Nucleotide diversity of the ZmPox3 maize peroxidase gene: Relationships between a MITE insertion in exon 2 and variation in forage maize digestibility

Carine Guillet-Claude*1,2, Christelle Birolleau-Touchard2, Domenica Manicacci3, Peter M Rogowsky4, Joan Rigau5, Alain Murigneux6, Jean-Pierre Martinant6 and Yves Barrière2

Address: 1Centre de recherche en Biologie Forestière, Université Laval, Sainte-Foy, Québec, G1K 7P4, Canada, 2Unité de Génétique et d’Amélioration des Plantes Fourragères, INRA, Route de Saintes, Lusignan, F86600, France, 3Station de Génétique Végétale, UMR INRA-UPS-CNRS-INAPG, Ferme du Moulon, Gif sur Yvette, F91190, France, 4Reproduction et Développement des Plantes, UMR 5667 INRA-CNRS-ENS-UCBL, IFR128 BioSciences Lyon-Gerland, ENS-Lyon, 46 allée d’Italie, Lyon, F69364, France, 5Departamento Genética Molecular, Instituto de Biología Molecular de Barcelona, Conseil Superior d’Investigacions Científiques, Jordi Girona 18–26, Barcelona, 08034, Spain and 6Biogemma S.A.S., Campus Universitaire des Cézeaux, 24 avenue des Landais, Aubière, F63170, France

Email: Carine Guillet-Claude* - cguillet@rsvs.ulaval.ca; Christelle Birolleau-Touchard - birollea@mons.inra.fr; Domenica Manicacci - manicacc@moulon.inra.fr; Peter M Rogowsky - peter.rogowsky@ens-lyon.fr; Joan Rigau - jrlgmr@ibmb.csic.es; Alain Murigneux - alain.murigneux@biogemma.fr; Jean-Pierre Martinant - jean-pierre.martinant@biogemma.fr; Yves Barrière - barriere@lusignan.inra.fr

* Corresponding author

Abstract

Background: Polymorphisms were investigated within the ZmPox3 maize peroxidase gene, possibly involved in lignin biosynthesis because of its colocalization with a cluster of QTL related to lignin content and cell wall digestibility. The purpose of this study was to identify, on the basis of 37 maize lines chosen for their varying degrees of cell wall digestibility and representative of temperate regions germplasm, ZmPox3 haplotypes or individual polymorphisms possibly associated with digestibility.

Results: Numerous haplotypes with high diversity were identified. Frequency of nucleotide changes was high with on average one SNP every 57 bp. Nucleotide diversity was not equally distributed among site categories: the estimated π was on average eight times higher for silent sites than for non-synonymous sites. Numerous sites were in linkage disequilibrium that decayed with increasing physical distance. A ZmPox3 mutant allele, carrying an insertion of a transposable element in the second exon, was found in lines derived from the early flint inbred line, F7. This element possesses many structural features of miniature inverted-repeat transposable elements (MITE). The mutant allele encodes a truncated protein lacking important functional sites. An ANOVA performed with a subset of 31 maize lines indicated that the transposable element was significantly associated with cell wall digestibility. This association was confirmed using an additional set of 25 flint lines related to F7. Moreover, RT-PCR experiments revealed a decreased amount of corresponding mRNA in plants with the MITE insertion.

Conclusion: These results showed that ZmPox3 could possibly be involved in monolignol polymerisation, and that a deficiency in ZmPox3 peroxidase activity seemingly has a negative effect on cell wall digestibility. Also, genetic diversity analyses of ZmPox3 indicated that this peroxidase could be a relevant target for grass digestibility improvement using specific allele introgressions.

Published: 16 July 2004

BMC Genetics 2004, 5:19 doi:10.1186/1471-2156-5-19

Received: 03 June 2004

Accepted: 16 July 2004

This article is available from: http://www.biomedcentral.com/1471-2156/5/19

© 2004 Guillet-Claude et al; licensee BioMed Central Ltd. This is an Open Access article: verbatim copying and redistribution of this article are permitted in all media for any purpose, provided this notice is preserved along with the article’s original URL.
Background

Because of its high energy content and good digestibility, silage maize is a major source of forage for dairy cattle. Its energy value depends mostly on the breakdown of cell walls by rumen micro-organisms. 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 digestibility [1].

Lignins are formed through dehydrogenative polymerization of two or three monolignols. Grass lignins are primarily made of two major units; guaiacyl (G, mono-methoxylated on the aromatic ring) and syringyl units (S, di-methoxylated), resulting of the polymerization of coniferyl and sinapyl alcohols, respectively. The last step in lignin biosynthesis transport of the monolignols to the cell wall thus involves the oxidation of monolignols within the plant cell wall matrix. The dehydrogenation to monolignols radicals is attributed to different classes of enzymes, such as peroxidases (oxidoreductases, EC 1.11.1.7), laccases (oxidoreductases, EC 1.10.3.2) and (poly)phenol oxidases [2,3]. However, the effects of modifying these enzymes on cell wall lignification and digestibility have not been investigated extensively in grasses. In tobacco, down-regulation of a peroxidase led to transformants with a reduced lignin content of up to 40–50% [4]. Down-regulation of an anionic peroxidase in transgenic aspen was associated with lower lignin content and modified lignin composition [5].

In plants, peroxidases belong to a multigene family. In maize, 13 peroxidase genes are described in The Maize Genetics and Genomics Database (http://www.maizegdb.org [6]). Three maize cDNAs coding for three different peroxidases (ZmPox1, ZmPox2 and ZmPox3) were isolated from a 9-day old tip-less root library [7]. Expression patterns of these peroxidases suggested that only ZmPox2 and ZmPox3 were involved in lignification. ZmPox3 was mapped by Biogemma SAS (A Murigneux, pers. com.) in the bins 6.05/07, and therefore co-localizes with a major QTL cluster for lignification and cell wall digestibility [8,9].

The breeding of forage crops for higher digestibility can involve specific genetic resources devised through genetic engineering of lignin biosynthesis [1]. In addition, the identification of natural allelic forms of candidate genes that correlate with higher digestibility, and their subsequent use in marker assisted selection (MAS) schemes, may become an interesting breeding tool. Gene sequence diversity has been primarily 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 [10-12]. However, Thornsberry et al [13] and Palaisa et al [14] 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 [15] found polymorphisms which could alter enzyme function in cinnamoyl CoA reductase (CCR) and cinnamyl alcohol dehydrogenase 2 (CAD2).

Allelic variation of the ZmPox3 peroxidase was studied in a set of 37 inbred lines or ecotypes representative of germplasm currently used in breeding programs and displaying a large range of variation in cell wall digestibility. A zmpox3 mutant allele, carrying an insertion of a transposable element in the second exon, was found in five lines and one ecotype. This element possesses many features of the MITE (miniature inverted repeat transposable element) class of mobile element [16]. This MITE insertion was further investigated to elucidate the relationship between this peroxidase mutant allele and cell wall digestibility.

Results

ZmPox3 gene diversity

The ZmPox3 gene, amplified for 37 diverse maize inbred lines or ecotypes (Table 1), is approximately 1.7 kb long, and the coding region contains two small introns of about 127 and 111 bp (Fig. 1A). The number of polymorphic sites and indel events for the complete ZmPox3 maize gene is reported in Table 2, along with two estimators of nucleotide diversity π [17] and θ [18]. Values are given for the complete sequence, as well as 3`UTR, introns, silent, synonymous and non-synonymous sites. Fourteen diverse haplotypes were found. There were 31 SNPs (i.e. an average of one SNP every 57 bp), and 17 indels which representing 20% of the whole sequence length. Of the 17 indels identified, 29% were single-bp, 23% were double-bp and 11% contained three-bp. Nucleotide diversity was not equally distributed among site categories. Values are given for the complete sequence, as well as 3`UTR, introns, silent, synonymous and non-synonymous sites. Fourteen diverse haplotypes were found. There were 31 SNPs (i.e. an average of one SNP every 57 bp), and 17 indels which representing 20% of the whole sequence length. Of the 17 indels identified, 29% were single-bp, 23% were double-bp and 11% contained three-bp. Nucleotide diversity was not equally distributed among site categories. The estimated π value was on average eight times higher for silent sites than for non-synonymous sites. Introns were particularly rich in SNPs and indels. Intronic regions had an average of one SNP every 14 bp while exons had only one SNP every 82 bp. The two ZmPox3 introns have 15 indels varying in size from 1 to 12 bp, while the 3`UTR region has 3 indels with size ranging from 2 to 5 bp.

For five inbred lines (F7, and its progenies F226, F227, F324 and F7012) and one ecotype (Québec28), the only indel found in the coding region was identified in the second exon of the ZmPox3 gene. This insertion corresponded to a transposable element sharing many features
of the MITE class of mobile element [16]. This element was short (321 bp), had 15 bp imperfect inverted repeats and was flanked by a 5 bp (CTCAG) direct repeat generated upon insertion.

**Amino acid changes**

According to the amino acid change classifications of Li *et al* [19] and Grantham [20], of the five non-synonymous SNPs found in ZmPox3, two changes were conservative (Val/Leu and Phe/Leu) and three were moderately conservative (Ser/Thr, Ala/Ser and Ala/Thr). The N terminal signal peptide (residue 1 to 22), and the active site as well as the heme-ligand site were not affected by any of these five amino acid replacements [7]. However, out of the 9 putative N-glycosylation sites identified [7] two were affected by one moderately conservative and one conservative amino acid changes.

**Description of the ZmPox3 MITE insertion**

The MITE insertion in the second exon disrupts the gene 129 bp downstream of the 3' splice site of the first intron (Fig. 1A) by introducing a stop codon 75 bp after the beginning of the insertion (Fig. 1A). The conceptual translation of the mutant allele results in a truncated protein of

---

**Table 1: Maize germplasm used in the association study.**

| Line     | Origin and Pedigree                                      | Cell wall digestibility |
|----------|---------------------------------------------------------|-------------------------|
| F1       | Lacaune (France)                                        | 3.5                     |
| F2       | Lacaune (France)                                        | 3.5                     |
| F7       | Lacaune (France)                                        | 4                       |
| F226     | Du101 × F7                                              | 4                       |
| F227     | F7 × F120                                               | 3.5                     |
| F7012    | F226 × F227                                             | 4.5                     |
| F4       | Etoile de Normandie (France)                            | 5                       |
| F66      | Soss (France)                                           | 3                       |
| Ep1      | Lizaragote (Spain)                                      | 4                       |
| F120     | F22 × Ep1 (with Fc22 from Chavannes)                    | 3                       |
| 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 F226, F227 and F286                           | 5                       |
| Line212  | Private Biogemma line                                   | 1                       |
| Line16   | Private Biogemma line                                   | 3.5                     |
| 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   | lodent                                                  | 2.5                     |
| F7025    | lodent × F113                                           | 2.5                     |
| B14      | Iowa Stoff Stalk Synthetic (Cl)                          | 1.5                     |
| B73      | Iowa Stoff Stalk Synthetic (BS13C5)                     | 1.5                     |
| Mo17     | Lancaster                                               | 2.5                     |
| Lan496   | Lancaster related                                        | 2                       |
| W64A     | W9 (from Reid Yellow dent) × Cl 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 | Rottaler silomais (Flint, Germany)                      | -                       |
|          | Noordlander VCI145 (Flint × dent, The Netherlands)      | -                       |
|          | Sibiriacka (Flint, Russia)                              | -                       |
|          | Rainbow flint (Canada)                                  | -                       |
|          | Polar dent (unknown)                                    | -                       |
|          | Québec28 (Flint, Canada)                                | -                       |
137 residues (Fig. 1B) the product of the translation of exon 1, the first third of exon 2 and the beginning of the MITE element. Since important functional sites (e.g. heme-ligand signature and putative N-glycosylation sites) are absent in the deduced truncated protein, maize inbred lines carrying this mutant allele are very likely deficient in ZmPox3 peroxidase activity.

Linkage disequilibrium and recombination

There were many occurrences of significant linkage disequilibrium (LD) among ZmPox3 polymorphic sites (20.6% of the tested LD are significant at P < 0.001) (Fig. 2). After correction for multiple tests using the Bonferroni procedure, 11.7% of LD remained significant (P < 0.001). LD concerned sites mainly located in the first intron. The transposable element (S0465) was in strong LD with six polymorphic sites located in the first two exons and in intron 1 and 2 (Fig. 2). These polymorphisms in the vicinity of the insertion site may have been linked to the mutation and inherited as such. LD decayed with increasing physical distance according to Remington’s model ([21], adjustment F (1.350) = 457, P < 0.0001) (Fig. 3). Three recombination events were detected in ZmPox3 when assessed as the minimum number of recombination events using the algorithm proposed by Hudson and Kaplan [22].

Relationship between ZmPox3 alleles and cell wall digestibility

ANOVA performed for a subset of 31 maize lines, for which digestibility values are known (Table 1), indicated that the insertion of the transposable element (MITE) in the second exon was significantly associated with cell wall digestibility (P = 0.032). This insertion characterized a group of five inbred lines related to F7 and having a high cell wall digestibility (F7, F7012, F324, F227 and F226). Based on the assumption that the lack of ZmPox3 peroxidase activity might lead to reduced or modified lignification, comparison of cell wall digestibility was investigated in 25 lines related to F7 and did or did not carry the zmpox3 mutant allele (Table 3). The cell wall digestibility was significantly higher in the 11 lines with the zmpox3 mutant allele. Lines having the MITE insertion were also more related to F7 than the lines having the normal allele, but this difference was not significant and the MITE effect remained significant when the percentage of F7 genome was added as a co-variable (Table 3). Moreover, it appeared that two genetic backgrounds were favorable to cell wall digestibility in this set of early flint lines, independently of MITE insertion; F7 (the Lacaze ecotype) and the germplasm from Argentina into which was bred F64 and is now found in lines F564, F286, F324, F7064 and F7065. Lines having a high cell wall digestibility (value ≥ 4), but not carrying the mutant allele, were all related to the Argentina germplasm (F286 and F564). In addition, line F324 carrying both the MITE insertion and Argentina germplasm, had a very high cell wall digestibility.

Table 2: Nucleotide polymorphism of the zmpox3 gene.

|              | Coding region | Non coding region | All silent sites | TOTAL |
|--------------|---------------|-------------------|-----------------|-------|
|              | Synonymous    | Non synonymous    | Introns 3’UTR   | Total |
| No of sites  | 272.12*       | 788.88*           | 238 135         | 373   | 645.12 | 1766 |
| S SNP (singl)| 8 5 (2)       | 17                | 1 18            | 26    | 31 (2) |
| S Indel (singl)| 0 1 **       | 14 (2)           | 3 17 (2)        | 17 (2) | 18 (2) |
| π             | 0.008         | 0.0018            | 0.032 0.0036    | 0.02  | 0.015 0.0075 |
| θ             | 0.007         | 0.0015            | 0.02 0.002     | 0.013 | 0.01 0.005 |
Expression profiles of the ZmPox3 gene and its mutant allele

To determine whether the mutant allele was transcribed, the ZmPox3 expression in F7012 (mutant zmPox3 allele) and Lan496 (normal ZmPox3 allele) were compared. Forward primer U19S1 located at the 5' end of the gene was combined either with primer U19R1 situated upstream or with primer U19MITEAS situated downstream of the MITE insertion in F7012 (Fig. 1A). The use of internal (18S) and external (pAW109) controls confirmed that both RNA quantity and RT-PCR were homogenous for samples obtained from stems with surrounding leaf sheaths of the upper or basal part of the plant (Fig. 4A, lines 2 and 3). No difference in signal intensity was observed between F7012 and Lan496 for the upstream part of ZmPox3 (Fig. 4A, line 4). On the contrary, the bands spanning the MITE insertion were barely visible in F7012 (Fig. 4A, line 5). This decrease for the mutant line could be explained by rapid RNA degradation of the non-translated part of the mutant allele, though we cannot exclude a decreased PCR efficiency due to possible secondary MITE structures. For experiments on leaf blades,
Decay of linkage disequilibrium with distance between nucleotide sites for ZmPox3. Curve shows non-linear regression of $r^2$ on distance, by using model described in Remington et al [21]. The regression coefficient b1 is 0.00462.

Moreover, ZmPox3 had a nucleotide diversity unequally distributed among site categories. There were 26 SNPs for all the silent sites and only five amino acids replacements for 357 residues (1.4%). Although the low frequency of non-synonymous polymorphisms suggested that at least some coding region of these genes are constrained, classical tests of neutrality showed no evidence of selection (data not shown). However, previous molecular diversity studies in maize have shown that departure from neutrality is rare [23]. A signature of positive selection could be revealed by an allele with unusually long-range LD [25]. In ZmPox3 however, LD decayed very rapidly with increasing physical distance, within 200 bp on average. This result is consistent with recent studies showing that LD typically decays rapidly within individual maize loci [10,21,23].

A moderately conservative amino acid changes, located in the second exon, affected a putative N-glycosylation site. This SNP was characteristic of a cluster of four high digestibility lines (F564, EP1, Wis94-443 and Wis93-3520) and was associated with cell wall digestibility (ANOVA; $P = 0.03$). It would be interesting to determine whether this amino acid replacement affects enzymatic activity. For example, in rice, Larkin and Park [26] have reported that two SNPs which caused amino acid substitutions in the waxy gene affected the endosperm apparent amylose content. The MITE insertion was also shown to be associated with cell wall digestibility (ANOVA; $P = 0.032$). Furthermore, a survey of an additional set of 25 lines related to F7 revealed significant association between MITE and cell wall digestibility. This MITE element of 321 bp is 93% identical to the 316 bp insertion observed in the B-M033 allele of the b gene, a regulatory gene involved in anthocyanin synthesis in maize [27]. It is also very similar to a transposon in intron 4 (92% identity) of the maize β-D-glucosidase gene (accession U60560), and to a 316 bp insertion in the 5’UTR region of the maize waxy gene (89% identity) [28]. It has been suggested that MITE insertions in non-coding regions such as the 5’UTR or 3’UTR of transcribed maize genes, provide regulatory sequences involved in transcription initiation or mRNA stability and are thus important tools of evolution [16]. Thornsberry et al [13] found significant association between flowering time traits and the 485 bp insertion in the 5’ non-coding region of the Dwarf8 gene caused by a putative MITE element. To our knowledge, the insertion of a MITE element in the second exon of ZmPox3 has never been previously reported in the coding region of maize genes. This insertion very likely resulted in transcription of a truncated and non-functional ZmPox3. Thereby, the high digestibility of numerous maize lines carrying the zmPox3 mutant allele could be considered as the result of a deficiency in the activity of this peroxidase probably involved in the lignin biosynthesis.
As reported by de Obeso et al [7], ZmPox3 mRNA was detected at low levels in lignifying tissues of young maize roots. Our RT-PCR experiments highlighted that the expression of the ZmPox3 gene was weaker in the first steps of stem lignification and in younger parts of the stem, than in older parts of the stem. These results, the well-known role of laccases and peroxidases at the final steps of lignin biosynthesis [7] and the identification of a ZmPox3 ortholog in the Zinnia elegans model system of lignifying tissues (S. Guillaumie, pers. com.) reinforce the possibility of an involvement of this peroxidase in maize monolignol polymerisation. The expression profile was clearly distinct from that of AldOMT indicating that lignification involves precise spatial and temporal regulation of the numerous genes involved in the different steps of this complex process.

Conclusions
Analysis of allelic diversity in relation with cell wall digestibility values likely validated ZmPox3 as a plausible candidate gene for silage maize digestibility improvement. This peroxidase could have a direct effect on plant cell wall digestibility, even more so given that this peroxidase is colocalized with a cell wall digestibility and lignification QTL. The comparison between ZmPox3 alleles that differ mostly by the presence or absence of the MITE element offers an opportunity to evaluate the impact of this kind of mutation event, as it was extensively done for

---

Figure 4
Expression profile of ZmPox3. Total RNA was isolated from the indicated tissues and amplified by RT-PCR. Amplification products were visualized by agarose gel electrophoresis. RNA transcribed from plasmid pAW109 was used as an external loading control. The ZmPox3 forward primer (U19S1) was combined with a reverse primer upstream (U19R1) or downstream (U19MITEAS) of the MITE insertion. PCR cycles used to evaluate ZmPox3 varied between panel A and panel B (indicated by numbers to the right of the gel pictures). A) ZmPox3 and AldOMT expression in lines F7012 (mutant allele) or Lan496 (normal allele). The stem with the leaf sheaths of 50 day old plants was divided into a basal part (4 inter-nodes) and an upper part (ending at the base of the tassel). In addition a 30 cm long leaf blade was harvested. B) ZmPox3 and AldOMT expression in different parts of line F2 (normal allele). At flowering the stem with the leaf sheaths was cut at the base of every node yielding units of node (N) and adjacent inter-node (i) with numbering starting at the base of the plant. In addition the ear leaves were harvested at defined stages expressed in days after pollination (DAP).
brown-midrib alleles [29]. Thus, either the targeted intro-
duction of the mutant allele in different genetic back-
grounds or the down-regulation in transgenic lines could
confirm the usefulness of $ZmPox3$ MITE disruption in
breeding maize for higher digestibility. Moreover, because
it seems to be a rare event, the search for the
$ZmPox3$ MITE
insertion in genetic resources could also help in determin-
ing the phylogeny of maize lines and introductions to
Europe.

Methods

Plant material

Thirty-seven inbred lines were chosen in order to display
a large variation in maize cell wall digestibility and to
represent a diversified sample of genotypes used both in
European breeding programs and representative of com-
monly used US germplasm (Table 1). Maize plants were
grown in vermiculite supplemented with a nutrient solu-
tion under a 16/8 h light/dark regime for 10 days. Because
an abnormal $zmPox3$ gene was first found in F7012, and
then in different parents of F7012 which were all related
to the INRA flint inbred F7, the existence of this mutant
allele was then investigated in 15 other private and public
lines related to F7 (Table 3). RT-PCR investigations of
$ZmPox3$ expression were done i) on successive nodes and
inter-nodes n° 5 to 14 in plants of line F2 harvested at
silking date, ii) on ear leaves of line F2 harvested 10, 21
and 40 days after silking, under field conditions, and iii)
in basal and upper plant parts (stem and sheaths), and in
the largest expended leaves of lines Lan496 and F7012
harvested 50 days after sowing in a greenhouse.

DNA extraction, primer design and PCR amplification

Genomic DNA was isolated from young maize leaves
using the plant DNAEASY miniprep kit (Qiagen).

A pair of primers was designed based on published cDNA
sequence (accession number AJ401276): U19S1 (forward;
5'-GACGAAGCGGCACTGCTTGCGCTTCACCA-3') and
U19AS1 (reverse; 5'-TGCCACAGTAACAAGCGAGCTTAC-
CAAGA-3'), respectively complimentary to positions 1-29
and to 1198-1170 of the maize $ZmPox3$ cDNA sequence.

Table 3: Comparisons of cell wall digestibility in 25 maize lines related to F7 with and without the MITE insertion in $ZmPox3$.

| Lines | Pedigree | % F7 | $ZmPox3$ MITE | Cell wall digestibility |
|-------|---------|------|----------------|------------------------|
| F7    | Lacaune ecotype | 100  | present        | 4                      |
| F192  | F7 × F2   | 50   | present        | 3                      |
| F226  | Du101 × F7 | 50   | present        | 3.5                    |
| F227  | F7 × F120 | 50   | present        | 3                      |
| F283  | F226 × F227 | 50 | present        | 3.5                    |
| F324  | (F282 × F283) × F286 | 62.5 | present  | 5                      |
| F7012 | F226 × F227 | 50   | present        | 4.5                    |
| F7032 | Early flint synthetic L | 25   | present        | 3.5                    |
| F7064 | F7012 × (BSSS × F286) | 43.7  | present        | 4.5                    |
| F7065 | F7012 × (BSSS × F286) | 43.7  | present        | 5                      |
| LGFS  | Early flint synthetic P | 12.5  | present        | 3.5                    |
| CPI718| Private line | 25   | absent         | 3                      |
| CPI1622| Private line | 25   | absent         | 3                      |
| F131  | F7 × CH10 | 50   | absent         | 2.5                    |
| F286  | F7 × F564 | 75   | absent         | 4.5                    |
| F268  | (F215 × F192) × (F2 × F160) | 12.5 | absent        | 1.5                    |
| F564  | F7 × F64 (F64 = Argentina flint) | 50   | absent         | 4                      |
| F7023 | Early flint synthetic T | 35   | absent         | 1.5                    |
| LG11  | Private line | 25   | absent         | 3                      |
| LG12  | Private line | 25   | absent         | 2                      |
| LG19  | Private line | 12.5  | absent         | 1.5                    |
| LGD3  | Private line | 20   | absent         | 2.5                    |
| R2n02 | Private line | 50   | absent         | 3.5                    |
| R2n12 | Private line | 25   | absent         | 2.5                    |
| R2n13 | Private line | 25   | absent         | 3                      |

Mean 11 lines with MITE present: 48.9% F7, 3.9
Mean 14 lines with MITE absent: 32.5% F7, 2.7
F value: 4.3 ns
F value with % F7 as covariable: 12.4 **
F value with % F7 as covariable: 6.6 **
Amplified fragments were about 1.4 to 1.7 kb long, and encompassed the coding region and 3’UTR (Table 4 and Fig. 1A).

A pair of primers flanking the MITE insertion was designed (Fig. 1A) as U19MITES (forward; 5’-GGCACTGGAGGCTCAGGGTGTGTT-3’) and U19MITEAS (reverse; 3’-AGGAGACAACGCCGGGGCAC-5’). If a maize line had the mutant allele, a 0.3 kb fragment corresponding to the MITE element was amplified. On plants having the normal allele, no amplification occurred (Table 4 and Fig. 1A). Combinations between pairs of primers U19S1 / U19AS1 and U19MITES / U19MITEAS allowed us to distinguish between homozygote and heterozygote genotypes (Table 4). The procedure to detect the zmPox3 mutant allele has been also described in detail in the French patent FR0302954.

PCR amplification reactions were performed in 50 µl containing 100 mM tris-HCl, 500 mM KCl, 15 mM MgCl₂ and 0.01% gelatin, 200 µM of each dNTP, 0.2 µM of 5’ oligo, 0.2 µM of 3’ oligo and 2.5 units of REDTaq DNA Genomic Polymerase (Sigma). 100 ng of genomic DNA was used as template. To amplify the ZmPox3 gene, 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. To genotype the MITE insertion in ZmPox3, the PCR amplification protocol was slightly modified and consisted of 5 min at 95°C followed by 25 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.

**DNA sequencing**

Sequencing was performed for each PCR fragments in both directions by Isoprim (France) and MWG-Biotech (Germany). Ecotype Québec28 was heterozygous at ZmPox3 locus and the two alleles have been sequenced. The sequences containing singletons were checked by re-amplifying genomic DNA and partially re-sequencing the appropriate alleles. Sequences were aligned using CLUSTALW [30]. Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under accessions n° AY500781-AY500812; AY508159-AY508163 and AY508516.

**RT-PCR analysis**

For isolation of total RNA, tissues were placed together with a stainless steel bead in 2 ml Eppendorf tubes and frozen in liquid nitrogen. Using pre-cooled holders the tissues were ground to powder in a Mixer Mill MM300 (Qiagen) by shaking two times 30 sec. The ground, frozen tissues were vortexed with 1 ml Trizol® Reagent (Invitrogen) at room temperature until melted. The aqueous phase resulting from a 10 min centrifugation at 18000 g and 4°C was re-extracted with 200 µl chloroform at room temperature. The RNA was precipitated with 500 µl isopropanol for 10 min at room temperature. The RNA pellet obtained by 10 min centrifugation at 18000 g and 4°C was washed with 1 ml 70% ethanol, dried and resuspended in 30 µl RNase free water. After treatment with RNase free DNase and inactivation of the DNase according to the instructions of the supplier (AMBIon) the RNA was quantified in a spectrophotometer at 260 nm. Approximately 5 µg of total RNA were reverse transcribed using random hexamers (Amersham) and reverse transcriptase without RNaseH activity (Fermentas). The 20 µl reverse transcription reaction also contained 2.5 x 10⁵ copies of GeneAmplimer pAW109 RNA (Applied Biosystems). The obtained cDNA was diluted 50 times in water and 5 µl used for amplification by PCR in a volume of 20 µl.

Part of ZmPox3 was amplified using primer U19S1 in combination with either primer U19R1 (5’-CGTCAGGTG-CCTACCGGTGTACGAC-3’) situated 84 bp upstream of the MITE insertion or primer U19MITEAS located downstream of the insertion. The constitutively expressed 18S rRNA gene (primers 5’ CCATCCCTCCG-TAGTTAGCTTCT 3’ and 5’ CCTGTCGGCGCCAG-GCTATAC 3’) was used as an internal control of RNA quantity and GeneAmplimer pAW109 RNA (primers 5’ CATGTCTACGTGTCATGC 3’ and 5’ TGACCACCCAGGCA-CTCC 3’) as positive control of the RT-PCR efficiency. In order to get semi-quantitative results, the

| Table 4: Pairs of primers used to obtain the ZmPox3 gene and to genotype the MITE insertion. |
|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|
|                               | U19S1  | U19AS1  | U19MITES | U19MITEAS |
| normal allele                 | 1.4 kb | none    | none     | 0.5 kb    |
| mutant allele                 | 1.7 kb | 0.3 kb  | 1.2 kb   | 0.8 kb    |
| heterozygote                  | 1.4 kb | 0.3 kb  | 1.2 kb   | 0.5 kb    |
|                               | 1.7 kb |         |          | 0.8 kb    |
number of cycles of the PCR reactions was adjusted for each gene to obtain barely visible bands in agarose gels. Aliquots of the PCR reactions were loaded on agarose gels and stained with ethidium bromide.

**Evaluation of maize 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 [8] 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 [31]. Data were gathered in order to obtain an index score of cell wall enzymatic solubility of the whole plant [31]. Values of lines F271 and F4, possessed the lowest and highest cell wall digestibility in this set of maize lines, respectively.

**Data analysis**

DNA sequences were analyzed using DnaSP [32]. Levels of nucleotide diversity were estimated as mean pairwise differences (\( \pi \); [17]) and number of segregating sites (\( \theta \); [18]). The minimum number of intragenic recombination events was estimated using the four-gamete test [22]. 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 [33]. 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's model [21] that considers potential low mutation rate and adjustment for sample size.

Each informative polymorphic site (SNPs and indels) was examined independently for an association with variation in cell wall digestibility using an ANOVA. The set of 31 lines for which digestibility values are known was used (Table 1).

A single factor variance analysis (MITE presence or absence) was used to compare the average digestibility value of lines, either having or not having the zmPox3 mutant allele, in the set of 25 lines related to F7 (Table 3). MITE effect was similarly investigated after adding the percentage of F7 genome in each line (expected value according to the pedigree) as a co-variable.

**List of abbreviations**

bp – base pairs, Indel – insertion/deletion, kb – kilo bases, MITE – miniature inverted-repeat transposable element, PCR – polymerase chain reaction, QTL – quantitative trait loci, SNP – single nucleotide polymorphism, UTR – untranslated region.

**Authors’ contributions**

CGC carried out most of the experimental studies and data analysis and participated in the discussion and preparation of the manuscript. CBT participated in the DNA sequencing and data analysis. DM carried out linkage disequilibrium measurement and participated in the discussion. PR carried out RT-PCR experiments. JR carried out the zmPox3 cloning. AM carried the zmPox3 mapping and participated in the design. JPM participated in the design and the coordination of the study. YB contributed to the design and coordination of the study and participated in the discussion and preparation of the manuscript.

**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 and Marie Fourmann for comments on the manuscript. We thank Jeremy I’Homedet, Richard Printemps and Stéphanie Barraud for skilful technical assistance. We also thank Pascal Condamine and Agnes Massonneau for their contribution to the RT-PCR analysis. Limagrain Genetics, Pau-Euralis and RAGT Semences are thanked for providing seeds of lines related to F7 and information on their cell wall digestibility. We are very grateful to Pierre-Philippe Claude for encouragement and help during manuscript preparation.

**References**

1. Barrière Y, Guillet C, Gofnner D, Pichon M: Genetic Variation and Breeding Strategies for Improved Cell Wall Digestibility in Annual Forage Crops. A Review. Annu Rev Plant Biol 2003, 54:519-546.
2. Boerjan W, Ralph J, Baucher M: Lignin biosynthesis. Annu Rev Plant Biol 2003, 54:519-546.
3. Boudet AM: Towards an understanding of the supramolecular organization of the lignified wall. The plant cell wall Edited by: J Rose. Blackwell publishing; 2003:155-182.
4. Blee KA, Choi JW, O’Connell AP, Schuch W, Lewis NG, Bolwell GP: A lignin-specific peroxidase in tobacco whose antisense suppression leads to vascular tissue modification. Phytochemistry 2003, 64:163-176.
5. Li Y, Kajita S, Kawai S, Katayama Y, Morohoshi N: Down-regulation of an anionic peroxidase in transgenic aspen and its effect on lignin characteristics. J Plant Res 2003, 116:175-182.
6. The Maize Genetics and Genomics Database [http://www.maizegdb.org/]
7. de Obeso M, Caparros-Ruiz D, Vignols F, Puigdomenech P, Rigau J: Characterisation of maize peroxidases having differential patterns of mRNA accumulation in relation to lignifying tissues. Gene 2003, 309:23-33.
8. Roussel V, Gibelin C, Fontaine AS, Barrière Y: 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 2002, 47:9-20.
9. Méchin V, Argillier O, Hébert Y, Guingo E, Moreau L, Charcosset A, Barrière Y: Genetic Analysis and QTL Mapping of Cell Wall Digestibility and Lignification in Silage Maize. Crop Sci 2001, 41:690-697.
10. Clark RM, Linton E, Messing J, Doebley JF. Pattern of diversity in the genomic region near the maize domestication gene tb1. *Proc Natl Acad Sci USA* 2004, 101:700-707.
11. Ching A, Caldwell KS, Jung M, Dolan M, Smith OS, Tingey S, Morgante M, Rafalski AJ. SNP frequency, haplotype structure and linkage disequilibrium in elite maize inbred lines. *BMC Genet* 2002, 3:19.
12. Buckler ES 4th, Thornberry JM. Plant molecular diversity and applications to genomics. *Curr Opin Plant Biol* 2002, 5:107-111.
13. Thornberry JM, Goodman MM, Doebley J, Kresovich S, Nielsen D, Thornsberry JM, Goodman MM, Doebley J, Kresovich S, Nielsen D. Dwarf8 polymorphisms associate with variation in flowering time. *Nat Genet* 2001, 28:286-289.
14. Palaisa KA, Morgante M, Williams M, Rafalski A. Contrasting effects of selection on sequence diversity and linkage disequilibrium at two phytocene synthase loci. *Plant Cell* 2003, 15:1795-1806.
15. Poke FS, Vaillancourt RE, C. Elliot R, Reid BR. Sequence variation in two lignin biosynthesis genes, cinnamyl CoA reductase (CCR) and cinnamyl alcohol dehydrogenase (CAD2). *Mol Breed* 2003, 12:107-118.
16. Wessler SR, Bureau TE, White SE. LTR-retrotransposons and MITEs: important players in the evolution of plant genomes. *Curr Opin Genet Dev* 1995, 5:814-821.
17. Tajima F. Evolutionary relationship of DNA sequences in finite populations. *Genetics* 1983, 105:437-460.
18. Watterson GA. On the number of segregating sites in genetical models without recombination. *Theor Popul Biol* 1975, 7:256-276.
19. Li WH, Wu CI, Luo CC. A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. *Mol Biol Evol* 1985, 2:150-174.
20. Grantham R. Amino acid difference formula to help explain protein evolution. *Science* 1974, 185:862-864.
21. Remington DL, Thornberry JM, Matsuoka Y, Wilson LM, Whitt SR, Doebley J, Kresovich S, Goodman MM, Buckler ES 4th. Structure of linkage disequilibrium and phenotypic associations in the maize genome. *Proc Natl Acad Sci U S A* 2001, 98:1479-1484.
22. Hudson RR, Kaplan NL. Statistical properties of the number of recombination events in the history of a sample of DNA sequences. *Genetics* 1985, 111:147-164.
23. Tenenon ML, Sawkins MC, Long AD, Gaut RL, Doebley JF, Gaut BS. Patterns of DNA sequence polymorphism along chromosome I of maize (Zea mays ssp. mays L.). *Proc Natl Acad Sci U S A* 2001, 98:9161-9166.
24. Rafalski A. Applications of single nucleotide polymorphisms in crop genetics. *Curr Opin Plant Biol* 2002, 5:94-100.
25. Rafalski A, Morgante M. Corn and humans: recombination and linkage disequilibrium in two genomes of similar size. *Trends Genet* 2004, 20:103-111.
26. Larkin PD, Park WD. Association of waxy gene single nucleotide polymorphism with starch characteristic in rice (Oriza sativa L.). *Mol Breed* 2003, 12:335-339.
27. Selinger DA, Chandler VL. Major recent and independent changes in levels and patterns of expression have occurred at the b gene, a regulatory locus in maize. *Proc Natl Acad Sci USA* 1999, 96:15007-15012.
28. Spell ML, Baran G, Wessler SR. An RFLP adjacent to the maize waxy gene has the structure of a transposable element. *Mol Gen Genet* 1988, 211:364-366.
29. Marita JM, Vermerris W, Ralph J, Hatfield RD. Variations in the cell wall composition of maize brown midrib mutants. *J Agric Food Chem* 2003, 51:1313-1322.
30. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucl Acid Res* 1994, 22:4673-4680.
31. Aufrère J, Michalet-Doreau B. In vivo digestibility and prediction of digestibility of some by-products. *ECC Seminar Belgique* 1983.
32. Rozas J, Sanchez-DeBariño JC, Meseguer X, Rozas R. DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* 2003, 19:2496-2497.
33. Sokal RR, Rohlf FJ. *Biometry*. Edited by: Ed W H Freeman. San Francisco; 1981.