Different mutations in the ZmCAD2 gene underlie the maize brown-midrib1 (bm1) phenotype with similar effects on lignin characteristics and have potential interest for bioenergy production

Yves Barrière*, Hélène Chavigneau¹, Sabrina Delaunay¹, Audrey Courtial¹,²,³, Mickael Bosio⁴, Hervé Lassagne⁴, Jérémy Derory⁴, Catherine Lapierre⁵, Valérie Méchin⁵, Christophe Tatout⁶

¹INRA, Unité de Génétique et d’Amélioration des Plantes Fourragères, 86600 Lusignan, France
²Université de Toulouse, UPS, UMR5546, LRSV, BP 42617, Auzelle, 31326 Castanet-Tolosan, France
³CNRS, UMR 5546, BP 42617, Auzelle, 31326 Castanet-Tolosan, France
⁴Biogemma, site de la Garenne, CS 90126, 63720 Chappes, France
⁵INRA, Institut Jean-Pierre Bourgin, 78026 Versailles, France
*Corresponding author: E-mail: yves.barriere@lusignan.inra.fr
⁶current address - CNRS UMR6293, INSERM U1103, GRED, les Cézeaux, 24 Avenue des Landais, BP 80026 Aubière cedex, France

Abstract
The maize ZmCAD2 gene has been fully sequenced in several normal and bm1 maize lines, highlighting a large diversity of mutations underlying the bm1 phenotype. Mutations in three bm1 lines (F2bm1, A619bm1, and 511Jbm1) were found corresponding to short InDels inducing premature stop codons and truncated proteins. In two lines (511Kbm1 and 5803Cbm1), mutations were limited to an only SNP or to a few SNPs, modifying the catalytic sites, and likely inactivating the proteins. Results also established that the 5803Ibm7 mutant was in fact a bm1 mutant, with a sequence fully identical to the 5803Cbm1 sequence. The two new F7803bm1 (natural mutant) and Ev2210bm1 (transposon tagging Mutator investigations) both had a transposon insertion in the ZmCAD2 DNA, resulting in a truncated protein, even if the mRNA was produced. The biochemical characteristics of the Ev2210bm1 lignins corroborated the signature of CAD2 deficiency in plants, with the presence of aldehydes and atypical compounds and linkages. Considering lignin structure and content, CAD2 is likely a good target for the improvement of energy production based on maize and grass lignocellulosic biomass, including a greater susceptibility to environmentally friendly pretreatments, as it was shown in bmr sorghum. The interest in maize bm1 hybrids for cattle feeding also should be considered because there seem to be little or limited negative effects of CAD2 mutations on other agronomical traits.

Keywords: maize, brown midrib, bm1, lignin, ferulic acid, cell wall, degradability, biofuel, biogas

Introduction
Maize brown-midrib (bm) mutants exhibit a reddish-brown pigmentation of the leaf midrib and stalk pith, associated with lignified tissues, after the plants have about five expanded leaves. The first bm mutants were described and investigated successively by Kiesselbach (1922), Eyster (1926), and Jorgenson (1931). The latter established that mutations of these first mutants were “due to identical factors” and that the “bm character” segregated as “a simple Mendelian recessive” trait. This gene was later named bm1 when the bm2 mutant was described by Burnham and Brinks (1932). A little later, two other genes inducing the bm phenotype were described as bm3 (Emerson, 1935) and bm4 (Burnham, 1947). The bm2, bm3, and bm4 genes also originated from natural mutations and segregated as simple Mendelian recessive traits. The bm7 mutation was assigned to maize chromosome 5 (bin 5.04) in position 77.5 cM, close to the centromer (Jorgenson, 1931; MaizeGDB, www.maizegdb.org). The three other bm2, bm3, and bm4 mutants were later mapped in bins 1.11, 4.05, and 9.07, respectively (MaizeGDB database). No new maize bm mutations were characterized for nearly 60 years after the Burnham’s paper (1947), despite the mention of additional natural bm mutants in the MaizeGDB database. Recent allelic tests of these latter mutants highlighted the three novel bm5, bm6, and bm7 loci (Haney et al., 2008; Ali et al., 2010). While bm5 and bm7 mutations are still uncharacterized, the bm6 mutations have been mapped to a 180 kb region of bin 2.01. Ten underlying candidates were proposed (Chen et al., 2012a), including a gene encoding a zinc finger CCCH-type protein, of which different family members also colocalized with cell wall degradability QTLs in RIL progenies of Arabidopsis (Chavigneau et al., 2012) and maize (Courtial et al., unpublished data).

In the final step of maize monolignol biosynthesis, the cinnamyl alcohol dehydrogenase (CAD) protein(s) catalyze(s) the reduction of p-coumaryl, coniferyl, and sinapyl aldehydes to their corresponding alcohols, using NADPH as a cofactor, prior to their trans-
port to the wall and their polymerization into the lignin polymer. Two types of CAD genes are currently considered in plants, based on original investigations in eucalyptus. EgCAD1-type enzymes are short-chain alcohol dehydrogenases (Jornvall et al., 1995; Goffner et al., 1998), which are active as monomers on coniferaldehyde, but not sinapaldehyde (Hawkins and Boudet 1994; Damiani et al., 2005). A ZmCAD1 activity has been described in maize (Kanazawa et al., 1999), corroborating the existence of the two types of CAD enzymes also in (this) grass species. EgCAD2-type enzymes are conversely zinc-containing long-chain alcohol dehydrogenases, which are active as dimers, and are considered the predominant CAD proteins involved in lignification (Jornvall et al., 1987; Goffner et al., 1992; Hawkins and Boudet, 1994). In maize bm1 plants, the activity of a CAD enzyme was shown to be reduced by 60 - 70% in the above-ground organs and by 90 - 97% in roots (Halpin et al., 1998). These authors therefore considered that “bm1 is not a null mutation of the ZmCAD, but affects its expression, possibly through alterations in upstream or downstream non-coding regions”. Linkage analyses also showed that the ZmCAD locus was closely linked to RFLP markers corresponding to the position of the bm1 mutation. Halpin et al (1998) thus concluded “that ZmCAD is very likely to be allelic to bm1”. This maize ZmCAD gene associated with the bm1 mutation was latter shown to be orthologous to EgCAD2, and it has thus been named ZmCAD2 (Guillaumie et al, 2007b). However, a lack of complete specificity of their polyclonal antibody left the possibility of reactions with other CAD or CAD-like proteins, and/or only partial reactions against the targeted ZmCAD2 protein. In addition, gene expression, which was investigated based on northern analysis with degenerate primers based on the tobacco CAD2 protein sequence, of which specificity against the ZmCAD2 sequence is likely partial, indicated a residual CAD expression in bm1 leaves and stems. In addition, based on gene expression investigations with the MaizeWall macro-array, the ZmCAD2 gene was shown to be under-expressed in bm1 plantlets with a residual expression of 0.36 as compared to the control F2 line (Guillaumie et al, 2007a). This agreed closely with the residual activity described by Halpin et al (1998). In addition, two ZmCAD2-like genes and the ZmCAD1 gene were simultaneously under-expressed in F2bm1 plantlets. Similarly, based on Suppression Subtractive Hybridization (SSH) and micro-array data, Shi et al (2006) also established that several ZmCAD and CAD-like genes were down-regulated in young bm1 plants (5-7 week-old plants).

The maize ZmCAD2 gene (GRMZM2G179981) is also located in bin 5.04, in a downstream position (129.940 Mbp), upstream of the estimated position of the centromere. The latter gene only has distant paralogs, and at least one of them encodes a dihydroflavonol-4-reductase. The ZmCAD2 also has no very close paralogs, and several of these ZmCAD2-like genes are now annotated as encoding mannitol dehydrogenase enzymes (maize sequence database, www.maizesequence.org). As a tentative conclusion, ZmCAD1 and ZmCAD2 are likely the two ZmCAD genes primarily involved in normal constitutive lignification of maize.

The consequences of the bm1 mutation on maize lignin content and structure were first described forty years after the description of the mutation (Kuc and Nelson, 1964; Gee et al, 1968). Mature maize bm1 plants have a lignin content that is reduced by 10 to 20%, a slight decrease in ferulic acid (FA) esters and substantially reduced contents (about 40%) in p-coumaric (pCA) esters and FA ethers (Provan et al, 1997; Barrière et al, 2004a). The frequency of p-hydroxyphenyl (H), guaiacyl (G) and syringyl (S) thioacidolysis monomers was similar in bm1 and normal plants, showing that the bm1 mutation does not specifically affect one of the lignin units. However, the reduced recovery of thioacidolysis monomers reveals that the frequency of lignin units involved only in β-O-4 bonds was about 50% lower in bm1 plants than in lignins of normal plants, indicating that lignins of bm1 plants were substantially enriched in carbon-carbon inter-unit linkages (Halpin et al, 1998; Barrière et al, 2004a). Lignins of bm1 plants are also typified by a substantial incorporation of coniferaldehyde and, to a lower extent, of sinapaldehyde and p-hydroxybenzaldehyde into the polymer (Jacquet, 1997; Kim et al, 2002; Kim et al, 2003; Barrière et al, 2004a). This notable incorporation of p-hydroxy-cinnamaldehyde-derived compounds in bm1 lignins was in agreement with the under-expression of the ZmCAD2 gene.

Concerns over global climate changes, together with a growing worldwide demand for energy, have highlighted the crucial need for alternative resources to replace fossil fuels. Second-generation biofuels, based on lignocellulose materials, have opened up new avenues, including a large valorization of agricultural and woody residues. The latter do not compete with food supplies, unlike biofuels based on cereal or oleaginous grains. The European Union commissioner for “Climate Action” Connie Hedegaard indeed recently said that “we cannot morally afford to build a very big industry on something that is not good for the environment or for food prices”. 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, as well as with p-hydroxycinnamic acids in grasses. A similar situation has existed for decades in breeding programs towards higher digestibility and energy value of forage plants. Improvement of maize
energy value, for both cattle nutrition and industrial purposes, can be based on crosses and breeding with lines for which high values of cell wall degradability have been shown (Barrière et al., 2009; Barrière et al., 2010). However, the use of mutants with similarly higher cell wall degradability is also a relevant strategy in plant breeding. A mutant gene with major effect can indeed be more easily and rapidly backcrossed in elite lines based on marker assisted selection than several quantitative traits originating from one or more genetic resources.

Based on the B73 ZmCAD2 sequence available from the Maize Genomics Sequencing Project (MGSP, www.maizesequence.org), the objective of this work was to sequence the ZmCAD2 gene in bm1 lines in comparison with their normal counterpart, including the reference line F2bm1, several new bm1 mutants (Ali et al., 2010), and a new natural bm1 mutant identified in an INRA nursery. In order to strengthen the involvement of the ZmCAD2 gene in the bm1 phenotype, a novel maize mutant obtained from transposon tagging in the ZmCAD2 gene was simultaneously considered and characterized during these investigations. Transcriptomic approaches were also performed in order to estimate the respective effects of different bm1 mutations on ZmCAD gene expression. In addition, because conflicting results were observed in new allelism tests, the bm7 mutant was added to the sequencing investigations. Finally, the potential of the bm1 mutation was discussed for further improvement of plant energy value for both cattle feeding and biofuel production.

Materials and Methods

*Origins of bm1 and ZmCAD2 mutant lines*
The F2bm1 line was obtained after seven backcrosses by the early flint INRA F2 line of the original bm1 source (95033-47) supplied to INRA Lusignan by LF Bauman of Purdue University (West Lafayette, Indiana, USA) in 1972. This INRA F2bm1 genetic resource was also used in investigations by Halpin et al. (1998). Seeds of A619 and A619bm1 were supplied by N de Leon of University of Wisconsin (Madison) to INRA Lusignan in 2009. Seeds of the recently identified new bm1 mutants 5803Cbm1, 511Kbm1 (or 5803Lbm1) and 511K, 511Jbm1 (or 5803Jbm1) were obtained in 2010 and 2011 from the Maize Genetics Cooperation Stock Center (MGSCC, www.maizegdb.org). The ZmCAD2 mutant was isolated by the Bio-gemma company from transposon tagging investigations with the *Mutator* element, in the framework of the French national Génoplanete program. Seeds of this mutant (Zmcd2-m2210:Mu named thereafter Ev2210bm1), and the corresponding normal control seeds (Ev2210), resulted from three generations of self-pollinations after five generations of backcrossing (BCSS3) the original mutant plant with an elite line of the Limagrain company. The new INRA F7803bm was found in 2009 in an S7 generation of a flint prog-
Table 1 - Primer pairs designed for amplification of the whole DNA ZmCAD2 gene, and for expression studies, based on the B73 DNA sequence.

<table>
<thead>
<tr>
<th>Forward primer (5'-&gt; 3')</th>
<th>Reverse primer (5'-&gt; 3')</th>
<th>Length (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gene sequencing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZmCAD2_F1</td>
<td>GACCTCCCTCAGAAGACGAAA</td>
<td></td>
</tr>
<tr>
<td>ZmCAD2_F2</td>
<td>TAATTGCGGGGACACTCTGT</td>
<td></td>
</tr>
<tr>
<td>ZmCAD2_F3</td>
<td>CTCCGTCGACTGTCGACTAAG</td>
<td></td>
</tr>
<tr>
<td>ZmCAD2_F4</td>
<td>CTTATGTCGGTACGGCAAGAG</td>
<td></td>
</tr>
<tr>
<td>ZmCAD2_F5</td>
<td>CGACCTCGCTGGACTACATCA</td>
<td></td>
</tr>
<tr>
<td>ZmCAD1aF</td>
<td>CTTCCTCTTCCGCAGAACAGTGT</td>
<td></td>
</tr>
<tr>
<td>ZmCAD2cF</td>
<td>GGTCACCAACAGCTCCTACACTG</td>
<td></td>
</tr>
<tr>
<td>ZmCAD4bF</td>
<td>TCAAGCCCAACGGCAAGATG</td>
<td></td>
</tr>
<tr>
<td>ZmUbiquitinF</td>
<td>CATTGTGCCTGTGGTGAATC</td>
<td></td>
</tr>
</tbody>
</table>

Expression studies

<table>
<thead>
<tr>
<th>Forward primer (5'-&gt; 3')</th>
<th>Reverse primer (5'-&gt; 3')</th>
<th>Length (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZmCAD1aR</td>
<td>AGCTTTCGAGTGGCTCTTGGACG</td>
<td></td>
</tr>
<tr>
<td>ZmCAD2cR</td>
<td>CCGGATCTCTCACCACAACTCTC</td>
<td></td>
</tr>
<tr>
<td>ZmCAD4bR</td>
<td>CAGGTCCTTGGCTCCAAATGAG</td>
<td></td>
</tr>
<tr>
<td>ZmUbiquitinR</td>
<td>AACAGCAACACCCACACACCA</td>
<td></td>
</tr>
</tbody>
</table>

Primer design and PCR amplifications

Pairs of primers used for the ZmCAD2 gene sequencing were designed based on the available maize B73 sequence (GRMZM2G544562, www.maizesequence.org, Table 1). Fragments of nearly 0.4 to 1.5 kb were amplified, encompassing the 5’UTR, the complete coding region, 1,000 bp before start codon and 1,000 bp after stop codon, giving a ZmCAD2 investigated region close to 5 kb long. PCR reactions were performed in a final volume of 50 µl containing 1X PCR buffer and 0.5 µU Taq DNA Pol (both MP Biomedicals, ref EPTQX925), 200 µM of dNTPs, 0.32 µM of 5’ oligo, and 0.32 µM of 3’ oligo. Fifty ng of genomic DNA was used as template. The program comprised 4 min at 94°C, followed by 40 cycles of 30 s at 95°C, 1 min at 60°C, and 1 min at 72°C, followed by 6 min at 72°C.

Expression analyses

For all eight normal (F2, 511K, Ev2210) and bm1 (F2bm1, 511Kbm1, 5803Cb1m, 511Jbm1, Ev2210bm1) genotypes, plants were grown in a glasshouse at INRA Lusignan in spring 2011. Lignin expression analyses 4,031 and 4,116 (relative to the transcriptional start site in B73), respectively. In addition, amino acid sequencing of the ZmCAD2 gene in normal and bm1 near-isogenic lines

F2 and F2bm1 lines

The F2 ZmCAD2 allele differed from the B73 ZmCAD2 allele in the 3’ region, with 12 SNPs and two 54- and 30-bp long deletions starting from positions 4,031 and 4,116 (relative to the transcriptional start site in B73), respectively. In addition, amino
acid 349 is an asparagine in F2, while it is an aspartic acid in B73. The ZmCAD2 alleles in B73 and F2bm1 were overall similar, likely related to the fact that the original bm1 mutation used in F2 backcrossing occurred in a dent Corn Belt line and not in a European flint one. However, a two basepair (AC) insertion was identified in exon 3 of the F2bm1 line after the C base located in B73 CDS position 392 (Supplementary Figure 1). The ORF was consequently disrupted in F2bm1 by a premature stop codon, resulting in a predicted truncated protein of 147 amino acids, in comparison to the 367 amino acid long proteins in B73 and F2 (Supplementary Figure 2). Such a truncated ZmCAD2 protein, in which the site involved in enzyme binding to its substrate, a catalytic site, and several binding sites to cofactors, were missing is very likely to be non-functional, even if produced. The AC base insertion may therefore explain the bm1 phenotype of the F2bm1 line. In addition, in F2, F2bm1, as in all investigated lines, in B73 CDS position 951, there was a synonymous substitution of an A base (B73) by a G base (all lines), the GGA and GGG codons both encoding for a glycine.

**A619 and A619bm1 lines**

Allele sequencing was only investigated in the area overlapping the AC insertion shown in F2bm1. The same mutation was highlighted in A619bm1 as the one identified in F2bm1. This suggests that the backcrosses of the bm1 mutation in the A619 line were made with the same bm1 mutant allele as the one used for the F2bm1 line, which is, given the origin of these lines, likely the original bm1-reference (bm1-ref) allele.

**511Kbm1 line**

A CGGG four base insertion was identified in the ZmCAD2 exon 3 of the 511Kbm1 line, which occurred after the CG bases located in positions 366-367 of the B73 CDS (Supplementary Figure 1, bm1-J allele). As observed in F2bm1, this insertion disrupted the ORF and induced a 127 amino acid long truncated protein (Supplementary Figure 2), which is very likely non-functional, even if produced.

**511Kbm1 line**

Allelic variation in the ZmCAD2 CDS of the 511Kbm1 line, in comparison to B73, was reduced to only one SNP in exon 3 substituting the GGA codon into a GAA codon (CDS bp 401), thus substituting a glycine (G) by a glutamic acid (E) as the 134th amino acid of the ZmCAD2 protein (Supplementary Figures 1 and 2, bm1-K allele). This unique mutation occurred in an area encoding the between-species conserved 130PTQGFA136 amino acid motif (130PTQGEFA136 in the mutant), located just before the j9 strand. However, a 21 bp deletion was also shown in the 5’UTR, in position 143-123 bp before the B73 ATG start codon, in which was inserted a 5 bp substitution establishing an extra TATATA box 128 bp before the ZmCAD2 ATG start codon of the 511K line, in addition to the TATATA box located 40 bp upstream. These two events were shown to be specific to the 511Kbm1 line and were not present in the 511K near-isogenic normal line. It was thus considered that either changes in protein structures inactivated its catalytic activity, or less likely that the transcription was reduced to a quasi-null level due to the extra TATA box.

**5803Cbm1 line**

Neither insertions nor deletions were shown in the ZmCAD2 CDS of the 5803Cbm1 line, in comparison to B73, but the two sequences differed by two synonymous and three non-synonymous SNPs. Synonymous SNPs occurred in CDS positions 192 bp (CCT/CCC, proline) and 237 bp (GGG/GGA, glycine). The first non-synonymous SNP was shown in exon 1 where a TCC codon was substituted by an ACC codon (CDS bp 76), encoding a threonine (T) instead of a serine (S) as the ZmCAD2 26th amino acid (Supplementary Figures 1 and 2). The second non-synonymous SNP occurred in exon 3, where the TCC codon was changed into TGC (CDS bp 410), encoding a cysteine (C) in place of a serine (S) as the 137th amino acid. The latter mutation, and the one shown in exon 1, likely did not greatly change activities of the encoded enzyme, as they both occurred in between-species genomic unconstrained areas (Youn et al., 2006). The third mutation also occurred in exon 3, with a codon CAC changed into CGC (CDS bp 205), which substituted a histidine (H) by an arginine (R) as 69th amino acid in the conserved amino acid motif 50GHEVQXXXGXXV52 leading to a 53GREVG motif in the mutant (Supplementary Figures 1 and 2, bm1-C allele). The latter motif is involved in the binding of the zinc ion at the catalytic site of the enzyme, and the GHE conserved amino acid sequence also corresponded to the area separating the j4 and j5 strands (Youn et al., 2006). This H to R mutation probably alters the catalytic activity of the encoded enzyme, and is consequently likely the cause of the bm1 phenotype. In addition, a very important polymorphism was shown in the ZmCAD2 line first intron of the 5803Cbm1 line, which differed from B73 first intron by 54 SNP and 18 Indels. These changes could also modify the ZmCAD2 gene expression in 5803Cbm1 mutant plants. Finally, the 5’UTR of the ZmCAD2 gene in 5803Cbm1 mutant line exhibited the same 15 bp deletion and extra TATATA box as observed in the 511Kbm1 mutant line.

**5803bm7 line**

The sequence of the ZmCAD2 gene in the 5803bm7 line was found to be fully identical to the sequence of the 5803Cbm1 line. This result was in agreement with new allelism tests done at INRA Lusignan showing that the 5803bm7 line gave hybrids with brown midribs in crosses with F2bm1 (Figure 1). Allelism between bm1 and bm7 was corroborated by another allelism test done by Hongjun Liu (Iowa state University, pers com) and it was also in agreement with unpublished previous observations of Sarah Hake (2007, pers com). The mutation in the 5803bm7
The F7803bm1 line is definitely allelic to \textit{bm1} (bm1-7 allele) and this fact is now specified in the maizegdb database (www.maizegdb.org, locus bm7).

\textbf{F7803bm1 line}

Brown-midrib plants were obtained in crosses of F7803bm with F2bm1 (Figure 1), but not in crosses with F2bm2, F2bm3, or F2bm4, and similarly not with the newly described \textit{bm5} and \textit{bm6} lines. However, the F7803bm line also gave plants with brown midribs in crosses with 5803bm7 (Figure 1). This fact further validated the allelism of the two \textit{bm1} and \textit{bm7} mutations. In contrast to investigations with the five previous \textit{bm1} mutants, the sequence of the \textit{ZmCAD2} gene in the F7803bm1 line was not fully obtained, with a failure in amplifying the whole second exon. PCR products corresponding to the second exon were small, with \textit{ZmCAD2} homologies reduced to a few base pairs, which most often cor-

\textbf{Figure 1} - Allelism tests between 5807bm7 and F2bm1 (A), F7803bm and F2bm1 (B), F7803bm and 5803bm7 (C), Ev2210bm and F2bm1 (E, F), all four crosses giving brown-midribs, and F7803bm and F2bm3 (D), giving normal midribs.
responded to primer sequence. The forward primer used was located 459 bp upstream of exon 2, and the reverse primer was located 629 bp downstream of exon 2 (these two primers enabled amplification of exon 2 in normal lines). The most probable situation in F7803bm1 would therefore be an insertion of a (retro-) transposon element in the exon 2 area, preventing PCR. In addition, there was an amino acid substitution in exon 4, with a glutamic acid replaced by an aspartic acid, corresponding in B73 CDS position 924 to a substitution of a G nucleotide by a T nucleotide (Supplementary Figures 1 and 2, allele bm1-F). Glutamic and aspartic acids are both diacid amino acids, and such a change is not expected to be responsible for the bm1 phenotype. Nevertheless, the probable insertion of a (retro-) transposon close to exon 2 would very likely result in a protein disrupted before the region encoded by exon 4.

Ev2210bm1 line

In contrast to previous spontaneously occurring bm1 mutations, the Ev2210 ZmCAD2 bm mutant line was obtained from transposon tagging investigations with the Mutator element. Allelism tests showed that brown-midrib plants were obtained in crosses with F2bm1, F2bm3, or F2bm4, corroborating the involvement of ZmCAD2 mutations in bm1 phenotype. As it was the case for F7803bm1, the full sequence of the ZmCAD2 gene in the Ev2210bm1 line was also not obtained. However, 24 bp of the upstream part and 26 bp of the downstream part of the Mutator element sequences were obtained, showing that the Mutator element was located in this allele in exon 4, between position 678 and 679 bp of the B73 CDS. Lengths of a single copy of Mutator transposons are nearly 1,500 bp long, and they have a 220 bp long conserved terminally inverted repeats (TIRs). Such a long insertion, and chiefly the presence of the TIRs which cause very stable hairpin structures, would prevent an efficient PCR running. In any case, the Mutator element insertion induced a stop codon and disrupted the ORF, resulting in a 237 amino acid long truncated protein (Supplementary Figures 1 and 2, allele bm1-M), which is likely to be inactive if it is produced.

Gene expression in bm1 mutants and their normal counterparts

The ZmCAD2 gene appeared to be significantly under-expressed in all investigated bm1 lines (F2bm1, 511Kbm1, 5803Cbm1, 511Jbm1, Ev2210bm1), in comparison to their normal counterparts (F2, 511K, Ev2210), corroborating the effect of bm1 mutations on the expression of the ZmCAD2 gene. In addition, the expression of the ZmCAD1 gene was significantly up-regulated in three bm1 mutants out of five, and especially in the Ev2210bm1 mutant line (Figure 2). These results did not corroborate the lower ZmCAD1 expression observed in bm1 plants (Shi et al, 2006; Guillaumie et al, 2007a), that was, however, observed in much younger plants. The ZmCAD4 gene, encoding a CAD-like protein with a mannitol dehydrogenase activity, had variable expression levels according to genotypes. ZmCAD4 was over-expressed in F2bm1 and Ev2210bm1 lines, while it was under-expressed in the three other bm1 lines considered.

Lignin content and structure in the ZmCAD2 Ev2210bm1 mutant line

The consequences of the bm1 mutation on maize lignin content and structure have been shown to date in F2bm1 (bm1-ref) mutant plants (with the AC insertion), before it was firmly established that the ZmCAD2 gene was affected in these plants. Relative to the wild-type plants, investigations in the Ev2210bm1 mutant plants, altered in the ZmCAD2

<table>
<thead>
<tr>
<th>Table 2 - Cell wall traits in whole plants without ear at silage maturity stage for the normal and ZmCAD2 Ev2210bm1 mutant plants.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>NDF</strong></td>
</tr>
<tr>
<td>Normal</td>
</tr>
<tr>
<td>Ev2210bm1</td>
</tr>
<tr>
<td><strong>F test</strong></td>
</tr>
</tbody>
</table>

NDF (neutral detergent fiber) as % dry matter (DM), ADL (acid detergent lignin) and Klason lignin (Ki) % NDF, in vitro NDF digestibility (IVNDFD) as weight % of NDF. Data as mean values from 9 independent samples, differences between normal and mutant plants were non significant (ns) or significant at P < 0.01 % (*) and 0.001 % (**).
gene after transposon insertion, established that this mutant also displayed substantial variations in lignins and cell wall phenolics (Table 2). The lignin contents estimated by contents the Klason procedure and Van Soest (ADL) method were reduced by 6 and 17%, respectively, in Ev2210bm1 plants, while the cell wall content (referred to as NDF) was not changed.

Substantial changes in lignin structure were also observed, with 30 - 40% fewer units only involved in β-O-4 bonds, as shown by the lower thioacidolysis yield (Table 3). Lignins in Ev2210bm1 plants are enriched in resistant inter-unit links, referred to as condensed linkages. Based on thioacidolysis investigations, the monomeric composition of the polymer was only slightly modified with a small increase in S unit frequency, a small decrease in G unit frequency, and no changes in H unit frequency. In addition to lignin-derived monomers, thioacidolysis of extractive-free samples released pCA and FA acids (Table 3). Previous studies have revealed that pCA is essentially ester-linked to lignins in maize cell walls (Ralph et al, 1994) and that most of these ester bonds do not survive the thioacidolysis procedure (Jacquet, 1997). Accordingly, the yield of thioacidolysis-released pCA accurately reflects the amount of p-coumarate esters. The origin of thioacidolysis-released FA is more complex because ferulate esters, which are linked to arabinoxylans, may be simultaneously linked to lignins by various bonding patterns. Anyway, obtained results suggested that the level of p-coumarate esters was substantially reduced in the cell walls of the Ev2210bm1 plants (nearly 30%), whereas ferulate units released by thioacidolysis were affected to a lower extent. Correlatively to lignin content and structure changes in Ev2210bm1 plants, their cell wall digestibility was improved by nearly 12%.

Following thioacidolysis and gas chromatography-mass spectrometry analysis of the lignin-derived monomers, signatures of CAD deficiency were identified in the Ev2210bm1 plants as previously observed in various CAD-deficient transgenic or mutant plants (Kim et al, 2002; Sibout et al, 2005; Thévenin et al, 2011). The mutant samples released 2 to 3 times more thioacidolysis monomers derived from coniferaldehyde end-groups than the control (Table 3). Moreover, the two G and S indene derivatives derived from coniferaldehyde and sinapaldehyde, which have been incorporated into lignins by 8-O-4 cross-coupling, were recovered in substantial amounts from Ev2210bm1 lignins and only as trace components from the control. These results corroborated those obtained in F2bm1 plants (Barrière et al, 2004a), even if aldehyde incorporations were shown to be lower in Ev2210bm1 plants than previously observed in F2bm1 plants. The difference between F2bm1 and Ev2210bm1 plants was probably related either to different penetrances of the two different types of ZmCAD2 mutations in F2bm1 and Ev2210bm1, and/or to the different genetic backgrounds of the two lines. Among the differences, the much higher ZmCAD1 over-expression was observed in Ev2210bm1, but not in F2bm1, plants (Figure 2).

Finally, when subjected to thioacidolysis and relatively to normal plants, Ev2210bm1 plants released 2 to 3 times higher amounts of the AG compound, which is a marker for the incorporation of free ferulic acid into lignins (Ralph et al, 2008). Despite this fact was not considered in previous years (Barrière et al, 2004a), a greater incorporation also occurred in F2bm1 lignins (data not shown), as well as in lignins of Arabidopsis CAD-C x CAD-D double mutant plants (Thévenin et al, 2011). The incorporation of ferulic acid into lignins allows new branching points with bi- and di-phenyl ether structures and consequently modifies the macro-molecular organization of the polymer (Ralph et al, 2008). The putative higher degree of branching of lignins in the Ev2210bm1 mutant was confirmed by the results obtained by thioacidolysis of permethylated samples which provides the proportion of terminal units with free phenolic groups relative to internal β-O-4 linked lignin units. In agreement with previous results obtained with maize lignins (Lapiere, 1993), thioacidolysis of the permethylated normal maize sample revealed that about 50% of G units are terminal units with free phenolic groups (Table 3), whereas this figure is higher for H

<table>
<thead>
<tr>
<th>Line</th>
<th>Thioacidolysis yield (H + G + S) (µmole g⁻¹ KL)</th>
<th>Relative frequency of thioacidolysis monomers (% molar)</th>
<th>Minor thioacidolysis monomers (µmole g⁻¹ KL) in β-O-4 linkage</th>
<th>Free phenolic groups (%)</th>
<th>Thioacidolysis release (p-H) cinnamic acids (µmole g⁻¹ KL)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>H units</td>
<td>G units</td>
<td>S units</td>
<td>coniferaldehyde end-groups</td>
<td>G indene</td>
</tr>
<tr>
<td>Normal</td>
<td>526.22</td>
<td>1.62</td>
<td>46.14</td>
<td>52.21</td>
<td>0.36</td>
</tr>
<tr>
<td>Ev2210bm1</td>
<td>294.89</td>
<td>1.60</td>
<td>43.57</td>
<td>54.82</td>
<td>1.18</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.2ns</td>
<td>12.0*</td>
<td>12.0*</td>
<td>172.4**</td>
</tr>
</tbody>
</table>

Thioacidolysis released H, G and S lignin monomers of coniferaldehyde (coniferal) end-groups, coniferaldehyde ether-linked at Cβ (releasing G indene), sinapaldehyde ether-linked at Cβ (releasing S indene), and of the marker compound (Aₜ) for free ferulic acid incorporation in lignins. Thioacidolysis-released p-coumaric acid (pCA) and ferulic acid (FA). Percentage (%) of free phenolic groups in β-O-4 linked H, G and S lignin units, as determined by thioacidolysis of permethylated samples. Significant differences between normal and mutant plants indicated as in Table 2.
the maize bm1 mutations units (about 87%) and much lower for S units (less than 7%). However, in the Ev2210bm1 mutant, more than 60% of G lignin units are terminal units with free phenolic groups. The substantially higher frequency of free phenolic groups in lignins is a structural specificity which is common to all the CAD-deficient plants analyzed so far, including the F2bm1 maize mutant \cite{Lapiere, 2010}, and the signature of a lignin network disorganized in smaller domains.

**Discussion**

While the ZmCOMT mutations were identified about 15 years ago in the maize bm3 mutants \cite{Vignols et al, 1995; Morrow et al, 1997}, no investigations had firmly established up until now what gene was affected in bm1 plants, even if it was greatly suspected that the CAD2 gene was involved in the mutation(s). The sequencing of the ZmCAD2 allele in eight bm1 lines highlighted the diversity of genomic events underly-ing the bm1 phenotype. Two- and four-base insertions, missense mutations, and transposon insertions are expected to lead to truncated or inactive CAD2 proteins, thereby causing the bm1 phenotype. The mutation of the ZmCAD2 gene in bm1 maize plants was also established in concomitant investigations in other genetic backgrounds \cite{Chen et al, 2012b}. The first intron of the ZmCAD2 bm1-das1 mutant (Dow AgroSciences bm1) was shown to contain a 3,444 bp insertion, resulting in a chimeric mRNA containing a premature stop codon, and a truncated protein of 48 amino acids. The mutation identified in the bm1 allele of the S15Dbm1 simultaneously investigated line was the same as the one observed in the F2bm1 and A619bm1 lines, consisting in an AC insertion in exon 3 that resulted in a 147 amino acid reduced protein. The AC insertion in the third ZmCAD2 exon could thus be considered to be the first described bm1 (bm1-ref) mutation.

In addition, based on observations in nurseries and allelism tests, the CAD2-bm1 mutation seems to occur or appear more frequently than other maize bm mutations. This fact considered from a very small number of events could actually correspond to a random effect, without any underlying biological mechanism. However, this fact could also be related i) to the complexity of the catalytic sites of the CAD2 protein rapidly altered by a slight modification of the amino acid sequence, ii) possibly to a DNA sequence or a chromosomal position favoring transposon insertion, and iii) to a low or null effect of the CAD2 mutation on agronomic value. \cite{Chen et al, 2012b} have indeed observed that the CAG to TAG nonsense mutation occurred in SbCAD2 of bmrm6 sorghum just one nucleotide apart from the AC insertion shown in maize bm1 mutants. Moreover, although plants contain numerous genes coding for CAD-like genes, only EgCAD2-type, and to a lesser extent EgCAD1-type CAD enzymes, were shown with a primary physiological role in lignin biosynthesis. This situation results from the presence in lignin-related CAD proteins of a few key residues permitting significant catalytic rates on monolignol precursors. The role of several of these key residues for efficient catalysis and monolignol biosynthesis had been especially established based on investigations in lead tree \cite{Leucaena leucocephala} and switchgrass \cite{Pandey et al, 2011; Saathoff et al, 2012}. Correlatively, while a highly truncated protein was shown to be the determinant of the SbCAD2 bmrm6-ref mutation in sorghum, the bmrm6-3 mutation indeed resulted from a unique mutation in the 188GXGG(V/L)/G189 motif changed into a 188GXGS(V/L) G189 motif, affecting the binding affinity for the NADP+ cofactor \cite{Saballos et al, 2009}. In addition, the rice gold-hull-and-internode-2 (gh2) mutant, which exhibits a reddish-brown pigmentation in the hull and internodes which become golden yellow at maturation, has an altered OsCAD2 gene with a G to A substitution in the fourth exon \cite{Zhang et al, 2006}. This substitution induced the replacement of a zero-charge glycine by an electro-negative aspartic acid, leading to an inactive protein, as confirmed by recombinant protein experiments. As observed in maize and sorghum, the CAD activity was drastically reduced in rice gh2 mutant, indicating that other CAD(-like) enzymes only partly substitute the OsCAD2 deficiency, at least in rice for conifer alkyl alcohol biosynthesis \cite{Zhang et al, 2006}.

All cell wall traits shown in plants of the Ev2210bm1 mutant were indeed very similar to the ones observed in the F2bm1 plants \cite{Barrière et al, 2004a}, corroborating the involvement of ZmCAD2 mutations in the bm1 phenotype. Both F2bm1 and the Ev2210bm1 plants had reduced contents in lignins and in p-coumarate esters, which are correlated with a higher cell wall digestibility. Both exhibited important structural alterations of their lignins, which correspond to the specific signature of CAD deficiency, namely a higher frequency of condensed bonds and of free phenolic groups and the increased incorporation of coniferaldehyde, sinapaldehyde and free ferulic acid in lignins \cite{Lapiere et al, 2004; Ralph et al, 2008}. However, in addition to the incorporation of aldehydes and atypical compounds into lignins, thioacidolysis yields indicated the synthesis of regular monolignols in CAD2 mutant plants, reaching nearly 55 % of the normal \beta-O-4-linked monomeric units in the Ev2210bm1 maize mutant. EgCAD1-type and CAD-like genes acting on \textit{p}-hydroxy-cinnamaldehydes, in addition to a possible residual CAD2 activity in mutant with modified but not truncated protein, should therefore be considered. Expression data showed a probable substitution role of the ZmCAD1 gene in Ev2210bm1 and, to a lesser extent, in F2bm1 and 511Kbm1. Corroborating the substitution by other CAD or CAD-like enzymes, the truncated SbCAD2 protein was not detected in extracts from bmrm6 sorghum plants, but CAD activity was still detectable in mutant plant tissues \cite{Pillonel et al, 1991; Saballos et al, 2009}. Nev-
nevertheless, if other CAD proteins (including EgCAD1-type) can be active on cinnamyl substrates, brown midrib or golden phenotypes and deep alterations in lignin composition establish that EgCAD2-type genes encode the main CAD proteins involved in the regular monolignol biosynthetic pathway of grasses.

The interest of CAD2-deficient plants in industrial processes was first shown with the demonstration of the significantly improved pulping capacity of poplar and pine CAD mutants (Baucher et al., 1996; Lapiere et al., 1999; Lapiere et al., 2000; O’Connell et al., 2002; Gill et al., 2003). In grasses, for bioenergy production, the efficiency of the EgCAD2-type mutation or deregulation towards significant increases of enzymatic hydrolysis and/or conversion rate into bioethanol (with or without pretreatment) has been shown in sorghum and switchgrass plants, without significant negative effects on agronomic value (Saballos et al., 2008; Corredor et al., 2009; Sattler et al., 2009; Sattler et al., 2010a; Sattler et al., 2010b; Scott et al., 2010; Saathoff et al., 2011). In maize, most investigations with bm genes have been done with the bm3 COMT mutant, because this latter induced the highest improvement in cell wall digestibility and energy value for cattle feeding (Barrière et al., 2004b). Similarly, high increases in the release of fermentable sugars were shown in bm3 mutant plants (Vermerris et al., 2007). However, an important decrease in agronomic value has been simultaneously shown in nearly all early bm3 investigated hybrids, a fact that could possibly be closely related to the germplasm of these ancient investigated hybrids. None of the considered hybrids were related to the lodent genetic resource, of high agronomic value and standability. Several bm3 hybrids are now indeed available in different markets, including the US market, with proven greater efficiency in dairy cow feeding (at least 28 published investigations between 1976 and 2012, most of them in Journal of Dairy Science). The feeding value of maize bm1 hybrids has not been extensively investigated, likely as the first digestibility measurements have shown the lower improvement of maize bm1 genotypes in comparison to bm3 near-isogenics (Barnes et al., 1971; Lechtenberg et al., 1972; Barrière and Argillier, 1993). Nevertheless, the agronomic value of maize bm1 hybrids (and CAD2 mutants) is seemingly less modified than the one of bm3 hybrids (and COMT mutants). Based on INRA Lusignan unpublished data from the two old INRA260 and LG11 hybrids, average in vivo NDF digestibility values were equal to 50.6, 55.7, and 59.4% in normal, bm1, and bm3 isogenics, while DM yields were equal to 12.7, 12.5, and 11.8 t ha⁻¹, respectively. On the contrary, recent investigations strengthened the interest of maize bm1-type mutants, based on CAD down-regulation, for industrial purposes. Transgenic CAD-RNAi maize plants had stem cell walls with a slight reduction of lignin content, but the stems of deregulated plants were nevertheless more degradable than their normal counterparts. Fermentation assays also revealed that CAD deregulated plants produced higher levels of ethanol compared to normal ones (Fornalé et al., 2012). Moreover, the degree to which lignin polymers incorporate various phenolic compounds in place of the three regular constitutive monolignols is surely underappreciated (Ralph, 2010). In addition to ferulic acid and hydroxy-cinnamaldehydes, unusual monomers including acylated hydroxycinnamyl alcohols, dihydro-hydroxycinnamyl alcohols, hydroxy-benzaldehydes and other hydroxy-cinnamatic acids, can be incorporated into lignins of wild-type plants (Vanholme et al., 2012). Because plants could tolerate shifts in lignin composition with no or lower impact on growth than observed for reduced lignin contents, the substitution of some fraction of the three regular monolignols by unusual or alternative monomers through breeding, directed mutagenesis, or genetic engineering is thus a relevant strategy to tailor lignins in bioenergy crops so that cell walls would be more susceptible to biomass pretreatments (Eudes et al., 2012; Vanholme et al., 2012). The interest of unusual monomers incorporation into lignins for bioenergy production was also highlighted with bmr6 sorghum. Glucose yields were indeed improved by at least 25% from sorghum biomass of bmr6 plants, compared to normal isogenics (Saballos et al., 2008; Dien et al., 2011). Several bmr6 (and “bmr”) sorghum hybrids have thus been registered, or are in registration process, for both European and American markets. The higher efficiency of bmr6 sorghum for animal feeding has also been shown from dairy cows experiments. Cows fed bmr6 silage in their diets had 16% higher milk yields in comparison to similar diet with normal sorghum (Oliver et al., 2003). However, milk yields were equal for bmr6 sorghum and normal maize silages, highlighting again the higher cell wall degradability in maize than in sorghum. In addition, the modification of the lignin structure could also be considered to increase cell wall degradability (Zhang et al., 2011). In a set of maize inbred lines, more condensed lignins were thus shown to be more favorable to an increased cell wall degradability than β-O-4 rich lignins. Finally, for both bioenergy production and cattle feeding, the successful breeding of improved maize (and C4 grasses) genotypes with more degradable cell walls, based on different lignin polymer organization, incorporation of alternative monomers, and reduced ferulate cross-linkages, requires more knowledge about phenolic compound biosynthesis in plants and about their coupling in the secondary walls.

Different tentative conclusions can be considered from the current investigations in bm1 maize. Several very different events are responsible for the ZmCAD2-bm1 mutation in maize. The ZmCAD2 gene encodes the predominant CAD for monolignol production, even though other CAD gene(s) could be simultaneously involved, or may have substitutive
activity. Comparing the biochemical traits in transposon tagged and disrupted mutant highlighted the relevance of the marker characteristics that were shown in CAD2 deficient plants, including aldehyde and atypical compounds and linkages. Even if further experiments need to be done, CAD2 is likely a good target for the improvement of energy production based on maize and grass lignocellulose biomass, including also a greater susceptibility to (environmentally friendly) pretreatments (Maehara et al, 2011; Wu et al, 2011). These specific properties are related to the presence of unusual compounds incorporated into lignins, inducing a modified structure of the polymer spread over carbohydrates in smaller domains. Despite the interest of bm1 maize hybrids for cattle feeding has not yet been established, contrarily to the proven efficiency of bm1 plants for bioenergy production, the breeding of maize bm1 hybrids is likely promising because the ZmCAD2 mutations seem to induce limited unfavorable consequences on other agronomical traits, even if opposite results could be observed with a few genetic backgrounds (Lorenz et al, 2009).

For the improvement of cell wall traits for both feeding and industrial uses, genetic targets should be considered in the whole set of genes involved in secondary wall biosynthesis. Genes involved in lignin production are upstream regulated by MYB and NAC transcription factors (Sonboi et al, 2009; Fornalé et al, 2010; Zhong et al, 2011; Gray et al, 2012). Ferulate driven cross-linkages are both dependent on ferulic acid (more probably of feruloyl-CoA) and arabinoxylan biosynthesis (Barrière et al, 2009; Hatfield and Marita, 2010; Jung and Phillips, 2010; Piston et al, 2010; Jung et al, 2011). Lignified tissue patterning is regulated during plant growth by members of several gene families, including members of the zinc finger and HD-ZIP families (Barrière et al, 2009). Search for transposon-tagging mutants with improved cell wall degradability should indeed be a relevant strategy for the discovery of genes with still unknown important role in secondary wall assembly (Vermerris et al, 2007; Jung and Phillips, 2010). Mutants and favorable alleles of all these genes have then to be considered as targets during maize (and grass) breeding for enhanced energy value. However, while plant breeding for animal feeding and bioenergy production both requested the knowledge of genetic mechanisms involved in secondary wall assembly, the breeding of bioenergy plants has also to take into account the lignin network susceptibility to mild-alkali pretreatments.

Acknowledgements

Authors thank Marty Sachs, Director of the maize stock center for the supply of bm1 mutants, Natalia de Leon and Jim Coors (University of Wisconsin) for the supply of A619bm1, and all three for their interest in this work. Authors also thank Thomas Lübberstedt, Hongjun Liu, and Paul Scott for discussions and interaction during investigations on the bm7 mutation. Christiane Minault, Dominique Denoue and Pascal Vernoux were in charge of plant cropping at INRA Lusignan. Laurent Cézard and Frédéric Legèe (INRA Versailles), and Corinne Melin (INRA Lusignan) were in charge of lignin content and composition analyses, and cell wall degradability estimates. Investigations on the Ev2210 bm1 were funded by the MaizeWall project of the French genomic network Génoplante.

References


Burnham CR, 1947, Maize Genetics Cooperation Newsletter 21, pp 36


Chavigneau H, Goué N, Courtil A, Jouanin L, Rey-
mond 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


Emerson RA, 1935. Cornell Univ Agric Exp Stn, Mem 180

Eyster WH, 1926 Chromosome VIII in maize Science 64: 22


Gee MS, Nelson OE, Kuc J, 1968. The abnormal lignins produced by the brown-midrib mutants of maize. II. Comparative studies on normal and brown-midrib-1 dimethylfornamide lignins. Arch Biochem Biophys 123: 403-408


35-45
Lechtenberg VL, Muller LD, Bauman LF, Rhayder CL, Barnes RF, 1972. Laboratory and in vivo evaluation of inbred and F1 populations of brown-midrib mutants of Zea mays L. Agron J 64: 657-680
Barrière et al.


Sonbol FM, Fornalé S, Cappellades M, Encina A, Tourino S, Torres JL, Rovira P, Ruel K, Puig-


**Supplementary Figure 1 - Nucleotide alignments of the ZmCAD2 CDS in normal and bm1 isogenic lines**

(Identical nucleotides are shaded in grey, X are missing data)
Supplementary Figure 2 - Alignments of the deduced amino-acid sequences of the ZmCAD2 protein in all investigated lines (Identical amino acid residues are shaded in grey, X are missing data).