.89	7.15	7.27	6.87
Unknown function	GRMZM2G075459	1.02	15.41	12.87	8.24	9.78	6.11	8.98	12.23	8.97	9.61	9.41	9.80
G protein alpha subunit	GRMZM2G064732	1.02	16.35	7.67	2.57	2.58	2.55	3.53	8.08	2.61	2.56	2.56	2.55
Dihydroflavonol-4-reductase	GRMZM2G109589	1.02	17.72	7.51	4.85	3.70	5.43	4.90	7.82	5.28	5.57	5.48	5.46
Unknown function	GRMZM2G075712	1.02	19.26	9.54	10.45	9.85	9.88	10.50	7.11	9.25	9.73	9.71	10.68
Receptor protein kinase CLAVATA1	GRMZM2G043584	1.03	30.26	6.99	8.15	7.90	8.05	8.09	5.52	7.18	7.51	7.41	7.70
Unknown function	GRMZM2G364178	1.03	30.35	2.32	8.07	8.16	7.58	8.38	2.79	8.63	8.78	8.92	8.29
Ubiquitin fusion protein	GRMZM2G116292	1.03	31.14	9.97	11.89	11.96	12.01	11.62	9.95	12.03	12.04	12.20	12.01
RNA binding protein	GRMZM2G158261	1.03	44.34	3.18	2.24	2.59	2.41	2.55	5.00	2.28	2.55	3.20	2.56
GILT protein	GRMZM2G307992	1.03	45.93	8.15	6.50	7.20	6.97	6.82	9.05	6.83	6.54	6.61	6.59
RRM containing protein	GRMZM2G153450	1.03	46.62	3.80	9.85	9.59	9.64	9.48	3.67	10.10	10.21	10.29	10.10
Pathogenesis-related protein	GRMZM2G112538	1.03	47.15	11.25	6.00	9.57	8.16	7.59	10.36	3.70	5.88	4.48	6.45
Pathogenesis-related protein	GRMZM2G112524	1.03	47.17	11.25	6.00	9.57	8.16	7.59	10.36	3.70	5.88	4.48	6.45
Unknown function	GRMZM2G140390	1.03	49.22	2.61	7.67	7.32	7.54	7.79	3.18	8.18	8.51	8.53	8.37

* The corresponding probe also targeted the closely located GRMZM2G054378 gene in position 22.09 Mbp on chromosome 5

numerous cellular processes including apoptosis, cell division, and intracellular trafficking, vesicle formation and membrane morphogenesis (Yang and Strittmatter, 2007; Nziengui and Schoefs, 2009). One member, GRMZM2G004828, annotated seed maturation RTN and located at the distal part of QTL support interval, was less expressed in RILs than in F271 at both stages. This result also pointed out an expression and a possible role of members of this gene family in growing stems.

LTP encoding genes

Plant lipid transfer proteins (LTP) are small, abundant lipid-binding proteins of biological, mostly unknown, functions (Arondel et al, 2000). Various biological roles for plant LTP have, nevertheless, been proposed, including defence against pathogens and modulation of plant development (Boutrot et al, 2008). In addition, xylogen is an extracellular arabinogalactan protein (AGP) which has a unique structure, containing a non-specific lipid transfer protein (nsLTP) domain and AGP domains which are involved in vascular development (Sieburth and Deyholos, 2005; Kobayashi et al, 2011). One member of the LTP family (GRMZM2G320373, position 162.66 Mbp) was less expressed in RILs than in F271 at stage S2.

Heat shock 60kDa protein

Heat shock proteins (Hsps) were first identified as proteins whose synthesis was enhanced by stresses. Recently, several major Hsps have been referred to as molecular chaperones of protein biogenesis. Hsp60 binds to unfolded proteins, preventing aggregation and facilitating protein folding (Craig et al, 1993). The GRMZM2G074790 gene located in position 161.73 Mbp was less expressed in RIL than F271 at stage S2, but its involvement in a pathway inducing differences in lignin traits did not appear very likely.

Genes of unknown function

The most differentially expressed gene (GRMZM2G129166, position 155.07 Mbp) in bin 6.06 was located a little downstream of the position of *EXPA5* gene. This gene, of unknown function, was more highly expressed in RILs than in F271 at both stages S1 and S2. Another gene of unknown function (GRMZM2G361064) located in position 154.99 Mbp was less expressed in RILs at the two stages. Finally, two other genes of unknown function were more expressed in RILs than in F271, one at stage S1 (GRMZM2G125037, position 153.76 Mbp), and one at stage S2 (GRMZM2G315199, position 156.12 Mbp).

Differentially expressed genes without known cell wall phenolic compound relationships and underlying QTL positions in bins 1.02 and 3.05

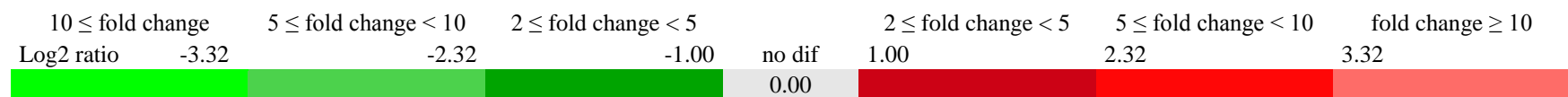
Thirty six genes located in the support intervals of the two other cell wall related QTL detected in the F288 x F271 progeny (bins 1.02 and 3.05) were differentially expressed between RILs and F271 (Supplementary tables 1E and 1F). Five and three of them, located in bin 1.02 and 3.05, respectively, had no de-

scribed function. While several of the differentially expressed genes were related to cell metabolism, signal transduction, and protein catabolism [serine/threonine kinase, RNA recognition motif (RRM) containing protein, F-box domain containing protein, guanine nucleotide-binding (G) proteins, hydroxymethylglutaryl-CoA synthase], several others had functions which could putatively be related to the phenotypic differences between F271 and the four RILs.

As other glutathione S-transferases, the *GST6* (GRMZM2G096247) and *GST3* (GRMZM2G146246) genes, located in bin 1.01 and 3.05, respectively, likely has a coupled activity with ABC transporters (Yazaki, 2005) and could be involved in transport of phenolics compounds used, or not, in cell wall assembly. These two GST were less expressed in RILs at stage S2. Exostosin-like proteins, which were initially described as encoded by a family of human tumour suppressor genes, are trans-membrane glycosyltransferases found in the endoplasmic reticulum. Such glycosyltransferases might also be involved in plant cell wall carbohydrate metabolism, such as the GT47 family xylan xylosyltransferase members that were hypothesized to play a role in the elongation of the xylan backbone (Brown et al, 2009; Wu et al, 2009). One *exostosin* gene (GRMZM2G158496) located in bin 3.05 was more expressed in RILs. Dihydroflavonol 4-reductases (DFR) are involved in the phenolic compound pathway and catalyze the reduction of dihydroflavonols to leucoanthocyanins. Leucoanthocyanidins can be utilized for the synthesis of compounds, such as catechins and oligomeric proanthocyanidins, which have anti-oxidant effects. The *DFR* gene (GRMZM2G109589) located in bin 1.02 was less expressed in RILs.

Pathogenesis-related proteins (PR proteins) were first defined as proteins that are not detectable in healthy tissues (or only at basal concentrations), but for which protein accumulation was shown upon pathological and stressed conditions. Although some PR proteins exhibit potential in vitro antimicrobial activities, a direct functional role in defense could not be demonstrated for all (Sels et al, 2008). Their expression in healthy tissue likely indicated other involvement in plant metabolism, and the two *PR* genes (GRMZM2G112538 and GRMZM2G112524) located in bin 1.03 were less expressed in RILs, especially at stage S2. Among PR proteins, γ -thionins were the first plant defense proteins described in literature, with growth inhibition activity toward pathogens (Pelegri and Franco, 2005). A flower-specific γ -thionin gene (GRMZM2G392863), located in bin 1.01, was less expressed in RILs and F271 at stage S2, indicating that such proteins are expressed elsewhere than in flowers, with roles differing from plant defense (a false annotation could also be considered). In addition, while chitin has been considered as an inducer for defence mechanisms in plants, the lower expression of an endochitinase gene (GRMZM2G099454) in RILs could

S_tables 1 - Gene expression, in below-ear internodes at two growing stages, S1 = emerging tassel stage and S2 = silking stage. GRMZM numbers are given according to the release 5b.60 of the www.maizesequence.org database. Range of colors according to the level of differential expression between F271 and RILs, and similarly for S_tables 1A to 1F, a log₂ sum intensity in green or in red indicates that the gene is less expressed or more expressed in RILs, respectively, other log₂ sum intensities (not in green or red) were not found to be statistically significant after Bonferroni correction (P < 0.05).



S_table 1F - Differentially expressed genes between F271 and four RILs underlying QTLs located in bin 3.05, in ear internodes, at two stages.

Gene	GRMZM gene	bin	Pos.	Emerging tassel stage (S1)					Silking stage (S2)				
				Mpb	F271	RIL39	RIL99	RIL54	RIL118	F271	RIL39	RIL99	RIL54
Unknown function	GRMZM2G037452	3.05	150.17	12.12	12.80	12.47	14.06	13.37	11.14	13.33	13.21	13.30	13.57
Glutathione S-transferase 3	GRMZM2G146246	3.05	154.97	9.94	9.27	9.34	9.27	9.13	9.87	8.14	8.21	8.12	8.25
Exostosin	GRMZM2G158496	3.05	160.02	5.03	8.63	8.88	7.93	8.57	3.24	8.36	8.27	7.83	7.77
F-box domain containing protein	GRMZM2G077143	3.05	163.09	8.53	5.79	5.98	5.62	6.58	8.52	6.47	6.30	6.46	6.52
Histone deacetylase	GRMZM2G057044	3.05	164.38	2.93	2.53	2.58	2.45	2.49	5.30	2.64	2.58	2.58	2.56
Unknown function	GRMZM2G004014	3.05	165.02	4.53	5.92	6.08	5.95	6.69	4.54	6.94	7.19	7.27	7.02
Ribonuclease 2	GRMZM2G064936	3.05	167.68	4.98	2.11	2.20	2.51	2.35	4.79	2.12	2.10	2.10	2.10
Chromatin assembly factor group	GRMZM2G096458	3.05	168.30	3.55	2.53	2.60	2.80	2.43	4.71	2.66	3.00	2.94	2.57
EMB1879	GRMZM2G081188	3.06	170.12	7.68	9.58	8.78	9.70	9.98	7.48	9.44	9.26	9.39	9.71
Cytochrome C	GRMZM2G081883	3.06	170.50	4.69	3.72	4.00	3.85	3.69	5.70	3.94	4.08	3.95	3.73
Unknown function	GRMZM2G147164	3.06	173.36	5.49	2.11	2.12	2.10	2.07	7.55	2.14	2.11	2.12	2.10
G10-like protein	GRMZM2G099007	3.06	177.37	11.61	5.97	5.8	5.66	5.59	11.06	5.94	5.74	5.81	5.74

cell wall gene expression in maize

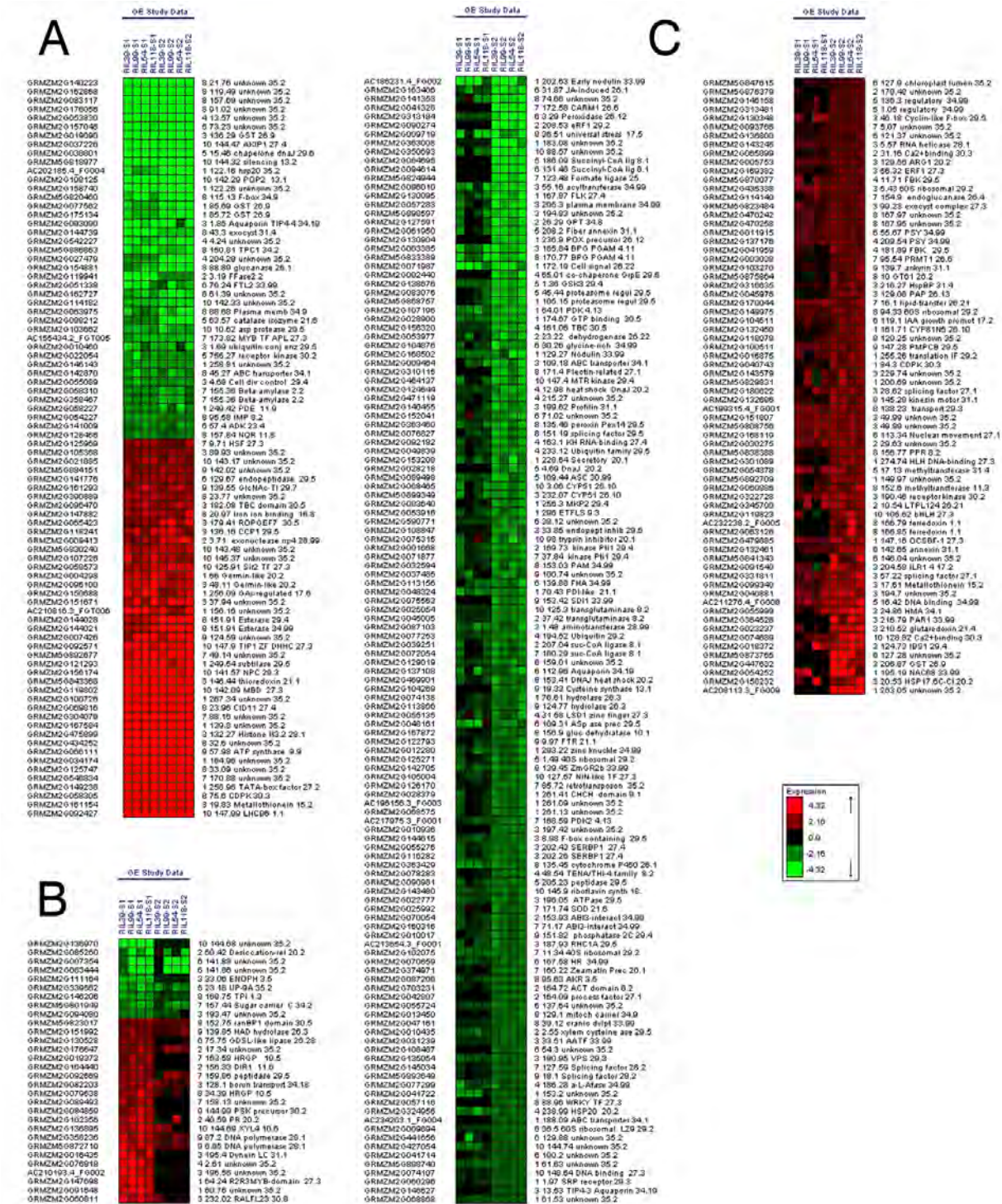


Figure 3 - Differentially expressed genes between F271 and the four RIL without known cell wall phenolic compound relationships and located elsewhere than in considered QTL positions. GRMZM numbers are given according to the reference genome B73 (Schnable et al, 2009, www.maizesequence.org release v2). Text on the right correspond to chromosome number, physical position in Mbp, annotation and Mapman bincode. (A) Differentially expressed at the two S1 and S2 stages. (B) Differentially expressed only at "emerging tassel" stage (S1). (C) Differentially expressed only at "silking date" stage (S2).

likely indicate a different function.

The CLAVATA1 (CLV1) receptor kinase regulates stem cell specification at shoot and flower meristems of Arabidopsis (Diévar et al, 2003; Durbak and Tax, 2011). CLV proteins restrict the expression domain of the stem cell-promoting homeodomain transcription factor WUSCHEL (van der Graaff et al, 2009). Similarly, FWL genes are a large gene family in plants, but with a strong conservation of key domains, suggesting a conservation of the core biochemical function of these proteins (Guo et al, 2010; Libault and Stacey, 2010). In maize, putative *fw2.2* orthologs were identified as *Cell Number Regulator* (CNR) genes, including the two closest orthologs CNR1 and CNR2. Over-expression of CNR1 reduced overall plant size while plant and organ size increased when its expression was co-suppressed or silenced, with changes in cell number, but not in cell size. CNR2 expression was found to be negatively correlated with tissue growth activity and hybrid seedling vigor (Guo et al, 2010). The CLV1 gene (GRMZM2G043584), located in bin 1.03, was more expressed in the four RILs at the two stages. Such differential expressions of genes putatively related to tissue differentiation and patterning could induce differences in lignified tissue areas and, consequently, on cell wall degradability.

Differentially expressed genes without known cell wall phenolic compound relationships and located elsewhere than in bins 6.06, 1.02 and 3.05 QTL positions

Finally, 312 differentially expressed genes between F271 and the four RIL were located elsewhere than in the three considered QTL positions and were not yet known to be involved in cell wall biosynthesis (Figure 3).

Highly differentially expressed genes

Forty two of these genes were more than 20 times differentially expressed between the four RILs and F271. No annotation was available for 18 of them. Among the 24 annotated genes, some genes could be more specifically considered as putatively involved in the differences observed between F271 and the four RILs, or with questionable role in maize. The *glutathione S-transferase GST18* (GRMZM2G019090) was less expressed in RILs at the two stages, and could be involved in the transport of phenolics compounds. The *pollen pistil incompatibility POP2* gene (GRMZM2G108125), less expressed in RILs at the two stages, encodes a transaminase that degrades GABA (Palanivelu et al, 2003), of which accumulation causes cell elongation defects and a decrease in expression of genes encoding cell wall-related proteins in Arabidopsis thaliana (Renault et al, 2011). The *Rapid Alkalinization Factor RALF23* (GRMZM2G060811), more expressed in RILs at stage S1, was shown to be down-regulated in the presence of brassinolide and over-expression of *AtRALF23* reduced plant growth, counteracting brassinolide growth-promoting effects (Srivastava et al, 2009). *AtRALF23* is synthesized as

a larger precursor protein, and is released in peptide form by the action of AtS1P, a subtilisin serine protease. One *subtilase* (GRMZM2G121293) was also more expressed in RILs at the two stages. The coactivator-associated *arginine methyltransferase, CARM1* (GRMZM2G041328), under-expressed in RILs at stage S2, modulates maintenance of embryonic stem cell pluripotency and early embryonic development in animal systems (Zhao et al, 2011). The homologous *AtPRMT4a* and *AtPRMT4b* genes were shown to be involved in regulation of flowering time (Niu et al, 2008) while the *AtPRMT5* is involved in vegetative growth and flowering time (Pei et al, 2007). The *eukaryotic peptide chain release factor subunit1-1 eRF1* (GRMZM2G090274), less expressed in RILs at stage S2, was shown to be involved in cell elongation and radial cell division. Its co-suppression in Arabidopsis has a major effect on plant morphology, with a reduction in internode elongation, reduced height of cells, and ectopic lignification (Petsch et al, 2005). Moreover, cell division in the fascicular cambial regions was altered. In addition, four other differentially expressed genes had probable involvements in transcription (*TATA box factor*, GRMZM2G149238), regulation of transcription (*MBD Methyl-CpG binding domain containing protein*, GRMZM2G119802) or RNA binding proteins (*AKIP1*, GRMZM2G037226, and *CTC-Interacting Domain CID11*, GRMZM2G069816).

Other differentially expressed genes

Among the genes less than 20 times differentially expressed between F271 and the four RILs, four genes were classified in the cell wall class, two genes encoding hydroxyproline-rich glycoprotein (HRGP), one gene encoding hydrolase XYL4 hydrolyzing O-glycosyl compound, and one gene encoding a dTDP-glucose 4,6-dehydratase, probably involved in cell wall metabolism. Two other GST and three other ABC transporters were also less expressed in RILs. The *TIP GROWTH DEFECTIVE1 S-acyl transferase (TIP1)* was strongly differentially expressed at the two stages in RILs. This gene regulates plant cell growth in Arabidopsis (Hemsley et al, 2005). Moreover, a MYB (AC155434.2_FGT005) orthologous to Arabidopsis MYR1 (*MYB related protein1*) and one NAC (GRMZM2G054252), putatively involved in cell division, were differentially expressed between the four RILs and F271. Finally, six genes, putatively involved in cell organization, were also differentially expressed, including the *annexin* (GRMZM2G061950) which was ortholog of Atg38760, a target of SND1 (Zhong et al, 2010).

Discussion

A lot of genes were shown differentially expressed between F271 and the four RILs, even if the main limit of such a transcriptomic approach using a 18,000 probes Affymetrix micro-array is obviously related to the fact that only nearly one third of the maize genes are present on the array. Several genes involved in

monolignol biosynthesis and polymerization, and a lot of transcription factors putatively in secondary wall biosynthesis and lignified tissue assembly, could thus not be investigated. In addition, differences between cell wall traits in lines could be related to genetic variation giving no difference in expression, such as variation in protein sequence and efficiency. Moreover, differences could be related to variation in regulation by miRNA, even if no miRNA related to lignification has yet been identified in bin 6.06.

Among the differentially expressed genes located in bin 6.06, the closest genes to QTL peak position were three genes, which encode proteins of unknown function (GRMZM2G037111, GRMZM2G129166 and GRMZM2G315199). None of the other genes could be easily considered as candidates underlying the effect of this detected QTL, except the *FKBP* gene and the α -*expansin 5* gene even if these genes were not clearly related to lignification but to cell division and extension, respectively. The differential expression could also be related to allelic variation in these genes and/or their co-regulation.

Only a few genes located outside the QTL positions in bin 6.06 and with a function related to cell wall assembly were differentially expressed between the four considered RILs and F271. Among genes involved in monolignol biosynthesis, a differential expression was only shown for the *ZmPAL2b*. Complementarily, two *ZRP4-like OMT*, one *ZmPox*, two *ZmLac*, and one *glycosyl hydrolase* gene were also differentially expressed in RILs and in F271. Their variable expression, related or not to a regulation factor located in bin 6.06, could contribute to explain the observed differences.

Three *ZmMYB* were differentially expressed when comparing the four RILs to F271. The *ZmMYB* located in bin 1.07, which is orthologous to AtMYB85 and EgMYB2 and positively regulates gene transcription, colocalized with several major lignin-related QTLs shown in the F838 x F286 progeny (Barrière et al, 2008; Barrière et al, 2010). Similarly, the *ZmMYB2* and the *ZmMYB8* located in bin 3.04 and 3.05, which is orthologous to AtMYB86 and AtMYB4 respectively and negatively regulates gene transcription, colocalized with several lignin-related QTLs shown in the Rlo x WM13 progenies (Barrière et al, 2012). They are also located upstream of other QTLs shown in bin 3.05 in the F288 x F271 progeny (Roussel et al, 2002; Thomas et al, 2010). These three *ZmMYB* could not be considered as the primary determinants of the observed variation as they are not colocalized with the major QTLs of the F288 x F271 progeny. However, the observed differences in expression strengthened their plausible involvement as target of an upstream regulation factor potentially colocalizing with the major QTL under investigation. Similarly, the differentially expressed *C3HC4 zinc finger* could be considered as a target factor. This gene was underlying QTLs shown in the F288 x F271 progeny (Roussel et

al, 2002; Thomas et al, 2010) and was also in close position, even a little upstream, to QTLs shown in the F838 x F286 progeny (Barrière et al, 2008). In addition, the differentially expressed NAC was localized under QTLs of two RIL progeny (F7025 x F4 and F116 x F2).

Finally, among the other genes differentially expressed, some genes were involved in plant growth in particular in the cell division and elongation. Such genes could be considered as probable candidates even if the modifications engendered by these genes have not yet been proved to be involved in cell wall lignification and degradability.

Conclusions

No definite conclusion was available from the current expression study for the discovery of the candidate gene(s) underlying cell wall related QTLs located in bin 6.06. In addition, and relative to the probes spotted on the array, the determinant did not appear as one of the genes previously listed as putatively involved in cell wall assembly. Corroborating results obtained on Arabidopsis, genes involved in monolignols biosynthesis are very likely not candidates underlying major QTLs for lignin content or cell wall degradability, despite the major effect of mutations in *ZmCCR1*, *ZmCOMT*, or *ZmCAD2* genes. Variation in expression of monolignol-related genes could, nevertheless, be a consequence of upstream regulation factors inducing differences in lignin content, structure and composition.

Observed differential gene expression between the four RILs and F271 was certainly partly not related to cell wall related traits, and was the consequence of allelic variations and linkage disequilibrium, even if several of them were likely co-regulated with, or targeted by, the primary trait determinants. In addition, a few genes could be considered as new members involved in secondary wall assembly. The differential expression of the three *ZmMYB* and the *C3HC4 zinc finger* could be considered with a reasonable level of confidence as related to variation in cell wall traits between F271 and the four RILs and, possibly, as a consequence of the genomic variation in bin 6.06 area. Assuming this hypothesis was true, the QTL underlying genetic determinant would, thus, be a gene involved in an upstream mechanism of lignin biosynthesis and lignified tissue assembly. In addition, the involvement of genes of still unknown function has likely to be considered.

Acknowledgements

This work has been funded by the maize breeding companies involved in the PROMAÏS - INRA network on maize cell wall lignification and digestibility (Advanta, Caussade Semences, Limagrain Europe, MaïsAdour, Monsanto SAS, Pioneer Génétique, Pau Euralis, R2n RAGT Semences, SDME KWS France, Syngenta seeds). Thanks to Christiane Minault, Dominique Denoue and Pascal Vernoux for their help in

experimentation carried out at INRA Lusignan.

References

- Andersen JR, Asp T, Lu YC, Kloiber-Maitz M, Ouzonova M, Luebbertstedt T, 2009. Development and mapping of gene-tagged SNP markers in laccases of maize (*Zea mays* L). *Plant Breeding* 128: 423-425
- Argillier O, Barrière Y, Hébert Y, 1995. Genetic variation and selection criterion for digestibility traits of forage maize. *Euphytica* 82: 175-184
- Arondel V, Vergnolle C, Cantrel C, Kader JC, 2000. Lipid transfer proteins are encoded by a small multigene family in *Arabidopsis thaliana*. *Plant Sci* 157: 1-12
- Aufrère J, Michalet-Doreau B, 1983. *In vivo* digestibility and prediction of digestibility of some by-products, pp. 25-33. In: EEC seminar, Melle Gontrode
- Barrett T, Troup DB, Wilhite SE, Ledoux P, Rudnev D, Evangelista C, Kim IF, Soboleva A, Tomashevsky M, Edgar R, 2007. NCBI GEO: mining tens of millions of expression profiles - database and tools update. *Nucleic Acids Res* 35: D760-765.
- Barrière Y, Gibelin C, Argillier A, Méchin V, 2001. Genetic analysis and QTL mapping in forage maize based on recombinant inbred lines descended from the cross between F288 and F271. I - Yield, earliness, starch and crude protein content. *Maydica* 46: 253-266
- Barrière Y, Guillet C, Goffner D, Pichon M, 2003. Genetic variation and breeding strategies for improved cell wall digestibility in annual forage crops. *Anim Res* 52: 193-228
- Barrière Y, Ralph J, Méchin V, Guillaumie S, Grabber JH, Argillier O, Chabbert B, Lapierre C, 2004. Genetic and molecular basis of grass cell wall biosynthesis and degradability. II. Lessons from brown-midrib mutants. *C R Biol* 327: 847-860
- Barrière Y, Thomas J, Denoue D, 2008. QTL mapping for lignin content, lignin monomeric composition, p-hydroxycinnamate content, and cell wall digestibility in the maize recombinant inbred line progeny F838 x F286. *Plant Sci* 175: 585-595
- Barrière Y, Méchin V, Lafarguette F, Manicacci D, Guillon F, Wang H, Laressergues D, Pichon M, Bosio M, Tatout C, 2009. Toward the discovery of maize cell wall genes involved in silage maize quality and capacity to biofuel production. *Maydica* 54: 161-198
- Barrière Y, Méchin V, Denoue D, Bauland C, Laborde J, 2010. QTL for yield, earliness and cell wall digestibility traits in topcross experiments of F838 x F286 RIL progenies. *Crop Sci* 50: 1761-1772
- Barrière Y, Méchin V, Lefevre B, Maltese S, 2012. QTLs for agronomic and cell wall traits in a maize RIL progeny derived from a cross between an old Minnesota13 line and a modern Iodent line. *Theor Appl Genet* (in press)
- Berthet S, Demont-Caulet N, Pollet B, Bidzinski P, Cezard L, Le Bris, P, Borrega N, Herve J, Blondet E., Balzergue S, Lapierre C, Jouanin L, 2011. Disruption of LACCASE4 and 17 results in tissue-specific alterations to lignification of *Arabidopsis thaliana* stems. *Plant Cell* 23: 1124-1137
- Boerjan W, Ralph J, Baucher M, 2003. Lignin biosynthesis. *Ann Rev Plant Biol* 54: 519-546
- Boudet AM, 2000. Lignins and lignification: selected issues. *Plant Physiol Biochem* 38: 81-96
- Boutrot F, Chantret N, Gautier MF, 2008. Genome-wide analysis of the rice and *Arabidopsis* non-specific lipid transfer protein (*nsLtp*) gene families and identification of wheat *nsLtp* genes by EST data mining. *BMC Genomics* 9: 86
- Brown DM, Zhang ZN, Stephens E, Dupree P, Turner SR, 2009. Characterization of IRX10 and IRX10-like reveals an essential role in glucuronoxylan biosynthesis in *Arabidopsis*. *Plant J* 57: 732-746
- Caparros-Ruiz D, Fornalé S, Civardi L, Puigdomènech P, Rigau P, 2006. Isolation and characterisation of a family of laccases in maize. *Plant Sci* 171: 217-225
- Cardinal AJ, Lee M, Moore KJ, 2003. Genetic mapping and analysis of quantitative trait loci affecting fiber and lignin content in maize. *Theor Appl Genet* 106: 866-874
- Carlsbecker A, Lee JY, Roberts CJ, Dettmer J, Lehesranta S, Zhou J, Lindgren O, Moreno-Risueno MA, Vaten A, Thitamadee S, Campilho A, Sebastian J, Bowman JL, Helariutta Y, Benfey PN, 2010. Cell signalling by microRNA165/6 directs gene dose-dependent root cell fate. *Nature* 465: 316-321
- Chavigneau H, Goué N, Courtial A, Jouanin L, Raymond M, Méchin V, Barrière Y, 2012. QTL for floral stem lignin content and degradability in three recombinant inbred line (RIL) progenies of *Arabidopsis thaliana* and search for candidate genes involved in cell wall biosynthesis and degradability. *OJGen* 2: 7-30
- Craig EA, Gambill BD, Nelson RJ, 1993. Heat shock proteins: molecular chaperones of protein biogenesis. *Microbiol Rev* 57(2): 402-414
- De Obeso M, Caparros-Ruiz, D Vignols F, Puigdomenech P, Rigau J, 2003. Characterisation of maize peroxidases having differential patterns of mRNA accumulation in relation to lignifying tissues. *Gene* 309: 23-33
- Dence CW, Lin SY, 1992. The determination of lignin, pp. 33-61. In: *Methods in lignin chemistry*. Lin SY and Dence CW eds. Springer-Verlag, Berlin
- Dias AP, Grotewold E, 2003. Manipulating the accumulation of phenolics in maize cultured cells using transcription factors. *Biochem Eng J* 14: 207-216
- Diévert A, Dalal M, Tax FE, Lacey AD, Huttly A, Li J, Clark SE, 2003. CLAVATA1 dominant-negative alleles reveal functional overlap between multiple receptor kinases that regulate meristem and organ development. *Plant Cell* 15: 1198-1211
- Di Lorenzo L, WysockaDiller J, Malamy JE, Pysh

- L, Helariutta Y, Freshour G, Hahn MG, Feldmann KA, Benfey PN, 1996. The SCARECROW gene regulates an asymmetric cell division that is essential for generating the radial organization of the Arabidopsis root. *Cell* 86: 423-433
- Dolstra O, Medema JH, 1990. An effective screening method for genetic improvement of cell-wall digestibility in forage maize, pp. 258-270. In: Proceedings 15th congress maize and sorghum section of Eucarpia. 4-8 June. Baden, Austria
- Durbak AR, Tax FE, 2011. CLAVATA signaling pathway receptors of Arabidopsis regulate cell proliferation in fruit organ formation as well as in meristems. *Genetics* 189: 177-194
- Escamilla-Trevino LL, Chen W, Card ML, Shih MC, Cheng CL, Poulton JE, 2006. Arabidopsis thaliana beta-glucosidases BGLU45 and BGLU46 hydrolyse monoglignol glucosides. *Phytochem* 67: 1651-1660
- Fagerstedt KV, Kukkola EM, Koistinen VV, Takahashi J, Marjamaa K, 2010. Cell wall lignin is polymerised by class III secretable plant peroxidases in Norway spruce. *J Integr Plant Biol* 52(2): 186-194
- Fornale S, Sonbol FM, Maes T, Capellades M, Puigdomenech P, Rigau J, Caparros-Ruiz D, 2006. Down-regulation of the maize and Arabidopsis thaliana caffeic acid O-methyl-transferase genes by two new maize R2R3-MYB transcription factors. *Plant Mol Biol* 62: 809-823
- Fornalé S, Shi X, Chai C, Encina A, Irar S, Capellades M, Fuguet E, Torres JL, Rovira P, Puigdomènech P, Rigau J, Grotewold E, Gray J, Caparrós-Ruiz D, 2010. ZmMYB31 directly represses maize lignin genes and redirects the phenylpropanoid metabolic flux. *Plant J* 64: 633-644
- Foucart C, Jauneau A, Gion JM, Amelot N, Martinez Y, Panegos P, Grima-Pettenati J, Sivadon P, 2009. Overexpression of EgROP1, a Eucalyptus vascular-expressed Rac-like small GTPase, affects secondary xylem formation in Arabidopsis thaliana. *New Phytologist* 183: 1014-1029
- Gagnot S, Tamby JP, Martin-Magniette ML, Bitton F, Taconnat L, Balzergue S, Aubourg S, Renou JP, Lechamy A, Brunaudet V, 2008. CATdb: a public access to Arabidopsis transcriptome data from the URGV CATMA platform. *Nucleic Acids Res* 36: D986-D990
- Ge Y, Dudoit S, Speed TP, 2003. Resampling-based multiple testing for microarray data analysis. *TEST* 12: 1-44
- Gentleman R, Carey V, 2002. Bioconductor. *RNews* 2: 1116
- Gion JM, Carouche A, Deweer S, Bedon F, Pichavant F, Charpentier JP, Bailleres H, Rozenberg P, Carocha V, Ognouabi N, Verhaegen D, Grima-Pettenati J, Vigneron P, Plomion C, 2011. Comprehensive genetic dissection of wood properties in a widely-grown tropical tree: Eucalyptus. *BMC Genomics* 12: 301
- Goicoechea M, Lacombe E, Legay S, Mihaljevic S, Rech P, Jauneau A, Lapierre C, Pollet B, Verhaegen D, Chaubet-Gigot N, Grima-Pettenati J, 2005. EgMYB2, a new transcriptional activator from Eucalyptus xylem, regulates secondary cell wall formation and lignin biosynthesis. *Plant J* 43: 553-567
- Goering HK, Van Soest PJ, 1970. Forage fiber analysis (Apparatus, reagents, procedures and some applications), pp 1-20. In: US Dept Agri Sc Handbook n° 379
- van der Graaff E, Laux T, Rensing SA, 2009. The WUS homeobox-containing (WOX) protein family. *Genome Biology* 10: 248
- Gray J, Caparros-Ruiz D, Grotewold E, 2012. Grass phenylpropanoids: Regulate before using! *Plant Science* 184: 112-120
- Guillaumie S, Pichon M, Martinant JP, Bosio M, Goffner D, Barrière Y, 2007a. Differential expression of phenylpropanoid and related genes in brown-midrib bm1, bm2, bm3, and bm4 young near-isogenic maize plants. *Planta* 226: 235-250
- Guillaumie S, San-Clemente H, Deswarte C, Martinez Y, Lapierre C, Murigneux A, Barrière Y, Pichon M, Goffner D, 2007b. MAIZEWALL. Database and developmental gene expression profiling of cell wall biosynthesis and assembly in maize. *Plant Physiol.* 143: 339-363
- Guillaumie S, Goffner D, Barbier B, Martinant JP, Pichon M, Barrière Y, 2008. Expression of cell wall related genes in basal and ear internodes of silking brown-midrib-3, caffeic acid O-methyltransferase (COMT) down-regulated, and normal maize plants. *BMC Plant Biol* 8: 71
- Guillet-Claude C, Birolleau-Touchard C, Manicacci D, Rogowsky PM, Rigau J, Murigneux A, Martinant JP, Barrière Y, 2004. Nucleotide diversity of the *ZmPox3* maize peroxidase gene: Relationships between a MITE insertion in exon 2 and variation in forage maize digestibility. *BMC Genetics* 5: 19
- Guo M, Rupe MA, Dieter JA, Zou J, Spielbauer D, Duncan KE, Howard RJ, Hou Z, Simmons CR, 2010. Cell Number Regulator1 affects plant and organ size in maize: implications for crop yield enhancement and heterosis. *Plant Cell* 22: 1057-103
- Harrar Y, Bellini C, Faure JD, 2001. FKBP: at the crossroads of folding and transduction. *Trends Plant Sci* 6:426-431
- Hatfield R, Vermerris W, 2001. Lignin formation in plants. The dilemma of linkage specificity. *Plant Physiol* 126: 1351-1357
- Hatfield R, Fukushima RS, 2005. Can lignin be accurately measured. *Crop Sci* 45: 832-839
- He X, Hall MB, Gallo-Meagher M, Smith RL, 2003. Improvement of forage quality by downregulation of maize O-methyltransferase. *Crop Sci* 43: 2240-2251
- Heine GF, Malik V, Dias AP, Grotewold E, 2007. Ex-

- pression and molecular characterization of Zm-MYB-IF35 and related R2R3 MYB transcription factors. *Mol Biotechnol.* 37(2): 155-64
- Held BM, Wang HQ, John I, Wurtele ES, Colbert JT, 1993. An messenger-RNA putatively coding for an O-methyltransferase accumulates preferentially in maize roots and is located predominantly in the region of the endodermis. *Plant Physiol* 102: 1001-1008
- Hemsley A, Kemp A, Grierson C, 2005. The TIP GROWTH DEFECTIVE1 S-Acyl Transferase Regulates Plant Cell Growth in Arabidopsis. *Plant Cell* 17(9): 2554-2563
- Ilegems M, Douet V, Meylan-Bettex M, Uyttewaal M, Brand L, Bowman JL, Stieger PA, 2010. Interplay of auxin, KANADI and Class III HD-ZIP transcription factors in vascular tissue formation. *Development* 137: 975-984
- Im KH, Cosgrove DT, Jones AM, 2000. Subcellular localization of expansin mRNA in xylem cells. *Plant Physiol* 123: 463-470
- Irizarry RA, Ooi SL, Wu Z, Boeke JD, 2003. Use of mixture models in a microarray based screening procedure for detecting differentially represented yeast mutants. *Stat Appl Genet Mol Biol* 2: Article1
- Juarez MT, Kui JS, Thomas J, Heller BA, Timmermans MCP, 2004. MicroRNA-mediated repression of rolled leaf1 specifies maize leaf polarity. *Nature* 428: 84-88
- Jung H, Mertens D, Payne A, 1997. Correlation of acid detergent lignin and Klason lignin with digestibility of forage dry matter and neutral detergent fiber. *J Dairy Sci* 80: 1622-1628
- Kaneda M, Schuetz M, Lin B, Chanis C, Hamberger B, Western T, Ehling J, Samuels A, 2011. ABC transporters coordinately expressed during lignification of Arabidopsis stems include a set of ABCBs associated with auxin transport. *Journal of Experimental Botany*, 62: 2063-2077
- Kawaoka A, Ebinuma H, 2001. Transcriptional control of lignin biosynthesis by tobacco LIM protein. *Phytochemistry* 57: 1149-1157
- Ko JH and Han KH, 2004. Arabidopsis whole-transcriptome profiling defines the features of coordinated regulations that occur during secondary growth. *Plant Molecular Biol* 55: 433-453
- Ko JH, Prassinis C, Han KH, 2006. Developmental and seasonal expression of PtaHB1, a Populus gene encoding a class III HD-Zip protein, is closely associated with secondary growth and inversely correlated with the level of microRNA (miR166). *New Phytologist* 169: 469-478
- Ko JH, Kim WC, Han KH, 2009. Ectopic expression of MYB46 identifies transcriptional regulatory genes involved in secondary wall biosynthesis in Arabidopsis. *Plant J* 60: 649-665
- Kobayashi Y, Motose H, Iwamoto K, Fukuda H, 2011. Expression and genome-wide analysis of the xylogen-type gene family. *Plant Cell Physiol* 52: 1095-1106
- Krakowsky MD, Lee M, Beeghly HH, Coors JG, 2003. Characterization of quantitative trait loci affecting fiber and lignin in maize (*Zea mays* L). *Maydica* 48: 283-292
- Krakowsky M, Lee M and Coors J, 2005. Quantitative trait loci for cell-wall components in recombinant inbred lines of maize (*Zea mays* L). I: Stalk tissue. *Theor Appl Genet* 111: 337-346
- Lanot A, Hodge D, Jackson R, George G, Elias L, Lim E, Vaistij F, Bowles D, 2006. The glucosyltransferase UGT72E2 is responsible for monolignol 4-O-glucoside production in Arabidopsis thaliana. *Plant Journal* 48: 286-295
- Lee H, Kim B, Song SK, Heo JO, Yu NI, Lee SA, Kim M, Kim DG, Sohn SO, Lim CE, Chang KS, Lee MM, Lim J, 2008. Large-scale analysis of the GRAS gene family in Arabidopsis thaliana. *Plant Mol Biol* 67: 659-670
- Legay S, Lacombe E, Goicoechea M, Briere C, Seguin A, MacKay J, Grima-Pettenati J, 2007. Molecular characterization of EgMYB1, a putative transcriptional repressor of the lignin biosynthetic pathway. *Plant Sci* 173: 542-549
- Legay S, Sivadon P, Blervacq AS, Pavy N, Baghdady A, Tremblay L, Levasseur C, Ladouce N, Lapierre C, Séguin A, Hawkins S, Mackay J, Grima-Pettenati J, 2010. EgMYB1, an R2R3 MYB transcription factor from eucalyptus negatively regulates secondary cell wall formation in Arabidopsis and poplar. *The New Phytologist* 188: 774-786
- Li Y, Jones L, McQueen-Mason S, 2003. Expansins and cell growth. *Curr Opin Plant Biol* 6: 603-610
- Li E, Bhargava A, Qiang W, Friedmann MC, Forneris N, Savidge RA, Johnson LA, Mansfield SD, Ellis BE, Douglas CJ, 2012. The Class II KNOX gene *KNAT7* negatively regulates secondary wall formation in Arabidopsis and is functionally conserved in Populus. *New Phytol* 194: 102-115
- Libault M, Stacey G, 2010. Evolution of FW2.2-like (*FWL*) and *PLAC8* genes in eukaryotes. *Plant Signal Behav* 5: 1226-1228
- Ma K, Xiao J, Li X, Zhang Q, Lian X, 2009. Sequence and expression analysis of the C3HC4-type RING finger gene family in rice. *Gene* 444: 33-45
- Matthews J, Bhati M, Lehtomaki E, Mansfield R, Cubeddu L, MacKay J, 2009. It takes two to tango: the structure and function of LIM, RING, PHD and MYND domains. *Cur Pharmaceutical Design* 15: 3681-3696
- Méchin V, Argillier O, Hébert Y, Guingo E, Moreau L, Charcosset A, Barrière Y, 2001. Genetic analysis and QTL mapping of cell wall digestibility and lignification in silage maize. *Crop Sci* 41: 690-697
- Méchin V, Argillier O, Rocher F, Hébert Y, Mila I, Pollet B, Barrière Y and Lapierre C, 2005. In search of a maize ideotype for cell wall enzymatic degradability using histological and biochemical lignin

- characterization. *J Agricul and Food Chem* 53: 5872-5881
- Miao YC, Liu CJ, 2010. ATP-binding cassette-like transporters are involved in the transport of lignin precursors across plasma and vacuolar membranes. *Proc Natl Acad Sci USA* 107: 22728-22733
- Milioni D, Sado PE, Stacey NJ, Domingo C, Roberts K, McCann MC, 2001. Differential expression of cell-wall-related genes during the formation of tracheary elements in the *Zinnia* mesophyll cell system. *Plant Mol Biol* 47: 221-238
- Minic Z, Jamet E, San-Clemente H, Pelletier S., Renou JP, Rihouey C, Okinyo DP, Proux C, Lerouge P, Jouanin L, 2009. Transcriptomic analysis of *Arabidopsis* developing stems: a close-up on cell wall genes. *BMC Plant Biology* 9: 6
- Muller B, Bourdais G, Reidy B, Bencivenni C, Massonneau A, Condamine P, Rolland G, Conéjéro G, Rogowsky P, Tardieu F, 2007. Association of specific expansins with growth in maize leaves is maintained under environmental, genetic, and developmental sources of variation. *Plant Physiol* 143: 278-290
- Niu L, Zhang Y, Pei Y, Liu C, Cao X, 2008. Redundant requirement for a pair of PROTEIN ARGININE METHYLTRANSFERASE4 Homologs for the proper regulation of *Arabidopsis* flowering time. *Plant Physiol* 148: 490-503
- Nziengui H, Schoefs B, 2009. Functions of reticulons in plants: What we can learn from animals and yeasts? *Cel and Molr Life Sci* 66: 584-595
- Onnerud H, Zhang L, Gellerstedt G, Henriksson G, 2002. Polymerization of monolignols by redox shuttle-mediated enzymatic oxidation: a new model in lignin biosynthesis I. *Plant Cell* 14: 1953-1962
- Palanivelu R, Brass L, Edlund AF, Preuss D, 2003. Pollen tube growth and guidance is regulated by *POP2*, an *Arabidopsis* gene that controls GABA levels. *Cell* 114: 47-59
- Parker G, Schofield R, Sundberg B, Turner S, 2003. Isolation of *COV1*, a gene involved in the regulation of vascular patterning in the stem of *Arabidopsis*. *Development* 130: 2139-2148
- Paz-Ares J Ghosal D, Wienand U, Peterson PA, Saedler H, 1987. The regulatory *c1* locus of *Zea mays* encodes a protein with homology to myb proto-oncogene products and with structural similarities to transcriptional activators. *EMBO J* 6: 3553-3558
- Pei Y, Niu L, Lu F, Liu C, JZhai J, Kong X, Cao X, 2007. Mutations in the Type II Protein Arginine Methyltransferase *AtPRMT5* result in pleiotropic developmental defects in *Arabidopsis*. *Plant Physiol.* 144: 1913-1923
- Pelegrini PB, Franco OL, 2005. Plant gamma-thionins: novel insights on the mechanism of action of a multi-functional class of defense proteins. *Int J Biochem Cell Biol* 37: 2239-2253
- Pesquet E, Korolev AV, Calder G, Lloyd CW, 2010. The microtubule-associated protein AtMAP70-5 regulates secondary wall patterning in *Arabidopsis* wood cells. *Cur Biol* 20: 744-749
- Petsch K, Mylne J, Botella J, 2005. Cosuppression of Eukaryotic Release Factor 1-1 in *Arabidopsis* Affects Cell Elongation And Radial Cell Division. *Plant Physiol.* 139: 115-126
- Pichon M, Deswartes C, Gerentes D, Guillaumie S, Lapierre C, Toppan A, Barrière B, Goffner D, 2006. Variation in lignin and cell wall digestibility in caffeic acid O-methyltransferase down-regulated maize half-sib progenies in field experiments. *Mol Breed* 18: 253-261
- Ranjan P, Yin T, Zhang X, Kalluri U, Yang X, Jawdy S, Tuskan G, 2010. Bioinformatics-based identification of candidate genes from QTLs associated with cell wall traits in *Populus*. *Bioenerg Res* 3: 172-182
- Ratcliffe OJ, Riechmann JL, Zhang JZ, 2000. *Interfacicular fiberless1* is the same gene as *REVOLUTA*. *Plant Cell* 12: 315-317
- Renault H, El Amrani A, Palanivelu R, Updegraff EP, Yu A, Renou JP, Preuss D, Bouchereau A, Deleu C, 2011. GABA accumulation causes cell elongation defects and a decrease in expression of genes encoding secreted and cell wall-related proteins in *Arabidopsis thaliana*. *Plant Cell Physiol* 52(5): 894-908
- Rengel D, San-Clemente H, Servant F, Ladouce N, Paux E, Wincker P, Couloux A, Sivadon P, Grima-Pettenati J, 2009. A new genomic resource dedicated to wood formation in *Eucalyptus*. *BMC Plant Biol* 9: 36
- Riboulet C, Fabre F, Dénoue D, Martinant JP, Lefevre B, Barrière Y, 2008a. QTL mapping and candidate gene research for lignin content and cell wall digestibility in a topcross of a flint recombinant inbred line progeny harvested at silage stage. *Maydica* 53: 1-9
- Riboulet C, Lefèvre B, Denoue D, Barrière Y, 2008b. Genetic variation in maize cell wall for lignin content, lignin structure, p-hydroxycinnamic acid content, and digestibility in a set of 19 lines at silage harvest maturity. *Maydica* 53: 11-19
- Riboulet C, Guillaumie S, Méchin V, Bosio M, Pichon M, Goffner D, Lapierre C, Pollet B, Lefèvre B, Martinant JP, Barrière Y, 2009. Kinetics of phenylpropanoid gene expression in maize growing internodes: Relationships with cell wall deposition. *Crop Sci* 49: 211-223
- Roussel V, Gibelin C, Fontaine AS, Barrière Y, 2002. Genetic analysis in recombinant inbred lines of early dent forage maize. II - QTL mapping for cell wall constituents and cell wall digestibility from per se value and top cross experiments. *Maydica* 47: 9-20
- Schnable PS, Ware D, Fulton RS, Stein JC, Wei F,

- Pasternak S, et al., 2009. The B73 maize genome: complexity, diversity, and dynamics. *Science* 326: 1112-1115
- Scott-Craig JS, Casida JE, Poduje L, Walton JD, 1998. Herbicide safener-binding protein of maize - Purification, cloning, and expression of an encoding cDNA. *Plant Physiol* 116: 1083-1089
- Sels J, Mathys J, De Coninck BM, Cammue BP, De Bolle MF, 2008. Plant pathogenesis-related (PR) proteins: A focus on PR peptides. *Plant Physiol Biochem* 46: 941-950
- Shen H, Yin YB, Chen F, Xu Y, Dixon RA, 2009. A bioinformatic analysis of NAC genes for plant cell wall development in relation to lignocellulosic bio-energy production. *BioEnergy Research* 2: 217-232
- Shin JH, Jeong DH, Park MC, An G, 2005. Characterization and transcriptional expression of the alpha-expansin gene family in rice. *Mol Cells* 20: 210-218
- Sieburth LE, Deyholos MK, 2005. Vascular development: the long and winding road. *Curr Opin Plant Biol*. 9: 48-54
- Sonbol FM, Fornalé S, Capellades M, Encina A, Tourino S, Torres JL, Rovira P, Ruel K, Puigdomenech P, Rigau J, Caparros-Ruiz D, 2009. The maize *ZmMYB42* represses the phenylpropanoid pathway and affects the cell wall structure, composition and degradability in *Arabidopsis thaliana*. *Plant Molecular Biology* 70: 283-296
- Srivastava R, Liu JX, Guo H, Yin Y, Howell SH, 2009. *Howell* Regulation and processing of a plant peptide hormone, AtRALF23, in *Arabidopsis*. *Plant J* 59: 930-939
- Struik P, 1983. Physiology of forage maize (*Zea mays* L) in relation to its production and quality, pp. 1-252. PhD Dissertation, Agricultural University, 6700 GW Wageningen, The Netherlands
- Tamagnone L, Merida A, Parr A, Mackay S, Culianez-Marcia FA, Roberts K, Martin C, 1998. The AmMYB308 and AmMYB330 transcription factors from *Antirrhinum* regulate phenylpropanoid and lignin biosynthesis in transgenic tobacco. *Plant Cell* 10: 135-154
- Tamasloukht B, Lam MS-JWQ, Martinez Y, Tozo K, Barbier O, Jourda C, Jauneau A, Borderie, G, Balzergue S, Renou JP, Huguet S, Martinant JP, Tatout C, Lapiere C, Barrière Y, Goffner D, Pichon M, 2011. Characterization of a cinnamoyl-CoA reductase1 (*CCR1*) mutant in maize: effects on lignification, fibre development, and global gene expression. *J Exp Bot* 62: 3837-3848
- Thomas J, Guillaumie S, Verdu C, Denoue D, Pichon M, Barrière Y, 2010. Cell wall phenylpropanoid-related gene expression in early maize recombinant inbred lines differing in parental alleles at a major lignin QTL position. *Mol Breeding* 25: 105-124
- Thimm O, Blasing O, Gibon Y, Nagel A, Meyer S, Kru-ger P, Selbig J, Muller LA, Rhee SY, Stitt M, 2004. MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. *Plant J* 37: 914-939
- Tokunaga N, Kaneta T, Sato S, Sato Y, 2009. Analysis of expression profiles of three peroxidase genes associated with lignification in *Arabidopsis thaliana*. *Physiol Plantarum* 136: 237-249
- Vittorioso P, Cowling R, Faure JD, Caboche M, Bellini C, 1998. Mutation in the *Arabidopsis PASTICCINO1* gene, which encodes a new fk506-binding protein-like protein, has a dramatic effect on plant development. *Mol Cell Biol* 18(5): 3034-3043
- Wilkins O, Nahal H, Foong J, Provart NJ, Campbell MM, 2009. Expansion and diversification of the *Populus* R2R3-MYB family of transcription factors. *Plant Physiol* 149: 981-993
- Wissenbach M, Uberlacker B, Vogt F, Becker D, Salamini F, Rohde W, 1993. MYB genes from *Hordeum vulgare* - Tissue-specific expression of chimeric MYB Promoter/Gus genes in transgenic tobacco. *Plant J* 4: 411-422
- Wu AM, Rihouey C, Seveno M, Hornblad E, Singh SK, Matsunaga T, Ishii T, Lerouge P, Marchant A, 2009. The *Arabidopsis* IRX10 and IRX10-LIKE glycosyltransferases are critical for glucuronoxylan biosynthesis during secondary cell wall formation. *Plant J* 57: 718-731
- Yamaguchi M, Kubo M, Fukuda H, Demura T, 2008. Vascular-related nac-domain7 is involved in the differentiation of all types of xylem vessels in *Arabidopsis* roots and shoots. *Plant J* 55: 652-664
- Yang YS, Strittmatter SM, 2007. The reticulons: a family of proteins with diverse functions. *Genome Biology* 8(12): 234
- Yazaki K, 2005. Transporters of secondary metabolites. *Curr Opin Plant Biol* 8: 301-307
- Zhao H, Zhang Y, Dai H, Zhang Y, Shen Y, 2011. *CARM1* Mediates Modulation of Sox2. *PLoS One*. 2011; 6: e27026
- Zhong RQ, Ye ZH, 1999. *IFL1*, a gene regulating interfascicular fiber differentiation in *Arabidopsis*, encodes a homeodomain-leucine zipper protein. *Plant Cell* 11: 2139-2152
- Zhong, RQ, Lee C, McCarty RL, Ye ZH, 2008. A battery of transcription factors involved in the regulation of secondary cell wall biosynthesis in *Arabidopsis*. *Plant Cell* 20: 2763-2782
- Zhong, RQ, Ye ZH, 2009. Transcriptional regulation of lignin biosynthesis. *Plant Signaling and Behavior* 4: 1-7
- Zhong R, Lee C, Ye ZH, 2010. Global analysis of direct targets of secondary wall NAC masters witches in *Arabidopsis*. *Molecular Plant* 3: 1087-1103
- Zhong R, Lee C, McCarthy RL, Reeves CK, Jones EG, Ye ZH, 2011. Transcriptional activation of secondary wall biosynthesis by rice and maize NAC and MYB transcription factors. *Plant Cell*

Physiol 52: 1856-1871

Zhou J, Lee C, Zhong RQ, Ye ZH, 2009. MYB58 and MYB63 are transcriptional activators of the lignin biosynthetic pathway during secondary cell wall formation in arabidopsis. Plant Cell 21: 248-266

CHAPITRE II - QTLs FANTÔMES ET QTLs EN POSITIONS
MULTIPLES DANS LE BIN 6.06 DE LA FAMILLE RIL F288 x
F271. APPROCHE PAR MARQUAGE CIBLÉ DE LA
DESCENDANCE EN S'APPUYANT SUR LE GÉNOME
SÉQUENCÉ B73. RECHERCHE DE GÈNES CANDIDATS

2.1. Avant propos

Après des travaux conduits sur hybrides inscrits et sur hybrides expérimentaux, les premières études systématiques de variabilité génétique sur des lignées de maïs pour le caractère de digestibilité des parois ont été, en France, conduites à Lusignan à partir de la fin des années 1980. Le terme de digestibilité est dans ce texte (et dans nombre de textes) considéré comme synonyme de dégradabilité, même si le premier se référerait plutôt à une digestion dans un animal, et le second à une dégradation par des microorganismes, y compris en fermenteur. Les deux lignées F271 et F288 ont été parmi les premières pour lesquelles des caractéristiques bien différentes en terme de dégradabilité des parois et de teneur en lignines aient été mises en évidence. Une famille recombinante de 131 RILs a alors été produite dans la descendance de leur croisement. Les essais de descendance en valeur propre et en topcross en plante entière ont été réalisés au cours des trois années 1998, 1999, et 2000.

La carte génétique obtenue en 2000 avec 108 marqueurs (Barrière et al, 2001) s'est avérée comporter de grandes régions monomorphes, une situation probablement liée en partie à une ascendance commune Co125 des deux lignées parentales F271 (50 % Co125) et F288 (25 % Co125). Le chromosome 10 est apparu très similaire entre les lignées F271 et F288, avec la mise en évidence de marqueurs polymorphes seulement dans les bins 10.06 et 10.07. Sept grandes régions monomorphes de plus de 60 cM de long, avaient été également trouvées sur les chromosomes 2, 5, 7, et 9. La consanguinité entre les lignées F271 et F288 serait ainsi proche de 2,5/8, plus du double de la valeur attendue 1/8. Les efforts de sélection auraient ainsi favorisé les allèles de Co125, lignée apportant une production en grain (et en plante entière) plus élevée. En conséquence, la recherche des QTLs dans les RILs F288 x F271 ne s'est faite que dans 70 à 75% du génome, mais cependant dans des régions fortement impliquées dans la variabilité des caractéristiques pariétales en raison des fortes différences observées pour la dégradabilité des parois entre les deux lignées.

Les analyses de variances des données en valeur propre et en topcross ont montré que les effets génotype étaient hautement significatifs ($P < 0,001$) pour les deux caractères de teneurs en lignines (ADL/NDF et KL/NDF) et de dégradabilité des parois (IVNDFD et DINAGZ), et beaucoup plus élevés que les effets d'interaction génotype x environnement (Roussel et al, 2002). Dans les essais en valeur propre, les teneurs en ADL/NDF varient de 4,8 à 7,8 % (5,4 et 7,6 % chez les parents) et les dégradabilités IVNDFD et DINAGZ respectivement de 25,4 à 39,0 % et de 45,7 à 55,7 % pour des valeurs parentales de 23,0 et 36,9 % pour IVNDFD et de 43,6 et 53,8 % pour DINAGZ, ce qui correspond à une gamme de variation fort importante pour tous ces caractères.

Les recherches de QTLs réalisées, sur les moyennes multiannuelles en raison des faibles interactions génotype x environnement, ont mis en évidence un cluster de QTLs dans le bin 6.06, avec

Tableau 1 - QTLs de teneurs en lignines et de dégradabilité des parois mis en évidence dans la famille RIL F288 x F271, en valeur propre et en topcross par F286 (Roussel et al, 2002).

QTLs	Essais valeur propre					Essais topcross				
	chr-pos	bin	intervalle support	LOD	R ²	chr-pos	bin	intervalle support	LOD	R ²
DINAGZ	1-64	1.02	32-100	2.1	7.5	-	-	-	-	-
IVNDFD	1-92	1.02	82-118	3.1	10.3	-	-	-	-	-
KL/NDF	1-108	1.02	86-120	3.6	11.8	-	-	-	-	-
ADL/NDF	-	-	-	-	-	2-126	2.08	120-142	3.1	10.3
DINAGZ	2-130	2.08	120-146	2.0	6.7	2-132	2.08	120-142	3.6	12.0
IVNDFD	-	-	-	-	-	2-166	2.08	152-182	3.6	12.8
IVNDFD	3-136	3.05	132-150	2.0	6.6	3-142	3.05	134-174	2.4	8.0
ADL/NDF	3-140	3.05	132-150	3.2	10.5	3-144	3.06	134-166	2.6	8.6
DINAGZ	3-142	3.05	136-150	6.6	20.7	3-146	3.06	134-168	2.7	8.9
DINAGZ	4-210	4.09	206-216	2.8	9.3	-	-	-	-	-
ADL/NDF	6-20	6.01	10-24	2.1	7.2	-	-	-	-	-
DINAGZ	6-20	6.01	12-24	4.2	13.8	6-14	6.01	6-24	2.0	6.7
DINAGZ	6-182	6.06	158-192	8.6	26.0	6-164	6.06	150-196	6.8	21.2
ADL/NDF	6-184	6.06	162-194	6.5	20.4	6-188	6.06	172-198	5.8	18.5
KL/NDF	6-184	6.06	158-192	8.6	26.1	-	-	-	-	-
IVNDFD	6-184	6.06	160-190	14.6	40.2	6-186	6.06	158-194	7.2	22.4
ADL/NDF	-	-	-	-	-	7-214	7.04	208-244	2.6	8.7
DINAGZ	-	-	-	-	-	8-38	8.05	26-44	3.6	12.0
IVNDFD	9-100	9.02	78-102	4.1	13.4	9-126	9.02	108-136	2.4	8.2
KL/NDF	9-100	9.02	66-102	3.6	12.0	9-130	9.02	122-144	3.8	12.3
ADL/NDF	9-110	9.02	98-122	3.3	11.0	-	-	-	-	-
DINAGZ	9-112	9.02	68-134	2.0	6.6	-	-	-	-	-

Tableau 2 - Analyses de variances marqueur par marqueur, et en combinant les deux marqueurs bnlg345 et phi089, pour les valeurs propres.

Modèle	Effet marqueur	ADL/NDF		KL/NDF		DINAGZ		IVNDFD	
		F	P (%)	F	P (%)	F	P (%)	F	P (%)
bnlg1702	bnlg1702	1.5	22.12	3.0	7.97	6.6	1.08	4.7	2.94
bnlg345	bnlg345	21.3	0.01	30.3	0.00	27.6	0.00	52.7	0.00
phi089	phi089	11.4	0.14	14.5	0.05	13.0	0.08	19.0	0.01
umc1127	umc1127	5.5	1.90	6.7	1.04	5.2	2.31	9.1	0.34
phi089+bnlg345	phi089	12.7	0.09	19.4	0.01	13.7	0.06	25.4	0.00
phi089+bnlg345	bnlg345	13.0	0.08	14.5	0.05	18.6	0.01	32.1	0.00

des pourcentages de variation phénotypique expliquée supérieurs à 20 %, et atteignant 40 % pour la dégradabilité des parois IVNDFD en valeur propre (Tableau 1). L'intervalle support de ces QTLs à effets majeurs s'étend sur plus de 30 cM correspondant à une région faisant environ 12 Mpb chez la lignée de référence B73, avec plus de 800 gènes présents (Schnable et al, 2009, www.maizesequence.org, release v2 5b60).

En vue de l'identification du ou des gènes explicatifs de ces QTLs majeurs, les premiers travaux présentés dans ce chapitre correspondent à des séquençages ciblés de gènes candidats sous-jacents ces QTLs à effets majeurs chez F288 et F271, approche complémentaire aux études transcriptomiques présentées au chapitre 1. Suite à cette recherche de polymorphisme allélique, nous avons procédé à une densification de la carte génétique de la famille RIL F288 x F271 afin (1) de réduire la taille de l'intervalle support des QTLs, (2) d'établir si ces QTLs à effets majeurs correspondent à un seul QTL ou bien à plusieurs QTLs liés avec des effets moins importants et (3) de rechercher les gènes candidats putatifs sous ces nouveaux QTLs en s'appuyant sur le génome de B73.

Ce travail est détaillé dans la publication qui suit « Targeted Linkage map densification improved positions of maize cell wall related qtls, puzzle out ghost QTLs, and highlighted significant epistatic QTL interactions » soumise à la revue *Theoretical and Applied Genetics*.

2.2. Résultats

De façon inattendue, le séquençage d'allèles de gènes putativement impliqués dans la formation des parois secondaires lignifiées sous-jacents aux QTLs majeurs du bin 6.06 a révélé que les gènes séquencés couvrant environ 50 % de l'intervalle support incluant les positions des QTLs étaient tous monomorphes entre les lignées parentales F288 et F271, suggérant que ces QTLs soient en fait des QTLs "fantômes".

Suite à ces résultats, nous avons réalisé une analyse de variance "marqueur par marqueur", non présentée dans la publication correspondant à ce chapitre. Cette analyse a montré un effet significatif important du marqueur bnl345, situé au dessus de la position estimée des QTLs, et un effet significatif, mais un peu plus faible du marqueur phi089, situé en dessous de la position estimée des QTLs (Tableau 2). Les marqueurs situés au dessus de bnl345 et au dessous de phi089 n'avaient pas d'effets significatifs sur les caractères. Les effets demeuraient significatifs pour chacun des marqueurs lors d'une analyse combinée avec les deux marqueurs, avec un effet plus fort de nouveau pour bnl345, sauf pour les lignines Klason (Tableau 2). Ces résultats laissaient supposer l'existence de deux positions de QTLs, au dessus et au dessous de la position actuellement estimée. Ils ont aussi renforcé la nécessité d'une nouvelle étape de densification en marqueurs de cette zone.

De nouveaux marqueurs ont ainsi été ajoutés dans le bin 6.06 mais également au niveau des autres régions présentant des QTLs, et dans plusieurs régions pauvres en marqueurs permettant d'établir une nouvelle carte génétique comportant 173 marqueurs. Plusieurs régions de grande taille demeurent sans marqueurs, aucun marqueur polymorphe entre F288 et F271 n'ayant été trouvé et laissant penser que ces régions sont monomorphes, probablement en liaison avec l'ascendance commune Co125. L'étude des lignées ancestrales de F288 (F186, Co125 et F188) et de F271 (Co125, W103) avait été prévue afin de confirmer l'origine du monomorphisme, mais n'a pas pu être réalisée dans le cadre de ce travail, faute de temps. Toutefois, nous avons pu montrer que dans la région polymorphe située à la base du bin 6.05, l'allèle défavorable à la dégradabilité des parois porté par F271 a pour origine la lignée parentale W103 et non pas la lignée Co125.

En s'appuyant sur la nouvelle carte génétique, 43 QTLs ont été détectés pour la teneur en lignines (ADL/NDF, KL/NDF) et la dégradabilité des parois (DINAGZ, IVNDFD) en prenant en compte les valeurs propres et topcross (Tableau 2 du projet de publication), tandis que seulement 32 avaient été mis en évidence avec les mêmes données phénotypiques en 2001. Cette différence de 11 QTLs n'est pas discutée dans le projet de publication mais correspond à la détection de 19 nouveaux QTLs et à l'absence de détection de huit QTLs détectés en 2001. Les derniers correspondaient à des QTLs à effets faibles. Parmi les nouveaux QTLs, quatre correspondent au fractionnement de QTLs détectés en 2001, les 15 autres QTLs sont quant à eux nouvellement détectés. En plus de la détection de nouvelles régions impliqués dans la dégradabilité et la teneur en lignines, la densification de la carte a permis d'affiner les positions de nombreux QTLs mis en évidence en 2001 comme c'est le cas pour les bins 1.01, 2.08, 3.05 et 6.06. Les régions avec des QTLs que je présente par la suite dans ce paragraphe ne sont pas présentées dans le projet de publication, qui est plus axé sur la présentation des deux régions majeures, bins 6.06 et 4.09, mais ces résultats soulignent l'importance d'avoir une carte génétique suffisamment densifiée pour rechercher les régions impliquées dans des caractères quantitatifs. Ainsi, les QTLs du bin 1.01 sont détectés en amont de la position initiale et s'étendent sur 38 cM contre 88 cM en 2001. Les QTLs du bin 2.08 sont détectés en aval de la position initiale avec des R^2 bien plus élevés (entre 12 et 19 % avec la nouvelle carte) et s'étendent comme auparavant sur 16 cM. Les QTLs du bin 3.05 semblent se répartir sur deux régions : le QTL de DINAGZ en valeur propre est ainsi fractionné en deux régions distantes de 1,8 cM, dont la somme des valeurs LOD équivaut environ à la valeur LOD de 2001. De façon inattendue, il y a moins de QTLs détectés (trois sur six) sur le chromosome 9, mais les trois QTLs restants ont des R^2 plus élevés. Ces QTLs se trouvent dans la région centromérique, avec des difficultés de marquage. Le QTL isolé de DINAGZ en topcross du chromosome 8 est confirmé, avec une valeur de LOD plus élevée, tandis que celui isolé d'ADL/NDF en topcross du chromosome 7 est confirmé avec une valeur LOD similaire. Sur le chromosome 1, une nouvelle position de QTLs est mise en évidence à l'extrémité distale en plus de ceux qui sont trouvés

dans la partie aval. Un nouveau QTL isolé, ADL/NDF en valeur propre, est également détecté sur le chromosome 5.

Comme cela est présenté dans le projet de publication, les QTLs "fantômes" du bin 6.06 se répartissent désormais sur deux positions, une en amont (bin 6.05) et l'autre en aval (bin 6.07) de la région considérée comme monomorphe. Le locus du bin 6.05 en amont du bin 6.06 présente les QTLs majeurs, qui expliquent de 37 % à 59 % de la variation pour les valeurs propres. Ces QTLs majeurs s'étendent maintenant sur seulement 6 cM, distance génétique correspondant à une distance physique de 2,2 Mpb (151,1 à 153,3 Mpb) chez la lignée de référence B73. Quarante-deux gènes (ou même seulement 50 si on ne prend pas en compte le QTL de KL/NDF en valeur propre légèrement décalé vers le bas) sont présents sous cet intervalle support de 2,2 Mpb chez la lignée B73. Parmi ces gènes, 39 % sont de fonction inconnue. Par ailleurs, neuf sont potentiellement impliqués dans la formation de la paroi secondaire lignifiée. Il s'agit des gènes *bHLH* (GRMZM2G061906), *FKBP* (GRMZM2G035922), *laccase* (GRMZM2G146152), *fascicline* (AC213621.5_FG004), *zinc finger de type C2H2* (AC206217.2_FG006), et *zinc finger de type C3HC4* (GRMZM2G075782 et GRMZM2G035601), mais aussi *NF-YB* (GRMZM5G804893) et *WRKY* (GRMZM2G169966).

Huit QTLs ont été détectés dans le bin 4.09, alors que seulement un QTL, expliquant 9% de la variation DINAGZ, avait été trouvé sur la carte de 2001. Ces QTLs sont répartis en trois régions successives non chevauchantes. La position centrale correspond aux QTL ayant les valeurs de R^2 les plus élevés, comprises entre 15 et 20 % (pour les QTLs en valeur propre). Cette région s'étend sur 10 cM, correspondant à 17 Mpb (entre 216,4 et 233,3 Mpb). D'importantes interactions épistatiques ont également été détectées entre ces nouveaux QTLs à effets majeurs (position 166 Mpb) et ceux du bin 6.05 pour les deux caractères dégradabilité des parois (IVNDFD et DINAGZ) et teneurs en lignines (ADL/NDF et KL/NDF). Cinq cent treize gènes sont sous-jacents aux QTLs majeurs du bin 4.09 dans le génome de la lignée B73. Parmi ces gènes, trois sont directement liées à la biosynthèse ou la mise en place de la paroi secondaire. Ces gènes sont le gène *ZmMYB42* (GRMZM2G419239), le gène *COVI-like* (GRMZM2G123790) et le gène *PAL-like* (GRMZM2G153871).

2.3. Conclusions et perspectives

La densification ciblée de la carte, en s'appuyant en particulier sur le génome de B73, a permis de préciser les positions de QTLs, d'augmenter les valeurs de LOD, et de réduire les intervalles support de ces QTLs. La densification de la carte a également permis de révéler de nouveaux QTLs dans des régions où seulement un QTL isolé était présent. Cette stratégie a aussi permis de démontrer la présence de plusieurs QTLs en cluster sous les QTLs à effets très forts préalablement mis en évidence. La réduction de la taille des intervalles support limite le nombre de gènes candidats et simplifie donc

les validations ultérieures de ces candidats (par transgénèse chez le maïs, *Arabidopsis*, *Brachypodium* ou *Setaria*, par transposon tagging ou encore par clonage positionnel).

A ce jour, l'addition de nouveaux marqueurs, sauf dans le cas de quelques localisations particulières comme la région des bins 1.01, 3.05/06, et 9.02/03, ne devrait pas permettre d'apporter plus d'informations en terme de QTLs, le facteur limitant essentiel étant maintenant l'effectif de la famille recombinante et le nombre de recombinaisons. En revanche, la réduction des intervalles support permet d'envisager, après la réalisation de banques BACs des lignées parentales, un séquençage ciblé des régions présentant des QTLs majeurs. Il est en effet important de ne pas se restreindre au génome de référence de B73 mais de se focaliser sur les lignées parentales en raison de l'existence fréquente de PAV (presence - absence variation) et de CNV (copy number variation), en plus des variations classiques en SNP et INDELS. Cette démarche sera présentée dans le chapitre 4 pour les QTLs du bin 6.05.

En parallèle, des étapes de cartographie fine et de validation de QTLs sont néanmoins inévitables. En effet, dû au faible effectif de la famille RILs, on ne peut pas exclure que les intervalles support des QTLs soient sous-estimés, conduisant à exclure des gènes clés comme le ZmMYB de type EgMYB1/Hv5 situé un peu au dessus de l'intervalle support pris en compte. Une autre situation défavorable, en l'état actuel du travail, serait que les variations soient liées à des facteurs situés de part et d'autre des positions estimées des QTLs, les QTLs détectés avec la nouvelle carte étant de nouveaux QTLs fantômes, correspondant encore à une estimation de position médiane incorrecte entre deux positions vraies. On pourrait par exemple faire l'hypothèse que les protéines codées par les gènes ZmMYB de type EgMYB1/Hv5 et Ovate situés de part et d'autre des intervalles support des QTLs du bin 6.05 puissent interagir entre elles, mais ceci reste à vérifier expérimentalement. Seules des démarches conjointes de clonage positionnel et de séquençage ciblé de BAC permettront de découvrir les déterminants de ces QTLs qui ont pourtant des effets majeurs.

Article 2 - Targeted linkage map densification improved positions of maize cell wall related qtls, puzzle out ghost QTLs, and highlighted significant epistatic QTL interactions (soumis à TAG)

Audrey Courtial^{1,2,3}, Justine Thomas^{1,2,3}, Matthieu Reymond⁴, Valérie Méchin⁴, Jacqueline Grima-Pettenati^{2,3} and Yves Barrière^{1*}.

¹ INRA, Unité de Génétique et d'Amélioration des Plantes Fourragères, 86600 Lusignan, France

² Université de Toulouse ; UPS ; UMR 5546, Laboratoire de Recherche en Sciences Végétales ; BP 42617 Auzeville, F-31326 Castanet-Tolosan, France

³ CNRS ; UMR 5546 ; BP 42617, F-31326 Castanet-Tolosan, France

⁴ INRA, Institut Jean-Pierre Bourgin, 78026 Versailles, France.

* corresponding author: E-mail: yves.barriere@lusignan.inra.fr

Key words: maize, degradability, lignin, cell wall, QTL

Abstract

Several QTLs for cell wall degradability and lignin content have been previously detected in the F288 x F271 maize RIL progeny, including a set of major QTLs located in bin 6.06 (Roussel et al. 2002). Unexpectedly, allelic sequencing of genes putatively involved in cell wall biosynthesis and assembly located within or at the close vicinity of the bin 6.06 QTL support intervals revealed that QTLs estimated positions were overlapping a monomorphous region, suggesting that these QTLs were likely "ghost" QTLs. Refining the positions of all QTLs detected in this population was thus considered, based on a linkage map densification in most important QTL regions, especially in bin 6.06 and in several large still unmarked regions. Putative ghost QTLs were then fractionating into two QTL positions located upstream and downstream the monomorphic region. The locus upstream bin 6.06 carried the major QTLs, explaining from 37% to 59% of the variation for *per se* values and extended on only 6 cM, corresponding to a physical distance of 2.2 Mbp. Among the 92 genes present in the corresponding area of the B73 maize reference genome, nine could putatively be considered as involved in the formation of the secondary cell wall [*bHLH*, *FKBP*, *laccase*, *fasciclin*, *zinc finger C2H2-type* and *C3HC4-type (two genes)*, *NF-YB*, and *WRKY*]. In addition, eight QTLs were detected in bin 4.09 while only one QTL was highlighted in the previous study. Among them, *per se* QTLs for all investigated traits colocalized in the same region and presented the highest R^2 values after those observed in bin 6.05. Moreover, these QTLs showed significant epistatic interaction with the major QTLs located in bin 6.05. Three genes related to secondary cell wall underlay QTL support intervals in this newly bin 4.09 identified region (*ZmMYB42*, *COV1-like*, *PAL-like*). This study, which aimed at

refining QTLs for cell wall degradability and lignin contents, is the first step towards the identification of suitable genes for marker-assisted selection, and will lead to the improvement of cell wall degradability for silage use or second generation bioethanol production.

Introduction

The currently decreasing reserves of fossil fuels facing the constant needs of people and industry for energy have stimulated a crucial interest in finding alternative and renewable energy resources. Lignocelluloses feedstock could significantly contribute to a sustainable supply of fuels (and also chemicals). During the previous decade, research efforts mainly focused on the bioconversion of lignocelluloses to produce bioethanol. The corresponding industrial processes generally begin by physical and/or chemical pre-treatment(s) in order to break down the recalcitrant structures of the lignified cell wall material. If technological improvements of pretreatments are essential to enhance enzymatic hydrolysis of cellulose, breeding plants with increased susceptibility of their cell walls to enzymatic hydrolysis is also effective and represents an alternative way towards commercializing profitable cellulosic ethanol. It is worth noting that since nearly 30 years, researches have been conducted to improve the digestibility/degradability of the lignocellulose cell walls of forage crops by ruminants in order to increase their energy value. Most bioenergy resources and forage plants are grasses, either C4 photosynthesis plants such as maize and panics, or C3 plants such as ryegrass or fescue. In addition to cropping devoted to bioethanol or methane production, straw of grain cereals is a significant resource for bioenergy production. However, for historical reasons mainly because cellulose from trees has been used since a long time in the paper pulping industry, more studies related to genetics, genomics and biochemistry of lignins have been performed on dicotyledonous species including the *Arabidopsis* model system.

The grass cell wall is a specific composite material including phenolic compounds, cellulose microfibrils, and an amorphous matrix consisting predominantly of glucurono-arabinoxylans. Phenolic compounds are comprised of lignins and cell wall-linked *p*-hydroxycinnamates, *p*-coumaric (*p*CA) and ferulic (FA) acid derivatives. Grass lignins include guaiacyl (G) units derived from coniferyl alcohol, syringyl (S) units derived from sinapyl alcohol, together with *p*-hydroxyphenyl units (H) derived from *p*-coumaryl alcohol. The low, but appreciable amount of H units, nearly five times higher than in dicotyledonous plants, significantly impacts the properties of grass cell walls as these units increase the frequency of resistant inter-unit bonds. The participation of *p*-hydroxycinnamates in the lignified cell wall is specific to grass species, and confers its original structure and properties. In the grass cell wall, a large proportion of S units are acylated by *p*CA, and extensive cross-linkages occur between feruloylated arabinoxylans and G units of lignins, as well as between arabinoxylan chains after ferulate

dimerization in mature cell walls (Ralph et al. 1992, 1995; Jacquet et al. 1995; Grabber et al. 2004; Ralph 2010). Lignins are essential for structural integrity of tissues and they impart hydrophobicity to vascular elements. Their association with other matrix components, together with the occurrence of linkages with and between cell wall carbohydrates, significantly impedes tissue properties towards higher stiffness and lower polysaccharide degradability, with negative effects on bioethanol production and also on silage energy value.

Interestingly, large ranges of genetic variation for cell wall degradability were shown in maize, opening ways for an efficient breeding for silage and/or for biofuel productions (Barrière et al. 2004a; Riboulet et al. 2008b; Barrière et al. 2009a). Cell wall degradability is the result of the combined effects of the cell wall composition in phenolic compounds as well as of the structural organization of lignified tissues. Therefore, the identification of genes involved in secondary wall formation and assembly deserves priority interest for maize biomass quality improvement. In addition, plant breeders have to choose cell wall trait combinations for minimizing negative effects on genotype agronomic value, including whole plant yield, standability, biotic and abiotic stress tolerance.

The characterization of maize mutants and/or of genetically engineered plants have highlighted a few genes able to affect maize cell wall degradability. It is the case of genes involved in monolignol biosynthesis (*CAD*, *COMT*, *CCR*; Vignols et al. 1995; Halpin et al. 1998; He et al. 2003; Barrière et al. 2004c; Pichon et al. 2006; Tamasloukht et al. 2011) as well as transcription factors of the MYB family regulating lignin biosynthesis (*ZmMYB31*, *ZmMYB42*; Fornalé et al. 2006; Sonbol et al. 2009; Fornalé et al. 2010). However, the involvement of these genes in the natural degradability variation between maize lines has not yet been established. In addition, the observed cell wall degradability improvements in these mutants often occurred together with negative effects on plant agronomic value. Efficient breeding of maize with higher cell wall degradability for silage use or second generation bioethanol production therefore demands the identification of the major determinants driving traits under selection. The search for candidate genes underlying QTLs for cell wall degradability and related traits is thus, one relevant strategy for further key-gene discovery and consequently efficient marker-assisted selection. QTLs for cell wall-related traits have been mostly detected in maize (reviewed in Barrière et al. 2007; 2009a; and references therein; Barrière et al. 2010b; 2012) and to a lesser extent in woody species such as pine (Sewell et al. 2002; Markussen et al. 2003; Pot et al. 2006), poplar (Yin et al. 2010), and *Eucalyptus* (Freeman et al. 2009; Thumma et al. 2010; Gion et al. 2011). However, no gene has yet been identified as being responsible for the effect of any cell wall QTL, even if colocalizations between candidate genes and QTLs were found in maize, *Arabidopsis*, poplar, and *Eucalyptus* (Barrière et al. 2010b; Ranjan et al. 2010; Thomas et al. 2010; Gion et al. 2011; Chavigneau et al. 2012). This is mainly due our current ignorance of the major genetic determinants of cell wall biosynthesis and assembly, even if exhaustive lists of candidate genes involved either in the biosynthesis of phenolic compounds or in the formation and assembly of

secondary walls as well as in the regulation of these processes have been proposed (Barrière et al. 2009a; Chavigneau et al. 2012). The frequent large size of the QTL support intervals, in average length 20 cM corresponding to 15 to 50 Mbp according to genomic location and recombination rate, is the second reason which render difficult the identification and validation of candidate genes.

In the F288 x F271 early maize RIL progeny, several QTLs were previously mapped, of which those located in bin 6.06 explained a highly significant part of the phenotypic variation for both lignin content and cell wall degradability, with R^2 values ranging from 20 to 40 % (Roussel et al. 2002; Thomas et al. 2010). However, the support intervals of these QTLs represent more than 20cM, which correspond to nearly 12 Mbp and more than 800 genes in the B73 reference genome.

Allelic variation of putative candidate genes in bin 6.06 was investigated as a first step towards the identification of the possible involvement of these genes in the effect of the detected QTLs. In addition, linkage map densification was performed to (i) reduce the size of QTL support intervals, (ii) to investigate whether the large effect QTLs corresponded to a single major QTLs or to closely linked QTLs with lower effects, and then to (iii) identify putative candidate genes taking advantage of the maize B73 genomic sequence (Schnable et al. 2009, www.maizesequence.org, release v2 5b60) and information on gene physical positions. Markers were therefore targeted within the QTL support intervals together with a marker densification on the whole F288 x F271 genetic map, especially in genetic regions with low marker density.

Material and methods

RIL production and RIL experiments

The set of 131 RIL was developed by single seed descent from the cross between the two early dent inbred lines F288 and F271 at INRA Lusignan, France (Barrière et al. 2001). F271 and F288 have low and medium-high cell wall degradability, respectively. F271 and F288 lines have both a Co125 common ancestry [F271 = (Co125 x W103), F288 = (F244 x F252) with F244 = (F186 x F188) and F252 = (F186 x Co125)]. As reported in Roussel et al. (2002), RIL progenies were evaluated in field experiments for their *per se* values in seven environments (three locations over 3 years), and for topcross experiments with F286 as flint tester of high cell wall degradability line in six environments (three locations over two years). Topcross and RIL *per se* were evaluated in generalized alpha-lattice designs including both parents with, in each location, three replicates for the tested RIL and nine replicates for the parents. Each experimental plot was a 5.2 m long single row of 37 plants. Row spacing was 0.75 m, and the resulting density was 95,000 plants/ha. Irrigation was applied in Lusignan during summer to prevent water stress. At the silage harvest stage [about 30 to 35% of whole plant dry

S1_table - Primers used for allele sequencing. Design was performed for amplified a sequence of nearly thousand base-pair based on the B73 genomic sequence (Schnable et al. 2009, www.maizesequence.org, release v2 5b60).

Gene Ids	Primer pair number	Forward primer	Reverse primer	Amplified genotype
GRMZM2G140817 (C3'H)	C1	GAACAGAATAGCGCAAAGGA	GGAAAGGAAGAGTGGAGAG	F288 and F271
	C2	GCTGAAAGAAAGTGGGAAGA	CAAATTAGCTGAAAACGCAAC	F288 and F271
	C3	AGCCAAACTTTATGTCAACC	TCCCATGCAATTTACTACCT	F288 and F271
	C4	TGTAGTTCCCTTGTGCAATC	ATTGGCTTTCCTTCATTACC	F288 and F271
	C5	TGCTCCATTAATGTATTTCA	TGGAAGATTAGTACTCGCAG	F288 and F271
	C6	AGGTGGGATTTGGAACTAT	ACTCAAGAACAGACCGGTAA	F288 and F271
	C7	TGCGTTTATATTCATTTACA	TCTATTTTCGATGGACTCTTT	F288 and F271
GRMZM2G101533 (COV1)	C1	TGCAGCTCGCCATGGATTAG	GACTCACTCCGCGCTCACAC	F288 and F271
	C2	GTGTGAGCGCGAGTGAGTC	CCCAACTGCGGACAACCTTA	F288 and F271
	C3	ACGCCGAGAGACTGTGACCC	AGAGTCATGCCAGGTTGC	F288 and F271
	C4	GATAAGAGTCACGCCACAGG	GATGTGACAAAGCCAAGACC	F288 and F271
	C5	GGACAAGCCTGGCATAGTGG	ATCTGCGGGTCGAGGGTTGA	F288 and F271
	C6	TGCCCAAATTCTGTCAACC	CTTCCGTACGCGCTGCTCT	F288 and F271
	C7	ACATGGATGGCTACTCCAGC	TGCAGCTAACCATCGACAAG	F288 and F271
	C8	GAAGTCTTCCCGAGGTGACA	GTTAATGAGGTGGCAGGTGG	F288 and F271
GRMZM2G077789 (MYB Hv5-like)	C1	CCGCTACCAACGCTCCATCA	CCCAGCAGCCAGAAAGCAA	F288 and F271
	C2	CCGGAACACCGTTCAGA	CGACTCTCCTCTCGGTTCTG	F271
	C3	ACAAGTCAGTGAGCGTGTGC	CGACTCTCCTCTCGGTTCTG	F271
	C4	ACCTCAACCTCGACCTGTCCGT	TGCTTGCATGCTGCCTGAGC	F288 and F271
	C5	AAAACGGCGGGTCAAAGAA	GGTTGTCACCTGTCAGCTCC	F288 and F271
	C6	CCGGAACACCGTTCAGA	CCCATATGAGAGACTAAAGCCAA	F288
	C7	CAGCAGCGGTTTCAGATTTTC	TCCCTCTAACCAAACACCTC	F288
GRMZM2G465835	C1	CGAGTACAAGGCTTTCGAGG	AACGTACGGGGCCTATCAAT	F288 and F271
GRMZM2G073415	C1	CAATGTTTCAGGCCTTGGAT	CCTTTTCACTGCCCCACTAA	F288 and F271
GRMZM2G066225	C1	ATGGACCAAGAAACCTCACG	ACCCTGAAAATCTGTGTGGC	F288 and F271
GRMZM2G061906	C1	ATGTCCACCGAGTATGACCA	TTGCAGGCCACAATGTTAGC	F288 and F271
GRMZM2G061906	C2	CTTCTCGCTGGATCGTCTTT	TCGCTTAGGATAGCGACTTTG	F288 and F271
AC206217.2_FGT006	C1	AGACGTGCAGGAGGAGGTT	GCTCATAAGCAGATGCACGA	F288 and F271
GRMZM2G035601	C1	TGGTAAGAGTGCCTCAATGT	AACTGAGGAGAGCACACCAA	F288 and F271
GRMZM2G169966	C1	GCAGGTCAAAGAAACCATCC	GGAAGACGTCCAAGTTTCCA	F288 and F271
GRMZM2G390436	C1	AACTCGCTGGCCAGGTTTC	GTAAGGCACTCGGTGGTTGT	F288 and F271
GRMZM2G701218	C1	AAACAGAATTTGGGTCCTGC	GGGTGCCCAAAAGAGAGG	F288 and F271
GRMZM2G157246	C1	ATCAATGCCTTCAACCAAGC	ATATCTGTACTCCGTGGCCG	F288 and F271
GRMZM2G117854	C1	ACTTCCGGTTACAGCCTGC	GGTGTAGCTCCTGAAGAAATATGA	F288 and F271
GRMZM2G159741	C1	AAAACCTTGTGCAGTGTGCG	TGGTAGCACAAGCGTTGAAG	F288 and F271
GRMZM2G175870	C1	TGCTTCTTGTGCTGGTGTTC	TCAGGAGGAATCAGGAATGG	F288 and F271
AC226373.2_FGT010	C1	ATGCCTTGATGTTGGAAAAGG	AGTCTGCAAAAACCAACCAGG	F288 and F271
GRMZM2G441325	C1	CAAGGCTCTTATGGCAGGAG	GGCATGCAAAAACAAACACAG	F288 and F271
GRMZM2G140996	C1	GTTTGAGAGCATTCCACCG	TGGCCTCCAACAACCATATC	F288 and F271
GRMZM2G141026	C1	ATGTTTCAGAGCATCCCAGC	AACCGAAACAACATGGCTC	F288 and F271
GRMZM2G401521	C1	TCAAAATCATGCATAGCCCA	CGGCTTCATTACAAAGGACG	F288 and F271
GRMZM2G074427	C1	AGCAGACCTGCAAGAGAAGC	ACGAACTTAAGCGTCGTGCT	F288 and F271
GRMZM2G175232	C1	GCCCACCATCTCTCAGAAAC	GCGTGGATCAATTCGCTTAT	F288 and F271
GRMZM2G031200	C1	TACACACACGGTGCCTGTTC	CAAGTCCTGGATCGTCTGTG	F288 and F271

matter (DM)], the plots were machine-harvested with a forage chopper. A representative sample of 1 kg chopped material per plot was collected for further analyses.

A highly significant genetic variation was shown for four cell wall related traits including lignin content (ADL/NDF and KL/NDF) and cell wall digestibility/degradability (IVNDFD and DINAGZ), with high broad sense heritability values (Roussel et al. 2002). The corresponding individual RIL *per se* and topcross mean values over locations obtained by Roussel et al. (2002) were used as phenotypic values in the present QTL investigation. NDF (Neutral Detergent Fiber) is an estimate of cell wall content and ADL (Acid Detergent Lignin) is an estimate of lignin content according to Goering and van Soest (1970). KL (Klason lignin) is another estimate of lignin content according to Dence and Lin (1992) including an acido-soluble part of lignin which is lost during the first step of the ADL procedure (Hatfield et al. 1994; Jung et al. 1997; Hatfield and Fukushima 2005). IVNDFD (*in vitro* NDF digestibility) is an estimate of cell wall digestibility, based on the enzymatic solubility of Aufrère and Michalet-Doreau (1983), which is computed according to Struik (1983) and Dolstra and Medema (1990), assuming that the non-NDF part is fully digestible. DINAGZ is another estimate of cell wall digestibility, similarly based on the same enzymatic solubility, which is computed according to Argillier et al. (1995) and Barrière et al. (2003b), assuming that starch, crude proteins, and soluble carbohydrates are fully digestible.

Allele sequencing

Genomic DNA was isolated from young maize leaves of F271 and F288 lines using the DNeasy Plant mini kit (Qiagen). Primer pairs were designed for amplified a sequence of nearly thousand base-pair based on the B73 genomic sequence (Schnable et al. 2009, www.maizesequence.org, release v2 5b60) using primer 3 software (Supplementary table 1). PCR amplification reactions were performed in 25 µl containing 1x buffer, 4X DMSO, 200 µM of each dNTP, 0.2 µM of each 5' oligo and 3' oligo, and 0.5 unit of Platinum Taq Polymerase (Invitrogen, 10966), 30 ng of genomic DNA was used as template. The PCR cycling program consisted of an initial denaturation of 2 min at 94°C, followed by 40 cycles for 30 s at 94°C, 30 s at 60°C, and 1 min at 72°C, and a final extension for 5 min at 72°C. Sequencing was performed for each of the PCR fragments in both directions by the Millegen company (31670 Labège, France). All the sequences were aligned with the B73 sequence using the CLUSTALX2 software.

Genotypic data and development of a new linkage map

The linkage map of the F288 x F271 RIL progeny was originally drawn based on 108 SSR markers and 131 RIL (Barrière et al. 2001). New markers were added with a targeted strategy in order to increase the marker density in the most important QTL regions, especially in bin 6.06, but also to set markers in several large still unmarked regions. New considered markers were first 113 SSR markers available in the MaizeGDB database (www.maizegdb.org) with unambiguous physical positions. In

addition, 128 new markers were designed specifically in genes chosen for physical positions and preferentially considering genes involved in cell wall biosynthesis, according to the list proposed by Barrière et al. (2009a). Primer pairs were designed based on the B73 genomic sequence (Schnable et al. 2009, www.maizesequence.org) using primer 3 software (sequences available upon request). Genotyping was also reinvestigated for 6 SSR markers present on the original map, but with missing data or discrepancies between genetic and physical positions. Finally, 3 SSR markers available from investigations in 2001 (Barrière et al. 2001), but not yet included in the published genetic map, were added to the renewed map.

Gene and marker polymorphisms were analyzed with High Resolution Melting (HRM) technology using the LightCycler 480 system (Roche Applied system). Genomic DNA was isolated from young maize leaves of the two parental F271 and F288 lines and of each RIL cropped at INRA, Lusignan (France), except for a few RIL for which DNA was directly isolated from grain after observation of their null or poor germination (adapted protocol from Dellaporta et al. 1983). Heterozygous DNA was produced by extracting DNA from fresh F288 and F271 mixture plant material. HRM polymorphisms were first revealed on the DNA from parental lines and this heterozygote DNA. If convenient results were obtained, markers were genotyped on the whole RIL progeny using the two F288 and F271 lines, and the constructed heterozygote as melting standard. PCR amplification reactions were performed in 10 µl containing 1x buffer, 200 µM of each dNTP, 0.2 µM of each primer, 4X DMSO, 3mM MgCl₂, 0.15 unit of Platinum Taq DNA Polymerase (Invitrogen, 10966) and 0.25X ResoLight (Roche Diagnostics, 0490964000). Approximately 20 ng of genomic DNA was used as template. The PCR cycling program consisted of 5 min at 94°C, followed by 50 cycles for 30 s at 94°C, 1 min at 60°C, and 1 min at 72°C. The HRM reaction was performed during 1min at 95°C, cooling to 40°C for 1min, raising the temperature to 65°C and then to 95°C with 25 fluorescent acquisitions per Celsius degree at this step. The following data analysis was performed with the gene scanning software module on the LightCycler[®] 480 instrument. The linkage map was developed using CarthaGene (version 1.2.2, De Givry et al. 2005).

QTLs identification and candidate genes

QTL detection was then performed following the method of Composite Interval Mapping (CIM, Zeng 1994) implemented in the PLABQTL software (Utz and Melchinger 1996) as previously performed by Roussel et al. (2002). PLABQTL uses the regression method (Haley and Knott 1992) in combination with markers which are selected by stepwise regression as cofactors. LOD support intervals are constructed in PLABQTL according to Lander and Botstein (1989) and are considered to be underestimated in the case of CIM. The percentage of phenotypic variance ascribed to an individual QTL was estimated with the approximate standard error of Kendall and Stuart (1961). The additive effects of QTL were estimated as half the difference between the phenotypic values of the respective homozygotes. Based on the permutation-test method of Churchill and Doerge (1994), LOD thresholds

Table 1 - Polymorphism of putative lignin-related candidate genes and marker genes underlying bin 6.06 QTL [Allele sequencing based on a 1000 bp long genomic fragment, except full length sequencing if *-quoted SNP and InDels. Marker and gene positions in Mbp are given according to the B73 line genomic sequence (www.maizesequence.org) release v2 5b60. Bins are defined region of approximately 20 cM delimited by two fixed *core markers* and are designated by a chromosome number followed by a two-digit decimal (www.maizegdb.org). Gene numbers are given according the same maizesequence database. Markers bnlg1702, bnlg345, and phi089 were shown as polymorphic between the two F271 and F288 lines in 2001 and mapped in the progeny (distance in cM) (Barrière et al., 2001); "na" are unavailable data].

Markers, genes and QTLs	Gene name	Bin	Phys pos	F288/F271	
				SNP	InDel
bnlg1702 (120.2 cM)		6.05	147.1	-	-
NAC domain-containing protein	GRMZM2G465835	6.05	149.4	0	1
Continuous vascular ring like (LCV2-like)	GRMZM2G073415	6.05	150.3	5	0
F288 x F271 QTL upper limit (150.0 cM)		6.05	150.4	-	-
MYB Hv5-like EgMYB1-like R2R3-type	GRMZM2G077789	6.05	150.7	8*	4*
Zinc finger C3HC4 transcription factor	GRMZM2G066225	6.05	150.8	10	3
bHLH transcription factor	GRMZM2G061906	6.05	151.1	21	2
ZmLac6 AtLac17 ortholog	GRMZM2G146152	6.05	151.5	na	na
Zinc finger C2H2 transcription factor	AC206217.2_FGT006	6.05	151.8	3	12
Zinc finger C3HC4 transcription factor	GRMZM2G075782	6.05	152.8	na	na
Zinc finger C3HC4 transcription factor	GRMZM2G035601	6.05	152.9	1	0
WRKY transcription factor OsWRKY70-like	GRMZM2G169966	6.05	153.3	24	9
bnlg345 (177.2 cM)		6.05	153.5	-	-
Zinc finger Ring H2 transcription factor	GRMZM2G390436	6.06	154.2	8	6
MYB transcription factor SHAQKYF class	GRMZM2G701218	6.06	155.2	19	6
Zinc finger C3HC4 transcription factor	GRMZM2G157246	6.06	155.2	0	0
MYB transcription factor SHAQKYF class	GRMZM2G117854	6.06	155.5	0	0
Cytochrome P450 CYP98A1 (C3'H2)	GRMZM2G140817	6.06	155.7	0*	0*
CCAAT-HAP5 transcription factor	GRMZM2G440949	6.06	155.7	na	na
Zinc finger C2H2 transcription factor	GRMZM2G159741	6.06	156.0	0	0
bZIP transcription factor	GRMZM2G175870	6.06	156.0	0	0
F288 x F271 QTL position (181.7 cM)		6.06	156.1	-	-
Zinc finger CCCH transcription factor	AC226373.2_FGT010	6.06	156.4	0	0
Ras small GTPase ROP6-like	GRMZM2G176217	6.06	158.4	na	na
Auxin response factor 4 (ARF)	GRMZM2G441325	6.06	158.7	0	0
ZRP4-like OMT	GRMZM2G140996	6.06	158.7	0	0
ZRP4-like OMT	GRMZM2G141026	6.06	158.7	0	0
ZRP4-like OMT	GRMZM2G102863	6.06	158.8	na	na
ZRP4-like OMT	GRMZM2G124799	6.06	158.8	na	na
Continuous vascular ring (COV1-like)	GRMZM2G101533	6.06	159.8	0*	0*
WRKY transcription factor OsWRKY58-like	GRMZM2G401521	6.06	159.8	0	0
Auxin-responsive Aux/IAA OsIAA18-like	GRMZM2G074427	6.06	160.2	0	0
ZmMYB AtMYB26-like	GRMZM2G175232	6.07	162.1	7	8
F288 x F271 QTL lower limit (198.0 cM)		6.07	162.6	-	-
WD40 repeat-like	GRMZM2G038032	6.07	162.9	na	na
NAC SND2/SND3-like	GRMZM2G031200	6.07	164.5	0	0
phi089 (213.4 cM)		6.07	166.3	-	-

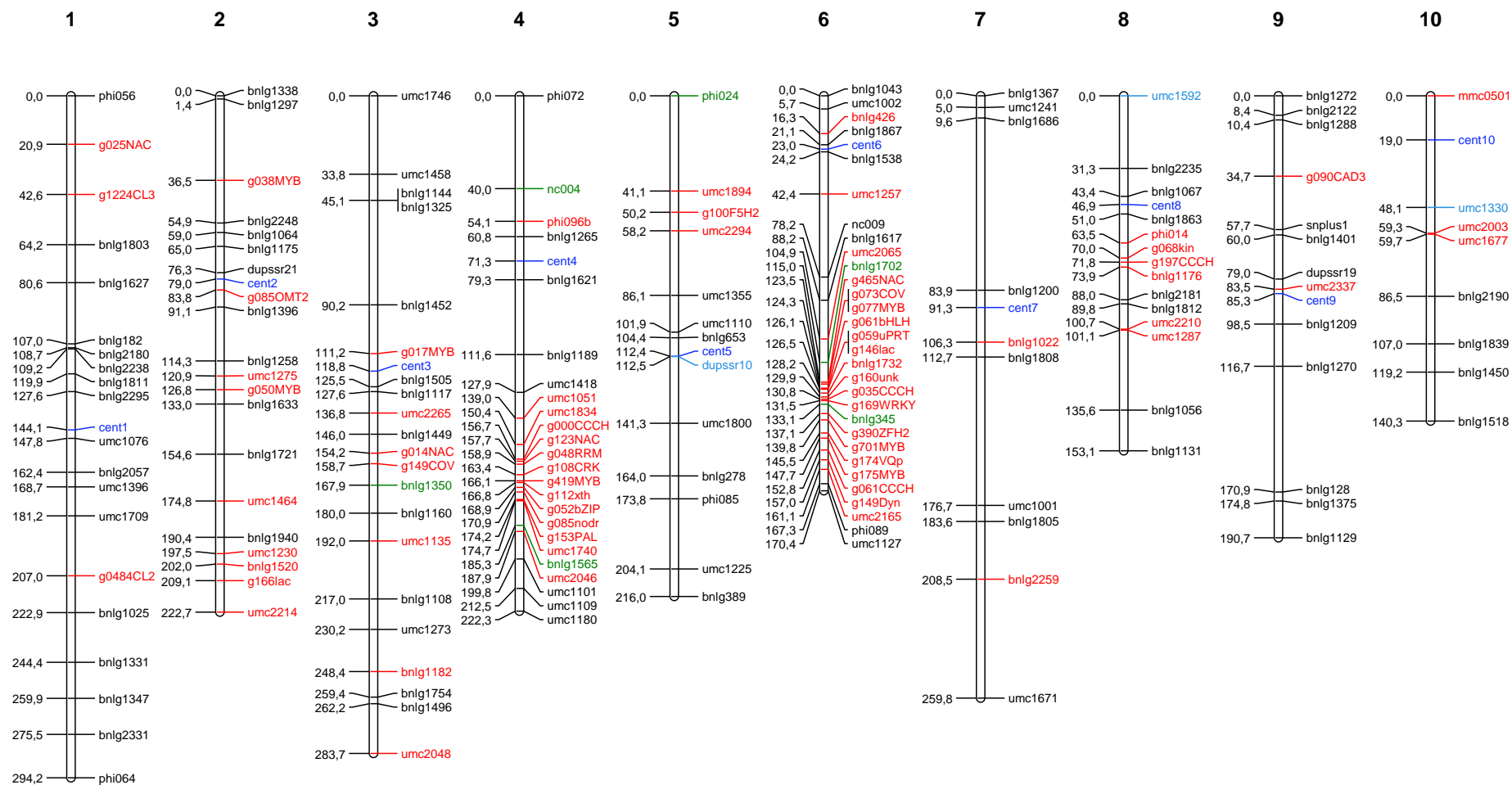
equal to, respectively 3.0, 3.5 and 4.6 allowed an experiment-wise error rates equal to 10, 5 and 1%. QTLs were finally considered for LOD thresholds higher or equal to 3.0, in order to highlight most of them. In addition, QTLs with lower LOD values were also considered when they were shown in colocalizing positions. Detection of epistatic interactions has been performed with R software using the *scantwo* procedure of the QTL library as described by Broman and Sen (2009). Physical QTL positions were estimated based on physical positions of the two flanking markers (B73 line sequence, release v2 5b60), assuming a linear relationship between recombination and physical distances within this interval. The lists of genes underlying QTLs were established according to the filtered gene set genes presented in the same reference genome (Maize Sequence database, Schnable et al. 2009).

Results

Allele sequencing reveals ghost QTLs in bin 6.06

Based on previous analyses highlighting lignin content and cell wall degradability QTLs with high R^2 values in the upstream part of bin 6.06 (Roussel et al. 2002; Thomas et al. 2010), allele sequencing was first carried out for genes putatively involved in cell wall biosynthesis and assembly within or at the close vicinity of the latter QTLs support intervals (*i.e.* from 149.4 Mbp to 164.5 Mbp; Table 1). Genomic sequence of three candidate genes considered of priority interest were full-length sequenced (nearly 500 bp before ATG and after stop codons), including the *cinnamate-3-hydroxylase C3H2* (GRMZM2G140817) involved in monolignol biosynthesis, the *MYB Hv5-like* (GRMZM2G077789) which is a R2R3 MYB repressor orthologous to barley *Hv5 MYB* genes (Wissenbach et al. 1993) and to *EgMYB1* (Legay et al. 2007; 2010), and the continuous vascular ring *COV1-like* (GRMZM2G101533), ortholog of an *Arabidopsis* gene involved in lignified tissue patterning (Parker et al. 2003). Twenty one other candidate genes were only partially sequenced (over 1000 bp length). No allelic polymorphism between the two parental lines F271 and F288 was found for the two fully sequenced *C3H2* and *COV1-like* genes (position 155.7 Mbp and 159.8 Mbp respectively). In addition, all other ten investigated genes located between the *zinc finger C3HC4* (GRMZM2G157246, position 155.2 Mbp) and the *OsIAA18-like* (GRMZM2G074427, position 160.2 Mbp), including these two genes, were also monomorphic between the two parental lines. In contrast, the ten investigated genes located upstream the *zinc finger C3HC4*, including the *MYB Hv5-like*, were polymorphic between F271 and F288. Similarly, downstream of the *OsIAA18-like* gene, the MYB *AtMYB26-like* (GRMZM2G175232, position 162.1 Mbp) was polymorphic. However, the sequence of the downstream *NAC SND2/SND3-like* (GRMZM2G031200, position 164.5 Mbp) gene was identical between both parental lines. As a tentative conclusion, no polymorphism was revealed in a 5 Mbp long region (12 investigated genes within this region), corresponding to nearly 50 % of the QTL support interval and overlapping all QTLs estimated positions (based on data of Roussel et al. 2002). This

Figure 1 - New linkage map of the F288 x F271 RIL progeny (with 173 makers). Black markers correspond to markers mapped on the previous map with 108 markers (2001 map), red markers to markers newly genotyped in 2011, light blue markers to markers genotyped in 2001 but not included in the previously published map, green markers to re-genotyped marker (data 2001 replaced by data 2011). Centromere (cent) positions were estimated based on physical positions in B73. Names on the right of the chromosome correspond to SSR names or gene names (abbreviated GRMZM2G identifier followed by annotation). Numbers on the left of the chromosome indicate marker positions in cM from the top of the chromosome.



region was therefore considered as monomorphic between the two parental lines, and consequently, detected QTLs at this locus were considered as putative "ghost" QTLs. Map densification, and generating a new linkage map, in this major QTL region, as well as in areas with low marker density, was therefore an obligatory step to precise QTL positions and to propose judicious candidate genes.

Revisiting of the polymorphism between F288 and F271 and generating a new linkage map

We tested 247 new markers of which 114 (46 %) were polymorphic between the two parental lines, 94 (38.0 %) were monomorphic, and 39 (16 %) failed during the HRM assay. Finally, 75 new polymorphic markers were analyzed on the whole RIL progeny, including the six re-investigated markers. The still unused 3 SSR markers genotyped in 2001 were added to the map, whereas 7 markers, which failed to be relevantly added to linkage groups during the map construction, were removed. The new map (Figure 1) was significantly densified with 173 markers as compared to the original map established in 2001 which only contained 108 SSR markers.

The new map, which spans a cumulative distance of 2,153 cM, shows an average distance between markers of 13.2 cM while it was 22.3 cM long on the 2001 map. The average distance between markers was however unevenly distributed, with the smallest distance (5.87 cM) found for chromosome 6 with only 2.1 cM between g465NAC et g149Dyn, and the largest for chromosome 7 (28.9 cM). Differences between chromosomes and between areas result both from the focus we gave on some areas, and also from the absence of polymorphic marker on other areas. Two large gaps, more than 60 cM long, were still present on chromosome 7, which likely correspond to large monomorphic areas. A similar situation was shown on chromosome 10 for which the first marker of the 2001 map (bnlg2190) was located in physical position 141.8 Mbp. The current investigations allowed to map mmc0501 at the upstream part of chromosome 10, in position 5.9 Mbp. Between the two markers mmc0501 and bnlg2190, only four markers out of the 34 tested were found polymorphic, one in position 8.8 Mbp and three surrounding the position 125 Mbp. These results strengthened a large monomorphic area of nearly 110 Mbp long on chromosome 10.

QTL detection

Overall, based on the new linkage map, 43 QTLs were detected for lignin content (ADL/NDF, KL/NDF) and cell wall degradability (DINAGZ, IVNDFD) traits in the F288 x F271 progeny for both *per se* and topcross experiments (Table 2), while only 32 were found in 2001. QTLs were distributed on nine chromosomes (since no QTL was detected on chromosome 10) and 17 genomic positions. Colocalizations between QTLs were observed on most positions, and only five QTLs were found in isolated positions, on chromosomes 4 (2 QTLs), 5, 7 and 8. A similar number of QTLs were mapped using data from *per se* (20 QTLs) and topcross (23 QTLs) experiments. Five QTLs were detected for ADL/NDF in *per se* experiments and six in topcross experiments. Four and six KL/NDF QTLs were detected for *per se* and topcross values, respectively. Eight QTLs for DINAGZ were observed for *per*

Table 2 - QTL analysis for cell wall investigated traits (IVNDFD: *In Vitro* Neutral Detergent Fiber Digestibility, DINAGZ: in vitro cell wall digestibility according to Argillier et al. (1995), ADL/NDF: Acid Detergent Lignin/Neutral Detergent Fiber, KL/NDF: Klason Lignin/NDF) in *per se* (ps) and topcross (tc) experiments.

LOD	exp wise	R2							
>2.35	0.25	7.9							
>3.00	0.10	10.0							
>3.49	0.05	11.6							
>4.61	0.01	15.0							
QTL	Chrom	Bin	Pos	Left_Mark	Supp.IV	LOD	R2	Add	Line +
KL/NDF tc	1	1.01	26	g025NAC	14-38	3.24	10.8	-0.117	F271
KL/NDF ps	1	1.01	32	g025NAC	24-42	5.74	18.3	-0.385	F271
DINAGZ ps	1	1.01	36	g025NAC	24-52	3.40	11.3	0.509	F288
IVNDFD tc	1	1.08	242	bnlg1025	230-254	3.28	10.9	-0.292	F271
KL/NDF tc	1	1.09	250	bnlg1331	230-260	3.00	10.1	0.114	F288
IVNDFD ps	2	2.08	194	bnlg1940	178-198	2.12	7.2	0.579	F288
KL/NDF tc	2	2.08	196	bnlg1940	190-202	4.54	14.8	-0.126	F271
ADL/NDF ps	2	2.08	198	umc1230	192-204	5.36	17.2	-0.166	F271
DINAGZ ps	2	2.09	200	umc1230	196-206	3.59	11.9	0.443	F288
ADL/NDF tc	2	2.09	202	umc1230	198-206	5.59	17.8	-0.071	F271
IVNDFD tc	2	2.09	202	umc1230	198-206	5.88	18.7	0.360	F288
DINAGZ tc	3	3.05	124	g017MYB	116-128	2.33	7.9	0.189	F288
DINAGZ ps	3	3.05	124	g017MYB	116-128	2.82	9.4	0.493	F288
ADL/NDF tc	3	3.05	134	bnlg1117	126-146	3.57	11.8	-0.060	F271
DINAGZ ps	3	3.06	142	umc2265	136-152	2.29	7.7	0.512	F288
IVNDFD tc	3	3.06	152	bnlg1449	144-160	4.03	13.2	0.309	F288
ADL/NDF tc	4	4.09	158	g123NAC	156-160	2.83	9.5	-0.049	F271
IVNDFD tc	4	4.09	158	g123NAC	156-160	4.16	13.6	0.296	F288
IVNDFD ps	4	4.09	168	g112xth	166-172	5.27	16.9	0.878	F288
KL/NDF ps	4	4.09	168	g112xth	166-172	5.79	18.4	-0.309	F271
ADL/NDF ps	4	4.09	168	g112xth	166-172	6.25	19.7	-0.180	F271
KL/NDF tc	4	4.09	168	g112xth	166-170	8.51	25.9	-0.169	F271
DINAGZ ps	4	4.09	174	g153PAL	170-176	4.62	15.0	0.528	F288
DINAGZ ps	4	4.09	198	umc2046	192-204	3.74	12.5	0.543	F288
ADL/NDF ps	5	5.08	210	umc1225	188-216	3.90	13.8	0.163	F288
DINAGZ ps	6	6.01	20	bnlg426	16-26	5.46	17.5	0.547	F288
ADL/NDF ps	6	6.01	22	bnlg1867	16-26	2.00	6.8	-0.099	F271
ADL/NDF tc	6	6.05	128	g146lac	126-130	2.38	8.0	-0.055	F271
IVNDFD tc	6	6.05	128	g146lac	126-130	2.99	10.0	0.318	F288
DINAGZ tc	6	6.05	128	g146lac	126-130	3.60	11.9	0.289	F288
ADL/NDF ps	6	6.05	128	g146lac	126-130	13.46	37.7	-0.308	F271
DINAGZ ps	6	6.05	128	g146lac	126-130	18.10	47.1	1.321	F288
IVNDFD ps	6	6.05	128	g146lac	126-130	25.04	58.5	2.533	F288
KL/NDF ps	6	6.05	130	g160unk	128-132	13.16	37.0	-0.576	F271
KL/NDF tc	6	6.06	142	g701MYB	136-146	5.08	16.3	-0.140	F271
DINAGZ tc	6	6.07	148	g175MYB	144-154	2.51	8.5	0.216	F288
ADL/NDF tc	6	6.07	148	g175MYB	144-152	4.05	13.3	-0.068	F271
IVNDFD tc	6	6.07	148	g175MYB	144-152	4.31	14.1	0.354	F288
ADL/NDF tc	7	7.03	190	bnlg1805	176-204	3.03	10.1	-0.063	F271
DINAGZ tc	8	8.04	66	phi014	56-70	5.08	16.4	0.289	F288
KL/NDF ps	9	9.03	86	umc2337	78-94	4.48	14.6	0.296	F288
IVNDFD tc	9	9.03	86	umc2337	80-94	5.47	17.5	-0.390	F271
KL/NDF tc	9	9.03	94	umc2337	86-112	5.04	16.2	0.141	F288

se values and four for topcross values. Finally, three and seven IVNDFD QTLs were detected for *per se* and topcross values, respectively. Marker densification of the genetic map allowed us to pinpoint some QTL positions and to reduce support intervals of QTLs.

Only eight QTLs out of 43 explained less than 10 % of the observed genetic variation ($LOD < 3.00$), while four QTLs explained more than 30 % of the observed variation. These major QTLs colocalized in bin 6.05, a little upstream of the revealed monomorphic region. Moreover, QTLs for all investigated traits in both *per se* and topcross experiments were observed in this region, except for the KL/NDF trait in topcross experiments. These bin 6.05 QTLs explained from 37% to 59% of the variation in lignin content and cell wall degradability observed in *per se* experiments, whereas the topcross QTLs explained only 8% to 12% of the observed variation. These bin 6.05 QTLs are all located in support intervals from 126 to 130 cM, except the *per se* value KL/NDF QTL found slightly downstream, between 128 and 132 cM. Downstream of this region and of the monomorphic area, topcross QTLs were also shown for all traits in position 148 cM (bin 6.07), except for KL/NDF QTL detected in position 142 cM. This latter position was still unexpected as it corresponded to the so-called monomorphic area. Downstream QTLs explained from 9 to 16% of the genetic variation. Finally, the seven "ghost" QTLs shown in 2001 in bin 6.06 could correspond to QTLs located in two positions, one upstream for both *per se* value and topcross observations (bin 6.05, position 151.8 Mbp), and one downstream for topcross observations (bin 6.07, position 162.1 Mbp), even if one was still located in the monomorphic area (Figure 2). The bin 6.06 QTLs, which expanded over a 48 cM long support interval corresponding to 12.2 Mbp (150.4 to 162.6 Mbp) based on the 2001 map, now only spans in its upstream part over a 6 cM long interval corresponding to 2.2 Mbp (151.1 to 153.3 Mbp) and in its downstream part over a 10 cM long interval corresponding to 3.7 Mbp (159.2 to 162.9 Mbp) if the KL/NDF topcross QTL is excluded. The upstream support interval extended only over 4 cM corresponding to a little more than 1.2 Mbp (151.1 to 152.3), if the not far downstream LK/NDF QTL was not considered.

In addition, eight QTLs were detected on chromosome 4, in bin 4.09, following the map densification (13 new markers) and reassessment in this region, while only one QTL, explaining 9 % of the DINAGZ variation, was found in 2001 (Figure 3). These QTLs were distributed in three successive non overlapping regions. One of these regions corresponded to QTLs with the highest R^2 value after those observed for bin 6.05 QTLs for all *per se* value traits (R^2 15-20 %). This region extended over 10 cM, corresponding to 17 Mbp (216.4 to 233.3 Mbp).

Furthermore, significant epistatic interactions were mostly detected between QTLs in bin 6.05 (average position 129 cM) and QTLs in bin 4.09 (166 cM) for ADL/NDF, KL/NDF, IVNDFD and DINAGZ in *per se* value and in topcross (Figure 4A for IVNDFD, other data not shown). These epistatic interactions revealed that, in addition to the main effects of both major QTLs located on chromosome 6 and 4, the effects of the QTLs on chromosome 6 were more pronounced when alleles from F271 are present at QTLs on chromosome 4 (Figure 4B).

Figure 2 - Improvement of bin 6.06 QTL positions. 2001 linkage map (left chromosome) versus 2012 linkage map (right chromosome) of the 131 RIL F288 x F271 progeny and QTL localizations for cell wall investigated traits (IVNDFD: *In Vitro* Neutral Detergent Fiber Digestibility, DINAGZ: *in vitro* cell wall digestibility according to Argillier et al. (1995), ADL/NDF: Acid Detergent Lignin/Neutral Detergent Fiber, KL/NDF: Klason Lignin/NDF) in *per se* (ps) and topcross (tc) experiments. Markers map in 2001 are written in black, re-evaluated genotyping markers in green and new markers in red. Numbers on the left of the chromosome indicate genetic position in cM from the top of the chromosome. Percentages under or above QTL bars correspond to R^2 values.

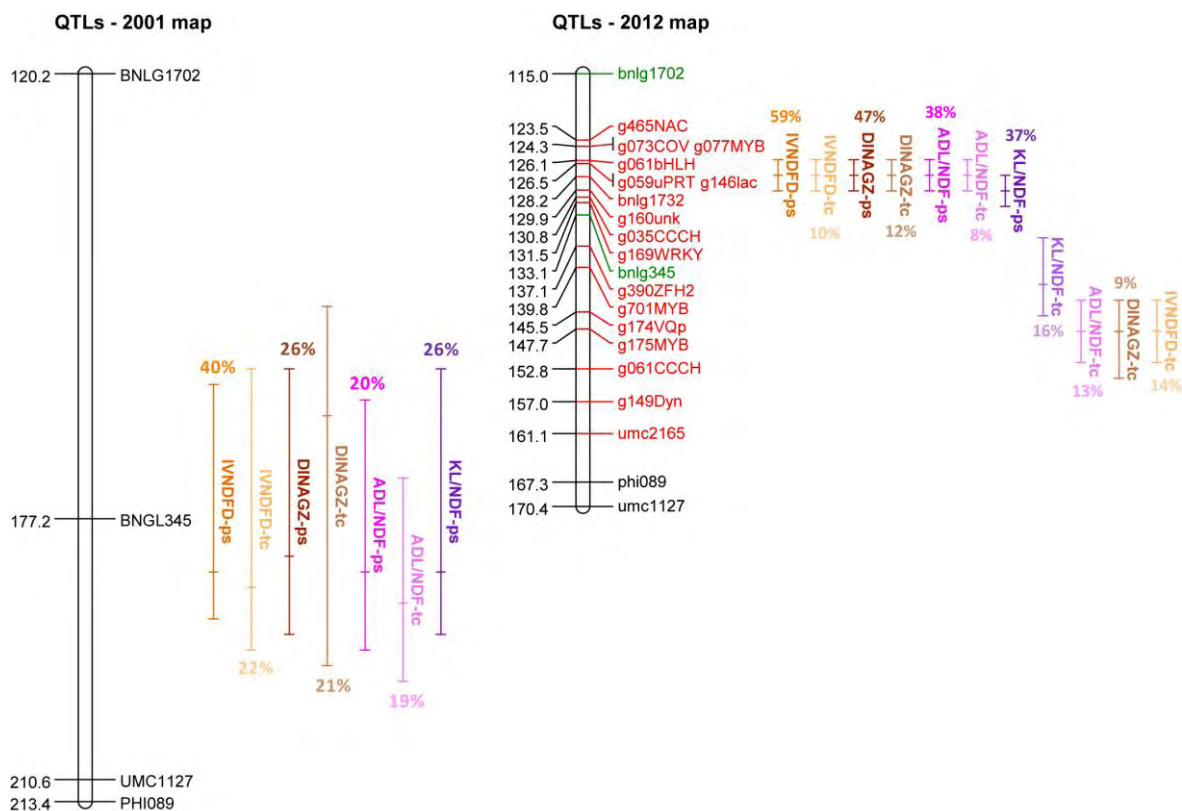
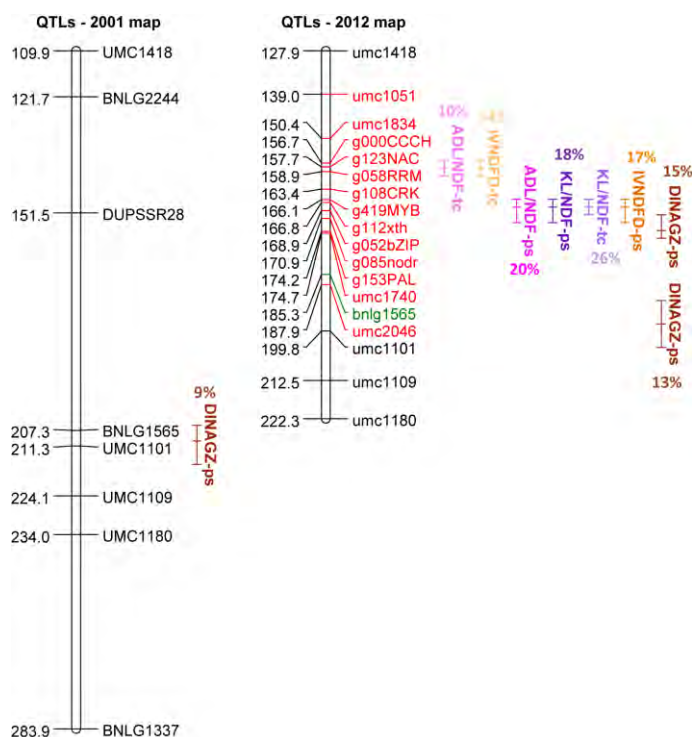


Figure 3 - Improvement of bin 4.09 QTL positions (legend as in figure 2).



Candidate genes underlying bins 6.05, 6.07 and 4.09 QTLs

Ninety two genes, or 50 genes when excluding the KL/NDF *per se* QTL, were present under the 2.2 Mbp support interval of the bin 6.06 upstream QTLs (bin 6.05) including the major *per se* QTLs. This is a strong reduction of the very high number of genes (800) present in the region defined in 2001. According to the MapMan BIN classification (Thimm et al. 2004), 39 % of the genes have an unknown function (Figure 5). Among genes with known function, only nine belong to families with members shown as involved or possibly involved in the biosynthesis and the assembly of the secondary cell wall and the regulation of these processes (*bHLH*, *FKBP*, *laccase*, *fasciclin*, *one zinc finger C2H2* and *two zinc finger C3HC4*, *NF-YB*, and *WRKY*). The *bHLH* (GRMZM2G061906), located in position 151.3 Mbp, is an ortholog of *AtbHLH105* transcription factor *ILR3* (At5g54680), belonging to the subclass IVc of the bHLH family (Pires et al. 2010). While functions of the plant group IV bHLH proteins are mostly unknown (Heim et al. 2003), a role has been shown for *ILR3* in modulating metal homeostasis and auxin-conjugate metabolism (Rampey et al. 2006). However, its role in secondary wall assembly is not clear, even if other bHLH were shown to be involved in the regulation of the phenylpropanoid metabolism (Ramsay and Glover 2005). The *FK506-binding proteins (FKBPs)* gene (GRMZM2G035922), located in position 151.5 Mbp, is the first ortholog of *AtFKBP20-1* (At3g55520). This gene belongs to the large family of peptidyl–prolyl *cis-trans* isomerases, which are known to be involved in growth and development (Harrar et al. 2001), based on observations in several mutants such as *pasticcino1 (pas1)* and *twisted dwarf (twd)*. Interestingly, this gene was less expressed in four RILs, with both the favorable allele and high cell wall degradability, compared to the parental line F271 with low degradability (Courtial et al. 2012a). However, its role in constitutive lignified tissue assembly is not yet established. The *laccase* gene (GRMZM2G146152), located at 151.5 Mbp, is one ortholog of *AtLac17* (At5g60020), which is involved in monolignol polymerization of *Arabidopsis* stems, and considered to be more particularly related to the deposition of G lignin units in fibers (Berthet et al. 2011). Moreover, *AtLac17* colocalizes with a QTL of lignin content in a RIL progeny of *Arabidopsis thaliana* (Chavigneau et al. 2012). The *fasciclin* gene (AC213621.5_FG004), located at 151.6 Mbp, is the first ortholog of the *SOS5 (Salt Overly Sensitive5)* gene, of which mutants have thinner cell walls. This gene encodes a putative cell-surface adhesion protein and is required for normal cell expansion (Shi et al. 2003). Three *zinc finger* genes, colocalize with bin 6.05 QTLs, one *zinc finger of the C2H2-type* (AC206217.2_FG006), located in position 151.8 Mbp and two *zinc finger of the C3HC4-type* (GRMZM2G075782 and GRMZM2G035601), in position 152.8 and 152.9 Mbp. *Zinc finger* genes encode members of one of the largest families of transcription factor regulatory proteins, which are involved in numerous regulations during plant development. The zinc finger C2H2 family was the most frequently represented in *Eucalyptus* secondary xylem libraries (Rengel et al. 2009). *Zinc finger C3HC4* genes belong to the transcription factor family that was the most differentially expressed in tension wood, deficient in lignin, in comparison with normal wood (Andersson et al. 2006). Moreover, plant over-expressing C3HC4

Figure 4 - Epistatic interactions for IVNDFD. A: Epistatic interactions present between the chromosome 4 and 6. $LOD_{f_{v1}}$ is plotted on the lower right triangle with color code for the LOD score presented on the right of the scale (Pvalue < 0.05 for LOD score > 4.32 according to permutation test performed - n=1000). B: IVNDFD phenotypic values of RILs according to their genotyping values for interacting markers *g419MYB* (QTLs bin 4.09) and *bnlg1732* (QTLs bin 6.05).

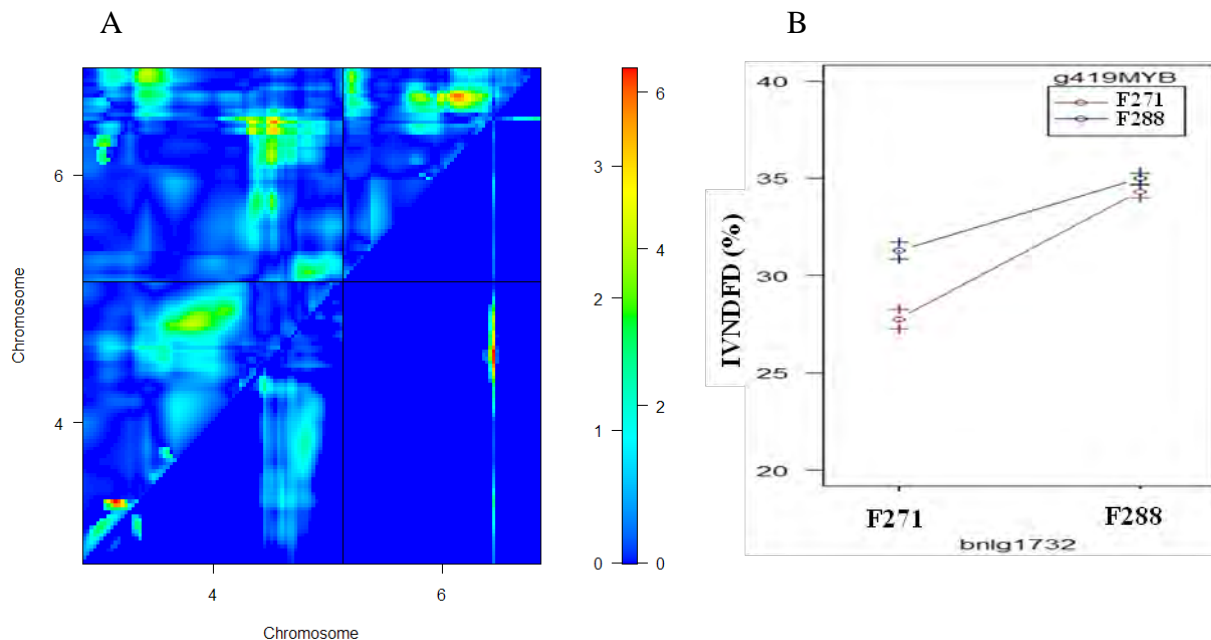
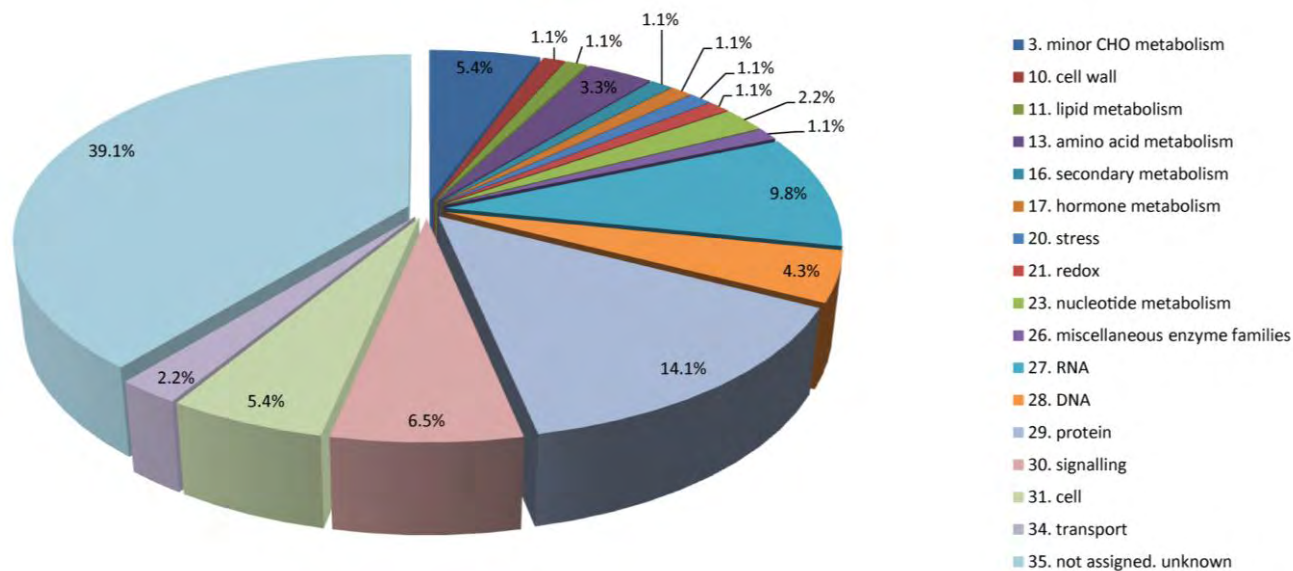


Figure 5 - Functional classification of the 92 genes underlying the major QTLs present in bin 6.05 (classification according to MapMan BIN, percentage of gene in each class).



protein has also increased cellulose and reduced lignin contents in *Eucalyptus* (Arruda et al. 2010). *Nuclear factor Y subunit B*, (GRMZM5G804893), located in position 152.2 Mbp, is an ortholog of *NF-YB8* expressed in vascular tissue in *Arabidopsis* (Siefers et al. 2009), of which the role is still unclear. Finally, the *WRKY* gene (GRMZM2G169966), located at 153.3 Mbp, belongs to a large family of transcription factors involved in development or in response to environmental signals, and especially to biotic and abiotic stress response. Its first *Arabidopsis* ortholog, *AtWRKY33*, is involved in defense reactions (Zheng et al. 2006).

One hundred seventy two genes colocalize with the 3.7 Mbp support interval of the bin 6.06 downstream topcross QTLs (bin 6.07), including two genes directly related to secondary cell wall biosynthesis and assembly. The *COVI-like* gene (GRMZM2G101533), in position 159.8 Mbp, is an ortholog of an *Arabidopsis* gene involved in lignified tissue patterning (Parker et al. 2003) but did not shown any polymorphism between the 2 parental lines F288 and F271. The *MYB* gene (GRMZM2G175232), just flanking QTL position (162.1 Mbp), is an ortholog of *AtMYB26*, of which loss of function induced a defect in the secondary wall thickening of anther walls resulting in indehiscent anthers (Yang et al. 2007).

In addition, 513 genes were shown underlying support intervals of QTLs located in bin 4.09, which are epistatic to QTLs located in bin 6.05 and QTLs with the highest R^2 value after those observed for bin 6.05 QTLs. Among these genes, three are directly related to secondary cell wall biosynthesis and assembly. *ZmMYB42* (GRMZM2G419239) is a transcriptional repressor of the maize lignin pathway genes (Fornalé et al. 2006; Sonbol et al. 2009; Gray et al. 2012). The *COVI-like* (GRMZM2G123790) gene is an ortholog of the *COVI Arabidopsis* gene, involved in the regulation of vascular patterning in the stem (Parker et al. 2003). The *PAL-like* (GRMZM2G153871) gene is paralogous to members of the PAL family, which are involved in the first step of monolignol biosynthesis. In addition, a single polymorphism in *ZmPAL* has been associated with *in vitro* degradability of organic matter (Andersen et al. 2007).

Discussion

A new linkage map to revisit line polymorphism and highlight probable monomorphic areas

Nearly one half (50.2 %) of the 494 successfully tested markers were shown to be polymorphic, which is a little more than the average of usual polymorphism rates (Yan et al. 2010). However, some areas remained with large distance between markers, despite the addition of numerous new markers on the F288 x F271 map. Because several markers tested did not shown any polymorphism, these regions were supposed to be monomorphic between the two parental lines, likely as a consequence of their common Co125 ancestry inducing a probable partial consanguinity between F288 and F271. In addition, it has been suspected that breeding efforts for grain yield have favored Co125 alleles, and

consequently induced larger monomorphic genomic areas than expected from their parentage (Roussel et al. 2002).

Ghost QTLs in bin 6.06 correspond likely to two distinct QTLs positions

Allelic sequencing revealed that QTLs estimated positions detected in 2001 in bin 6.06 were probably in a monomorphic region. This situation could result from erroneous QTLs positions as a consequence of insufficient markers coverage and/or RIL progeny numbers (57 cM between bnlg1702 and bnlg345, 36.2 cM between bnlg345 and phi089, and only 131 RIL). It could also be the consequence of the joint influence of two dependent QTLs for the investigated traits (Martinez and Curnow 1992, Studer and Doebley 2011). One QTL position could likely be located in the vicinity of bnlg345 while another QTL position could be located a little downstream, in the vicinity of phi089. In addition, several candidate genes and/or polymorphisms jointly responsible for variable phenotypes could underlie each QTL position. Corroborating possibly two slightly different QTL positions, Krakowsky et al. (2003) have located one QTL affecting fiber and lignin in stalk content located near the 155 Mbp position in the B73 x De811 F3 progeny, while two QTLs were found later near the 150 Mbp and 165 Mbp positions in the B73 x De811 RIL progeny (Krakowsky et al. 2005).

Finally, the addition of 17 new markers between bnlg1702 and phi089 allowed to distinguish two successive chromosomal regions involved in degradability and lignin content in bin 6.05 and 6.07 for topcross values, but only one region for *per se* values just upstream of the supposed monomorphic area. Ghost QTL detected in 2001 for topcross value were thus fractionated into two QTLs separated by 20 cM and situated on both sides of the previous position, confirming the joint influence of two linked QTLs. These two QTLs had quite similar R^2 values, of which sum was closed to 2001 R^2 values. As for *per se* QTLs, the downstream position for 2001 QTLs should probably be the consequence of the low marker density of the original map in this region.

Other contributions of the densified map for the detection of QTLs and limits

Marker densification of the genetic map allowed us to pinpoint QTL positions and to reduce support intervals of QTL as was shown by Darvasi et al. (1993). Furthermore, new QTLs were shown in bin 4.09, a region where the two initial markers (bnlg2244 and dupssr28) were removed and replaced by 13 new markers and one marker (bnlg1565) was re-genotyped. This point highlights the impact of marker spacing on QTL detection. This improved map which allowed to reduce the number of genes underlying the QTLs is a significant progress towards the identification of candidate genes. However, the 6 cM and 10 cM interval supports, corresponding to bin 6.05 QTLs and bin 4.09 major QTLs, respectively, seem now to have reached their minimum sizes with 7 and 6 markers underlying them, pointing out the limits imposed by the population size (Darvasi et al. 1993).

In addition, QTL effects were still high in these two regions. However, the estimates of phenotypic variances associated with correctly identified QTL were shown to be overestimated when the number

of progenies was low (Beavis 1998; Xu 2003). The high R^2 values observed in this experiment should thus be considered in light of the so-called Beavis effect. In addition, large R^2 values could be a consequence of the presence of multiple linked QTLs under the interval support (Studer and Doebley 2011).

Taking into account limits induced by the RIL population size, one could consider that the two genes and one microRNA, located just upstream and downstream of the interval supports of the major QTLs in bin 6.05 (151.1 to 153.3 Mbp), could be responsible for the presence of the QTLs. Just upstream, the *MYB* (GRMZM2G077789) in position 150.7 Mbp is an ortholog of *EgMYB1*, which negatively regulates secondary cell wall formation (Legay et al. 2007; Legay et al. 2010). Just downstream, the *OVATE* gene (GRMZM2G026927), in position 153.8 Mbp, is an *OFPI* ortholog, which was shown to interact with the negative regulator of secondary wall biosynthesis *KNAT7*, itself targeted by *SND1/VND6* and *MYB46* (Grima-Pettenati et al. 2012). Finally, the mir164c, located in position 153.4 Mbp, was shown to target several NAC and one laccase mRNA, and thus was involved in lignin processes (Zhang et al. 2009). Assuming a possible relationship between the *MYB* and *OVATE* genes located upstream and downstream of the QTL positions, it is conceivable either that the underlying determinants would be at least these two genes, or that these QTLs should be again fractioned into at least two positions.

Conclusion

The objective of this study was to identify the candidate genes underlying the major QTLs for cell wall degradability and lignin content previously detected in bin 6.06 in the F288 x F271 RIL progeny (Roussel et al. 2002). These QTLs expanded over a 48 cM long support interval corresponding to 12.2 Mbp, and more than 800 genes in the B73 reference genome. As a first step identification, allelic variation was investigated for putative candidate genes involved in cell wall biosynthesis and assembly. This targeted sequencing highlighted a monomorphic region between the two parental lines, including the QTL estimated positions and consequently suggests that QTLs were likely "ghost" QTLs. Targeted map densification, using HRM, was thus performed to pinpoint QTL positions and to reduce support intervals of QTL. Markers were mapped in the most important QTLs regions, especially within or at the close vicinity of the major bin 6.06 QTLs support intervals, and in several large still unmarked regions. Finally, only few areas remained with large distance between markers, as a consequence of the absence of polymorphism in these regions.

Forty-three QTLs were detected for lignin (ADL/NDF, KL/NDF) and cell wall degradability (DINAGZ, IVNDFD) traits in the F288 x F271 progeny for both *per se* and topcross experiments with the improved map, while only 32 were highlighted in 2001. This increase of the QTL number could be

explained by the fact that some large effect QTLs were fractionated in two QTLs (bin 6.06) and that new region involved in degradability were revealed (bin 4.09). The putative ghost QTLs in bin 6.06 indeed corresponded to two regions, with cell wall traits QTLs detected at the two positions, one located upstream (bin 6.05) and one downstream (bin 6.07) of the monomorphic region. Major *per se* QTLs in this progeny, colocalized now in bin 6.05 with support intervals only extending on 6 cM corresponding to 2.2 Mbp. This decrease of QTL support interval length came along with a strong reduction of gene number underlying QTLs (currently 92 genes versus 800 genes in 2001). Among these 92 genes, nine could putatively be considered as involved in the biosynthesis and the assembly of the secondary cell wall and the regulation of these processes, but genes of unknown function, or genes for which roles in lignification have not yet been established could be the real candidate(s) or could belong to the group of underlying determinants.

In addition to the improvement of bin 6.06 QTL positions, new areas involved in cell wall degradability were detected after the marker densification in bin 4.09. *Per se* QTLs with high R^2 values colocalized in this region and showed significant epistatic interactions with the major QTLs located in bin 6.05. Three genes greatly related to secondary cell wall were present in this region (*ZmMYB42*, *COV1-like*, *PAL-like*).

This study allowed refining QTLs for cell wall degradability and lignin content, and also to fractionate large effect QTLs, corroborating the questioning proposed by Studer and Doebley (2011). For breeding purposes, this work allowed to specify markers which could be used in MAS towards higher maize cell wall degradability for silage use or second generation bioethanol production. Moreover, following the length reduction of QTL support interval, the identification of the only gene or clustered genes responsible of the QTLs was made easier, with a great reduction of gene number underlying QTL areas. Finally, this work opened avenues for the establishment of BAC libraries of the two parental lines F271 and F288, followed by the targeted sequencing of BACs overlapping bin 6.05 QTL support intervals.

Acknowledgements

This work has been funded by the maize breeding companies involved in the PROMAÏS - INRA network on maize cell wall lignification and digestibility (Advanta, Caussade Semences, Limagrain Europe, MaïsAdour, Monsanto SAS, Pioneer Génétique, Pau Euralis, R2n RAGT Semences, SDME KWS France, Syngenta seeds). We thank Christiane Minault, Dominique Denoue and Pascal Vernoux for seed multiplications and plant cropping and sampling at INRA Lusignan. Thanks also to Jacques Laborde, who managed seed multiplications at INRA, St Martin de Hinx, France. We are grateful to Fabien Mounet for his fruitful criticisms.

CHAPITRE III - RECHERCHE DE QTLs DE DÉGRADABILITÉ
DES PAROIS, DE TENEUR EN LIGNINES, DE STRUCTURE
DES LIGNINES, ET DE TENEUR EN ACIDES *P*-
HYDROXYCINNAMIQUES DANS LA DESCENDANCE
RECOMBINANTE F288 x F271 (ESSAIS MAÏS EN PLANTES
SANS ÉPIS RÉALISÉS EN 2010 ET 2011, AVEC RÉCOLTE AU
STADE ENSILAGE)

3.1. Avant propos

Les essais initiaux qui ont fondé ce travail ont été conduits à la fin des années 1990 sur les 131 RILs développées à partir du croisement des lignées F288 x F271. La synthèse de ces essais (Roussel et al, 2002) avait ainsi permis la mise en évidence des QTLs de dégradabilité des parois et de teneur en lignines dont ceux à effets majeurs au niveau du bin 6.06. Seul le caractère de teneur en lignines pouvait être prédit en spectroscopie dans le proche infrarouge (NIRS) en parallèle de la dégradabilité des parois au moment où ces essais ont été conduits avec des récoltes en plantes entières (avec épi). Récemment de nouvelles équations de prédiction NIRS ont été développées, sur des plantes sans épi, pour ces caractères mais aussi pour d'autres caractères influençant la dégradabilité des parois tels que la composition en monomères des lignines (mesurée après attaque au nitrobenzène alcalin et non après thioacidolyse et rupture spécifique des liaisons β -O-4), et la teneur des parois en acides *p*-hydroxycinnamiques (Barrière et al, 2008; Riboulet et al, 2008b). Ce chapitre fait ainsi l'objet de recherches de QTLs pour l'ensemble de ces caractères pariétaux et permet ainsi de renforcer les études précédentes faites sur cette famille recombinante. Cette approche QTLs faite pour l'ensemble des caractères biochimiques influençant la dégradabilité des parois, devrait également contribuer à la compréhension des mécanismes génétiques impliqués dans la dégradabilité des parois cellulaires et à l'identification des déterminants génétiques.

Ce chapitre fera l'objet de publication une fois la publication précédente acceptée et inclura les essais mis en place en 2012 sur les sites INRA du Moulon et de Ploudaniel.

3.2. Résultats

Au cours de ces nouveaux essais, une variabilité importante et significative a été mise en évidence pour tous les caractères pariétaux étudiés, avec des effets génotypes bien supérieurs aux effets d'interactions génotype x environnement. Par ailleurs, nous avons pu montrer que la dégradabilité des parois était corrélée négativement à la teneur en unités syringaldehyde (Sga) libérée, représentative de celle en sous-unités S, mais aussi aux teneurs en lignines et en etherFA libéré (qui correspond à une partie des FA pontant arabinoxylanes et lignines). A l'inverse, la dégradabilité est corrélée positivement à la teneur en diférulates libérés confortant les hypothèses émises lors des études précédentes (Riboulet et al, 2008b; Barrière et al, 2009a).

En prenant en compte l'ensemble des caractères étudiés, 42 QTLs ont été mis en évidence, répartis sur sept chromosomes, avec notamment des clusters de QTLs au niveau des bins 3.06, 4.09 et 6.05, confirmant les résultats obtenus en 2001 sur cette même famille. Le cluster le plus important de part le

nombre de QTLs et leurs effets (R^2) demeure celui du bin 6.05 (Courtial et al, 2012d) avec dix QTLs détectés sur un intervalle de 12 cM, regroupés en deux sous-clusters, dont sept QTLs en position 128 cM (dégradabilité des parois, teneur en lignines, et Sga), et trois QTLs de férulate et diférulates en positions 132 - 134 cM. Le deuxième cluster présent au niveau du bin 3.06 comprend, quant à lui, cinq QTLs (dégradabilité des parois, de teneurs en ADL/NDF, en étherFA, et en syringaldéhyde) avec des valeurs de R^2 également élevées. Enfin, un cluster de quatre QTLs comprenant des QTLs de dégradabilité des parois colocalisant avec des QTLs de teneur ADL/NDF et de *p*-hydroxybenzaldéhyde a été mis en évidence dans le bin 4.09, avec toutefois des valeurs de R^2 un peu moins élevées. De façon cumulée, les QTLs situés dans les bins 3.06, 4.09, et 6.05 expliquent respectivement près de 70 et 60 % de la variation phénotypique pour la teneur en ADL/NDF et la dégradabilité des parois, et 40 % de la variation phénotypique pour le Sga et les étherFA libérés. Dans tous les cas sur ces trois positions, les "allèles" défavorables à la dégradabilité des parois sont ceux de la lignée F271.

La recherche des gènes candidats sous-jacents ces QTLs s'appuyant sur la lignée de référence B73 a permis de mettre en évidence que les QTLs du bin 4.09 colocalisent avec le gène *ZmMYB42*, un gène de type COV1, un gène de type PAL, et un gène de type KNAT7. D'autres gènes de types MYB et NAC colocalisent avec d'autres QTLs. D'autre part, les QTLs d'étherFA du bin 3.06 et 3.09 colocalisent avec trois gènes codant des glycosyl transférases de la famille GT47 et un gène codant pour une arabinoxylane CoA-acyltransférase de la famille PF02458, respectivement. Ces transférases sont actuellement les seuls candidats putativement identifiés pour assurer le transfert d'un groupe feruloyl sur une chaîne d'arabinoxylanes dans les parois cellulaires de graminées.

3.3. Conclusions et perspectives

Un premier résultat important de cette nouvelle démarche d'expérimentation au champ et d'analyses des parois est la robustesse de la mise en évidence des QTLs du bin 6.05 (ex 6.06), mais aussi des bins 3.06 et 4.09, en dépit d'un nombre un peu faible de RILs et de seulement deux années d'expérimentation dans des conditions climatiques nettement différentes des conditions de la fin des années 1990. Le second point le plus notable est sans doute le fait qu'il y a de nouveau des colocalisations multiples entre caractères (dégradabilité des parois, teneur en lignines, structure des lignines, férulate et diférulates pontant les composés pariétaux), et donc entre des caractères dont les déterminants moléculaires sont, *a priori* et avec les connaissances actuelles, différents. Il faut alors, semble-t-il, imaginer des déterminants multiples proches et liés, ou des déterminants très amonts dont la nature n'est pas connue, et/ou des déterminants de type gènes de régulation déjà identifiés, mais dont l'ensemble des fonctions n'est pas connu.

En terme de critères de sélection, le critère global de mesure de la dégradabilité des parois est certainement un outil efficace. Toutefois, les progrès à la fois en sélection et en compréhension, ainsi que la valorisation des ressources génétiques, viendront, au-delà des mesures de dégradabilité de parois, des mesures de teneur en lignines (ADL), de la détermination de la composition monomérique des lignines, et de la prise en compte de la réticulation au sein des parois par le férulate et les diférulates. Et même si cela n'est pas apparu au cours de cette étude, l'importance de l'acylation des unités S par l'acide *p*-coumarique ne devrait pas être sous-estimée. Une autre raison de la nécessité d'une prise en compte des déterminants de façon individuelle est liée au fait que, pour l'alimentation animale, les critères de dégradabilité des parois ou des plantes entières n'expliquent que 50 % de la variabilité observée sur animaux. De plus, les aspects liés au comportement mécanique des tissus, qui conditionnent entre autres l'ingestibilité du fourrage, sont complètement ignorés. En revanche, différents résultats laissent penser que la variabilité de l'ingestibilité pourrait être très significativement liée à l'intensité des réticulations entre les composants pariétaux par les ponts férulate et diférulates. Enfin, si des données non publiées (Génoplande projet GrassBioFuel) montrent que la variabilité génétique demeure, mais de façon atténuée après un prétraitement alcalin des échantillons, la sensibilité des génotypes à des traitements modérés, peu coûteux en énergie et peu polluants, est très probablement liée à des caractères pariétaux non directement mesurés par les dégradations *in vitro* actuellement utilisées, confortant également pour cette cible industrielle, la nécessité d'arriver aux déterminants génétiques des différents facteurs impliqués dans la mise en place des parois secondaires du maïs et plus largement des graminées.

Article 3 - QTLs for cell wall degradability, lignin content and structure, and *p*-hydroxycinnamic acid contents in the F288 x F271 early maize RIL progeny (in progress)

Audrey Courtial^{1,2,3}, Valérie Méchin⁴, Matthieu Reymond⁴, Jacqueline Grima-Pettenati^{2,3}, Yves Barrière^{1*}

¹ INRA, Unité de Génétique et d'Amélioration des Plantes Fourragères, 86600 Lusignan, France

² Université de Toulouse ; UPS ; UMR 5546, Laboratoire de Recherche en Sciences Végétales ; BP 42617 Auzeville, F-31326 Castanet-Tolosan, France

³ CNRS ; UMR 5546 ; BP 42617, F-31326 Castanet-Tolosan, France

⁴ INRA, Institut Jean-Pierre Bourgin, 78026 Versailles, France.

* corresponding author

Key words: maize, degradability, lignin, *p*-hydroxycinnamic acid, cell wall, QTL

Abstract

In a two year experiment, cell wall degradability and several cell wall component traits including lignin content, *p*-hydroxycinnamic acid contents, and lignin monomeric structure were investigated in the F288 x F271 RIL progeny. Genotype effects were highly significant for all cell wall investigated traits ($P < 0.001$), and much higher than genotype \times environment (GE) interaction effects. In addition to more isolated QTLs observed in six positions, three large clusters gathering 19 QTLs were shown in bins 3.06, 4.09, and 6.05. Eleven and 23 QTLs out of 42 explained more than 15 and 20 % of the observed trait phenotypic variation, respectively. According to positions, colocalizations of cell wall degradability QTLs occurred with lignin QTLs, *p*-coumaric acid QTLs, etherified ferulic acid QTLs and/or syringaldehyde QTLs. Candidate genes regulating cell wall degradability were searched for colocalizing positions, pointing out the possible involvement of *ZmMYB42*, COV1-like, and a KNAT7-like gene in bin 4.09. MYB, NAC, and zinc finger genes also colocalized with the other cell wall degradability QTLs, but candidates also possibly included colocalizing genes of unknown functions. Finally, the etherFA QTL located in bin 3.09 was in close position with an arabinoxylan CoA-acyltransferase of the PF02458 family.

Introduction

Large variations in cell wall degradability of maize lines are currently clearly established, that are of interest for maize breeding either for silage use or for second generation biofuel production. While large variations in maize cell wall composition are also established, the relationships between the different cell wall component traits and cell wall degradability are not univocal, nor established for all components. The core-lignin content (or acido-resistant fraction of lignins, estimated by the ADL/NDF trait, Goering and van Soest, 1971) is probably the first trait for which a negative relationship with the degradability of the cell walls has been shown, even if correlations were significantly dependent on the genetic backgrounds. The content in *p*-coumaric acid (pCA) acylating syringyl (S) units of lignins regularly appeared negatively related to cell wall degradability, and even at a higher level than the ADL/NDF lignin content (Riboulet al, 2008b), but without clear explanations. Different assumptions regarding the maturity of the secondary tissues and the consequences of the acylation on the geometric structure of lignins, and thus the accessibility of cell wall carbohydrates by enzymes, have been proposed (Zhang et al, 2011). The content in etherified ferulic acid (etherFA), corresponding to ferulic acid bridging arabinoxylans and guaiacyl (G) units of lignins, was also shown significantly negatively related to cell wall degradability. Such cross-linkages were considered as explaining up to 50% of the negative effects of phenolic compounds on the cell wall degradability (Grabber et al, 2009). Contrarily, the diferulate (diFA) bridging of arabinoxylans chains appeared positively related to cell wall degradability, while such bridges are unfavorable. This is currently explained by the fact that the measured diFA are accessible diFA, and not an accurate estimate for the total amount of diFA. Similarly, it is unclear how much of etherFA are released, and there are many reasons to believe that this share varies between genotypes. Finally, the effect of ratios between H (*p*-hydroxyphenyl), G, and S monomeric units of lignins is also variable considering different genetic backgrounds, although there is a trend that more H and/or S are unfavorable to the degradability of the cell walls (Riboulet et al, 2008b; Barrière et al, 2009b). While degradability of the cell walls is related to lignin content, the lignin monomeric composition and the ferulate cross-linkages likely have a considerable impact on this trait as well. Moreover, a given value of cell wall degradability can result from different secondary wall composition and organization, which could have more or less negative effects on plant agronomic performances, including drought and pest tolerance. In addition, there is also a large variation in the intensity of lignification between the different tissues within the same line and between lines. Thus, the understanding of the genetic mechanisms involved in cell wall degradability is also supported by the understanding of the genetic basis of the different traits which contribute to a given value of cell wall degradability.

One way to decipher the relationships between cell wall degradability and the secondary wall related traits at the genetic level is to investigate the colocalizations of QTLs for these sets of traits.

However, colocalizations do not give insights if an only genetic trait is responsible for all observed variations (pleiotropy) or if this was due to linked genes. To address this question and to identify the gene(s) responsible for the presence of the clustered QTLs, different approaches would be required such as fine mapping and validation with mutants of candidate genes. The identification of genetic determinism underlying QTLs is also necessary for the definition of robust markers for breeding of lines with more degradable cell walls and high agronomic value. In the F288 x F271 RIL progeny, QTLs with large effects have been mapped for both lignin content and cell wall degradability traits (Roussel et al, 2002). However, during these previous experiments, *p*-hydroxycinnamic acid contents and monomeric structure of lignins were not investigated. The objective of the present work was to strengthen the previous study in investigating also genetic basis for lignin structure and *p*-hydroxycinnamic acids content in relation with previously and currently detected QTLs in this early maize progeny. Moreover, further insights were considered for investigating relationships between biomass degradability and phenolic cell wall components.

Materials and methods

F288 x F271 RIL progeny

The set of the 131 RILs was initially developed nearly 15 years ago by single seed descent from the cross between the two early dent inbred lines F271 and F288, which are of low and medium-high cell wall degradability, respectively (Barrière et al, 2001). Due to poor or null germination of several progenies, only 119 lines were currently re-investigated, without relationships between cell wall traits and lack of germination.

Field experiments

Field experiments were carried out over two years (2010 and 2011) for RIL *per se* evaluation in one location at INRA Lusignan (a two-location experiment is planned for 2012 in le Moulon and Ploudaniel). Each year, RILs were evaluated using a generalized alpha-lattice design with two replicates, with three extra replicates of a 32 RIL core collection in 2011. Each experimental plot was a 5.2 m long single row of 37 plants, with a row spacing equal to 0.75 m. Irrigation was applied during the summer to prevent water stress. In order to have an estimate of stalk ferulate content without the confusing effect of grain ferulates, ears were removed by hand from all plants the day before and/or the day of harvest in all plots. Plots were machine-harvested with a forage chopper at silage maturity stage. A representative sample of 1 kg chopped material per plot was collected for dry-matter (DM) content, DM yield estimates, and biochemical and NIRS analyses. Samples were dried in a ventilated oven (65°C), and dry samples were ground with a hammer mill to pass through a 1 mm screen for further analyses.

Cell wall investigations

Neutral Detergent Fiber (NDF), Acid Detergent Fiber (ADF), and Acid Detergent Lignin (ADL) were investigated according to Goering and Van Soest (1970). Klason lignin (KL) was estimated according to Dence and Lin (1992). Both ADL and KL values were expressed as percentage of NDF. KL values, which are two to four times greater in grasses than ADL estimates, contain the lost acid-soluble part of lignins during the first step of the ADL procedure (Hatfield et al, 1994; Jung et al, 1997; Hatfield and Fukushima, 2005). The soluble fraction of lignin ($KL - ADL = KL/NDF - ADL/NDF$) was thus estimated by the difference between the two lignin traits (Zhang et al, 2011). The *in vitro* dry matter digestibility (IVDMD) was estimated according to Aufrère and Michalet-Doreau (1983). Cell wall digestibility was investigated according to Struik (1983) and Dolstra and Medema (1990). *In vitro* NDF digestibility (IVNDFD) was thus computed assuming that the non-NDF part of plant material was completely digestible [$IVNDFD = 100 \times (IVDMD - (100 - NDF)) / NDF$]. Cell wall digestibility was also estimated as the *in vitro* digestibility of the "non starch, non soluble carbohydrates, and non crude protein" part (DINAGZ, Argillier et al, 1995; Barrière et al, 2003b). The latter is computed assuming these three constituents are completely digestible [$DINAGZ = 100 \times (IVDMD - ST - SC - CP) / (100 - ST - SC - CP)$] where ST, SC and CP are starch, soluble carbohydrate, and crude protein contents, respectively (and considering a null value for starch content in plant without ears).

p-Hydroxycinnamic acid contents were measured after treating NDF fractions with NaOH according to the double procedure previously described by Morrison et al, (1993) and used by Méchin et al, (2000). This procedure involves a mild alkaline treatment allowing the release of esterified ferulic (esterFA) and *p*-coumaric (pCA) acids, and a severe alkaline treatment allowing the release of both esterified and etherified ferulic acid (etherFA). Because almost all the pCA is esterified (Ralph et al, 1994), only esterified pCA content was investigated. The concentration of etherFA was calculated as the difference between FA amounts released by the severe and mild alkaline treatments, as all FA involved in etherified linkages are also (and previously) involved in esterified linkages. Besides esterFA and etherFA contents, two FA dimers were also reported as the content in the 5-5 and 8-O-4 FA dimers. The latter is the predominant FA dimer out of the six shown in maize cell wall (Lindsay and Fry, 2008). In order to investigate the monomeric composition of lignins, oxidation of cell wall residues with alkaline nitrobenzene was performed according to a method adapted from Roadhouse and MacDougall (1956) according to Higuchi and Kawamura (1966). During alkaline nitrobenzene oxidation, *p*-hydroxyphenyl (H), guaiacyl (G), and syringyl (S) monomers, resulting from *p*-coumaryl, coniferyl and syringyl alcohol polymerization, respectively, are oxidized into *p*-hydroxybenzaldehyde (pHb), vanillin (Van) and syringaldehyde (Sga), respectively, with disruption of C α and C β linkages. Extracted pHb, Van, and Sga were analyzed by HPLC. Parts of pCA and FA are also oxidized into pHb and Van, respectively. Consequently, amounts of H and G units of lignins may be partly overestimated. In addition, as proposed by Zhang et al, (2011), the proportion of S units acylated by

pCA was estimated assuming that all of the *p*-coumaric acid was esterified to syringyl lignin units and extrapolating the percentage of syringyl units released by alkaline nitrobenzene oxidation to the whole lignin [$Sga\text{-}pCA = pCA / (Sga \% \times KL/NDF)$, with $Sga \% = Sga / (pHb + Van + Sga)$].

Due to the large number of samples, all biochemical traits were estimated using near infrared reflectance spectroscopy (NIRS, NIRS system 6500 spectrophotometer), with the specific calibrations, previously developed at INRA Lusignan for plants without ears (Barrière et al, 2008; Riboulet et al, 2008b). Calibration equations have been validated by a laboratory analysis of 20 samples per location with a sample choice based on spectral data.

Data and QTL analyses

Variance analyses were carried out using the standard procedure of a fixed model with genotype, environment (year-location), block, sub-block and interaction effects, as $Y_{ijkl} = \mu + E_j + B_k.E_j + SB_l.B_k.E_j + G_i + G_i.E_j + \varepsilon_{ijkl}$, with Y_{ijkl} = observed value for a given trait, μ = grand mean, E_j = environment effect, $B_k.E_j$ = block within environment effect, $SB_l.B_k.E_j$ = sub-block within block and environment effect, G_i = genotype effect, $G_i.E_j$ = genotype x environment interaction, and ε_{ijkl} = residual error (= σ_r^2), using Modli (Least Square Means, Kobilinsky 1983) and Splus Softwares (Venables and Ripley, 1994). Pearson's phenotypic correlations between traits were computed on the mean basis over environments.

The linkage map was recently developed based on 173 markers and 131 RILs, after a targeted addition of markers (Courtial et al, 2012d), completing the original map (Barrière et al, 2001) in areas without markers or close to previously shown QTL positions. QTL mapping was based on trait means over years and locations for the 119 RILs, using the Composite Interval Mapping (CIM, Zeng, 1994) method implemented in the PLABQTL computer package (Utz and Melchinger, 1996). PLABQTL uses the regression method (Haley and Knott, 1992) in combination with markers which are selected by stepwise regression as cofactors. LOD support intervals are constructed in PLABQTL according to Lander and Botstein (1989) and are likely underestimated in the case of CIM. The percentage of phenotypic variance ascribed to an individual QTL was estimated with the approximate standard error of Kendall and Stuart (1961). The additive effects of QTLs were estimated as half the difference between the phenotypic values of the respective homozygotes. LOD threshold estimates were based on the permutation-test method of Churchill and Doerge (1994) implemented in PLABQTL, with 1000 permutations. The maize bin was given for each QTL as the position of the left-flanking marker, according to MaizeGDB. QTL physical positions were estimated based physical positions of the two flanking markers on the reference genome B73 (Schnable et al, 2009, www.maizegdb.org and www.maizesequence.org, release v2 5b60), assuming a linear relationship between recombination and physical distances within this interval. The physical lengths of QTL support intervals were difficult to

Table 1 - Variance analysis, mean, maximum, and minimum values of traits for the F288 x F271 RIL progenies in *per se* experiments and plant without ear at silage harvest (all genotype and genotype x environment mean-square (MS) effects were significant at $P < 0.001$).

Traits	genotype MS	gen x env MS	σ_r^2	mean	mini	maxi	F271	F288
ADL/NDF	1.59	0.20	0.08	5.32	4.33	6.73	6.00	4.85
KL/NDF	8.28	0.88	0.37	14.1	11.3	17.4	15.6	14.1
KL-ADL	5.53	0.71	0.32	8.80	6.02	11.80	9.63	9.25
IVNDFD	38.5	4.01	2.32	25.7	17.5	33.0	22.9	28.2
DINAGZ	39.8	5.74	2.54	45.5	38.6	53.1	43.8	48.7
pCA	6.94	0.87	0.52	12.5	10.1	15.8	12.5	12.5
EsterFA	0.86	0.10	0.047	5.87	4.66	6.52	5.48	6.03
EtherFA	0.102	0.019	0.008	1.43	1.13	1.81	1.52	1.29
5-5diFA	0.001	0.000	0.000	0.13	0.08	0.16	0.11	0.14
8-O-4diFA	0.003	0.001	0.000	0.29	0.24	0.35	0.29	0.30
pHb	0.156	0.029	0.014	1.76	1.35	2.18	1.74	1.73
Van	1.98	0.32	0.16	6.90	5.84	8.36	6.46	6.86
Sga	3.93	0.48	0.26	8.82	7.11	11.3	9.28	8.23
Sga-pCA	16.3	2.44	0.98	17.7	12.9	21.3	15.1	18.2

NDF = Neutral Detergent Fiber, ADL = Acid Detergent Lignin, KL = Klason Lignin, IVNDFD = in vitro NDF digestibility, DINAGZ = in vitro cell wall digestibility according to Argillier et al (1995), KL-ADL = KL/NDF - ADL/NDF, pCa = *p*-coumaric acid, EsterFA and EtherFA = esterified and etherified ferulic acids, 5-5 diFA and 8-O-4 diFA = 5-5 and 8-O-4 diferulic acids, pHb = *p*-hydroxybenzaldehyde, Van = vanillin, Sga = syringaldehyde, Sga-pCA = proportion of S units acylated by pCA (pCA, FA, pHb, Van, and Sga as mg/g NDF, other traits as percentages).

Table 2 - Correlation between cell wall degradability traits and cell wall traits in the F288 x F271 RIL progeny (legend as in table 1).

Traits	IVNDFD	DINAGZ
ADL/NDF	-0.74	-0.69
KL/NDF	-0.69	-0.34
KL-ADL	-0.43	-0.05
pCA	-0.50	-0.34
EsterFA	0.29	0.41
EtherFA	-0.63	-0.69
5-5diFA	0.58	0.57
8-O-4diFA	0.49	0.45
pHb	-0.42	-0.61
Van	-0.47	-0.58
Sga	-0.80	-0.74
Sga-pCA	0.30	0.07

estimate in several instances, especially when there was an overlapping of centromer positions, and were thus not reported. Their gross values could be considered to range between 10 and 50 Mbp.

The search for candidate genes underlying cell wall related traits focused on genes involved in secondary wall biosynthesis and assembly, and was based on the list previously proposed by Barrière et al, (2009a). The list also included maize orthologs of newly published transcription or regulation factors in *Arabidopsis* or woody plants (poplar, *Eucalyptus*, ...), and currently comprises nearly 300 genes. Gene physical positions were based on the maize B73 sequence (www.maizesequence.org, release v2 5b60), assuming a similar genomic structure in RIL parental lines and B73 within QTL support intervals.

Results

Trait mean, variance, and correlation estimates

Average RIL stover DM content was equal to 35.8 %, corresponding to an advanced maturity stage, nevertheless fully suitable for investigations of lignification traits, especially for biofuel production based on straw harvest. Lower lignin content and higher cell wall degradability were observed in F288 in comparison to F271, with important transgressive values in the RIL progenies (Table 1). While the two parental lines had similar *p*-coumaric acid content, transgressions were shown in the progeny for both higher and lower values. EsterFA, 5-5diFA, and 8-O-4diFA release was higher in F288 than in F271, and the reverse for etherFA release, with again transgressive values in the progeny. *p*-Hydroxybenzaldehyde and vanillin releases were similar in the two parental lines, with transgressive values in the progeny, especially for higher vanillin release. A higher release of syringaldehyde was shown in F271 than in F288. The high values observed for *p*-hydroxybenzaldehyde, nearly 10 % of the total aldehyde release, illustrated a partial oxidation of *p*-coumaric acid into *p*-hydroxybenzaldehyde during the alkaline-nitrobenzene attack of cell wall components. The *p*-hydroxyphenyl monomer indeed usually represents only 3 % of maize lignin monomers released after thioacidolysis. A little less than 20 % of S units were estimated to be acylated by pCA. However, the proportion of acylated S units varied from 12.9% to 21.3% between RILs, which seemed considerable because this parameter was considered to have limited variation between maize inbred lines (Méchin, com pers). Estimates of cell wall degradability were nearly 20 percentage points lower with IVNDFD than with DINAGZ, corroborating previous investigations. Nonetheless, both IVNDFD and DINAGZ traits are similarly correlated with *in vivo* NDF digestibility estimated with sheep and explained nearly 60 % of the *in vivo* trait, whose average value was shown to be close to 48 % with a 40-60 % range (Barrière et al, 2004a).

Genotype effects were highly significant for all investigated traits ($P < 0.001$), and much higher than genotype \times environment (GE) interaction effects (Table 1). Based on mean-square (MS) values,

Table 3 - QTL analysis for cell wall traits in the F288 x F271 progeny (cell wall degradability, lignin content and structure, *p*-hydroxycinnamic acid contents, legends as in table 1).

	chr	pos cM	pos Mbp	Left marker	bin	supp int cM	LOD	R2	add value	Line +
Sga-pCA	1	226	226.9	bnlg1025	1.08	212-238	3.51	12.7	-0.50	F271
Van	2	42	16.2	g038MYB	2.03	22-50	5.63	19.6	0.26	F288
pCA	2	202	223.6	umc1230	2.08	198-206	4.46	15.8	0.36	F288
pHb	2	204	225.4	bnlg1520	2.08	198-210	4.15	14.9	0.07	F288
Sga-pCA	2	204	225.4	bnlg1520	2.08	198-208	5.65	19.6	0.59	F288
Van	3	112	54.3	g017MYB	3.04	104-120	4.95	17.4	-0.22	F271
pHb	3	118	97.9	g017MYB	3.04	106-126	3.76	13.5	-0.07	F271
ADL/NDF	3	134	158.0	bnlg1117	3.05	128-144	5.44	19.0	-0.17	F271
Sga	3	138	161.0	umc2265	3.06	128-146	3.03	11.1	-0.23	F271
IVNDFD	3	138	161.8	umc2265	3.06	132-142	7.09	24.0	0.97	F288
DINAGZ	3	140	164.0	umc2265	3.06	132-144	7.26	24.5	1.10	F288
EtherFA	3	140	164.0	umc2265	3.06	134-144	7.44	25.0	-0.06	F271
EtherFA	3	248	218.2	umc1273	3.08	240-254	6.19	21.3	-0.05	F271
ADL/NDF	4	168	224.6	g112xth	4.09	166-172	3.04	11.1	-0.12	F271
IVNDFD	4	172	230.0	g085nodr	4.09	168-174	1.94	7.2	0.47	F288
pHb	4	174	231.5	g153PAL	4.09	170-176	3.99	14.3	-0.06	F271
DINAGZ	4	174	231.5	g153PAL	4.09	170-176	4.30	15.3	0.73	F288
Van	5	42	12.1	umc1894	5.01	28-48	4.21	15.1	-0.20	F271
Sga-pCA	5	64	42.8	umc2294	5.03	52-78	3.71	13.4	-0.53	F288
DINAGZ	5	212	215.0	umc1225	5.08	204-216	3.01	11.8	-0.72	F271
KL/NDF	6	4	1.04	bnlg1043	6.01	0-6	3.89	15.0	0.35	F288
pCA	6	4	1.04	bnlg1043	6.01	0-6	7.46	26.8	0.48	F288
EsterFA	6	12	15.8	umc1002	6.01	4-20	2.99	10.9	0.12	F288
KL-ADL	6	14	20.4	umc1002	6.01	6-20	3.80	13.7	0.36	F288
8-O-4diFA	6	116	147.2	bnlg1702	6.05	108-122	3.42	12.4	0.01	F288
KL-ADL	6	126	151.1	g073COV	6.05	124-128	3.25	11.8	-0.35	F271
ADL/NDF	6	128	151.8	g146lac	6.05	126-130	14.30	42.5	-0.33	F271
KL/NDF	6	128	151.8	g146lac	6.05	126-130	7.80	26.1	-0.560	F271
Sga	6	128	151.8	g146lac	6.05	126-130	8.83	29.0	-0.45	F271
Sga-pCA	6	128	151.8	g146lac	6.05	126-130	11.90	36.9	0.97	F288
IVNDFD	6	128	151.8	g146lac	6.05	126-130	10.38	33.1	1.38	F288
DINAGZ	6	128	151.8	g146lac	6.05	126-130	4.82	17.0	0.92	F288
EsterFA	6	132	153.4	g169WRKY	6.05	130-134	9.40	30.5	0.25	F288
5-5diFA	6	132	153.4	g169WRKY	6.05	130-134	5.25	18.4	0.01	F288
EtherFA	6	134	153.7	bnlg345	6.05	130-136	4.44	15.8	-0.05	F271
pCA	6	152	162.5	g175MYB	6.07	146-156	2.48	9.1	-0.26	F271
KL/NDF	6	154	162.8	g061CCCH	6.07	148-158	2.09	7.8	-0.26	F271
IVNDFD	6	160	164.0	g149Dyn	6.07	156-164	3.24	11.8	0.66	F288
DINAGZ	6	160	164.0	g149Dyn	6.07	156-168	2.67	9.8	0.65	F288
KL-ADL	9	82	27.9	dupssr19	9.03	80-84	4.09	14.7	0.45	F288
KL/NDF	9	84	31.2	umc2337	9.03	80-90	3.87	13.9	0.34	F288
pCA	9	84	31.2	umc2337	9.03	78-90	4.44	15.8	0.35	F288

GE interaction effects were thus at least 7 times (and often nearly 10 times) lower than genotype effects for most traits. QTL analyses could be therefore relevantly investigated based on means over environments for each trait, even if some specific environmental information was likely lost. Based on genotype MS, genetic variation was greatly higher for KL lignin than for ADL lignin, even it has been shown that the ADL lignin trait was the most correlated with cell wall degradability (Table 2).

The correlation between the two cell wall degradability traits (IVNDFD and DINAGZ) was high and equal to 0.87, nevertheless indicating that nearly 75 % of the trait phenotypic variation was not similarly estimated by the two traits. Correlations shown between these two traits and the investigated cell wall biochemical components were nearly similar (Table 2), with negative effects of syringaldehyde and etherFA releases, and ADL/NDF contents, but no effect of esterFA release, corroborating previous observations (Méchin et al, 2000; Barrière et al, 2008; Riboulet et al, 2008b). The ratio of S unit acylated by pCA (estimated by the Sga-pCA trait) was (unexpectedly) weakly related to cell wall degradability, and a little more to lignin contents ($r = -0.31$ and -0.58 with ADL/NDF and KL/NDF, respectively). The soluble fraction of lignins (estimated by the difference between KL/NDF and ADL/NDF) was not related to the ADL lignin content ($r = 0.08$), but was highly correlated with the KL lignin content ($r = 0.91$). As was shown for KL/NDF, the solubilized fraction of lignins during the van Soest attack was not related to cell wall degradability.

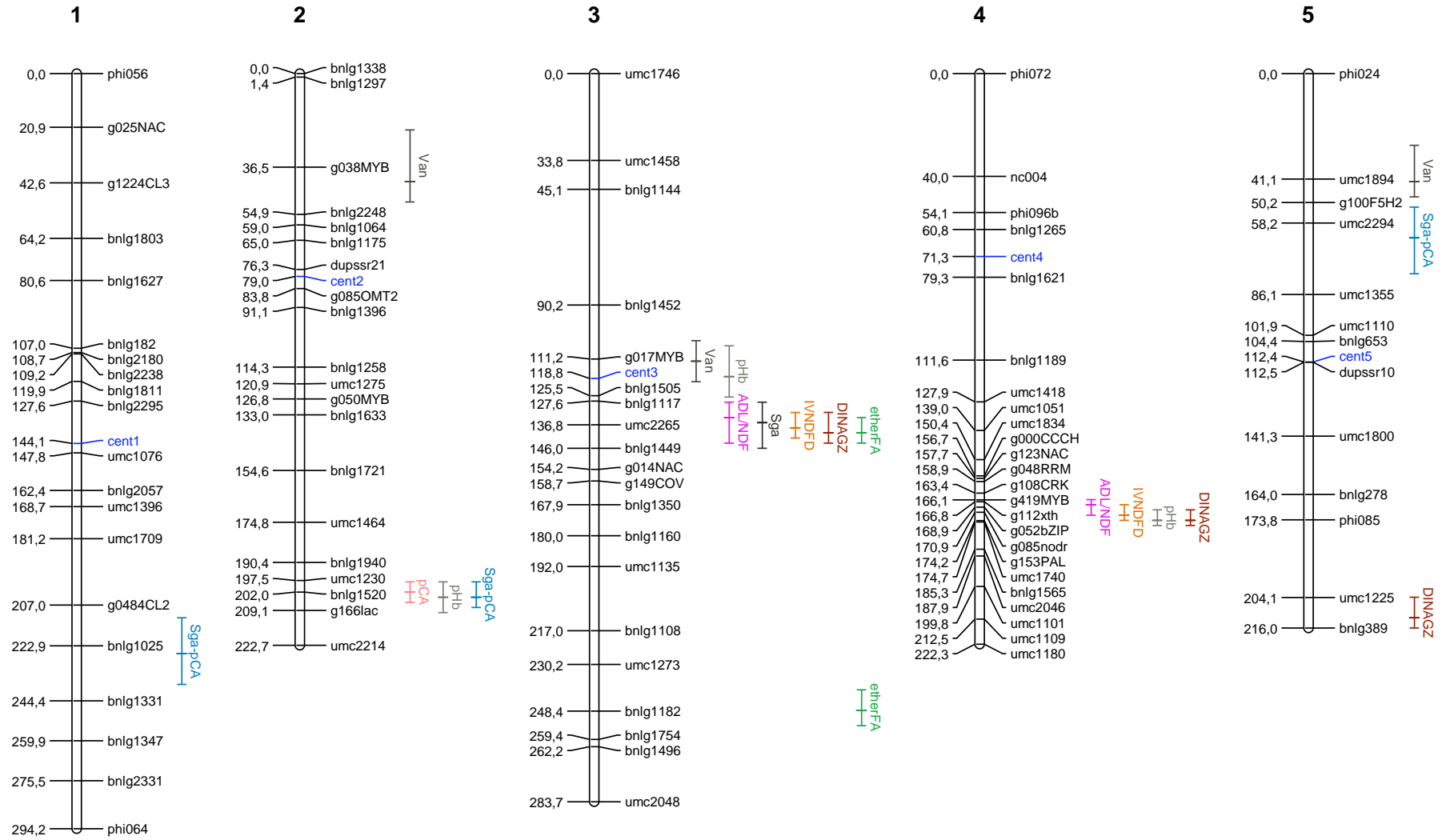
Map of the F288 x F271 progeny

The genetic map of the F288 x F271 progeny was improved, based on 173 markers and had a total length of 2,153 cM (Courtial et al, 2012d), that is a little longer than the 1759 cM of the UMC98 standard reference map in the MaizeGDB database. This fact could be partly related to large monomorphous areas, especially on chromosomes 7 and 10, likely a consequence of the common Co125 ancestry of the two F271 and F288 parental lines. The average distance between markers, which was equal to 13.2 cM, corresponded to variable situations. Distances between consecutive markers thus ranged from 83 cM between bnlg1686 and bnlg1200 on chromosome 7, to only 2.1 cM between g465NAC and g149Dyn on chromosome 6. Such differences between chromosomes and areas also resulted from the targeted marker mapping (Courtial et al, 2012d) in areas where QTLs were previously shown (Roussel et al, 2002).

LOD thresholds and QTL analysis

Based on the permutation-test method (Churchill and Doerge, 1994), LOD threshold estimates equal to 3.0, 3.5, and 4.5 yielded experiment-wise error rates equal to 10, 5 and 1 %, respectively. QTLs were therefore considered for LOD higher or equal to 3.0. However, a LOD threshold comprised between 2.4 ($P < 0.25$) and 3.0 was considered for QTLs colocalizing with more significant ones (an IVNDFD QTL was however considered with a still lower LOD value in bin 4.09). When including all considered traits, 42 QTLs were observed on seven chromosomes (Table 3, Figure 1). No QTLs were

Figure 1 - QTL analysis for cell wall related traits in the F288 x F271 RIL progeny (legends as in table 1).



observed on chromosomes 7, 8, 10, while largest monomorphous areas were present on chromosomes 7 and 10. Conversely, most QTLs were detected on chromosomes 6 (19 QTLs) and 3 (8 QTLs), with only 4 QTLs on chromosomes 2 and 4, 3 QTLs on chromosomes 5 and 9 and one QTL on chromosome 1.

QTLs for lignin content and lignin structure

Three and four QTLs were mapped for ADL/NDF and KL/NDF, respectively, which colocalized only in bin 6.05. Each ADL/NDF QTLs explained from 42 to 11 % of the observed phenotypic variation in this lignin trait, while KL/NDF QTLs explained from 26 to 8 %. All ADL/NDF QTLs had increasing alleles originating from F271, while it was the case for only two KL/NDF QTL, located in bin 6.05 and 6.07. An epistatic interaction was significant between ADL/NDF QTLs located in bins 4.09 and 6.05 (additive value +0.10). The three QTLs for the soluble fraction of lignin (KL-ADL) colocalized with KL/NDF QTLs, corroborating the greater genetic variation for KL lignins than for ADL lignins. Eight QTLs related to the monomeric composition of lignin were detected, out of which one QTL for vanillin and one QTL for syringaldehyde explained up to 20 and 29 % of the phenotypic trait variation, respectively. A colocalization between lignin structure QTLs was mapped only one time, in bin 3.04, for *p*-hydroxybenzaldehyde and vanillin QTLs. Four QTLs for the ratio of acylated S units were shown, two of which were in isolated positions, another one colocalizing with a pCA QTL in bin 2.08, and one colocalizing with a Sga QTL and other trait QTLs in the bin 6.05 cluster of QTLs.

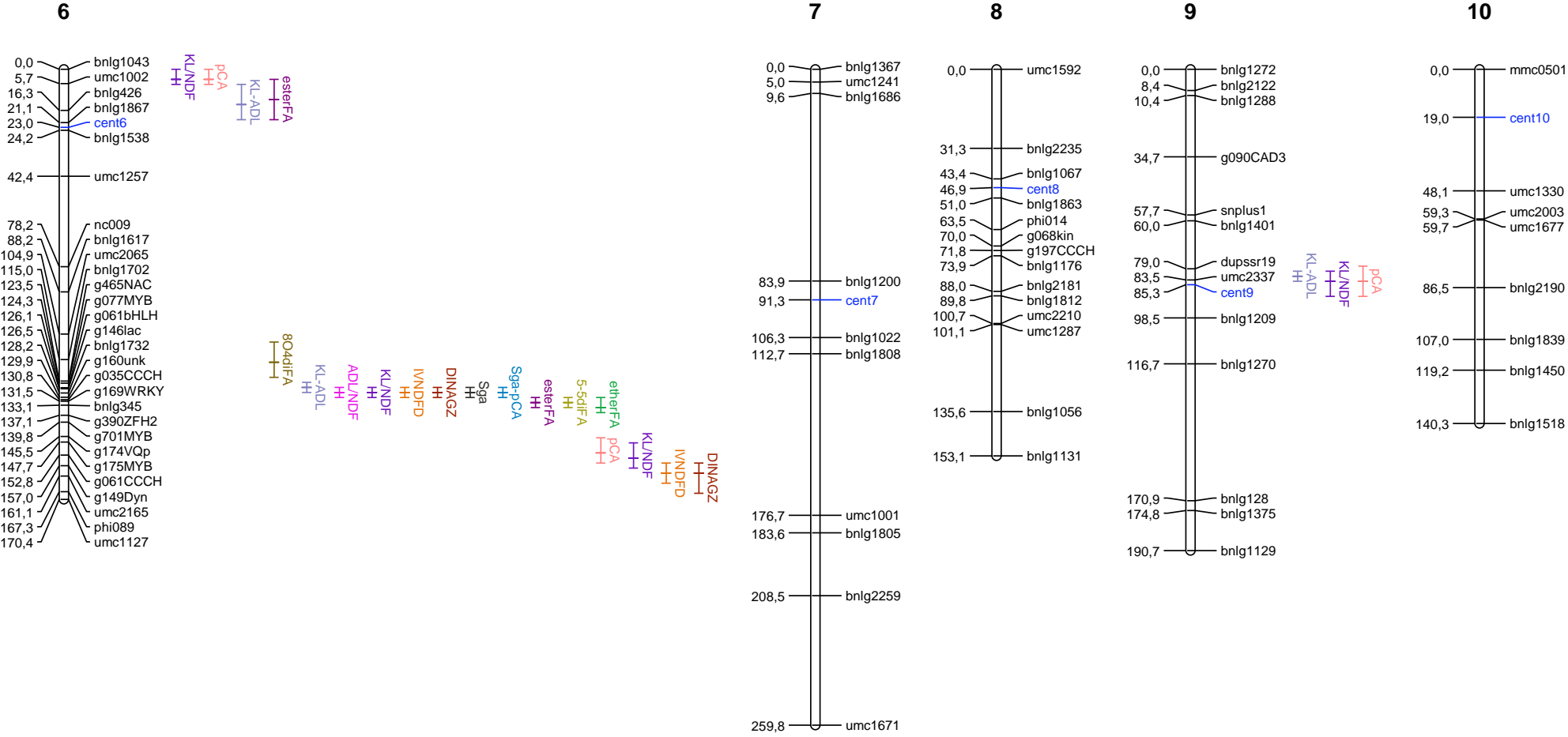
QTLs for *p*-hydroxycinnamic acids

Eleven QTLs of *p*-hydroxycinnamic acid content were detected, including four QTLs of *p*-coumaric acid and seven QTLs of ferulic and diferulic acids. Only one QTL of *p*-coumaric acid colocalized with QTLs of *p*-hydroxybenzaldehyde, in bin 2.08, illustrating possibly a limited oxidation of part of *p*-coumaric acid into *p*-hydroxybenzaldehyde during the alkaline nitrobenzene attack. QTLs of *p*-coumaric acid explained between 9 and 27 % of the trait phenotypic variation. The three QTLs of etherFA explained between 16 and 25 % of the trait phenotypic variation, while the two QTLs of esterFA individually explained 11 and 31 % of the phenotypic variation. EsterFA and etherFA QTLs colocalized only in bin 6.05, simultaneously with a 5-diFA QTL. EsterFA alleles increasing trait originated from F288 while the reverse was shown for etherFA with increasing alleles originating from F271. Alleles increasing diferulate releases originated from F288.

QTLs for cell wall digestibility

Nine QTLs for cell wall digestibility, including four IVNDFD QTLs and five DINAGZ QTLs, were detected in only five positions as all the IVNDFD QTLs colocalized with DINAGZ QTLs (Table 3). Four of the cell wall digestibility QTLs located in bins 3.06 and 6.05 explained at least 17 % of the variation, up to 33 %. Alleles increasing cell wall digestibility originated from F288 in four genomic

Figure 1 - continued- QTL analysis for cell wall related traits in the F288 x F271 RIL progeny (legends as in table 1).



locations and from F271 only for the isolated DINAGZ QTL in bin 5.08, corroborating the decreasing effect on cell wall digestibility of the F271 alleles. Cell wall digestibility QTLs colocalized with ADL/NDF QTLs three times, in bins 3.06, 4.09, and 6.05. Colocalizations also occurred two times with syringaldehyde and etherFA releases, in bins 3.06 and 6.05.

Cluster of QTLs in bin 3.06, 4.09, and 6.05 areas

Ten QTLs were mapped in a 12 cM long area of bin 6.05, which clustered in two locations, with seven QTLs in position 126 - 128 cM for cell wall degradability, lignin content and syringaldehyde QTLs, and three ferulate and diferulate QTLs in positions 132 - 134 cM. All these QTLs had major effects, with R² values ranging between 12 and 43 % whatever the trait. A little upstream was shown a QTL for 8-O-4 diFA (position 116 cM), while a little downstream was shown a cluster of three cell wall degradability and *p*-coumaric acid QTLs (positions 152 - 160 cM). It could thus be considered that four close QTL positions existed in bin 6.05 and 6.07 area (positions 116, 128, 133, and 156 cM, respectively).

Similarly, a cluster of five colocalizing QTLs that also comprised QTLs for the two cell wall degradability traits, for ADL/NDF content, and for etherFA and syringaldehyde releases was mapped in bin 3.06. These QTLs also had high R² values. In addition, a cluster of four QTLs including cell wall degradability QTLs was also shown in bin 4.09, with colocalizations with ADL/NDF and *p*-hydroxybenzaldehyde QTLs. These QTLs had little less R² values. Finally, QTLs located in bins 3.06, 4.09, and 6.05 cumulatively explained nearly 70 and 60 % of the phenotypic variation for ADL/NDF and cell wall traits, respectively, and 40 % of the phenotypic variation for syringaldehyde and etherFA releases. In both QTL positions, alleles increasing lignin content, lignin monomer release, pCA and etherFA, and decreasing cell wall degradability, originated in the three bins from F271. In bin 6.05, esterFA and etherFA QTLs colocalized, but with increasing alleles originating from F288 and F271, respectively.

Comparison between the two sets (2002 and 2012) of QTL investigations

Results obtained from the current experiments for lignin content and degradability traits mainly corroborated those previously obtained (Roussel et al, 2002; Courtial et al, 2012d), even if the latter experiments were based on plants without ears, whereas whole plants were harvested for the previous ones. However, for such cell wall related traits, the presence of ears could lead to different estimates only because of the presence of husks and highly lignified cob. In addition, climatic conditions were also different in the beginning and end of the 2000 decade, especially with warmer and dryer May and June. Contrary to what was shown in the older analyses, no QTLs for lignin or degradability traits were shown in the current one on chromosomes 1 and 2, while QTLs were similarly shown on chromosomes 4, 6, and 9. Moreover, mapped QTLs were estimated either at the same position (chromosome 6), or at positions differing by less than 4 cM. Colocalizations between cell wall

degradability QTLs and lignin structure or *p*-hydroxycinnamic acid QTLs in the current experiment gave new clues in the search for underlying determinants.

Candidate genes underlying colocalizing QTLs

The search for candidate genes underlying QTLs for lignin content and cell wall degradability located in bins 6.05 and 6.07 has also been previously investigated (Courtial et al, 2012c; Courtial et al, 2012d), pointing out the possible role of the *ZmMYB hv5-like* gene (GRMZM2G077789) and the *AtLac17* ortholog (GRMZM2G146152) in position 151 Mbp, the role of one MYB in position 162 Mbp, as well as the role of one WD40 gene in position 163 Mbp, of one NAC and two CCCH zinc finger genes in position 165 Mbp, and finally of two OVATE genes in positions 154 and 159 Mbp., Three xylan α -arabinosyl transferase of the GT47 family (GRMZM2G898668, GRMZM2G056702 and GRMZM2G448834) and one MYB (AC203535.4_FG001) colocalized with the cluster of QTLs located in bin 3.06 (position 163 Mbp) and were thus possible candidate genes. In bin 4.09 (position 230 Mbp), the cluster of four QTLs colocalized with the *ZmMYB42* gene, a COV1-like gene, a PAL-like gene, and a KNAT7-like gene (GRMZM2G419239, GRMZM2G123790, GRMZM2G153871, GRMZM2G433591, respectively, in positions 217 to 235 Mbp). In addition, in bin 3.09, the etherFA QTL (position 218 Mbp) colocalized with the GRMZM2G094428 arabinoxylan CoA-acyltransferase (PF02458 family, position 225 Mbp). These latter acyltransferase genes are currently the putative strongest candidates for encoding enzymes catalyzing the transfer of a feruloyl group onto an arabinoxylan chain in grass cell walls. No clear candidate genes were shown for the cluster of ferulate QTLs located just downstream of the cell wall degradability QTLs in bin 6.05.

Discussion

Comparison between the older and current sets of QTL investigations

Despite different climatic conditions, number of locations, number of map markers, ear presence or absence in harvested plants, and a lower RIL set in the current experiment, QTL detection for commonly investigated traits was in good agreement for analyses carried out with almost a 10 year difference. Most of detected QTLs were thus stable across environment and therefore provided relevant targets for breeding processes. QTLs of chromosomes 1 and 2 that were not shown in the current experiments were also those having among the lower LOD values. The four ADL/NDF, KL/NDF, IVNDFD, and DINAGZ QTLs in bin 6.05 were located at the same 128 cM (151.75 Mbp) position with high LOD values in the two sets of experiments, a fact greatly strengthening that at least one major factor involved in cell wall biosynthesis and assembly was located in this genomic area.

QTLs and cell wall trait variation in the F288 x F271 progeny

Based on phenotypic correlations, lignin content only explained 50 % of the variation for cell wall degradability, corroborating previous observations. Among the other cell wall considered traits, the vanillin release, as an estimate of the G unit proportion in lignins, did not appear as an important factor in this progeny, based on the fact that no colocalizations were shown between any of the three vanillin QTLs and any cell wall degradability QTL. Similarly and unexpectedly, only one pCA QTL out of four colocalized with cell wall degradability QTLs, while on the contrary, in the two F838 x F286 and RIo x WM13 progenies (Barrière et al, 2008; Barrière et al, 2012), all pCA QTLs colocalized with cell degradability QTLs. Corroborating the limited relationships between pCA and cell wall degradability in this progeny, the only colocalization between Sga-pCA and cell wall degradability QTLs was shown in bin 6.05, a position with a syringaldehyde QTL, but no pCA QTL. In addition, significant high negative correlations were also shown between pCA and cell degradability in different RIL progenies (Méchin et al, 2000; Riboulet et al, 2008b; Barrière et al, 2009a). The two syringaldehyde QTLs colocalizing with cell wall degradability QTLs could illustrate a cell wall maturity effect, with a longer and a greater lignin deposition typified by the later deposition of S units. This colocalization could also be related to the fact that the doubly methoxylated S units are more involved in β -O-4 linkages, which were shown to induce a more extended conformation of the lignin polymer inclined to adsorb all along the cellulose fibers and preventing their accessibility to enzymes (Zhang et al, 2011).

Based on currently available results, esterFA release was not considered as influencing cell wall degradability, while etherFA and diFA cross-linkages impeded cell wall degradability. However, it is not known how well released etherFA reflects total ferulate cross-linking in cell walls. Released etherified ferulate may thus account for as little as 15% of total cross-linkages because C-C, 8-O-4 styryl ether, and biphenyl ether coupling of ferulate to lignins are not yet determined due to the limitations of current solvolytic methods (Grabber et al, 2000). Anyway, the correlation between etherFA content and cell wall degradability was always shown as negative, both in maize and in other grasses (Casler and Jung, 1999; Méchin et al, 2000; Lam et al, 2003; Riboulet et al, 2008b; Jung and Phillips, 2010). Conversely, the positive relationship between diFA release and cell wall degradability was considered as explaining not the intensity of the corresponding cross-linkages, but the accessibility and susceptibility of diFA bridges in the cell wall towards alkaline hydrolysis (Riboulet et al, 2008b; Barrière et al, 2009a).

Questioning the number of QTL underlying determinants

Colocalizations between cell wall degradability, ADL/NDF and/or lignin structure, etherFA and/or diFA QTLs, as was shown in bins 3.05/06 and 6.05, were also observed in bins 1.02, 1.05, and 1.07 in the F838 x F286 progeny (Barrière et al, 2008) and in bins 2.06 and 3.04 in the RIo x WM13 progeny (Barrière et al, 2012). The current experiment strengthened the idea that the cell wall degradability

QTLs are probably driven by different interrelated underlying traits. In bin 6.05, a lignin content and structure determinant could be assumed in position 128 cM, with secondly a ferulate cross-linkage determinant in position 132 cM. Similarly, in bin 3.06, the cell wall degradability QTLs were probably simultaneously related to etherFA cross-linkage, to syringaldehyde, and to ADL/NDF QTLs. Colocalizations between cell wall degradability, ADL/NDF, and syringaldehyde could correspond to a shared mechanism involved in lignin biosynthesis and duration of lignin deposition. A greater proportion of S units in the lignin polymer could also correspond to a more linear polymer, with a greater masking effect on carbohydrates. Complementarily, colocalizations between cell wall degradability, etherFA, and diFA could correspond to another genetic mechanism involved in FA biosynthesis, cross-linkages between arabinoxylan chains, and cross-linkages between arabinoxylans and lignins. Simultaneous colocalizations between cell wall degradability and both lignin and ferulate related traits led to question the possible underlying genetic determinant(s). A cluster of linked genes involved in the different mechanisms of cell wall biosynthesis and assembly is likely the simplest situation to consider, and was strengthened by the presence of two close QTLs in bin 6.05. However, a single co-regulating "master" factor located upstream in the pathway of cell wall assembly cannot be definitely ruled out, even though no candidates are currently conceived for both lignin and ferulate components. The involvement of different genomic determinants in cell wall degradability variation strengthen the possibilities of breeding for higher value of this trait without deleterious effects on agronomic value (yield, biotic and abiotic stress tolerance).

MYB genes as candidates underlying cell wall content and degradability QTLs

From investigations in several species, NAC and MYB genes have been shown to be the major regulators of secondary wall lignification and assembly (Zhong et al, 2008; Gray et al, 2012; Grima-Pettenati et al, 2012). The *ZmMYB42*, which colocalized with lignin and cell wall degradability QTLs in bin 4.09, has been shown to negatively regulate the expression of several genes of the lignin pathway (Fornalé et al, 2006; Sonbol et al, 2009). *ZmMYB31*, that was also shown with a strong repressing effect on lignin-related gene expression (Fornalé et al, 2006; Fornalé et al, 2010), also colocalized with cell wall trait QTLs in three other RIL progenies (Courtial et al, 2012c). In addition to *ZmMYB31* and *ZmMYB42*, several other MYB factors are very likely involved in the regulation of maize lignification, including orthologs of *EgMYB2* which are transcriptional activators (Goicoechea et al, 2005). Considering six RIL progeny and 35 "meta"-QTL positions (Courtial et al, 2012c), 41 *ZmMYB* putatively involved in lignification colocalized with QTL in 26 positions. Two of these *ZmMYB*, located in bins 1.07 and 3.04 (GRMZM2G037650 and GRMZM2G017520), were also differentially expressed between F271 and four RILs of higher degradability (Courtial et al, 2012a), strengthening their involvement in maize cell wall assembly.

Arabinosyl-transferase genes as candidate genes underlying cell wall ferulate QTLs

The colocalization between the PF02458 arabinose-CoA-acyltransferase gene and the etherFA QTL located in bin 3.09 strengthened the probable involvement of members of the CoA-dependent acyltransferase PF02458 (Finn et al, 2008) gene family in the feruloylation of arabinoxylan chains (Mitchell et al, 2007; Piston et al, 2010). This family has nine members in maize, located in bins 3.03 (2), 3.09, 6.03, 6.05, 8.03 (2), and 8.05 (2). An esterFA QTL located in bin 8.05 was also located in the vicinity of two PF02458 arabinose-CoA-acyltransferase genes (GRMZM2G050072, GRMZM2G050270) in the RIo x WM13 progeny (Barrière et al, 2012). Similarly, another member of this family (GRMZM2G060210) was located between an etherFA and an esterFA QTLs in bins 8.03 in the F838 x F286 progeny (unpublished data). In addition, glycosyltransferases of the GT47 family, are considered in grasses with both xylan α -1,2- and α -1,3-arabinosyl transferase activities,. These enzymes thus allow the transfer of an arabinosyl residue onto a xylan chain. Missing or reduced GT47 activity would theoretically compromise the grafting of arabinose on xylan chains. As a consequence, a lower intensity of ferulate cross-linkages could be expected, with colocalizations of encoding genes with ferulate QTLs. Among the eight orthologs of GT47 α -1,3-arabinosyl transferase genes of the maize genome (Courtial et al, 2012c), a cluster of three duplicated genes colocalized with QTLs in bin 3.06. However, no candidate genes were shown for several other ferulate and diferulate QTLs, a fact that was also observed several times in the two F838 x F286 and RIo x WM13 progenies. The low-ferulate *sfe* mutant (Jung and Phillips, 2010) is currently one of the rare available tracks towards the elucidation of cross-linkage assembly in grasses. Similarly, no candidate genes were identified underlying any of the QTLs for the three aldehydes representative of the three H, G, and S lignin monomeric units. In addition to the limited number of colocalizations between candidate genes and cell wall degradability traits also shown in the F288 x F271 RIL progeny (Courtial et al, 2012d), this situation illustrates how little is known about phenolic component biosynthesis and assembly pathways in grasses.

Conclusions

Colocalizations were thus detected between cell wall degradability QTLs, lignin content QTLs, *p*-coumarate and/or lignin structure QTLs, and ferulate-driven cross-linkage QTLs. This trait colocalization could likely correspond to different genetic organizations, with either a common and unique upstream co-regulating factor or more probably the clustering of different genes and regulation factors. QTLs located in bin 3.05/06, 4.09, and 6.05 explained nearly 65 % of the genetic variation for both ADL/NDF and cell wall degradability. Moreover, cell wall degradability variations are not only related to lignin content, but also to other cell wall traits including lignin structure and S unit acylation by pCA, and ferulate cross-linkages. Considering all observed QTLs, no colocalizations occurred in

this RIL progeny experiment between cell wall degradability trait QTLs and any member of the gene families involved in monolignol biosynthesis, except with a putative PAL-like gene in bin 4.09. Based on present knowledge and the currently observed colocalizations, the most probable candidate genes were probably the *ZmMYB42*, COV1-like and KNAT7-like genes in bin 4.09, the *ZmMYB hv5*-like gene and COV-LCV2-like gene located a little downstream QTL positions in bin 3.06, and the *ZmMYB hv5*-like gene in bin 6.05, while other proposed candidates were likely more questionable. In addition, the GT47 glycosyltransferase and PF02458 arabinose-CoA-acyltransferase gene located in bins 3.06 and 3.09 should be considered as a relevant candidate for arabinose feruloylation. Nevertheless, the nearly 300 candidate gene list involved in phenolic compound biosynthesis and lignified tissue assembly, which included members of several families of transcription factors and tissue patterning regulation factors, appeared as partly incomplete for the understanding of the cell wall degradability variation in maize lines. The involvement of genes of still unknown function cannot be excluded, and similarly for genes of which any role in cell wall assembly has not yet been described. Moreover, gene identification was currently only based on the sequence of the reference B73 genome. Because large variation in genome sequences have been highlighted between maize lines, including PAV (presence - absence variation) and CNV (copy number variation) in addition to the well-known SNP and INDEL variation (Springer et al, 2009; Beló et al, 2010; Lai et al, 2010; Swanson-Wagner et al, 2010), it is must also be considered that the genetic determinant is not present at QTL position in the B73 sequence, correspond to a missing gene or cluster of gene in F271 or F288, as well as gene duplication in the investigated lines.

Different strategies towards the discovery of gene underlying QTL position could then be considered. Following several QTL investigations for cell wall related traits, a meta-analysis of these data should make more accurate both QTL positions and values of their respective support intervals. Positional cloning and the use of NIL lines should also allow reducing greatly QTL support intervals. With the decreasing costs of genome sequencing, targeted BAC sequencing and comparison of allelic variation between the parental lines in the QTL support intervals could also be considered. Considering that differences in cell wall degradability between lines could be related to differences in gene expression, a quantification of transcript levels among chosen progenies, based on massive sequencing, will point out the impact of each QTL on the overall transcriptomic profile. Such transcriptomic variation should give clues in the search for the gene function of the candidate to be discovered. Finally, association mapping based on large panels of lines with similar maturity could point out genes or genomic locations related to lignin traits and cell wall degradability.

Acknowledgements

This work has been funded by the maize breeding companies (Advanta, Caussade Semences, Limagrain Europe, MaisAdour, Monsanto SAS, Pioneer Génétique, Pau Euralis, R2n RAGT Semences, SDME KWS France, Syngenta seeds) involved in the ProMais - INRA "ZeaWall" network on maize cell wall lignification and degradability. Christiane Minault, Dominique Denoue and Pascal Vernoux were in charge of seed multiplications and field experiments carried out at INRA Lusignan. Seed multiplications were also managed at INRA St Martin de Hinx by Jacques Laborde.

CHAPITRE IV - SÉQUENÇAGE CIBLÉ DE BACs SITUÉS DANS
LE BIN 6.05 SOUS DES QTLs LIÉS À LA DÉGRADABILITÉ
DES PAROIS DÉTECTÉS DANS LA FAMILLE
RECOMBINANTE F288 x F271

4.1. Avant propos

L'amélioration de la carte génétique des RILs F288 x F271, à la suite de l'addition ciblée de marqueurs dans les régions des QTLs ainsi que dans les régions insuffisamment marquées, a permis de préciser les positions des QTLs, l'importance de leurs effets, et de réduire la taille de leurs intervalles support, avec les limites liées au faible effectif de RILs et celles liées à la méthode de détection des QTLs par cartographie par intervalle composite CIM (chapitre 2). Parmi l'ensemble des QTLs de dégradabilité des parois et des caractères associés (teneur en lignines, en acides *p*-coumarique, férulique, diféruliques, teneur en sous-unités H, G et S), un cluster de QTLs à effets forts a été localisé dans le bin 6.05 (chapitres 2 et 3), avec l'essentiel des QTLs colocalisant entre 126 et 130 cM, correspondant aux positions physiques 151,1-152,3 Mbp chez le génome de référence B73. Cette région de 1,2 Mpb est ainsi appelée par la suite "hot spot".

Afin d'identifier les déterminants sous-jacents à ces QTLs majeurs, une stratégie de séquençage de cette région a été entreprise avec la création de deux banques BAC des parents F288 et F271, et le séquençage ciblé des BACs ayant inséré une partie de cette région d'intérêt.

Cette stratégie est apparue nécessaire pour accéder à la variabilité structurale et allélique existant entre F288 et F271 en raison (1) de l'absence de gène candidat "évident" parmi les gènes présents dans la région correspondante dans la lignée de référence maïs B73, (2) de l'existence chez le maïs d'une forte variabilité entre génotypes, plus complexe que les SNP (Single Nucleotide Polymorphism) et INDEL (Insertion/DEletion) classiquement décrits, et constituée de nombreuses variations de type PAV (Presence Absence Variation) et CNV (Copy Number Variation). La mise en évidence de séquences de génome insérées ou perdues, de même que celle de gènes répétés, ne peut être réalisée qu'à partir du séquençage de grands fragments. Il en est de même pour l'insertion de transposon(s) supplémentaire(s), dans un génome déjà constitué de près de 80 % d'éléments transposables, qui peuvent par ailleurs avoir un effet régulateur ou inactiver des gènes.

4.2. Résultats

A partir de banques BAC ayant une couverture de 1,3 X pour la lignée F288 et 1,4 X pour la lignée F271, il n'a pas été à ce jour possible de trouver l'ensemble des BACs permettant de couvrir toute la région des QTLs du bin 6.05 chez les deux parents. De plus, l'assemblage *de novo* des BACs séquencés en 454 se heurte à la fréquence des éléments répétés dont les séquences sont extrêmement proches et bien souvent plusieurs contigs sont obtenus. Il manquerait ainsi, respectivement, environ 68 % et 55 % de la région "hot spot" chez les deux lignées F288 et F271. Les régions séquencées étant souvent à des positions différentes entre les deux génotypes, seulement 18 % de la région a pu être comparé entre les deux lignées. Malgré ces limitations, 72 % des régions chevauchantes entre F288 et

F271 présentent seulement des homologies partielles, voire même aucune homologie, révélant ainsi un grand polymorphisme entre F288 et F271. Ces différences correspondent à la fois à des différences structurales (inversions, translocation de séquence, PAV) mais aussi à des polymorphismes de séquence génique (SNP et INDEL). Des différences similaires sont observées entre les génomes des deux lignées étudiées et la lignée de référence B73.

4.3. Conclusions et perspectives

L'approche de banques BAC à faible couverture a tout d'abord été choisie pour accéder rapidement aux séquences correspondant à la région à caractériser. Cette approche, dès les premiers séquençages de BAC, a permis de mettre en évidence une forte variabilité entre les deux génotypes étudiés. L'augmentation de la couverture des banques BAC en utilisant notamment une nouvelle enzyme de restriction et la définition de nouvelles amorces dans les gènes présents aux extrémités des BACs séquencés devraient permettre d'accéder à l'ensemble de la région étudiée chez F288 et F271, et donc de comparer cette région dans son intégralité. La technique de séquençage paired-end permettra, quant à elle, très probablement d'avoir un assemblage des BACs plus optimal. Il semble même que son utilisation sera complètement nécessaire à l'achèvement des comparaisons de séquences de lignées chez le maïs. L'analyse des données obtenues et les travaux de séquençage n'étant pas terminés, une analyse complète de l'ensemble des données sera nécessaire pour pouvoir identifier, parmi l'ensemble des variations génomiques, lesquelles peuvent réellement contribuer aux variations de dégradabilité des parois et des caractères phénotypiques associés. Au vu du grand polymorphisme existant entre ces lignées, cette identification risque toutefois de s'avérer difficile. En plus des gènes, il ne faudra pas exclure que les déterminants soient d'autres structures impliquées dans la régulation comme des miRNA ou des rétrotransposons. L'insertion d'un transposon dans une partie régulatrice devra ainsi être prise en compte, y compris très en amont du gène comme c'est le cas pour le gène de "domestication" *tb1* (Studer et al, 2011). La cartographie fine de cette région à partir d'une population en cours d'obtention devrait permettre d'affiner à nouveau la position des QTLs et ainsi faciliter l'identification des déterminants génétiques impliqués dans la formation des parois secondaires lignifiées.

Article 4 -Targeted BAC sequencing during search for candidate genes underlying a major cell wall related QTL in an early maize progeny (in progress)

Audrey Courtial^{1,2,3}, Hélène San Clemente^{2,3}, Genséric Beydon⁵, Clémentine Vitte⁴, Stéphane Cauet⁵, Hélène Bergès⁵, Jacqueline Grima-Pettenati^{2,3}, Yves Barrière^{1*}

¹ INRA, Unité de Génétique et d'Amélioration des Plantes Fourragères, 86600 Lusignan, France

² Université de Toulouse; UPS; UMR 5546, Laboratoire de Recherche en Sciences Végétales; BP 42617 Auzeville, 31326 Castanet-Tolosan, France

³ CNRS; UMR 5546; BP 42617, 31326 Castanet-Tolosan, France

⁴ INRA, UMR de Génétique Végétale, Ferme du Moulon, 91190 Gif sur Yvette, France.

⁵ INRA, CNRGV, CS 52627, 31326 Castanet-Tolosan, France.

* corresponding author: E-mail: yves.barriere@lusignan.inra.fr

Key words: maize, degradability, lignin, cell wall, genome, BAC, sequencing

Abstract

Major QTLs for cell wall degradability and related traits have been detected using the RIL (recombinant inbred lines) progeny of the cross between the maize lines F288 x F271, (Roussel et al, 2002; Courtial et al, 2012b; Courtial et al, 2012c). In order to identify the genes underlying these QTLs located in the bin 6.05 region, we have decided to compare the genomic sequence of this region between the two parental lines. We first have constructed a low-coverage BAC library for each line: F288 (~ 1.3 X) and F271 (~ 1.4 X) and performed a pool screening using primers designed approximately each 100 kb based the maize B73 reference genome. Sequencing of the BAC clones allowed to access to ~ 32 % and ~ 45% of the sequences corresponding to the region of interest in the F288 and F271 genotype, respectively. Large differences between these two lines were highlighted as 72 % of the overlapping region between F288 and F271 presented either partial homologies or no homology, corresponding to INDEL (Insertion/ DELetion) of long sequence, sequence order change, as well as PAV (Presence Absence Variation) and allelic polymorphisms. The complete sequencing of this region and further analyses could allow identifying, among these genome differences, which could contribute to the large part variation for cell wall degradability and related traits between the two genotypes. The highlighted genetic determinants will permit to better understand the lignified cell wall formation in grasses and to set up selection programs based on these genes for maize degradability improvement for cattle feeding, as well as for the production of second-generation biofuel.

Introduction

Concerns over global climate changes, together with a growing worldwide demand for energy, pointed out the crucial need of alternative resources that can displace fossil fuels. First generation biofuels based on cereal grains compete with food supplies, while second generation biofuels based on lignocellulose materials open new avenues valorizing for a large part agricultural and woody residues. However, the biological conversion of cell wall carbohydrates, mainly located in the secondary lignified plant cell walls, into fermentable sugars is hindered by their association with lignins, a phenolic polymer rather resistant to chemical and microbiological attacks. A similar situation was faced since decades in breeding programs towards higher digestibility and energy value of forage plants, and traits related to cell wall degradability were then especially investigated in maize for its silage use.

Lignification and the development of a vascular system have allowed plants to leave aquatic habitats and to acquire erect growth, even if the earliest vascular thickenings were not lignified and if the cell wall lignification appeared in the outer cortex of Devonian land plants, 400 million years ago (mya), probably as a defense mechanism against pathogens (Boyce et al, 2003). Lignins, with cellulose, contributed to the mechanical properties and structural integrity of tissues. Lignins also impart hydrophobicity to vascular elements allowing water and nutriment transportation. Finally lignins are involved in mechanisms related to disease, pest, and abiotic stress tolerance, and lignin-deficient tracheary elements may collapse as their cell wall cannot stand the negative pressure generated during long and/or intensive transpiration periods (Cochard, 2002; Cochard et al, 2008). However, lignins association with other matrix components greatly influences cell wall properties, including the degradability of structural polysaccharides and their energy value for cattle feeding or biofuel production.

Most studies related to histology, biochemistry, genetics and genomics of lignification have been conducted in woody dicotyledonous or gymnosperm species. This focus is due to the major interest in reducing the cost and environmental impacts of lignin removing from woody plants during pulping for paper production. However, nearly one third of the earth's vegetative cover is indeed dominated by grass ecosystems (Jacobs et al, 1999), and grasslands are thus a major source of nutrients for wild and domesticated herbivores, with also important grasses and maize silage production in more developed countries. Similarly, plants considered for second generation biofuel production also are grasses, including C4 species (maize, sorghum, panicum, sugarcane, miscanthus, ...) and straw of C3 small grain cereals (rice, wheat, barley, triticale, ...). In addition to studies in woody species and model dicotyledonous plants (*Arabidopsis*, *zinnia*, tobacco), most information devoted to grass lignification is based on maize which is both one of the most economically important crop and also a classical

model for plant research in genetics and genomics. Only rare and scattered information is yet available concerning lignification in rice or in the two new model grass species *Brachypodium distachyon* for C3 plants and *Setaria italica* for C4 plants, even if increasing cell wall investigations are expected in the latters.

The divergence between monocotyledons and dicotyledonous plants is very ancient and dated to 200 mya, while the first grasses, appeared at forest margins nearly 65 mya, in late Cretaceous or early Tertiary, likely a little later after dinosaurs extinction (Wolfe et al, 1989; Herendeen and Grane, 1995; Jacobs et al, 1999; Kellogg, 2001). Vascular organization is remarkably different between monocotyledons and dicotyledonous plants. In contrast to dicotyledonous plants, the vascular system of monocotyledons is characterized by the absence of bifacial cambium and secondary growth. In monocotyledons, lignification proceeds from an intercalary meristem in each internode, with vascular bundles scattered, penetrating radially and present in medulla and cortex. However, the evolution of lignified vascular tissues is considered to derive in all plants from a preexisting poly-phenolic pathway, with differential regulations and targeted depositions in different cell types of different species (Boyce et al, 2003). Tremendous specificities in grass cell wall organization have nevertheless emerged over the millennia of evolution and compensate for the absence of woody tissue. The first major difference is certainly the crucial role of ferulate and diferulate cross-linkages in secondary walls, between guaiacyl (G) units of lignins and arabinoxylans, and between arabinoxylans chains, that greatly contributed to plant standability and tissue stiffness. The second difference is likely related to the importance of *p*-hydroxyphenyl (H) units in the lignin polymer, which reach nearly 3 % of the units released by thioacidolysis and have consequences on the frequency of resistant condensed linkages. Thirdly, grass cell wall are typified by the high frequency of syringyl (S) lignin unit acylated by *p*-coumaric residues, inducing likely a more linear and spanning polymer (Zhang et al, 2011).

Despite the currently observed large diversity in species and growth habits, plants of the grass family have well conserved genomes and a monophyletic origin (Devos, 2005). From an ancestral grass karyotype with 5 chromosomes, an intermediate grass ancestor has emerged with 12 chromosomes, which is at the origin of all actual grasses through the differentiation of a Pooideae ancestor (12 chromosomes) and a Panicoideae ancestor (10 chromosomes), after different duplication, reciprocal translocation, and nested chromosome fusion events (Murat et al, 2010). The maize genome has also undergone an additional whole-duplication amphipolyploidization event between 5 and 12 mya, distinguishing it from its close Panicoideae relatives (Schnable et al, 2009). According to Salse (2012), the maize karyotype history could be summarized by the equation $(((12 + 2 - 4) \times 2) + 7 - 17]$, reflecting that the monocotyledonous ancestral intermediate with 12 chromosomes went through 2 fissions and 4 fusions to a Panicoideae ancestor with 10 chromosomes, that underwent a whole genome duplication, forming a 20 chromosome intermediate, that have finally undergone 17

chromosome fusions and 7 fissions, reaching the modern maize genome structure with 10 chromosomes. In addition to its segmental allotetraploid structure, the size of the maize genome has greatly increased during the last 3 million years, with an extensive proliferation of long terminal repeat (LTR) retrotransposons (San Miguel et al, 1998) that compose more than 75 % of the maize B73 genomic sequence (Schnable et al, 2009).

Understanding lignification in grasses has a fundamental interest and will contribute to the understanding of grass and plant organization, evolution and phylogeny. Understanding lignification has also a practical interest, as the improvement of plant energy value for cattle feeding and for biofuel production are both fully related to lignin quantity, monomeric composition and linkages, and to the complex relationships between lignins and cell wall carbohydrates. Identification of genomic areas involved in maize lignified cell wall assembly and degradability has been especially based on QTL approaches. Among numerous locations with cell wall related QTLs (Courtial et al, 2012a), QTLs explaining major parts of the genetic variation for numerous traits have been shown in the bin 6.05 after investigations in the F288 x F271 RIL progeny (Roussel et al, 2002; Courtial et al, 2012b; Courtial et al, 2012c). The main part of these QTLs colocalized between 126 - 130 cM, corresponding to the physical positions 151.1 - 152.3 Mbp in the reference genome B73. We refer to this 1.2 Mbp region as the bin 6.05 “hot spot QTL”. Underlying candidate genes have been searched for, based on physical QTL positions, and based on the evidenced genes shown after the whole sequencing of the maize line B73 (Schnable et al, 2009; www.maizesequence.org). Sequencing of some candidate genes could be performed by PCR to investigate polymorphism between the parental lines and search the genetic determinant underlying the QTLs. However, it has been shown that gene expression can be regulated by the presence of a noncoding region far upstream of gene transcription start site, as shown for the maize *Teosinte branched1 (tb1)* gene (Studer et al, 2011; Zhou et al, 2011) and *APETALA2-like* gene (Salvi et al, 2007). In the case of *tb1*, one transposable element, located nearly 65 kb upstream the *tb1* gene, acts as an enhancer of this gene expression and increases the apical dominance in maize compared to its progenitor (teosinte). On the contrary, a Mite transposon in the non-coding region *Vgt1*, located 70 kb upstream of the Ap2-like transcription factor, disrupts the enhancement of this gene expression and delays the flowering time. Moreover, the particular structure of the maize genome limits the relevancy of the identification of candidate genes based on reference genome. Genome sequence comparison of maize lines indeed established that several hundred CNVs (copy number variation), as well as several thousand PAVs (presence-absence variation) could exist between two maize lines (Springer et al, 2009; Beló et al, 2010; Lai et al, 2010; Swanson-Wagner et al, 2010). Haplotype-specific PAVs between B73 and Mo17 thus contained hundreds of single-copy, expressed genes, a fact considered as contributing either to the large phenotypic diversity of maize, and to the great heterosis between lines (Springer et al, 2009). BAC libraries of the two parental lines F271 and F288 have therefore been constructed, followed by the identification of BACs underlying the major

QTL area of bin 6.05, and the whole sequencing of these considered BACs, in order to evidence the genetic and genomic variation really existing between these two lines, and possibly explaining the observed QTLs.

Materials and methods

Plant material

F288 and F271 plants were grown for 11 days in a growth chamber under 16 hours of light per day and with a temperature of 22 °C. Plants were maintained in complete darkness for 72 h before leaf sampling. Collected young leaves were frozen in liquid nitrogen, and then stored at - 80°C before extracting high molecular weight genomic DNA (HMW DNA).

BAC library construction

BAC library construction was performed at CNRGV (Toulouse, France). For each extraction, approximately 20 g of frozen leaf tissue, ground in liquid nitrogen with a mortar and pestle, were used to prepare megabase-size DNA embedded in agarose plugs. HMW DNA of the two genotypes was prepared as described by Peterson et al (2002) with modifications described in Gonthier et al (2010). Embedded HMW DNA was partially digested with *HindIII* (New England Biolabs, Ipswich, Massachusetts). Pulsed-field migration programs, electrophoresis buffer, and ligation desalting conditions were performed according to Chalhoub et al (2004). Restriction fragments were submitted to two size selection steps by pulsed-field gel-electrophoresis (PFGE), using a BioRad CHEF Mapper system (Bio-Rad Laboratories, Hercules, California). Size-selected partially digested DNA samples [from 100 to > 300 kilobases (kb)] were recovered from the agarose by electroelution and ligated to pIndigoBAC-5 *HindIII*-Cloning Ready vector (Epicentre Biotechnologies, Madison, Wisconsin). Competent *Escherichia coli* DH10B cells (Invitrogen, Carlsbad, CA) were transformed with each ligation mix by electroporation and transformants were selected on LB-X-gal- IPTG plates containing chloramphenicol. To evaluate the average insert size of each library, DNA was isolated from about 50 randomly selected BAC clones for each library using the NucleoSpin® 96 Plasmid kit (Macherey-Nagel, Düren, Germany). Digestion was performed with the rare cutter *NotI*, and digestion products were analyzed by PFGE. All fragments generated by *NotI* digestion contained the 7.5 kb vector band and various insert fragments. According to the average insert size and to the maize genome size, the number of BAC clones required to obtain a low-coverage BAC library (~ 1x haploid genome equivalents) was estimated and white colonies were picked using a QPix (Genetix, New Milton, Hampshire,UK) into 384-well plates (Genetix) containing LB freezing medium with chloramphenicol supplemented with glycerol. After incubation of 16 h, plates were stored at -80°C.

Table 1 - Primers used for BAC screening, based on the B73 genomic sequence (Schnable et al. 2009, www.maizegenome.org, release v2 5b60).

chromosome	B73 position average (bp)	Position difference (bp)	Gene identifier	Description	Primer names	Primer sequences (5' → 3')	Tm (°C)	Amplicon size (bp)
6	151 126 489	104 071	GRMZM2G061906	bHLH transcription factor	GRMZM2G061906_F2 GRMZM2G061906_R2	ACGAAAGGTGCGTTTCAGAG CACAGAGCTCGTTGAACCTG	60 60	109
6	151 230 560	105 676	GRMZM2G150471	integral membrane Yip1 family protein	GRMZM2G150471_F1 GRMZM2G150471_R1	TTCGCCTGAGGTTTGCTTGATG AACTGGTCTGCCATGCCCTTTC	60 60	208
6	151 336 236	106 469	-	-	chr6v2_151336236_F1 chr6v2_151336236_R1	TCTATGTTCCGCAGTCGTTG GCACGTGTTTGTCCACTGAC	60 60	222
6	151 442 704	101 761	GRMZM2G059121	uracil phosphoribosyltransferase	GRMZM2G059121_F1 GRMZM2G059121_R1	TCCACGTGTACACAGGAATGA AAAGATTCTCGCAGGTCGAA	60 60	281
6	151 544 465	94 288	GRMZM2G146152	Laccase LAC17	GRMZM2G146152_F1 GRMZM2G146152_R1	CAAAATCGTTTGGCCTGTGCAAC ATGGCGCAACAACACATGACC	60 60	129
6	151 638 752	67 814	GRMZM5G850966	60S ribosomal protein L36 (RPL36B)	GRMZM5G850966_F1 GRMZM5G850966_R1	GGCTGCCAGAGATTATGTTCCAG ACCCTGGATGCATTGGAAGGAG	60 60	121
6	151 706 566	71 243	GRMZM2G701193	CysteinyI-tRNA synthetase	GRMZM2G701193_F1 GRMZM2G701193_R1	GCTATGACCCACCTTGAAA TTCAACCGCAACTGGCTATT	60 60	67
6	151 777 809	95 244	bnlg1732 (GRMZM2G012052)	UBC7 (Ubiquitin Carrier Protein 7)	bnlg1732_F bnlg1732_R	AACTTTTGGCATTGCACTGG CGTAAGTGCACACGGCATT	60 60	108
6	151 873 053	101 623	GRMZM2G143745	Glycine-rich cell wall structural protein 2	GRMZM2G143745_F1 GRMZM2G143745_R1	ATAAAGGGCCCGCTGTTC AAG ACTGGCAGGCTTACATCTTCGC	60 60	87
6	151 974 675	98 130	GRMZM2G108723	plant-specific domain TIGR01589 family protein	GRMZM2G108723_F1 GRMZM2G108723_R1	TCAACCTGAGCAAGGAAGAGTGC TTGACGTTGGCATGCTTCTCCAG	60 60	55
6	152 072 805	95 452	GRMZM2G399495	unknown protein	GRMZM2G399495_F1 GRMZM2G399495_R1	CTGCAAGAAAGTCCAGCAGCTT TCCTCTGCAAATCTCGGTGAGGT	60 60	136
6	152 168 257	89 143	GRMZM2G066432	protein kinase APK1B family protein	GRMZM2G066432_F1 GRMZM2G066432_R1	AGCAATGATAAGGGCGGGTGTG TGCACGAAGCCACTTCATCTGC	60 60	79
6	152 257 400	62 273	GRMZM2G160354	unknown protein	GRMZM2G160354_F2 GRMZM2G160354_R1	ACACGCACACTGCCAACGAG TCGAGTCCCAAGCGGAGTTCT	60 60	302
6	152 319 673	113 670	GRMZM2G318992	KIP1 (protein kinase interacting protein 1)	GRMZM2G318992_F1 GRMZM2G318992_R1	TGCAAGAAGCCACCGAACT GGGAACACCACCTTTCGTTGAC	60 60	61
6	152 433 343	26 068	GRMZM2G064679	unknown protein	GRMZM2G064679_F1 GRMZM2G064679_R1	AACATTGGGTCCAAGTACGG CTGAAGAAGCCCCACGAAC	60 60	140
6	152 459 411	75 377	GRMZM2G088409	protein kinase family protein	GRMZM2G088409_F2 GRMZM2G088409_R2	CGTACGTACAGACCCGAATG CTCTCACCACCACTCCTCA	60 60	104
6	152 534 787	111 235	GRMZM2G069737	protein hypothetical protein with F box domain	GRMZM2G069737_F1 GRMZM2G069737_R1	CGGAGCGAGTTCCTTCTTGAAC ACAAACTTGCCGTCTCAAGGC	60 60	64
6	152 646 022	101 119	GRMZM2G372390	unknown protein	GRMZM2G372390_F1 GRMZM2G372390_R2	TTGGATTCTGGACGCAGGATG TTGTCTCCATCAGCGCAACCAG	60 60	55
6	152 747 141	102 341	GRMZM2G067198	protein expressed with F box domain	GRMZM2G067198_F1 GRMZM2G067198_R1	AATTGAGCCGGTITGGAAACCG ACGAGGATATGCCACCTTGTG	60 60	84
6	152 849 481	98 589	GRMZM2G052939	protein hypothetical with F box domain	GRMZM2G052939_F1 GRMZM2G052939_R1	TCCCTTGAGATCGCAGGTTAGG ACCCATCGAAGACGAACACAC	60 60	84
6	152 948 070	101 854	GRMZM2G035601	zinc finger C3HC4-type RING finger	GRMZM2G035601_F3 GRMZM2G035601_R2	AAATCCTGAGCCATTTGGTG TGGTGCAACCTGCTAGCTTC	60 60	273
6	153 049 924	86 047	GRMZM2G443444	unknown protein	GRMZM2G443444_F1 GRMZM2G443444_R1	AGCAGGCGACTGGTACTATTG AAATGCTCCGCGACAAGGAGAC	60 60	102
6	153 135 971	97 018	GRMZM2G061105	protein ribonuclease P	GRMZM2G061105_F1 GRMZM2G061105_R1	GCCACGCCGTCATTGTTAACTG ACCGCTGAAGGCAGATTCTGTAG	60 60	118
6	153 232 989	33 157	IDP4165 (AC209257.4_FG004)	unknown protein	IDP4165_F IDP4165_R	ATGGCTTCCAGTCTCGTCG CCGTCCTGAAGCAGAAGC	60 60	256
6	153 266 146		GRMZM2G169966	WRKY (WRKY3/OsWRKY70)	GRMZM2G169966_F1 GRMZM2G169966_R1	ACTGCCCATCCTTTGACT CTGCCGCTTAGTCTGTTCT	60 60	261

BAC library screening and BAC clone characterization

In order to obtain BAC clones spanning the region of interest, primer pairs were chosen approximately each 100 kb based on the reference genome B73 (Schnable et al. 2009, www.maizesequence.org, RefGen v2) (Table 1). Primers were preferentially designed in genes using QuantPrime software (Ardivissov et al, 2008).

Positive BAC clones were identified according to the following procedure. For the first screen, the 384 clones contained in one 384-well plate were pooled together to obtain a plate pool (PP). PCR amplification of the different PPs allowed identifying the plates containing at least one insert located within the targeted region. Then, row pools (RP) and column pools (CP) were prepared from the positive plates and screened by PCR in order to determine BAC coordinates (plate, row, and column numbers). Finally, positive BAC clones detected in pools were individually validated by PCR amplification. For all these identification steps, PCR were performed in 10 µl containing 1X EvaGreen master mix (Biorad), 0.125 µM of each primer on a CFX384 Real-Time PCR Detection System (Biorad). The PCR cycling program consisted of an initial denaturation of 2 min at 95°C, then 50 cycles for 2s at 95°C and 3s at 60°C, followed by melting curve program of 10s at 95°C, then raising of temperature between 65°C to 95°C with an increment of 0.5°C between each fluorescence acquisition. F288 and F271 genomic DNA were used as positive controls. Melting peaks were analyzed with the biorad software and positive clones were validated by visualization of PCR products on agarose gels. DNA of each positive BAC clone was isolated using NucleoFast® 96 PCR (Macherey-Nagel, Düren, Germany) in order to determine insert size and to sequence BAC-ends. Insert sizes were estimated after digestion with *NotI* and migration of digestion products in PFGE. BAC-ends were sequenced on the Applied Biosystems 3730 after amplification of clone insert using universal primers T7 and M13 reverse.

BAC clone sequencing and assembly

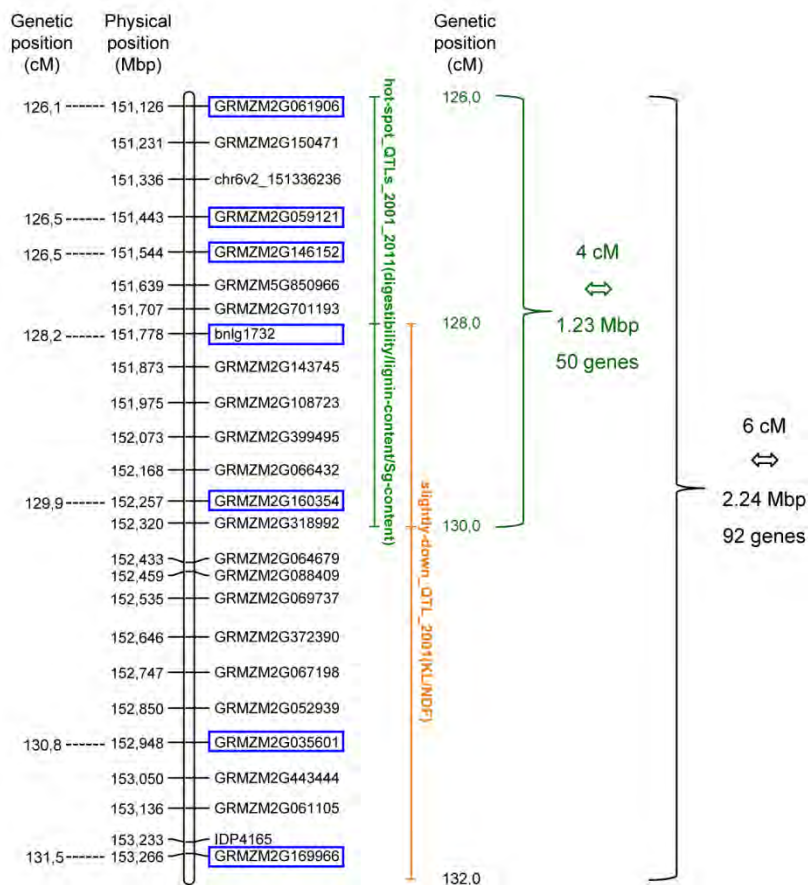
DNA of each positive BAC clone was isolated using NucleoBond® Xtra Midi kit (Macherey-Nagel, Düren, Germany) at CNRGV (Toulouse, France). The Roche/454 library construction and sequencing were performed by the Genomic Platform of Toulouse (France, <http://genomique.genotoul.fr/>). All clones were sequenced using a Roche GS FLX Life Sciences instrument (Margulies et al, 2006). *De-novo* assembly was performed using the Newbler software (454 Life Sciences, Branford, CT, USA). BAC-end sequences obtained from PCR amplification (see above) were aligned to corresponding BAC sequences obtained by 454 sequencing using BLAST ([blast2seq, www.ncbi.nlm.nih.gov/](http://blast2seq.ncbi.nlm.nih.gov/)) to evaluate the quality of the BAC extremities sequences (~ 1000 bp). Only contigs with size over 500 bp were considered for further analysis.

Table 2 - Characteristics of F288 and F271 BAC libraries

Libraries	F288	F271
Plate number	86	79
Clone number	33024	30336
Average insert size (kb)	100	100
Nuclear haploid genome equivalents	1.4 X	1.3 X
Cloning Site	HindIII	HindIII
Vector	pIndigoBAC-5 Hind III	pIndigoBAC-5 Hind III

Figure 1 - Genetic and physical positions of the primers used for the screening of the BACs in the region of interest, according to physical positions of B73 genome reference.

Within the region of interest (2.24 Mbp), we have focused our analysis on the so called “hot spot” region (1.23 Mbp) represented in green which corresponds to the main QTLs for cell wall degradability and related cell wall traits. Markers in a blue box correspond to the markers present on the genetic map of F288 x F271 RIL progeny and used for QTL detection. Genetic positions of these seven markers and of QTLs are shown (cM).



Gene prediction and annotation

De novo gene predictions were performed for all contigs using the gene structure prediction program FGeneSH (<http://linux1.softberry.com/all.htm>) with monocot matrix. Annotation of the predicted genes was then performed using BLASTP and BLASTX (Altschul et al, 1997) against the SwissProt database. The first significant hit (with a maximum E-value of e^{-20}) was selected for gene annotation.

Sequence comparisons

Overall sequences of the F271 and F288 genotypes were compared to the B73 reference genome through to BLASTN alignment of each BAC sequence to the orthologous B73 region (non filtered BLAST searches with an E-value of e^{-50}). The set of predicted mRNA and encoded proteins were also compared between F271, F288, and B73 by performing a BLAST search against the B73 reference genome protein set contained in the region of interest, using BLASTX (E-value of e^{-30}) and BLASTP (E-value of e^{-50}), respectively.

Sequence comparisons between F271 and F288 were carried out by pairwise aligning BACs screened with the same marker or with two consecutive markers (BLASTN, E-value $\leq e^{-50}$). F271 and F288 gene comparisons were based on a BLASTX of predicted RNA against predicted proteins with an E-value of e^{-10} . Sequence alignments of two similar genes were performed with DIALIGN (Morgenstern, 2004). A script was written to add exon translations in the three reading frames and protein sequence alignments.

Sequence arrangement and visualization

Position of the vector sequence as well as results from BLAST analyses described above were used to propose contigs organization and orientations within a given BAC. Alignments of two BACs screened with the same or two successive primers, for the same genotypes (BLASTN, 100 % identity) also allowed organizing contigs.

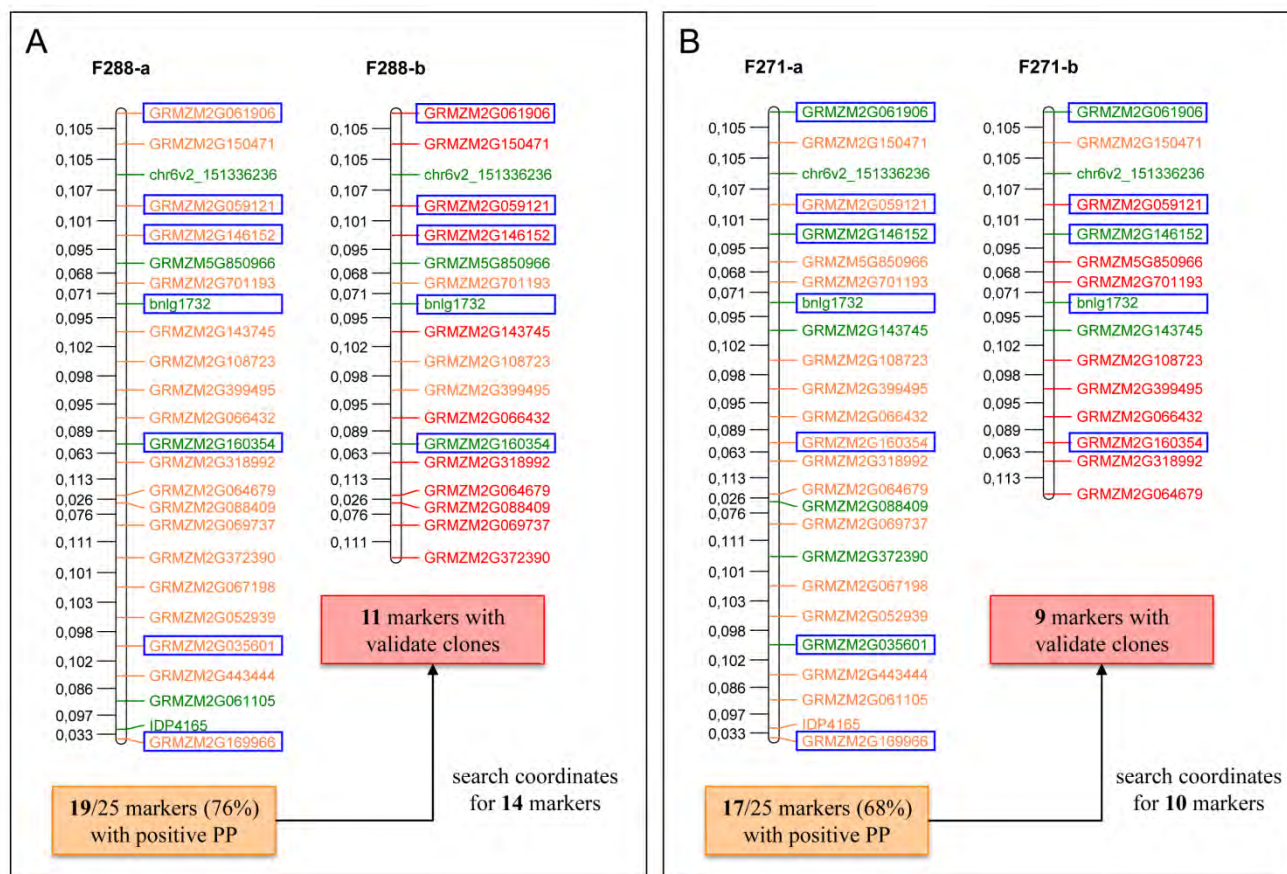
A website (www.polebio.lrsv.ups-tlse.fr/synthviewer) was especially developed to display contig organization within a BAC and BAC alignments between lines. This representation can include markers used for BAC screening, predicted gene positions, and homologies between BAC sequences.

Results and discussion

BAC library characterization

Two BAC libraries were constructed for the two parental lines F288 and F271, which exhibited differences in cell wall degradability and related traits. The main characteristics of the two BAC libraries called ZmaF288 and ZmaF271 BAC libraries are summarized in Table 2. The average insert

Figure 2 - Screening of F288 (A) and F271 (B) BAC libraries with markers highlighting bin 6.05 region involved in degradability. a and b correspond to pool plate (PP) screening and row/column pool (RCP) screening, respectively. Numbers at the left of chromosome correspond to the physical distance scale (Mbp) in the reference genome B73. Color of markers is given according to their screening status: green for markers with no positive pool plate (PP), orange for markers with at least one positive PP and red for markers with validated BAC clones. Markers in a blue box correspond to the markers present on the genetic map of F288 x F271 RIL progeny.



size of each library was estimated to 100 kb. Finally, 33024 and 30336 BAC clones were arranged in 384-well plates allowing to obtain low-coverage BAC libraries of 1.3 X and 1.4 X haploid genome equivalents of F288 and F271, respectively, considering that the haploid maize genome has a nearly 2.3 Gb physical length.

BAC library screening

We designed a set of 25 PCR-based molecular markers spaced approximately every 100 kb throughout the 2.24 Mbp region (Courtial et al, 2012c) of the B73 maize reference sequence corresponding to the region of interest, to screen BAC libraries of our two genotypes (Table 1; Figure 1). All the PCR markers except chr6v2_151336236, which was designed in an intergenic region due to the absence of genes between positions 151.25 and 152.44 Mbp, successfully amplified portions of genes. Among the targeted genes, seven corresponded to genes mapped in this region on the genetic map of F288 x F271 RIL progeny (Courtial et al, 2012c and chapter 3). Moreover, 14 primer pairs covered the 1.2 Mbp “hot spot” QTL region (positions 151.1-152.3 Mbp / 126-130 cM).

Positives PPs were found for 76% and 68 % of the primers, for F288 and F271, respectively (Figure 2). BAC coordinates were investigated only for markers located within or at the close vicinity of the hot spot QTL region (Figure 2), *i.e.* 14 markers for F288 and 10 markers for F271. Positive BAC clones were found for 9 and 11 markers in ZmaF271 BAC library and ZmaF288 BAC library, respectively. Three markers for F288 (GRMZM2G701193, GRMZM2G108723 and GRMZM2G399495) and one marker for F271 (GRMZM2G150471), for which BAC coordinates were searched, did not amplified row pools and column pools. The positive PP, first identified, could be false positives due to the absence of the screened gene in the 384 BAC clones and the presence of a close paralog as many genes belong to multigene families. Although 3D pools are known to generate false-positives (Yim et al, 2007; Luo et al, 2009; You et al, 2010), the rate is still low compared to conventional hybridization techniques (review in Ariyadasa and Stein, 2012). Screening BAC libraries by PCR assays remains one of the earliest methods of library screening and has been extensively used in early physical mapping projects (McPherson et al, 2001; Gregory et al, 2002).

Sequencing of selected BAC clones

The nine and eight positive BAC clones screened in the vicinity of the hot spot QTLs region, above the GRMZM2G372390 marker, for ZmaF288 BAC library and ZmaF271 BAC library, respectively were sequenced with Roche GS FLX sequencer and assembled using the Newbler assembly methodology (Figure 3 and Table 3). Excluding the 1K14 BAC for which sequencing did not work well, the sequencing produced on average 14,079 reads per sequenced BAC clone with a mean read length of 440 nucleotides. These clones were sequenced at different levels of sequence coverage ranging from 21X to 79X. On average 72 % of these reads were used for assembling the full sequences of each clone into a minimum of one and a maximum of 27 contigs, for ZmaF288_23M12

Figure 3 - Positions of the sequenced BACs in the bin 6.05 region involved in degradability. The names of the 454 sequenced BAC clones are noted in red in front of the corresponding markers. Numbers on the left of chromosome correspond to physical position (Mbp) of markers in the B73 reference genome.

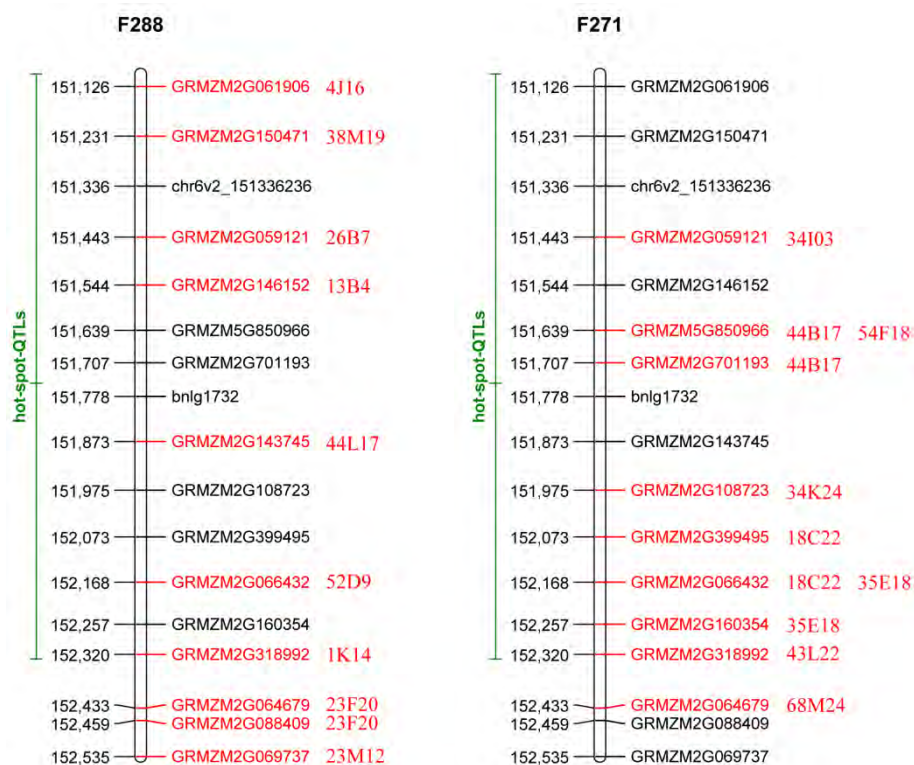


Table 3 - Sequencing (454FLX) and assembly of selected BAC clones.

Library	BAC	Nb of reads	Total reads length (kb)	% of total reads length in the assembly	DNA insert coverage (X)	Nb of large contigs (> 500bp)	Observed length (kb)	Assembly length (kb)
F288	23M12	13 083	5 635	74,1	39	18	105	97,5
	1K14	415	181	77,2	1	22	100	20,2
	04J16	15 828	6 973	72,4	43	9	115	117,0
	13B4	18 731	8 143	72,5	45	2	130	107,4
	26B7	9 722	4 312	71,4	38	1	80	92,7
	23F20	7 116	3 131	70,8	36	3	60	48,8
	38M19	21 194	9 346	71,9	79	7	85	81,5
	44L17	11 716	5 079	69,0	33	6	105	94,1
52D9	16 784	7 299	72,5	58	2	90	84,3	
F271	34I03	12 129	5 274	72,4	36	7	105	98,8
	34K24	19 508	8 539	67,8	41	12	140	127,9
	35E18	12 571	5 505	72,0	34	4	115	110,1
	54F18	13 111	5 630	72,2	40	6	100	99,4
	18C22	13 203	5 757	72,5	32	1	130	128,6
	68M24	17 936	7 810	76,3	49	13	120	137,8
	43L22	6 943	3 019	75,4	21	8	105	84,1
44B17	15 683	6 965	75,8	52	6	100	108,8	

and ZmaF288_26B7, respectively. The number of contigs with sizes over 500 bp varied among clones from 1 (for clone ZmaF288_23M12) to 18 (for clone ZmaF288_26B7).

Alignment of the BAC-end sequences (~ 1000 bp) obtained after PCR amplification with the corresponding 454 assembly sequence and alignment between overlapping BAC clones sequences of the same genotype validated the nucleotide sequences obtained from pyrosequencing (100 % of identity was obtained).

Sequence and structure comparison of the investigated lines to the reference genome B73

Assembled BACs were physically mapped onto the B73 genome thanks to the genes found in common and to the anchoring of the BAC-ends. This mapping allowed to estimate the region coverage (Figure 4). Sixty five percent of the BAC extremity positions were estimated only based on the presence of common genes since 35 % of the BAC-end corresponded to transposable elements matching with numerous positions in the maize genome. Several gaps remained in each genotype due to an absence of BAC coverage. These regions correspond to 68 % and 55 % of the “hot spot” QTL region in F288 and F271, respectively. The absence of several BAC clones in the library was observed even for corresponding markers mapped on the parental lines. This could be first related to the low coverage of the haploid maize sequence. This strategy was in fact used for rapid access to the large investigated region, limiting BAC clone massive picking steps and screening efforts. The absence of BAC clones in the library could also be the consequence of the use of only one restriction enzyme for library construct. When using a BAC library originated from partial digestion with a single enzyme, physical map may indeed contain gaps, due to cloning bias and unstable segments in the *E. coli* host (Yang et al, 2003; Ammiraju et al, 2005). Different restriction enzymes are therefore generally used to generate BAC libraries (Coe et al, 2002; Schulte et al, 2011).

Moreover, considered gap lengths are estimated based on the reference genome. It is then possible that these gaps are under- or over-estimated due to differences between parental line and B73 sequences. In the region close to 151.5 Mbp (as referred to the B73 sequence), the F288 genomic sequence seemed to be twice as long as the corresponding one in the B73 genome. For instance, the Zma26B7 BAC (position 151.44 Mbp) spanned nearly 93 kb in the F288 genome while it covered only 42 kb in the B73 genome. The same size difference was observed for the F288 107 kb long Zma13B4 BAC (position 151.54 Mbp) which only covered 65 kb in the B73 genome. Larger genome size in the investigated lines than in the reference line could make difficult the whole region screening.

In addition, large structural differences were observed between investigated lines and the B73 reference genome (RefGen v2) such as insertions, deletions, inversions of sequence positions (ordering), reverted sequences (Figure 5). However, considering the available results, filtered B73 genes were all found in F288 and F271 lines, except one gene in F271 in ZmaF271_18C22 (GRMZM2G415025, position 152.12 Mbp, aldose reductase).

Figure 4 - Estimated physical positions (Mbp) of ZmaF288 (orange) and ZmaF271 (pink) BAC on B73 genome. For each BAC clone, name and assembly length are noted. Gene identifiers on the left of the chromosome correspond to the B73 genes with genes written in blue being used as markers for BAC screening. Chromosome region coloured in green correspond to the hot-spot QTL region (151.1-152.3Mbp).

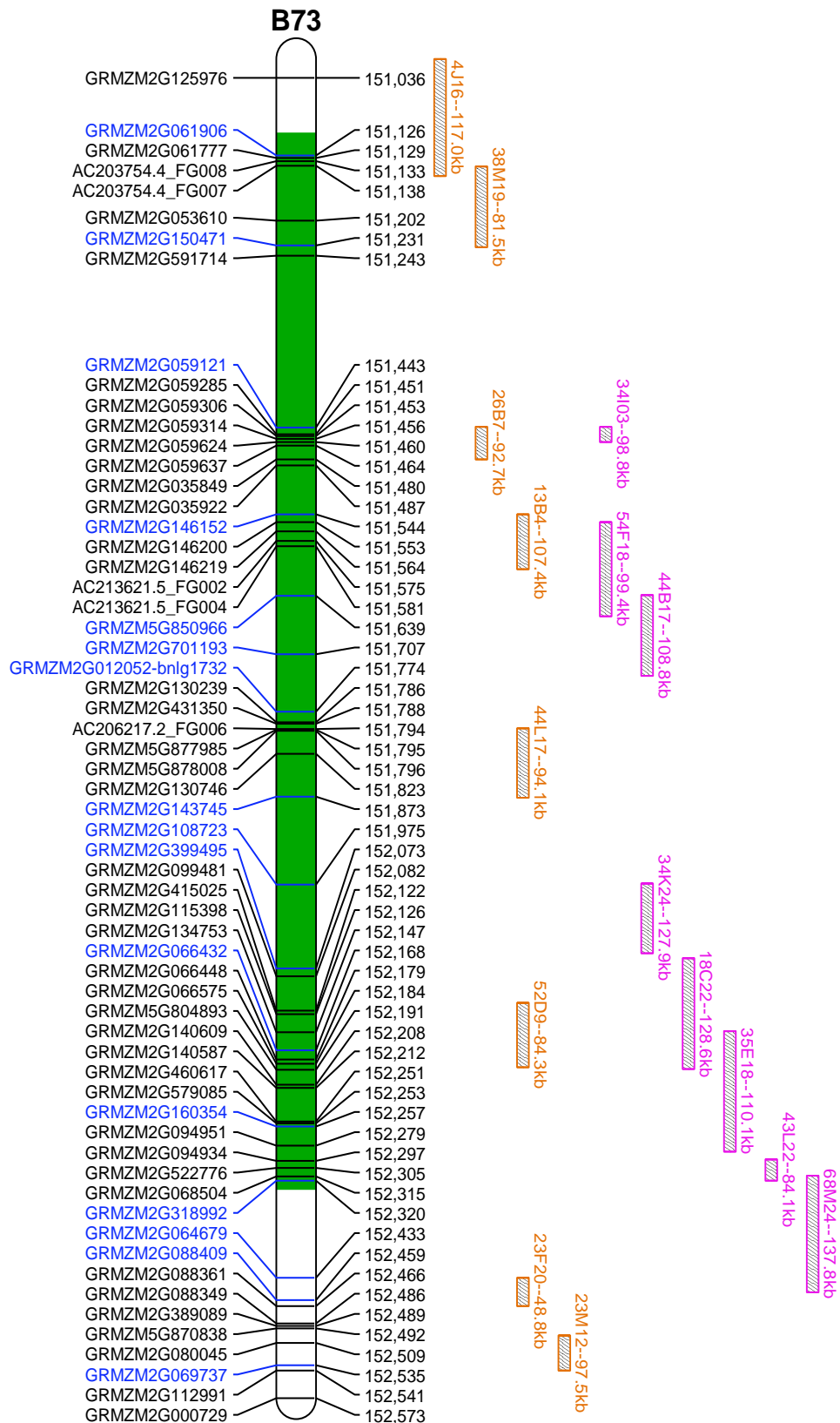


Figure 5 – Example of genome comparison between investigated genotype and B73 reference genome. Comparison of ZmaF288_26B7 BAC clone versus B73 reveals structural differences of genome. Other BAC alignments with B73 shown similar differences (figures not shown).

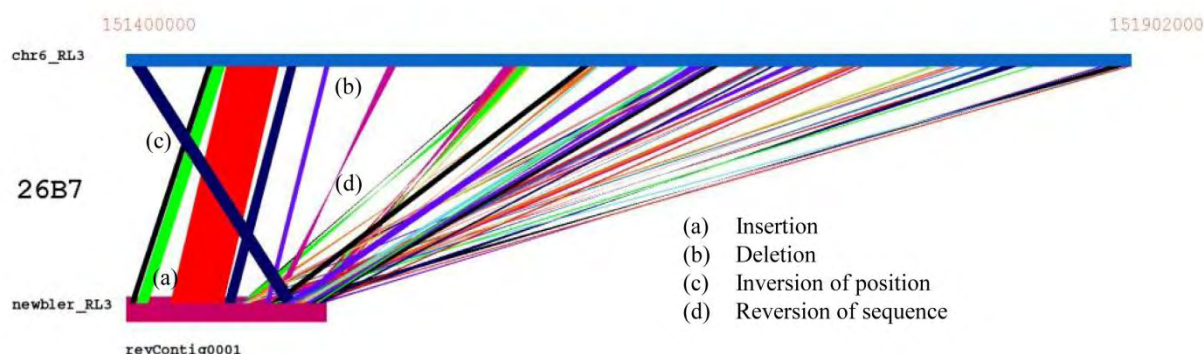
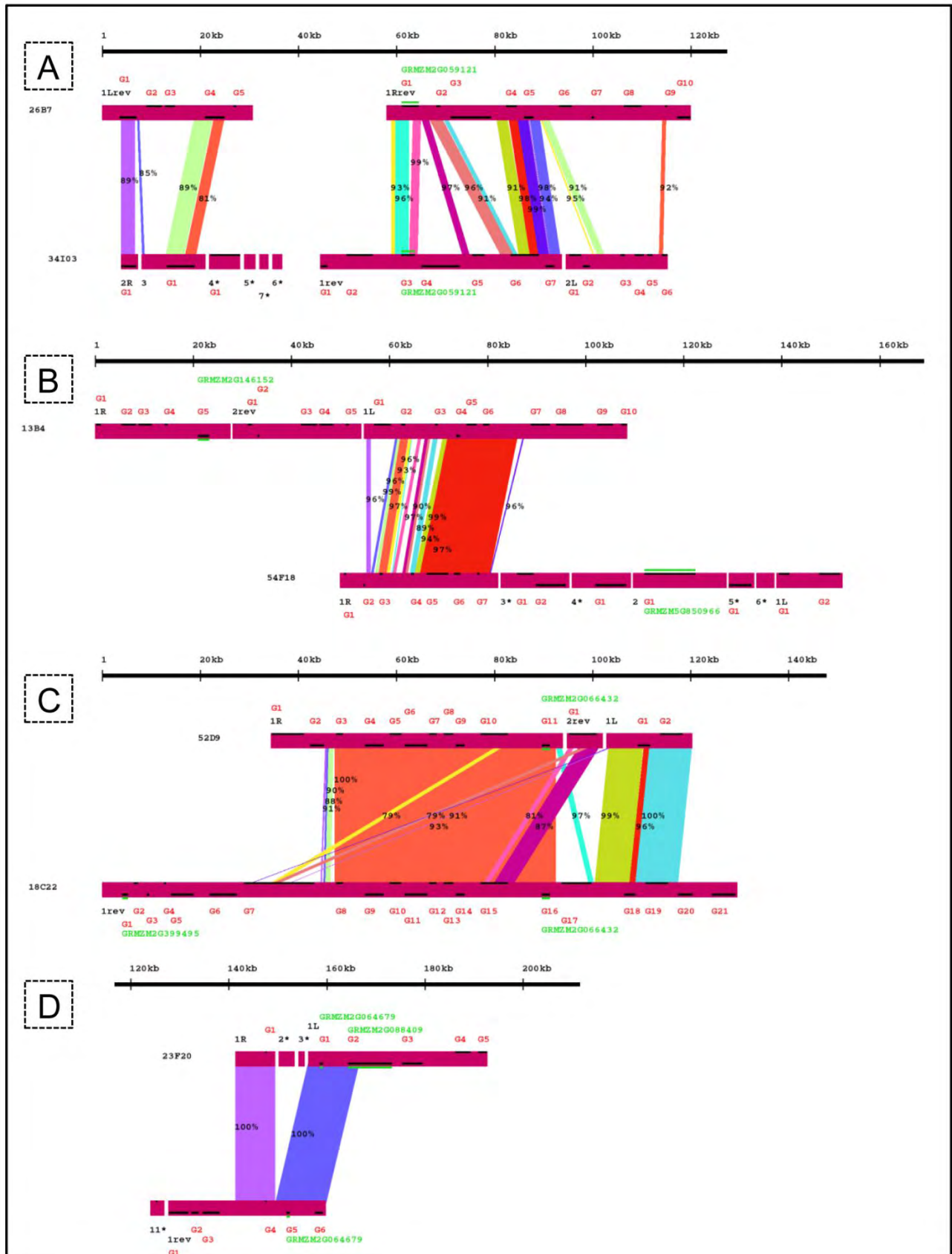


Table 4 - Comparison of the overlapping regions between F288 and F271. Homologies were considered when E-value is $\leq e^{-50}$ (BLASTN). Id : identity

Overlapping BAC	Overlapping length (kb)	Region length with 100 % Id (kb)	Region length with E-value $\leq e^{-50}$ and Id < 100 % (kb)	Region length with E-value > e^{-50} (kb)
ZmaF288_26B7 vs ZmaF271_34I03	98,8	0	32,2	66,6
ZmaF288_13B4 vs ZmaF271_54F18	58,8	0	23,5	35,3
ZmaF288_52D9 vs zmaF271_18C22	84,3	53,8	18,7	11,8
ZmaF288_23F20 vs ZmaF271_68M24	24	19	0	5
Total	265,9	72,8	74,4	118,7

Figure 6 - Overlapping BAC comparison between F288 and F271. This representation includes markers used for BAC screening (in green), predicted gene positions (G1, G2 ...), and homologies between BAC sequences.



Sequence and structure comparison of two parental lines

The sequences and structure of the F288 and F271 genotypes were compared for the four couples of BACs screened with the same marker or with two successive markers (*i.e.* ZmaF288_26B7 *versus* ZmaF271_34I03, ZmaF288_13B4 *vs* ZmaF271_54F18, ZmaF288_52D9 *vs* zmaF271_18C22, ZmaF288_23F20 *vs* ZmaF271_68M24). We did not compare ZmaF271_35E18 to ZmaF288_52D9, as ZmaF271_35E18 sequence is the same as ZmaF271_18C22 BAC sequence which is assembled in only one contig. The cumulated overlapping regions between the two genotypes covered only 266 kb of the 1.5 Mbp B73 region (18 %) located between the first and the last sequenced BAC.

Global sequence alignments (Figure 6 and Table 4) revealed that four zones showed 100 % identity between F288 and F271 and corresponded approximately to 27.4 % of the overlapping region length. Indeed, in the vicinity of the B73 position 152.433 Mbp, the two completely identical zones covered a large part (79 %) of the overlapping sequence between the two genotypes (Figure 6D). The other 100 % identical zones covered approximately 64 % of the region in common close to the position 152.168 Mbp (Figure 6C). The other remaining zones, which represented 72.6 % of the overlapping region length, presented either partial homologies or no homology at all. Around the B73 position 151.443 Mbp, 33 % of the overlapping region between ZmaF288_26B7 and ZmaF271_34I03, distributed in 17 zones, were presenting an E-value $\leq e^{-50}$, and an identity superior to 81% (Figure 6A). Around position 151.544 Mbp, 13 zones, which spanned on 40 % of the overlapping region between ZmaF288_13B4 and ZmaF271_54F18, showed an alignment with an E-value $\leq e^{-50}$ and had more than 89% identity. Around position 152.168 Mbp, ZmaF288_52D9 and ZmaF271_18C22 sequences were aligned, with an E-value $\leq e^{-50}$ and a percentage of identity comprised between 81-99%, on 12 zones which covered 22 % of the two overlapping BACs (Figure 6C). The remaining overlapping zones (45 %), not cited before, had an E-value $> e^{-50}$ and low identities ($< 79\%$) even no homology in the case of INDEL of long sequence. Several INDEL were found between F288 and F271, as for example the INDEL of 5 kb that distinguished the two overlapping BAC ZmaF288_23F20 and ZmaF271_68M24. One 5 kb part was also deleted in F271 compared to F288 (Figure 6B) and B73. This genomic part located between 151558067 and 151563237 bp in B73 was close to the QTLs position (151.75 Mbp). One gene (F288-13B4_contig00001_left_FGenesh1.1) was predicted at this INDEL position in F288 but was not predict in B73 (maizesequence.org, release v2) even if these two lines shown 99.5 % identity. Orthologs of this protein (731 amino acids) were only found in the first part of the protein (between 129-341 amino acids) and were uncharacterized proteins (first hit: uncharacterized protein LOC100383069 with an E-value of $6e^{-71}$).

Other differences between lines have been investigated at the gene level, including PAV (Presence Absence Variation), CNV (Copy Number Variation) and allelic polymorphisms (Table 5). Seventy two genes and/or transposable elements were predicted in both F288 and F271 genotypes in the overlapping regions. Among these genes, four genes were specific to F288 and two were specific to F271, as they were not found in F271 and F288 lines, respectively. Some other genes were missing in

Table 5 - Comparison of F288 and F271 genes. % Id = % Identity; %pos = % positives

BACF288	contig	gene	BAC F271	contig	gene	Remarks	annotation	Score	E-value	% Id	% pos	Query coverage (%)	Hit coverage (%)
26B7	1L_rev	G1	34I03	2R	G1		-	473	1,00E-161	70.2	74.4	88.7	44.4
26B7	1L_rev	G2	-	-	-	incomplete assembly	sp Q02817 MUC2_HUMAN Mucin-2	-	-	-	-	-	-
26B7	1L_rev	G3	-	-	-	incomplete assembly	sp Q9UR07 RTF23_SCHPO Retrotransposable element Tf2 155 kDa prot	-	-	-	-	-	-
26B7	1L_rev	G4	34I03	3	G1		sp P10978 POLX_TOBAC Retrovirus-related Pol polyprotein from transposon	311	5,00E-97	75.0	77.8	42.1	13.7
26B7	1L_rev	G5	-	-	-	incomplete assembly	-	-	-	-	-	-	-
26B7	1R_rev	G1	34I03	1_rev	G3		sp P93394 UPP_TOBAC Uracil phosphoribosyltransferase	362	1,00E-132	99.4	99.4	99.4	99.4
-	-	-	34I03	1_rev	G4		sp Q99315 YG31B_YEAST Transposon Ty3-G Gag-Pol polyprotein	-	-	-	-	-	-
-	-	-	34I03	1_rev	G5		-	-	-	-	-	-	-
26B7	1R_rev	G2	34I03	1_rev	G6		sp Q84TF0 AKRCA_ARATH Aldo-keto reductase family 4 member C10	136	2,00E-42	97.0	100.0	73.0	16.4
26B7	1R_rev	G3	-	-	-		sp P11369 POL2_MOUSE Retrovirus-related Pol polyprotein LINE-1	-	-	-	-	-	-
26B7	1R_rev	G4	34I03	1_rev	G6		sp Q84TF0 AKRCA_ARATH Aldo-keto reductase family 4 member C10	630	0.0	98.1	98.7	99.4	77.3
26B7	1R_rev	G5	34I03	1_rev	G7		sp Q84TF0 AKRCA_ARATH Aldo-keto reductase family 4 member C10	686	0.0	83.8	84.8	96.8	99.7
26B7	1R_rev	G6	-	-	-	incomplete assembly	sp Q9CAY3 GPAT5_ARATH Glycerol-3-phosphate acyltransferase 5	-	-	-	-	-	-
26B7	1R_rev	G7	-	-	-	incomplete assembly	-	-	-	-	-	-	-
26B7	1R_rev	G8	-	-	-	incomplete assembly	sp Q93VV0 ZDHC6_ARATH Probable S-acyltransferase At3g09320	-	-	-	-	-	-
26B7	1R_rev	G9	34I03	2L	G6		-	146	4,00E-46	80.2	84.9	95.9	35.0
13B4	1L	G1	-	-	-		-	-	-	-	-	-	-
13B4	1L	G2	54F18	1R	G3		-	434	1,00E-158	97.8	97.8	77.1	99.6
13B4	1L	G3	54F18	1R	G4		-	327	1,00E-118	97.1	97.6	99.2	99.4
13B4	1L	G4	54F18	1R	G5		-	84	1,00E-54	61.2	74.8	51.7	6.3
13B4	1L	G5	54F18	1R	G5		sp Q8W486 Y1491_ARATH DUF246 domain-containing protein At1g04910	426	1,00E-150	99.5	99.5	95.4	35.1
13B4	1L	G6	54F18	1R	G3		sp Q9SNC3 FLA4_ARATH Fasciclin-like arabinogalactan protein 4	812	0.0	98.1	98.8	99.7	99.8
13B4	1L	G7	-	-	-	incomplete assembly	sp P10978 POLX_TOBAC Retrovirus-related Pol polyprotein from transposon	-	-	-	-	-	-
13B4	1L	G8	54F18	2	G1		sp P10978 POLX_TOBAC Retrovirus-related Pol polyprotein from transposon	972	0.0	76.0	84.5	37.1	34.2
13B4	1L	G9	-	-	-	incomplete assembly	sp Q6ZQ06 WDR87_HUMAN WD repeat-containing protein 87	-	-	-	-	-	-
13B4	1L	G10	-	-	-	incomplete assembly	-	-	-	-	-	-	-
52D9	1R	G1	18C22	1_rev	G7		sp P10978 POLX_TOBAC Retrovirus-related Pol polyprotein from transposon	576	1,00E-176	47.3	57.9	44.5	23.1
52D9	1R	G2	-	-	-		sp P23901 ALDR_HORVU Aldose reductase OS=Hordeum vulgare	-	-	-	-	-	-
52D9	1R	G3	18C22	1_rev	G8		sp Q8BUV8 GP107_MOUSE Protein GPR107 OS=Mus musculus	938	0.0	99.8	99.8	99.8	99.8
52D9	1R	G4	18C22	1_rev	G9		sp Q32NW2 LENG8_XENLA Leukocyte receptor cluster member 8 homolog	1177	0.0	100.0	100.0	99.8	99.8
52D9	1R	G5	18C22	1_rev	G10		-	797	0.0	100.0	100.0	99.7	99.7
52D9	1R	G6	18C22	1_rev	G11		-	841	0.0	100.0	100.0	99.8	99.8
52D9	1R	G7	18C22	1_rev	G12		sp Q02817 MUC2_HUMAN Mucin-2 OS=Homo sapiens	968	0.0	99.6	99.6	99.8	99.8
52D9	1R	G8	18C22	1_rev	G13		sp P04146 COPIA_DROME Copia protein OS=Drosophila melanogaster	831	0.0	90.5	90.9	99.6	97.5
52D9	1R	G9	18C22	1_rev	G14		-	571	0.0	97.0	98.3	97.7	96.8
52D9	1R	G10	18C22	1_rev	G15		sp P10978 POLX_TOBAC Retrovirus-related Pol polyprotein from transposon	2761	0.0	97.7	97.8	99.9	99.9
52D9	1R	G11	18C22	1_rev	G16		sp Q9SFT7 Y3707_ARATH Serine/threonine-protein kinase At3g07070	800	0.0	99.5	99.5	99.7	99.7
-	-	-	18C22	1_rev	G17	incomplete assembly	-	-	-	-	-	-	-
52D9	2_rev	G1	-	-	-		sp P10978 POLX_TOBAC Retrovirus-related Pol polyprotein from transposon	-	-	-	-	-	-
52D9	1L	G1	18C22	1_rev	G18		p Q9SF32 IQD1_ARATH Protein IQ-DOMAIN 1	1112	0.0	93.9	94.3	99.8	99.8
52D9	1L	G2	18C22	1_rev	G19		sp Q8TFG9 YL61_SCHPO Putative GPI-anchored protein PB15E9.01c	2755	0.0	100.0	100.0	99.9	99.9
23F20	1R	G1	68M24	1_rev	G4		-	142	1,00E-48	100.0	100.0	98.6	98.6
23F20	1L	G1	68M24	1_rev	G5		sp Q9STN5 Y4833_ARATH Uncharacterized protein At4g08330, chloroplastic	308	1,00E-112	99.3	99.3	99.3	99.3
23F20	1L	G2	68M24	1_rev	G6		sp P10978 POLX_TOBAC Retrovirus-related Pol polyprotein from transposon	1064	0.0	100.0	100.0	40.6	99.8

one genotype but these results were questionable due to an incomplete assembly in the overlapping zone. Other genes with no transposable element function presented polymorphism between F288 and F2271 protein. For example, the fasciclin gene (F288-13B4_contig00001_left_FGenesh6.1 and F271-54F18_contig00001_right_FGenesh6.1), which is one candidate gene (Courtial et al, 2012c), encoded two proteins with 98.1 % identity.

Even if large gaps remain for each genotype at different physical positions, the already done sequencing allowed having an outline of the differences between B73, F288 and F271. All these differences have to be considered in a more comprehensive way and some have to be confirmed using PCR. For example, inversion of sequence positions could be due in some cases to a wrong position of contigs, absence of gene could reflect the absence of the corresponding sequence in the 454 sequencing reads and assemblies. When differences are observed within a contig, these differences are less doubtful.

Conclusion

Even if genome differences between F288 and F271 were evaluated in only approximately 18 % of the region involved in cell wall degradability traits, large differences between these two lines were highlighted as 72 % of the overlapping regions differed in the two lines. All these observed differences, absence of a region in a genotype or variations of nucleotide sequences could contribute to the phenotypic differences between F288 and F271. Such differences were also observed when each investigated lines were compared to the reference genome B73. Together, these results strengthened the need not to limit investigations of candidate genes in comparison with the reference genome, and pointed out the necessity to get the F288 and F271 whole genome sequences in this region of interest.

In the future, this project has to be continued to complete the sequencing of this region in the two genotypes and compare them in order to evidence the genetic and genomic variation really existing between these two lines, and possibly explaining the observed QTLs. On one hand, missing sequences will be searched focusing on several strategies. First, BAC library coverage will be improved by using a new restriction enzyme such as *Bam*HI. New primers could be designed for each genotype in genes located at BAC extremities to complete gaps. On the other hand, data analyses have to be continued and considered in a more comprehensive way. In addition, paired-end sequencing will be used to improve BAC assembly. Finally, simultaneous transcriptomic investigations of selected RILs differing by their parental alleles in the bin 6.05 QTL position could also contribute to point out the relevant candidate gene(s).

Acknowledgements

This work has been funded by the maize breeding companies (Advanta, Caussade Semences, Limagrain Europe, MaisAdour, Monsanto SAS, Pioneer Génétique, Pau Euralis, R2n RAGT Semences, SDME KWS France, Syngenta seeds) involved in the ProMais - INRA "ZeaWall" network on maize cell wall lignification and degradability. We thank Jacques Laborde for seed production at INRA St Martin de Hinx. We are grateful to Nathalie Rodde (CNRGV, Toulouse) for BAC DNA extractions. The sequencing and assembly of BAC clones using 454 Technology has been done in collaboration with the Sequencing platform (<http://genomique.genotoul.fr/>, Toulouse) and bioinformatics platform (<http://bioinfo.genotoul.fr/>, Toulouse) respectively.

CHAPITRE V - RECHERCHE DE GÈNES CANDIDATS
PUTATIFS IMPLIQUÉS DANS LA MISE EN PLACE ET LA
FORMATION DES PAROIS LIGNIFIÉES DU MAÏS ET
COLOCALISATIONS AVEC DES QTLs DE DÉGRADABILITÉ
DE PAROIS ISSUS D'UNE SYNTHÈSE DES DONNÉES SUR
SIX FAMILLES DE LIGNÉES RECOMBINANTES DE MAÏS
PRÉCOCES ET DEMI-PRÉCOCES

5.1. Avant propos

L'objectif des travaux présentés dans ce chapitre était d'établir une liste de gènes candidats pouvant expliquer des QTLs de dégradabilité de parois. Pour cela, d'une part, les gènes potentiellement impliqués dans la formation des parois secondaires du maïs ont été recherchés dans le génome de référence B73 (Schnable et al, 2009, www.maizesequence.org, release v2 5b.60) en se basant sur des connaissances disponibles sur le maïs, mais aussi sur d'autres espèces telles qu'*Arabidopsis*, le riz, et l'*Eucalyptus* qui sont des plantes significativement travaillées sur cette thématique. D'autre part, une synthèse des QTLs parois actuellement disponibles à partir de six familles RILs, a été réalisée, en s'appuyant sur une carte physique "consensus", construite à partir des positions physiques des marqueurs sur le génome de référence de B73. Cette synthèse inclut des QTLs de dégradabilité de parois et de façon corrélative des QTLs des caractères pariétaux liés incluant la teneur en lignines ainsi que, quand cela était disponible, des QTLs de structures des lignines (estimées par les teneurs en chacun des monomères H, G, et S libérés après une attaque au nitrobenzène alcalin) et de teneur en acides *p*-hydroxycinnamiques.

5.2. Résultats

Au cours de ces trois années de travail, une liste de 287 gènes potentiellement impliqués dans la formation des parois secondaires lignifiées a été établie chez la lignée de référence maïs B73. Cette liste inclut les gènes de la voie de biosynthèse et de polymérisation des composés phénoliques (lignines et acides hydroxycinnamiques) et les facteurs de transcription régulant ces voies de biosynthèse et/ou la mise en place des tissus lignifiés.

La carte consensus construite à partir des six familles RILs et basée sur le génome de B73 comprend 173 marqueurs répartis environ toutes les 10 Mpb. D'après les positions physiques estimées des 122 QTLs de dégradabilité des parois (IVNDFD et DINGAZ détectés pour des valeurs propres et topcross), 35 régions différentes du génome semblent impliquées dans la dégradabilité des parois. Ces régions, qui ne couvrent "que" près de 40 % des régions physiques du maïs, colocalisent avec des caractères pariétaux influençant la dégradabilité, à l'exception de trois régions.

Soixante dix neuf pour cent des gènes appartenant à la liste établie sont sous-jacent à ces QTLs, incluant les gènes impliqués dans la voie de biosynthèse des monolignols et dans leur polymérisation et des facteurs de transcription de type MYB, NAC.

5.3. Conclusions et perspectives

Les colocalisations fréquentes entre gènes candidats putatifs et QTLs de dégradabilité des parois observées semblent conforter la pertinence de la liste de gènes candidats établie. Ceci étant, il est clair que de nombreux gènes de fonction inconnue se trouvent sous les QTLs, de même également que des gènes reliés à des groupes fonctionnels pour lesquels aucun rôle dans la mise en place des parois lignifiés n'a été mis en évidence à ce jour. Enfin, si des candidats probables ou très probables colocalisent parfois, il n'est pas observé de cas où l'on puisse conclure *a priori* de façon quasi-certaine. Le rôle probable de gènes MYB et NAC, isolément, associés entre eux, ou associés à des gènes d'autres familles, dans les variations de dégradabilité des parois est toutefois nettement conforté. De même, l'existence de plusieurs gènes comme facteurs explicatifs des QTLs apparaît à ce niveau aussi probable que celle d'un gène majeur unique. Ceci complexifiera très probablement les démarches de validation des candidats, quelque soit la ou les démarche(s) envisagée(s).

La détermination de régions consensus impliquées dans la dégradabilité des parois reste toutefois à valider par une méta-analyse de QTLs avec une démarche statistique grâce au logiciel biomercator (comme cela a été fait par Truntzler et al, 2011), dont une toute nouvelle version prend en compte les informations issues de la séquence physique de B73.

Article 5 - Search for cell wall degradability consensus QTLs based on a synthesis of data from six RIL progenies, and search for underlying candidate genes (in progress)

Audrey Courtial^{1,2,3}, Marçal Soler^{2,3}, Valérie Méchin⁴, Matthieu Reymond⁴, Hua Wang^{2,3}, Jacqueline Grima-Pettenati^{2,3}, Yves Barrière^{1*}

¹ INRA, Unité de Génétique et d'Amélioration des Plantes Fourragères, F-86600 Lusignan, France

² Université de Toulouse, UPS, UMR 5546, Laboratoire de Recherche en Sciences Végétales, BP 42617 Auzeville, F-31326 Castanet-Tolosan, France

³ CNRS, UMR 5546, BP 42617, F-31326 Castanet-Tolosan, France

⁴ INRA, Institut Jean-Pierre Bourgin, F-78026 Versailles, France

* corresponding author

Key words - maize, cell wall, QTL, degradability, lignin, candidate gene.

Abstract

QTL for cell wall related traits, including QTLs for cell wall degradability, have been mapped in six RIL progenies investigated under homogeneous conditions. A consensus QTL approach was thus considered, taking advantage of the release of the first maize line genomic sequence, allowing to draw a physical map for all progenies, and to estimate the physical positions of each QTL. This was done assuming a similar genetic organization in the B73 sequenced line and in all lines used in the QTL analyses considered. The consensus physical map was drawn, based on the 731 markers available from the six RIL progeny, and considering only one marker every 10 Mbp and a total of 173 markers. Based on the six RIL progenies tested for their per se and/or topcross values, and considering the two cell wall IVNDFD and DINAGZ traits, 122 cell wall degradability QTLs were shown, corresponding to 35 non-overlapping positions. However, the fact that QTLs overlapped centromere positions in bins 3.04, 4.05, 5.04, 6.01, 8.03, 9.02, and 10.03, has led to a certain degree of inaccuracy in the estimated corresponding physical positions, with a risk of estimating one QTL position instead of two (at least) in these areas. Candidate genes involved in secondary cell wall and lignin metabolisms were colocalized with obtained "consensus" QTLs. Based on a list of 287 candidate genes, 228 candidates were shown to colocalize with QTLs, including nearly all genes involved in monolignol biosynthesis, many MYB genes of the EgMYB1 and EgMYB2 type, many NAC gene orthologs of *Arabidopsis* master genes regulating secondary wall assembly, and many orthologs of genes which were shown to play a key role in different model or woody species. For every QTL position, while a major unique genetic factor could be the only underlying determinant, several clustered genes are likely to be the co-

determinants of observed QTLs, possibly also including genes whose involvement in the secondary wall assembly has not yet been determined, and/or genes of still unknown function.

Introduction

Cell wall degradability is the basic trait determining plant energy value both for cattle feeding and second generation biofuel production. Consequently, deciphering the genetic determinants of this trait is a strategic issue for grass breeding, both for animal feed and for the biofuel industry. Given the genetic and genomic knowledge available, the technical possibilities mostly related to separate inflorescences, large seed production per plant, and the economic importance of the species, most grass cell wall research has been done on maize, which was also a model species. Important genetic variations for cell wall degradability have been shown in different sets of lines and hybrids (Argillier et al, 1995; Barrière and Argillier, 1997; Barrière et al, 2004a), with values nearly ranging from 22 to 40 %, based on NDF (Neutral Detergent Fiber) enzymatic degradability (Barrière et al, 2009a). These variations were later shown to be related to cell wall phenolic composition and structure (Méchin et al, 2000; Riboulet et al, 2008b; Grabber et al, 2009). The lignified maize secondary cell wall is a composite material with phenolic components, cellulose microfibrils, an amorphous matrix consisting predominantly of hemicelluloses (mainly glucurono-arabinoxylans), and very few pectins. Phenolic components are first comprised of lignins, including guaiacyl (G) units derived from coniferyl alcohol, syringyl (S) units derived from sinapyl alcohol, together with a lower amount of *p*-hydroxyphenyl units (H) derived from *p*-coumaryl alcohol. Phenolic components also comprised *p*-coumaric acid (pCA) acylated to S units of lignins, and ferulic acid (FA) derivatives, involved in extensive cross-linkages occurring between feruloylated arabinoxylans and G units of lignins, and between arabinoxylan chains after ferulate dimerization in mature cell walls (Ralph et al, 1992; Ralph et al, 1995; Jacquet et al, 1995; Grabber et al, 2004; Ralph, 2010). Lignins impart hydrophobicity to vascular elements, allowing water and nutriment transportation. The association of lignins and other matrix components, including the occurrence of linkages with and between cell wall carbohydrates, significantly modifies tissue properties towards higher stiffness and contributes to biotic and abiotic stress tolerance. However, cell wall hardness also induces lower polysaccharide degradability in animal rumen or industrial fermenters, with negative effects on silage energy value and bioethanol or biogas production.

QTL for cell wall related traits have been mapped in many studies, and a first meta-analysis investigation was done by Truntzler et al, (2010), based on all available published experiments giving lignin or plant degradability QTLs. However, this investigation was not targeted to cell wall degradability QTLs, did not include the complete set of QTL investigations currently available, and

Table 1 - RIL progenies and experiments considered in the QTL synthesis (nbr = number, env = environment, including year and location).

RIL progeny	RIL nbr	Whole Plant env	Without ear env	Topcross env	Reference
(1) Io x F2	100	2	0	2	Méchin et al, 2001
(2) F288 x F271	135	7	2	6	Roussel et al, 2002; Courtial et al, 2012b,d
(3) F7025 x F4	284	5	0	5	INRA and Génoplante unpublished data
(4) F838 x F286	242	0	6	3	Barrière et al, 2008, 2010
(5) FI16 x F2	140	3	0	6	Riboulet et al, 2008a, and unpublished data
(6) RIo x WM13	163	3	4	5	Barrière et al, 2012, and unpublished data

did not take advantage of consensus approaches allowed by the release of the first maize line genomic sequence (Schnable et al, 2009). The objective of the current work was thus to gather data targeted to the search for cell wall degradability QTLs, and to take secondly into account correlative cell wall component QTLs. Estimates of physical position of every QTL was done assuming a similar genetic organization in the B73 sequenced line and in all lines used in considered QTL analyses. Candidate genes involved in cell wall related phenolic compound biosynthesis and in regulations of their biosynthesis and deposition were colocalized with QTLs, based on their physical positions in the B73 line.

Materials and methods

RIL progenies and QTL analyses

Cell wall degradability QTLs and colocalizing cell wall component QTLs were gathered from analyses in six RIL progenies after RIL *per se* value experiments, either with whole plant harvest and/or harvest of plants without ear (ear removed the day of harvest to avoid confusing effect of stem and grain ferulates) and whole plant topcross experiments (Table 1).

ADL (acid detergent lignin, acido-resistant part of the lignin polymer) and NDF (neutral detergent fiber, cell wall content) were estimated according to Goering and van Soest (1971). Cell wall degradability estimates were based on the IVNDFD (Struik, 1983; Dolstra and Medema, (1990) and DINAGZ (Argillier et al, 1995, Barrière et al, 2003b) traits, computed from an enzymatic solubility of dry-matter (IVDMD, Aufrère and Michalet-Doreau, 1983) and cell component contents (NDF, crude protein, soluble carbohydrates). Contents in *p*-hydroxycinnamic acids [*p*-coumaric acid (pCA), esterified and etherified ferulic acid (esterFA and etherFA), 5-5 and 8-O-4 diferulic acids (diFA)] were estimated according to Morrison et al, (1993) and Méchin et al, (2000). Lignin structure was considered based on aldehyde releases after alkaline nitrobenzene oxidation of cell wall residues (Roadhouse and MacDougall, 1956; Higuchi and Kawamura, 1966). *p*-Hydroxyphenyl (H), guaiacyl (G), and syringyl (S) lignin monomers were thus oxidized into *p*-hydroxybenzaldehyde (pHb), vanillin (Van) and syringaldehyde (Sga), respectively. Due to the large number of samples, all traits were estimated using near infrared reflectance spectroscopy (NIRS, NIRS system 6500 spectrophotometer) after validation of calibration equations by a laboratory analysis of nearly 20 samples per location.

In all investigations, QTL mapping was based on trait means over years and locations for all RIL investigations, using the Composite Interval Mapping (CIM, Zeng, 1994) method implemented in the PLABQTL computer package (Utz and Melchinger, 1996). QTL physical positions were estimated based on physical positions of the two flanking markers available in MaizeGDB database (B73 release 5b60), and assuming a constant relationship between recombination and physical distances within this interval. For each trait, QTLs from different progenies were considered as colocalizing when their

interval supports were overlapping, and average physical positions of "consensus" QTLs were then estimated. The physical lengths of consensus QTL support intervals were difficult to estimate in several instances due to their overlapping centromer positions and/or variable genetic and physical distances between progenies. Consequently, the latter were not reported. Their gross values could be considered to be close to 20 Mbp, but could reach nearly 50 Mbp, and even more, in centromere locations.

Search for candidate genes

The search for candidate genes underlying cell wall related traits focused on genes involved in secondary wall biosynthesis and assembly. The candidate gene list was based on the previously proposed list (Barrière et al, 2009a) to which were added maize orthologs of newly described transcription or regulation factors in *Arabidopsis* or woody plants (poplar, *Eucalyptus*, ...). An e-value at least lower than e^{-50} was retained during ortholog searches. Three groups of genes were considered for colocalizations with QTLs, including i) monolignol and *p*-hydroxycinnamic acid biosynthesis and polymerization related genes, ii) transcription factor involved in regulation of these pathways (MYB, NAC and other transcription factors), and iii) genes involved in lignified tissue patterning and assembly. Genes involved in cell wall carbohydrate biosynthesis, and their regulation factors, were not added to the list because it was considered that they were likely not involved in mechanisms explaining cell wall degradability variation. However, genes involved in arabinoxylan biosynthesis and feruloylation were searched for, due to the major importance of feruloylated arabinoxylans in cross-linkages and degradability of secondary cell walls. Gene physical positions were based on the maize B73 sequence (www.maizesequence.org), assuming a similar genomic structure in RIL parental lines and B73 within QTL support intervals. Genes were considered as possibly colocalizing with QTLs when they were mapped within a nearly ± 10 Mbp interval around QTLs positions.

Results and discussions

Consensus map and QTL positions

A consensus physical map was drawn, based on the 731 markers available from the six RIL progeny, considering only one marker every 10 Mbp and then 173 markers (Figure 1). Corroborating the very few recombinations occurring in regions surrounding maize centromeres (Liu et al, 2009), only few markers were mapped in these areas. According to RILs and chromosomes, physical distances between markers ranged around centromeres between 24 and 56 Mbp, with an average value equal to 39 Mbp. Conversely, marker density was higher in both distal parts of chromosomes. The average ratio between physical and genetic distances was thus equal to 6.4 Mbp/cM in centromere positions, while it was only 0.8 Mbp/cM elsewhere.

Figure 1 - Cell wall degradability QTL in 6 RIL progenies (Consensus map and physical distances as Mbp).

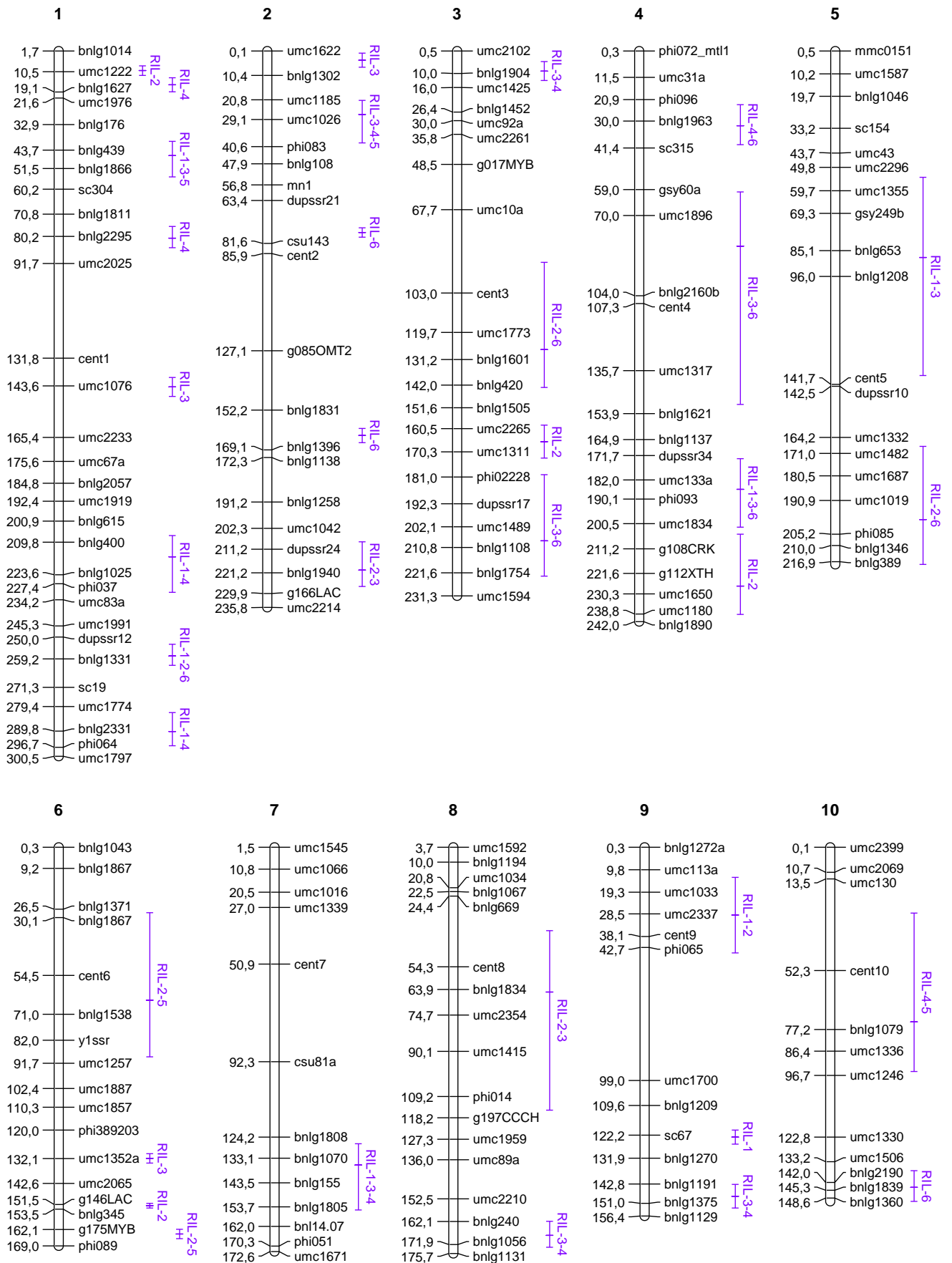


Table 2 - Cell wall degradability (IVNDFD and DINAGZ) QTLs and colocalizing cell wall component QTLs in six RIL progenies [Io x F2 (1), F288 x F271 (2), F7025 x F4 (3), F838 x F286 (4), F116 x F2 (5), and RIo x WM13 (6); nbr = QTL number, pos = Mbp position, (RIL) = RIL progeny number, from-to = extreme positions of QTLs in the considered position].

QTL		Cell wall degradability				Cell wall components [pos (RIL)]						
bin	nbr	pos (RIL)	from-to	ADL/NDF	pCA	esterFA	etherFA	5-5diFA	8-O-4diFA	pHb	Van	Sga
1.01	1	10 (2)	-	9 (4)	-	-	-	-	-	3 (6)	-	-
1.02	2	16 (4)	15-17	15 (4)	17 (4)	-	17 (4)	-	-	17 (4)	17 (4)	17 (4)
1.03	6	46 (1, 3, 5)	41-54	42 (1, 3)	-	-	-	-	-	-	-	-
1.05	3	81 (4)	79-81	79 (4)	79 (4)	-	69 (4)	-	88 (4)	-	76 (4)	76 (4)
1.05	1	144 (3)	-	-	-	-	-	-	-	-	-	-
1.07	4	216 (1, 4)	211-227	208 (4, 6)	211 (4)	-	208 (4)	217 (4)	-	-	-	209 (4)
1.08	3	258 (1, 2, 6)	255-260	-	-	-	-	-	-	-	-	-
1.10	3	290 (1, 4)	284-294	288 (4, 6)	287 (6)	-	-	-	-	284 (6)	-	-
2.01	1	4 (3)	-	9 (3, 4)	-	-	-	-	-	-	-	-
2.02	5	27 (3, 4, 5)	23-37	26 (1, 3, 6)	-	-	-	-	-	-	26 (2, 4)	-
2.04	2	77 (6)	76-77	113 (6)	-	77 (4)	-	-	-	-	-	-
2.06	1	163 (6)	-	-	-	-	154 (4)	163 (6)	163 (6)	174 (6)	185 (6)	185 (6)
2.08	4	220 (2, 3)	211-224	218 (2, 3, 4)	224 (2)	223 (4)	-	-	-	-	218 (4)	218 (4)
3.03	2	9 (3, 4)	6-11	8 (3)	-	-	-	-	17 (4)	-	10 (4)	-
3.04	5	127 (2, 6)	102-141	-	-	133 (4)	80 (6)	136 (6)	136 (6)	-	-	-
3.06	4	166 (2)	161-171	158 (2)	165 (4)	-	163 (2)	-	-	-	-	162 (2)
3.07	4	208 (3, 6)	182-221	206 (6)	-	-	218 (2)	-	-	198 (2)	-	-
4.04	2	32 (4, 6)	25-38	38 (1, 3, 6)	37 (6)	-	38 (4)	-	-	55 (6)	35 (4)	35 (4)
4.05	7	83 (3, 6)	63-147	-	-	-	-	-	-	-	-	-
4.08	8	186 (1, 3, 6)	176-199	193 (3, 6)	-	189 (4)	-	-	-	-	-	-
4.09	6	227 (2)	208-236	219 (2)	-	-	-	-	214 (6)	231 (2)	-	-

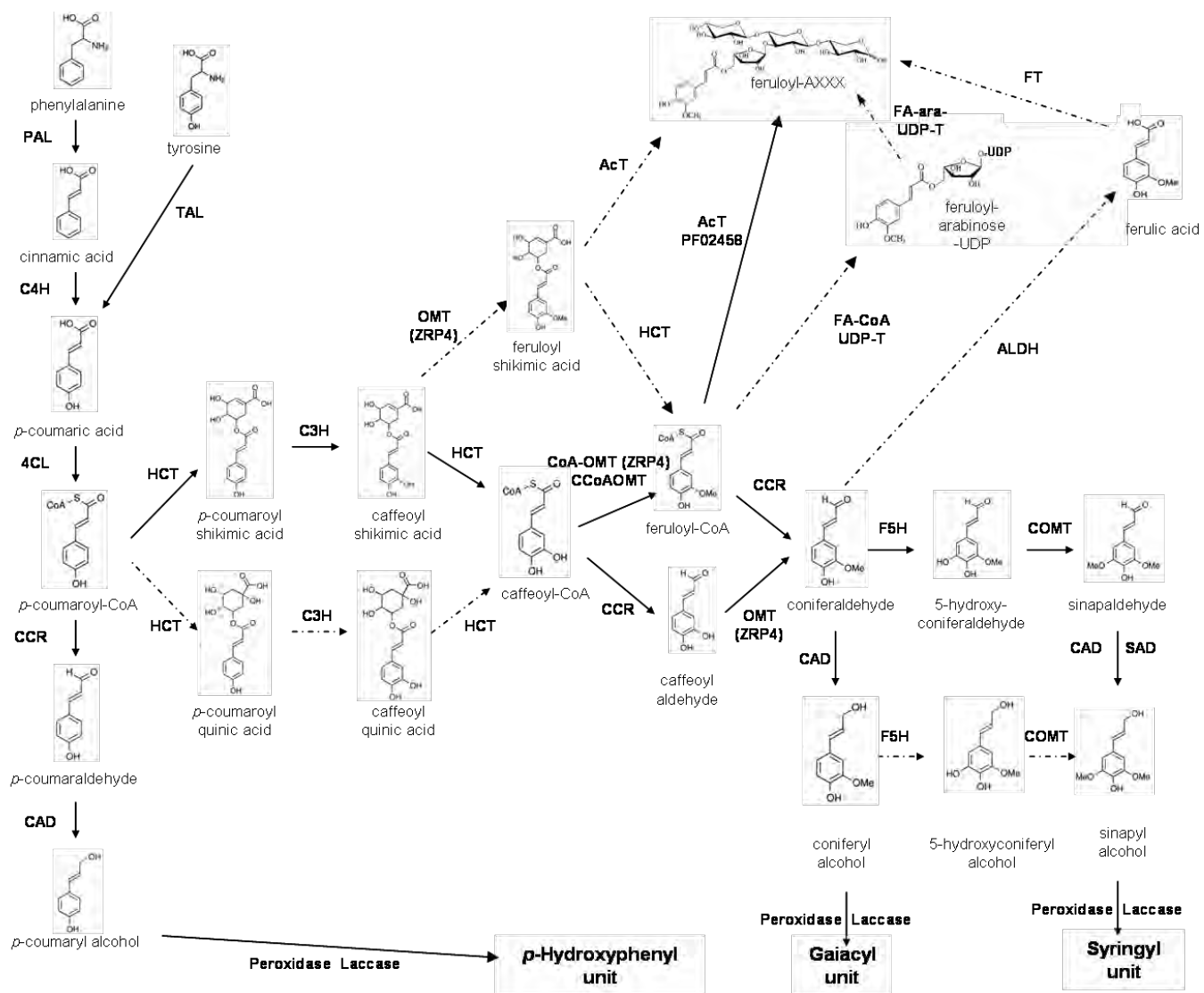
Table 2 - continued

bin	nbr	pos (RIL)	from-to	ADL/NDF	pCA	esterFA	etherFA	diFA	diFA	pHb	Van	Sga
5.04	3	88 (1, 3)	57-135	72 (5)	-	-	-	-	-	-	-	-
5.05	4	199 (2, 6)	171-215	196 (2, 4, 6)	-	-	-	194 (4)	200 (4)	-	-	160 (6)
6.01	3	65 (2, 5)	30-87	64 (2, 5)	-	-	-	-	-	-	-	-
6.04	1	132 (3)	-	132 (3)	-	-	-	-	-	-	-	-
6.05	6	152 (2)	151-152	152 (2)	-	153 (2)	154 (2)	153 (2)	147 (2)	-	-	152 (2)
6.06	5	164 (2, 5)	162-167	162 (2, 3)	162 (2)	-	-	-	-	-	-	-
7.03	6	136 (1, 3, 4)	130-152	140 (2, 4)	-	-	-	-	-	-	-	131 (4)
8.03	3	65 (2, 3)	42-112	74 (4, 6)	-	-	78 (4)	46 (2)	59 (6)	-	66 (6)	-
8.07	4	168 (3, 4)	165-170	168 (3, 4)	167 (4)	-	168 (4)	-	-	-	-	169 (4)
9.02	2	29 (1, 2)	16-42	-	31 (2)	-	-	-	-	-	-	-
9.04	1	123 (1)	-	-	-	122 (4)	-	-	-	-	-	-
9.06	3	148 (3, 4)	146-150	147 (3, 4)	-	-	-	-	-	-	142 (6)	-
10.03	4	74 (4, 5)	31-92	85 (4, 5)	31 (4)	96 (4)	-	85 (4)	85 (4)	31 (4)	82 (4)	47 (4)
10.07	3	144 (6)	140-147	143 (6)	-	144 (4)	-	-	-	-	-	-

Based on the six RIL progenies experimented for their per se and/or topcross values, and considering the two cell wall IVNDFD and DINAGZ traits, 122 cell wall degradability QTLs were shown, likely corresponding to 35 non-overlapping positions (Table 2). QTLs were shown in only one RIL progeny for 11 positions, while QTLs were simultaneously shown for three RIL progenies in five positions, in bins 1.03, 1.08, 2.02, 4.08, and 7.03. QTLs were shown only for the IVNDFD trait in five positions, and only for the DINAGZ trait in one position. The fact that QTLs overlapped centromere positions in bins 3.04, 4.05, 5.04, 6.01, 8.03, 9.02, and 10.03, resulted in less accurate estimates of the corresponding physical positions. The risk of estimating one QTL position instead of a minimum of two in these areas should therefore be considered. If the gene number per Mb is substantially reduced in the centromere region, these regions indeed have a 4-fold higher density of genes per cM (Eichten et al, 2011).

Colocalizations of cell wall degradability and ADL/NDF QTLs occurred in 28 out of the 35 non-overlapping positions shown for the ADL/NDF QTLs. However, in each position, lignin and cell wall degradability QTLs were not always mapped in all progenies (Table 2). At the same time, no colocalizations between cell wall degradability and lignin QTLs were shown in seven positions, including three positions where cell wall degradability QTLs did not colocalized with any trait. One of these latter positions corresponded to single QTLs shown in only one progeny (bins 1.05). However, the two other positions corresponded to clusters of cell wall degradability QTLs mapped in two (bin 4.05) and three (bin 1.08) progenies. More scattered QTL colocalizations were observed with the other considered cell wall component traits, which were however investigated in only three progenies. Colocalizations with pCA occurred in positions where colocalizations with ADL/NDF QTLs also occurred, except in bin 9.02 where pCA was the only shown trait that could explain variation in cell wall degradability. Colocalizations between cell wall degradability QTLs and etherFA and/or diFA QTLs were shown in 15 positions, likely corroborating the significant role of cross-linkages in cell wall degradability variations. Lignin structure QTLs colocalized with lignin content QTLs, and nearly half of each aldehyde QTL did not colocalize with any other aldehyde QTL. Multi-trait QTL colocalizations were shown in seven positions (bins 1.05, 1.07, 2.06, 4.04, 6.05, 8.03, and 10.03), highlighting either a cluster of cell wall related genes or the presence of an upstream regulation factor acting on all the cell wall components. Conversely, no colocalization was shown in three positions, including in bin 4.05 for a consensus QTL gathering seven QTLs from two progenies. This fact might indicate that other traits should be considered as involved in cell wall degradability, and/or that some analytical procedures used for trait estimates could be partly inaccurate. The latter hypothesis could be especially considered for etherFA and diFA estimated contents. Released etherified ferulate may thus account for as little as 15% of total cross-linkages because several ether and biphenyl ether coupling of ferulate to lignins have not yet been determined due to the limitations of the solvolytic methods (Grabber et al. 2000). Similarly, the positive relationship observed between diFA release and cell wall degradability was considered as explaining not the intensity of the corresponding cross-linkages which

Figure 2 - Monolignol biosynthesis pathway and hypothetical ways towards arabinoxylan feruloylation in maize (more probable or known pathways are shown with solid lines, while hypothetical ones are shown with dotted lines).



have negative effects on cell wall degradability, but the accessibility and susceptibility of diFA bridges in the cell wall towards alkaline hydrolysis (Riboulet et al. 2008a; Barrière et al, 2009). In addition, the R^2 values of NIRS calibrations are only equal to 0.65 for ferulate related traits, while they are higher than 0.80 for the other investigated traits (Riboulet et al. 2008a).

Candidate genes colocalizing with cell wall degradability QTLs

When considering genes involved in cell wall phenolic compound metabolism and assembly, and thus excluding genes involved in cell wall carbohydrate biosynthesis, the considered candidate gene list comprised 287 genes (Supplementary table S1). Out of these genes, 98 were related to monolignol and hydroxycinnamic acid biosynthesis and polymerization, 107 were transcription factors putatively involved in the regulation of cell wall phenolic genes, and 82 were related to the regulation of lignified tissue assembly. The search for candidate genes was done in the cumulative area of all consensus QTL estimated support intervals (Figure 1), corresponding to nearly 40 % of the total genome physical length. Consequently, the candidates listed for each QTL position are oversized, while the true candidate could nevertheless be missing if it was either a gene of unknown function, or a gene of which function was not yet related to cell wall assembly.

Genes involved in monolignol biosynthesis

Monolignol biosynthesis (Figure 2) occurred downstream the shikimate pathway which links the carbohydrate metabolism to the biosynthesis of aromatic amino acids (phenylalanine, tyrosine, and tryptophan) and, consequently, to the phenylpropanoid pathway. The first step of monolignol biosynthesis is the deamination of L-phenylalanine into cinnamic acid by a phenylalanine ammonia lyase (PAL) enzyme. Maize PAL enzymes also have a tyrosine ammonia lyase (TAL) activity (Roesler et al, 1997; Higuchi et al, 1967), catalyzing tyrosine deamination into pCA. Successive steps including hydroxylation and methylation on the aromatic ring further lead to the production of the three monolignols which are then exported to the cell wall and polymerized into lignins. Most genes involved in monolignol biosynthesis belong to small multigene families, with different members possibly involved in different metabolons corresponding to each monolignol biosynthesis and/or to their biosynthesis in each type of lignified tissue. Gene families involved in all steps of monolignol biosynthesis had a member, and often all members, colocalizing with cell wall degradability QTLs. However, these genes were most often not located very close to QTL positions (Table 3).

Seven PAL genes out of the ten considered in the maize genome, which are clustered in four genomic positions (Barrière et al, 2009a), colocalized with QTL positions in bins 2.03, 5.06, and 4.09. The hydroxylation of cinnamic acid is catalyzed by cinnamate 4-hydroxylase (C4H) and one of the two C4H genes, located in bin 8.08, colocalized with cell wall degradability QTLs. The resulting *p*-coumaric acid is then converted into coumaroyl-CoA by two 4-CoA ligases (4CL), both colocalizing with QTLs in bin 1.07 and 5.04. The conversion of coumaroyl-CoA into caffeoyl-CoA through the

Table 3 - Colocalizations of IVNDFD and DINAGZ QTLs with genes putatively involved in biosynthesis and polymerization of cell wall phenolic components and their regulation, in six RIL progenies [Io x F2 (1), F288 x F271 (2), F4 x F7025 (3), F838 x F286 (4), F116 x F2 (5), and RIo x WM13 (6)].

Table 3a - Chromosome 1.

bin	RIL	QTL pos Mbp	ZmGene annotation	GRMZM number (v2 5b60)	Gene pos Mbp
1.01	2	10	ATHB8 HD-ZIPIII RLD2/IFL1	GRMZM2G042250	2.80
			KNAT7-like ELK-KNOX1	GRMZM2G159431	5.01
			ZmNAC041 VND1-like ZmSWN5	GRMZM2G025642	5.64
			Zinc finger C3HC4-CHY-like	GRMZM2G062724	10.04
1.02	4	16	Zinc finger C2H2-type (id1)	GRMZM2G171073	23.57
1.03	1, 3	46	ZmLac AtLac12-like	GRMZM2G305526	39.93
			Zinc finger C2H2-type	GRMZM2G048154	42.12
			ZmLac AtLac17-like	GRMZM5G814718	46.41
			UDP-xylose-4-epimerase (MUR4-like)	GRMZM2G000632	47.47
			Ovate family protein OFP1/4-like	GRMZM2G075988	56.31
			ZmPox2	GRMZM2G040638	63.65
1.05	4	81	ZmMYB019 EgMYB2-like (G13)	GRMZM2G147698	64.24
1.05	3	144	-	-	-
1.07	1, 4	216	ZmMYB010 EgMYB1-like (G4)	GRMZM2G084583	206.85
			ZmMYB003 AtMYB85-like (G8)	GRMZM2G037650	207.18
			Zm4CL2	GRMZM2G048522	210.10
			ZmCCR1	GRMZM2G131205	211.52
			ZmMYB012 AtMYB85-like (G8)	GRMZM2G106558	215.34
			ZmF5H1	AC210173.4_FG005	224.04
1.08	1, 4, 6	258	ZmLac AtLac17-like	GRMZM2G164467	260.07
1.10	1, 4	290	ABC transporter	GRMZM2G135199	299.98
			ABC transporter ATPDR6-like	GRMZM2G391815	300.53

formation of shikimate (or quinate) esters involves a hydroxycinnamoyl-CoA shikimate/quinate hydroxycinnamoyl transferase (HCT) encoded by two genes with one colocalizing in bin 2.03, and a *p*-coumaroyl-shikimate/quinate 3'-hydroxylase (C3'H) also encoded by two genes with one colocalizing in bin 6.06. Caffeoyl-CoA and feruloyl-CoA are the two key-compounds towards the synthesis of coniferyl and sinapyl alcohol (and also ferulate derivatives). The conversion of caffeoyl-CoA into feruloyl-CoA is catalyzed by caffeoyl-CoA O-methyltransferases (CCoAOMT) with five genes in the maize genome, and three colocalizing with QTLs in bins 4.08 and 9.02.

Conversions of activated *p*-coumaroyl-, caffeoyl-, and feruloyl-CoA compounds are mainly driven by the ZmCCR1 cinnamoyl-CoA reductase, even if a ZmCCR2 gene is present in the maize genome. Similarly, the reduction of *p*-hydroxy-cinnamaldehydes into alcohols is also mainly catalyzed by the ZmCAD2 cinnamyl alcohol dehydrogenase, while the ZmCAD1 gene and activity have been shown in maize and that other CAD-like genes might encode enzymes with CAD activity (Guillaumie et al, 2007a; Guillaumie et al, 2007b; Barrière et al, 2009a). The major roles of both ZmCCR1 and ZmCAD2 in the two last steps of monolignol biosynthesis are highlighted by the effects of corresponding mutants. ZmCAD2 mutations have been associated with the maize brown-midrib *bm1* phenotype, inducing higher cell wall degradability, lower lignin content, an incorporation of aldehydes in the lignin polymer, and no change in the syringyl/guaiacyl (S/G) ratio (Halpin et al, 1998; Barrière et al, 2004c). Even if no brown-midrib phenotype was shown, a transposon-tagging mutation in the ZmCCR1 gene also induced reduced lignin content and higher cell wall degradability. In addition, H units were released in lower amounts from the ZmCCR1 mutant plants compared with the normal ones, with simultaneously an increase in the S/G ratio in mutants (Tamasloukht et al, 2011). Both ZmCCR1 and ZmCAD2 genes colocalized with cell wall degradability QTLs, in bins 1.07 and 5.04, respectively.

Ferulate 5-hydroxylase (F5H) catalyzes the 5-hydroxylation of coniferaldehyde (and probably, but to a lesser extent, coniferyl alcohol) into 5-hydroxyconiferaldehyde (5-hydroxyconiferyl alcohol, respectively). Two F5H genes are present in maize genome, F5H1 with a strong expression in maize stalks and F5H2 mostly expressed in roots (Guillaumie et al, 2007a; Riboulet et al, 2009). The F5H1 gene colocalized with QTLs in bin 1.07. 5-Hydroxyconiferaldehyde (and 5-hydroxyconiferyl alcohol) is methylated into sinapaldehyde by the caffeic acid O-methyltransferase (COMT), which is the only gene of the maize monolignol pathway that does not belong to a small multigene family. The maize COMT gene exactly colocalized with QTLs in bin 4.04, and the importance of this gene in lignin biosynthesis was previously proved from mutant and transformed plant investigations, and association studies. The brown-midrib *bm3* mutation induced a reduced COMT activity (Grand et al, 1985; Guillaumie et al, 2007a; Guillaumie et al, 2008) and occurred in the COMT gene (Vignols et al, 1995). *Bm3* plants had greatly improved cell wall degradability, lower lignin content, and reduced S/G ratio with a reduction to 40 % of S units released after thioacidolysis and a significant incorporation of 5-

Table 3b - Chromosome 2.

bin	RIL	QTL pos Mbp	ZmGene annotation	GRMZM number (v2 5b60)	Gene pos Mbp
2.01	3	4	UDP-xylose-4-epimerase (MUR4-like)	GRMZM2G040397	9.50
			ZmCAD EgCAD2like	GRMZM2G046070	10.53
			ZmMYB026 AtMYB58-like (G3)	GRMZM2G038722	13.30
2.02	3, 4, 5	27	Zinc finger WRKY AtWRKY12-like	GRMZM2G123387	21.05
			ZmNAC169 VND6-like	GRMZM2G178998	22.70
			Zinc finger LIM-type	GRMZM2G153268	23.13
			ABC transporter	GRMZM2G177812	23.88
			ZmNAC168 VND7-like ZmSWN4	AC212859.3_FG008	25.75
			ZmNAC AtNAC082-VNI1-like	GRMZM2G176677	26.43
			ZmPAL2b	GRMZM2G118345	28.05
			ZmPAL2a	GRMZM2G441347	28.12
			ZmMYB027 AtMYB58-like (G2)	GRMZM2G048295	29.54
			ZmHCT2	GRMZM2G158083	31.83
			WD40-like	GRMZM2G022627	36.94
2.04	6	77	ZmNAC091 SND1-like	GRMZM2G099144	47.05
			ZmNAC096 XND1-like	GRMZM2G316840	48.88
			Xylan α -arabinosyl transferase GT47	GRMZM2G023020	62.92
			ZmOMT ZRP4-like	GRMZM2G147491	118.54
			ZmOMT ZRP4-like	GRMZM2G036048	126.76
			ZmOMT ZRP4-like SBP1	GRMZM2G085924	127.09
			ZmMYB035 EgMYB1-like (G4)	GRMZM2G124715	131.35
2.06	6	163	ZmMYB024 (G1)	AC165178.2_FG004	181.42
			Zinc finger CCCH-type	GRMZM2G168163	186.90
			Basic helix-loop-helix bHLH	GRMZM2G009478	188.53
			ZmCCoAOMT3	GRMZM2G004138	189.28
2.08	2, 3	220	ZmMYB028 (MYB31) EgMYB1-like (G4)	GRMZM2G050305	196.29
			GRAS SCARECROW-like	GRMZM2G431309	207.84
			ZmNAC080 SND2/SND3-like	GRMZM2G137546	220.00
			Zinc finger C2H2-type	GRMZM5G887286	222.16
			ZmLac (putative laccase)	GRMZM2G166857	229.90
			Zinc finger CCCH-type	GRMZM2G025014	232.64

hydroxy-coniferaldehyde, results which were corroborated in COMT down-regulated plants (He et al, 2003; Pichon et al, 2006).

Biosyntheses of coniferyl and sinapyl alcohols are likely based on two different preferential routes starting from caffeoyl-CoA. Coniferyl alcohol could mostly originate from a synthesis of coniferaldehyde with a methoxylation on the activated compound and production of feruloyl-CoA. Syringyl alcohol could be mostly produced with first a reduction of caffeoyl-CoA into caffeoyl aldehyde, and then a methoxylation on this last compound. The methoxylation of caffeoyl aldehyde into coniferaldehyde has been considered to be catalyzed by the caffeic acid O-methyltransferase (COMT) in several studies on dicotyledonous plants (Lee et al, 1997; Guo et al, 2001a; Parvathi et al, 2001; Chen et al, 2006; Do et al, 2007; Lee et al, 2011; Zhao and Dixon, 2011; Gray et al, 2012). The route duality and the involvement of COMT in another methoxylation step have not been established in grasses. However, since the disruption of the ZmCOMT gene in bm3 mutants did not completely prevent the synthesis of syringyl alcohol, an alternative methoxylation pathway should exist in maize. On the contrary, in the *Arabidopsis* AtOMT1 mutant, the lignin content in S units is reduced to a value close to zero (Goujon et al, 2003). Because CCoAOMT enzymes have a strict affinity for CoA-esters (Martz et al, 1998; Meng and Campbell, 1998; Parvathi et al, 2001), they cannot be considered as candidates involved in 5-hydroxy-coniferaldehyde methoxylation. Conversely, several ZRP4-like OMT are expressed in maize lignifying tissue (Guillaumie et al, 2007a; Guillaumie et al, 2007b) and their role is very likely not limited to methylation of suberin sub-unit precursors in plant roots as initially described (Held et al, 1993). At least one ZRP4-like OMT could thus contribute to methoxylate the C5 position of the phenolic ring during monolignol biosynthesis in maize. Five ZRP4-like OMT colocalized with QTLs, in bins 4.08 and 6.06.

Genes involved in monolignol transport and polymerization

After their biosynthesis, the three monolignols are transported from the cytosol to the apoplast likely as monolignol glucosides, which are also probably their storage form (Lim et al, 2005; Escamilla-Trevino et al, 2006; Liu, 2012). Maize orthologs of *Arabidopsis* uridine-diphosphate-glucosyltransferases (UGT) are involved in glucosylation of coniferyl and sinapyl alcohols (Lim et al, 2005; Lanot et al, 2006), and *Arabidopsis* β -glucosidases (β -Glu45 and β -Glu46) encode proteins with narrow specificity towards the three monolignol glucosides (Escamilla-Trevino et al, 2006). However, no maize ortholog colocalized with any of the cell wall degradability QTL. ABC (ATP-binding cassette) transporters are probably involved in the transport of monolignols across membranes (Sanchez-Fernandez et al, 2001; Samuels et al, 2002; Ehling et al, 2005). The *Arabidopsis* AtABCG29 ABC transporter has indeed been shown to act as a *p*-coumaryl alcohol transporter (Alejandro et al, 2012). Out of the 10 considered ABC transporters either expressed in maize stem (Guillaumie et al, 2007b) or as *Eucalyptus* (Rengel et al, 2009) and *Arabidopsis* orthologs, six colocalized with cell wall degradability QTLs.

Table 3c - Chromosome 3.

bin	RIL	QTL pos Mbp	ZmGene annotation	GRMZM number (v2 5b60)	Gene pos Mbp
3.03	3, 4	9	ZmNAC SND2/SND3-like	GRMZM2G166721	6.94
			Ara-CoA-acylT (PF02458)	GRMZM2G314898	8.30
			Ara-CoA-acylT (PF02458)	GRMZM2G305900	9.85
			ABC transporter ATPDR7-like	GRMZM2G118243	10.27
3.04	2, 6	127	ZmMYB044 EgMYB2-like (G13)	GRMZM2G052377	13.20
			ATHB-8 HD-ZIPIII	GRMZM2G178102	123.27
			Zinc finger LIM-type	GRMZM2G175761	134.61
			MADSbox SHP1-like	GRMZM2G160687	137.23
3.06	2	166	ZmMYB041 EgMYB1-like (G4)	GRMZM2G041415	140.12
			Xylan α -arabinosyl transferase GT47	GRMZM5G898668	157.02
			Xylan α -arabinosyl transferase GT47	GRMZM2G056702	157.05
			Xylan α -arabinosyl transferase GT47	GRMZM2G448834	157.08
3.07	3, 6	208	Zinc finger C2H2-type	GRMZM2G019266	172.63
			ZmMYB052 EgMYB1-like (G4)	GRMZM2G160838	173.52
			COV-LCV2-like	GRMZM2G149662	175.56
			Zinc finger DOF-type	GRMZM2G135703	175.97
			ZmLac laccase 1	GRMZM5G842071	179.12
			ZmLac AtLac17-like	GRMZM2G072780	180.68
			ZmLac AtLac12-like	GRMZM2G132169	183.65
			Ovate family protein OFP1/4-like	GRMZM2G330159	184.31
			ZmC3H1	GRMZM2G138074	187.71
			Ovate family protein OFP4-like	GRMZM2G127431	198.04
			ABC transporter AtABCG29-like	GRMZM2G000614	201.54
			ZmMYB055 (G14)	GRMZM2G167829	201.93
Zinc finger C3HC4-CHY-like	GRMZM2G077307	201.99			
ZmMYB047 EgMYB2-like (G29)	GRMZM2G088783	204.15			
FRA2-like AAA ATPase	GRMZM2G054715	208.05			
ZmNAC101 SND2/SND3-like	GRMZM2G058518	210.23			
Zinc finger CCCH-AtC3H14-like	GRMZM2G157927	214.75			
ZmALDH RF2D	GRMZM2G071021	221.69			
ZmALDH RF2C	GRMZM2G097699	221.71			
Ara-CoA-acylT (PF02458)	GRMZM2G094428	225.34			
ZmMYB045 EgMYB2-like (G13)	GRMZM2G064744	227.39			

While class III peroxidases have long been considered as the unique class of oxidases involved in lignin polymerization, EST sequencing and expression studies based on lignifying tissues, and mutant investigations, have shown that both laccases and peroxidases are involved in cell wall lignification (Boudet, 2000; Boerjan et al, 2003; McCaig et al, 2005; Cai et al, 2006; Sato and Whetten, 2006; Tokunaga et al, 2009; Fagerstedt et al, 2010; Berthet et al, 2011). Class III peroxidases and laccases belong to multigene families, and consequently, redundancy in their activity has often been suspected. However, the importance of oxidase redundancy is greatly reduced by the fact that many peroxidases or laccases have specific spatio-temporal expression patterns. When considering genes expressed in maize vascular and lignifying tissues, and orthologs of *Arabidopsis* genes expressed in lignifying stems, (de Obeso et al, 2003; Bakalovic et al, 2006; Caparrós-Ruiz et al, 2006; Guillaumie et al, 2007b; Andersen et al, 2009; Barrière et al, 2009a; Riboulet et al, 2009), seven peroxidase and fifteen laccase genes were currently considered in the maize genome. Whether all these latter genes are effectively involved in constitutive lignification, and whether some other members are still unidentified, especially for peroxidases, is not known. However, several investigations suggested that only a few members would be involved in secondary wall assembly. This fact is more likely a consequence of a regulated spatio-temporal expression of peroxidase and laccase genes, rather than a specificity of several family members towards monolignols.

Down-regulation of the tobacco peroxidase TP60 led to plants with lignin reductions of 20 and up to 40-50 % of control plants. In the line with the most robust changes in lignin content through several generations, plants have thin cell walls and limited secondary wall thickening with an abnormal S2 layer (Blee et al, 2003; Kavousi et al, 2010). Similar results were shown in aspen after deregulation of the PrxA3a peroxidase (Li et al, 2003b). In addition, a MITE insertion disrupting the ZmPox3 gene was associated with higher cell wall degradability in a set of related maize European flint lines (Guillet-Claude et al, 2004), whereas ZmPox2 was considered as involved at a greater extent than ZmPox3 in maize vascular vessels and epidermis lignification, (de Obeso et al, 2003). While the lignin content was not changed, the down-regulation of the poplar Lac3 laccase gene induced an important alteration of xylem fiber cell walls, with an increase in soluble phenolic compounds (Ranocha et al, 2002). In *Arabidopsis*, laccase AtLac4 and AtLac17 double mutants had lignin content that were 35 % lower than in control plants, with higher saccharification yields, while the reduction was nearly 13 % in single mutants (Berthet et al, 2011). Over-expression of a cotton laccase in poplar plants induced an increase in lignin content in all tested transgenic lines in varying degrees, but as highest as 21.5% (Wang et al, 2008). Observed effects on lignin content in (double) peroxidase or laccase mutants or transformants were thus of the same order of magnitude than the one observed with monolignol genes, such as in the maize bm3/COMT mutant, and even higher than in the maize bm1/CAD mutant.

Colocalizations between cell wall degradability QTLs and peroxidase genes only occurred one time, in bin 6.04, with the ZmPox3 gene. Conversely, colocalizations between cell wall degradability QTLs and laccase genes occurred nine times, including seven colocalizations with orthologs of

Table 3d - Chromosome 4.

bin	RIL	QTL pos Mbp	ZmGene annotation	GRMZM number (v2 5b60)	Gene pos Mbp
4.04	4, 6	32	ZmCOMT	AC196475.3_FG004	32.25
			Basic helix-loop-helix bHLH	GRMZM2G017586	35.58
			ZmNAC143 SND1/NST1-like ZmSWN2	GRMZM2G069047	38.03
4.05	3, 6	83	ZmNAC VND7-like	GRMZM2G048826	59.34
			ZmMYB069 AtMYB58-like (G2)	GRMZM2G127857	142.64
			ZmPAL3d	GRMZM2G160541	143.38
			ZmPAL3c	GRMZM2G081582	143.47
			ZmPAL	GRMZM2G063917	143.51
			ZmNAC VND6-like	GRMZM2G354151	148.82
			Zinc finger WRKY AtWRKY12-like	GRMZM2G377217	151.07
			MADSbox SHP1-like ZmZAG5	GRMZM2G003514	156.09
			ZmMYB075 AtMYB58-like (G3)	GRMZM5G833253	157.77
4.08	1, 3, 6	186	Zinc finger DOF-type HCA2-like	GRMZM2G589696	160.28
			ZmMYB064 AtMYB85-like (G8)	GRMZM2G055158	166.91
			ZmCCR CCR-like	GRMZM2G099420	171.78
			ZmLac (putative laccase)	GRMZM2G169033	181.57
			ROP GTPase AtROP3-EgROP1-like	GRMZM2G001953	182.19
			ZmLac AtLac12-like	GRMZM2G388587	183.70
			ZmMYB072 (G1)	GRMZM2G162434	185.65
			ZmOMT ZRP4-like	GRMZM2G097297	192.92
			ZmCCoAOMT4	GRMZM2G033952	198.08
			ZmCCoAOMT5	GRMZM2G332522	198.08
4.09	2	227	ZmMYB065 EgMYB1-like (G4)	GRMZM2G089244	206.87
			ZmNAC AtNAC083-VNI2like	GRMZM2G123667	207.98
			ZmMYB071 AtMYB85-like (G8)	GRMZM2G138427	216.21
			ZmMYB073 (MYB42) EgMYB1-like (G4)	GRMZM2G419239	216.61
			COV1-like	GRMZM2G123790	222.80
			ABC transporter ATPDR6-like	GRMZM2G003411	231.40
			ZmPAL PAL-like	GRMZM2G153871	231.70
			KNAT7-like ELK-KNOX1	GRMZM2G433591	235.21
			ROP GTPase AtROP3-EgROP1-like	GRMZM5G846811	238.05

AtLac17 and AtLac12, and two colocalizations with putative laccase genes. AtLac12 is a gene expressed in *Arabidopsis* stems, mostly in the medium part, at a lower extent than AtLac4 and AtLac17 (Berthet et al, 2011).

MYB transcription factors putatively involved in secondary wall assembly

The regulation of phenylpropanoid biosynthesis was the first role identified for a R2R3-MYB transcription factor in plants (Paz-Ares et al, 1987; Tamagnone et al, 1998). Afterwards, MYB transcription factors containing the two-repeat R2R3 DNA binding domains were shown to be major regulators in the biosynthesis of secondary wall compounds. Based on experimental and phylogenetic approaches, R2R3-MYB genes involved in regulation of secondary wall biosynthesis can be clustered in different subgroups, depending on activities of the encoded proteins. Thus, a few of R2R3-MYB have a master function on the transcription program regulating secondary wall biosynthesis. R2R3-MYB were also shown to be either repressors or activators of the phenylpropanoid pathway and cell wall assembly. Finally, their activity can target phenolics, carbohydrates, or both components of the cell wall (Zhong and Ye, 2009; Grima-Pettenati et al, 2012; Zhong and Ye, 2012).

Only ZmMYB31 and ZmMYB42 have been proven to be related to the lignin pathway in maize. Both have a repressive effect on the expression of several genes of the lignin pathway (Fornalé et al, 2006; Sonbol et al, 2009; Fornalé et al, 2010). Other MYB factors are obviously involved in the regulation of maize lignification, which have been researched as orthologs of lignin-related R2R3 MYB genes described in *Eucalyptus* [EgMYB1, (Legay et al, 2007), EgMYB2 (Goicoechea et al, 2005)], barley [MYB hv5 and hv33, (Wissenbach et al, 1993)], and *Arabidopsis* [AtMYB46, AtMYB83, (Zhong et al, 2007; Zhong and Ye, 2012) AtMYB4, AtMYB7, AtMYB32 (Zhong and Ye, 2009; Zhou et al, 2009; Zhong and Ye, 2010; Zhong et al, 2010b; Zhong and Ye, 2012), AtMYB58, AtMYB63, AtMYB85 (Zhong et al, 2007; Zhou et al, 2009), and AtMYB75 (Bhargava et al, 2010)].

Moreover, a complete overview of the R2R3-MYB gene family in maize has been recently proposed, with a comprehensive classification of all family members, including subgroups involved in regulation of lignified cell wall biosynthesis and deposition (Du et al, 2012). Maize MYB orthologous to EgMYB1, which was shown to be a negative regulator of lignin gene expression, were gathered in the G4 subgroup "phenylpropanoid pathway" (Du et al, 2012). This group included the two ZmMYB31 and ZmMYB42 genes (Table 4), and members of this group are also orthologs of hv5 MYB. Orthologs of AtMYB58 were gathered in the G3 subgroup "lignin biosynthesis", and in the G2 subgroup with genes considered to be involved in "defense" processes. Orthologs of AtMYB85 were classified in the G8 subgroup "lignin deposition". Orthologs of EgMYB2, AtMYB46, AtMYB83, which were shown to be transcriptional activators of lignification, were unexpectedly not related to any subgroup described as involved in lignified tissue assembly by Du et al, (2012). Only ZmMYB146 was classified by Du et al (2012) with AtMYB46 and AtMYB85 in the G31 "metabolism" subgroup, together with EgMYB2 and PtMYB4 [ZmMYB146 (or GRMZM2G052606,

Table 3e - Chromosome 5.

bin	RIL	QTL pos Mbp	ZmGene annotation	GRMZM number (v2 5b60)	Gene pos Mbp
5.04	1, 3	88	ROP GTPase AtROP3-EgROP1-like	GRMZM5G803949	70.88
			Zm4CL1	GRMZM2G075333	89.15
			KNAT7-like ELK-KNOX1	GRMZM2G370332	94.45
			ZmCAD2	GRMZM5G844562	99.00
			ERF/AP2 SHINE-like	GRMZM2G085678	114.10
			ZmCAD1	GRMZM2G179981	130.66
			COV1-like	GRMZM2G146511	136.77
			ZmNAC SND1/NST1-like	GRMZM2G155816	142.18
5.06	2, 6	199	ZmNAC XND1-like	GRMZM2G094067	172.92
			ZmHCT1	GRMZM2G035584	183.61
			ZmMYB083 AtMYB58-like (G2)	GRMZM2G095904	186.41
			ZmPAL3a ZmPAL	GRMZM2G074604	186.68
			ZmPAL3b	GRMZM2G029048	186.73
			ZmPAL PAL3-like	GRMZM2G334660	186.73
			ZmPAL PAL3-like	GRMZM2G170692	186.80
			ZmNAC VND6-like	GRMZM2G315140	188.92
			MADSbox SHP1-like	GRMZM2G160565	196.55
			Zinc finger DOF-type HCA2-like	GRMZM2G140694	201.38
			GRAS SCARECROW-like	GRMZM5G885274	205.08
			AtMAP70-1-like	GRMZM5G832989	205.31
			ROP GTPase AtROP3-EgROP1-like	GRMZM2G415327	206.15
			ZmLac (putative laccase)	GRMZM2G320786	207.89
			ROP GTPase AtROP3-EgROP1-like	GRMZM2G375002	217.60
			ROP GTPase AtROP3-EgROP1-like	GRMZM2G073609	217.61

bin 10.03) is also named ZmMYB46 by Zhong et al, (2011)]. In *Arabidopsis*, the two AtMYB46 and AtMYB83 genes, together with their NAC regulators and their direct targets, were shown to be master genes regulating an array of downstream genes and thereby activating the secondary wall lignin and carbohydrate biosynthetic programs, in a multileveled feed-forward loop regulatory structure (Zhong and Ye, 2012). Several other ZmMYB that were also considered as orthologous to AtMYB46 were classified in the G13 "metabolism" subgroup (Du et al, 2012), which included the hv33 MYB gene expressed in lignifying tissue of barley (Wissenbach et al, 1993). In addition, the ZmMYB130 gene was the only maize MYB belonging to the G28 subgroup "phenylpropanoid pathway", and the closest *Arabidopsis* ortholog of ZmMYB130 is AtMYB5, a gene involved in anthocyanin metabolism. However, ZmMYB130 also has orthologous similarities with AtMYB75. This latter *Arabidopsis* gene, assigned to the G6 "anthocyanin biosynthesis" subgroup by Du et al, (2012), was shown to have a role in stem lignification (Bhargava et al, 2010) and it was also considered to be "the first transcription factor reported so far that functions as a repressor of the entire secondary cell wall program" (Zhao and Dixon, 2011).

Colocalizations between cell wall degradability QTLs and 41 MYB genes putatively involved in secondary wall lignification were shown in 26 QTL positions. Fourteen ZmMYB of the G3 and G4 subgroups (EgMYB1-like and AtMYB58-like) colocalized with cell wall degradability QTLs, corresponding to all subgroup members, except the G4 ZmMYB106, located in bin 7.05. Five ZmMYB of the G8 subgroup out of eight colocalized with QTLs, the three non-colocalizing ones being located in bins 6.04 and 7.02 (two genes). The ZmMYB146, which is the closest ortholog of EgMYB2, AtMYB46, and AtMYB83, was the only ZmMYB in the G31 subgroup. This latter ZmMYB146 colocalized in the upstream position with an important cluster of QTLs overlapping centromere area in bin 10.03, and consequently should be considered a major candidate. All three ZmMYB in the G27 subgroup, which included only maize MYB, colocalized with QTLs, while, out of the ten ZmMYB in the G13 subgroup, colocalizations with QTLs were shown for six closely related genes and a more distant one.

NAC transcription factors putatively involved in secondary wall assembly

The first demonstrations that NAC transcription factors were involved in secondary wall assembly were likely the roles of NST1 and NST2 in secondary wall thickening (Mitsuda et al, 2005) and the roles of VND6 and VND7 in vessel xylem formation (Kubo et al, 2005). Later, different *Arabidopsis* NAC proteins were shown to be master actors regulating the expression of several transcription factors and/or genes involved in secondary cell wall biosynthesis (ie Zhong and Ye, 2009; Zhong et al, 2010b). NAC factors thus include NAC SECONDARY WALL THICKENING PROMOTING FACTORS (NST 1-3, with NST3 = SND1) which are specific to fibres, and the VASCULAR-RELATED NAC-DOMAIN (VND 1-7) which are specific to vascular vessels (ie Grima-Pettenati et al, 2012). Both NST and VND factors belong to the same NAC subfamily (Yamaguchi, 2010) and

Table 3f - Chromosome 6.

bin	RIL	QTL pos Mbp	ZmGene annotation	GRMZM number (v2 5b60)	Gene pos Mbp			
6.01	2, 5	65	ZmMYB099 EgMYB2-like (G27)	GRMZM2G102790	57.54			
			ZmNAC020 SND1-like	GRMZM2G091490	66.03			
			ZmCCoAOMT1	GRMZM2G127948	79.19			
6.04	3	132	ZmPox3	GRMZM2G135108	125.01			
			Ara-CoA-acylIT (PF02458)	GRMZM2G108714	128.05			
			MADSbox SHP1-like	GRMZM2G052890	131.83			
6.05	2	152	GRAS SCARECROW-like	GRMZM2G153333	147.91			
			ZmNAC AtNAC082-VNI1-like	GRMZM2G456568	147.92			
			Zinc finger DOF-type	GRMZM2G371058	149.13			
			COV-LCV2-like	GRMZM2G073415	150.26			
			ZmMYB094 EgMYB1-like (G4)	GRMZM2G077789	150.69			
			ZmLac AtLac12/17-like	GRMZM2G146152	151.54			
			Ovate family protein OFP1/4/17-like	GRMZM2G026927	153.84			
			ZmC3H2	GRMZM2G140817	155.65			
			6.06	2, 5	164	ROP GTPase AtROP3-EgROP1-like	GRMZM2G176217	158.45
						ZmOMT ZRP4-like	GRMZM2G140996	158.66
ZmOMT ZRP4-like	GRMZM2G141026	158.69						
ZmOMT ZRP4-like	GRMZM2G102863	158.76						
ZmOMT ZRP4-like	GRMZM2G124799	158.84						
Ovate family protein OFP1/4-like	GRMZM2G127680	159.01						
COV1-like	GRMZM2G101533	159.77						
ZmMYB100 EgMYB2-like (G13)	GRMZM2G175232	162.11						
Xylan α -arabinoxyl transferase GT47	GRMZM2G059825	162.55						
Xylan α -arabinoxyl transferase GT47	GRMZM2G059845	162.56						
WD40-like	GRMZM2G038032	162.88						
ZmNAC123 SND2/SND3-like	GRMZM2G031200	164.50						
Zinc finger CCCH-type	GRMZM2G020928	164.57						
Zinc finger CCCH-AtC3H14-like	GRMZM2G149347	165.85						

function as transcriptional activators. In addition, VNI2 (VND-INTERACTING2, and to a lesser extent VNI1) is a transcriptional repressor of vessel-specific genes regulated by VND7 (Yamaguchi et al, 2010). Interactions mostly occur between VNI2 and VND7, but also exist to a lesser extent with other VND proteins and possibly other NAC factors. In maize, four orthologs of *Arabidopsis* master NAC (ZmSWN2, 3, 6, 7) were investigated and additional analysis has demonstrated that the latter ZmSWNs were functional orthologs of SND1 capable of activating the secondary wall biosynthetic program (Zhong et al, 2011). Colocalizations between cell wall degradability QTLs and maize NAC genes putatively involved in the regulation of the lignified cell wall assembly occurred for 28 genes (out of 39) in 20 QTL positions. In addition, all ZmSWN NAC colocalized with QTLs, except ZmSWN3 located in the upstream part of chromosome 6, and taking into account that ZmSWN7 corresponded in fact to two closely duplicated genes in positions 28.04 and 28.14 Mbp on chromosome 9.

Partner and target genes of MYB or NAC transcription factors y

The class II KNAT7 *Arabidopsis* gene, first described as IRX11 (Brown et al, 2005), is one of the direct targets of SND1, VND6 and AtMYB46 (Zhong et al, 2008; Ko et al, 2009). The KNAT7 gene was later shown to be a transcriptional repressor of secondary cell wall biosynthesis, in interaction with OFP1 and OFP4 (OVATE FAMILY PROTEIN 1 and 4, Li et al, 2011), and AtMYB75 (Li et al, 2012). OFP4 loss-of-function mutants also have an irregular xylem phenotype and thinner interfascicular fiber cell walls (Li et al, 2011). *Arabidopsis* KNAT7 was considered to be involved in "a negative feedback loop within these regulatory networks governing secondary cell wall biosynthesis, working antagonistically to NAC and MYB positive regulators" (Li et al, 2012). Protein-protein interactions in the phenylpropanoid metabolism also involve at least basic helix-loop-helix (bHLH), and WD40 proteins (Grima-Pettenati et al, 2012). bHLH transcription factors belong to a protein family for which many different functions have first been identified in animals, including the control of cell proliferation and development of specific cell lineages. In the rice genome, 167 bHLH genes have been shown (Li et al, 2006), likely indicating that more than 200 are present in the maize genome, with only a few members involved in lignified tissue assembly. WD40 proteins are regulatory proteins which contain a domain of nearly 40 amino acids often terminating with tryptophan (W) and aspartic acid (D), these specific traits giving their "WD40" denomination (Ramsay and Glover, 2005). Four of the five maize KNAT7 orthologs colocalized with QTL positions. All six bHLH and all four WD40 of the candidate gene list colocalized with QTLs, while seven OVATE genes out of the nine considered also colocalized with QTLs.

Zinc finger regulation factors putatively involved in lignified tissue assembly

Zinc finger proteins constitute one of the largest families of transcription factor regulatory proteins. They are involved in many regulations during plant development, including lignified tissue assembly.

Table 3g – Chromosomes 7.

bin	RIL	QTL pos Mbp	ZmGene annotation	GRMZM number (v2 5b60)	Gene pos Mbp
7.03	1, 3, 4	136	Zinc finger CCCH-type	GRMZM2G467907	127.15
			Basic helix-loop-helix bHLH	GRMZM2G082586	128.71
			Basic helix-loop-helix bHLH	GRMZM2G089501	130.24
			Zinc finger CCCH-type	GRMZM2G177229	133.53
			ZmCCR CCR-like	GRMZM2G050076	133.61
			ZmMYB111 AtMYB85-like (G8)	GRMZM2G104551	143.21
			ZmMYB116 EgMYB2-like (G29)	GRMZM2G172327	150.09
			ZmNAC AtNAC083-VNI2-like	GRMZM2G179885	158.24

Table 3h – Chromosome 8.

bin	RIL	QTL pos Mbp	ZmGene annotation	GRMZM number (v2 5b60)	Gene pos Mbp			
8.03	2, 3	65	ZmMYB123 EgMYB2-like (G13)	GRMZM2G119693	63.55			
			WD40-like	GRMZM2G040477	65.75			
			Zinc finger CCCH-type	GRMZM2G031827	70.14			
			ZmC4H1	GRMZM2G139874	83.79			
			Ara-CoA-acylT (PF02458)	GRMZM2G060210	90.58			
			Ovate family protein OFP1/4-like	GRMZM2G067376	91.86			
			COV-LCV2-like	GRMZM2G052855	100.96			
			ZmMYB132 EgMYB1-like (G4)	GRMZM2G431156	101.78			
			ZmNAC AtNAC082-VNI1-like	GRMZM2G104400	102.54			
			8.07	3, 4	168	ZmMYB130 AtMYB75-like (G28)	GRMZM2G395672	152.77
						ZmMYB117 EgMYB2-like (G13)	GRMZM2G003406	153.46
						ZmMYB124 (G14)	GRMZM2G151205	155.64
						Xylan α -arabinosyl transferase GT47	GRMZM2G000581	160.34
						Zinc finger C2H2-type	GRMZM5G898314	163.12
Zinc finger C2H2-type	GRMZM2G005236	165.55						
ZmMYB131 EgMYB1-like (G4)	GRMZM2G405094	165.88						
COV-LCV2-like	GRMZM2G046098	166.41						
Zinc finger DOF-type	GRMZM2G042218	166.58						
ZmLac AtLac17-like	GRMZM2G447271	168.76						
ZmLac AtLac12-like	GRMZM2G336337	170.04						
Ovate family protein OFP1/4-like	AC204502.4_FGP006	170.28						
ZmC4H2	GRMZM2G010468	170.45						
ZmNAC AtNAC083-VNI2-like	GRMZM2G068973	170.86						

Zinc-finger C2H2 genes were the most frequently represented transcription factors in *Eucalyptus* secondary xylem libraries (Rengel et al, 2009). The AtC3H14 zinc finger protein has been shown to activate all of the secondary wall phenolics and carbohydrate related genes tested. Both SND1 and AtMYB46 proteins bound to the AtC3H14 promoter, and AtC3H14 may function as master regulator of secondary wall biosynthesis, located downstream of AtMYB46 (Ko et al, 2009; Kim et al, 2012). In addition, DOF type (DNA-binding with one finger) domain proteins, which are plant-specific zinc finger transcription factors involved as transcriptional activators or repressors in diverse plant growth and development processes (Yanagisawa, 2004; Kushwaha et al, 2011), are also involved in lignified tissue assembly. The high cambial activity of the HCA2 *Arabidopsis* mutant resulted from an elevated expression of a DOF transcription factor (AtDOF34) preferentially expressed in the cambium, phloem, and interfascicular parenchyma cells of stems (Guo et al, 2009). Ectopic lignification was also related to variation in DOF gene expression in pom1, eli1 (ectopic lignification 1) and det3 (de-etiolated 3) *Arabidopsis* mutants, in addition to expression variation of MYB genes (Rogers et al, 2005). Finally, LIM zinc finger proteins can also be involved in the regulation of plant lignification, as shown with the tobacco NtLIM1 acting as a positive regulator of the lignin pathway (Kawaoka and Ebinuma, 2001). LIM proteins are characterized by zinc-binding domains that ligate two zinc ions. Unlike the classical zinc fingers, these domains do not bind DNA, but mediate interactions with other proteins (Matthews et al, 2009). WRKY zinc finger proteins have highly conserved WRKYGQK amino acid sequences in their N-terminal part, followed by the C2H2 or C2HC zinc-finger motifs. WRKY proteins are involved in diverse physiological and developmental processes, especially including defense against biotic stresses. However, their role in cell wall constitutive lignification has not yet been established (Wu et al, 2005; Guillaumie et al, 2010; Rushton et al, 2010; Tripathi et al, 2012). Nonetheless, the AtWRKY12 (At2g44745) gene was highly expressed in lignifying stems (<http://genecat.mpg.de/> database), and its mutation induced secondary wall formation of pith cell (Wang et al, 2010). Colocalizations occurred between cell wall degradability QTLs and 32 zinc finger genes (out of 37) putatively involved in lignified tissue assembly. None of the three maize orthologs of the At3g21270 zinc finger colocalized with QTLs, likely indicating that these family members were not related to plant cell wall degradability.

Other genes and regulation factors putatively involved in secondary wall assembly

The *Arabidopsis* COV1 (CONTINUOUS VASCULAR RING) gene encodes an integral membrane protein of unknown function which is supposed to be involved in a mechanism that negatively regulates the differentiation of stem vascular tissue by a mechanism independent of auxin (Parker et al, 2003). In addition to the firstly described COV1 gene (At2g20120), two COV1-like paralogs were later identified as the LCV2 and LCV3 genes [Like-COV-2 (At1g43130) and Like-COV-3 (At2g18460), TAIR database (<http://arabidopsis.org/>)]. Ten orthologous genes were found in the maize genome, including four, five, and one ortholog for COV1, LCV2, and LCV3, respectively.

Table 3i - Chromosome 9.

bin	RIL	QTL pos Mbp	ZmGene annotation	GRMZM number (v2 5b60)	Gene pos Mbp
9.02	1, 2	29	ZmCCoAOMT2	GRMZM2G099363	16.32
			ZmNAC032 SND1-like	GRMZM2G171395	23.28
			ZmMYB136 EgMYB2-like (G27)	GRMZM2G044824	26.63
			ZmNAC VND6-like ZmSWN1	GRMZM2G171395	23.28
			ZmNAC NST1-like ZmSWN7a	GRMZM2G041668	28.04
			ZmNAC NST1-like ZmSWN7b	GRMZM2G440219	28.14
9.04	1	123	Basic helix-loop-helix bHLH	GRMZM2G049229	112.01
			ZmOMT ZRP4-like	GRMZM2G093092	119.78
			ZmOMT ZRP4-like	GRMZM2G106172	119.84
			Brittle-stalk2 COBRA-like	GRMZM2G109326	123.24
			WD40-like (EgMYB2 Prt)	GRMZM2G123709	126.19
			Zinc finger LIM-type	GRMZM2G024887	132.70
9.06	3, 4	148	Zinc finger C2H2-type	GRMZM2G112251	139.17
			Zinc finger C2H2-type	GRMZM2G165355	143.86
			ZmNAC AtNAC083-VNI2-like	GRMZM2G126936	145.84
			GRAS SCARECROW-like	GRMZM2G028039	149.17
			ZmNAC VND6-like	GRMZM2G172053	153.53
			KNAT7-like ELK-KNOX1	GRMZM2G060507	153.74
			ATHB8 HD-ZIPIII RLD1/IFL1	GRMZM2G109987	154.65
			Zinc finger CCCH-type	GRMZM2G148090	155.98

Table 3j - Chromosome 10.

bin	RI L	QTL pos Mbp	ZmGene annotation	GRMZM number (v2 5b60)	Gene pos Mbp
10.03	4, 5	74	Zinc finger C3HC4-CHY-like	GRMZM2G056270	25.30
			ZmMYB146 EgMYB2-like (G31)	GRMZM2G052606	28.60
			MADSbox SHP1-like	GRMZM2G010669	30.87
			Zinc finger C2H2-type	GRMZM2G095323	37.64
			Basic helix-loop-helix bHLH	GRMZM2G093744	83.59
10.07	6	144	ZmMYB150 EgMYB2-like (G13)	GRMZM2G127490	87.79
			ZmMYB144 AtMYB58-like (G2)	AC206901.3_FG005	130.72
			Zinc finger LIM-type	GRMZM2G024887	132.70
			Zinc finger DOF-type	GRMZM2G010290	137.21
			ZmMYB147 EgMYB2-like (G27)	GRMZM2G081557	140.05
			ZmMYB148 AtMYB58-like (G3)	GRMZM2G097636	140.19
			ZmMYB149 AtMYB58-like (G3)	GRMZM2G097638	140.19
			UDP-xylose-4-epimerase (MUR4-like)	GRMZM2G145460	142.59

Colocalizations of cell wall degradability QTLs were shown with seven COV-like genes (out of nine), the only maize ortholog of COV LCV3 being one of the non-colocalizing genes. This fact puts into question whether this latter gene is involved in cell wall assembly.

Members of a small class III homeodomain-leucine zipper family, including AtHB8, AtHB9 (PHAVOLUTA), AtHB14 (PHABULOSA), AtHB15 (CORONA), and IFL1 (REVOLUTA), are expressed in vascular tissues and they have been considered to play regulatory roles in vascular differentiation (Talbert et al, 1995; Ratcliffe et al, 2000; Baima et al, 2001; McConnell et al, 2001; Green et al, 2005). The IFL1 *Arabidopsis* gene has two maize orthologs of which mutants have rolled leaf phenotypes (RLD1 and RLD2). The maize RLD1 gene is regulated by the ZmmiR166 miRNA (Juarez et al, 2004). The expression of the aspen PtaHB1 gene, which is also orthologous to IFL1, is also inversely correlated with the level of miR166 miRNA (Ko et al, 2006). In addition, interactions between HDZIP III and KANADI gene family members were shown to be involved in the establishment of the spatial arrangement of phloem, cambium and xylem. It was considered that HDZIP III and KANADI transcription factors control cambium activity, with KANADI proteins acting on auxin transport, and HDZIP III proteins promoting axial cell elongation and xylem differentiation (Ilegems et al, 2010). The GRAS SCARECROW and SCARECROW-like proteins belong to a plant-specific transcription factor family which contains basic leucine zipper regions and are involved in complex regulatory pathways regulating tissue patterning and differentiation (DiLaurenzio et al, 1996; Lee et al, 2008). SCARECROW proteins are thus involved in bidirectional cell signaling mediated by miRNA165/166. These proteins interfere with the transcription factor SHORT ROOT (SHR) equally expressed in stem and root, and HDZIP III proteins, towards the control of xylem patterning (Carlsbecker et al, 2010). Several colocalizations between these genes and QTLs for cell wall degradability were shown, including colocalizations with the two IFL1 orthologs (RLD1 and RLD2) in bins 1.01 and 9.06. These genes could be considered as candidates even if they are involved upstream in the pathway.

A member of the plant ROP family (EgROP1) was shown to be preferentially expressed in the cambial zone and differentiating xylem of *Eucalyptus* (Rengel et al, 2009). Its over-expression in *Arabidopsis* altered vessel formation and fibre growth in secondary xylem (Foucart et al, 2009). ROP/RAC/RAB genes encode geranyl-geranylated GTP-binding proteins (GTPases) involved in the auxin proteolysis pathway. The latter are thought to provide a universal mechanism in the control of extracellular signal transmission to intracellular metabolic pathways related to growth, differentiation, development and defense responses (Gu et al, 2004; Nibau et al, 2006). Several of them are involved in autophagy and xylem development (Kwon et al, 2010). Seven colocalizations were shown between cell wall degradability QTLs out of nine ROP1-like genes. Consequently, the latter could be considered as putative candidates.

Candidate genes were also considered in the ERF/AP2 (ethylene responsive factor/APETALA2) SHINE family. After over-expression investigations in rice, an ERF/AP2 gene was considered as an

Table 4 - Classification of colocalizing lignin-related ZmMYB, subgroups and numbers according to Du et al (2012)

MYB type	Subgroup	ZmMYB number
Master switch MYB AtMYB46-, AtMYB83-, and EgMYB2-like	G13	ZmMYB045, ZmMYB100, ZmMYB117, ZmMYB123
	G13	ZmMYB019, ZmMYB044, ZmMYB150
	G27	ZmMYB099, ZmMYB136
	G29	ZmMYB047, ZmMYB116
	G31	ZmMYB146
Lignin biosynthesis MYB AtMYB58- and AtMYB63-like	G2	ZmMYB027, ZmMYB069, ZmMYB083, ZmMYB144
	G3	ZmMYB026, ZmMYB075, ZmMYB148, ZmMYB149
Phenylpropanoid pathway MYB AtMYB4-, AtMYB7-, AtMYB32-, and EgMYB1-like	G4	ZmMYB028 (Zm31), ZmMYB073 (Zm42), ZmMYB010 (Zm38)
	G4	ZmMYB010, ZmMYB035, ZmMYB041, ZmMYB052
	G4	ZmMYB065, ZmMYB094, ZmMYB131, ZmMYB132
Phenylpropanoid pathway MYB AtMYB75-like	G28	ZmMYB130
Lignin deposition MYB AtMYB43- and AtMYB85-like	G8	ZmMYB012, ZmMYB064
	G8	ZmMYB003, ZmMYB071, ZmMYB111
Other putatively lignin-related R2R3 MYB	G1	ZmMYB024, ZmMYB072
	G14	ZmMYB055, ZmMYB124

upstream transcriptional regulator of both master and secondary target genes involved in the biosynthesis of cell wall phenolic and carbohydrate components. This ERF/AP2 transcription factor was supposed to directly bind promoter regions of NAC and MYB genes involved in regulation of cell wall assembly. Rice plants over-expressing *Arabidopsis* SHINE2 gene thus had lower lignin and higher cellulose and hemicellulose contents, without changes in plant strength or overall performances (Ambavaram et al, 2011). Only one colocalization was shown between cell wall degradability QTLs and SHINE orthologs, in bin 5.04, possibly indicating that genes of this family are weakly involved in cell wall degradability.

The *Arabidopsis* SHP1 (SHATTERPROOF MADS-box) gene, which has been shown to specify with SHP2 the lignified valve margin of mature siliques (Liljegren et al, 2000), likely has other roles in tissue lignification as it is also expressed in stems and down-regulated at the maturing stage (Ko and Han, 2004). Moreover, there are orthologs in maize including the ZmZAG5 gene, of which the functions are not known (Guillaumie et al, 2008; Barrière et al, 2009a). Colocalizations were shown between cell wall degradability QTLs and four SHP1 orthologs, but not with ZmZAG5 and APETALA1-like orthologs which are thus likely specific to floral development.

Several mutants have been shown to be altered in fiber and/or microfibril deposition. The fragile fiber FRA1 *Arabidopsis* mutant and the brittle culm 12 mutant of rice are altered in kinesin proteins (Zhong et al, 2002, Zhang et al, 2010). Kinesins are ATP-driven microtubule-based motor proteins with diverse functions in plant growth and developmental processes. These functions include the mediation of cortical microtubule activity and the orientation of cellulose microfibrils during differentiation of xylem cells. The fragile fiber *Arabidopsis* FRA2 mutant, lacking distinct S1, S2, and S3 layers in secondary wall, is altered in a gene encoding a katanin-like protein involved in fiber cell length and wall thickness (Burk and Ye, 2002). Rice brittle culm 1 and maize brittle stalk 2 mutants, which have reduced mechanical strengths, are affected in COBRA-like genes encoding putative glycosyl-phosphatidyl-inositol-anchored proteins (Li et al, 2003c; Ching et al, 2006; Sindhu et al, 2007; Dai et al, 2011). COBRA-like proteins were considered to be involved in patterning of lignin-cellulose interactions that maintain organ flexibility rather than having a direct role in cellulose biosynthesis, even if the cellulose content could be reduced in mutant plants. In addition, *Arabidopsis* MAP70 microtubule-associated proteins were shown to be essential for defining where secondary cell wall polymers were positioned and for determining the overall pattern of xylem vessel secondary cell walls (Pesquet et al, 2010; Pesquet et al, 2011). Colocalizations between maize orthologs of these different genes and cell wall degradability QTLs were shown. If these genes can induce differences in mechanical stem stiffness, it is still unclear whether such variations can induce differences in intrinsic cell wall degradability.

Genes involved p-hydroxycinnamate biosynthesis and their transfer onto cell wall components

In addition to the importance of H units, a specific trait of grass lignification is the participation of *p*-hydroxycinnamates in cell wall composition and organization of the lignified tissues, both of which have significant consequences on cell wall degradability. High negative correlations have been shown between pCA release after alkaline hydrolysis and cell wall degradability (Riboulet et al, 2008b). Furthermore, the role of ferulate cross-linkage was "tentatively estimated to account for nearly one half of the inhibitory effects of lignin on cell wall fermentation" (Grabber et al, 2009). Moreover, ferulate cross-linkages were also considered to be involved in stalk stiffness (Grabber et al, 1995; Grabber et al, 2000; MacAdam and Grabber, 2002). Consequently, the latter were thought to have negative effects on tissue friability and silage intake (Barrière et al, 2009a), in addition to their effects on cell wall degradability. As observed for pCA, the correlation between etherified FA content and cell wall degradability was always shown to be negative, both in maize and in other grasses (Casler and Jung, 1999; Méchin et al, 2000; Lam et al, 2003; Riboulet et al, 2008b; Jung and Phillips, 2010; Jung et al, 2011). In addition, despite the fact that they are often referred to as branched three-dimensional polymers, lignins are in fact largely linear. The two only known branching structures in lignins are the 5-5 and 4-O-5 bonding patterns, which cannot be formed without the participation of at least one G or H unit (Ralph et al, 2008a). However, the incorporation of free FA in lignins through bis-8-O-4 cross-coupling provides a third branching point, which occurs at low levels in normal plants, but may build up in CCR-deficient angiosperms (Ralph et al, 2008b). In addition, the degree to which the lignin polymer incorporates various phenolics in place of the three regular constitutive monolignols has likely been underestimated (Ralph, 2010). This is also true in terms of the consequences on polymer structure and cell wall degradability.

p-Coumarate is mainly esterified to the γ -position of the phenylpropane side chain of S lignin units. Most *p*-coumarate accretion occurs in tandem with lignification and 25 to 50% of maize S lignin units are acylated by pCA. This acylation occurs at the monolignol level and occurs before the transfer of S-pCA in the cell wall space (Ralph et al, 1994; Lu and Ralph, 1999; Grabber and Lu, 2007; Martinez et al, 2008). Acylation of sinapyl alcohol has a marked influence on the bonding mode of S lignin units, on the spatial organization of lignins and on their capacity to interact with polysaccharides. During monolignol polymerization, sinapyl alcohol is only slowly oxidized by maize peroxidases. Conversely, an oxidation shuttle operates in acylated conditions as the pCA component of the S-pCA conjugate is readily oxidized, with the subsequent transfer of its oxidation state to sinapyl alcohol (Boudet, 2000; Hatfield et al, 2008). The process of *p*-coumaroylation of grass lignin is dependent upon the production of the activated *p*-coumaroyl-CoA (pCA-CoA) which is a key-compound in the phenylpropanoid pathway, and upon specific transferase(s) acylating sinapyl alcohol with an activated *p*-coumaric acid. However, the corresponding acyltransferases(s) are still unidentified.

Ferulate is the major hydroxycinnamic derivative in young grass cell walls and maize cell walls can contain up to 5% ferulate monomers plus dimers (Grabber et al, 2004). Ferulic units are primarily

esterified to non-cellulosic polysaccharides, such as glucurono-arabinoxylans. Lignins and arabinoxylans are secondarily bridged through FA ether-linkages at the β -position of G units. Ferulates thus provide points of growth for the lignin polymer, act as lignin nucleation sites, and direct cell wall cross-linking (Ralph et al, 1992; Ralph et al, 1995; Jacquet et al, 1995; Ralph, 2010). In addition, the presence of ferulates linked to arabinosyl side chains of arabinoxylans provides a convenient and reliable way of cross-linking these polysaccharide chains (Ralph et al, 2004). Over 50% of wall ferulates can undergo dehydrodimerization and arabinoxylans are thus extensively cross-linked by ferulate dimerization in mature cell walls (Grabber et al, 2004). The way for biosynthesis of feruloylated arabinoxylans is still not fully understood. However, the *p*-coumarate to caffeate and ferulate conversions do not occur at the free acid level, but involve conjugates of the acids which are likely CoA-esters (Fry et al, 2000). Corroborating this fact, a putative feruloyl-CoA-arabinoxylan-trisaccharide O-hydroxycinnamoyl transferase activity (Yoshida-Shimokawa et al, 2001) has been found in suspension-cultured rice cells fed feruloyl-CoA and arabinoxylan-trisaccharide (AXX), allowing the formation of feruloyl arabinoxylan-trisaccharide (FAXX). A reaction between feruloyl-CoA and UDP-arabinose was also considered, giving a FA-Ara-UDP intermediate which could be transferred to the arabinoxylan chain in a reaction catalyzed by a feruloyl-arabinose-UDP transferase (Buanafina, 2009). Another, even less probable, hypothesis for ferulate biosynthesis in grasses has been considered from investigations in the *Arabidopsis* REF1 mutant plants, which have a reduced content in soluble sinapate esters. This mutant is affected in a sinapaldehyde dehydrogenase gene and the REF1 protein exhibited *in vitro* both sinapaldehyde and coniferaldehyde dehydrogenase activities (Nair et al, 2004). The formation of free sinapic and probably ferulic acids in *Arabidopsis* is thus catalyzed via the oxidation of the corresponding cinnamaldehydes. In maize, three mitochondrial and two cytosolic ALDH (ALDH2C and 2D) orthologs of the REF1 gene have been described (Skibbe et al, 2002; Nair et al, 2004). However, their physiological role has not yet been determined. It is indeed not yet known if such an ALDH pathway exists in maize and grasses, and especially if this pathway is functional for cell wall linked ferulate biosynthesis.

Genes involved in arabinoxylan feruloylation (Figure 2) were thus tentatively identified as acyltransferase (AcT) encoding genes specifically expressed in grasses in contrast to dicotyledons in which this particular function is supposed to be missing (Mitchell et al, 2007). One of the most differentially expressed groups of grass genes included members of the Pfam family PF02458 encoding CoA-dependent AcT including hydroxycinnamyl transferases. A study of gene deregulation in rice (Piston et al, 2010) supported the involvement of these putative feruloyl-transferases. Plants with an individual reduced expression of four members of this family had a reduced content of ester-linked ferulate in leaves and/or stems. Members of the glycosyltransferase GT47 family specifically more expressed in (C3) grasses were also shown by Mitchell et al, (2007). These genes were considered as encoding in grasses enzymes with both xylan α -1,2- and α -1,3-arabinosyl transferases, allowing the transfer of an arabinosyl residue onto an X(X) chain. Lack or reduction of this enzymatic

activity would theoretically compromise the grafting of arabinose on xylan chains, and consequently the intensity of ferulate cross-linkages. Eight orthologs of GT47 α -1,3-arabinosyl transferase genes were found in maize, including a cluster of three duplicated genes in bin 3.05. In addition, the biosynthesis pathway of hemicellulose polysaccharides could strengthen the (simultaneous) involvement of UDP-arabinose in feruloylated arabinoxylan formation. UDP-xylose is formed from UDP-glucose in the Golgi apparatus and exported to the external surface of the membrane in Golgi vesicles (Dennis and Blakeley, 2000). Afterwards, UDP-xylose can be converted into UDP-arabinose in a reversible reaction catalyzed by an UDP-xylose-4-epimerase. In *Arabidopsis*, the MUR4 mutant was shown to be affected in a UDP-xylose-4-epimerase gene (Burget et al, 2003), and four MUR4 orthologous genes were shown in maize.

Five PF02458 genes out of nine colocalized with cell wall degradability QTLs, including the three closest orthologs of rice genes investigated by Piston et al, (2010). Seven GT47 genes out of eight and three MUR4 orthologous genes out of four colocalized with cell wall degradability QTLs. Colocalizations between cell wall degradability QTL and a gene of the PF02458, GT47, and MUR4 families sometimes occurred unexpectedly without any colocalization with any FA QTL. Conversely, several esterFA and etherFA QTLs colocalized with cell wall degradability QTLs without colocalization with any considered family member gene, indicating that other mechanisms related to arabinoxylan feruloylation are still unknown. The two cytosolic ALDH also colocalized with the bin 3.07 QTL, but their putative involvement in cell wall linked ferulate biosynthesis should be supported by other data before drawing any definite conclusions. Anyway, observed colocalizations together with previously described enzymatic approaches, confirmed the probable role of feruloyl-CoA and UDP-arabinose in ferulate biosynthesis and arabinoxylan feruloylation.

Finally, it must be reminded that released and measured etherFA did not reflect total ferulate and diferulate cross-linkages in secondary cell walls because several resistant coupling of ferulate and diferulates to lignins have not yet been determined, due to the limitations of current solvolytic methods (Grabber et al, 2000). The *sfe* mutant (Jung and Philips, 2010; Jung et al, 2011) is thus probably one of the best ways to track mechanisms involved in arabinoxylan feruloylation.

Conclusions

The transfer of genetic and genomic data related to lignification and lignified tissue patterning from *Arabidopsis* and dicotyledons to maize and grasses is difficult due the vascular specific traits in grasses. In contrast to dicotyledonous plants, the vascular system of non-woody monocotyledons is characterized by the absence of bifacial cambium and secondary growth, and the major role of *p*-hydroxycinnamates in secondary wall structure. Because the emergence and evolution of lignified tracheids and vascular tissues was indeed based on the expression of a preexisting poly-phenolic

Table 5 - Colocalizations of IVNDFD and DINAGZ QTLs in 24 positions out of 34 with priority genes putatively involved in biosynthesis and polymerization of cell wall phenolic components and their regulation, in six RIL progenies [QTL pos = QTL average position, FA QTL = number of ferulate-related QTLs with - = no FA QTL or missing data, gene pos = candidate gene position (Mbp)].

bin	QTL pos	FA QTL	candidate gene	gene pos
1.01	10	-	Zinc finger C3HC4-CHY-like	10.04
1.03	46	-	ZmLac AtLac17-like	46.41
1.07	216	2	Zm4CL2	210.10
			ZmCCR1	211.52
			ZmMYB012 AtMYB85-like (G8)	215.35
			ZmF5H1	224.04
2.02	27	-	ZmNAC168 VND7-like ZmSWN4	25.75
			ZmPAL2b	28.05
			ZmPAL2a	28.12
2.04	77	1	ZmNAC096 XND1-like	48.88
			ZmMYB035 EgMYB1-like (G4)	131.35
2.08	220	1	ZmMYB028 (MYB31) EgMYB1-like (G4)	196.29
			ZmNAC080 SND2/SND3-like	220.00
3.03	9	1	Ara-CoA-acylT (PF02458)	8.30
			Ara-CoA-acylT (PF02458)	9.85
			ZmMYB044 EgMYB2-like (G13)	13.20
3.06	166	1	ZmMYB052 EgMYB1-like (G4)	173.52
			COV-LCV2-like	175.56
3.07	208	1	Zinc finger C3HC4-CHY-like	201.99
			ZmMYB047 EgMYB2-like (G29)	204.15
			FRA2-like AAA ATPase	208.05
			ZmNAC101 SND2/SND3-like	210.23
			Zinc finger CCCH-AtC3H14-like	214.75
4.04	32	1	ZmCOMT	32.25
			ZmNAC143 SND1/NST1-like ZmSWN2	38.03
4.05	83	-	ZmNAC VND7-like	59.34
			ZmMYB069 AtMYB58-like (G2)	142.64
			ZmPAL3d	143.38
			ZmPAL3c	143.47
			ZmPAL	143.51
			ZmNAC VND6-like	148.82
4.08	186	-	Zinc finger DOF-type HCA2-like	160.28
			ZmMYB064 AtMYB85-like (G8)	166.91
			ZmCCoAOMT5	198.08
4.09	227	-	ZmMYB073 (MYB42) EgMYB1-like (G4)	216.61
			COV1-like	222.80
			KNAT7-like ELK-KNOX1	235.21

pathway (Boyce et al, 2003), there is a large commonality in cell wall component biosynthesis in existing grass and "woody" plants. However, orthologs of genes involved in dicotyledonous secondary growth may have acquired distinct roles and targets during plant evolution, or may not exist in maize, while specific grass genes have appeared, as shown for arabinoxylan biosynthesis and feruloylation (Mitchell et al, 2007).

As demonstrated in investigations on the six RIL progenies, 35 cell wall degradability "consensus" QTL positions were shown. This highlights the fact that cell wall degradability of a maize plant results from the cumulative effects of many factors. In addition, the frequent clustering of QTLs for cell wall related traits raised the question of whether the underlying genetic determinant corresponds to a unique factor, or to a small set of highly linked and co-regulated genes. Colocalizations between cell wall degradability, core lignin content, and syringaldehyde QTLs could correspond to a shared mechanism involved in lignin biosynthesis and duration of lignin deposition. In addition, a greater proportion of S units in lignins could also correspond to a polymer richer in β -O-4 linkages and thus more linear, with greater masking effects on carbohydrates. Colocalizations between cell wall degradability, etherFA, and diFA QTLs could correspond to another genetic mechanism(s) involved in FA biosynthesis, in cross-linkages between arabinoxylan chains and between arabinoxylans and lignins. Occurrences of the simultaneous colocalizations between cell wall degradability, lignin content and structure, and ferulate related trait QTLs complicate the understanding and identification of the possible underlying genetic determinant(s). A cluster of linked genes involved in the different mechanisms of cell wall biosynthesis and assembly is likely the simplest situation to consider, but a single co-regulating "master" factor located upstream in the pathway of cell wall assembly can also be considered. Depending on colocalizing traits and QTLs, the two types of situation probably coexist in the maize genome. The fact that different linked genomic determinants are involved in cell wall degradability variation strengthens the possibilities of breeding for higher values of this trait without negative effects on agronomic value (yield, biotic and abiotic stress tolerance). Nevertheless, marker-assisted selection is essential in order to correctly identify the favorable recombinations, which require prior identification of the genes involved.

Out of the 228 genes colocalizing with cell wall degradability QTLs (79 % of considered candidates), it would be of interest to tentatively highlight the most probable candidates, based on their position and putative role in secondary wall assembly. No highly probable candidates were retained in ten QTL positions out of 34. In the 24 remaining QTL positions, 72 priority candidate genes were considered, based on their function in cell wall metabolism, with one to six genes per position (Table 5). At least two possible genes were present in 21 positions, the polymorphisms of which could partly explain observed differences in cell wall degradability. This illustrates the possible or probable association of several cell wall related genes as an underlying determinant of every QTL. The

Table 5 - continued

bin	QTL pos		candidate gene	gene pos
5.04	88	-	Zm4CL1	89.15
			ZmCAD2	99.00
			ZmCAD1	130.66
			COV1-like	136.77
5.05	199	2	ZmPAL3a ZmPAL	186.68
			ZmPAL3b	186.73
			ZmNAC VND6-like	188.92
			MADSbox SHP1-like	196.55
			Zinc finger DOF-type HCA2-like	201.38
6.05	152	4	COV-LCV2-like	150.26
			ZmMYB094 EgMYB1-like (G4)	150.69
			Ovate family protein OFP1/4/17-like	153.84
6.06	164	-	ZmMYB100 EgMYB2-like (G13)	162.11
			WD40-like	162.88
			ZmNAC123 SND2/SND3-like	164.50
			Zinc finger CCCH-AtC3H14-like	165.85
7.03	136	-	ZmMYB111 AtMYB85-like (G8)	143.21
8.03	65	3	ZmMYB123 EgMYB2-like (G13)	63.55
			WD40-like	65.75
8.07	168	1	ZmMYB132 EgMYB1-like (G4)	101.78
			ZmMYB130 AtMYB75-like (G28)	152.77
			ZmMYB117 EgMYB2-like (G13)	153.46
			ZmMYB131 EgMYB1-like (G4)	165.88
			COV-LCV2-like	166.41
9.02	29	-	ZmNAC NST1-like ZmSWN7a	28.04
9.06	148		ZmNAC NST1-like ZmSWN7b	28.14
			ZmNAC AtNAC083-VNI2-like	145.84
			ATHB8 HD-ZIPIII RLD1/IFL1	154.65
10.03	74	2	Zinc finger C3HC4-CHY-like	25.30
			ZmMYB146 EgMYB2-like (G31)	28.60
			ZmMYB150 EgMYB2-like (G13)	87.79
10.07	144	-	ZmMYB147 EgMYB2-like (G27)	140.05
			ZmMYB148 AtMYB58-like (G3)	140.19
			ZmMYB149 AtMYB58-like (G3)	140.19

association of MYB and/or NAC genes, and genes involved in monolignol biosynthesis could explain QTLs located in bin 1.07, 4.05, 4.08, 5.04, and 5.05. The MYB31 and MYB42 genes (Fornalé et al, 2006; Sonbol et al, 2009; Fornalé et al, 2010) in association with a NAC and COV1-like genes, respectively, could be the underlying determinant of QTLs located in bins 2.08 and 4.09. Clustering of ZmMYB and interacting protein encoding genes such as KNAT7, WD40, and/or OVATE occurred in bins 4.09, 6.05, 6.06, and 8.03, which could explain observed QTLs in these positions. AtMYB58, AtMYB75, IFL1, and AtC3H14 zinc finger genes were considered to explain observed cell wall related QTLs in the Bur0 x Col0 progeny of *Arabidopsis*, especially in light of the fact that allelic polymorphisms were shown between the two parental lines (Chavigneau et al, 2012). Similarly, maize orthologs of these genes were possibly among the candidates underlying QTLs in bins 4.05 and 10.07 for AtMYB58, in bin 8.07 for AtMYB75, in bin 9.06 for IFL1, and in bins 3.07 and 6.06 for AtC3H14 zinc finger. Genes of the PF02458, GT47, and MUR4 families, which were supposed to be involved in feruloylated arabinoxylan biosynthesis, colocalized with etherFA or diFA QTLs in only five positions out of 18. This fact again illustrated again how much the pathway of arabinoxylan feruloylation remains unknown in grasses and/or how estimates of ferulates are only partly satisfying with the currently available solvolytic methods. Correlatively, both for wall ferulates or other phenolic wall constituents, the considered list is obviously not exhaustive. Genes of still unknown function, and/or genes whose involvement in secondary wall metabolism has not yet been established, could be candidates or co-candidates underlying cell wall degradability QTLs. Such a situation should be considered at least for the EXPA5 genes, colocalizing with QTLs located in bin 6.05, that was shown differentially expressed in RILs with the favorable allele from F288 and the parental line F271 (Courtial et al, 2012a, and chapter 1).

Consequently, different strategies to determine and/or validate the gene(s) underlying QTL positions could be considered. Positional cloning and the use of NIL lines should also make it possible to greatly reduce QTL support intervals, but phenotypic variation could also be reduced after recombination occurring between favorable determinants. With the decreasing costs of genome sequencing, targeted BAC sequencing and comparison of allelic variation between the parental lines in the QTL support intervals should be a major avenue to candidate gene identification. Considering that differences in cell wall degradability between lines could be related to differences in gene expression, a quantification of transcript levels among chosen progenies, based on massive sequencing, will highlight the impact of each QTL on the overall transcriptomic profile. These transcriptomic variations should help identify functions of the candidate(s) to be discovered, and would be especially fruitful if several genes underlay QTL positions. Finally, association mapping based on large panels of lines with similar maturity could highlight genes or genomic locations related to lignin traits and cell wall degradability.

Acknowledgements

This work has been funded by the maize breeding companies (Advanta, Caussade Semences, Limagrain Europe, MaisAdour, Monsanto SAS, Pioneer Génétique, Pau Euralis, R2n RAGT Semences, SDME KWS France, Syngenta seeds) involved in the ProMais - INRA "ZeaWall" network on maize cell wall lignification and degradability. Christiane Minault, Dominique Denoue and Pascal Vernoux were in charge of seed multiplications and field experiments carried for INRA Lusignan during the past 10 years. Seed multiplications were also managed at INRA St Martin de Hinx by Jacques Laborde.

Chapitre V - Synthèse QTLs et recherche de gènes candidats

Supplementary table S1 - Candidate gene list for maize cell wall biosynthesis and assembly, established from data in maize, rice, arabidopsis, eucalyptus, and poplar (type 1 = monolignol related genes, type 2 = lignification related transcription factors, and type 3 = genes involved in lignified tissue patterning and assembly; chr = chromosome, pos Mbp = position (Mbp) in the B73 genome, colocalization with cell wall degradability QTLs = 1, no colocalization = 0)

type	gene	numgenerelease5b.60	chr	pos Mbp	colocalization	Bin
1	4CL1 (4CL2 Lubberstedt)	GRMZM2G075333	5	89,15	1,00	5.04
1	4CL2 (4CL1 Lubberstedt)	GRMZM2G048522	1	210,10	1,00	1.07
1	ABC transporter AtABCG29-like	GRMZM2G479018	2	39,75	0,00	2.04
1	ABC transporter AtABCG29-like	GRMZM2G000614	3	201,54	1,00	3.07
1	ABC transporter AtABCG29-like	GRMZM2G014282	10	4,90	0,00	10.01
1	ABC transporter ATPDR6-like	GRMZM2G391815	1	300,53	1,00	1.12
1	ABC transporter ATPDR6-like	GRMZM2G003411	4	231,40	1,00	4.09
1	ABC transporter ATPDR6-like	GRMZM2G319138	10	99,21	0,00	10.04
1	ABC transporter ATPDR7-like	GRMZM2G118243	3	10,27	1,00	3.03
1	ABC transporter EucaWood ctg6414/ PGP13	GRMZM2G085111	8	0,40	0,00	8.00
1	ABC transporter MaizeWall 3024030.2.1	GRMZM2G135199	1	299,98	1,00	1.12
1	ABC transporter MaizeWall 3871923.2.1	GRMZM2G177812	2	23,88	1,00	2.03
1	C3'H1	GRMZM2G138074	3	187,71	1,00	3.06
1	C3'H2	GRMZM2G140817	6	155,65	1,00	6.06
1	C4H1	GRMZM2G139874	8	83,79	1,00	8.03
1	C4H2	GRMZM2G010468	8	170,45	1,00	8.08
1	CAD ZmCAD1 (EgCAD1 like)	GRMZM2G179981	5	130,66	1,00	5.04
1	CAD ZmCAD2 (EgCAD2 like)	GRMZM5G844562	5	99,00	1,00	5.04
1	CAD10 (EgCAD2-like) AtCAD3/6/9	GRMZM2G700188	7	108,34	0,00	7.02
1	CAD6/SAD (EgCAD2 like) AtCAD3/6/9	GRMZM2G046070	2	10,53	1,00	2.02
1	CAD9 (EgCAD2-like) AtCAD3/6/9	AC234163.1_FG002	7	108,25	0,00	7.02
1	CCoAOMT1 ex6	GRMZM2G127948	6	79,19	1,00	6.01
1	CCoAOMT2	GRMZM2G099363	9	16,32	1,00	9.02
1	CCoAOMT3	GRMZM2G004138	2	189,28	1,00	2.07
1	CCoAOMT4	GRMZM2G033952	4	198,08	1,00	4.08
1	CCoAOMT5	GRMZM2G332522	4	198,08	1,00	4.08
1	CCR1 (ZmCCR1)	GRMZM2G131205	1	211,52	1,00	1.07
1	CCR2 (ZmCCR2)	GRMZM2G131836	7	47,72	0,00	7.02
1	CCR8 AtCCR2-like	GRMZM2G099420	4	171,78	1,00	4.06
1	CCR-like	GRMZM2G050076	7	133,61	1,00	7.03
1	COMT	AC196475.3_FG004	4	32,25	1,00	4.04
1	F5H1 (Cytochrome P450 CYP84A33v1)	AC210173.4_FG005	1	224,04	1,00	1.07
1	F5H2	GRMZM2G100158	5	22,66	0,00	5.03
1	HCT1	GRMZM2G035584	5	183,61	1,00	5.05
1	HCT2	GRMZM2G158083	2	31,83	1,00	2.04
1	OMT ZRP4-like1	GRMZM2G408458	4	18,37	0,00	4.03
1	OMT ZRP4-like2a	GRMZM2G036048	2	126,76	1,00	2.05
1	OMT ZRP4-like2b SBP1	GRMZM2G085924	2	127,09	1,00	2.05
1	OMT ZRP4-like2c	GRMZM2G147491	2	118,54	1,00	2.05
1	OMT ZRP4-like3a	GRMZM2G093092	9	119,78	1,00	9.04
1	OMT ZRP4-like3b	GRMZM2G106172	9	119,84	1,00	9.04
1	OMT ZRP4-like4	GRMZM2G097297	4	192,92	1,00	4.08
1	OMT ZRP4-like5a	GRMZM2G140996	6	158,66	1,00	6.06
1	OMT ZRP4-like5b	GRMZM2G141026	6	158,69	1,00	6.06
1	OMT ZRP4-like5c	GRMZM2G102863	6	158,76	1,00	6.06
1	OMT ZRP4-like5d	GRMZM2G124799	6	158,84	1,00	6.06
1	PAL	GRMZM2G063917	4	143,51	1,00	4.06
1	PAL more distant	GRMZM2G153871	4	231,70	1,00	4.09
1	PAL pal3-like	GRMZM2G334660	5	186,73	1,00	5.05
1	PAL pal3-like	GRMZM2G170692	5	186,80	1,00	5.05
1	PAL2a (pal2 locus)	GRMZM2G441347	2	28,12	1,00	2.03
1	PAL2b (pal2 locus)	GRMZM2G118345	2	28,05	1,00	2.03
1	PAL3a (pal3 locus ZmPAL)	GRMZM2G074604	5	186,68	1,00	5.05
1	PAL3b (pal3 locus)	GRMZM2G029048	5	186,73	1,00	5.05
1	PAL3c (pal3 locus)	GRMZM2G081582	4	143,47	1,00	4.06
1	PAL3d (pal3 locus)	GRMZM2G160541	4	143,38	1,00	4.06
1	Transferase Ara-CoA-acylT1 (PF02458)*	GRMZM2G314898	3	8,30	1,00	3.03
1	Transferase Ara-CoA-acylT2 (PF02458)	GRMZM2G305900	3	9,85	1,00	3.03
1	Transferase Ara-CoA-acylT3 (PF02458)	GRMZM2G094428	3	225,34	1,00	3.09
1	Transferase Ara-CoA-acylT4 (PF02458)*	GRMZM2G107027	6	103,32	0,00	6.03
1	Transferase Ara-CoA-acylT5 (PF02458)*	GRMZM2G108714	6	128,05	1,00	6.05
1	Transferase Ara-CoA-acylT6 (PF02458)	GRMZM2G375159	8	18,44	0,00	8.03
1	Transferase Ara-CoA-acylT7 (PF02458)*	GRMZM2G060210	8	90,58	1,00	8.03
1	Transferase Ara-CoA-acylT8a (PF02458)	GRMZM2G050072	8	143,28	0,00	8.05
1	Transferase Ara-CoA-acylT8b (PF02458)	GRMZM2G050270	8	143,28	0,00	8.05
1	Transferase xylan α -arabinosyl transferase GT47	GRMZM2G023020	2	62,92	1,00	2.04
1	Transferase xylan α -arabinosyl transferase GT47	GRMZM5G898668	3	157,02	1,00	3.05
1	Transferase xylan α -arabinosyl transferase GT47	GRMZM2G056702	3	157,05	1,00	3.05
1	Transferase xylan α -arabinosyl transferase GT47	GRMZM2G448834	3	157,08	1,00	3.05
1	Transferase xylan α -arabinosyl transferase GT47	GRMZM2G152029	5	6,04	0,00	5.01

type	gene	numgenerelease5b.60	chr	pos Mbp	colocalization	Bin
1	Transferase xylan α -arabinosyl transferase GT47	GRMZM2G059825	6	162,55	1,00	6.07
1	Transferase xylan α -arabinosyl transferase GT47	GRMZM2G059845	6	162,56	1,00	6.07
1	Transferase xylan α -arabinosyl transferase GT47	GRMZM2G000581	8	160,34	1,00	8.06
1	UDP-xylose-4-epimerase (MUR4-like)	GRMZM2G000632	1	47,47	1,00	1.03
1	UDP-xylose-4-epimerase (MUR4-like)	GRMZM2G040397	2	9,5	1,00	2.02
1	UDP-xylose-4-epimerase (MUR4-like)	GRMZM2G830983	9	136,62	0,00	9.05
1	UDP-xylose-4-epimerase (MUR4-like)	GRMZM2G145460	10	142,59	1,00	10
1	ZmALDH RF2C	GRMZM2G097699	3	221,71	1,00	3.09
1	ZmALDH RF2D	GRMZM2G071021	3	221,69	1,00	3.09
1	ZmLac Atlac2 o4 Atlac17 o4	GRMZM2G305526	1	39,93	1,00	1.03
1	ZmLac putative	GRMZM2G166857	2	229,90	1,00	2.09
1	ZmLac putative	GRMZM2G169033	4	181,53	1,00	4.07
1	ZmLac putative	GRMZM2G320786	5	207,89	1,00	5.07
1	ZmLac1a Atl17 o1	GRMZM2G072780	3	180,68	1,00	3.06
1	ZmLac1a' Atlac17 o2 Atlac2 o2	GRMZM2G164467	1	260,07	1,00	1.09
1	ZmLac1a Atlac2 o3 Atlac17 o3	GRMZM2G447271	8	168,76	1,00	8.07
1	ZmLac2 (multicopper oxidase 3)	GRMZM2G141376	9	77,55	0,00	9.03
1	ZmLac3 Atlac12 o5	GRMZM2G388587	4	183,73	1,00	4.08
1	ZmLac4 Atlac12 o3	GRMZM2G814718	1	46,41	1,00	1.03
1	ZmLac5 lac1	GRMZM2G842071	3	179,12	1,00	3.06
1	ZmLac6 Atlac2 o5 Atlac12/17 o6/6	GRMZM2G146152	6	151,54	1,00	6.05
1	ZmLac7 Atlac12-like1 o1	GRMZM2G132169	3	183,65	1,00	3.06
1	ZmLac8 Atlac12-like2 o2	GRMZM2G336337	8	170,04	1,00	8.08
1	ZmPox12	GRMZM2G103342	3	146,52	0,00	3.05
1	ZmPox2	GRMZM2G040638	1	63,65	1,00	1.04
1	ZmPox3	GRMZM2G135108	6	125,01	1,00	6.05
1	ZmPox39	GRMZM2G085967	5	47,58	0,00	5.03
1	ZmPox54	GRMZM2G088765	1	110,49	0,00	1.05
2	Basic helix-loop-helix bHLH	GRMZM2G017586	4	35,58	1,00	4.05
2	Basic helix-loop-helix bHLH	GRMZM2G082586	7	128,71	1,00	7.03
2	Basic helix-loop-helix bHLH	GRMZM2G089501	7	130,24	1,00	7.03
2	Basic helix-loop-helix bHLH	GRMZM2G049229	9	112,01	1,00	9.04
2	Basic helix-loop-helix bHLH	GRMZM2G093744	10	83,59	1,00	10.03
2	Basic helix-loop-helix bHLH	GRMZM2G009478	2	188,53	1,00	2.07
2	Basic-leucine zipper (bZIP)	GRMZM2G157722	4	57,68	0,00	4.05
2	Basic-leucine zipper (bZIP)	GRMZM2G133331	6	116,75	0,00	6.04
2	Basic-leucine zipper (bZIP) ABF2	GRMZM2G478417	1	199,49	0,00	1.07
2	KNAT7-like1 At1g62990 (ELK KNOX1)	GRMZM2G159431	1	5,01	1,00	1.01
2	KNAT7-like2 At1g62990	GRMZM2G055243	9	92,87	0,00	9.03
2	KNAT7-like3 (ELK KNOX1)	GRMZM2G060507	9	153,74	1,00	9.07
2	KNAT7-like4 (ELK KNOX1)	GRMZM2G433591	4	235,21	1,00	4.09
2	KNAT7-like5 (ELK KNOX1)	GRMZM2G370332	5	94,45	1,00	5.04
2	MYB (G1) ZmMYB024 R2R3Myb	AC165178.2_FG004	2	181,42	1,00	2.06
2	MYB (G1) ZmMYB072MaizeWall R2R3Myb	GRMZM2G162434	4	185,65	1,00	4.08
2	MYB (G13) ZmMYB019 EgMYB2-like	GRMZM2G147698	1	64,24	1,00	1.03
2	MYB (G13) ZmMYB044 EgMYB2-like	GRMZM2G052377	3	13,2	1,00	3.01
2	MYB (G13) ZmMYB045 EgMYB2-like	GRMZM2G064744	3	227,39	1,00	3.09
2	MYB (G13) ZmMYB100 EgMYB2-like	GRMZM2G175232	6	162,11	1,00	6.07
2	MYB (G13) ZmMYB117 EgMYB2-like	GRMZM2G003406	8	153,46	1,00	8.05
2	MYB (G13) ZmMYB123 EgMYB2-like	GRMZM2G119693	8	63,55	1,00	8.03
2	MYB (G13) ZmMYB150 EgMYB2-like	GRMZM2G127490	10	87,79	1,00	10.03
2	MYB (G14) ZmMYB055 MaizeWall R2R3Myb	GRMZM2G167829	3	201,93	1,00	3.07
2	MYB (G14) ZmMYB124 R2R3Myb	GRMZM2G151205	8	155,64	1,00	8.00
2	MYB (G2) ZmMYB027 AtMYB58-like	GRMZM2G048295	2	29,54	1,00	2.02
2	MYB (G2) ZmMYB069 AtMYB58-like	GRMZM2G127857	4	142,64	1,00	4.06
2	MYB (G2) ZmMYB083 AtMYB58-like	GRMZM2G095904	5	186,41	1,00	5.05
2	MYB (G2) ZmMYB144 AtMYB58-like	AC206901.3_FG005	10	130,72	1,00	10.05
2	MYB (G27) ZmMYB099 EgMYB2-like	GRMZM2G102790	6	57,54	1,00	6.01
2	MYB (G27) ZmMYB136 EgMYB2-like	GRMZM2G044824	9	26,63	1,00	9.03
2	MYB (G27) ZmMYB147 EgMYB2-like	GRMZM2G081557	10	140,05	1,00	10.07
2	MYB (G28) ZmMYB130 AtMYB75-like	GRMZM2G395672	8	152,77	1,00	8.05
2	MYB (G29) ZmMYB047 EgMYB2-like	GRMZM2G088783	3	204,15	1,00	3.07
2	MYB (G29) ZmMYB116 EgMYB2-like	GRMZM2G172327	7	150,09	1,00	7.03
2	MYB (G3) ZmMYB026 AtMYB58-like	GRMZM2G038722	2	13,30	1,00	2.02
2	MYB (G3) ZmMYB075 AtMYB58-like	GRMZM2G833253	4	157,77	1,00	4.06
2	MYB (G3) ZmMYB148 AtMYB58-like	GRMZM2G097636	10	140,19	1,00	10.07
2	MYB (G3) ZmMYB149 AtMYB58-like	GRMZM2G097638	10	140,19	1,00	10.07
2	MYB (G31) ZmMYB146 EgMYB2-like ZmSWMYB46	GRMZM2G052606	10	28,60	1,00	10.03
2	MYB (G4) ZmMYB010 EgMYB1-like	GRMZM2G084583	1	206,86	1,00	1.07
2	MYB (G4) ZmMYB028 (MYB31) EgMYB1-like	GRMZM2G050305	2	196,29	1,00	2.07
2	MYB (G4) ZmMYB035 EgMYB1-like	GRMZM2G124715	2	131,35	1,00	2.05
2	MYB (G4) ZmMYB041 EgMYB1-like	GRMZM2G041415	3	140,12	1,00	3.05
2	MYB (G4) ZmMYB052 EgMYB1-like	GRMZM2G160838	3	173,52	1,00	6.00
2	MYB (G4) ZmMYB065 EgMYB1-like	GRMZM2G089244	4	206,87	1,00	4.09
2	MYB (G4) ZmMYB073 (MYB42) EgMYB1-like	GRMZM2G419239	4	216,61	1,00	4.09

Chapitre V - Synthèse QTLs et recherche de gènes candidats

type	gene	numgenerelease5b.60	chr	pos Mbp	colocalization	Bin
2	MYB (G4) ZmMYB094 EgMYB1-like	GRMZM2G077789	6	150,69	1,00	6.05
2	MYB (G4) ZmMYB106 EgMYB1-like	GRMZM2G000818	7	176,13	0,00	7.05
2	MYB (G4) ZmMYB131 EgMYB1-like	GRMZM2G405094	8	165,88	1,00	8.07
2	MYB (G4) ZmMYB132 EgMYB1-like	GRMZM2G431156	8	101,78	1,00	8.03
2	MYB (G8) ZmMYB003 AtMYB85-like	GRMZM2G037650	1	207,18	1,00	1.07
2	MYB (G8) ZmMYB012 AtMYB85-like	GRMZM2G106558	1	215,34	1,00	1.07
2	MYB (G8) ZmMYB064 AtMYB85-like	GRMZM2G055158	4	166,91	1,00	4.06
2	MYB (G8) ZmMYB071 AtMYB85-like	GRMZM2G138427	4	216,21	1,00	4.09
2	MYB (G8) ZmMYB111 AtMYB85-like	GRMZM2G104551	7	143,21	1,00	7.03
2	NAC ZmNAC AtNAC082 VNI1 At5g09330	GRMZM2G340305	1	203,09	0,00	1.07
2	NAC ZmNAC AtNAC082 VNI1 At5g09330	GRMZM2G176677	2	26,43	1,00	2.03
2	NAC ZmNAC AtNAC082 VNI1 At5g09330	GRMZM2G125777	4	50,10	0,00	4.05
2	NAC ZmNAC AtNAC082 VNI1 At5g09330	GRMZM2G456568	6	147,92	1,00	6.05
2	NAC ZmNAC AtNAC082 VNI1 At5g09330	GRMZM2G104400	8	102,54	1,00	8.03
2	NAC ZmNAC AtNAC083 VNI2 At5g13180	GRMZM2G123667	4	207,98	1,00	4.09
2	NAC ZmNAC AtNAC083 VNI2 At5g13180	GRMZM2G336533	5	2,89	0,00	5.00
2	NAC ZmNAC AtNAC083 VNI2 At5g13180	GRMZM2G179885	7	158,24	1,00	7.04
2	NAC ZmNAC AtNAC083 VNI2 At5g13180	GRMZM2G068973	8	170,86	1,00	8.08
2	NAC ZmNAC AtNAC083 VNI2 At5g13180	GRMZM2G126936	9	145,84	1,00	9.06
2	NAC ZmNAC AtNAC083 VNI2 At5g13180	GRMZM2G430849	7	173,52	0,00	7.05
2	NAC ZmNAC AtXND1-like ZmNAC037/079	GRMZM2G094067	5	172,92	1,00	5.05
2	NAC ZmNAC AtXND1-like ZmNAC096	GRMZM2G316840	2	48,82	1,00	2.04
2	NAC ZmNAC NST1-like ZmNAC076 ZmSWN7a	GRMZM2G041668	9	28,04	1,00	9.03
2	NAC ZmNAC NST1-like ZmNAC169 ZmSWN6	GRMZM2G178998	2	22,70	1,00	2.03
2	NAC ZmNAC NST1-like ZmNAC176 ZmSWN7b	GRMZM2G440219	9	28,14	1,00	9.03
2	NAC ZmNAC SND1/NST1-like	GRMZM2G155816	5	142,18	1,00	5.04
2	NAC ZmNAC SND1/NST1-like	GRMZM2G104074	9	64,93	0,00	9.03
2	NAC ZmNAC SND1/NST1-like ZmNAC014	GRMZM2G092465	6	4,31	0,00	6.00
2	NAC ZmNAC SND1/NST1-like ZmNAC143 ZmSWN2	GRMZM2G069047	4	38,03	1,00	4.05
2	NAC ZmNAC SND1-like	GRMZM2G435824	10	119,13	0,00	10.04
2	NAC ZmNAC SND1-like ZmNAC020	GRMZM2G091490	6	66,03	1,00	6.01
2	NAC ZmNAC SND1-like ZmNAC091	GRMZM2G099144	2	47,05	1,00	2.04
2	NAC ZmNAC SND1-like ZmNAC5	GRMZM2G041746	6	106,25	0,00	6.03
2	NAC ZmNAC SND2/SND3 ANAC010	GRMZM2G166721	3	6,94	1,00	3.03
2	NAC ZmNAC SND2/SND3 ARGOS ZmNAC080/156	GRMZM2G137546	2	220,00	1,00	2.08
2	NAC ZmNAC SND2/SND3-like	GRMZM2G112681	8	20,83	0,00	8.03
2	NAC ZmNAC SND2/SND3-like ZmNAC101/182	GRMZM2G058518	3	210,23	1,00	3.08
2	NAC ZmNAC SND2/SND3-like ZmNAC123	GRMZM2G031200	6	164,50	1,00	6.07
2	NAC ZmNAC SND3/SND2-like ZmNAC019	GRMZM2G134717	8	149,81	0,00	8.05
2	NAC ZmNAC VND1 ZmNAC184 ZmSWN5	GRMZM2G025642	1	5,64	1,00	1.01
2	NAC ZmNAC VND6-like	GRMZM2G172053	9	153,53	1,00	9.07
2	NAC ZmNAC VND6-like ANAC012	GRMZM2G315140	5	188,92	1,00	5.05
2	NAC ZmNAC VND6-like ZmNAC178	GRMZM2G354151	4	148,82	1,00	4.06
2	NAC ZmNAC VND7-like ZmNAC168/186 ZmSWN4	AC212859.3_FG008	2	25,75	1,00	2.03
2	NAC ZmNAC VND7-like ZmNAC032/118 ZmSWN1	GRMZM2G171395	9	23,28	1,00	9.03
2	NAC ZmNAC VND7-like ZmNAC076 ZmSWN3	GRMZM2G052239	6	0,78	0,00	6.00
2	NAC ZmNAC VND7-like ZmNAC115	GRMZM2G048826	4	59,34	1,00	4.05
2	Ovate family protein OFP1/4/17-like	GRMZM2G026927	6	154	1,00	6.06
2	Ovate family protein OFP1/4-like	GRMZM2G075988	1	56,31	1,00	1.02
2	Ovate family protein OFP1/4-like	GRMZM2G330159	3	184	1,00	3.08
2	Ovate family protein OFP1/4-like	GRMZM2G127680	6	159	1,00	6.07
2	Ovate family protein OFP1/4-like	GRMZM2G067376	8	92	1,00	8.05
2	Ovate family protein OFP1/4-like	AC204502.4_FGP006	8	170	1,00	8.08
2	Ovate family protein OFP4-like	GRMZM2G127431	3	198	1,00	2.07
2	Ovate family protein OFP4-like	GRMZM5G845472	7	172	0,00	7.05
2	Ovate family protein OFP4-like	GRMZM2G312221	8	109	0,00	8.03
2	WD40 repeat-like EW CCAAT-HAP5	GRMZM2G038032	6	162,88	1,00	6.07
2	WD40 repeat-like EW CCAAT-HAP5	GRMZM2G040477	8	65,75	1,00	8.03
2	WD40 repeat-like EW WD40	GRMZM2G022627	2	36,94	1,00	2.04
2	WD40-like (EgMYB2 Prt)	GRMZM2G123709	9	126,00	1,00	9.00
3	AtMAP70-1 At1g68060	GRMZM5G832989	5	205,31	1,00	5.06
3	AtMAP70-1 At1g68060	GRMZM2G008556	6	115,64	0,00	6.04
3	AtMAP70-1 At1g68060	GRMZM2G039325	9	0,57	0,00	9.00
3	Brittle stalk 2 COBRA-like	GRMZM2G109326	9	123,24	1,00	9.04
3	COV LCV2-like	GRMZM2G149662	3	175,56	1,00	3.06
3	COV LCV2-like	GRMZM2G048150	5	20,68	0,00	5.03
3	COV LCV2-like	GRMZM2G073415	6	150,26	1,00	6.05
3	COV LCV2-like	GRMZM2G052855	8	100,96	1,00	8.03
3	COV LCV2-like	GRMZM2G046098	8	166,41	1,00	8.07
3	COV LCV3-like	GRMZM2G125985	10	13,55	0,00	10.02
3	COV1-like	GRMZM2G123790	4	222,80	1,00	4.09
3	COV1-like	GRMZM2G146511	5	136,77	1,00	5.04
3	COV1-like	GRMZM2G101533	6	159,77	1,00	6.06
3	COV1-like	GRMZM2G052200	8	125,01	0,00	8.05
3	FRA1-like1 "kinesin motor region"	GRMZM2G026218	7	47,93	0,00	7.02

type	gene	numgenerelease5b.60	chr	pos Mbp	colocalization	Bin
3	FRA2-like1 "AAA ATPase" ERH3	GRMZM2G054715	3	208,05	1,00	3.08
3	FRA2-like2 "AAA ATPase"/ ERH3	GRMZM2G017305	8	151,26	0,00	8.05
3	GRAS SCARECROW-like1 At1g21450	GRMZM2G431309	2	207,84	1,00	2.08
3	GRAS SCARECROW-like1 At1g21450	GRMZM5G885274	5	205,08	1,00	5.06
3	GRAS SCARECROW-like1 At1g21450	GRMZM2G153333	6	147,91	1,00	6.05
3	GRAS SCARECROW-like1 At1g21450	GRMZM2G098517	7	161,29	0,00	7.04
3	GRAS SCARECROW-like1 At1g21450	GRMZM2G028039	9	149,17	1,00	9.07
3	HD-ZIP ATHB8 HD-ZIPIII	GRMZM2G469551	1	230,54	0,00	1.05
3	HD-ZIP ATHB-8 HD-ZIPIII	GRMZM2G178102	3	123,27	1,00	3.04
3	HD-ZIP ATHB-8 HD-ZIPIII bZIP	GRMZM2G003509	1	173,23	0,00	1.05
3	HD-ZIP IFL1 (ATHB8) HD-ZIPIII rld1	GRMZM2G109987	9	154,65	1,00	9.07
3	HD-ZIP IFL1 HD-ZIPIII rld2 (ATHB8) bZIP	GRMZM2G042250	1	2,80	1,00	1.01
3	ROP AtROP3 EgROP1 At2g17800	GRMZM2G001953	4	182,19	1,00	4.07
3	ROP AtROP3 EgROP1 At2g17800	AC209819.3_FG012	8	122,95	0,00	8.05
3	ROP AtROP3 EgROP1 At2g17800	GRMZM2G102946	9	3,07	0,00	9.00
3	ROP family GTPase ROP2 AtROP3 EgROP1	GRMZM5G846811	4	238,05	1,00	4.1
3	ROP family GTPase ROP9 AtROP3 EgROP1	GRMZM5G803949	5	70,88	1,00	5.03
3	ROP Ras small GTPase Rho type AtROP3 EgROP1	GRMZM2G375002	5	217,60	1,00	5.09
3	ROP Ras small GTPase Rho type AtROP3 EgROP1	GRMZM2G415327	5	206,15	1,00	5.06
3	ROP Ras small GTPase Rho type EU968843.1*	GRMZM2G073609	5	217,61	1,00	5.09
3	ROP Ras small GTPase ROP6 AtROP3 EgROP1	GRMZM2G176217	6	158,45	1,00	6.06
3	SHINE1/ SHINE2/ SHINE3	GRMZM2G085678	5	114,10	1,00	5.04
3	SHINE1/ SHINE2/ SHINE3/ AP2/ERF	GRMZM2G106591	6	102,42	0,00	6.03
3	SHP1 MADSbox-like AGL11 STK (SEEDSTICK)	GRMZM2G052890	6	131,83	1,00	6.05
3	SHP1 MADSbox-like Kbox AP1 (APETALA1)	GRMZM2G072582	7	2,07	0,00	7.00
3	SHP1 MADSbox-like MADSbox	GRMZM2G160687	3	137,23	1,00	3.05
3	SHP1 MADSbox-like MADSbox	GRMZM2G160565	5	196,55	1,00	5.06
3	SHP1 MADSbox-like MADSbox	GRMZM2G359952	8	22,98	0,00	8.03
3	SHP1 MADSbox-like MADSbox	GRMZM2G010669	10	30,87	1,00	10.03
3	SHP1 ZmZAG5 Shatterproof Agamous	GRMZM2G003514	4	156,09	0,00	4.06
3	Zinc finger C2H2-like	GRMZM2G048154	1	42,12	1,00	1.03
3	Zinc finger C2H2-like	GRMZM2G112251	9	139,17	1,00	9.05
3	Zinc finger C2H2-like	GRMZM2G165355	9	143,86	1,00	9.06
3	Zinc finger C2H2-like	GRMZM2G095323	10	37,64	1,00	10.03
3	Zinc finger C2H2-like indeterminate growth1 (id1)	GRMZM2G171073	1	23,57	1,00	1.02
3	Zinc finger C2H2-like lot of est	GRMZM5G887286	2	222,16	1,00	2.08
3	Zinc finger C2H2-like lot of est	GRMZM5G898314	8	163,12	1,00	8.06
3	Zinc finger C2H2-like RNA recognit motif RNP-1	GRMZM2G019266	3	172,63	1,00	3.06
3	Zinc finger C2H2-like RNA recognit motif RNP-1	GRMZM2G005236	8	165,55	1,00	8.06
3	Zinc finger C3HC4-type RING	GRMZM2G077307	3	201,99	1,00	3.07
3	Zinc finger C3HC4-type RING	GRMZM2G056270	10	25,30	1,00	10.03
3	Zinc finger C3HC4-type RING lot EST	GRMZM2G062724	1	10,04	1,00	1.01
3	Zinc finger CCCH-type AtC3H14-like	GRMZM5G830949	8	146,98	0,00	8.05
3	Zinc finger CCCH-type AtC3H14-like1	GRMZM2G157927	3	214,75	1,00	3.09
3	Zinc finger CCCH-type AtC3H14-like1	GRMZM2G149347	6	165,86	1,00	6.07
3	Zinc finger CCCH-type AtU2AF35b	GRMZM2G025014	2	232,64	1,00	2.09
3	Zinc finger CCCH-type AtU2AF35b	GRMZM2G020928	6	164,57	1,00	6.07
3	Zinc finger CCCH-type AtU2AF35b	GRMZM2G177229	7	133,53	1,00	7.03
3	Zinc finger CCCH-type AtU2AF35b	GRMZM2G031827	8	70,14	1,00	8.03
3	Zinc finger CCCH-type C3H	GRMZM2G148090	9	155,98	1,00	9.07
3	Zinc finger CCCH-type RNA recognit motif RNP-1	GRMZM2G168163	2	186,90	1,00	2.07
3	Zinc finger CCCH-type RNA recognit motif RNP-1	GRMZM2G467907	7	127,15	1,00	7.02
3	Zinc finger DOF-typa HCA2 At5g62940	GRMZM2G589696	4	160,28	1,00	4.06
3	Zinc finger DOF-typa HCA2 At5g62940	GRMZM2G140694	5	201,38	1,00	5.06
3	Zinc finger DOF-type At1g21340	GRMZM2G135703	3	175,97	1,00	3.06
3	Zinc finger DOF-type At1g21340	GRMZM2G371058	6	149,13	1,00	6.05
3	Zinc finger DOF-type At1g21340	GRMZM2G042218	8	166,58	1,00	8.07
3	Zinc finger DOF-type At1g64620	GRMZM2G010290	10	137,21	1,00	10.06
3	Zinc finger DOF-type At3g21270	GRMZM2G011832	1	246,48	0,00	1.08
3	Zinc finger DOF-type At3g21270	GRMZM2G084130	5	6,11	0,00	5.01
3	Zinc finger DOF-type At3g21270	GRMZM2G178767	5	19,53	0,00	5.03
3	Zinc finger Lim-type	GRMZM2G024887	10	132,70	1,00	10.05
3	Zinc finger Lim-type mLim2 Ntlim1-like	GRMZM2G175761	3	134,61	1,00	3.05
3	Zinc finger Lim-type Ntlim1-like	GRMZM2G485184	6	21,97	0,00	6.01
3	Zinc finger Lim-type ZmLim3 MaizeWall	GRMZM2G153268	2	23,13	1,00	2.03
3	Zinc finger WRKY AtWRKY12-like	GRMZM2G123387	2	21,05	1,00	2.02
3	Zinc finger WRKY AtWRKY12-like	GRMZM2G377217	4	151,07	1,00	4.05

CHAPITRE VI - APPROCHE DE LA VARIABILITÉ
HISTOLOGIQUE DE LA RÉPARTITION DES TISSUS
LIGNIFIÉS DANS UNE TIGE DE MAÏS

6.1 Avant propos

L'amélioration de la valeur énergétique du maïs fourrage, comme celle de l'augmentation de sa dégradabilité pour la production de biocarburants de seconde génération, est fondée, comme cela a été illustré dans les chapitres précédents, sur différentes approches. Elles visent à comprendre la variabilité génétique des teneurs et des caractéristiques des différents composants phénoliques des parois d'une part, de leurs liaisons avec les composants glucidiques pariétaux d'autre part. Ces approches ont pour but d'identifier les gènes à la base de ces caractères et de leur variabilité, à des fins de connaissance, mais aussi à des fins appliquées de sélection assistée par marqueurs. L'ensemble des travaux précédemment décrits s'est basé sur l'étude des caractères biochimiques des parois sur des descendances recombinantes. Ces caractères incluent en particulier les teneurs en chacun des composés phénoliques ainsi que des estimations de solubilité enzymatique. Or, les caractères biochimiques ne sont pas les seuls facteurs responsables de la variabilité de la dégradabilité. La structure anatomique des cellules et des tissus, et notamment le ratio entre tissus lignifiés et non lignifiés des tiges jouent également un rôle.

Il a donc été décidé d'avoir de façon complémentaire aux travaux en génétique, génomique, et transcriptomique faits sur la descendance RIL F288 x F271, une démarche d'histologie sur les parents et les RILs, ou au moins une core-collection des RILs pour affiner les déterminants de la variabilité des caractères "parois" étudiés. Dans un premier temps, un protocole d'histologie a d'abord été mis au point afin de visualiser les tissus lignifiés de façon pratique, rapide, facile et surtout analysable sur un grand nombre d'échantillons. Dans un deuxième temps, cette étude a porté sur la mise en place des tissus lignifiés au sein de l'entre-nœud sous épi et au cours du développement de la plante chez les lignées parentales F288 et F271. En revanche, l'étude histologique des RILs n'a pas pu être réalisée au cours de ce travail de thèse.

Approche de la variabilité histologique de la répartition des tissus lignifiés dans une tige de maïs

Audrey Courtial^{1,2,3,§}, Yves Martinez⁴, Yu Zhang⁵, Valérie Méchin⁵, Alain Jauneau⁴, Jacqueline Grima-Pettenati^{2,3}, Yves Barrière^{1*}

¹ INRA, Unité de Génétique et d'Amélioration des Plantes Fourragères, 86600 Lusignan, France

² Université de Toulouse ; UPS ; UMR 5546, Laboratoire de Recherche en Sciences Végétales ; BP 42617 Auzeville, F-31326 Castanet-Tolosan, France

³ CNRS ; UMR 5546 ; BP 42617, F-31326 Castanet-Tolosan, France

⁴ Plateforme de microscopie, FRAIB - FR3450 CNRS, 24 chemin de Borde-Rouge BP 42617 Auzeville 31326 Castanet Tolosan

⁵ INRA, Institut Jean-Pierre Bourgin, 78026 Versailles, France.

* corresponding author

Abréviations - Ar : Surface rouge; Ab : Surface Bleue; BAS : Un centimètre du bas de l'entrenœud; G : Gaïacyle; H : *p*-hydroxyphényle; IVNDFD : Digestibilité des parois in vitro; KL : Teneur en lignine Klason; M ENE : Milieu de l'entrenœud; M HAUT : Milieu haut de l'entrenœud; M BAS : Milieu bas de l'entrenœud; NCS : Nombre de couches cellulaires dans le sclérenchyme cortical; NDF : Neutral detergent fiber; S : Syringyle.

Mots clés - maïs, histologie, lignines, fasga, phloroglucinol, autofluorescence

6.2. Résumé

La lignification des tissus étant un facteur limitant important de la dégradabilité des parois, ce travail a tout d'abord consisté à définir un protocole histologique qui permette de mettre en évidence l'ensemble des tissus lignifiés d'une tige de maïs, et cela de façon pratique, rapide, et facile à mettre en place sur un grand nombre d'échantillons. La coloration différentielle Fasga, qui permet de différencier les tissus lignifiés des tissus non lignifiés grâce à deux colorants, la safranine 0 qui colore les lignines en rouge, et le bleu alcian qui colore la cellulose en vert/ bleu, a été retenue. La présence de composés phénoliques au niveau des tissus colorés en rouge par cette technique a par ailleurs été validée en autofluorescence UV. Ensuite, un protocole automatisé d'acquisition des images et de leur analyse a été mis au point, permettant d'avoir accès au ratio entre tissus lignifiés et non lignifiés.

Le suivi cinétique de la mise en place des tissus lignifiés au niveau de l'entrenœud sous épi à trois stades de développement chez les deux lignées F288 et F271 montre, comme attendu, que la

lignification augmente de façon ascendante au sein de l'entrenœud, et plus précisément dans la région située à un cm du bas jusqu'au milieu haut de l'entrenœud. Les différences entre les stades de développement ne sont pas si nettes à partir du milieu de l'entrenœud jusqu'à sa partie supérieure, où la lignification atteint son niveau maximal. Il n'a pas été mis en évidence des différences de surfaces lignifiées (aire rouge) et non lignifiées (aire bleue) entre F288 et F271, contrairement à ce qui était attendu en fonction d'études antérieures. D'autres caractères semblent toutefois variables entre les deux lignées, dont le nombre et la taille de faisceaux vasculaires, le nombre de cellules de sclérenchyme entourant les faisceaux, la taille de ces cellules et l'épaisseur de leur paroi. A ce jour, il n'a pas encore été développé de méthodes d'analyses d'images automatisées sur ces caractères.

A terme, l'objectif est de parvenir à un niveau d'automatisation de l'acquisition de ce type de données phénotypiques tel que des recherches de QTLs de ces caractères histologiques puissent être réalisées. L'accès à ce type d'information donnera les outils permettant la mise en place de programme de sélection assistée par marqueurs permettant de sélectionner pour l'intensité de la lignification de façon tissu-spécifique.

6.3. Introduction

Le maïs, qui est aujourd'hui mondialement cultivé, est devenu la première céréale produite devant le riz et le blé. Son importance économique est liée à sa facilité de production, de récolte et de stockage, ainsi qu'à sa forte valeur énergétique. Actuellement la dégradabilité de la biomasse du maïs est un critère d'intérêt économique à la fois pour l'alimentation des ruminants ainsi que pour la production de biocarburant de seconde génération et de biogaz. En France en particulier, suite à la diminution de la valeur alimentaire et de dégradabilité des parois du maïs fourrage, après des cycles de sélection trop orientés sur des critères grain, résistance à la verse et tolérance aux stress, la compréhension des facteurs sous-tendant la dégradabilité des parois est devenue une priorité, en vue de la définition de nouveaux critères de sélection du maïs.

La tige et les feuilles de maïs sont constituées de différents tissus, dont l'épiderme, les fibres périvasculaires, les faisceaux vasculaires et le parenchyme, qui présentent chacun des intensités de lignification très variables. Il a été démontré que la structure anatomique des cellules et des tissus, et notamment le ratio entre tissus lignifiés et non lignifiés influence également la dégradabilité, en plus des variations des caractères biochimiques étudiées principalement jusqu'à ce jour. La structure anatomique influencerait d'abord l'accessibilité aux parois par les enzymes de dégradation (Wilson et Mertes, 1995), et l'épaisseur des parois limiterait le déplacement des enzymes suite à une limitation de la surface à l'intérieur des cellules (Wilson et Hatfield, 1997). L'étude microscopique des différents tissus au cours d'un processus de digestion dans le rumen montre que le phloème et le parenchyme

Figure 1 - Variabilité histologique de cinq lignées de maïs (coupes et photos Méchin, 2000, coloration Fasga, en rouge tissus lignifiés, et bleu tissus non lignifiés).

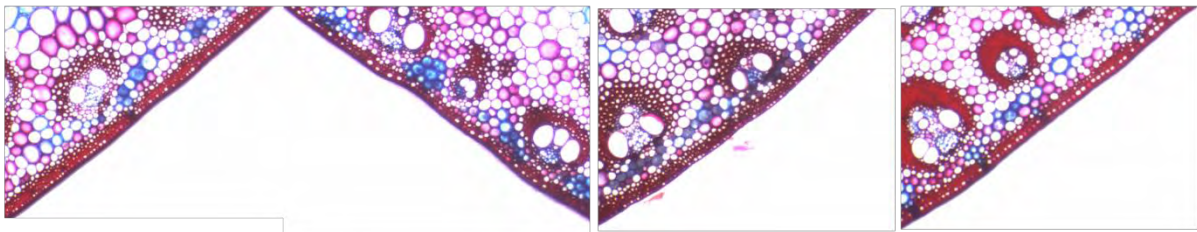


Cm484

F4

W117

F271



W117

Cm484

F271

Lh224

sont dégradés et disparaissent les premiers, tandis que les tissus lignifiés sont peu (sclérenchyme) ou pas (xylème) dégradés. La dégradation des tissus des feuilles est aussi plus rapide que celle des tissus de tiges. De plus, lorsque la plante vieillit, la dégradation des tissus diminue en même temps que les parois se lignifient (Grenet, 1997). Par ailleurs, chez le maïs, d'autres facteurs tels que la proportion de tissus lignifiés ainsi que et la surface de cortex, l'épaisseur et le nombre de couches cellulaires dans le sclérenchyme hypodermique (appelé sclérenchyme cortical par Méchin, 2000, Méchin et al, 2005) sont variables entre génotypes et auraient un impact sur la dégradabilité des plantes (Méchin, 2000; Méchin et al, 2005; Riboulet et al, 2008b). La figure 1 illustre la variabilité génétique de caractères histologiques à partir des travaux de thèse de V Méchin (2000). Les coupes de cette figure illustrent l'existence d'une variabilité à plusieurs niveaux, avec une lignification quasi-nulle du parenchyme chez les lignées Cm484 ou F4, une lignification intermédiaire chez W117, et une lignification intense chez F271. L'épaisseur du sclérenchyme cortical est également variable, plus important chez W117 et Lh224. Enfin, la lignification autour des vaisseaux est également variable, très faible chez W117, très forte chez Lh224, et intermédiaire pour Cm484 et F271. Ceci montre aussi qu'il y aurait une certaine indépendance de l'intensité de la lignification dans les différents types de tissus. L'addition de caractères histologiques aux caractères biochimiques améliore fortement la qualité de la prédiction de la valeur de dégradabilité des parois (Méchin et al, 2005). Ainsi, une équation de régression incluant des caractéristiques histologiques explique 89 % des variations de digestibilité des parois contre 57 % quand seuls les facteurs biochimiques sont pris en compte (Méchin et al, 2005). Enfin, tous les tissus ne se lignifient pas en même temps, ni à la même vitesse, au sein d'une plante et entre les différents génotypes, impliquant une nécessaire prise en compte cinétique des phénomènes. A titre d'illustration, la surface de tissus lignifiés d'une coupe transversale passe ainsi de 5 % à la base de l'entre-nœud à 17% au sommet de ce même entre-nœud (Migné et al, 1996). Tout ceci conforte l'importance de la prise en compte des deux types de caractères biochimiques et histologiques pour la compréhension de la variabilité génétique de la dégradabilité des parois des maïs.

La lignine, est un composé pariétal nécessaire à la croissance de la plante et à sa survie et constitue par ailleurs un obstacle à la dégradabilité des parois. Il est maintenant établi qu'une partie de la sensibilité des plantes au stress hydrique et aux températures élevées est liée à des collapsus des vaisseaux conducteurs qui ne supportent pas les pressions rencontrées dans ces situations de transpiration intense (Cochard, 2002; Cochard et al, 2008). De plus, une partie de la résistance aux insectes foreurs de tiges (pyrale, sésamie) est liée à la résistance mécanique des tissus périphériques vis-à-vis des larves tentant de pénétrer la tige. L'échec agronomique et la dépression de rendement observée sur les plantes *bm3* (mutation de la COMT), qui ont une dégradabilité des parois augmentée de plus de 25 %, sont ainsi essentiellement liés à la plus grande sensibilité de ce mutant aux stress abiotiques et biotiques. A contrario, la sélection forte pour le rendement dans une large gamme d'environnement, avec le choix de génotypes tolérants aux stress, et en l'absence de mesure de la

Tableau 1 - Caractérisation des deux lignées F271 et F288 (Méchin et al, 2005; IVNDFD, In Vitro NDF Digestibility ; KL (%NDF), Klason Lignin ; Ar, Aire rouge (lignifiée) ; Ab, Aire bleue (non lignifiée) régions colorées par la méthode Fasga ; Tcs, épaisseur des couches cellulaire du sclérenchyme cortical; Ncs, nombre de couches cellulaires dans le sclérenchyme cortical; As+Ag, aire totale colorée en Maïle attribuée aux tissus lignifiés contenant des sous unités S et des sous unités G.

Génotype	IVNDFD	KL (% NDF)	Ar (mm ²)	Ab (mm ²)	Ar/ (Ar + Ab)	Tcs	Ncs	As+Ag (mm ²)
F271	25,66	16,64	29,97	10,97	0,73	0,052	3,5	26,20
F288	28,78	18,53	19,30	14,64	0,57	0,067	4,0	16,42

valeur énergétique, a conduit entre les années 1975 et 2000 à une diminution moyenne de la dégradabilité des parois et de la valeur énergétique de 10 % (et très probablement à une diminution au moins égale des quantités ingérées). La réduction de la teneur en lignine ne peut donc pas être faite de façon drastique dans tous les tissus et doit être ciblée. L'étude de la répartition des tissus lignifiés est ainsi d'autant plus importante que la "fragilisation" des tiges favorables à une meilleure digestibilité dans le rumen, à une meilleure ingestibilité par les animaux, et à une meilleure dégradabilité et fermentation en bioéthanol, ne doit pas être conduite de façon semblable dans tous les tissus.

Les mécanismes génétiques régulant la mise en place des surfaces lignifiées et l'intensité de leur lignification dans les tissus d'une tige de maïs ne sont que très imparfaitement connus. Il est tout à fait probable que ces mécanismes soient tissu-spécifiques. La connaissance des variations génétiques des caractères histologiques contribuera alors à la mise en évidence des déterminants régulant la mise en place des tissus lignifiés chez une plante de maïs. Avant de pouvoir étudier une grande collection de RILs, l'objectif de ce travail était d'étudier la répartition spatiale et quantitative des tissus lignifiés et non lignifiés chez les deux lignées parentales F288 et F271 au cours de leur développement, lignées dont la digestibilité des parois est bien différente. La première partie du travail a consisté (1) à choisir la méthode histologique qui permet de révéler l'ensemble des tissus lignifiés tout en étant la plus adaptée pour un grand nombre d'échantillons (rapide, facilement analysable, reproductible), puis (2) à mettre au point l'acquisition d'image et leur analyse. La deuxième partie du travail a ensuite été d'étudier l'évolution des tissus lignifiés au niveau de l'entre-nœud sous épi à trois stades de développement chez les deux lignées retenues.

6.4. Matériel et méthodes

Description générale des lignées

Les deux lignées F271 et F288, qui ont des valeurs contrastées pour la digestibilité des parois cellulaires (Argillier et al, 1995; Argillier et al, 2000; Méchin, 2000) sont les lignées parentales d'une famille de lignées recombinantes (RIL) utilisée pour détecter des QTLs (Quantitative Trait Loci) de dégradabilité des parois, de teneur en lignines, de composition en monomère (H, G et S) des lignines, et de teneur en acide *p*-hydrocinnamiques (Roussel et al, 2002; chapitres 2 et 3). Cette famille recombinante fait actuellement l'objet de travaux visant à identifier les gènes candidats expliquant les QTLs des différents caractères pariétaux et à comprendre l'importance des gènes de régulation et de tissu-patterning dans la variabilité de dégradabilité des parois du maïs. Une première caractérisation histologique des deux lignées F271 et F288 avait été faite à la fois en biochimie et en histologie (Tableau 1) par Méchin et al, (2005).

Dispositif expérimental

Les deux lignées F271 et F288 ont été semées au champ à l'INRA de Lusignan (Vienne), durant l'été 2010, dans un essai comprenant deux blocs. Les lignes étaient séparées de 0,75 mètre et la densité était de 90 000 plantes par hectare. L'essai a été irrigué afin d'éviter un stress hydrique.

Afin de pouvoir suivre la mise en place des tissus lignifiés au cours du développement, la récolte des échantillons de maïs a été faite aux trois stades de développement panicule visible émergeant du cornet (stade 1), floraison femelle (stade 2), et maturité ensilage (stade 3). Seul l'entrenœud sous épi a été étudié, en fonction des résultats antérieurs montrant l'intérêt de cet entrenœud pour représenter l'ensemble de la plante (Guillaumie et al, 2007a; Riboulet et al, 2009). Deux entrenœuds sous épi pour chaque lignée (soit deux plantes), chaque stade et chaque bloc ont été prélevés et conservés dans de l'alcool à 70 %. Tous les entrenœuds ont été mesurés en taille (cm) pour vérifier s'il n'y avait pas d'éventuelles différences de croissance.

Détermination et coupes des régions de l'entrenœud à étudier

A la base de chaque entrenœud se situe un méristème intercalaire, responsable de l'allongement de l'entrenœud. Celui-ci génère de nouvelles cellules à la base et les cellules plus matures sont ainsi situées à l'extrémité supérieure. Quatre régions ont alors été étudiées comprenant (1) "BAS" à 1 cm du bas de l'entrenœud (région plus proche du méristème), (2) "M BAS" milieu du bas de l'entrenœud, (3) "M ENE" milieu de l'entrenœud, et (4) "M HAUT" milieu haut de l'entrenœud. Chaque entrenœud a été coupé en sections transversales de 1cm, conservées dans de l'alcool 70 % avant d'effectuer des coupes fines au vibratome.

Les coupes histologiques de chaque section ont été faites à l'aide d'un vibratome VT1000S (Leica). Cet appareil permet de couper du matériel frais ou fixé, à l'aide d'une lame rasoir qui avance progressivement tout en vibrant latéralement sur l'échantillon, préalablement fixé sur un support métallique, le tout immergé dans de l'eau. Pour ce travail, la lame de rasoir a été remplacée pour une lame en saphir synthétique (oxyde de zirconium, Delaware Diamond Knives) qui a permis de faire des coupes transversales complètes d'épaisseur de 100 µm.

Colorations et autofluorescence

Coloration Fasga - Le protocole Fasga utilisé a été mis au point par Tolivia et Tolivia en 1987. Les solutions sont préparées séparément pour chaque colorant [solution safranine à 1 % (pour 100 ml : 1g de Safranine O, 1 g d'acétate de sodium, 75 ml d'éthanol absolu, 25 ml d'H₂O et 2 ml de formaldéhyde); solution bleu alcian 8GX à 0,5 % (pour 100ml : 500mg de Bleu alcian 8GX et 100 ml d'éthanol absolu)], avant d'être mélangées pour donner la solution mère Fasga (pour 65 ml : 3 ml de solution safranine à 1 %, 11 ml de solution bleu alcian à 0,5%, 30 ml de glycérine à 99%, 20ml d'H₂O et 1 ml d'acide acétique glacial). Les coupes sont colorées pendant 16 heures dans la solution Fasga diluée au 1/8ème (soit 7 ml d'H₂O et 1 ml de solution mère). Les coupes sont ensuite rincées à l'eau et

conservées dans cette dernière. La coloration Fasga est une coloration différentielle des tissus basée sur le mélange de deux colorants. Cette méthode permet de différencier les structures en différentes couleurs (Tolivia et Tolivia, 1987). La safranine O colore spécifiquement les lignines en rouge et le bleu alcian colore la cellulose en vert/ bleu donnant un couleur bleu aux tissus non lignifiés. Le bleu Alcian, colorant textile, a été adopté en histochimie en 1950. C'est un colorant basique polyvalent, soluble dans l'eau. Il réagit avec les groupements anioniques et peut s'employer à différents pH (le plus souvent 1,0 et 2,5). C'est à un pH 1,0, que le bleu alcian colore la cellulose (vert à bleu). La safranine O, est un colorant spécifique des lignines, qui sont alors colorées en rouge (Johansen, 1940). Les avantages du bleu alcian sont la compatibilité avec la safranine O et sa forte affinité pour les structures qui ont peu d'affinité avec la safranine. Ainsi l'affinité de la safranine est augmentée par compétition. Ces propriétés permettent une double coloration efficace des tissus, dont les constituants pariétaux sont révélés de la manière suivante : la cellulose est verte ou bleue (différentes tonalités), le sclérenchyme est rose ou bien rouge, le collenchyme est vert ou violet- bleu, la lignine est rouge avec différentes tonalités, la subérine et la cutine sont rouges à brun orangé, et les tannins sont rougeâtres.

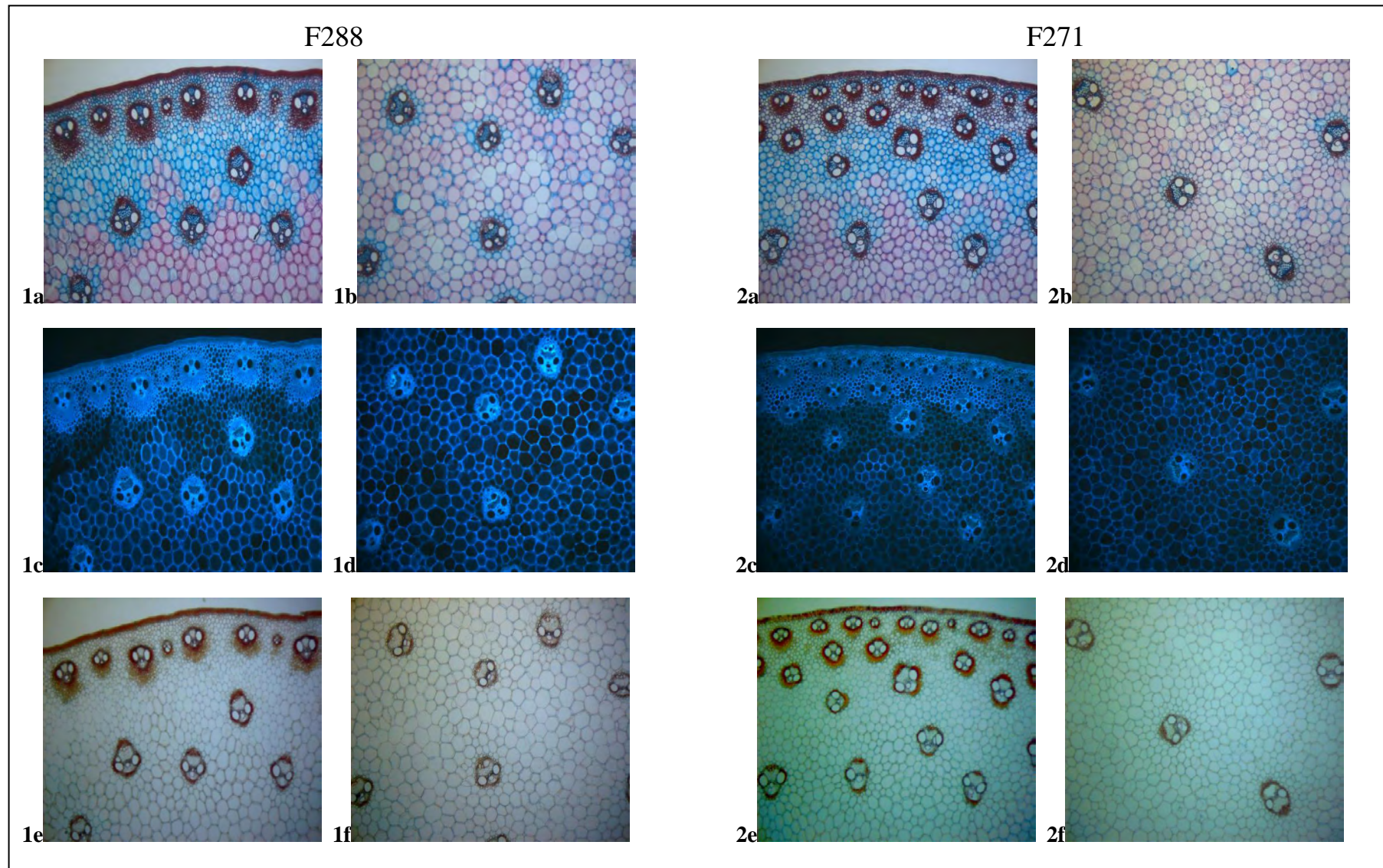
Coloration phloroglucinol/HCl (Réaction de Weisner) - La coloration au phloroglucinol est instantanée (3 min). Seules quelques gouttes de phloroglucinol en solution chlorhydrique suffisent pour démarrer la réaction. Cette coloration est instable dans le temps, la couleur rouge vire rapidement. La prise d'image doit donc être rapide. Le réactif phloroglucinol-HCl permet de révéler les structures aromatiques des aldéhydes contenues dans les lignines. La réaction au phloroglucinol est caractérisée par une attaque nucléophile sur les groupes aldéhydes et forme un intermédiaire quinone par résonance. Le phloroglucinol réagit avec le coniféraldéhyde, le sinapaldéhyde en donnant une coloration rose, mais aussi avec la vanilline et le syringaldéhyde en donnant une coloration rouge marron (Pomar et al, 2002). Le groupe fonctionnel cinnamaldéhyde non substitué est responsable de cette coloration. Cette technique permet donc de visualiser rapidement les parois des tissus lignifiés contenant des coniféraldéhydes et des sinapaldéhydes.

Autofluorescence - Aucune étape préalable n'est nécessaire avant observations des coupes. La microscopie à fluorescence est une technique utilisée pour visualiser les composés phénoliques contenus dans les parois. De nombreux composés de plantes comme la chlorophylle, les lignines, la subérine, la cutine et d'autres composés phénoliques peuvent autofluorescer grâce à leurs propriétés intrinsèques (Yeung, 1998). Après excitation dans l'UV (360 nm), les lignines apparaissent en bleu, la cutine et la subérine apparaissent en blanc argenté, les composés phénoliques autres que les lignines varient du vert au bleu, et la chlorophylle apparaît en rouge. Cette technique permet de détecter des composés fluorescents à des concentrations très faibles, sans besoin de coloration.

Observations à la loupe binoculaire et au microscope photonique

A la suite des colorations, les coupes ont été observées d'une part à la loupe binoculaire avec l'objectif x 0,8, qui a permis d'observer la coupe complète en une seule fois ou en deux fois (ce

Figure 2 - Comparaison des différentes techniques histologiques entre F288 et F271 au milieu de l'entreœud (mENE) au stade panicule visible (stade 1). Les observations ont été faites au microscope x5. 1a et 2a - cortex observé après coloration Fasga; 1b et 2b - moelle observée après coloration Fasga; 1c et 2c, cortex observé par autofluorescence; 1d et 2d - moelle observée par autofluorescence; 1e et 2e - cortex observé après coloration phloroglucinol/HCl; 1f et 2f - moelle observée après coloration phloroglucinol/HCl.



grossissement a permis d'observer la répartition des différents tissus lignifiés), et d'autre part au microscope photonique avec l'objectif x 5 (cet objectif a permis l'observation d'une partie de la coupe). Seule la partie corticale située à l'opposé de l'insertion de l'épi et la partie centrale de la coupe ont été prises en images. Ces observations ont permis d'avoir plus de détails sur ces zones. Les coupes pour lesquelles deux images ont été acquises à la loupe binoculaire ont été par la suite assemblées grâce au "plugin mosaïque" du logiciel Image J.

Analyses d'images

Seules les images colorées au Fasga et acquises à la loupe binoculaire ont été analysées grâce au logiciel d'analyse et de traitement d'images, ImageJ. Après une étape de délimitation du contour, la surface de tissus lignifiés (rouges) et la surface de tissus celluloseux (bleus) ont été mesurées en ajustant les niveaux de colorimétrie, saturation et brillance dans la fenêtre "color threshold".

6.5. Résultats

Approches méthodologiques

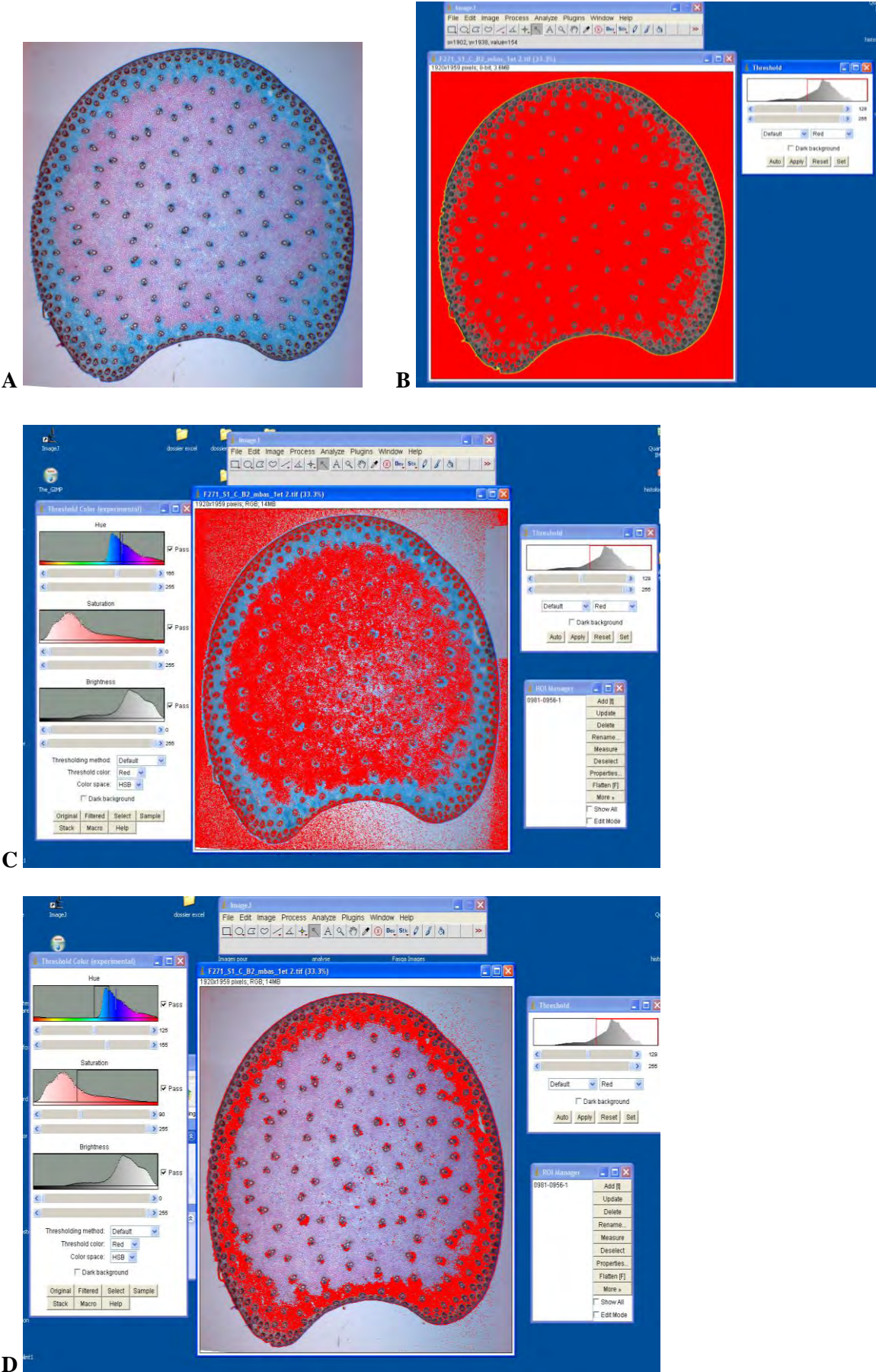
Choix de la méthode histologique

Plusieurs techniques histologiques permettent de révéler les tissus lignifiés (voir matériel et méthodes). La première étape du travail a été de choisir la méthode histologique qui permettait de révéler l'ensemble des tissus lignifiés tout en étant la plus adaptée à un grand nombre d'échantillons (rapide, facilement analysable, reproductible).

La coloration Fasga, la coloration phloroglucinol/HCl et la technique d'autofluorescence ont été analysées sur une seule plante par génotype, au milieu de l'entreœud sous épi pour les trois stades de développement afin de comparer les résultats entre eux. Au total pour chaque entreœud, deux coupes successives faites au vibratome ont été utilisées pour faire la comparaison des trois méthodes, la coupe observée en autofluorescence étant ensuite colorée au phloroglucinol/HCl. La partie corticale et la partie centrale de la coupe a été observée au microscope objectif x 5. Pour chaque entreœud, les images prises correspondent exactement à la même région.

Les deux colorations phloroglucinol/HCl et Fasga, et l'autofluorescence, montrent bien des points communs mais également des différences (Figure 2). Alors que le sclérenchyme cortical et le sclérenchyme des faisceaux libéro-ligneux sont colorés en rouge avec le Fasga et le phloroglucinol/HCl et autofluorescent, le parenchyme médullaire n'est pas coloré en rouge avec le phloroglucinol/HCl mais est rouge avec le Fasga et il autofluoresce. Le parenchyme cortical et quelquefois la première assise de parenchyme entourant les faisceaux libéro-ligneux qui apparaissent bleu/vert pâle en UV (nuance de bleu différente), sont bleus avec le Fasga et ne présentent pas de

Figure 3 - Analyse des images avec le logiciel Image J. (A) Original Fasga. (B) Délimitation de la zone à traiter. (C) Calcul de l'aire rouge. (D) Calcul de l'aire bleue.



coloration avec le phloroglucinol/HCl. Le Fasga et l'autofluorescence permettraient ainsi de différencier le parenchyme lignifié d'un parenchyme cellulosique non lignifié alors que le phloroglucinol ne le permet pas. Cette possibilité de distinction est néanmoins quelquefois moins nette en autofluorescence.

Au vu de ces résultats et de sa stabilité dans le temps, la coloration Fasga a été retenue pour l'étude de l'évolution des tissus lignifiés au niveau de l'entrenœud sous épi aux trois stades de développement chez les deux lignées de digestibilité différente.

Mise au point du protocole choisi, acquisition et analyse d'image

Les premières étapes de cette étude histologique ne posent pas de problème. L'ensemble des échantillons sont conservés dans de l'éthanol 70%. Les coupes réalisées avec la lame saphir sont parfaitement homogènes. Le temps de coloration est le même pour tous les échantillons. L'ensemble des étapes suivant la coloration a en revanche nécessité des mises au point plus difficiles. L'acquisition des images a été effectuée à la loupe binoculaire avec l'objectif x 0,8. Cet objectif permet d'observer l'intégralité de la coupe en une fois ou deux fois tout en ayant une résolution suffisante pour pouvoir quantifier la surface rouge et la surface bleue. La répartition des lignines pour l'ensemble de la coupe transversale peut donc être étudiée. Afin de pouvoir analyser l'ensemble de la coupe en une fois, les coupes, pour lesquelles deux images ont été acquises, ont été par la suite assemblées grâce au plugin mosaïque du logiciel Image J. La dernière étape est la quantification des surfaces lignifiées (rouges) et des surfaces non lignifiées (bleues). La mise au point de l'analyse d'images avec ImageJ a été faite sur plusieurs coupes de F288 et F271 afin de s'assurer des réglages et de leur reproductibilité. Les coupes ont été choisies selon plusieurs critères incluant un niveau de lignification varié (coupes les moins matures et les plus matures), et des coupes avec des régions non colorées (coupes abîmées, ou cellules non colorées). Les valeurs retenues pour déterminer l'aire rouge et l'aire bleue dans la fenêtre "color threshold " sont respectivement 165-255 et 125-155 pour les niveaux de colorimétrie, 0-255 et 90-255 pour la saturation, 0-255 et 0-255 pour la brillance (Figure 3).

Evolution des tissus lignifiés au cours du développement chez les lignées F271 et F288

Vérification de l'homogénéité des tailles des entrenœuds

Les mesures ont confirmé une homogénéité convenable de la taille des entrenœuds de chaque lignée (Tableau 2). Les entrenœuds matures de F271 sont plus allongés que ceux de F288, correspondant d'ailleurs au fait que cette lignée est plus haute au champ. En revanche, la lignée F271 a des entrenœuds ayant un diamètre un peu plus faible (données non présentées).

Tableau 2 - Taille des entrenœuds prélevés pour les études histologiques. Le nom de l'entrenœud (ENE) est désigné par une lettre. La longueur est donnée en centimètres. Les différents stades de développement sont : panicule visible (stade 1), floraison femelle (stade 2), et maturité ensilage (stade 3).

Lignées	ENE	Bloc	Longueur ENE (cm) Stade 1	Longueur ENE (cm) Stade 2	Longueur ENE (cm) Stade 3
F288	A	1	14,0	15,5	15,0
F288	B	1	16,0	16,5	15,0
F288	C	2	17,0	16,5	16,0
F288	D	2	15,0	16,0	17,5
F271	A	1	15,5	17,5	14,0
F271	B	1	18,5	18,5	18,5
F271	C	2	16,0	16,0	16,5
F271	D	2	17,0	19,0	19,0

Tableau 3 - Valeurs moyennes de l'importance relative des tissus lignifiés Ar/(Ar+Ab) pour les lignées F271 et F288 aux différents stades de développement (Bas - région un centimètre du bas de l'entrenœud; M Bas - milieu bas de l'entrenœud; M ENE - milieu de l'entrenœud; M Haut - milieu haut de l'entrenœud).

	Rapport Ar/(Ar+Ab) stade 1	Rapport Ar/(Ar+Ab) stade 2	Rapport Ar/(Ar+Ab) stade 3
F271			
Bas	0,36	0,60	0,73
M Bas	0,65	0,79	0,87
M ENE	0,81	0,85	0,89
M. Haut	0,82	0,85	0,88
F288			
Bas	0,50	0,61	0,79
M Bas	0,78	0,83	0,84
M ENE	0,85	0,87	0,87
M. Haut	0,88	0,88	0,88

Evolution des tissus lignifiés au sein de l'entrenœud et au cours du développement

Les résultats pour chaque entrenœud sont donnés en annexe 1. Les valeurs moyennes, pour les quatre échantillons par génotype, de l'importance relative des tissus lignifiés, $Ar/(Ar+Ab)$ sont, quant à elles, données dans le tableau 3 pour les deux lignées F271 et F288. Au sein d'un entrenœud, la surface de tissus lignifiés (aire rouge) évolue et devient de plus en plus importante au fur et à mesure que l'on monte dans l'entrenœud, la croissance s'effectuant à partir de la base. A l'inverse, l'aire bleue (tissus non lignifiés) diminue (Figure 4). Cette évolution s'atténue avec l'âge de la plante. Il en est de même quand on observe l'évolution par rapport aux stades de développement, pour une même région de l'entrenœud, l'aire rouge augmente alors que l'aire bleue diminue. Cette évolution est moins marquée à partir du milieu de l'entrenœud et est même absente pour le milieu haut de l'entrenœud chez F288.

L'évolution de lignification la plus remarquable se produit donc de façon ascendante à partir de la région du bas jusqu'au milieu haut de l'entrenœud. Ces différences ne sont pas si marquées à partir du milieu de l'entrenœud jusqu'au milieu haut où la lignification atteint son niveau maximal.

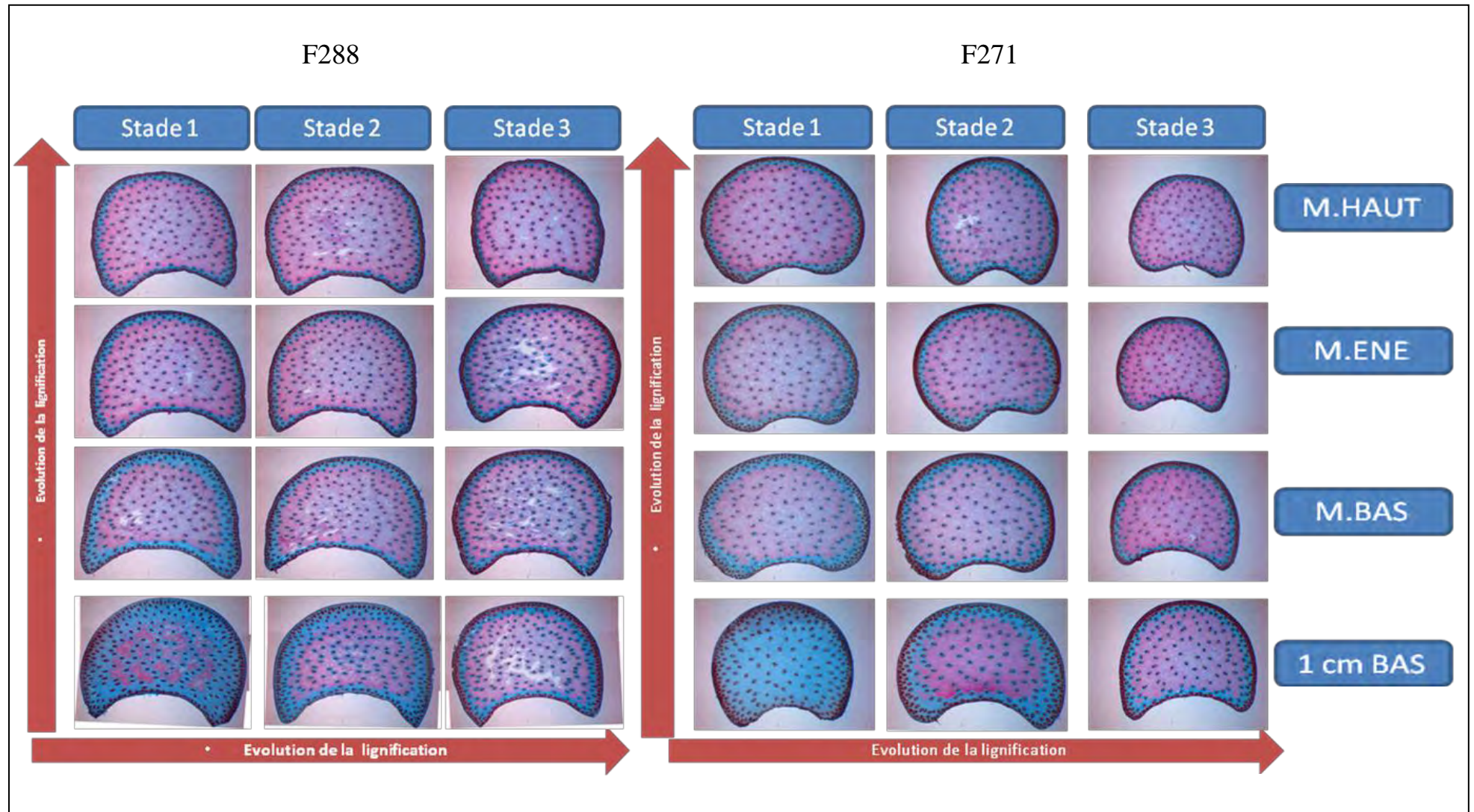
Comparaison des deux lignées F271 et F288

Entre les deux lignées F271 et F288, il n'y a pas de différence évidente de surface rouge et bleue lors d'une observation de l'ensemble de la coupe (Figure 4). L'analyse d'images donne une estimation plus précise de cette absence de différences (annexe 1). Par exemple, les valeurs $Ar/(Ar+Ab)$ chez F288 au stade 3, maturité ensilage et au milieu de l'entrenœud (M ENE) ont comme moyenne 0,87 avec comme valeur minimale 0,84 et comme valeur maximale 0,91. Chez F271, pour le même stade et la même région, la moyenne $Ar/(Ar+Ab)$ est de 0,89 avec comme valeur minimale 0,81 et comme valeur maximale 0,92. Les valeurs moyennes sont donc très similaires entre les deux lignées F288 et F271 (Tableau 3), ce qui est contraire aux observations antérieures et à ce qui était attendu. Il y avait en effet une différence de 0,16 dans le rapport $Ar/(Ar+Ab)$ entre les deux lignées dans les études réalisées précédemment (Tableau 1).

6.6. Discussion*Mise en place du protocole d'étude histologique des tissus lignifiés du maïs*

Les différences de coloration observées entre le Fasga et le phloroglucinol peuvent s'expliquer par le fait que ces colorations ne s'adressent pas aux mêmes constituants de la paroi cellulaire. La coloration Fasga, coloration différentielle, permet de différencier les tissus lignifiés des tissus non lignifiés grâce à deux colorants (la safranine O qui colore les lignines en rouge et le bleu alcian qui colore la cellulose en vert/bleu). Le phloroglucinol ne révèle que les structures aromatiques des cinnamaldéhydes (coniféraldéhyde, sinapaldéhyde) contenus dans les lignines des parois en donnant

Figure 4 - Comparaison des différentes régions étudiées de l'entrenœud sous épi entre les lignées F288 et F271, aux trois stades de développement. Coloration Fasga et observations à la loupe binoculaire x 0.8.



une coloration rose (Pomar et al, 2002). Cette technique, en colorant uniquement les aldéhydes des lignines, est donc limitée. Diverses études histochimiques révèlent que les unités coniféraldéhydes sont plus présentes dans les étapes jeunes des tissus notamment dans la paroi cellulaire du xylème (Pomar et al, 2002). De plus, la coloration rose peut être faible voir même absente dans les lignines qui contiennent de forte teneur en sous unités syringyle (Nakano et Meshitsuka. 1992).

La présence de composés phénoliques, au niveau des régions colorées en rouge par le Fasga et non colorées au phloroglucinol, a pu être validée par la technique d'autofluorescence qui montre la présence de composés phénoliques, ces régions étant autofluorescentes suite à une excitation dans l'UV. L'absence de coloration rouge avec la technique phloroglucinol peut signifier un manque d'aldéhydes et non pas une absence de lignines. En 2005, Méchin et al, avaient déjà préféré la coloration Fasga, qui donne surtout plus de contraste pour les analyses automatisées.

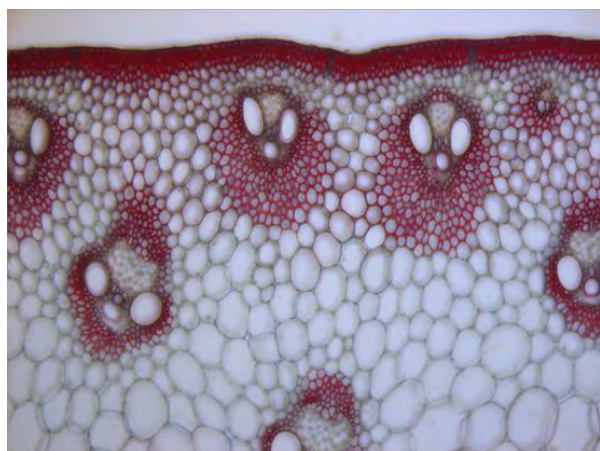
Evolution des tissus lignifiés au cours du développement chez les lignées F271 et F288

Pour un même génotype au même stade et à la même région, des différences pour les différents critères mesurés existent entre les quatre échantillons. Ces différences sont très importantes pour la région à 1cm au bas de l'entre-nœud et au milieu bas, au stade le plus jeune étudié "panicule visible". Par ailleurs, la région "1cm du bas" présente le plus de variation inter-plante, cette région devant être plus ou moins éloignée du méristème et donc étant plus ou moins différenciée. Les différences de lignification entre les quatre échantillons d'un même génotype au même stade et même région s'atténuent avec la maturité des tissus. Ces approches montrent aussi que des études de transcriptomique peuvent ainsi donner des résultats différents en fonction des zones de tissus prélevés.

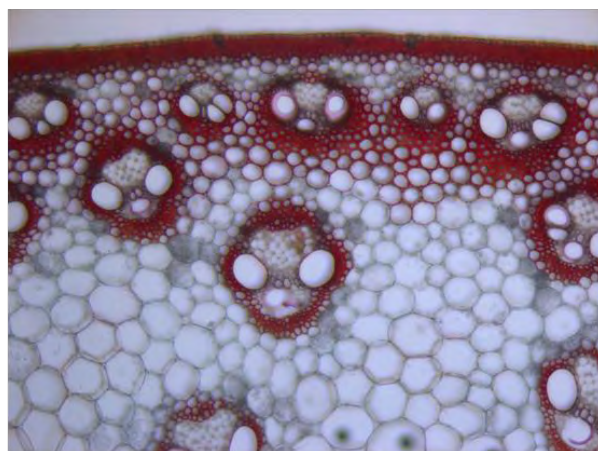
L'évaluation des données obtenues après l'analyse des images montre une faible différence au niveau de la lignification $Ar/(Ar+Ab)$ entre les lignées F271 et F288, pourtant de digestibilité différente. Ce résultat diffère des résultats obtenus par Méchin et al, en 2005 (Tableau 1). Il n'y a pas de certitudes sur les raisons pouvant expliquer ces résultats différents. Toutefois, il semblerait que la lignée Cm484, qui a été caractérisée comme ayant un parenchyme non lignifié, puisse présenter un parenchyme effectivement lignifié en conditions chaudes avec un certain niveau de déficit hydrique (Méchin, comm. pers). Il est par ailleurs connu que la lignification est un mécanisme classique de réaction aux stress, de nature biotique ou abiotique. Il est alors possible et même probable que les conditions de culture rencontrées par les plantes à Lusignan en 2010 aient généré un stress, en particulier lié à une irrigation un peu limitée. Si cette hypothèse est correcte, cela complexifiera à un niveau supplémentaire la recherche des déterminants moléculaires, au sens où il faudra séparer ce qui est lignification constitutive du parenchyme, de ce qui pourrait être plus une lignification réactive de stress, dépendante d'une autre voie de régulation.

Figure 5 - Comparaison des deux lignées F288 et F271 au milieu de l'entreœud (mENE) au stade maturité ensilage (S3). Coloration phloroglucinol et observations au microscope x 10.

F288



F271



6.7. Conclusions et perspectives

Ce travail a permis tout d'abord de mettre en place un protocole histologique utilisant la coloration différentielle Fasga qui permet de mettre en évidence l'ensemble des tissus lignifiés, facteur limitant la dégradabilité, de façon pratique, rapide, facile et surtout analysable en condition partiellement automatisée sur un grand nombre d'échantillons. Dans la mesure où la lignification évolue au cours du développement, notamment dans la moitié inférieure de l'entrenœud, des études comparatives sur de plus grand effectif de génotypes devront être réalisées en se focalisant sur des coupes faites dans la partie supérieure où l'évolution de lignification est moindre, voire absente, la lignification étant terminée à ce niveau au moment du prélèvement. Si, dans les conditions de culture des plantes, aucune différence dans les surfaces rouges et bleues n'a été mise en évidence entre F288 et F271, d'autres caractères semblent toutefois varier comme le nombre et la taille des faisceaux, le nombre de cellules de sclérenchyme entourant les faisceaux, la taille de ces cellules et l'épaisseur de leur paroi (Figure 5). Il faudrait tenter alors de développer pour ces caractères une automatisation en analyse d'images et estimer alors leurs relations avec les différences de dégradabilité.

A terme, l'objectif est de parvenir à un niveau d'automatisation tel que des recherches de QTLs de ces caractères histologiques puissent être réalisées. S'il n'est sans doute pas réaliste d'envisager ce type de recherche sur des effectifs de 250 individus avec plusieurs répétitions, il est possible en revanche de dériver des familles recombinantes ciblées sur ces caractères, et/ou d'utiliser des familles existantes au sein desquelles des différences histologiques seraient nettes, et de rechercher ces QTLs particuliers sur une core-collection définies à partir d'un marquage suffisamment dense de toutes les descendances RILs. Il a été décidé de suspendre la démarche d'étude de la collection de RILs F288 x F271 au vu des résultats sur les surfaces lignifiées, et cela d'autant plus que ce type de travail est très coûteux en temps. Toutefois, les entrenœuds de la core-collection, prélevés en 2011 à l'INRA de Lusignan, au stade maturité ensilage ont été conservés, dans la mesure où une variabilité peut être observée dans la descendance, même en l'absence de variation chez les parents. Une autre descendance RIL spécifiquement destinée à l'étude de la variabilité histologique est en cours de production. Les lignées parentales de cette descendance, Cm484 et F271, sont bien plus extrêmes que ne le sont F288 et F271. La figure 1 illustre les différences entre ces deux lignées, qui ont par ailleurs des valeurs de dégradabilité des parois bien différentes (24.6 et 35.8 % en IVNDFD, Méchin et al, 2000). Cette population, d'effectif supérieur à celui de la famille recombinante F288 x F271, était semée cet été à St Martin de Hinx où 250 RILs étaient en multiplication au stade F6 pour F7. Certaines descendances de cette population seront aussi un support privilégié pour des approches de transcriptomique après une capture laser de certains tissus (laser capture microdissection, LCM) afin de mettre en évidence les gènes différenciellement exprimés de façon tissu-spécifique. L'accès à des régulations spécifiques de la lignification des vaisseaux et à ceux de la lignification du parenchyme donnera ainsi des outils clés

pour la mise en place de programme de sélection assistée par marqueurs de génotypes plus dégradables, tout en ayant des vaisseaux résistants aux contraintes hydriques.

Remerciements

Nous remercions Paola Paredes pour sa contribution à ces travaux, ainsi que Marie-Françoise Devaux (INRA BIA Nantes) pour ses conseils dans la réalisation des coupes et des colorations histologiques. Merci à nos collègues de l'INRA Lusignan Christiane Minault, Dominique Denoue, Pascal Vernoux, pour la mise en place des essais et la réalisation des prélèvements. Ces travaux ont été financés par l'INRA et le groupe de travail ProMaïs- INRA "ZeaWall" qui rassemble les entreprises semencières maïs implantées en France (Advanta, Caussade Semences, Limagrain Europe, MaïsAdour, Monsanto SAS, Pioneer Génétique, Pau Euralis, R2n RAGT Semences, SDME KWS France, Syngenta seeds).

Chapitre VI - Approche histologique

Annexe 1 - Mesures de surfaces rouges et de surface bleues pour les lignées F271 et F288 aux différents stades de développement. Le périmètre (perim.) de chaque coupe est également précisé. (Bas - région un centimètre du bas de l'entrenœud; M Bas - milieu bas de l'entrenœud; M ENE - milieu de l'entrenœud; M Haut - milieu haut de l'entrenœud).

Stade	région ENE	Lignée	Bloc	Plante	Perim.	Aire totale mm2	Aire rouge mm2	Aire bleue mm2	Aire rouge %	Aire bleue %	Ar +Ab	Ar/(Ar+Ab)	Ab/(Ar+Ab)	moy %Ar	moy %Ab	moy Ar/(Ar+Ab)	moy Ab/(Ar+Ab)
Stade 1	bas	F288	1	A	62,13	186,39	89,95	66,40	48,26	35,62	156,35	0,58	0,42	40,67	40,56	0,50	0,50
Stade 1	bas	F288	1	B	66,20	187,11	68,19	81,15	36,45	43,37	149,35	0,46	0,54				
Stade 1	bas	F288	2	C	72,81	249,39	96,79	108,33	38,81	43,44	205,13	0,47	0,53				
Stade 1	bas	F288	2	D	70,93	203,72	79,76	81,12	39,15	39,82	160,87	0,50	0,50				
Stade 1	bas	F271	1	A	49,92	144,56	31,44	94,27	21,75	65,21	125,71	0,25	0,75	28,87	53,56	0,36	0,64
Stade 1	bas	F271	1	B	54,97	169,75	59,86	71,66	35,26	42,22	131,52	0,46	0,54				
Stade 1	bas	F271	2	C	70,70	249,68	37,67	182,22	15,09	72,98	219,89	0,17	0,83				
Stade 1	bas	F271	2	D	68,88	243,86	105,80	82,53	43,39	33,84	188,33	0,56	0,44				
Stade 1	mbas	F288	1	A	57,52	169,43	113,16	30,18	66,79	17,81	143,34	0,79	0,21	66,54	18,22	0,78	0,22
Stade 1	mbas	F288	1	B	60,60	173,48	99,20	44,90	57,18	25,88	144,10	0,69	0,31				
Stade 1	mbas	F288	2	C	70,22	231,95	149,88	40,18	64,62	17,32	190,06	0,79	0,21				
Stade 1	mbas	F288	2	D	60,42	179,92	139,58	21,34	77,58	11,86	160,92	0,87	0,13				
Stade 1	mbas	F271	1	A	51,23	141,05	78,67	36,94	55,78	26,19	115,62	0,68	0,32	46,45	27,11	0,65	0,35
Stade 1	mbas	F271	1	B	61,25	175,47	34,56	101,44	19,69	57,81	136,00	0,25	0,75				
Stade 1	mbas	F271	2	C	61,03	217,36	118,06	8,91	54,31	4,10	126,97	0,93	0,07				
Stade 1	mbas	F271	2	D	60,58	218,70	122,47	44,48	56,00	20,34	166,95	0,73	0,27				
Stade 1	mENE	F288	1	A	55,24	165,85	123,79	21,42	74,64	12,91	145,21	0,85	0,15	73,31	13,44	0,85	0,15
Stade 1	mENE	F288	1	B	54,86	162,54	119,94	22,82	73,79	14,04	142,76	0,84	0,16				
Stade 1	mENE	F288	2	C	63,95	213,21	152,23	29,85	71,40	14,00	182,08	0,84	0,16				
Stade 1	mENE	F288	2	D	60,76	180,48	132,52	23,14	73,43	12,82	155,66	0,85	0,15				
Stade 1	mENE	F271	1	A	54,10	138,74	102,61	20,09	73,95	14,48	122,70	0,84	0,16	66,14	15,51	0,81	0,19
Stade 1	mENE	F271	1	B	50,67	157,60	100,19	31,32	63,57	19,87	131,51	0,76	0,24				
Stade 1	mENE	F271	2	C	57,68	192,33	110,13	29,50	57,26	15,34	139,63	0,79	0,21				
Stade 1	mENE	F271	2	D	58,47	204,95	142,96	25,31	69,75	12,35	168,27	0,85	0,15				
Stade 1	mhaut	F288	1	A	56,02	182,57	141,24	19,65	77,36	10,76	160,89	0,88	0,12	77,27	10,78	0,88	0,12
Stade 1	mhaut	F288	1	B	55,90	169,78	129,57	19,93	76,32	11,74	149,50	0,87	0,13				
Stade 1	mhaut	F288	2	C	64,64	213,35	165,56	20,85	77,60	9,77	186,42	0,89	0,11				
Stade 1	mhaut	F288	2	D	57,44	192,03	149,36	20,85	77,78	10,86	170,21	0,88	0,12				
Stade 1	mhaut	F271	1	A	47,52	137,33	84,31	20,72	61,39	15,09	105,03	0,80	0,20	67,65	14,55	0,82	0,18
Stade 1	mhaut	F271	1	B	53,65	168,53	117,14	28,76	69,51	17,07	145,90	0,80	0,20				
Stade 1	mhaut	F271	2	C	53,75	188,43	117,95	29,48	62,59	15,65	147,43	0,80	0,20				
Stade 1	mhaut	F271	2	D	57,39	207,38	159,91	21,55	77,11	10,39	181,46	0,88	0,12				
Stade 2	bas	F288	1	A	74,37	229,72	116,80	80,22	50,84	34,92	197,02	0,59	0,41	50,45	32,14	0,61	0,39
Stade 2	bas	F288	1	B	83,88	231,46	105,45	90,58	45,56	39,13	196,03	0,54	0,46				
Stade 2	bas	F288	2	C	69,69	219,33	115,66	59,69	52,73	27,22	175,35	0,66	0,34				
Stade 2	bas	F288	2	D	68,58	211,44	111,37	57,70	52,67	27,29	169,07	0,66	0,34				
Stade 2	bas	F271	1	A	64,59	163,62	87,17	62,26	53,28	38,05	149,43	0,58	0,42	54,15	35,53	0,60	0,40
Stade 2	bas	F271	1	B	56,58	168,88	98,45	49,65	58,29	29,40	148,10	0,66	0,34				
Stade 2	bas	F271	2	C	65,03	253,14	146,20	85,69	57,75	33,85	231,88	0,63	0,37				
Stade 2	bas	F271	2	D	60,21	183,91	86,92	75,08	47,26	40,83	162,00	0,54	0,46				
Stade 2	mbas	F288	1	A	66,09	206,04	140,79	30,58	68,33	14,84	171,37	0,82	0,18	70,98	14,49	0,83	0,17
Stade 2	mbas	F288	1	B	64,18	200,87	138,35	30,57	68,87	15,22	168,92	0,82	0,18				
Stade 2	mbas	F288	2	C	66,00	187,01	124,84	34,03	66,75	18,20	158,87	0,79	0,21				
Stade 2	mbas	F288	2	D	70,48	278,45	222,64	26,99	79,96	9,69	249,63	0,89	0,11				
Stade 2	mbas	F271	1	A	49,64	134,30	86,33	18,03	64,28	13,43	104,36	0,83	0,17	63,37	17,13	0,79	0,21
Stade 2	mbas	F271	1	B	49,39	132,23	91,51	15,54	69,20	11,75	107,05	0,85	0,15				
Stade 2	mbas	F271	2	C	60,79	198,32	124,59	36,52	62,82	18,41	161,10	0,77	0,23				
Stade 2	mbas	F271	2	D	56,99	168,68	96,41	42,04	57,15	24,92	138,44	0,70	0,30				
Stade 2	mENE	F288	1	A	58,99	188,89	145,24	22,02	76,89	11,66	167,25	0,87	0,13	75,94	11,48	0,87	0,13
Stade 2	mENE	F288	1	B	59,22	186,85	140,38	22,37	75,13	11,97	162,75	0,86	0,14				
Stade 2	mENE	F288	2	C	57,45	177,21	133,39	21,27	75,27	12,00	154,66	0,86	0,14				
Stade 2	mENE	F288	2	D	71,19	279,02	213,39	28,69	76,48	10,28	242,08	0,88	0,12				
Stade 2	mENE	F271	1	A	48,02	118,34	89,52	11,86	75,65	10,03	101,39	0,88	0,12	70,60	12,29	0,85	0,15
Stade 2	mENE	F271	1	B	48,17	130,14	89,67	13,97	68,90	10,74	103,64	0,87	0,13				
Stade 2	mENE	F271	2	C	54,54	188,14	122,84	27,91	65,29	14,83	150,75	0,81	0,19				
Stade 2	mENE	F271	2	D	48,33	148,63	107,84	20,18	72,56	13,57	128,02	0,84	0,16				
Stade 2	mhaut	F288	1	A	61,07	201,16	157,64	22,94	78,37	11,40	180,58	0,87	0,13	76,72	10,81	0,88	0,12
Stade 2	mhaut	F288	1	B	58,64	197,66	156,12	18,28	78,99	9,25	174,40	0,90	0,10				
Stade 2	mhaut	F288	2	C	57,77	186,65	141,80	21,24	75,97	11,38	163,04	0,87	0,13				
Stade 2	mhaut	F288	2	D	69,37	276,11	203,12	30,94	73,57	11,20	234,06	0,87	0,13				
Stade 2	mhaut	F271	1	A	52,19	128,42	91,04	12,30	70,89	9,58	103,34	0,88	0,12	70,70	11,97	0,85	0,15
Stade 2	mhaut	F271	1	B	47,01	134,16	106,27	11,81	79,21	8,80	118,07	0,90	0,10				
Stade 2	mhaut	F271	2	C	54,50	187,88	126,51	23,14	67,34	12,31	149,65	0,85	0,15				
Stade 2	mhaut	F271	2	D	49,43	149,92	98,00	25,75	65,37	17,17	123,75	0,79	0,21				
Stade 3	bas	F288	1	A	67,26	220,98	142,69	45,12	64,57	20,42	187,81	0,76	0,24	67,87	18,21	0,79	0,21
Stade 3	bas	F288	1	B	54,40	138,31	98,22	25,25	71,01	18,25	123,46	0,80	0,20				
Stade 3	bas	F288	2	C	64,79	179,53	121,93	30,75	67,92	17,13	152,68	0,80	0,20				
Stade 3	bas	F288	2	D	62,92	172,53	117,32	29,40	68,00	17,04	146,72	0,80	0,20				
Stade 3	bas	F271	1	A	57,56	194,93	142,71	37,01	73,21	18,99	179,72	0,79	0,21	65,93	24,60	0,73	0,27
Stade 3	bas	F271	1	B	50,20	137,33	92,10	25,81	67,07	18,80	117,92	0,78	0,22				

Stade	région ENE	Lignée	Bloc	Plante	Perim.	Aire totale mm2	Aire rouge mm2	Aire bleue mm2	Aire rouge %	Aire bleue %	Ar +Ab	Ar/(Ar+Ab)	Ab/(Ar+Ab)	moy %Ar	moy %Ab	moy Ar/(Ar+Ab)	moy Ab/(Ar+Ab)
Stade 3	bas	F271	2	C	59,43	204,80	151,32	36,33	73,89	17,74	187,66	0,81	0,19				
Stade 3	bas	F271	2	D	57,15	183,81	91,07	78,84	49,55	42,89	169,91	0,54	0,46				
Stade 3	mbas	F288	1	A	62,46	194,41	127,28	33,26	65,47	17,11	160,55	0,79	0,21	71,59	13,98	0,84	0,16
Stade 3	mbas	F288	1	B	49,05	121,57	98,48	9,79	81,00	8,05	108,26	0,91	0,09				
Stade 3	mbas	F288	2	C	61,92	173,71	128,87	21,58	74,19	12,43	150,45	0,86	0,14				
Stade 3	mbas	F288	2	D	58,84	184,03	120,88	33,78	65,68	18,35	154,66	0,78	0,22				
Stade 3	mbas	F271	1	A	52,36	153,04	120,67	15,36	78,85	10,03	136,02	0,89	0,11	74,94	11,34	0,87	0,13
Stade 3	mbas	F271	1	B	46,27	119,53	95,07	12,73	79,54	10,65	107,80	0,88	0,12				
Stade 3	mbas	F271	2	C	58,13	148,23	110,44	14,31	74,50	9,65	124,74	0,89	0,11				
Stade 3	mbas	F271	2	D	57,80	146,48	97,96	21,99	66,87	15,01	119,95	0,82	0,18				
Stade 3	mENE	F288	1	A	59,37	200,15	145,89	27,07	72,89	13,53	172,96	0,84	0,16	75,64	11,19	0,87	0,13
Stade 3	mENE	F288	1	B	47,64	125,87	99,61	10,33	79,14	8,21	109,94	0,91	0,09				
Stade 3	mENE	F288	2	C	55,56	155,13	119,42	14,76	76,98	9,51	134,18	0,89	0,11				
Stade 3	mENE	F288	2	D	56,80	174,16	128,11	23,56	73,56	13,53	151,67	0,84	0,16				
Stade 3	mENE	F271	1	A	52,30	150,09	115,03	10,33	76,64	6,88	125,36	0,92	0,08	75,02	9,48	0,89	0,11
Stade 3	mENE	F271	1	B	40,18	95,66	79,60	7,67	83,21	8,02	87,27	0,91	0,09				
Stade 3	mENE	F271	2	C	49,36	147,95	112,77	11,52	76,22	7,79	124,30	0,91	0,09				
Stade 3	mENE	F271	2	D	47,45	141,11	90,33	21,51	64,02	15,24	111,84	0,81	0,19				
Stade 3	mhaut	F288	1	A	70,77	210,34	151,86	26,26	72,20	12,48	178,12	0,85	0,15	75,30	10,53	0,88	0,12
Stade 3	mhaut	F288	1	B	48,91	142,23	114,23	11,66	80,31	8,20	125,89	0,91	0,09				
Stade 3	mhaut	F288	2	C	55,79	165,08	122,90	17,55	74,45	10,63	140,45	0,88	0,12				
Stade 3	mhaut	F288	2	D	54,80	159,47	118,41	17,21	74,25	10,79	135,62	0,87	0,13				
Stade 3	mhaut	F271	1	A	50,17	151,54	107,83	15,61	71,16	10,30	123,44	0,87	0,13	74,58	9,88	0,88	0,12
Stade 3	mhaut	F271	1	B	43,04	104,75	82,24	9,33	78,51	8,91	91,58	0,90	0,10				
Stade 3	mhaut	F271	2	C	48,99	147,41	118,31	10,03	80,26	6,81	128,35	0,92	0,08				
Stade 3	mhaut	F271	2	D	48,02	147,77	101,04	19,97	68,37	13,52	121,01	0,83	0,17				

CONCLUSION GÉNÉRALE ET PERSPECTIVES

Le cadre de ce travail était de s'appuyer sur un cluster de QTLs à effets forts, liés à la dégradabilité et à la structure biochimique des parois secondaires du maïs, pour mettre en évidence les déterminants génétiques majeurs impliqués dans la formation et la mise en place des parois lignifiées, en prenant en compte les liaisons entre les composants biochimiques de ces parois qui sont une caractéristique spécifique des graminées. Cette démarche a été mise en place avec un double objectif. Le premier, plus fondamental, visait à contribuer à la connaissance des gènes impliqués chez les graminées dans la mise en place des parois secondaires lignifiées, et corrélativement de ceux impliqués dans la biosynthèse des composés phénoliques, en se focalisant sur le modèle maïs. Le second, plus appliqué, visait à donner aux professionnels des semences de maïs et de graminées des outils moléculaires pour le typage des ressources génétiques et la sélection assistée par marqueurs en vue de l'amélioration de la valeur énergétique de ces espèces à la fois pour l'alimentation des ruminants et la production de biocarburants de seconde génération.

Le choix de réaliser ce travail sur le cluster de QTLs du bin 6.06 des RILs F288 x F271 a été fait en raison des valeurs de LOD et de R^2 très élevées obtenues à cette position, tout en relativisant cette situation par le nombre un peu faible de RILs étudiées. Les effets additifs des QTLs étaient également élevés, atteignant ainsi 2,6 points de dégradabilité des parois en IVNDFD (Roussel et al, 2002). Cette base génétique permettait aussi d'élargir l'approche à l'impact de la répartition des lignines au sein des différents tissus de la tige, d'autres travaux ayant en effet montré que les deux lignées avaient des structures histologiques de tiges contrastées (Méchin et al, 2000; 2005). De plus, des analyses de transcriptomique comparée entre RILs présentant une bonne dégradabilité de parois et le parent F271 avaient été mises en place. Enfin, un investissement sur ces lignées permettait de compléter les données déjà disponibles en mettant en place des essais des RILs en valeur propre avec retrait des épis à la récolte. Les équations NIRS correspondantes permettant d'élargir la gamme des caractéristiques pariétales étudiées (Riboulet et al, 2008b) n'étaient en effet pas disponibles lors des premières années d'essais. Des données sur les structures des lignines (teneurs en chacun des monomères H, G, et S libérés par une attaque au nitrobenzène alcalin), de même que des données sur les teneurs en acide *p*-coumarique acylant les unités S des lignines, et les teneurs en férulate et diférulates pontant les chaînes d'arabinoxylanes et les arabinoxylanes aux unités G des lignines devenaient alors accessibles.

L'étude s'est avérée plus complexe que ce qui avait été initialement imaginé, et ce à plusieurs niveaux, même si la mise en évidence du polymorphisme génétique responsable d'un caractère quantitatif est toujours un challenge dans la durée limitée d'un travail de thèse. Il n'a pas été possible d'aller aussi loin que ce qui avait été espéré dans l'identification des déterminants génétiques de la variabilité de la dégradabilité des parois du maïs. Toutefois, ce travail a ouvert des perspectives dont certaines n'étaient pas attendues et il a modifié significativement des convictions, et sans doute des *a*

priori, sur les déterminants génétiques qui sous-tendent la variabilité génétique des caractéristiques biochimiques d'une paroi secondaire de graminées, et en conséquence ceux qui déterminent la variabilité de son niveau de dégradabilité. Ce travail a aussi modifié le regard à avoir sur les résultats des analyses QTL, et à prendre en compte de façon plus constante les limites de ces démarches.

De façon imprévue, mais comme pouvait le laisser envisager des analyses juste préliminaires, il a été assez rapidement confirmé que les positions estimées des QTLs à effets majeurs du bin 6.06, ainsi que l'essentiel des intervalles supports correspondants, se situaient dans une région apparemment monomorphe entre les deux parents. La densification ciblée de la carte, ainsi que les nouvelles expérimentations réalisées au champ, ont permis de montrer que ces QTLs, que l'on pouvait alors qualifier de fantômes, correspondaient en fait à des QTLs se répartissant à la fois sur la partie inférieure du bin 6.05 et sur la partie supérieure du bin 6.07, ainsi que, éventuellement, sur la partie inférieure du bin 6.06. Ils correspondaient ainsi à deux ou trois positions distinctes. Le marquage plus dense a aussi permis de conforter les positions des autres QTLs, et d'en enrichir le nombre. Une carte dense apparaît donc comme un prérequis à l'estimation de l'ensemble des QTLs, avec la nécessité d'un nombre de marqueurs plus important que ce qui est habituellement réalisé, ainsi compris entre 150 et 200. Mais surtout, seul un marquage très dense semble permettre la détection de QTLs proches. Cette démarche ne peut être envisagée qu'*a posteriori*, ou bien par l'utilisation *a priori* de techniques de génotypage utilisant des puces SNP haute densité. De plus, pour la recherche des gènes candidats à partir d'un génome de référence, le nombre de marqueurs retenus devra assurer un marquage homogène à la fois par rapport aux distances physiques et génétiques, ce qui n'est toujours facile en raison des grandes variations du rapport entre distance physique et distance génétique, surtout autour des centromères.

La recherche des gènes candidats sous les QTLs a conduit à mettre en évidence une situation également plus complexe et sans doute plus longue à expliciter que prévu. Il y a clairement une forte variabilité phénotypique dans la descendance RIL F288 x F271, avec des QTLs à effets forts ou très forts, et la liste proposée et relativement longue de gènes impliqués ou putativement impliqués dans la mise en place des parois secondaires, conduit à un nombre significatif de colocalisations. En revanche, pour chaque position de QTL de dégradabilité des parois, il y a une difficulté certaine à trouver un candidat qui serait qualifié de "très probable", en particulier sous les intervalles support des QTLs des bins 6.05/06/07. Pour les QTLs majeurs de la base du bin 6.05 (position 151.8 Mbp), il pourrait aussi de nouveau exister une situation de QTLs fantômes, correspondant à des gènes ayant une liaison fonctionnelle entre eux, et situés en amont et en aval de la position estimée des QTLs après densification de la carte. Ceci pourrait ainsi être le cas des gènes MYB (type EgMYB1/Hv5) et Ovate, situés respectivement à 150.7 et 153.8 Mbp, quand l'intervalle support de la position des QTLs est de 151.1 à 153.3 Mbp. Un autre fait à prendre en compte est une possible sous-estimation de la taille des

intervalles support, situation qui doit être particulièrement prise en compte en CIM, et en particulier avec un faible nombre de RILs (Melchinger et al, 1998). Dans ce cas, le gène MYB ou le gène *Ovate* pourrait être le candidat recherché.

Cette situation de complexité se retrouve lors de la recherche des déterminants génétiques après la synthèse des données QTLs sur six familles RILs, avec de plus, à ce niveau, certains QTLs qui demeurent sans aucun candidat. Au vu de la liste de gènes candidats établie et des colocalisations mises en évidence, l'existence de plusieurs gènes comme facteurs explicatifs d'un QTL ou d'un cluster de QTLs donné apparaît, pour une part non négligeable des positions, au moins aussi probable que celle d'un gène majeur unique. En considérant l'ensemble des colocalisations, le rôle dans les variations de dégradabilité de paroi des gènes de type MYB, en tant que facteur unique à une position, ou bien associés entre eux, associés à des gènes de type NAC, ou associés à des gènes d'autres familles (COV-like, zinc finger, ...), correspond à la situation la plus fréquemment probable. On ne peut toutefois pas exclure un biais à ce niveau, puisque les gènes de type MYB et NAC sont aussi ceux qui ont été le plus étudiés en tant que régulateurs de la mise en place des parois lignifiées. Si la plupart des gènes de la voie de biosynthèse des monolignols colocalisent aussi avec des QTLs, ils n'en sont très probablement pas les déterminants majeurs dans la plupart des cas, d'autant qu'ils se situent assez souvent aux extrémités des intervalles support. Leurs variations éventuelles de niveau de transcrits correspondraient plutôt à l'action de facteurs de régulation situés en amont. Ceci étant, il est clair que des gènes reliés à des groupes fonctionnels sans rôle décrit dans la mise en place des parois lignifiées, ainsi que de nombreux gènes de fonction inconnue, se trouvent sous les QTLs, et peuvent en être les (co)-déterminants. Par ailleurs, le fait qu'il y ait des QTLs à effets (très) forts peut paraître contradictoire avec une hypothèse de cluster de gènes sous les positions des QTLs, dans la mesure où des recombinaisons dans les descendances RILs auraient dû conduire à la mise en évidence de plusieurs QTLs consécutifs, ceci à condition que le marquage soit suffisamment dense. Ceci étant, des recombinaisons dans un espace de 2 ou 3 cM dans une descendance de 150 à 250 individus ont aussi une probabilité faible d'être mises en évidence en terme de phénotypes et donc de QTLs. Les résultats obtenus conduisent à conclure à de nécessaires démarches de validation des gènes candidats, qui incluent la création des descendances portant des événements de recombinaison dans les régions étudiées.

L'exploration de la variabilité au niveau des QTLs par un séquençage ciblé des BACs correspondants chez les lignées parentales avec des techniques à haut débit s'est avéré difficile en raison de l'importance des séquences répétées de type rétrotransposons, ainsi que de la grande variabilité de séquence entre les génotypes qui ne permet pas complètement de s'appuyer sur le génome de B73. Cette approche permet néanmoins de comparer les séquences des gènes entre les parents des RILs, mais plus difficilement de mettre en évidence un possible déterminant de la

variabilité hors des gènes. Si cette démarche permet également de mettre en évidence des variations de type présence / absence ou gènes dupliqués, les difficultés de contigage pourront demander en cas de doute une validation des résultats par des démarches de PCR classiques. Cette démarche de séquençage ciblé des BACs suppose aussi la fiabilité des positions des QTLs, et celle de la taille des intervalles support qui sont plutôt sous-estimés en CIM.

La démarche de transcriptomique a été limitée par l'utilisation de la puce 17.000 sondes, qui était celle alors disponible. Cela a toutefois permis de cibler comme candidats un nombre réduit de gènes différentiellement exprimés et situés sous les QTLs des bins 6.05/06, 1.01/02 et 3.05. Mais à nouveau, aucun de ces gènes candidats n'a pu être considéré comme étant "très probablement" le gène recherché. En revanche, trois gènes MYB, un gène NAC, et un gène zinc finger C3HC4, tous potentiellement impliqués dans la mise en place des parois lignifiées et situés dans le génome ailleurs que dans les bins 6.05/06, étaient différentiellement exprimés. Tous ces gènes colocalisent avec des QTLs de dégradabilité des parois dans d'autres familles RILs (chapitre 5). Ces résultats conforteraient leur implication dans la mise en place des parois, et (peut-être) en conséquence l'existence d'un régulateur situé en amont, qui pourrait correspondre au gène candidat recherché dans le bin 6.05. A ce titre, il faut aussi considérer l'implication possible de gènes codant pour des protéines de fonction non connues ou pour des protéines dont le rôle dans la mise en place (ou la régulation de la mise en place) des parois lignifiées n'est pas encore établie (gènes EXPA5, FK506 binding protein situées dans le bin 6.05/06).

La mise en évidence du ou des déterminant(s) des QTLs à forts effets observés autour des bins 6.06, 4.09, et 3.05 des RILs F288 x F271, en tant que modèle le plus avancé sur cette thématique, semble maintenant une priorité pour avancer dans la compréhension des déterminants de la variabilité de dégradabilité des parois du maïs et des graminées. Il se dégage à ce niveau plusieurs questionnements. Il faudra confirmer si ces QTLs majeurs correspondent en fait, ou non, à un ensemble de sous-QTLs très proches. En parallèle, il faudra savoir si ces QTLs ou sous-QTLs correspondent chacun à un gène ou à un petit groupe de gènes. Il faudra alors découvrir si les déterminants de la variabilité correspondent à une variabilité allélique des gènes, et/ou correspondent plutôt à des variations dans la régulation liés à des insertions de transposons, voire à des mécanismes de type CNV ou PAV fréquents dans le génome du maïs (Springer al, 2009; Beló et al, 2010; Lai et al, 2010; Swanson-Wagner et al, 2010). Il faudra aussi mesurer l'implication possible de gènes de fonctions inconnues ou non encore reliées à la mise en place des parois, ainsi que celle, possible, de miRNA.

Les démarches à conduire, pour élucider les QTLs des RILs F288 x F271, et plus largement pour comprendre les différents systèmes génétiques impliqués dans la mise en place des parois secondaires du maïs et des graminées, pourraient s'appuyer sur différentes stratégies.

i) Sans qu'il soit exclu de passer de nouveaux marqueurs pour affiner la position de certains QTLs et réduire la taille des intervalles supports, le séquençage ciblé des BACs situés dans les intervalles support des QTLs des bins 6.06, 4.09, et 3.05 devra être poursuivi. A ce titre, il faudra augmenter la couverture des banques BAC de F288 et F271, sans doute en utilisant d'autres enzymes de restriction afin de couvrir toutes les régions recherchées. Il sera également indispensable de réaliser du séquençage de type "paired-end" pour faciliter le contigage des "reads" obtenus.

ii) La réalisation de descendance hautement recombinantes entre RILs (RIL5 x RIL35) et la création de NILs dans le parent F271 (avec la RIL122) devraient permettre de valider et d'affiner la position du ou des QTL(s) dans la région du bin 6.05, mais aussi au niveau des bins 4.08/09, et pour partie 3.05/06. Des effectifs (très) importants de descendance seront dans ce cas nécessaires pour obtenir les recombinants dans les régions d'intérêt. Il faudra cependant prévoir simultanément la mise en essai d'un nombre suffisant de répétitions (et/ou d'environnements) pour mettre en évidence des différences phénotypiques qui seront probablement de plus faible niveau entre plantes ayant recombiné pour des déterminants (très) proches.

iii) La réalisation d'une nouvelle étude d'expression comparée avec une puce comprenant des sondes de l'ensemble des gènes du maïs, et des lignées choisies en prenant en compte l'ensemble des marqueurs maintenant disponibles, ainsi que les NILs ou des lignées HRILs issues des démarches de clonage positionnel, devrait permettre à la fois de mieux décrypter les différentiels d'expression de gènes situés sous les QTLs, mais aussi de mettre en évidence l'ensemble des gènes différentiellement exprimés et donc de cibler les voies régulés par les facteurs majeurs situés sous les QTLs.

iv) Les nouvelles techniques de séquençage haut débit permettent de quantifier le niveau de transcrits pour l'ensemble des gènes exprimés dans un tissu par RNA sequencing (RNA-seq). Dès lors, le transcriptome n'est plus contraint par les sondes présentes sur la puce utilisée. La quantification des niveaux de transcrits entre RILs choisies, et entre RILs et le parent F271, entre descendance HRILs ou NILs, réalisée au niveau de l'ensemble du génome par RNA-seq, permettra de déterminer l'impact d'un QTL sur le profil transcriptomique global. Cette image du différentiel d'expression permettra également de mieux appréhender la fonction du ou des gène(s) candidat(s) recherché(s).

v) Même si les résultats d'histologie comparative n'ont pu être confirmés dans le cadre de ce travail, des variations d'intensité de la lignification entre les différents tissus des tiges semblent exister

entre F271 et F288. Par ailleurs, il n'est pas possible de diminuer trop fortement la teneur en lignines et/ou les qualités mécaniques de tous les tissus pour maintenir la tolérance des vaisseaux conducteurs aux conditions de déficit hydrique, la résistance des plantes aux ravageurs, et pour partie, la résistance à verse en végétation et à la verse à maturité. Dans ce contexte, pour améliorer la dégradabilité des parois, il pourra être recherché des RILs ayant des profils histologiques typés, en particulier sur des différences de lignification du parenchyme, des vaisseaux conducteurs, et des tissus périphériques des tiges. L'utilisation de la microdissection laser (Laser Capture Microdissection, LCM), qui permet d'isoler les différents tissus, pourrait alors permettre de découvrir quels sont les gènes différentiellement exprimés dans chacun des tissus. Sous réserve de mise en évidence de déterminants génétiques "simples", ceci permettrait ensuite de définir des cibles fines de sélection assistée par marqueurs pour une meilleure dégradabilité des parois, sans perte de valeur agronomique.

vi) Les mécanismes génétiques conduisant à la formation des pontages entre lignines et arabinoxyanes, et entre chaînes d'arabinoxyanes, restent très grandement inconnus, alors que cette réticulation des parois peut avoir un effet majeur sur leur dégradabilité, ainsi que sur les qualités mécaniques des tissus régulant l'ingestibilité par un ruminant. Ces pontages sont réalisés par des férulates (en particulier entre unités G des lignines et arabinoxyanes) et diférulates (en particulier entre chaînes d'arabinoxyanes). Des acyltransférases de type PF02458 sont quasi-certainement impliquées dans la mise en place des liaisons entre l'acide férulique et les arabinoxyanes, mais nombre de QTLs de férulate ou diférulates ne colocalisent pas avec des gènes codant des enzymes de cette famille. Il n'est d'ailleurs pas complètement clair si il y a, ou non, une formation spécifique du feruloyl-CoA impliqué dans ces pontages. Des glycosyltransférases de la famille des GT47 sont également probablement impliquées, ces gènes codant pour des enzymes de type xylane α -1,2- et α -1,3-arabinosyle-transférases, permettant le transfert d'un résidu arabinosyle sur une chaîne de xylane. L'absence ou la réduction de cette activité enzymatique compromettrait le greffage de l'arabinose au xylane chaînes, et par conséquent l'intensité des possibilités de réticulations par les férulates. Il faut enfin imaginer un rôle possible d'un UDP-arabinose dans la formation des arabinoxyanes féruloylés. Les expérimentations sur plantes sans épis ont permis la mise en évidence de QTLs d'étherFA et de diFA qui pourraient servir de support à une recherche ciblée des candidats correspondants, et par là même à une meilleure connaissance de cette voie encore obscure de synthèse et de couplage de l'acide férulique. La colocalisation de QTLs de férulate avec des QTLs de teneurs en lignines ou de structure des lignines pose de nouveau la question des déterminismes de ces caractères résultant soit de l'activité d'un gène majeur en amont, soit du fonctionnement indépendant ou co-régulé d'un cluster de gènes impliqués dans la formation des parois secondaires. Concernant les acides *p*-hydroxycinnamiques, les déterminants de l'acylation des unités S par l'acide *p*-coumarique sont également encore inconnus, alors que des QTLs de pCA et de S-pCA ont été clairement mis en évidence.

vii) Une étape clé sera la validation des candidats mis en évidence, qui seront sans doute relativement nombreux, et qui devront probablement être testés simultanément sous forme de couple de gènes co-responsables du QTL, voire peut-être de triplet. L'accès à une banque de mutants, comme celle des descendances de transposon tagging qui ont montré leur intérêt sur des gènes liés aux parois (CCR, CAD2, ...), donnera des réponses plus rapidement que la création de plantes de maïs dérégulées par transgénèse. Toutefois, il n'est pas obligatoire que les démarches de première validation doivent se faire directement sur maïs. Des mutants ou des transformants d'*Arabidopsis*, comme cela a été fait pour le ZmMYB42 (Sonbol et al, 2009), donnent des réponses tout à fait déterminantes, au moins pour les gènes ou les mécanismes non spécifiques des graminées. Les deux plantes modèles de la famille des graminées, *Brachypodium* (C3) et sétaire (C4), n'ont été que peu ou pas utilisées pour des gènes liés aux parois, mais les outils de transformation sont disponibles et les banques de mutants en cours de construction ou de caractérisation. Des résultats sur sorgho, qui possède une structure génétique plus simple que celle du maïs, pourraient également être valorisés, en particulier pour les approches de production de biocarburants.

viii) A plus long terme, les deux lignées parentales F288 et F271, et sans doute plus particulièrement la lignée F271, peuvent être utilisées pour la production de nouvelles familles recombinantes. Une descendance de 250 RILs issues du croisement entre F271 et Cm484, lignée dont le parenchyme semble, au moins dans certaines conditions de milieu, non lignifié au contraire de F271, sera ainsi disponible à la récolte des pépinières de 2012. Il serait également intéressant de disposer de descendances entre une lignée de type Iodent élite et une lignée de dégradabilité des parois élevée apparentée au groupe Minnesota13 comme W117, F7019 ou F7082. Enfin, il ne faudrait pas exclure les deux lignées à dégradabilité de parois élevée, issues d'écotypes et créées dans le cadre de ProMaïs, F7106 issues de Gelderland, et F7104, issues de matériel argentin mais qui est toutefois plus tardive (Barrière et al, 2010a). Ce type de démarche impliquant plusieurs descendances RIL connectées par un parent ou des parents communs permettrait de conforter significativement l'inventaire des loci majeurs impliqués dans la dégradabilité des parois, et en conséquence le choix optimisé des allèles à introduire dans les lignées élites.

La mise en évidence des gènes impliqués dans les variations de dégradation des parois de graminées devra aussi s'appuyer sur l'élucidation d'autres QTLs à effets forts mis en évidence dans d'autres familles RIL. A ce titre, il y a en particulier les QTLs de dégradabilité des parois (IVNDFD en valeur propre) issus d'autres études (Méchin et al, 2001; Barrière et al, 2008; Barrière et al, 2012; Génoplante, données non publiées). Il serait ainsi possible de tenter d'élucider les gènes correspondants aux QTLs des bins 1.07 et 10.02/03 des RILs F838 x F286 (R^2 de 22 et 23 %), ceux des bins 3.04 et 4.05 des RILs Rio x WM13 (R^2 de 18 et 16 %), ceux des bins 1.04 et 2.01 des RILs F7025 x F4 (R^2 de 16 et 17 %), et celui du bin 4.08 des RIL Io x F2 (R^2 de 20 %). Ce travail était pour

partie envisagé au départ de ce projet, en particulier pour deux QTLs des RILs F838 x F286. Mais l'élucidation simultanée de plusieurs QTLs dans différentes familles s'est vite avérée irréaliste dans la durée d'une thèse. L'utilisation d'outils de séquençage à haut débit, qui ont particulièrement vu leur essor au cours de cette thèse, devrait accélérer le processus de découverte des gènes candidats. Toutefois, cela ne sera vrai pour le maïs que si les séquences obtenues ont des longueurs et des qualités suffisantes pour permettre leur contigage, sachant que la complexité du déterminisme génétique des caractères liés à la paroi restera dans la plupart des cas un obstacle. Le couplage de l'ensemble des approches envisagées devrait permettre d'obtenir des indices convergents vers l'identification et la validation des gènes candidats les plus probables. Enfin, il est clair que la découverte des deux ou trois premiers gènes candidats donnera des indices sur les mécanismes impliqués dans les variations de dégradabilité des parois, et facilitera sans doute la recherche des candidats suivants.

Même si les expérimentations peuvent être rendues difficiles par les variations de précocité des lignées, les approches de génétique d'association peuvent aussi permettre d'identifier des régions, co-localisant (ou non) avec les QTLs déjà mis en évidence, plus fines que les intervalles support des QTLs, surtout ceux obtenus sur des familles de 100 ou 200 RILs. L'étude du polymorphisme au sein des panels de lignées utilisé en génétique d'association devrait ainsi permettre d'accélérer l'identification de gènes candidats. La diversité des fonds génétiques importante entre les lignées impliquées dans ce type d'expérimentation, et les différences phénotypiques parfois peu élevées entre ces lignées pour les caractères liés aux parois, contraignent aussi à faire ce type d'étude avec un nombre suffisant de répétitions et d'environnements.

Dans un autre domaine, il y a beaucoup à attendre des avancées en biochimie des parois pour progresser sur les cibles de sélection et la compréhension des mécanismes génétiques en jeu. Un travail important va devoir être développé à ce niveau là pour rendre accessible des équations de prédictions NIRS pour les caractères de biochimie fine ayant un effet sur la dégradabilité des parois. Il est indispensable de faire ainsi évoluer nos connaissances sur la structure physique des lignines en fonction de leur ratio en monomères constitutifs et en fonction du type de liaison entre ces monomères, avec la quantification de l'effet de la structure physique sur la dégradabilité. Les travaux récents montrent également que la présence d'acide *p*-coumarique sur les unités S des lignines semble jouer un rôle important sur la structure des parois et leur dégradabilité. Il faudra donc réussir à élucider l'influence de cette acylation des unités S par l'acide *p*-coumarique sur la conformation du polymère lignines. Là encore, les mécanismes et les gènes recrutés pour permettre l'accrochage de cet acide *p*-coumarique sur les unités S pourront être une cible majeure de l'amélioration de la dégradabilité des parois. Ce caractère pourrait être ciblé sans que la teneur en lignine ne soit pour autant trop réduite, permettant ainsi aux plantes de conserver des qualités agronomiques importantes. Une stratégie

contraire pourrait également être envisagée où l'on diminuerait assez fortement la teneur en lignine en privilégiant une plus forte acylation des unités S par l'acide *p*-coumarique pour conférer aux parois des qualités suffisantes au bon maintien de la plante. Seule la connaissance de ces différents mécanismes permettra de définir la situation de compromis au sein des parois pour en améliorer la dégradabilité tout en assurant la tenue et l'intégrité des plantes. Enfin, il faudra réussir à améliorer les méthodes biochimiques de libération du férulate et des diférulates au sein d'une matrice pariétale complexe. En effet, les méthodes actuelles ne donnent pas accès à l'intégralité des férulates, ce qui peut biaiser l'évaluation de son impact et ne pas permettre la mise en évidence de gènes clés. Une autre voie prometteuse et complémentaire à ce niveau réside dans l'utilisation d'anticorps permettant de cibler spécifiquement les férulates et diférulates au sein des parois sur coupes. Ces anticorps sont en cours de production et leur disponibilité va permettre d'accéder à une information à l'échelle tissulaire de la réticulation par les férulates au sein des parois. Sous réserve de la mise en évidence effective des déterminants génétiques correspondants, cela permettrait là aussi d'imaginer une sélection pour une régulation ciblant des tissus clés au sein des tiges.

Finalement, les travaux réalisés sur le modèle des RIL F288 x F271, au-delà des résultats obtenus en terme de gène candidats, donnent une illustration de la complexité des mécanismes génétiques qui peuvent sous-tendre un caractère quantitatif tel que la dégradabilité d'une paroi secondaire lignifiée. Ces travaux illustrent aussi la nécessité d'une compréhension des phénomènes en amont, avec une démarche de recherche fondamentale, pour déboucher en aval, dans un cadre industriel appliqué, sur des stratégies pertinentes et efficaces de sélection et de valorisation des ressources génétiques. L'intérêt d'une démarche multidisciplinaire est clairement apparu, avec un besoin de compétences en génétique quantitative, en génomique, et en biochimie. Un cadre et une méthodologie pour les travaux à venir sur la mise en évidence des déterminants génétiques de la variabilité génétique de la dégradabilité des parois de maïs et de graminées est maintenant disponible, avec des démarches qui devront être priorisées en fonction de l'évolution des technologies en séquençage, en bioanalyse, et également en analyse d'images.

REFERENCES BIBLIOGRAPHIQUES

- Alejandro S, Lee Y, Tohge T, Sudre D, Osorio S, Park J, Bovet L, Lee Y, Geldner N, Fernie AR, Martinoia E, 2012. AtABCG29 is a monolignol transporter involved in lignin biosynthesis. *Curr Biol* 22:1207-1212
- Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ, 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25:3389-3402
- Ambavaram MM, Krishnan A, Trijatmiko KR, Pereira A, 2011. Coordinated activation of cellulose and repression of lignin biosynthesis pathways in rice. *Plant Physiol* 155:916-931
- Ammiraju JS, Yu Y, Luo M, Kudrna D, Kim H, Goicoechea JL, Katayose Y, Matsumoto T, Wu J, Sasaki T, Wing RA, 2005. Random sheared fosmid library as a new genomic tool to accelerate complete finishing of rice (*Oryza sativa* spp. Nipponbare) genome sequence: sequencing of gap-specific fosmid clones uncovers new euchromatic portions of the genome. *Theor Appl Genet* 111(8):1596-607
- Andersen JR, Asp T, Lu YC, Kloiber-Maitz M, Ouzonova M, Lueberstedt T, 2009. Development and mapping of gene-tagged SNP markers in laccases of maize (*Zea mays* L.). *Plant Breeding* 128:423-425
- Andersen JR, Zein I, Wenzel G, Krutzfeldt B, Eder J, Ouzonova M, Lübberstedt T, 2007. High levels of linkage disequilibrium and associations with forage quality at a Phenylalanine Ammonia-Lyase locus in European maize (*Zea mays* L.) inbreds. *Theor Appl Genet* 114:307-319
- Andersson-Gunneras S, Mellerowicz EJ, Love J, Segerman B, Ohmiya Y, Coutinho PM, Nilsson P, Henrissat B, Moritz T, Sundberg B, 2006. Biosynthesis of cellulose-enriched tension wood in *Populus*: global analysis of transcripts and metabolites identifies biochemical and developmental regulators in secondary wall biosynthesis. *Plant J* 45:144-165
- Andrieu J, Aufrère J, 1996. Prévision à partir de différentes méthodes (physique, chimique et biologique) de la digestibilité et de la valeur énergétique de la plante de maïs à l'état frais. Colloque Maïs Ensilage, Nantes, 17-18 septembre 1996
- Andrieu J, Barrière Y, Demarquilly C, 1999. Digestibilité et valeur énergétique des ensilages de maïs : le point sur les méthodes de prévision au laboratoire. *INRA Prod Anim* 12:391-396
- Argillier O, Barrière Y, Dardenne P, Emile JC, Hébert Y, 1998. Genotypic variation for in vitro criteria and relationships with in vivo digestibility in forage maize hybrids. *Plant Breeding* 117:437-441
- Argillier O, Barrière Y, Hébert Y, 1995. Genetic variation and selection criterion for digestibility traits of forage maize. *Euphytica* 82:175-184
- Argillier O, Méchin V, Barrière Y, 2000. Inbred line evaluation and breeding for digestibility-related traits in forage maize. *Crop Sci* 40:1596-1600
- Ariyadasa R, Stein N, 2012. Advances in BAC-Based Physical Mapping and Map Integration Strategies in Plants. *Journal of Biomedicine and Biotechnology* 2012:184854
- Aronel V, Vergnolle C, Cantrel C, Kader JC, 2000. Lipid transfer proteins are encoded by a small multigene family in *Arabidopsis thaliana*. *Plant Sci* 157:1-12
- Arruda P, Gerhardt IR, 2010. Nucleic acid molecules encoding plant proteins in the C3HC4 family and methods for the alteration of plant cellulose and lignin content. US Patent App. 12/520, 266 US 2010/0077509 A1
- Arvidsson S, Kwasniewski M, Riaño-Pachón DM, Mueller-Roeber B, 2008. QuantPrime - a flexible tool for reliable high-throughput primer design for quantitative PCR. *BMC Bioinformatics* 9:465-480
- Aufrère J, Michalet-Doreau B, 1983. In vivo digestibility and prediction of digestibility of some by-products. In: EEC seminar, Melle Gontrode. 26-29 September. pp. 25-33

- Baima S, Possenti M, Matteucci A, Wisman E, Altamura MM, Ruberti I, Morelli G, 2001. The Arabidopsis ATHB-8 HD-zip protein acts as a differentiation-promoting transcription factor of the vascular meristems. *Plant Physiol* 126:643-655
- Bakalovic N, Passardi F, Ioannidis V, Cosio C, Penel C, Falquet L, Dunand C, 2006. PeroxiBase: a class III plant peroxidase database. *Phytochem* 67:534-539
- Barbazuk WB, Emrich SJ, Chen HD, Li L, Schnable PS, 2007. SNP discovery via 454 transcriptome sequencing. *Plant J* 51:910-918
- Barrett T, Troup DB, Wilhite SE, Ledoux P, Rudnev D, Evangelista C, Kim IF, Soboleva A, Tomashevsky M, Edgar R, 2007. NCBI GEO: mining tens of millions of expression profiles--database and tools update. *Nucleic Acids Res* 35:D760-765
- Barrière Y, Argillier O, 1997. In vivo silage feeding value of early maize hybrids released in France between 1958 and 1994. *Euphytica* 99:175-182
- Barrière Y, Charcosset A, Denoue D, Madur D, Bauland C, Laborde J, 2010a. Genetic variation for lignin content and cell wall digestibility in early maize lines derived from ancient landraces *Maydica* 55:65-74
- Barrière Y, Emile JC, Surault F, 2003a. Genetic variation of maize silage ingestibility in dairy cattle. *Animal Research* 52:489-500
- Barrière Y, Emile JC, Traineau R, Hébert Y, 1995. Genetic variation in the feeding efficiency of maize genotypes evaluated from experiments with dairy cows. *Plant Breed* 114:144-148
- Barrière Y, Emile JC, Traineau R, Surault F, Briand M, Gallais A, 2004a. Genetic variation for organic matter and cell wall digestibility in silage maize. Lessons from a 34-year long experiment with sheep in digestibility crates. *Maydica* 49:115-126
- Barrière Y, Gibelin C, Argillier O, Méchin V, 2001. Genetic analysis and QTL mapping in forage maize based on recombinant inbred lines descended from the cross between F288 and F271. I - Yield, earliness, starch and crude protein content. *Maydica* 46:253-266
- Barrière Y, Goncalves G, Emile J, Lefevre B, 2004b. Higher intake of DK265 corn silage by dairy cattle. *Journal of Dairy Science* 87:1439-1445
- Barrière Y, Guillet C, Goffner D, Pichon M, 2003b. Genetic variation and breeding strategies for improved cell wall digestibility in annual forage crops. A review. *Animal Research* 52:193-228
- Barrière Y, Méchin V, Denoue D, Bauland C, Laborde J, 2010b. QTL for yield, earliness and cell wall digestibility traits in topcross experiments of F838 x F286 RIL progenies. *Crop Sci* 50:1761-1772
- Barrière Y, Méchin V, Lafarguette F, Manicacci D, Guillon F, Wang H, Laressergues D, Pichon M, Bosio M, Tatout C, 2009a. Toward the discovery of maize cell wall genes involved in silage maize quality and capacity to biofuel production. *Maydica* 54:161-198
- Barrière Y, Méchin V, Lefevre B, Maltese S, 2012. QTLs for agronomic and cell wall traits in a maize RIL progeny derived from a cross between an old Minnesota13 line and a modern Iodent line. *Theor Appl Genet* 125:531-549
- Barrière Y, Méchin V, Riboulet C, Guillaumie S, Thomas J, Bosio M, Fabre F, Goffner D, Pichon M, Lapierre C, Martinant JP, 2009b. Genetic and genomic approaches for improving biofuel production from maize. *Euphytica* 170:183-202
- Barrière Y, Ralph J, Méchin V, Guillaumie S, Grabber JH, Argillier O, Chabbert B, Lapierre C, 2004c. Genetic and molecular basis of grass cell wall biosynthesis and degradability. II. Lessons from brown-midrib mutants. *C R Biol* 327:847-860
- Barrière Y, Riboulet C, Méchin V, Maltese S, Pichon M, Cardinal AJ, Lapierre C, Lübberstedt T, Martinant JP, 2007. Genetics and genomics of lignification in grass cell walls based on maize as a model system. *Genes, Genomes and Genomics* 1:133-156

- Barrière Y, Thomas J, Denoue D, 2008. QTL mapping for lignin content, lignin monomeric composition, p-hydroxycinnamate content, and cell wall digestibility in the maize recombinant inbred line progeny F838 x F286. *Plant Sci* 175:585-595
- Baumont R, Aufrère J, Niderkorn V, Andueza D, Surault F, Peccatte JR, Delaby L, Pelletier P, 2008. La diversité spécifique dans le fourrage : conséquences sur la valeur alimentaire. *Fourrages* 194:189-206
- Baumont R, Champciaux P, Agabriel J, Andrieu J, Aufrère J, Michalet-Doreau B, Demarquilly C, 1999. Une démarche intégrée pour prévoir la valeur des aliments pour les ruminants : PrévAlim pour INRAtion. *INRA Production Animale* 12:183-194
- Beavis WD, 1998. QTL analyses: power, precision, and accuracy. In: Paterson AH (ed) *Molecular dissection of complex traits*. CRC Press, Boca Raton, pp 145-162
- Beló A, Beatty MK, Hondred D, Fengler KA, Li B, Rafalski A, 2010. Allelic genome structural variations in maize detected by array comparative genome hybridization. *Theor Appl Genet* 120:355-367
- Bennetzen JL, Schmutz J, Wang H, Percifield R, Hawkins J, Pontaroli AC, Estep M, Feng L, Vaughn JN, Grimwood J, Jenkins J, Barry K, Lindquist E, Hellsten U, Deshpande S, Wang X, Wu X, Mitros T, Triplett J, Yang X, Ye CY, Mauro-Herrera M, Wang L, Li P, Sharma M, Sharma R, Ronald PC, Panaud O, Kellogg EA, Brutnell TP, Doust AN, Tuskan GA, Rokhsar D, Devos KM, 2012. Reference genome sequence of the model plant *Setaria*. *Nat Biotechnol* 13:555-561
- Berthet S, Demont-Caulet N, Pollet B, Bidzinski P, Cezard L, Le Bris, P, Borrega N, Herve J, Blondet E, Balzergue S, Lapierre C, Jouanin L, 2011. Disruption of LACCASE4 and 17 results in tissue-specific alterations to lignification of *Arabidopsis thaliana* stems. *Plant Cell* 23:1124-1137
- Besombes S, Mazeau K, 2005. The cellulose/lignin assembly assessed by molecular modeling. Part 2: seeking for evidence of organization of lignin molecules at the interface with cellulose. *Plant Physiol Biochem* 43:277-286
- Bhargava A, Mansfield SD, Hall HC, Douglas CJ, Ellis BE, 2010. MYB75 functions in regulation of secondary cell wall formation in the *Arabidopsis* inflorescence stem. *Plant Physiol* 154:1428-1438
- Blee KA, Choi JW, O'Connell AP, Schung W, Lewis NG, Bolwell GP, 2003. A lignin specific peroxidase in tobacco whose antisense suppression leads to vascular tissue modification. *Phytochem* 64:163-176
- Boerjan W, Ralph J, Baucher M, 2003. Lignin biosynthesis. *Ann Rev Plant Biol* 54:519-546
- Boudet AM, 2000. Lignins and lignification: selected issues. *Plant Physiol Biochem* 38:81-96
- Boutrot F, Chantret N, Gautier MF, 2008. Genome-wide analysis of the rice and *Arabidopsis* non-specific lipid transfer protein (nsLtp) gene families and identification of wheat nsLtp genes by EST data mining. *BMC Genomics* 9:86
- Boyce CK, Cody GD, Fogel ML, Hazen RM, Alexander CMO, Knoll AH, 2003. Chemical evidence for cell wall lignification and the evolution of tracheids in early Devonian plants. *Int J Plant Sci* 164:691-702
- Broman KW, Sen S, 2009. *A Guide to QTL Mapping with R/qlt*. Springer, New York ISBN: 978-0-387-92124-2
- Brown DM, Zeef LA, Ellis J, Goodacre R, Turner SR, 2005. Identification of novel genes in *Arabidopsis* involved in secondary cell wall formation using expression profiling and reverse genetics. *Plant Cell* 17:2281-2295
- Brown DM, Zhang ZN, Stephens E, Dupree P, Turner SR, 2009. Characterization of IRX10 and IRX10-like reveals an essential role in glucuronoxylan biosynthesis in *Arabidopsis*. *Plant J* 57:732-746
- Buanafina MM, 2009. Feruloylation in grasses: Current and future perspectives. *Mol Plant* 2:861-872

- Buckler E, Gaut B, McMullen M, 2006. Molecular and functional diversity of maize. *Current Opin Plant Biol* 9:172-176
- Burget EG, Verma R, Molhoj M, Reiter WD, 2003. The biosynthesis of L-arabinose in plants: Molecular cloning and characterization of a Golgi-localized UDP-D-xylose 4-epimerase encoded by the MUR4 gene of Arabidopsis. *Plant Cell* 15:523-531
- Burk D, Ye Z, 2002. Alteration of oriented deposition of cellulose microfibrils by mutation of a katanin-like microtubule-severing protein. *Plant Cell* 14:2145-2160
- Cai X, Davis EJ, Ballif J, Liang M, Bushman E, Haroldsen V, Torabinejad J, Wu Y, 2006. Mutant identification and characterization of the laccase gene family in Arabidopsis. *J Exp Bot* 57:2563-2569
- Caparrós-Ruiz D, Fornalé S, Civardi L, Puigdoménech P, Rigau P, 2006. Isolation and characterisation of a family of laccases in maize. *Plant Sci* 171:217-225
- Cardinal AJ, Lee M, Moore KJ, 2003. Genetic mapping and analysis of quantitative trait loci affecting fiber and lignin content in maize. *Theor Appl Genet* 106:866-874
- Carlsbecker A, Lee JY, Roberts CJ, Dettmer J, Lehesranta S, Zhou J, Lindgren O, Moreno-Risueno MA, Vaten A, Thitamadee S, Campilho A, Sebastian J, Bowman JL, Helariutta Y, Benfey PN, 2010. Cell signalling by microRNA165/6 directs gene dose-dependent root cell fate. *Nature* 465:316-321
- Carpita NC, Gibeaut DM, 1993. Structural models of the primary cell walls of flowering plants. Consistency of structure with the physical and biochemical changes during growth. *Plant J* 3:1-30
- Carroll A, Somerville C, 2009. Cellulosic biofuels. *Annual Review of Plant Biology* 60:165-182
- Casler MD, Jung HJG, 1999. Selection and evaluation of smooth bromegrass clones with divergent lignin or etherified ferulic acid concentration. *Crop Sci* 39:1866-1873
- Chalhoub B, Belcram H, Caboche M, 2004. Efficient cloning of plant genomes into bacterial artificial chromosome (BAC) libraries with larger and more uniform insert size. *Plant Biotechnology Journal* 2:181-188
- Chavigneau H, Goué N, Courtial A, Jouanin L, Reymond M, Méchin V, Barrière Y, 2012. QTL for floral stem lignin content and degradability in three recombinant inbred line (RIL) progenies of Arabidopsis thaliana and search for candidate genes involved in cell wall biosynthesis and degradability. *OJGen* 2:7-30
- Chen F, Reddy MSS, Temple S, Jackson L, Shadle G, Dixon RA, 2006. Multi-site genetic modulation of monolignol biosynthesis suggests new routes for formation of syringyl lignin and wall-bound ferulic acid in alfalfa (*Medicago sativa* L.). *Plant J* 48:113-124
- Ching A, Caldwell KS, Jung M, Dolan M, Smith OS, Tingey S, Morgante M, Rafalski AJ, 2002. SNP frequency, haplotype structure and linkage disequilibrium in elite maize inbred lines. *BMC Genet* 3:19
- Ching A, Dhugga K, Appenzeller L, Meeley R, Bourett T, Howar R, Rafalski A, 2006. Brittle stalk 2 encodes a putative glycosylphosphatidylinositol-anchored protein that affects mechanical strength of maize tissues by altering the composition and structure of secondary cell walls. *Planta* 224:1174-1184
- Christensen JH, Bauw G, Welinder KG, Van Montagu M, Boerjan W, 1998. Purification and characterization of peroxidases correlated with lignification in poplar xylem. *Plant Physiol* 118:125-135
- Churchill GA, Doerge RW, 1994. Empirical threshold values for quantitative trait mapping. *Genetics* 138:963-971
- Cochard H, 2002. Xylem embolism and drought-induced stomatal closure in maize. *Planta* 215:466-471
- Cochard H, Barigah ST, Kleinhentz M, Eshel A, 2008. Is xylem cavitation resistance a relevant criterion for screening drought resistance among *Prunus* species? *J Plant Physiol* 165:976-982

- Coe E, Cone K, McMullen M, Chen SS, Davis G, Gardiner J, Liscum E, Polacco M, Paterson A, Sanchez-Villeda H, Soderlund C, Wing R, 2002. Access to the maize genome: an integrated physical and genetic map. *Plant Physiol* 128(1):9-12
- Courtial A, Jourda C, Arribat S, Balzergue S, Huguet S, Reymond M, Grima-Pettenati J, Barrière Y, 2012a. Comparative expression of cell wall related genes in four maize RILs and one parental line of variable lignin content and cell wall degradability. *Maydica* 57:56-74
- Courtial A, Reymond M, Grima-Pettenati J, Barrière Y, 2012b. QTLs for cell wall degradability, lignin content and structure, and p-hydroxycinnamic acid contents in the F288 x F271 early maize RIL progeny. In progress
- Courtial A, Soler M, Méchin V, Matthieu Reymond, Wang H, Grima-Pettenati J, Barrière Y, 2012c. Colocalizations between cell wall degradability QTLs and cell wall component QTLs based on a synthesis of data from six RIL progenies, and search for underlying candidate genes. In progress
- Courtial A, Thomas J, Reymond M, Méchin V, Grima-Pettenati J, Barrière Y, 2012d. Targeted linkage map densification improved positions of maize cell wall related qtls, puzzle out ghost QTLs, and highlighted significant epistatic QTL interactions. Submitted to TAG
- Craig EA, Gambill BD, Nelson RJ, 1993. Heat shock proteins: molecular chaperones of protein biogenesis. *Microbiol Rev* 57(2):402-414
- Dai X, You C, Chen G, Li X, Zhang Q, Wu C, 2011. OsBC1L4 encodes a COBRA-like protein that affects cellulose synthesis in rice. *Plant Mol Biol* 75:333-345
- Darvasi A, Weinreb A, Minke V, Weller JI, Soller M, 1993. Detecting marker-QTL linkage and estimating QTL gene effect and map location using a saturated genetic-map. *Genetics* 134:943-951
- De Givry S, Bouchez M, Chabrier P, Milan D, Schiex T, 2005. CARTHAGENE: multipopulation integrated genetic and radiated hybrid mapping. *Bioinformatics* 21:1703-1704
- De Obeso M, Caparrós-Ruiz D, Vignols F, Puigdomenech P, Rigau J, 2003. Characterisation of maize peroxidases having differential patterns of mRNA accumulation in relation to lignifying tissues. *Gene* 309:23-33
- Dellaporta J, Hicks JB, 1983. A plant DNA miniprep: version II. *Plant Mol Biol Rep* 1:19-21
- Demura T, Fukuda H, 2007. Transcriptional regulation in wood formation. *Trends in Plant Science* 12:64-70
- Dence CW, Lin SY, 1992. The determination of lignin. In SY Lin and CW Dence (Eds.), *Methods in lignin chemistry*. Springer-Verlag, Berlin, Germany, pp 33-61
- Dennis DT, Blakeley SD, 2000. Carbohydrate metabolism. In B. Buchanan, W. Gruissem, R.L. Jones (Eds.), *Biochemistry and Molecular Biology of Plants*. American Society of Plant Biologist, Rockville, MD, pp 630-675
- Devos KM, 2005. Updating the 'Crop circle'. *Current Opin Plant Biol* 8:155-162
- Dias AP, Grotewold E, 2003. Manipulating the accumulation of phenolics in maize cultured cells using transcription factors. *Biochem. Eng J* 14:207-216
- Diévert A, Dalal M, Tax FE, Lacey AD, Huttly A, Li J, Clark SE, 2003. CLAVATA1 dominant-negative alleles reveal functional overlap between multiple receptor kinases that regulate meristem and organ development. *Plant Cell* 15:1198-1211
- DiLaurenzio L, WysockaDiller J, Malamy JE, Pysh L, Helariutta Y, Freshour G, Hahn MG, Feldmann KA, Benfey PN, 1996. The SCARECROW gene regulates an asymmetric cell division that is essential for generating the radial organization of the Arabidopsis root. *Cell* 86:423-433
- Dixon RA, Chen F, Guo D, Parvathi K, 2001. The biosynthesis of monolignols: a ‘‘metabolic grid’’, or independent pathways to guaiacyl and syringyl units. *Phytochem* 57:1069-1084.

- Do CT, Pollet B, Thevenin J, Sibout R, Denoue D, Barrière Y, Lapierre C, Jouanin L, 2007. Both caffeoyl Coenzyme A 3- O -methyltransferase 1 and caffeic acid O -methyltransferase 1 are involved in redundant functions for lignin, flavonoids and sinapoyl malate biosynthesis in Arabidopsis. *Planta* 226:1117-1129
- Dolstra O, Medema JH, 1990. An effective screening method for genetic improvement of cell-wall digestibility in forage maize. In: Proceedings 15th congress maize and sorghum section of Eucarpia. 4-8 June . Baden, Austria. pp 258-270
- Doust A, 2007. Architectural evolution and its implications for domestication in grasses. *Annals of Botany* 100:941-950
- Du H, Feng BR, Yang SS, Huang YB, Tang YX, 2012. The R2R3-MYB transcription factor gene family in maize. *PLoS one* 7:e37463
- Durbak AR, Tax FE, 2011. CLAVATA signaling pathway receptors of Arabidopsis regulate cell proliferation in fruit organ formation as well as in meristems. *Genetics* 189:177-194
- Ehltng J, Mattheus N, Aeschliman DS, Li EY, Hamberger B, Cullis IF, Zhuang J, Kaneda M, Mansfield SD, Samuels L, Ritland K, Ellis BE, Bohlmann J, Douglas CJ, 2005. Global transcript profiling of primary stems from Arabidopsis thaliana identifies candidate genes for missing links in lignin biosynthesis and transcriptional regulators of fiber differentiation. *Plant J* 42:618-640
- Eichten SR, M. Foerster JM, de Leon N, Kai Y, Yeh CT, Liu S, Jeffrey A, Jeddloh JA, Schnable PS, Kaeppler SM, Springer NM, 2011. B73-Mo17 near-isogenic lines demonstrate dispersed structural variation in maize. *Plant Physiol* 156:1679-1690
- Escamilla-Trevino LL, Chen W, Card ML, Shih MC, Cheng CL, Poulton JE, 2006. Arabidopsis thaliana beta-glucosidases BGLU45 and BGLU46 hydrolyse monolignol glucosides. *Phytochem* 67:1651-1660
- Fagerstedt KV, Kukkola EM, Koistinen VV, Takahashi J, Marjamaa K, 2010. Cell wall lignin is polymerised by class III secretable plant peroxidases in Norway spruce. *J Integr Plant Biol* 52(2):186-194
- Fernandez I, Martin C, Champion M, Michalet-Doreau B, 2004. Effect of corn hybrid and chop length of whole-plant corn silage on digestion and intake by dairy cows. *Journal of Dairy Science* 87:1298-1309
- Finn RD, Tate J, Mistry J, Coghill PC, Sammut SJ, Hotz HR, Ceric G, Forslund K, Eddy SR, Sonnhammer ELL, Bateman A, 2008. The Pfam protein families database. *Nucleic Acids Res* 36: D281-D288
- Fontaine AS, Briand M, Barrière Y, 2003. Genetic variation and QTL mapping of para-coumaric and ferulic acid contents in maize stover at silage harvest. *Maydica* 48:75-82
- Fornalé S, Shi X, Chai C, Encina A, Irar S, Capellades M, Fuguet E, Torres JL, Rovira P, Puigdomènech P, Rigau J, Grotewold E, Gray J, Caparrós-Ruiz D, 2010. ZmMYB31 directly represses maize lignin genes and redirects the phenylpropanoid metabolic flux. *Plant J* 64:633-644
- Fornalé S, Sonbol FM, Maes T, Capellades M, Puigdomenech P, Rigau J, Caparrós-Ruiz D, 2006. Down-regulation of the maize and Arabidopsis thaliana caffeic acid O-methyl-transferase genes by two new maize R2R3-MYB transcription factors. *Plant Mol Biol* 62:809-823
- Foucart C, Jauneau A, Gion JM, Amelot N, Martinez Y, Panegos P, Grima-Pettenati J, Sivadon P, 2009. Overexpression of EgROP1, a Eucalyptus vascular-expressed Rac-like small GTPase, affects secondary xylem formation in Arabidopsis thaliana. *New Phytologist* 183:1014-1029
- Freeman JS, Whittock SP, Potts BM, Vaillancourt RE, 2009. QTL influencing growth and wood properties in Eucalyptus globulus. *Tree Genet Genomes* 5:713-722

- Fry SC, Willis S, Paterson A, 2000. Intraprotoplasmic and wall-localised formation of arabinoxylan-bound diferulates and larger ferulate coupling-products in maize cell-suspension cultures. *Planta* 211:679-692
- Fu Y, Wen TJ, Ronin YI, Chen HD, Guo L, Mester DI, Yang Y, Lee M, Korol AB, Ashlock DA, Schnable PS, 2006. Genetic dissection of intermated recombinant inbred lines using a new genetic map of maize. *Genetics* 174:1671-1683
- Gagnot S, Tamby JP, Martin-Magniette ML, Bitton F, Taconnat L, Balzergue S, Aubourg S, Renou JP, Lecharny A, Brunaudet V, 2008. CATdb: a public access to Arabidopsis transcriptome data from the URGV CATMA platform. *Nucleic Acids Res* 36:D986-D990
- Gardiner JM, Coe EH, Melia-Hancock S, Hoisington DA, Chao S, 1993. Development of a core RFLP map in maize using an immortalized F2 population. *Genetics* 134:917-930
- Garvin DF, 2007. Brachypodium: a new monocot model plant system emerges. *J Sci Food Agric* 87:1177-1179
- Ge Y, Dudoit S, Speed TP, 2003. Resampling-based multiple testing for microarray data analysis. *TEST* 12:1-44
- Gentleman R, Carey V, 2002. Bioconductor. *R News* 2:1116
- Gion JM, Carouche A, Deweer S, Bedon F, Pichavant F, Charpentier JP, Bailleres H, Rozenberg P, Carocha V, Ognouabi N, Verhaegen D, Grima-Pettenati J, Vigneron P, Plomion C, 2011. Comprehensive genetic dissection of wood properties in a widely-grown tropical tree: Eucalyptus. *BMC Genomics* 12:301
- Goering HK, Van Soest PJ, 1970. Forage fiber analysis (Apparatus, reagents, procedures and some applications). U.S. Dept. Agri. Sci. Handbook n°379, pp 1-20
- Goicoechea M, Lacombe E, Legay S, Mihaljevic S, Rech P, Jauneau A, Lapierre C, Pollet B, Verhaegen D, Chaubet-Gigot N, Grima-Pettenati J, 2005. EgMYB2, a new transcriptional activator from Eucalyptus xylem, regulates secondary cell wall formation and lignin biosynthesis. *Plant J* 43:553-567
- Gonthier L, Bellec A, Blassiau C, Prat E, Helmstetter N, Rambaud C, Huss B, Hendriks T, Bergès H, Quillet MC, 2010. Construction and characterization of two BAC libraries representing a deepcoverage of the genome of chicory (*Cichorium intybus* L., Asteraceae). *BMC Research Notes* 3:225
- Goujon T, Sibout R, Maba B, Nussaume L, Bechtold N, Lu F, Ralph J, Pollet B, Mila I, Charpentier JP, Barrière Y, Lapierre C, Jouanin L, 2003. A new Arabidopsis mutant deficient in the expression of O-methyltransferase 1: Impact on lignin and sinapic esters. *Plant Mol Biol* 51:973-989
- Grabber JH, 2005. How do lignin composition, structure, and cross-linking affect degradability? A review of cell wall model studies. *Crop Sci* 45:820-831
- Grabber JH, Hatfield RD, Ralph J, 1998a. Diferulate cross-links impede the enzymatic degradation of non-lignified maize walls. *J Sci Food Agric* 77:193-200
- Grabber JH, Hatfield RD, Ralph J, Zon J, Amrhein N, 1995. Ferulate cross-linking in cell-walls isolated from maize cell-suspensions. *Phytochem* 40:1077-1082
- Grabber JH, Lu FC, 2007. Formation of syringyl-rich lignins in maize as influenced by feruloylated xylans and p-coumaroylated monolignols. *Planta* 226:741-751
- Grabber JH, Mertens DR, Kim H, Funk C, Lu F, Ralph J, 2009. Cell wall fermentation kinetics are impacted more by lignin content and ferulate cross-linking than by lignin composition. *J Sci Food Agric* 89:122-129
- Grabber JH, Ralph J, Hatfield RD, 1998b. Ferulate cross-links limit the enzymatic degradation of synthetically lignified primary walls of maize. *J Agric Food Chem* 46:2609-2614

- Grabber JH, Ralph J, Hatfield RD, 2000. Cross-linking of maize walls by ferulate dimerization and incorporation into lignin. *J Agric Food Chem* 48:6106-6113
- Grabber JH, Ralph J, Hatfield RD, Quideau S, 1997. p-hydroxyphenyl, guaiacyl, and syringyl lignins have similar inhibitory effects on wall degradability. *J Agric Food Chem* 45:2530-2532
- Grabber JH, Ralph J, Lapiere C, Barrière Y, 2004. Genetic and molecular basis of grass cell-wall degradability. I. Lignin-cell wall matrix interactions. *CR Biologie* 327:455-465
- Grabber JH, Schatz PF, Kim H, Lu F, Ralph J, 2010. Identifying new lignin bioengineering targets: 1. Monolignol-substitute impacts on lignin formation and cell wall fermentability. *BMC Plant Biol* 10:114
- Grand C, Parmentier P, Boudet A, Boudet AM, 1985. Comparison of lignins and of enzymes involved in lignification in normal and brown midrib (bm3) mutant corn seedlings. *Physiol Veg* 23:905-911
- Gray J, Caparrós-Ruiz D, Grotewold E, 2012. Grass phenylpropanoids: Regulate before using! *Plant Science* 184:112-120
- Green KA, Prigge MJ, Katzman RB, Clark SE, 2005. CORONA, a member of the class III homeodomain leucine zipper gene family in Arabidopsis, regulates stem cell specification and organogenesis. *Plant Cell* 17:691-704
- Gregory SG, Sekhon M, Schein J, Zhao S, Osoegawa K, Scott CE, Evans RS, Burrige PW, Cox TV, Fox CA, Hutton RD, Mullenger IR, Phillips KJ, Smith J, Stalker J, Threadgold GJ, Birney E, Wylie K, Chinwalla A, Wallis J, Hillier L, Carter J, Gaige T, Jaeger S, Kremitzki C, Layman D, Maas J, McGrane R, Mead K, Walker R, Jones S, Smith M, Asano J, Bosdet I, Chan S, Chittaranjan S, Chiu R, Fjell C, Fuhrmann D, Girn N, Gray C, Guin R, Hsiao L, Krzywinski M, Kutsche R, Lee SS, Mathewson C, McLeavy C, Messervier S, Ness S, Pandoh P, Prabhu AL, Saedi P, Smailus D, Spence L, Stott J, Taylor S, Terpstra W, Tsai M, Vardy J, Wye N, Yang G, Shatsman S, Ayodeji B, Geer K, Tsegaye G, Shvartsbeyn A, Gebregeorgis E, Krol M, Russell D, Overton L, Malek JA, Holmes M, Heaney M, Shetty J, Feldblyum T, Nierman WC, Catanese JJ, Hubbard T, Waterston RH, Rogers J, de Jong PJ, Fraser CM, Marra M, McPherson JD, Bentley DR, 2002. A physical map of the mouse genome. *Nature* 418(6899):743-50
- Grenet E, 1997. Aspects microscopiques de la dégradation microbienne des tissus végétaux dans le rumen. *INRA Productions Animales* 10:241-249
- Grima-Pettenati J, Soler M, Camargo E, Wang H, 2012. Transcriptional Regulation of the Lignin Biosynthetic Pathway Revisited: New Players and Insights, In: Lise Jouanin and Catherine Lapiere, Editor(s), *Advances in Botanical Research*, Academic Press, Volume 61, Chapter 6, pages 173-218
- Gu Y, Wang Z, Yang Z, 2004. ROP/RAC GTPase: an old new master regulator for plant signaling. *Current Opin Plant Biol* 7:527-536
- Guillaumie S, Goffner D, Barbier B, Martinant JP, Pichon M, Barrière Y, 2008. Expression of cell wall related genes in basal and ear internodes of silking brown-midrib-3, caffeic acid O-methyltransferase (COMT) down-regulated, and normal maize plants. *BMC Plant Biol* 8:71
- Guillaumie S, Mzid R, Méchin V, Léon C, Hichri I, Destrac-Irvine A, Trossat-Magnin C, Delrot S, Lauergeat V, 2010. The grapevine transcription factor WRKY2 influences the lignin pathway and xylem development in tobacco. *Plant Mol Biol* 72:215-234
- Guillaumie S, Pichon M, Martinant JP, Bosio M, Goffner D, Barrière Y, 2007a. Differential expression of phenylpropanoid and related genes in brown-midrib bm1, bm2, bm3, and bm4 young near-isogenic maize plants. *Planta* 226:235-250
- Guillaumie S, San-Clemente H, Deswarte C, Martinez Y, Lapiere C, Murigneux A, Barrière Y, Pichon M, Goffner D, 2007b. MAIZEWALL. Database and developmental gene expression profiling of cell wall biosynthesis and assembly in maize. *Plant Physiol* 143:339-363
- Guillet-Claude C, Birolleau-Touchard C, Manicacci D, Rogowsky PM, Rigau J, Murigneux A, Martinant JP, Barrière Y, 2004. Nucleotide diversity of the ZmPox3 maize peroxidase gene:

- Relationships between a MITE insertion in exon 2 and variation in forage maize digestibility. *BMC Genetics* 5:19
- Guo D, Chen F, Inoue K, Blount JW, Dixon RA, 2001a. Downregulation of caffeic acid 3-O-methyltransferase and caffeoyl CoA 3-O-methyltransferase in transgenic alfalfa: Impacts on lignin structure and implications for the biosynthesis of G and S lignin. *Plant Cell* 13:73-88
- Guo D, Chen F, Wheeler J, Winder J, Selman S, Peterson M, Dixon RA, 2001b. Improvement of in-rumen digestibility of alfalfa forage by genetic manipulation of lignin O-methyltransferases. *Transgenic Res* 10:457-464
- Guo M, Rupe MA, Dieter JA, Zou J, Spielbauer D, Duncan KE, Howard RJ, Hou Z, Simmons CR, 2010. Cell Number Regulator1 affects plant and organ size in maize: implications for crop yield enhancement and heterosis. *Plant Cell* 22:1057-103
- Guo Y, Qin G, Gu H, Qu LJ, 2009. Dof5.6/HCA2, a Dof transcription factor gene, regulates interfascicular cambium formation and vascular tissue development in Arabidopsis. *Plant Cell* 21:3518-3534
- Haley CS, Knott SA, 1992. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity* 69:315-324
- Halpin C, Holt K, Chojecki J, Oliver D, Chabbert B, Monties B, Edwards K, Barakate A, Foxon GA, 1998. Brown-midrib maize (bm1) - a mutation affecting the cinnamyl alcohol dehydrogenase gene. *Plant J* 14:545-553
- Handakumbura PP, Hazen SP, 2012. Transcriptional regulation of grass secondary cell wall biosynthesis: playing catch-up with Arabidopsis thaliana. *Frontiers in Plant Science* 3:74
- Harrar Y, Bellini C, Faure JD, 2001. FKBP: at the crossroads of folding and transduction. *Trends Plant Sci* 6:426-431
- Harris D, DeBolt S, 2010. Synthesis, regulation and utilization of lignocellulosic biomass. *Plant Biotechnol J* 8:244-262
- Hatfield R, Fukushima RS, 2005. Can lignin be accurately measured. *Crop Sci* 45:832-839
- Hatfield R, Ralph J, Grabber JH, 2008. A potential role for sinapyl p-coumarate as a radical transfer mechanism in grass lignin formation. *Planta* 228:919-928
- Hatfield R, Vermerris W, 2001. Lignin formation in plants. The dilemma of linkage specificity. *Plant Physiol* 126:1351-1357
- Hatfield RD, Jung HJG, Ralph J, Buxton DR, Weimer PJ, 1994. A comparison of the insoluble residues produced by the Klason lignin and acid detergent lignin procedures. *J Sci Food Agric* 65:51-58
- He X, Hall MB, Gallo-Meagher M, Smith RL, 2003. Improvement of forage quality by downregulation of maize O-methyltransferase. *Crop Sci* 43:2240-2251
- Heim MA, Jakoby M, Werber M, Martin C, Weisshaar B, Bailey PC, 2003. The basic helix-loop-helix transcription factor family in plants: a genome-wide study of protein structure and functional diversity. *Mol Biol Evol* 20:735-747
- Heine GF, Malik V, Dias AP, Grotewold E, 2007. Expression and molecular characterization of ZmMYB-IF35 and related R2R3 MYB transcription factors. *Mol Biotechnol* 37(2):155-64
- Held BM, Wang HQ, John I, Wurtele ES, Colbert JT, 1993. An messenger-RNA putatively coding for an O-methyltransferase accumulates preferentially in maize roots and is located predominantly in the region of the endodermis. *Plant Physiol* 102:1001-1008
- Hemsley A, Kemp A, Grierson C, 2005. The TIP GROWTH DEFECTIVE1 S-Acyl Transferase Regulates Plant Cell Growth in Arabidopsis. *Plant Cell* 17(9): 2554-2563

- Herendeen PS, Grane PR, 1995. The fossil history of the monocotyledons. In: Rudall PJ, Cribbb PJ, Cutler DF, Humphries CJ (Eds) *Monocotyledons, Systematic and Evolution*, Royal Botanical Gardens, Kew, pp 1-21
- Higuchi T, Ito Y, Kawamura I, 1967. p-Hydroxyphenyl component of grass lignin and the role of tyrosine ammonia-lyase in its formation. *Phytochem* 6:875-881
- Higuchi T, Kawamura I, 1966. Occurrence of p-hydroxyphenylglycerol-beta-aryl ether structure in lignins. *Holzforschung* 20:16-21
- Iiyama K, Lam TBT, Stone BA, 1990. Phenolic acid bridges between polysaccharides and lignin in wheat internodes. *Phytochem* 29:733-737
- Ilegems M, Douet V, Meylan-Bettex M, Uyttewaal M, Brand L, Bowman JL, Stieger PA, 2010. Interplay of auxin, KANADI and Class III HD-ZIP transcription factors in vascular tissue formation. *Development* 137:975-984
- Im KH, Cosgrove DT, Jones AM, 2000. Subcellular localization of expansin mRNA in xylem cells. *Plant Physiol* 123:463-470
- Irizarry RA, Ooi SL, Wu Z, Boeke JD, 2003. Use of mixture models in a microarray based screening procedure for detecting differentially represented yeast mutants. *Stat Appl Genet Mol Biol* 2:Article1
- Jacobs BF, Kingston JD, Jacobs LL, 1999. The origin of grass-dominated ecosystems. *Annals of the Missouri Botanical Garden* 86:590-643
- Jacquet G., Pollet B, Lapierre C, 1995. New ether-linked ferulic acid-coniferyl alcohol dimers identified in grass straws. *J Agric Food Chem* 43:2746-2751
- Johansen DA, 1940. *Plant microtechnique*. McGraw Hill editor, London, UK, 523 pages
- Juarez MT, Kui JS, Thomas J, Heller BA, Timmermans MCP, 2004. MicroRNA-mediated repression of rolled leaf1 specifies maize leaf polarity. *Nature* 428:84-88
- Jung HG, Allen MS, 1995. Characteristics of plant cell wall affecting intake and digestibility of forages by ruminants. *J Anim Sci* 73: 2774-2790
- Jung HG, Deetz DA, 1993. Cell wall lignification and degradability. In: H. G. Jung, D. R. Buxton, R. D. Hatfield, and J. Ralph (Ed.) *Forage Cell Wall Structure and Digestibility*. ASA-CSSA-SSSA, Madison, WI. pp. 315-346
- Jung HG, Mertens D, Payne A, 1997. Correlation of acid detergent lignin and Klason lignin with digestibility of forage dry matter and neutral detergent fiber. *J Dairy Sci* 80:1622-1628
- Jung HG, Mertens D, Phillips RL, 2011. Effect of reduced ferulate-mediated lignin/arabinoxylan cross-linking in corn silage on feed intake, digestibility, and milk production. *J Dairy Sci* 94:5124-5137
- Jung HG, Morrison TA, Buxton DR, 1998. Degradability of cell-wall polysaccharides in maize internodes during stalk development. *Crop Sci* 38:1047-1051
- Jung HG, Phillips RL, 2010. Putative seedling ferulate ester (sfe) maize mutant: morphology, biomass yield, and stover cell wall composition and rumen degradability. *Crop Sci* 50:403-418
- Jung HJ, Samac DA, Sarath G, 2012. Modifying crops to increase cell wall digestibility. *Plant Science* 185-186:65-77
- Kajita S, Hishiyama S, Tomimura Y, Katayama Y, Omori S, 1997. Structural characterization of modified lignin in transgenic tobacco plants in which the activity of 4-coumarate:coenzyme A ligase is depressed. *Plant Physiol* 114:871-879
- Kaneda M, Schuetz M, Lin B, Chanis C, Hamberger B, Western T, Ehlting J, Samuels A, 2011. ABC transporters coordinately expressed during lignification of Arabidopsis stems include a set of ABCBs associated with auxin transport. *Journal of Experimental Botany* 62:2063-2077

- Kato Y, Shiozawa R, Takeda S, Ito S, Matsuda K, 1982. Structural investigation of b-glucan and xyloglucan from bambooshoot cell-walls. *Carbohydr Res* 109:233-248
- Kavousi B, Daudi A, Cook CM, Joseleau JP, Ruel K, Devoto A, Bolwell GP, Blee KA. 2010. Consequences of antisense down-regulation of a lignification-specific peroxidase on leaf and vascular tissue in tobacco lines demonstrating enhanced enzymic saccharification. *Phytochem* 71:531-542
- Kawaoka A, Ebinuma H, 2001. Transcriptional control of lignin biosynthesis by tobacco LIM protein. *Phytochem* 57:1149-1157
- Kellogg EA, 2001. Evolutionary history of the grasses. *Plant Physiol* 125:1198-1205
- Kendall MG, Stuart A, 1961. The advanced theory of statistics. Inference and relationship. Charles Griffin and Co (eds), London, vol. II, 3rd edition
- Kim WC, Ko JH, Han KH, 2012. Identification of a cis-acting regulatory motif recognized by MYB46, a master transcriptional regulator of secondary wall biosynthesis. *Plant Mol Biol* 78:489-501
- Ko JH, Han KH, 2004. Arabidopsis whole-transcriptome profiling defines the features of coordinated regulations that occur during secondary growth. *Plant Mol Biol* 55:433-453
- Ko JH, Kim WC, Han KH, 2009. Ectopic expression of MYB46 identifies transcriptional regulatory genes involved in secondary wall biosynthesis in Arabidopsis. *Plant J* 60:649-665
- Ko JH, Prassinis C, Han KH, 2006. Developmental and seasonal expression of PtaHB1, a Populus gene encoding a class IIIHD-Zip protein, is closely associated with secondary growth and inversely correlated with the level of microRNA (miR166). *New Phytol* 169:469-478
- Kobayashi Y, Motose H, Iwamoto K, Fukuda H, 2011. Expression and genome-wide analysis of the xylogen-type gene family. *Plant Cell Physiol* 52:1095-1106
- Kobilinsky A, 1983. MODLI, logiciel d'analyse de modèles linéaires. INRA
- Krakowsky M, Lee M, Beeghly H, Coors J, 2003. Characterization of quantitative trait loci affecting fiber and lignin in maize (*Zea mays* L.). *Maydica* 48:283-292
- Krakowsky M, Lee M, Coors J, 2005. Quantitative trait loci for cell-wall components in recombinant inbred lines of maize (*Zea mays* L.). I: Stalk tissue. *Theor Appl Genet* 111:337-346
- Kubo M, Udagawa M, Nishikubo N, Horiguchi G, Yamaguchi M, Ito J, Mimura T, Fukuda H, Demura T, 2005. Transcription switches for protoxylem and metaxylem vessel formation. *Genes Dev* 19:1855-1860
- Kushwaha H, Gupta S, Singh VK, Rastogi S, Yadav D, 2011. Genome wide identification of Dof transcription factor gene family in sorghum and its comparative phylogenetic analysis with rice and Arabidopsis. *Mol Biol Rep* 38:5037-5053
- Kwon SI, Cho HJ, Jung JH, Yoshimoto K, Shirasu K, Park OK, 2010. The Rab GTPase RabG3b functions in autophagy and contributes to tracheary element differentiation in Arabidopsis. *Plant J* 64:151-164
- Lai J, Li R, Xu X, Jin W, Xu M, Zhao H, Xiang Z, Song W, Ying K, Zhang M, Jiao Y, Ni P, Zhang J, Li D, Guo X, Ye K, Jian M, Wang B, Zheng H, Liang H, Zhang X, Wang S, Chen S, Li J, Fu Y, Springer NM, Yang H, Wang J, Dai J, Schnable PS, Wang J, 2010. Genome-wide patterns of genetic variation among elite maize inbred lines. *Nat Genet* 42:1027-1030
- Lam TBT, Iiyama K, Stone BA, 2003. Hot alkali-labile linkages in the wall of the forage grass *Phalaris aquatica* and *Lolium perenne* and their relation to in vitro wall digestibility. *Phytochem* 64:603-607
- Lander ES, Botstein D, 1989. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 121:185-199

- Lanot A, Hodge D, Jackson R, George G, Elias L, Lim E, Vaistij F, Bowles D, 2006. The glucosyltransferase UGT72E2 is responsible for monolignol 4-O-glucoside production in *Arabidopsis thaliana*. *Plant J* 48:286-295
- Lapierre C, 1993. Applications of new methods for the investigation of lignin structure. In: Jung HG, Buxton DR, Hatfield RD (Eds) *Forage Cell Wall Structure and Digestibility*, ASA-CSSA-SSSA: Madison, WI, pp 133-136
- Lawrence CJ, Walbot V, 2007. Translational Genomics for Bioenergy Production from Fuelstock Grasses: Maize as the Model Species. *Plant Cell* 19:2091-2094
- Lee D, Meyer K, Chapple C, Douglas CJ, 1997. Antisense suppression of 4-coumarate: coenzyme a ligase activity in *Arabidopsis* leads to altered lignin subunit composition. *Plant Cell* 9:1985-1998
- Lee H, Kim B, Song SK, Heo JO, Yu NI, Lee SA, Kim M, Kim DG, Sohn SO, Lim CE, Chang KS, Lee MM, Lim J, 2008. Large-scale analysis of the GRAS gene family in *Arabidopsis thaliana*. *Plant Mol Biol* 67:659-670
- Lee Y, Chen F, Gallego-Giraldo L, Dixon RA, Voit EO, 2011. Integrative analysis of transgenic alfalfa (*Medicago sativa* L.) suggests new metabolic control mechanisms for monolignol biosynthesis. *Plos Comput Biol* 7:e1002047
- Lee, H, Kim B, Song SK, Heo JO, Yu NI, Lee SA, Kim M, Kim DG, Sohn SO, Lim CE, Chang KS, Lee MM, Lim J, 2008. Large-scale analysis of the GRAS gene family in *Arabidopsis thaliana*. *Plant Mol Biol* 67:659-670
- Legay S, Lacombe E, Goicoechea M, Briere C, Seguin A, MacKay J, Grima-Pettenati J, 2007. Molecular characterization of EgMYB1, a putative transcriptional repressor of the lignin biosynthetic pathway. *Plant Sci* 173:542-549
- Legay S, Sivadon P, Blervacq AS, Pavy N, Baghdady A, Tremblay L, Levasseur C, Ladouce N, Lapierre C, Séguin A, Hawkins S, Mackay J, Grima-Pettenati J, 2010. EgMYB1, an R2R3 MYB transcription factor from eucalyptus negatively regulates secondary cell wall formation in *Arabidopsis* and poplar. *The New Phytologist* 188:774-786
- Li E, Bhargava A, Qiang W, Friedmann MC, Forneris N, Savidge RA, Johnson LA, Mansfield SD, Ellis BE, Douglas CJ, 2012. The Class II KNOX gene KNAT7 negatively regulates secondary wall formation in *Arabidopsis* and is functionally conserved in *Populus*. *New Phytol* 194:102-115
- Li E, Wang S, Liu Y, Chen JG, Douglas CJ, 2011. OVATE FAMILY PROTEIN4 (OFP4) interaction with KNAT7 regulates secondary cell wall formation in *Arabidopsis thaliana*. *Plant J* 67: 328-341
- Li X, Duan X, Jiang H, Sun Y, Tang Y, Yuan Z, Guo J, Liang W, Chen L, Yin J, Ma H, Wang J, Zhang D, 2006. Genome-wide analysis of basic/helix-loop-helix transcription factor family in rice and *Arabidopsis*. *Plant Physiol* 141:1167-1184
- Li X, Ximenes E, Kim Y, Slininger M, Meilan R, Ladisch M, Chapple C, 2010. Lignin monomer composition affects *Arabidopsis* cell-wall degradability after liquid hot water pretreatment. *Biotechnol Biofuels* 3:27
- Li Y, Jones L, McQueen-Mason S, 2003a. Expansins and cell growth. *Curr Opin Plant Biol* 6:603-610
- Li Y, Kajita S, Kawai S, Katayama Y, Morohoshi N, 2003b. Down-regulation of an anionic peroxidase in transgenic aspen and its effect on lignin characteristics. *J Plant Res* 116:175-182
- Li Y, Qian Q, Zhou Y, Yan M, Sun L, Zhang M, Fu Z, Wang Y, Han B, Pang X, Chen M, Li J, 2003c. Brittle culm1, which encodes a COBRA-like protein, affects the mechanical properties of rice plants. *Plant Cell* 15:2020-2031
- Libault M, Stacey G, 2010. Evolution of FW2.2-like (FWL) and PLAC8 genes in eukaryotes. *Plant Signal Behav* 5:1226-1228

- Liljegren SJ, Ditta GS, Eshed HY, Savidge B, Bowman JL, Yanofsky MF, 2000. SHATTERPROOF MADS-box genes control seed dispersal in Arabidopsis. *Nature* 404:766-770
- Lim EK, Jackson RG, Bowles DJ, 2005. Identification and characterization of Arabidopsis glycosyltransferases capable of glucosylating coniferyl aldehyde and sinapyl aldehyde. *FEBS Lett* 579:2802-2806
- Lindsay SE, Fry SC, 2008. Control of diferulate formation in dicotyledonous and gramineous cell-suspension cultures. *Planta* 227:439-452
- Liu CJ, 2012. Deciphering the enigma of lignification: Precursor transport, oxidation, and the topochemistry of lignin assembly. *Mol Plant* 5:304-317
- Liu F, Schnable PS, 2002. Functional specialization of maize mitochondrial aldehyde dehydrogenases. *Plant Physiol* 130:1657-1674
- Liu S, Yeh CT, Ji T, Ying K, Wu H, Tang HM, Fu Y, Nettleton D, Schnable PS, 2009. Mu transposon insertion sites and meiotic recombination events co-localize with epigenetic marks for open chromatin across the maize genome. *PLoS Genet* 5: e1000733
- Lu FC, Ralph J, 1999. Detection and determination of p-coumaroylated units in lignins. *J Agric Food Chem* 47:1988-1992
- Lübberstedt T, Melchinger AE, Schön CC, Utz HF, Klein D, 1997. QTL mapping in testcrosses of European flint lines of maize: I. Comparison of different testers for forage yield traits. *Crop Sci* 37:921-931
- Luo MC, Xu K, Ma Y, Deal KR, Nicolet CM, Dvorak J, 2009. A high-throughput strategy for screening of bacterial artificial chromosome libraries and anchoring of clones on a genetic map constructed with single nucleotide polymorphisms. *BMC Genomics* 10:28
- Ma K, Xiao J, Li X, Zhang Q, Lian X, 2009. Sequence and expression analysis of the C3HC4-type RING finger gene family in rice. *Gene* 444:33-45
- MacAdam JW, Grabber JH, 2002. Relationship of growth cessation with the formation of diferulate cross-links and p-coumaroylated lignins in tall fescue leaf blades. *Planta* 215:785-793
- Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen Z, Dewell SB, Du L, Fierro JM, Gomes XV, Godwin BC, He W, Helgesen S, Ho CH, Irzyk GP, Jando SC, Alenquer ML, Jarvie TP, Jirage KB, Kim JB, Knight JR, Lanza JR, Leamon JH, Lefkowitz SM, Lei M, Li J, Lohman KL, Lu H, Makhijani VB, McDade KE, McKenna MP, Myers EW, Nickerson E, Nobile JR, Plant R, Puc BP, Ronan MT, Roth GT, Sarkis GJ, Simons JF, Simpson JW, Srinivasan M, Tartaro KR, Tomasz A, Vogt KA, Volkmer GA, Wang SH, Wang Y, Weiner MP, Yu P, Begley RF, Rothberg JM, 2006. Genome sequencing in microfabricated high-density picolitre reactors (vol 437, pg 376, 2005). *Nature* 441(7089):120-120
- Markussen T, Fladung M, Achere V, Favre JM, Faivre-Rampant P, Aragonés A, Perez DD, Harvenge L, Espinel S, Ritter E, 2003. Identification of QTLs controlling growth, chemical and physical wood property traits in *Pinus pinaster* (Ait.). *Silvae Genetica* 52:8-15
- Martinez AT, Rencoret J, Marques G, Gutierrez A, Ibarra D, Jimenez-Barbero J, del Rio JC, 2008. Monolignol acylation and lignin structure in some nonwoody plants: A 2D NMR study. *Phytochem* 69:2831-2843
- Martinez O, Curnow RN, 1992. Estimating the locations and the sizes of the effects of quantitative trait loci using flanking markers. *Theor Appl Genet* 85:480-488
- Martz F, Maury S, Pincon G, Legrand M, 1998. cDNA cloning, substrate specificity and expression study of tobacco caffeoyl-CoA 3-O-methyltransferase, a lignin biosynthetic enzyme. *Plant Mol Biol* 36:427-437

- Matsuoka Y, Vigouroux Y, Goodman MM, Sanchez GJ, Buckler E, Doebley J, 2002. A single domestication for maize shown by multilocus microsatellite genotyping. *Proc Natl Acad Sci USA* 99:6080-6084
- Matthews J, Bhati M, Lehtomaki E, Mansfield R, Cubeddu L, MacKay J, 2009. It takes two to tango: the structure and function of LIM, RING, PHD and MYND domains. *Cur Pharmaceutical Design* 15:3681-3696
- McCaig BC, Meagher RB, Dean JF, 2005. Gene structure and molecular analysis of the laccase-like multicopper oxidase (LMCO) gene family in *Arabidopsis thaliana*. *Planta* 221:619-636
- McConnell JR, Emery J, Eshed Y, Bao N, Bowman J, Barton MK, 2001. Role of PHABULOSA and PHAVOLUTA in determining radial patterning in shoots. *Nature* 411:709-713
- McPherson J D, Marra M, Hillier L, Waterston R H, Chinwalla A, Wallis J, Sekhon M, Wylie K, Mardis E R, Wilson R K, Fulton R, Kucaba T A, Wagner-McPherson C, Barbazuk W B, Gregory S G, Humphray S J, French L, Evans R S, Bethel G, Whittaker A, Holden J L, McCann O T, Dunham A, Soderlund C, Scott C E, Bentley D R, Schuler G, Chen H C, Jang W, Green E D, Idol J R, Maduro V V, Montgomery K T, Lee E, Miller A, Emerling S, Kucherlapati, Gibbs R, Scherer S, Gorrell J H, Sodergren E, Clerc-Blankenburg K, Tabor P, Naylor S, Garcia D, de Jong P J, Catanese J J, Nowak N, Osoegawa K, Qin S, Rowen L, Madan A, Dors M, Hood L, Trask B, Friedman C, Massa H, Cheung V G, Kirsch I R, Reid T, Yonescu R, Weissenbach J, Bruls T, Heilig R, Branscomb E, Olsen A, Doggett N, Cheng J F, Hawkins T, Myers R M, Shang J, Ramirez L, Schmutz J, Velasquez O, Dixon K, Stone N E, Cox D R, Haussler D, Kent W J, Furey T, Rogic S, Kennedy S, Jones S, Rosenthal A, Wen G, Schilhabel M, Gloeckner G, Nyakatura G, Siebert R, Schlegelberger B, Korenberg J, Chen X N, Fujiyama A, Hattori M, Toyoda A, Yada T, Park H S, Sakaki Y, Shimizu N, Asakawa S, Kawasaki K, Sasaki T, Shintani A, Shimizu A, Shibuya K, Kudoh J, Minoshima S, Ramser J, Seranski P, Hoff C, Poustka A, Reinhardt R, Lehrach H, International Human Genome Mapping Consortium, 2001. A physical map of the human genome. *Nature* 409:934-941
- Méchin V, 2000. Etude des facteurs biochimiques et génétiques explicatifs de la variabilité pour la valeur alimentaire du maïs fourrage. Thèse INA-PG 17 janvier 2000, 205 pages
- Méchin V, Argillier O, Hébert Y, Guingo E, Moreau L, Charcosset A, Barrière Y, 2001. Genetic analysis and QTL mapping of cell wall digestibility and lignification in silage maize. *Crop Sci* 41:690-697
- Méchin V, Argillier O, Menanteau V, Barrière Y, Mila I, Pollet B, Lapierre C, 2000. Relationship of cell wall composition to in vitro cell wall digestibility of maize inbred line stems. *J Sci Food Agric* 80:574-580
- Méchin V, Argillier O, Rocher F, Hébert Y, Mila I, Pollet B, Barrière Y, Lapierre C, 2005. In search of a maize ideotype for cell wall enzymatic degradability using histological and biochemical lignin characterization. *J Agricul and Food Chem* 53:5872-5881
- Melchinger AE, Utz HF, Schön CC, 1998. Quantitative trait locus (QTL) mapping using different testers and independent population samples in maize reveals low power of QTL detection and large bias in estimates of QTL effects. *Genetics* 149:383-403
- Meng H, Campbell WH, 1998. Substrate profiles and expression of caffeoyl coenzyme A and caffeic acid O-methyltransferases in secondary xylem of aspen during seasonal development. *Plant Mol Biol* 38:513-520
- Messing J, Dooner H, 2006. Organization and variability of the maize genome. *Current Opin Plant Biol* 9:157-163
- Meyers BC, Tingey SV, Morgante M, 2001. Abundance, Distribution, and Transcriptional Activity of Repetitive Elements in the Maize Genome. *Genome Res* 11:1660-1676
- Miao YC, Liu CJ, 2010. ATP-binding cassette-like transporters are involved in the transport of lignin precursors across plasma and vacuolar membranes. *Proc Natl Acad Sci (USA)* 107:22728-22733

- Migné C, Grenet E, Jamot J, 1996. Microbial degradation of the apical internode of Co125 and W401 maize in the rumen. *Anim Feed Sci Technol* 58:165-185
- Migné C, Prensier G, Cornu A, Grenet E, 1996. Immunocytochemical localization of arabinoxylans in the cell wall of maize apical internode after microbial degradation in the rumen. *Biol Cell* 88:137-144
- Milioni D, Sado PE, Stacey NJ, Domingo C, Roberts K, McCann MC, 2001. Differential expression of cell-wall-related genes during the formation of tracheary elements in the Zinnia mesophyll cell system. *Plant Mol Biol* 47:221-238
- Minic Z, Jamet E, San-Clemente H, Pelletier S, Renou JP, Rihouey C, Okinyo DP, Proux C, Lerouge P, Jouanin L, 2009. Transcriptomic analysis of Arabidopsis developing stems: a close-up on cell wall genes. *BMC Plant Biology* 9:6
- Minson DJ, Wilson JR, 1994. Prediction of intake as an element of forage quality. *Forage Quality, Evaluation, and Utilization*. G. C. Fahey, ed. Am. Soc. Agron. Inc, Madison, WI 533-563
- Mitchell RAC, Dupree P, Shewry PR, 2007. A novel bioinformatics approach identifies candidate genes for the synthesis and feruloylation of arabinoxylan. *Plant Physiol* 144:43-53
- Mitsuda N, Seki M, Shonozaki K, Ohme-Takagi M, 2005. The NAC transcription factors NST1 and NST2 of Arabidopsis regulate secondary wall thickenings and are required for anther dehiscence. *Plant Cell* 17:2993-3006
- Morgenstern B, 2004. DIALIGN: multiple DNA and protein sequence alignment at BiBiServ. *Nucleic Acids Res* 32:W33-36
- Morrison TA, Kessler JR, Buxton DR, 1994. Maize internode elongation patterns. *Crop Sci* 34:1055-1060
- Morrison WH, Akin DE, Himmelsbach DS, Gamble GR, 1993. Investigation of the ester-linked and ether-linked phenolic constituents of cell-wall types of normal and brown-midrib pearl-millet using chemical isolation, microspectrophotometry and C-13 NMR-spectroscopy. *J Sci Food Agric* 63:329-337
- Muller B, Bourdais G, Reidy B, Bencivenni C, Massonneau A, Condamine P, Rolland G, Conéjéro G, Rogowsky P, Tardieu F, 2007. Association of specific expansins with growth in maize leaves is maintained under environmental, genetic, and developmental sources of variation. *Plant Physiol* 143:278-290
- Murat F, Xu JH, Tannier E, Abrouk M, Guilhot N, Pont C, Messing J, Salse J, 2010. Ancestral grass karyotype reconstruction unravels new mechanisms of genome shuffling as a source of plant evolution. *Genome Res* 20:1545-1557
- Murigneux A, Martinant JP, Barrière Y, 2006. Apports de la génomique à l'amélioration des plantes. L'exemple du maïs fourrage. *Biofutur* 266:34-41
- Nair RB, Bastress KL, Ruedger MO, Denault JW, Chapple C, 2004. The Arabidopsis thaliana reduced epidermal fluorescence1 gene encodes an aldehyde dehydrogenase involved in ferulic acid and sinapic acid biosynthesis. *Plant Cell* 16:544-554
- Nakano J, Meshitsuka N, 1992. The detection of lignin. *Methods in Lignin Chemistry*. Springer-Verlag: Berlin, pp 23-32
- Ni W, Phillips RL, Jung HG, 1998. Quantitative trait loci for cell-wall traits in maize. 8th International Cell Wall Meeting, Sept. 1-5 Norwich, UK, p 1.38
- Nibau C, Wu HM, Cheung AY, 2006. RAC/ROP GTPases: 'hubs' for signal integration and diversification in plants. *Trends Plant Sci* 11:309-315
- Nielsen KL, Indiani C, Henriksen A, Feis A, Becucci M, Gajhede M, Smulevich G, Welinder KG, 2001. Differential activity and structure of highly similar peroxidases. Spectroscopic, crystallographic and enzymatic analyses of lignifying Arabidopsis thaliana peroxidase A2 and horseradish peroxidase A2. *Biochemica* 40:11013-11021

- Nishitani K, Nevins DJ, 1991. Glucuronoxylan xylanohydrolase. A unique xylanase with the requirement for appendant glucuronosyl units. *J Biol Chem* 266:6539-6543
- Niu L, Zhang Y, Pei Y, Liu C, Cao X, 2008. Redundant requirement for a pair of PROTEIN ARGININE METHYLTRANSFERASE4 Homologs for the proper regulation of Arabidopsis flowering time. *Plant Physiol* 148:490-503
- Nziengui H, Schoefs B, 2009. Functions of reticulons in plants: What we can learn from animals and yeasts? *Cell Mol Life Sci* 66:584-595
- Onnerud H, Zhang L, Gellerstedt G, Henriksson G, 2002. Polymerization of monolignols by redox shuttle-mediated enzymatic oxidation: a new model in lignin biosynthesis I. *Plant Cell* 14:1953-1962
- Ostergaard L, Teillum K, Mirza O, Mattsson O, Petersen M, Welinder KG, Mundy J, Gajhede M, Henriksen A, 2000. Arabidopsis ATP A2 peroxidase. Expression and high resolution structure of a plant peroxidase with implications for lignification. *Plant Molecular Biology* 44:231-243
- Palanivelu R, Brass L, Edlund AF, Preuss D, 2003. Pollen tube growth and guidance is regulated by POP2, an Arabidopsis gene that controls GABA levels. *Cell* 114:47-59
- Parker G, Schofield R, Sundberg B, Turner S, 2003. Isolation of COV1, a gene involved in the regulation of vascular patterning in the stem of Arabidopsis. *Development* 130:2139-2148
- Parvathi K, Chen F, Guo DJ, Blount JW, Dixon RA, 2001. Substrate preferences of O-methyltransferases in alfalfa suggest new pathways for 3-O-methylation of monolignols. *Plant J* 25:193-202.
- Pauly M, Keegstra K, 2010. Plant cell walls as precursors for biofuels. *Current Opin Plant Biol* 13:305-312
- Paz-Ares J, Ghosal D, Wienand U, Peterson PA, Saedler H, 1987. The regulatory c1 locus of Zea mays encodes a protein with homology to myb proto-oncogene products and with structural similarities to transcriptional activators. *EMBO J* 6:3553-3558
- Pei Y, Niu L, Lu F, Liu C, Zhai J, Kong X, Cao X, 2007. Mutations in the Type II Protein Arginine Methyltransferase AtPRMT5 result in pleiotropic developmental defects in Arabidopsis. *Plant Physiol* 144:1913-1923
- Pelegrini PB, Franco OL, 2005. Plant gamma-thionins: novel insights on the mechanism of action of a multi-functional class of defense proteins. *Int J Biochem Cell Biol* 37:2239-2253
- Pesquet E, Korolev AV, Calder G, Lloyd CW, 2010. The microtubule-associated protein AtMAP70-5 regulates secondary wall patterning in Arabidopsis wood cells. *Cur Biol* 20:744-749
- Pesquet E, Korolev AV, Calder G, Lloyd CW, 2011. Mechanisms for shaping, orienting, positioning and patterning plant secondary cell walls. *Plant Signal Behav* 6:843-849
- Peterson DG, Tomkins JP, Frisch DA, Wing RA, Paterson AH, 2002. Construction of plant bacterial artificial chromosome (BAC) libraries: An illustrated guide. Second Edition
- Petsch K, Mylne J, Botella J, 2005. Cosuppression of Eukaryotic Release Factor 1-1 in Arabidopsis Affects Cell Elongation And Radial Cell Division. *Plant Physiol* 139:115-126
- Pichon M, Deswartes C, Gerentes D, Guillaumie S, Lapierre C, Toppan A, Barrière B, Goffner D, 2006. Variation in lignin and cell wall digestibility in caffeic acid O-methyltransferase down-regulated maize half-sib progenies in field experiments. *Mol Breeding* 18:253-261
- Piquemal J, Lapierre C, Myton K, O'Connell A, Schuch W, Grima-Pettenati J, Boudet AM, 1998. Down-regulation of cinnamoyl-CoA reductase induces significant changes of lignin profiles in transgenic tobacco plants. *Plant J* 13:71-83
- Pires N, Dolan L, 2010. Origin and diversification of basic-helix-loop-helix proteins in plants. *Molecular Biology and Evolution* 27:862-874

- Piston F, Uauy C, Fu L, Langston J, Labavitch J, Dubcovsky J, 2010. Down-regulation of four putative arabinoxylan feruloyl-transferase genes from family PF02458 reduces ester-linked ferulate content in rice cell walls. *Planta* 231:677-691
- Pomar F, Merino F, Ros Barceló A, 2002. O-4-Linked coniferyl and sinapyl aldehydes in lignifying cell walls are the main targets of the Wiesner (phloroglucinol-HCl) reaction. *Protoplasma* 220:17-28
- Pot D, Rodrigues JC, Rozenberg P, Chantre G, Tibbits J, Cahalan C, Pichavant F, Plomion C, 2006. QTLs and candidate genes for wood properties in maritime pine (*Pinus pinaster* Ait.). *Tree Genetics Genomes* 2:10-24
- Ralph J, 2010. Hydroxycinnamates in lignification. *Phytochem Rev* 9:65-83
- Ralph J, Brunow G, Harris PJ, Dixon RA, Schatz RF, Boerjan W, 2008a. Lignification: are lignins biosynthesized via simple combinatorial chemistry or via proteinaceous control and template replication?. In F Daayf, A El Hadrami, L. Adam, GM Ballance (Eds), *Recent advances in polyphenol research*. Wiley-Blackwell Publishing, Oxford, UK, pp 36-66
- Ralph J, Bunzel M, Marita JM, HatWeld RD, Lu F, Kim H, Schatz PF, Grabber JH, Steinhart H, 2004. Peroxidase-dependent cross-linking reactions of p-hydroxycinnamates in plant cell walls. *Phytochem Rev* 3:79-96
- Ralph J, Grabber JH, Hatfield RD, 1995. Lignin-ferulate cross-links in grasses - Active incorporation of ferulate polysaccharide esters into ryegrass lignins. *Carbohydr Res* 275:167-178
- Ralph J, Hatfield RD, Piquemal J, Yahiaoui N, Pean M, Lapiere C, Boudet AM, 1998. NMR characterization of altered lignins extracted from tobacco plants down-regulated for lignification enzymes cinnamyl-alcohol dehydrogenase and cinnamoyl-CoA reductase. *Proc Natl Acad Sci USA* 95:12803-12808
- Ralph J, HatWeld RD, Quideau S, Helm RF, Grabber JH, Jung HJG, 1994. Pathway of p-coumaric acid incorporation into maize lignin as revealed by NMR. *J Am Chem Soc* 116:9448-9456
- Ralph J, Helm RF, Quideau S, Hatfield RD, 1992. Lignin feruloyl ester cross-links in grasses. Part 1. Incorporation of feruloyl esters into coniferyl alcohol dehydrogenation polymers. *J Chem Soc, Perkin Trans 1* 1:2961-2969
- Ralph J, Kim H, Lu F, Grabber JH, Leple JC, Berrio-Sierra J, Derikvand MM, Jouanin L, Boerjan W, Lapiere C, 2008b. Identification of the structure and origin of a thioacidolysis marker compound for ferulic acid incorporation into angiosperm lignins (and an indicator for cinnamoyl CoA reductase deficiency). *Plant J* 53:368-379
- Rampey RA, Woodward AW, Hobbs BN, Tierney MP, Lahner B, Salt DE, Bartel B, 2006. An *Arabidopsis* basic helix-loophelix leucine zipper protein modulates metal homeostasis and auxin conjugate responsiveness. *Genetics* 174:1841-1857
- Ramsay NA, Glover BJ, 2005. MYB-bHLH-WD40 protein complex and the evolution of cellular diversity. *Trends Plant Sci* 10:63-70
- Ranjan P, Yin T, Zhang X, Kalluri U, Yang X, Jawdy S, Tuskan G, 2010. Bioinformatics-based identification of candidate genes from QTLs associated with cell wall traits in *Populus*. *Bioenerg Res* 3:172-182
- Ranocha P, Chabannes M, Chamayou S, Danoun S, Jauneau A, Boudet AM, Goffner D, 2002. Laccase down-regulation causes alterations in phenolic metabolism and cell wall structure in poplar. *Plant Physiol* 129:145-155
- Ratcliffe OJ, Riechmann JL, Zhang JZ, 2000. Interfascicular fiberless1 is the same gene as REVOLUTA. *Plant Cell* 12:315-317
- Rebourg C, Chastanet M, Gouesnard B, Welcker C, Dubreuil P, Charcosset A, 2003. Maize introduction into Europe: the history reviewed in the light of molecular data. *Theor Applied Genet* 106:895-903

- Renault H, El Amrani A, Palanivelu R, Updegraff EP, Yu A, Renou JP, Preuss D, Bouchereau A, Deleu C, 2011. GABA accumulation causes cell elongation defects and a decrease in expression of genes encoding secreted and cell wall-related proteins in *Arabidopsis thaliana*. *Plant Cell Physiol* 52 (5):894-908
- Rengel D, San Clemente H, Servant F, Ladouce N, Paux E, Wincker P, Couloux A, Sivadon P, Grima-Pettenati J, 2009. A new genomic resource dedicated to wood formation in *Eucalyptus*. *BMC Plant Biology* 9:36
- Riboulet C, 2007. Recherche des déterminants biochimiques et moléculaires de la réticulation des parois et de l'ingestibilité du maïs fourrage. Thèse : Sciences agronomiques, biotechnologies et agro-alimentaires soutenue le 20 décembre 2007 à Poitiers, 185 pages
- Riboulet C, Fabre F, Dénoue D, Martinant JP, Lefevre B, Barrière Y, 2008a. QTL mapping and candidate gene research for lignin content and cell wall digestibility in a topcross of a flint recombinant inbred line progeny harvested at silage stage. *Maydica* 53:1-9
- Riboulet C, Guillaumie S, Méchin V, Bosio M, Pichon M, Goffner D, Lapierre C, Pollet B, Lefèvre B, Martinant JP, Barrière Y, 2009. Kinetics of phenylpropanoid gene expression in maize growing internodes: Relationships with cell wall deposition. *Crop Sci* 49:211-223
- Riboulet C, Lefèvre B, Denoue D, Barrière Y, 2008b. Genetic variation in maize cell wall for lignin content, lignin structure, p-hydroxycinnamic acid content, and digestibility in a set of 19 lines at silage harvest maturity. *Maydica* 53:11-19
- Roadhouse FE, MacDougall D, 1956. A study of the nature of plant lignin by means of alkaline nitrobenzene oxidation. *Biochem J* 63: 33-39
- Roesler J, Krekel F, Amrhein N, Schmid J, 1997. Maize phenylalanine ammonia-lyase has tyrosine ammonia-lyase activity. *Plant Physiol* 113:175-179
- Rogers LA, Dubos C, Surman C, Willment J, Cullis IF, Mansfield SD, Campbell MM, 2005. Comparison of lignin deposition in three ectopic lignification mutants. *New Phytol* 168:123-140
- Rokhsar D, Chapman J, Mitros T, Goodstein D, 2009. Update on the Mo17 genome sequencing project *Maize Genetics Conference Abstracts* 51:T13
- Roussel V, Gibelin C, Fontaine AS, Barrière Y, 2002. Genetic analysis in recombinant inbred lines of early dent forage maize. II - QTL mapping for cell wall constituents and cell wall digestibility from per se value and top cross experiments. *Maydica* 47:9-20
- Rushton PJ, Somssich IE, Ringler P, Shen QJ, 2010. WRKY transcription factors. *Trends Plant Sci* 15:247-258
- Salse J, 2012. In silico archeogenomics unveils modern plant genome organisation, regulation and evolution. *Curr Opin Plant Biol* 15(2):122-30
- Salvi S, Sponza G, Morgante M, Tomes D, Niu X, Fengler KA, Meeley R, Ananiev EV, Svitashv S, Bruggemann E, Li B, Hainey CF, Radovic S, Zaina G, Rafalski JA, Tingey SV, Miao GH, Phillips RL, Tuberosa R, 2007. Conserved noncoding genomic sequences associated with a flowering-time quantitative trait locus in maize. *Proc Natl Acad Sci USA* 104:11376-11381
- Samuels AL, Rensing KH, Douglas CJ, Mansfield SD, Dharmawardhana DP, Ellis BE, 2002. Cellular machinery of wood production: differentiation of secondary xylem in *Pinus contorta* var. *latifolia*. *Planta* 216:72-82
- Sanchez-Fernandez R, Davies TGE, Coleman JOD, Rea PA, 2001. The *Arabidopsis thaliana* ABC protein superfamily, a complete inventory. *J Bio Chem* 276:30231-30244
- SanMiguel P, Gaut BS, Tikhonov A, Nakajima Y, Bennetzen JL, 1998. The paleontology of intergene retrotransposons of maize. *Nat Genet* 20:43-45

- Sasaki S, Baba K, Nishida T, Tsutsumi Y, Kondo R, 2006. The cationic cell-wall-peroxidase having oxidation ability for polymeric substrate participates in the late stage of lignification of *Populus alba* L. *Plant Molecular Biology* 62:797-807
- Sato Y, Whetten RW, 2006. Characterization of two laccases of loblolly pine (*Pinus taeda*) expressed in tobacco BY-2 cells. *J Plant Res* 119:581-588
- Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, Zhang J, Fulton L, Graves TA, Minx P, Reily AD, Courtney L, Kruchowski SS, Tomlinson C, Strong C, Delehaunty K, Fronick C, Courtney B, Rock SM, Belter E, Du F, Kim K, Abbott RM, Cotton M, Levy A, Marchetto P, Ochoa K, Jackson SM, Gillam B, Chen W, Yan L, Higginbotham J, Cardenas M, Waligorski J, Applebaum E, Phelps L, Falcone J, Kanchi K, Thane T, Scimone A, Thane N, Henke J, Wang T, Ruppert J, Shah N, Rotter K, Hodges J, Ingenthron E, Cordes M, Kohlberg S, Sgro J, Delgado B, Mead K, Chinwalla A, Leonard S, Crouse K, Collura K, Kudrna D, Currie J, He R, Angelova A, Rajasekar S, Mueller T, Lomeli R, Scara G, Ko A, Delaney K, Wissotski M, Lopez G, Campos D, Braidotti M, Ashley E, Golser W, Kim H, Lee S, Lin J, Dujmic Z, Kim W, Talag J, Zuccolo A, Fan C, Sebastian A, Kramer M, Spiegel L, Nascimento L, Zutavern T, Miller B, Ambroise C, Muller S, Spooner W, Narechania A, Ren L, Wei S, Kumari S, Faga B, Levy MJ, McMahan L, Van Buren P, Vaughn M,W, Ying K, Yeh C,-T, Emrich S,J, Jia Y, Kalyanaraman A, Hsia A-P, Barbazuk WB, Baucom RS, Brutnell TP, Carpita NC, Chaparro C, Chia J-M, Deragon J-M, Estill JC, Fu Y, Jeddelloh JA, Han Y, Lee H, Li P, Lisch DR, Liu S, Liu Z, Nagel DH, McCann MC, SanMiguel P, Myers AM, Nettleton D, Nguyen J, Penning BW, Ponnala L, Schneider KL, Schwartz DC, Sharma A, Soderlund C, Springer NM, Sun Q, Wang H, Waterman M, Westerman R, Wolfgruber TK, Yang L, Yu Y, Zhang L, Zhou S, Zhu Q, Bennetzen JL, Dawe RK, Jiang J, Jiang N, Presting GG, Wessler SR, Aluru S, Martienssen RA, Clifton SW, McCombie WR, Wing RA, Wilson RK, 2009. The B73 maize genome: complexity, diversity, and dynamics. *Science* 326:1112-1115
- Schopfer P, 1996. Hydrogen peroxide-mediated cell-wall stiffening in vitro in maize coleoptiles. *Planta* 199:43-49
- Schulte D, Ariyadasa R, Shi B, Fleury D, Saski C, Atkins M, Dejong P, Wu CC, Graner A, Langridge P, Stein N, 2011. BAC library resources for map-based cloning and physical map construction in barley (*Hordeum vulgare* L.). *BMC Genomics* 12:247
- Scobbie L, Russell W, Provan GJ, Chesson A, 1993. The newly extended maize internode: a model for the study of secondary cell wall formation and consequences for digestibility. *J Sci Food Agric* 61:217-225
- Scott-Craig JS, Casida JE, Poduje L, Walton JD, 1998. Herbicide safener-binding protein of maize - Purification, cloning, and expression of an encoding cDNA. *Plant Physiol* 116:1083-1089
- Sels J, Mathys J, De Coninck BM, Cammue BP, De Bolle MF, 2008. Plant pathogenesis-related (PR) proteins: A focus on PR peptides. *Plant Physiol Biochem* 46:941-950
- Sewalt V, Ni W, Blount JW, Jung HG, Masoud SA, Howles PA, Lamb C, Dixon RA, 1997. Reduced lignin content and altered lignin composition in transgenic tobacco down-regulated in expression of phenylalanine ammonia-lyase or cinnamate 4-hydroxylase. *Plant Physiol* 115:41-50
- Sewell MM, Davis MF, Tuskan GA, Wheeler NC, Elam CC, Bassoni DL, Neale DB, 2002. Identification of QTLs influencing wood property traits in loblolly pine (*Pinus taeda* L.). II. Chemical wood properties. *Theor Appl Genet* 104:214-222
- Shane MW, McCully ME, Canny MJ, 2000. The Vascular System of Maize Stems Revisited: Implications for Water Transport and Xylem Safety. *Annals Botany* 86:245-258
- Shen H, Yin YB, Chen F, Xu Y, Dixon RA, 2009. A bioinformatic analysis of NAC genes for plant cell wall development in relation to lignocellulosic bioenergy production. *BioEnergy Research* 2:217-232

- Shi C, Uzarowska A, Ouzunova M, Landbeck M, Wenzel G, Lübberstedt T, 2007. Identification of candidate genes associated with cell wall digestibility and eQTL (expression quantitative trait loci) analysis in a Flint x Flint maize recombinant inbred line population. *BMC Genomics* 8:22
- Shi H, Kim YS, Guo Y, Stevenson B, Zhu JK, 2003. The Arabidopsis SOS5 locus encodes a putative cell surface adhesion protein and is required for normal cell expansion. *Plant Cell* 15:19-32
- Shin JH, Jeong DH, Park MC, An G, 2005. Characterization and transcriptional expression of the alpha-expansin gene family in rice. *Mol Cells* 20:210-218
- Sieburth LE, Deyholos MK, 2005. Vascular development: the long and winding road. *Curr Opin Plant Biol.* 9:48-54
- Siefers N, Dang KK, Kumimoto RW, Bynum IV WE, Tayrose G, Holt BF 3rd, 2009. Tissue-specific expression patterns of Arabidopsis NF-Y transcription factors suggest potential for extensive combinatorial complexity. *Plant Physiol* 149:625-641
- Sindhu A, Langewisch T, Olek A, Multani DS, Mccann MC, Vermerris W, Carpita NC, Johal G, 2007. Maize Brittle stalk2 encodes a COBRA-like protein expressed in early organ development but required for tissue flexibility at maturity. *Plant Physiol* 145:1444-1459
- Skibbe D, Liu F, Wen T, Yandeu M, Cui X, Cao J, Simmons C, Schnable P, 2002. Characterization of the aldehyde dehydrogenase gene families of *Zea mays* and Arabidopsis. *Plant Mol Biol* 48:751-764
- Sonbol FM, Fornalé S, Cappellades M, Encina A, Tourino S, Torres JL, Rovira P, Ruel K, Puigdomenech P, Rigau J, Caparrós-Ruiz D, 2009. The maize ZmMYB42 represses the phenylpropanoid pathway and affects the cell wall structure, composition and degradability in Arabidopsis thaliana. *Plant Mol Biol* 70:283-296
- Springer NM, Stupar RM, 2007. Allelic variation and heterosis in maize: How do two halves make more than a whole? *Genome Res* 17:264-275
- Springer NM, Ying K, Fu Y, Ji T, Yeh CT, Jia Y, Wu W, Richmond T, Kitzman J, Rosenbaum H, Iniguez AL, Barbazuk WB, Jeddeloh JA, Nettleton D, Schnable PS, 2009. Maize inbreds exhibit high levels of copy number variation (CNV) and presence/absence variation (PAV) in genome content. *PLoS Genet* 5:e1000734
- Srivastava R, Liu JX, Guo H, Yin Y, Howell SH, 2009. Howell Regulation and processing of a plant peptide hormone, AtRALF23, in Arabidopsis. *Plant J* 59:930-939
- Struik PC, 1983. Physiology of forage maize (*Zea mays* L.) in relation to its production and quality. Ph. Dissertation, Agricultural University, 6700 GW Wageningen, The Netherlands, pp 1-252
- Studer A, Zhao Q, Ross-Ibarra J, Doebley J, 2011. Identification of a functional transposon insertion in the maize domestication gene *tb1*. *Nat Genet* 43:1160-1163
- Studer AJ, Doebley JF, 2011. Do large effect QTL fractionate? A case study at the maize domestication QTL *teosinte branched1*. *Genetics* 188(3):673-681
- Swanson-Wagner RA, Eichten SR, Kumari S, Tiffin P, Stein JC, Ware D, Springer NM, 2010. Pervasive gene content variation and copy number variation in maize and its undomesticated progenitor. *Genome Res* 20:1689-1699
- Taboada A, Novo-Uzal E, Flores G, Loureda M, Ros Barceló A, Masa A, Pomar F, 2010. Digestibility of silages in relation to their hydroxycinnamic acid content and lignin composition. *J Sci Food Agric* 90:1155-1162
- Talbert PB, Adler HT, Parks DW, Comai L, 1995. The REVOLUTA gene is necessary for apical meristem development and for limiting cell divisions in the leaves and stems of Arabidopsis thaliana. *Development* 121:2723-2735

- Tamagnone L, Merida A, Parr A, Mackay S, Culianez-Marcia FA, Roberts K, Martin C, 1998. The AmMYB308 and AmMYB330 transcription factors from *Antirrhinum* regulate phenylpropanoid and lignin biosynthesis in transgenic tobacco. *Plant Cell* 10:135-154
- Tamasloukht B, Lam MS-JWQ, Martinez Y, Tozo K, Barbier O, Jourda C, Jauneau A, Borderie, G, Balzergue S, Renou JP, Huguet S, Martinant JP, Tatout C, Lapierre C, Barrière Y, Goffner D, Pichon M, 2011. Characterization of a cinnamoyl-CoA reductase 1 (CCR1) mutant in maize: effects on lignification, fibre development, and global gene expression. *J Exp Bot* 62:3837-3848
- Tenaillon MI, Sawkins MC, Long AD, Gaut RL, Doebley JF, Gaut BS, 2001. Patterns of DNA sequence polymorphism along chromosome 1 of maize (*zea mays* ssp. *mays* L.). *Proc Natl Acad Sci USA* 98:9161-9166
- Terashima N, Fukushima K, 1993. Comprehensive model of the lignified plant cell wall. In: Jung HG, Buxton D, Hatfield R, Ralph J (Eds) *Forage Cell Wall Structure and Digestibility*, Madison, Wisconsin, pp 247-270
- The Angiosperm Phylogeny Group, 2009. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Botanical Journal of the Linnean Society* 161:105-121
- Thimm O, Blasing O, Gibon Y, Nagel A, Meyer S, Kruger P, Selbig J, Muller LA, Rhee SY, Stitt M, 2004. MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. *Plant J* 37:914-939
- Thomas J, Guillaumie S, Verdu C, Denoue D, Pichon M, Barrière Y, 2010. Cell wall phenylpropanoid-related gene expression in early maize recombinant inbred lines differing in parental alleles at a major lignin QTL position. *Mol Breeding* 25:105-124
- Thumma B, Southerton S, Bell J, Owen J, Henery M, Moran G, 2010. Quantitative trait locus (QTL) analysis of wood quality traits in *Eucalyptus nitens*. *Tree Genet Genomes* 6:305-317
- Tokunaga N, Kaneta T, Sato S, Sato Y, 2009. Analysis of expression profiles of three peroxidase genes associated with lignification in *Arabidopsis thaliana*. *Physiol Plantarum* 136:237-249
- Tolivia D, Tolivia J, 1987. Fasga: a new polychromatic method for simultaneous and differential staining of plants tissues. *Journal of Microscopy* 148:113-117
- Tomlinson PB, 1995. Non homology of vascular organization in monocotyledons and dicotyledons. In: Rudall PJ, Cribbb PJ, Cutler DF, Humphries CJ (Eds) *Monocotyledons, Systematic and Evolution*, Royal Botanic Gardens, Kew, pp 589-622
- Tripathi P, Rab ara RC, Langum TJ, Boken AK, Rushton DL, Boomsma DD, Rinerson CI, Rabara J, Reese RN, Chen X, Rohila JS, Rushton PJ, 2012. The WRKY transcription factor family in *brachypodium distachyon*. *BMC Genomics* 13:270
- Truntzler M, Barrière Y, Sawkins MC, Lespinasse D, Betran J, Charcosset A, Moreau L, 2010. Meta-analysis of QTL involved in silage quality of maize and comparison with the position of candidate genes. *Theor Appl Genet* 121:1465-1482
- Utz H, Melchinger A, 1996. PLABQTL: a program for composite interval mapping of QTL. *J Agric Genomics* 2:1-6
- van der Graaff E, Laux T, Rensing SA, 2009. The WUS homeobox-containing (WOX) protein family. *Genome Biology* 10:248
- Venables W, Ripley BB, 1994. *Modern applied statistics with Splus*. Springer-Verlag, New York
- Vielle-Calzada JP, Martínez de la Vega O, Hernández-Guzmán G, Ibarra-Laclette E, Alvarez-Mejía C, Vega-Arreguín JC, Jiménez-Moraila B, Fernández-Cortés A, Corona-Armenta G, Herrera-Estrella L, Herrera-Estrella A, 2009. The Palomero genome suggests metal effects on domestication. *Science* 326:1078

- Vignols F, Rigau J, Torres MA, Capellades M, Puigdomenech P, 1995. The brown midrib 3 (bm3) mutation in maize occurs in the gene encoding caffeic acid O-methyltransferase. *Plant Cell* 7:407-416
- Vigouroux Y, Jaqueth JS, Matsuoka Y, Smith OS, Beavis WD, Smith JSC, Doebley J, 2002. Rate and pattern of mutation at microsatellite loci in maize. *Mol Biol Evol* 19:1251-1260
- Vittorioso P, Cowling R, Faure JD, Caboche M, Bellini C, 1998. Mutation in the Arabidopsis PASTICCINO1 gene, which encodes a new fk506-binding protein-like protein, has a dramatic effect on plant development. *Mol Cell Biol* 18(5):3034-3043
- Vogel J, 2008. Unique aspects of the grass cell wall. *Curr Opin Plant Biol* 11:301-307
- Vroh Bi I, McMullen MD, Sanchez-Villeda H, Schroeder S, Gardiner J, Soderlund C, Wing R, Fang Z, Coe EH, 2006. Single nucleotide polymorphisms and insertion-deletions for genetic markers and anchoring the maize fingerprint contig physical map. *Crop Sci* 46:12-21
- Wang H, Avci U, Nakashima J, Hahn MG, Chen F, Dixon RA, 2010. Mutation of WRKY transcription factors initiates pith secondary wall formation and increases stem biomass in dicotyledonous plants. *Proceedings Natl Acad Sci USA* 107:22338-22343
- Wang H, Dixon RA, 2012. On-Off switches for secondary cell wall biosynthesis. *Molecular Plant* 5:297-303
- Wang J, Zhu ML, Wei ZM, 2008. Cotton laccase gene overexpression in transgenic *Populus alba* var. *pyramidalis* and its effects on the lignin biosynthesis in transgenic plants. *Fen Zi Xi Bao Sheng Wu Xue Bao* 41:11-18 (Article in Chinese)
- White S, Doebley J, 1998. Of genes and genomes and the origin of maize. *Trends Genet* 14:327-332
- Wilkins O, Nahal H, Foong J, Provart NJ, Campbell MM, 2009. Expansion and diversification of the *Populus* R2R3-MYB family of transcription factors. *Plant Physiol* 149:981-993
- Wilson JR, Hatfield RD, 1997. Structural and chemical changes of cell wall types during stem development: Consequences for fibre degradation by rumen microflora. *Aust J Agric Res* 48:165-180
- Wilson JR, Mertens DR, 1995. Cell-wall accessibility and cell structure limitations to microbial digestion of forage. *Crop Sci* 35:251-259
- Wissenbach M, Uberlacker B, Vogt F, Becker D, Salamini F, Rohde W, 1993. MYB genes from *Hordeum vulgare* - Tissue-specific expression of chimeric MYB Promoter/Gus genes in transgenic tobacco. *Plant J* 4:411-422
- Wolfe KH, Gouy M, Yang YW, Sharp PM, Li WH, 1989. Date of the monocot-dicot divergence estimated from chloroplast DNA sequence data. *Proc Natl Acad Sci* 86:6201-6205
- Wu AM, Rihouey C, Seveno M, Hornblad E, Singh SK, Matsunaga T, Ishii T, Lerouge P, Marchant A, 2009. The Arabidopsis IRX10 and IRX10-LIKE glycosyltransferases are critical for glucuronoxylan biosynthesis during secondary cell wall formation. *Plant J* 57:718-731
- Wu KL, Guo ZJ, Wang HH, Li J, 2005. The WRKY family of transcription factors in rice and Arabidopsis and their origins. *DNA Research* 12:9-26
- Xu S, 2003. Theoretical basis of the Beavis effect. *Genetics* 165:2259-2268
- Yamaguchi M, 2010. Transcriptional regulation of secondary wall formation controlled by NAC domain proteins. *Plant Biotechnol* 242:237-242
- Yamaguchi M, Kubo M, Fukuda H, Demura T, 2008. Vascular-related nac-domain7 is involved in the differentiation of all types of xylem vessels in arabidopsis roots and shoots. *Plant J* 55:652-664
- Yamaguchi M, Ohtani M, Mitsuda N, Kubo M, Ohme-Takagi M, Fukuda H, Demura T, 2010. VND-INTERACTING2, a NAC domain transcription factor, negatively regulates xylem vessel formation in Arabidopsis. *Plant Cell* 22:1249-1263

- Yan J, Yang X, Shah T, Sánchez-Villeda H, Li J, Warburton M, Zhou Y, Crouch JH, Xu Y, 2010. High-throughput SNP genotyping with the GoldenGate assay in maize. *Mol Breeding* 25(3):441-451
- Yanagisawa S, 2004. Dof domain proteins: plant-specific transcription factors associated with diverse phenomena unique to plants. *Plant Cell Physiol* 45:386-391
- Yang C, Xu Z, Song J, Conner K, Vizcay Barrena G, Wilson ZA, 2007. Arabidopsis MYB26/MALE STERILE35 regulates secondary thickening in the endothecium and is essential for anther dehiscence. *The Plant Cell* 19:534-548
- Yang T, Xue L, An L, 2007. Functional diversity of miRNA in plants. *Plant Sci* 172:423-432
- Yang TJ, Yu Y, Nah G, Atkins M, Lee S, Frisch D, Wing RA, 2003. Construction and utility of 10-kb libraries for efficient clone-gap closure for rice genome sequencing. *Theor Appl Genet* 107:652-660
- Yang YS, Strittmatter SM, 2007. The reticulons: a family of proteins with diverse functions. *Genome Biology* 8(12):234
- Yazaki K, 2005. Transporters of secondary metabolites. *Curr Opin Plant Biol* 8:301-307
- Yeung E, 1998. A beginner's guide to the study of plant structure in Tested studies for laboratory teaching, Volume19 (S.J.Karcher, editor) 365 pages
- Yim YS, Moak P, Sanchez-Villeda H, Musket TA, Close P, Klein PE, Mullet JE, McMullen MD, Fang Z, Schaeffer ML, Gardiner JM, Coe EH Jr, Davis GL, 2007. A BAC pooling strategy combined with PCR-based screenings in a large, highly repetitive genome enables integration of the maize genetic and physical maps. *BMC Genomics* 8:47
- Yin T, Zhang X, Gunter L, Priya R, Sykes R, Davis M, Wulschleger SD, Tuskan GA, 2010. Differential detection of genetic Loci underlying stem and root lignin content in Populus. *PLoS One* 5(11):e14021
- Yokoyama R, Nishitani K, 2004. Genomic basis for cell-wall diversity in plants. A comparative approach to gene families in rice and Arabidopsis. *Plant Cell Physiol* 45:1111-1121
- Yoshida-Shimokawa T, Yoshida S, Kakegawa K, Ishii T, 2001. Enzymic feruloylation of arabinoxylan-trisaccharide by feruloyl-CoA: arabinoxylan-trisaccharide O-hydroxycinnamoyl transferase from *Oryza sativa*. *Planta* 212:470-474
- You FM, Luo MC, Xu K, Deal KR, Anderson OD, Dvorak J, 2010. A new implementation of high-throughput five dimensional clone pooling strategy for BAC library screening. *BMC Genomics* 11:692
- Zeng ZB, 1994. Precision mapping of quantitative trait loci. *Genetics* 136:1457-1468
- Zhang BH, Pan XP, Anderson TA, 2006. Identification of 188 conserved maize microRNAs and their targets. *Febs Letters* 580:3753-3762
- Zhang G, Liu X, Quan Z, Cheng S, Xu X, Pan S, Xie M, Zeng P, Yue Z, Wang W, Tao Y, Bian C, Han C, Xia Q, Peng X, Cao R, Yang X, Zhan D, Hu J, Zhang Y, Li H, Li H, Li N, Wang J, Wang C, Wang R, Guo T, Cai Y, Liu C, Xiang H, Shi Q, Huang P, Chen Q, Li Y, Wang J, Zhao Z, Wang J, 2012. Genome sequence of foxtail millet (*Setaria italica*) provides insights into grass evolution and biofuel potential. *Nat Biotechnol* 13:549-554
- Zhang L, Chia JM, Kumari S, Stein JC, Liu Z, Narechania A, Maher CA, Guill K, McMullen MD, Ware D, 2009. A Genome-Wide Characterization of MicroRNA Genes in Maize. *PLoS Genet* 5(11):e1000716
- Zhang M, Zhang B, Qian Q, Yu Y, Li R, Zhang J, Liu X, Zeng D, Li J, Zhou Y, 2010. Brittle Culm 12, a dual-targeting kinesin-4 protein, controls cell-cycle progression and wall properties in rice. *Plant J* 63:312-328

- Zhang Y, Culhaoglu T, Pollet B, Melin C, Denoue D, Barrière Y, Baumberger S, Méchin V, 2011. Impact of Lignin Structure and Cell Wall Reticulation on Maize Cell Wall Degradability. *J Agric Food Chem* 59:10129-10135
- Zhao H, Zhang Y, Dai H, Zhang Y, Shen Y, 2011. CARM1 Mediates Modulation of Sox2. *PLoS One* 6: e27026
- Zhao Q, Dixon RA, 2011. Transcriptional networks for lignin biosynthesis: more complex than we thought? *Trends Plant Sci* 16:227-233
- Zhao Q, Wang H, Yin Y, Xu Y, Chen F, Dixon RA, 2010. Syringyl lignin biosynthesis is directly regulated by a secondary cell wall master switch. *Proc Natl Acad Sci USA* 107:14496-14501
- Zheng Z, Qamar SA, Chen Z, Mengiste T, 2006. Arabidopsis WRKY33 transcription factor is required for resistance to necrotrophic fungal pathogens. *Plant J* 48:592-605
- Zhong R, Burk DH, Morrison WH, Ye ZH, 2002. A kinesin-like protein is essential for oriented deposition of cellulose microfibrils and cell wall strength. *Plant Cell* 14:3101-3117
- Zhong R, Lee C, McCarthy RL, Reeves CK, Jones EG, Ye ZH, 2011. Transcriptional activation of secondary wall biosynthesis by rice and maize NAC and MYB transcription factors. *Plant Cell Physiol* 52:1856-1871
- Zhong R, Lee C, Ye ZH, 2010a. Functional characterization of poplar wood-associated NAC domain transcription factors. *Plant Physiology* 152:1044-1055
- Zhong R, Lee C, Ye ZH, 2010b. Global analysis of direct targets of secondary wall NAC master switches in Arabidopsis. *Molecular Plant* 3:1087-1103
- Zhong R, Lee C, Ye ZH, 2010c. Evolutionary conservation of the transcriptional network regulating secondary cell wall biosynthesis. *Trends Plant Sci* 15:625-632
- Zhong R, Lee C, Zhou J, McCarthy RL, Ye ZH, 2008. A battery of transcription factors involved in the regulation of secondary cell wall biosynthesis in Arabidopsis. *Plant Cell* 20:2763-2782
- Zhong R, Richardson EA, Ye ZH, 2007. The MYB46 transcription factor is a direct target of SND1 and regulates secondary wall biosynthesis in Arabidopsis. *Plant Cell* 19:2776-2792
- Zhong R, Ye ZH, 1999. IFL1, a gene regulating interfascicular fiber differentiation in Arabidopsis, encodes a homeodomain-leucine zipper protein. *Plant Cell* 11:2139-2152
- Zhong R, Ye ZH, 2009. Transcriptional regulation of lignin biosynthesis. *Plant Signaling Behavior* 4:1-7
- Zhong R, Ye ZH, 2010. The poplar PtrWNDs are transcriptional activators of secondary cell wall biosynthesis. *Plant Signaling Behavior* 5:469-72
- Zhong R, Ye ZH, 2012. MYB46 and MYB83 bind to the SMRE sites and directly activate a suite of transcription factors and secondary wall biosynthetic genes. *Plant Cell Physiol* 53:368-380
- Zhou J, Lee C, Zhong RQ, Ye ZH, 2009. MYB58 and MYB63 are transcriptional activators of the lignin biosynthetic pathway during secondary cell wall formation in Arabidopsis. *Plant Cell* 21:248-266
- Zhou L, Zhang J, Yan J, Song R, 2011. Two transposable element insertions are causative mutations for the major domestication gene teosinte branched 1 in modern maize. *Cell Res* 21:1267-1270

Author - Audrey COURTIAL

Title - Towards the identification of maize genes involved in lignified secondary cell wall degradability through the elucidation of QTLs with strong effects.

PhD co-supervisors - Jacqueline GRIMA-PETTENATI et Yves BARRIÈRE

Abstract

Discovering the genetic determinants of the lignified cell wall assembly in grasses is a major challenge for both basic research and for plant breeding based on marker-assisted selection. Cell wall degradability is a limiting factor of plant energy value for cattle feeding, as well as for the production of second-generation biofuel. The research conducted thus aimed at identifying genes involved in cell wall related traits, taking as model a cluster of strong effect QTLs located in the bin 6.06 of the maize recombinant inbred line (RIL) progeny F288 x F271.

Having shown that these QTL positions were located in a monomorphic area between the two parental lines, targeted densification of the genetic map revealed that these "ghost" QTLs correspond in fact to QTLs located on two close positions (bins 6.05 and 6.07). New major QTLs in bin 4.09 have also been detected. New QTL detection from new field experiments has also allowed to consolidate the involvement of *p*-hydroxycinnamic acids and of the lignin monomeric composition, in the variation of cell wall degradability.

In order to identify the candidate genes underlying these QTLs, expression studies and sequencing were undertaken, besides the *a priori* search for genes potentially involved in the lignified cell wall formation, from the bibliography. The expression studies between the F271 parental line and four RILs carrying favorable alleles for the cell wall degradability (F288) at the major QTLs of bin 6.06 allowed to highlight 360 differentially expressed genes. The targeted sequencing of BACs carrying the QTL region of interest, for F271 and F288, underlined the great polymorphism between these parental maize lines. Further sequencing targeted to bin 6.05 should contribute to identify the gene or genes responsible for these major QTLs in the RIL progeny F288 x F271. Candidate genes should be latter validated through functional analyses.

Key words - *Zea mays*, degradability, cell wall, lignins, *p*-hydroxycinnamic acids, QTL, candidate genes, transcriptomic, BAC, 454 sequencing.

Auteur - Audrey COURTIAL

Titre - Vers l'identification de gènes contrôlant la dégradabilité de la paroi secondaire lignifiée chez le maïs à travers l'élucidation de QTLs à effets forts.

Directeurs de thèse - Jacqueline GRIMA-PETTENATI et Yves BARRIÈRE

Lieu et date de soutenance - 30 Octobre 2012, Salle Marc Ridet, Campus INRA-Auzeville

Résumé

La dégradabilité des parois des plantes fourragères est un facteur limitant à la fois pour l'alimentation des ruminants et pour la production de biocarburants de seconde génération. Les recherches conduites ont donc visé à identifier les gènes contrôlant des propriétés des parois lignocellulosiques chez le maïs en prenant comme modèle un cluster de QTLs à effets forts localisé dans le bin 6.06 des lignées recombinantes (RILs) F288 x F271.

Après avoir mis en évidence que les positions de ces QTLs se situaient dans une région monomorphe entre les deux lignées parentales, une densification ciblée de la carte génétique a permis de révéler que ces QTLs "fantômes" correspondaient en fait à des QTLs localisés sur deux positions proches (bins 6.05 et 6.07). De nouveaux QTLs majeurs au bin 4.09 ont également été détectés. La détection de QTLs avec de nouvelles mesures phénotypiques a également permis de conforter l'implication des acides *p*-hydroxycinnamiques et de la composition monomériques des lignines dans la dégradabilité des parois.

Afin d'identifier les gènes candidats présents sous ces QTLs, des études d'expression et du séquençage ont été entreprises, en plus de la recherche *a priori* de gènes potentiellement impliqués dans la formation des parois lignifiées, à partir de la bibliographie. L'étude d'expression entre le parent F271 et quatre RILs porteuses des allèles favorables à la dégradabilité des parois (F288) aux QTLs majeurs du bin 6.06 a permis de mettre en évidence 360 gènes différentiellement exprimés. Le séquençage ciblé de BACs porteurs de la région d'intérêt chez F271 et F288, quant à lui, a souligné le très grand polymorphisme entre ces deux lignées. La poursuite du séquençage ciblé du bin 6.05 devrait contribuer à l'identification du gène ou des gènes responsables de ces QTL majeurs chez la famille RIL F288 x F271. Les gènes candidats devront également être validés grâce à une analyse fonctionnelle.

Mots clés - *Zea mays*, dégradabilité, paroi, lignines, acides *p*-hydroxycinnamiques, QTL, gènes candidats, transcriptome, BAC, séquençage.

Discipline administrative - Développement des plantes

Intitulés et adresses des laboratoires -

UMR 5546 UPS/CNRS, Laboratoire de Recherche en Sciences Végétales, Pôle de Biotechnologies Végétales, 24 Chemin de Borde Rouge, BP 42617 Auzeville, 31326 Castanet-Tolosan, France.

INRA, Unité de Génétique et d'Amélioration des Plantes Fourragères, Le Chêne, RD 150, BP 80006, 86600 Lusignan, France.