Theor Appl Genet (2004) 110: 126–135
DOI 10.1007/s00122-004-1808-4

ORIGINAL PAPER

C. Guillet-Claude · C. Birolleau-Touchard
D. Manicacci · M. Fourmann · S. Barraud · V. Carret
J. P. Martinant · Y. Barrière

Genetic diversity associated with variation in silage corn digestibility for three *O*-methyltransferase genes involved in lignin biosynthesis

Received: 26 March 2004 / Accepted: 24 August 2004 / Published online: 22 October 2004

Abstract Polymorphisms within three candidate genes for lignin biosynthesis were investigated to identify alleles useful for the improvement of maize digestibility. The allelic diversity of two caffeoyl-CoA 3-*O*-methyltransferase genes, *CCoAOMT2* and *CCoAOMT1*, as well as that of the aldehyde *O*-methyltransferase gene, *AldOMT*, was evaluated for 34 maize lines chosen for their varying degrees of cell wall digestibility. Frequency of nucleotide changes averaged one SNP every 35 bp. Ninety-one indels were identified in non-coding regions and only four in coding regions. Numerous distinct and highly diverse haplotypes were identified at each locus. Numerous sites were in linkage disequilibrium that declined rapidly within a few hundred bases. For F4, an early flint French line with high cell wall digestibility, the *CCoAOMT2* first exon presented many non-synonymous polymorphisms. Notably we found an 18-bp indel, which resembled a microsatellite and was associated with cell wall digestibility variation. Additionally, the *CCoAOMT2* gene co-localized with a QTL for cell wall digestibility and lignin content. Together, these results suggest that genetic diversity investigated on a broader genetic basis could contribute to the identification of favourable alleles to be used in the molecular breeding of elite maize germplasm.

Communicated by H.H. Geiger

C. Guillet-Claude · C. Birolleau-Touchard
V. Carret · Y. Barrière
Unité de Génétique et d'Amélioration des Plantes Fourragères,
INRA, Route de Saintes, 86600 Lusignan, France

D. Manicacci · M. Fourmann · S. Barraud
Station de Génétique Végétale,
UMR INRA-UPS-CNRS-INAPG,
Ferme du Moulon, 91 190 Gif sur Yvette, France

J. P. Martinant
Biogemma S.A.S., Campus Universitaire des Cézeaux,
24 avenue des Landais, 63170 Aubiere, France

C. Guillet-Claude (✉)
Centre de recherche en biologie forestière, Université Laval,
Sainte-Foy, Québec, Canada, G1K 7P4
E-mail: cguillet@rsvs.ulaval.ca
Tel.: +1-418-6562131
Fax: +1-418-6567493

Introduction

Silage corn is a major forage in dairy cattle feeding because of its high energy content and good ingestibility. Its energy value depends on grain content, but mostly on cell wall breakdown by rumen microorganisms. Lignin interferes with the digestion of cell wall polysaccharides by acting as a physical barrier to microbial enzymes. In forage crops, lignin content, lignin structure, and cross-linking between cell wall components influence cell wall digestibility (reviewed in Barrière et al. 2003).

Gene sequence diversity has been mainly studied to understand the impact of selection during maize domestication, and to identify the levels and patterns of genetic variation in a large sample of maize loci (reviews in Buckler and Thornsberry 2002). However, Thornsberry et al. (2001) and Palaisa et al. (2003) demonstrated that nucleotide diversity analysis of candidate genes for flowering date and endosperm colour, respectively, allowed for the identification of alleles responsible for variation of these quantitative agronomic traits. Recently, in *Eucalyptus globulus*, Poke et al. (2003) found cinnamoyl CoA reductase (*CCR*) and cinnamyl alcohol dehydrogenase 2 (*CAD2*) polymorphisms that could alter enzyme function. Furthermore, in a previous report on the nucleotide diversity analyses of a maize peroxidase gene (*ZmPox3*), a mutant allele was shown to be associated with cell wall digestibility (Guillet-Claude et al. 2004). This method can thus be considered promising in terms of detection of useful alleles for the improvement of silage corn digestibility.

Caffeoyl-CoA 3-*O*-methyltransferase (CCoAOMT) plays an essential role in the synthesis of guaiacyl lignin units as well as in the supply of substrates for the synthesis of syringyl lignin units. Caffeic acid 3-*O*-

methyltransferase (COMT), or more precisely, 5-hydroxyconiferaldehyde *O*-methyltransferase (AldOMT), controls the biosynthesis of syringyl lignin units (for a review, Boerjan et al. 2003). Two CCoAOMT genes were identified in maize (Collazo et al. 1992). The *CCoAOMT1* and *CCoAOMT2* genes were mapped, respectively, on chromosomes 6 (Génoplante project, data unpublished) and 9 (bin 9.02, this work) and co-localized with major QTL clusters for lignification and cell wall digestibility (Méchin et al. 2001; Roussel et al. 2002). Only one *AldOMT* was evidenced in the maize genome and mapped on chromosome 4 short arm in bin 4.05 (as reported in the Maize Database, http://www.maizegdb.org), where a QTL for cell wall lignin content and digestibility were found (Méchin et al. 2001). Transgenic analyses confirmed that these three genes are preferential targets for maize silage digestibility improvement (reviewed in Barrière et al. 2003). Digestibility of CCoAOMT down-regulated alfalfa plants was greatly improved (Guo et al. 2001). *AldOMT* down-regulated maize plants had altered lignification and improved digestibility as compared with a *bm3* mutant (Piquemal et al. 2002). The *bm3* mutation is the result of deletion of an *AldOMT* gene (Vignols et al. 1995) and has a drastic effect on cell wall digestibility (reviewed in Barrière et al. 2003).

The purpose of this study was to identify in *CCoAOMT2*, *CCoAOMT1*, and *AldOMT*, based on 34 maize lines chosen for their varying degrees of cell wall digestibility and representative of temperate regions germplasm, haplotype or individual polymorphisms associated with digestibility. This study should provide data illustrating how genetic diversity analyses of lignification candidate genes could contribute to the identification of alleles to be used in the molecular breeding of elite forage maize germplasm.

Materials and methods

Plant material and DNA extraction

Thirty-four inbred lines were chosen in order to display a large variation in cell wall digestibility and to represent a diversified sample of genotypes used both in European and US breeding programs (Table 1). Maize plants were grown in vermiculite supplemented with a nutrient solution under a 16/8-h light/dark regime for 10 days. Genomic DNA from young maize leaves was isolated using the plant DNAEASY miniprep kit (Qiagen, Courtaboeuf, France).

Primer design and PCR amplifications

For *CCoAOMT1* and *CCoAOMT2*, pairs of primers were designed based on published cDNA sequences (accession numbers AJ242980 and AJ242981, respectively, Table 2). For both genes, fragments of about 1.2 kb to1.3 kb were amplified, encompassing the 5′UTR and the complete coding region (Fig. 1a). Even with several different PCR primers and primer combinations, it was not possible to amplify the *CCoAOMT2* gene for two genotypes (F7 and Rottaler Silomaïs). Because of its size, the *AldOMT* gene (accession M73235) was amplified using three overlapping pairs of primers. A total region of 2.2 kb, which encompassed the 5′UTR and the complete coding region, was amplified for this gene (Fig. 1b).

Reactions were performed in 50 µl containing 1× PCR buffer (Sigma, Saint-Quentin Fallavier, France), 200 µM of dNTPs, 0.2 µM of 5′ oligo, 0.2 µM of 3′ oligo, and 2.5 U REDTaq DNA Genomic Polymerase (Sigma). One hundred nanograms of genomic DNA were used as template. The program was 5 min at 95°C, followed by 30 cycles of 30 s at 95°C, 30 s at 60°C, and 1.5 min at 72°C, followed by 5 min at 72°C. For the two *AldOMT* exons, which are GC-rich sequences, we used the Advantage-GC Genomic Polymerase Mix (BD Biosciences Clontech, Palo Alto, Calif., USA) to facilitate PCR amplification. DNA amplifications were performed in a volume of 50 µl. The reactions contained 100 ng genomic DNA, 0.2 µM of each primer, 200 µM of dNTPs, 1.1 mM $Mg(Oac)_2$, 5–6 U *Tth* DNA polymerase, 1× Genomic PCR reaction buffer, and 1 M GC-melt (BD Biosciences Clontech). The program was 1 min at 95°C, followed by 30 cycles of 30 s at 94°C, and 2 min at 64°C, followed by a terminal elongation at 64°C for 3 min.

All PCR products were purified with Minelute columns (Qiagen, Courtaboeuf, France) prior to be cloned into the pGEM-T vector using the pGEM-T Easy Vector System (Promega, Charbonnières, France).

Isolation of the maize *AldOMT* promoter region

Because the *AldOMT* gene promoter sequence was not accessible in databases, a walking-PCR procedure was used to amplify the 5′-flanking region. Two maize Genome Walker libraries (*Pvu*II, *Dra*I) were constructed using a Universal GenomeWalker Kit (BD Biosciences Clontech) for six maize genotypes (F4, F7012, F2, Lan496, F288, and W117). Two specific reverse primers and two nested adaptor primers (AP1 and AP2) provided in the kit were used in two nested PCR reactions to amplify the 5′ regulatory region (Table 2; Fig. 1b). Reactions were performed in the same reaction buffer as previously described for the REDTaq DNA polymerase (Sigma), but with 0.2 µM 5′ oligo (adaptor primer) and 0.2 µM 3′ oligo (*AldOMT* gene-specific primer). One microlitre of each DNA library was used as template for the primary PCR amplification, and 1 µl of these PCR products was used for secondary PCR amplifications. The PCR program was 5 min at 95°C, followed by 30 cycles of 30 s at 95°C, 30 s at 61°C, and 2 min at 72°C, followed by 5 min at 72°C. After secondary PCR walking, about 1.1-kb fragments were cloned prior to sequencing as described above for the

Table 1 List of maize germplasms indicating cell wall digestibility values

Line	Origin and pedigree	Cell wall digestibility[a]
Flint		
F1	Lacaune (France)	3.5
F2	Lacaune (France)	3.5
F7	Lacaune (France)	4
F7012	Related to European flint	4.5
F4	Etoile de Normandie (France)	5
F66	Sost (France)	3
Ep1	Lizaragote (Spain)	4
Du101	Related to Jaune de Bade	3.5
F64	Argentina PI186-223	4
F564	F7 × F64	4.5
F286	F7 × F564	4
F324	Related to European and Argentina flint	5
Line212	Private Biogemma line	1
Line16	Private Biogemma line	3.5
Early dent		
F113	Spooner473 = (W37a × W37) × (W47 × EK43)	4
W117	W643 (from Golden Krug) × Minnesota13	4
F271	Co125 (unknown origin) × W103 (from Golden Glow)	1
F288	Related to US early dent (7/8) and Blanc de Chalosses (1/8)	3
MBS847	Iodent	2.5
F7025	Iodent × F113	2.5
Late dent		
B14	Iowa Stiff Stalk Synthetic (C1)	1.5
B73	Iowa Stiff Stalk Synthetic (BS13C5)	1.5
Mo17	Lancaster	2.5
Lan496	Lancaster related	2
W64A	Wf9 (from Reid Yellow dent) × CI 187-2 (from Krug Reid)	2.5
Wis93-3520	Wisconsin dent material	4
Wis94-443	Wisconsin dent material	4
DE811	Related to US late dent (BSSS, Lancaster, ...)	1.5
Ecotypes[b]		
	Rottaler silomaïs (Flint, Germany)	-[c]
	Noordlander VC145 (Flint × dent, The Netherlands)	-
	Sibiriacka (Flint, Russia)	-
	Rainbow flint (Canada)	-
	Polar dent (unknown)	-
	Québec28 (Flint, Canada)	-

[a] A synthetic note from 1 to 2.5 corresponds to low to medium cell wall digestibility; values between 2.5 and 5 indicate medium to high cell wall digestibility

[b] Ecotypes are chosen because of apparent high cell wall digestibility

[c] A *dash* means that the cell wall digestibility is not yet evaluated

three *O*-methyltransferase (OMT) genes. Based on these new promoter sequences, a forward primer was designed from its 5′ end (Table 2) to obtain by classical PCR the promoter region (approximately 850 bp) on the other genotypes sampled (except F7025, MBS847, Du101 and F66, for which PCR amplifications failed). The amplification conditions were the same as those described for the nested PCR, except that 100 ng genomic DNA was used as template, and the annealing temperature was 55°C.

DNA sequencing

Sequencing was performed for each PCR fragments in both directions by Isoprim (Toulouse, France) and MWG-Biotech (Ebersberg, Germany). The sequences containing singletons were checked by re-amplifying genomic DNA and partially re-sequencing the appropriate alleles. Contigs were constructed using SeqWeb (GCG Wisconsin Package). Sequences were aligned using CLUSTAL W (Thompson et al. 1994).

Sequence data from this article have been deposited with the EMBL/GenBank data libraries under accessions nos. AY323238-AY323271 (*CCoAOMT1* alleles), AY279004-AY279035 (*CCoAOMT2* alleles), and AY323272-AY323305 (*AldOMT* alleles).

Promoter sequence analysis

Putative *cis*-acting regulatory elements were identified using the data base PlantCARE (http://intra.psb.ugent.be:8080/PlantCARE) and MatInspector (http://www.genomatix.de/cgi-bin/matinspector/matinspector.pl/). The promoter structure was analyzed using Promoter Prediction software (Neural Network; http://www.fruitfly.org/seq_tools/promoter.html).

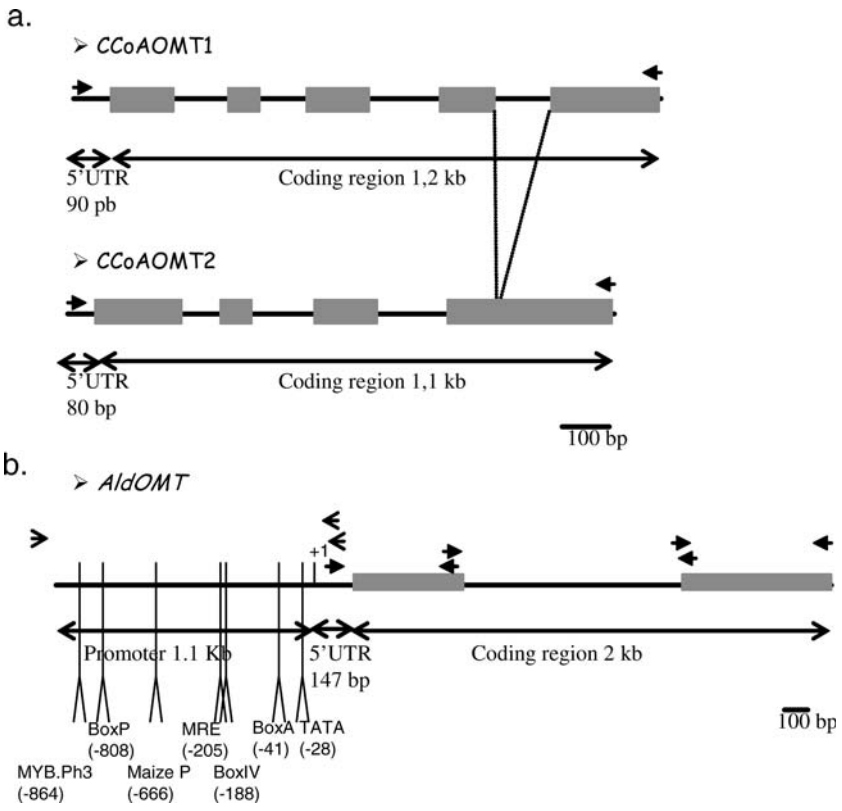

Fig. 1 A schematic diagram of the three *O*-methyltransferase genes. **a** Comparison of the exon–intron organization of the two maize caffeoyl-CoA 3-*O*-methyltransferase genes, *CCoAOMT1* and *CCoAOMT2*. *Boxes* represent the exon. *Arrowheads* indicate positions of the primers used to amplify each gene. **b** Organisation of the aldehyde *O*-methyltransferase (*AldOMT*) gene. The coding region has an unique intron. *Arrowheads* indicate positions of the primers used to amplify the 5′UTR and coding region. *Thin arrows* correspond to primers used to obtain the promoter region. The putative transcriptional start site is indicated by *+1*. A putative TATA box is located 28 bp downstream of this site. Motifs with significant homologies to known plant *cis*-acting element are the MYB.ph3 binding site (ATAAGTTAGTGAT), found in the gene *CHS* promoter (Solano et al. 1995); Box P (CAACAACCTA) identified in peroxidase; *PAL*, *4-CL*, *CCoAOMT*, and *CAD* gene promoters (Feuillet et al. 1995; Grimmig and Matern 1997; Neustaedter et al. 1999; Ito et al. 2000; Whitbred and Schuler 2000); Maize P (ACCAACCCG), also located in peroxidase gene promoter (Ito et al. 2000); MRE (AACCTAA), present in the *CHS* gene promoter (Feldbrügge et al. 1997); Box IV (ATTAT), found in the *PAL* gene promoter (Whitbred and Schuler 2000); and box A (CCGTCC), identified in the *CCoAOMT* gene promoter (Grimmig and Matern 1997). Their positions are given with respect to the consensus sequence and correspond to the 5′ first nucleotide in the motif numbered from the transcription initiation site

Mapping of the *CCoAOMT2* gene

CCoAOMT2 was mapped using a RIL population developed at INRA Lusignan (France) from the cross between F271 and F288 lines. Detailed results on the linkage map developed in the F288 × F271 RIL progeny were previously reported by Barrière et al. (2001). Specific primers for each parental allele were designed based on an indel polymorphism (named *lus1*) identified in the second *CCoAOMT2* intron. Amplification reactions were performed on 131 RILs using two pairs of primers: a forward primer specific to either F271 (5′-ACAGACCGAGATCTGACTGAGAAGTGA-3′) or F228 (5′-CGTGGACAGACCGAGTCTGAGAAC-3′) and a reverse primer common to both alleles (Table 2). The PCR amplification protocol was the same as that used to obtain the two CCoAOMT genes. When the specific primer corresponded to its *CCoAOMT2* allele, a 900-bp fragment was amplified. If not, the PCR failed. The success of PCR amplifications was checked by electrophoresis in a 1% agarose gel containing ethidium bromide.

Evaluation of corn silage digestibility

In vitro cell wall digestibility of lines (per se values) was investigated through different multi-year and multi-local experiments managed by the different partners. Values of cell wall digestibility were estimated as described in Roussel et al. (2002) through the DINAGZ criterion [in vitro digestibility of the "non starch (ST), non-soluble carbohydrates (SC) and non crude protein (CP) part"] based on the enzymatic solubility of the whole plant (Aufrère and Michalet-Doreau 1983). Data were gathered in order to obtain an index score of cell wall digestibility for each line from 1 (40%) to 5 (60%), according to the DINAGZ trait, whose heritability is most often close to 0.75 (Barrière et al. 2003). Lines F271 and F4 possessed the

Table 2 Pairs of primer used to amplify the complete coding region for two caffeoyl-CoA 3-*O*-methyltransferase genes, *CCoAOMT1*, *CCoAOMT2*, and the aldehyde *O*-methyltransferase gene, *AldOMT*, as well as 850 kb to 1 kb of the *AldOMT* gene 5′ regulatory region

Gene	Forward primer 5' to 3'	Position on cDNA	Reverse primer 5' to 3'	Position on cDNA	Reference cDNA
CCoAOMT1					
	CTCGTGCCCCAACGCGCTAGCTAGTTCTAT	1-30	GGCCAGGCCAGGGCATGTTTTCACTTGA	885-858	AJ242980
CCoAOMT2					
	CGCAAGCCAGTGCCGCGCCCAGATCTC	1-27	CTTGACGCGGCGGCAGAGCGTGACGCC	862-830	AJ242981
AldOMT					
5'UTR and Exon1	GGTGAGCCGTCCGGCCCAATAAAACCCCT	1-30	CATGAGGACCTTGTCCTGGTTCAT	544-520	M73235
Intron	ATGAACCAGGACAAGGTCCTCATG	521-544	GTTGAACGGGATGCCGCCGTC	1514-1494	M73235
Exon2	GACGGCGGCATCCCGTTCAAC	1494-1514	GCCCAGGCGTTGGCGTAGATG	2122-2102	M73235
AldOMT promoter					
AP1	GTAATACGACTCACTATAGGGC		CAGCTCGATGGCGTTCTTCAGCGTCAT	247-221	M73235
AP2	ACTATAGGGCACGCGTGGT		ACGTCGCCGGCGGTGGAGCCCATGGC	153-128	M73235
PCR	TCCACGGCAGCTGCCACCGTCGCTATCGC		ACGTCGCCGGCGGTGGAGCCCATGGC	153-128	M73235

lowest and highest cell wall digestibility in this set of maize lines, respectively.

Data analysis

DNA sequences were analyzed using DnaSP (Rozas et al. 2003). Levels of nucleotide diversity were estimated as mean pairwise differences (π) (Tajima 1983) and number of segregating sites (θ) (Watterson 1975). The minimum number of intra-genic recombination events was estimated using the four-gamete test (Hudson and Kaplan 1985). The Tajima (1989) and Fu and Li (1993) tests of selection were conducted without specifying an outgroup. Insertions and deletions (indels) were excluded from the estimates. The significance of pairwise linkage disequilibrium (LD) among polymorphic sites (SNPs and indels) was tested using Fisher's exact test excluding non-informative sites (singletons) and corrected for multiple analyses using the Bonferroni procedure (Sokal and Rohlf 1981). The decay of LD with physical distance along genes was evaluated by non-linear regression (PROC NLIN in SAS software, SAS Institute 1999) following Remington et al.'s model (2001) that considers potential low mutation rate and adjustment for sample size.

Statistical associations between variation in cell wall digestibility and each informative polymorphic sites (SNPs and indels) in the three OMT genes were tested using stepwise multiple linear regression (PROC REG in SAS software, SAS Institute 1999) with 1% criterion for entry and 1% for staying in the model. Population structure was not incorporated in these analyses.

Results

Organization of the three *OMT* genes

The *CCoAOMT1* coding region (about 1.2 kb) had four short introns of 108, 83, 134, and 112 bp (Fig. 1a). The *CCoAOMT2* coding region (about 1.1 kb) contained only three short introns (respectively, 78, 118, and 136 bp) at positions similar to those of introns I, II, and III in *CCoAOMT1*, indicating that the last intron was missing (Fig. 1a). *CCoAOMT1* and *CCoAOMT2* were approximately 83% and 95% identical at the nucleotide and amino acid levels, respectively. Intron positions were also well conserved compared to other published plant CCoAOMT genomic sequences (poplar, AJ223620 and AJ223621; parsley, Z33878).

As previously described for the inbred line W64A (Collazo et al. 1992), the *AldOMT* gene had an unique intron (approximately 949 bp, Fig. 1b). A 850–1.1-kb fragment of the *AldOMT* gene promoter was successfully amplified and sequenced in 30 maize lines. The putative transcription initiation site was found 147 bp upstream of the initiation codon, and a TATA box was found at position −28 (Fig. 1b). Numerous putative *cis*-acting regulatory elements, conserved in phenylpropanoid and flavonoid biosynthetic genes, were identified and are presented in Fig. 1b.

Nucleotide diversity

Nucleotide diversity for *CCoAOMT1*, *CCoAOMT2*, and *AldOMT* genes in maize is reported in Table 3. Values were estimated for the complete sequence and were also evaluated separately for 5′UTR, introns, silent, synonymous, and non-synonymous sites. The observed nucleotidic polymorphism was characterized by few singletons (polymorphic sites present only once in the sample), but numerous haplotypes (12 to 22). In addition, nucleotide diversity was not equally distributed among site categories: a higher diversity was observed in non-coding regions, especially in the introns as compared to coding regions where non-synonymous sites showed a very low diversity (Table 3). For each gene, the estimated π and θ were on average ten times higher for silent sites than for non-synonymous sites. Moreover, a large number of indels were present in non-coding regions. Out of the 95 indels identified for the three genes, 19% are single base pairs, 11% double base pairs, 10% triple base pairs, and 10% quadruple base pairs. Five to 10-bp indels were also frequent (32%).

Table 3 Nucleotide polymorphism at the *CCoAOMT1*, *CCoAOMT2*, and *AldOMT* genes

	Coding regions		Non-coding regions			All silent sites	Total
	Synonymous	Non-synonymous	5′UTR	Introns	Total		
CCoAOMT1 ($n = 34$, haplotype = 12)[a]							
Number of sites	186.27	587.73	90	404	515	701.27	1,328
S SNP (singl)	2	1	1	14 (3)	16 (3)	18 (3)	19 (3)
S Indel (singl)	0	0	1	10 (4)	11 (4)	11 (4)	11 (4)
π	0.0051	0.0007	0.0048	0.012	0.011	0.0094	0.0055
θ	0.0026	0.0004	0.0028	0.0084	0.076	0.0063	0.0036
CCoAOMT2 ($n = 32$, haplotype = 14)[a]							
Number of sites	188.11	591.89	80	331	413	601.11	1,221
S SNP (singl)	3	6 (1)	1	22	23	26	32 (1)
S Indel (singl)	0	4	5 (2)	18	23 (2)	23 (2)	27 (2)
π	0.0055	0.0019	0.0064	0.0244	0.0205	0.0157	0.0084
θ	0.004	0.0025	0.0038	0.222	0.018	0.014	0.0078
AldOMT[b] (Coding region $n = 34$, haplotype = 22; Total n = 30, haplotype = 26)[a]							
Number of sites	266.57	807.43	876	949	1,825	2,091.57	2,876
S SNP (singl)	8	7	22 (2)	50 (5)	72 (7)	80 (7)	87 (7)
S Indel (singl)	0	0	26 (10)	31 (3)	57 (13)	57 (13)	57 (13)
π	0.011	0.002	0.0097	0.0236	0.0208	0.0184	0.011
θ	0.0073	0.0021	0.008	0.017	0.015	0.013	0.009

[a] *n* Number of sequences in the sample, *S SNP* total number of SNPs, *S Indel* total number of insertions/deletions, π average rate of pairwise difference among sequences (Tajima 1983), θ number of segregating sites (Waterson 1975), *singl* number of singletons, which are indicated in *parentheses*.

[b] For the *AldOMT* gene, the 5′UTR region covers the promoter region sequenced for 30 lines. Consequently polymorphism data about 5′UTR, total non-coding region, all silent sites, and Total are provided from for 30 maize lines.

The nucleotide diversity values of *CCoAOMT2* were consistently higher than those of *CCoAOMT1*. The *CCoAOMT2* value for θ_{TOTAl} was 2.1 times that of *CCoAOMT1*, whereas its π_{TOTAL} value was 1.5 times that of *CCoAOMT1* (Table 3). These differences were mainly due to a larger diversity in introns and numerous replacements for *CCoAOMT2* (Table 3). More precisely, the *CCoAOMT1* gene had a θ value (0.0076) similar to that of the regulatory locus *c1* (Hanson et al. 1996), one of the least polymorphic loci studied in maize to date. However, haplotype frequency was similar between *CCoAOMT1* and *CCoAOMT2* (data not shown).

AldOMT was also twice as diverse as *CCoAOMT1*. For instance, we counted 65 SNPs in the *AldOMT* coding region (i.e. an average of one SNP every 31 bp) and 31 indels representing 14% of the whole sequence. Also, 20% of the *AldOMT* promoter region sites consisted of indels and 3.1% of SNPs, possibly explaining why, for some lines, the PCR failed. This high level of polymorphism also affected three putative regulatory *cis*-elements: MYB.ph3 binding site, and boxes P and IV (Fig. 1b).

Amino acid changes

Contrasting diversity between *CCoAOMT1* and *CCoAOMT2* was also found at the amino acid level. For the CCoAOMT1 protein, we found only one conservative amino acid change (Ile/Val) located in the second exon. The CCoAOMT2 protein had six amino acid replacements (Ala/Gly, Ala/Thr, Arg/Gln, Arg/His, Asp/Asn, and Asn/Ser). These polymorphisms were unevenly distributed along the *CCoAOMT2* sequence, most replacements being found in the first exon (Table 3). Moreover, four indels were found in first *CCoAOMT2* exon (Tables 2, 3), but none of them induces a frameshift. Polymorphic sites located in the first *CCoAOMT2* exon defined six haplotypes, among which two were unique (haplotypes 5 and 6, Table 3) to this set of lines.

Of the seven amino acid changes detected in *AldOMT* (Ser/Gly, Ala/Ser, Ala/Thr, Ala/Val, Ala/Pro, and two Asp/Asn), the moderately conservative amino acid change Ala/Val occurs in a highly conserved region (region IV, Ibrahim et al. 1998).

LD and recombination

An LD (non-random association of allele) study revealed many occurrences of significant LD among *CCoAOMT1* polymorphic sites (25.5% at $P < 0.001$, after Bonferroni's correction, Fig. 2b). Most of the significant associations within the four introns involved pairs of sites belonging to the same regions (Fig. 2b). For the *CCoAOMT2* and *AldOMT* genes, at $P < 0.001$, 1.4% and 0.85% of the pairs showed significant LD, respectively (Fig. 2a, c). For the *CCoAOMT2* gene, pairs of polymorphic sites in LD were evenly distributed along the 5′UTR and the three introns (Fig. 2c). In contrast, for the *AldOMT* gene, clusters of non-randomly associated sites were located only in the 5′-end promoter region and in the 5′ part of the unique intron (Fig. 2a). No LD between the three OMT genes was found (data not shown). The rate of LD decay with increasing physical distance was surprisingly rapid. This decay rate was more rapid in *CCoAOMT2* ($r^2 = 0.2$, distance ~200 bp) and *AldOMT* ($r^2 = 0.2$, distance ~255 bp) than in *CCoAOMT1* ($r^2 = 0.23$, distance ~1,227 bp). LD

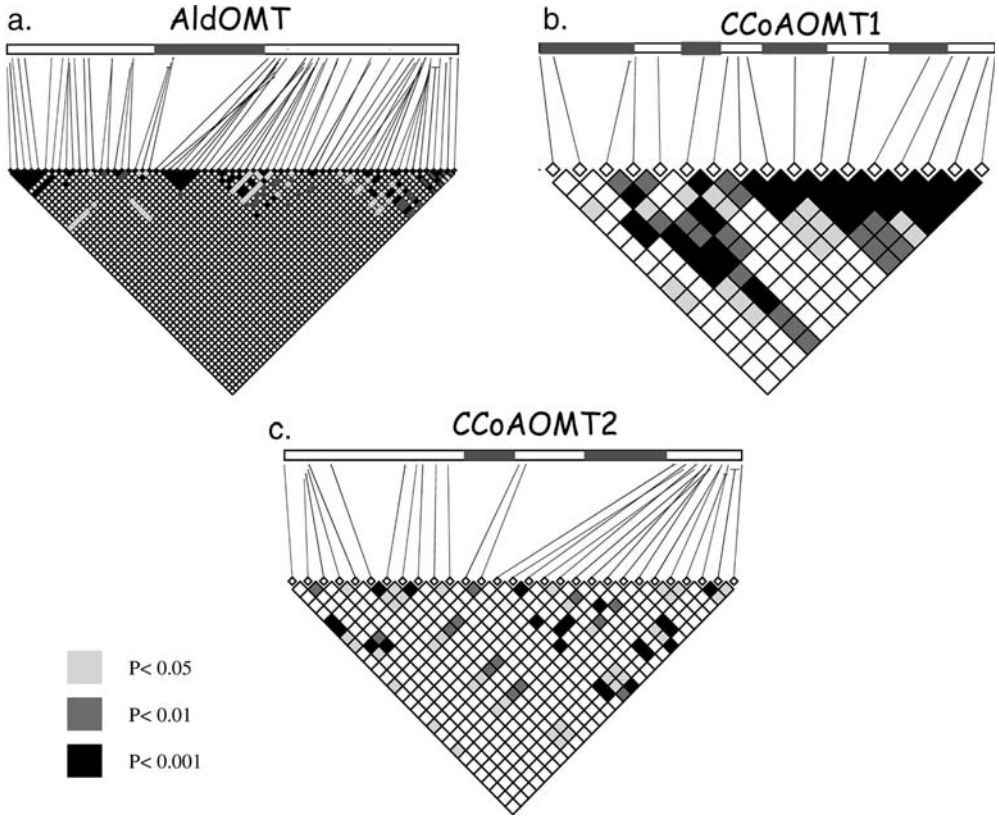

Fig. 2 Linkage disequilibrium matrix among *AldOMT* (**a**) *CCoAOMT1* (**b**), and *CCoAOMT2* (**c**) informative polymorphisms (SNPs and insertions/deletions) for 30 to 34 maize sequences. The *top part of graph* indicates the position of each polymorphic site along the three genes sequenced regions (exons represented as *black boxes*). *P*-values are the probabilities obtained using Bonferroni correction on Fisher's exact test

decay rate could have been influenced by the recombination frequencies along the gene (Rafalski and Morgante 2004). Therefore, the minimum number of intra-genic recombination events was estimated using Hudson and Kaplan's (1985) method. Similar recombination frequencies were found within each gene; six and four recombination events were detected from the *CCoAOMT1* and *CCoAOMT2* data sets, respectively. A minimum of nine recombination events were inferred from the 34 *AldOMT* coding region sequences, while fewer recombination events ($Rm = 3$) were detected in the *AldOMT* gene promoter region data set.

Mapping of the *CCoAOMT2* gene

Using an indel polymorphism (*lus1*) as a genetic marker, the *CCoAOMT2* gene was mapped on chromosome 9, in bin 9.02, where a cluster of QTL for cell wall digestibility and lignin content (measured by two standard methods: acid detergent lignin and Klason lignin) was located (Roussel et al. 2002). The *CCoAOMT2* gene could effectively have an essential role in lignification and maize cell wall digestibility.

Analysis of *CCoAOMT2* alleles in relation with cell wall digestibility

The 18-bp indel located in the first *CCoAOMT2* exon (S0159-S0176) resembled as SSR marker for which the increasing number of the 6-bp motif (AAC-GGC) appeared to be correlated with improved cell wall digestibility (Table 4). Indeed, association studies performed for a subset of 27 maize lines indicated that this indel was linked to cell wall digestibility ($P = 0.038$). Furthermore, the introduction of one repeat of this 6-bp motif was found in 15 genotypes with good cell wall digestibility. Two genotypes, F4 and Sibiriacka, have the 18-bp insertion (three repeats of the 6-bp motif, Table 4). Though the cell wall digestibility of Sibiriacka is not yet evaluated, F4 is an early flint French line with a high cell wall digestibility. Moreover, F4 and Sibiriacka had a 9-bp indel (S0099-S0107) leading to the loss of three amino acids (Glu-Ala-Thr, Table 4). The F4 allele had a singleton polymorphic site (S0208) that led to the conservative replacement of a His by an Arg (Table 4). Because of the interesting characteristics of F4, it was hypothesised that these amino acid modifications in the N-terminal region contributed to the constitution of a favourable *CCoAOMT2* allele for silage digestibility (Guillet and Barrière 2001, patent WO 03/054229).

Discussion

Variation at *CCoAOMT1*, *CCoAOMT2*, and *AldOMT* genes

CCoAOMT1, *CCoAOMT2*, and *AldOMT* genes are strategic targets for maize silage digestibility improvement. Values of nucleotide diversity at each gene were

Table 4 Nucleotide and amino acid polymorphisms in the *CCoAOMT2* gene first exon. Name of sites corresponds to their position on the *CCoAOMT2* consensus sequence

Site name	S0099 – S0104	S0105 – S0107	S0113 – S0111	S0114	S0127	S0159 - S0164	S0165 – S0176	S0226	S0208	S0236
Change[a]	I	n I	I	n	n	I	I	n[d]		
Haplotype1[b]	G A G G C G	A C C	A C G	A	A	– – – – – –	– – – – – – – – – – – –	C	A	C
Haplotype2	G A G G C G	A C C	A C G	A	A	A A C G G C	– – – – – – – – – – – –	C	A	C
Haplotype3	G A G G C G	**G** C C	– – –	**G**	A	– – – – – –	– – – – – – – – – – – –	C	A	**G**
Haplotype4	G A G G C G	A C C	A C G	A	**G**	A A C G G C	– – – – – – – – – – – –	C	A	C
Haplotype5 (Sibiriacka)	– – – – – –	– – –	A C G	A	A	A A C G G C	A A C G G C A A C G G C	**G**	A	C
Haplotype6 (F4)	– – – – – –	– – –	A C G	A	A	A A C G G C	A A C G G C A A C G G C	**G**	**G**	C
Amino acid modifications[c]	E A/– –	T/A/–	T/–	T/A	Q/R	N G/– –	– – – –/NGNG		H/R	

[a] *I* Indel, *n* non-synonymous SNP, a *dash* indicates a base or amino acid absence

[b] Haplotype 1 corresponds to the maize inbred lines B14, DE811, Mo17, W64A, F64, Lan496. Haplotype 2 corresponds to the maize inbred lines F271, F7025, F2, F286, F564, Wis93, Wis94, Line16, F324, F66, F1, EP1, Line212. The lines B73, F113, F288, F7012, MBS847, W117, Du101, Noordlander and Rainbow Flint correspond to haplotype 3. Polar Dent and Québec 28 have the haplotype 4

[c] The different amino acids encoded by replacement polymorphisms are shown *below*, with the most predominant amino acid shown *first*. Except for the deletions, all the amino acid replacements are conservative

[d] Singleton

comparable to those previously reported for other nuclear genes in maize, or even slightly higher. For instance, Tenaillon et al. (2001) and Rafalski (2002) found one SNP approximately every 48 bp and 130 bp in the 3′UTR and coding regions, respectively. In the flanking region of maize microsatellites, one SNP is found on average every 40 bp (Mogg et al. 2002). For the three complete OMT genes, we counted on average one SNP every 35 bp. Moreover, in the case of introns and non-coding regions, we observed a large number of indels varying from 1 bp to 109 bp in size. Indels are useful as genetic markers (Bhattramakki et al. 2002; Rafalski 2002), as demonstrated in the mapping of *CCoAOMT2* locus in a F288 × F271 RIL progeny using an indel (*lus1*) located in the second intron.

Based on comparative mapping analyses between maize and rice, both *CCoAOMT1* and *CCoAOMT2* map positions correspond to duplicated regions of the maize genome (Wilson et al. 1999). Despite the fact that the two CCoAOMT genes (1) are probably derived from gene duplication, (2) have the same function, and (3) are 83% similar at the DNA sequence level, they nevertheless displayed contrasting levels of nucleotide diversity. Different nucleotide diversity levels were reported for other duplicated genes (Small and Wendel 2002; Zhang et al. 2002) and may be explained by different evolutionary rates (Hudson et al. 1987). *CCoAOMT1* had low genetic diversity as compared to the colourless aleurone regulatory locus *c1*, suspected to have been under selection in maize (Hanson et al. 1996). However, Tajima's (1989) and Fu and Li's (1993) neutrality tests failed to show any evidence of selection for *CCoAOMT1*, and also for *CCoAOMT2* and *AldOMT* (data not shown). Previous studies of random maize loci have shown that departure from neutrality is rare (Tenaillon et al. 2001).

Similar recombination frequencies were reported for several maize genes (Hilton and Gaut 1998; Cummings and Clegg 1998), and these were in agreement with recent studies in maize, indicating that recombination occurs primarily in genes (for review, see Rafalski and Morgante 2004). Tenaillon et al. (2002) found that SNP diversity was correlated with the recombination rate estimated from DNA sequence data. Therefore, the high haplotype diversity observed in OMT genes could be due in part to the amount of recombination. For the three genes, LD decayed very rapidly, within a few hundred base pairs on average. A rapid LD decline was already described in maize [$r^2 = 0.2$, distance ~400 bp (Tenaillon et al. 2001); $r^2 = 0.2$, distance ~200–1,500 bp (Remington et al. 2001)]. As the extent of LD could depend on recombination frequency (Rafalski and Morgante 2004), numerous recombination events could explain why the sites in LD were mainly arranged within *AldOMT* in two main separate blocks. The high density of polymorphic sites and the rapid LD breakdown observed within these three candidate genes can favour the accurately identification of functional polymorphisms that could contribute to variation in cell wall digestibility (Remington et al. 2001; Thornsberry et al. 2001).

Potential implication of DNA polymorphism for variation in cell wall digestibility and gene function

The *CCoAOMT2* gene displayed six amino acid changes: a deletion of four amino acids and the insertion of six amino acids (i.e. 5.9% polymorphism). The latter involved stretches of a repetitive Asn/Gly motif and was reminiscent of microsatellites (Table 4). None of these modifications affected residues that can correspond to putative binding sites for Mg^{2+}, substrate or *S*-adenosyl-L-methionine (Hoffmann et al. 2001). In addition, most of the modifications observed for CCoAOMT2 were located in the N-terminal region, and the analysis of its secondary structure showed that this cluster of amino acid polymorphisms was responsible for some

structure modifications (data not shown). Because this N-terminal region is not well conserved between species (Hoffmann et al. 2001), the impact on the expression or the functionality of CCoAOMT2 was not obvious. However, the unique *CCoAOMT2* haplotype identified for F4, a line with a high cell wall digestibility, contained major polymorphisms located in the N-terminal region (Table 4). In *Brassica nigra*, the study of allelic variation for a major flowering-time locus also revealed one indel located in the coding regions that was strongly associated with early flowering (Osterberg et al. 2002). In *CCoAOMT2*, association studies showed that different non-synonymous SNPs had significantly associated with cell wall digestibility (data not shown), though these results need to be validated using a larger set of lines.

Of the seven amino acid replacements observed for AldOMT, only one moderately conservative amino acid change (Ala/Val) affected a highly conserved residue in an important region (region IV). The latter was proposed as a signature for plant OMT and was believed to be involved in metal binding (Ibrahim et al. 1998). Because this modification involved two inbred lines with different genetic background and silage digestibility, the real impact of this change on lignification was difficult to size. Association studies allowed us to identify sites associated with variation in cell wall digestibility, but all were located in non-coding regions (data not shown). The indel with the most significant association ($P = 0.002$; $R^2 = 0.32$) was a single-base pair deletion located in the intron and was found in a set of eight lines (Line 16, F7012, EP1, F324, F4, Du101, F66, and F113) having high cell wall digestibility (Table 1). This indel affected a UA-rich motif; yet, UA-rich sequences, typically distributed throughout the entire length of plant introns, have been showed to be required for efficient processing and splice-site selection (Lorkovic et al. 2000).

Lignification is a tightly regulated and dynamic process subject to modulation during development and in response to different stresses (Boerjan et al. 2003). The *AldOMT* gene promoter region had putative *cis*-acting elements frequently found in phenylpropanoid genes (Feuillet et al. 1995; Grimmig and Matern 1997; Neustaedter et al. 1999; Ito et al. 2000; Whitbred and Schuler 2000), confirming the work of Capellades et al. (1996), which showed that the *AldOMT* promoter was active in lignifying tissues and responded to wounding and elicitors. Changes observed in *cis*-regulatory elements (MYB.ph3 binding site, boxes P and IV) seem to have no impact in terms of lignification and silage digestibility, though the functionality of these putative elements remains to be proven. However, association studies realized on a subset of 24 inbred lines pointed out that one SNP, though it did not concern already identified *cis*-elements, was significantly associated with cell wall digestibility ($P = 0.0001$, $R^2 = 0.32$). This SNP was characteristic of six inbred lines (Line212, DE811, B14, B73, Mo17, and W64A) with low cell wall digestibility (Table 1).

Conclusions and perspectives

Analysis of allelic diversity in relation with cell wall digestibility values very likely validated *CCoAOMT2* and *AldOMT* as candidates for silage maize digestibility breeding, whereas the involvement of the *CCoAOMT2* gene was not so clearly evidenced. However, to further investigate the functional significance of the described SNPs or indels, investigations have to be considered on a significantly broader-based maize germplasm, including modern and old genotypes. Moreover, to take into account the existing population structure in association studies, this large collection of lines should be characterized using neutral SSR markers. Nevertheless, our results open an avenue for the further investigation of association studies. When relationships between a SNP or an indel and cell wall digestibility is thus confirmed, this favourable allele could be introgressed (using MAS) into elite germplasm, thus greatly accelerating silage maize improvement.

Acknowledgements This work was supported by grants from the Genoplante maize program, coordinated by Alain Charcosset (INRA Le Moulon) and Alain Murigneux (Biogemma, Les Cézeaux). We thank Sabine Guillaumie, V. Méchin and M. Pichon for comments on the manuscript. We thank Jeremy l'Homedet for skilful technical assistance. We are very grateful to P-Ph. Claude for encouragement and help during manuscript preparation.

References

Aufrére J, Michalet-Doreau B (1983) In vivo digestibility and prediction of digestibility of some by-products. IECC seminar, Belgique

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, Guillet C, Goffner D, Pichon M (2003) Genetic variation and breeding strategies for improved cell wall digestibility in annual forage crops. A review. Anim Res 52:193–228

Bhattramakki D, Dolan M, Hanafey M, Wineland R, Vaske D, Register JC III, Tingey SV, Rafalski A (2002) Insertion-deletion polymorphisms in 3′ regions of maize genes occur frequently and can be used as highly informative genetic markers. Plant Mol Biol 48:539–547

Boerjan W, Ralph J, Baucher M (2003) Lignin biosynthesis. Annu Rev Plant Biol 54:519–546

Buckler ES IV, Thornsberry J (2002) Plant molecular diversity and applications to genomics. Curr Opin Plant Biol 5:107–111

Capellades M, Torres MA, Bastisch I, Stiefel V, Vignols F, Bruce WB, Peterson D, Puigdomenech P, Rigau J (1996) The maize caffeic acid O-methyltransferase gene promoter is active in transgenic tobacco and maize plant tissues. Plant Mol Biol 31:307–322

Collazo P, Montoliu L, Puigdomenech P, Rigau J (1992) Structure and expression of the lignin O-methyltransferase gene from *Zea mays L.* Plant Mol Biol 20:857–867

Cummings MP, Clegg MT (1998) Nucleotide sequence diversity at the alcohol dehydrogenase 1 locus in wild barley (*Hordeum vulgare ssp. spontaneum*): an evaluation of the background selection hypothesis. Proc Natl Acad Sci USA 95:5637–5642

Feldbrugge M, Sprenger M, Hahlbrock K, Weisshaar B (1997) PcMYB1, a novel plant protein containing a DNA-binding do-

main with one MYB repeat, interacts in vivo with a light-regulatory promoter unit. Plant J 11:1079–1093
Feuillet C, Lauvergeat V, Deswarte C, Pilate G, Boudet A, Grima-Pettenati J (1995) Tissue- and cell-specific expression of a cinnamyl alcohol dehydrogenase promoter in transgenic poplar plants. Plant Mol Biol 27:651–667
Fu Y-X, Li W-H (1993) Statistical tests of neutrality of mutations. Genetics 133:693–709
Grimmig B, Matern U (1997) Structure of the parsley caffeoyl-CoA O-methyltransferase gene, harbouring a novel elicitor responsive cis-acting element. Plant Mol Biol 33:323–341
Guillet C, Barrière Y (2001) Polymorphismes du gène *CCoAOMT2* de maïs, et leurs utilisations pour améliorer la digestibilité des plantes. Patent WO 03/054229
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 Genet 5:19
Guo D, Chen F, Wheeler J, Winder J, Selman S, Peterson M, Dixon RA (2001) Improvement of in-rumen digestibility of alfalfa forage by genetic manipulation of lignin O-methyltransferases. Transgenic Res 10:457–464
Hanson MA, Gaut BS, Stec AO, Fuerstenberg SI, Goodman MM, Coe EH, Doebley JF (1996) Evolution of anthocyanin biosynthesis in maize kernels: the role of regulatory and enzymatic loci. Genetics 143:1395–1407
Hilton H, Gaut BS (1998) Speciation and domestication in maize and its wild relatives: evidence from the *globulin-1* gene. Genetics 150:863–872
Hoffmann L, Maury S, Bergdoll M, Thion L, Erard M, Legrand M (2001) Identification of the enzymatic active site of tobacco caffeoyl-coenzyme A O-methyltransferase by site-directed mutagenesis. J Biol Chem 276:36831–36838
Hudson RR, Kaplan NL (1985) Statistical properties of the number of recombination events in the history of a sample of DNA sequences. Genetics 111:147–164
Hudson RR, Kreitman M, Aguade M (1987) A test of neutral molecular evolution based on nucleotide data. Genetics 116:153–159
Ibrahim RK, Bruneau A, Bantignies B (1998) Plant O-methyltransferases: molecular analysis, common signature and classification. Plant Mol Biol 36:1–10
Ito H, Hiraga S, Tsugawa H, Matsui H, Honma M, Otsuki Y, Murakami T, Ohashi Y (2000) Xylem-specific expression of wound-inducible rice peroxidase genes in transgenic plants. Plant Sci 155:85–100
Lorkovic ZJ, Wieczorek Kirk DA, Lambermon MH, Filipowicz W (2000) Pre-mRNA splicing in higher plants. Trends Plant Sci 5:160–167
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
Mogg R, Batley J, Hanley S, Edwards D, O'Sullivan H, Edwards J (2002) Characterization of the flanking regions of *Zea mays* microsatellites reveals a large number of useful sequence polymorphisms. Theor Appl Genet 105:532–543
Neustaedter DA, Lee SP, Douglas CJ (1999) A novel parsley 4CL1 cis-element is required for developmentally regulated expression and protein–DNA complex formation. Plant J 18:77–88
Osterberg MK, Shavorskaya O, Lascoux M, Lagercrantz U (2002) Naturally occurring indel variation in the *Brassica nigra COL1* gene is associated with variation in flowering time. Genetics 161:299–306
Palaisa KA, Morgante M, Williams M, Rafalski A (2003) Contrasting effects of selection on sequence diversity and linkage disequilibrium at two phytoene synthase loci. Plant Cell 15:1795–1806
Piquemal J, Chamayou S, Nadaud I, Beckert M, Barriere Y, Mila I, Lapierre C, Rigau J, Puigdomenech P, Jauneau A, Digonnet C, Boudet AM, Goffner D, Pichon M (2002) Down-regulation of caffeic acid o-methyltransferase in maize revisited using a transgenic approach. Plant Physiol 130:1675–1685
Poke FS, Vaillancourt RE, C ER, Reid BR (2003) Sequence variation in two lignin biosynthesis genes, cinnamoyl CoA reductase (*CCR*) and cinnamyl alcohol dehydrogenase (*CAD2*). Mol Breed 12:107–118
Rafalski A (2002) Applications of single nucleotide polymorphisms in crop genetics. Curr Opin Plant Biol 5:94–100
Rafalski A, Morgante M (2004) Corn and humans: recombination and linkage disequilibrium in two genomes of similar size. Trends Genet 20:103–111
Remington DL, Thornsberry JM, Matsuoka Y, Wilson LM, Whitt SR, Doebley J, Kresovich S, Goodman MM, Buckler ES IV (2001) Structure of linkage disequilibrium and phenotypic associations in the maize genome. Proc Natl Acad Sci USA 98:11479–11484
Roussel V, Gibelin C, Fontaine AS, Barrière Y (2002) Genetic analysis in recombinant inbred lines of early dent forage. II. QTL mapping for cell wall constituents and cell wall digestibility from per se value and top cross experiments. Maydica 47:9–20
Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19:2496–2497
SAS Institute (1999) PROC REG software. SAS Institute, Cary
Small RL, Wendel JF (2002) Differential evolutionary dynamics of duplicated paralogous Adh loci in allotetraploid cotton (*Gossypium*). Mol Biol Evol 19:597–607
Sokal RR, Rohlf FJ (1981) Biometry. Freeman, San Francisco
Solano R, Nieto C, Avila J, Canas L, Diaz I, Paz-Ares J (1995) Dual DNA binding specificity of a petal epidermis-specific MYB transcription factor (MYB.Ph3) from *Petunia hybrida*. EMBO J 14:1773–1784
Tajima F (1983) Evolutionary relationship of DNA sequences in finite populations. Genetics 105:437–460
Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123:585–595
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
Tenaillon MI, Sawkins MC, Anderson LK, Stack SM, Doebley J, Gaut BS (2002) Patterns of diversity and recombination along chromosome 1 of maize (*Zea mays ssp. mays L.*). Genetics 162:1401–1413
Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680
Thornsberry JM, Goodman MM, Doebley J, Kresovich S, Nielsen D, Buckler ES IV (2001) *Dwarf8* polymorphisms associate with variation in flowering time. Nat Genet 28:286–289
Vignols F, Rigau J, Torres MA, Capellades M, Puigdomenech P (1995) The brown midrib3 (bm3) mutation in maize occurs in the gene encoding caffeic acid O-methyltransferase. Plant Cell 7:407–416
Watterson GA (1975) On the number of segregating sites in genetical models without recombination. Theor Popul Biol 7:256–276
Whitbred JM, Schuler MA (2000) Molecular characterization of *CYP73A9* and *CYP82A1 P450* genes involved in plant defense in pea. Plant Physiol 124:47–58
Wilson WA, Harrington SE, Woodman WL, Lee M, Sorrells ME, McCouch SR (1999) Inferences on the genome structure of progenitor maize through comparative analysis of rice, maize and the domesticated panicoids. Genetics 153:453–473
Zhang L, Peek AS, Dunams D, Gaut BS (2002) Population genetics of duplicated disease-defense genes, *hm1* and *hm2*, in maize (*Zea mays ssp. mays L.*) and its wild ancestor (*Zea mays ssp. parviglumis*). Genetics 162:851–860