

# Unraveling the Role of Maize Cell Wall Hydroxycinnamates in Digestibility, Biofuel Production and Pest Resistance Using A Multi-Parent Advanced Generation Intercross (MAGIC) Approach.

Ana López-Malvar (✉ [alopezmalvar@uvigo.es](mailto:alopezmalvar@uvigo.es))

University of Vigo

**Rosa Ana Malvar**

Misión Biológica de Galicia

**Ana Butrón**

Misión Biológica de Galicia

**Pedro Revilla**

Misión Biológica de Galicia

**Xose Carlos Souto**

University of Vigo

**Rogelio Santiago**

University of Vigo

---

## Research Article

**Keywords:** Zea mays, MAGIC, GWAS, cell wall hydroxycinnamates, ethanol, digestibility

**Posted Date:** December 11th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-121747/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

---

## Abstract

**Background:** Mechanical resistance due to higher hydroxycinnamate content makes maize tissues more recalcitrant to damage by insects, less digestible by ruminants, and less suitable for biofuel production. The integrated study of the maize functional genetic variability for each hydroxycinnamate component could be crucial to identify relevant genetic variants that may be incorporated into selection programs to breed maize varieties for multiple uses. A Genome Wide Association study was carried out in a maize Multiparent-Advanced Intercross (MAGIC) Population to identify Single Nucleotide Polymorphisms (SNPs) associated with cell wall bound hydroxycinnamates; and we checked thereafter their relationship with SNPs significantly associated with saccharification efficiency, digestibility of organic matter and corn borer damage.

**Results:** We found 24 SNPs, corresponding to 15 QTL, significantly associated with cell wall bound hydroxycinnamates. Each SNP explained between 6 and 8% of the total variability. We define new genomic regions and genes involved in polysaccharide synthesis and modifications, and the oxidative coupling associated to cell wall hydroxycinnamates content.

**Conclusions:** SNPs explained a small proportion of the variability for hydroxycinnamates, saccharification efficiency, digestibility or insect damage, therefore we recommend a genomic selection approach for future breeding programs of these traits. In addition, no colocalizations were found between hydroxycinnamates and final-use-related traits so breeding strategies can be focus on each particular trait with no side effects on the others.

## Background

The structural and functional properties of the plant cell wall are controlled by the composition and organization of each of its individual components [1]. In this sense, cell walls are mainly composed by microfibrils of celluloses embedded in a matrix of hemicellulose, lignin and other phenolic compounds (mostly hydroxycinnamic acids). Particularly, the role of stiffening, strengthening and fortification that hydroxycinnamic acid play within the cell wall has been either positively or negatively correlated with economically important uses and characteristics of cereal grasses such as resistance to insects, ethanol production and feedstock digestibility [2–4].

The most common hydroxycinnamates in those grasses are *p*-coumaric (PCA) and ferulic (FA) acid. PCA is involved in the radical coupling of S units of the lignin polymer and is known to be ester bond to the  $\gamma$  position of the side chains of the lignin S units [5]. On the other hand, FA is ester bond to arabinose residue of arabinoxylans chains, later crosslinked through ether bond and C-C bonds to the G unit of lignin, working as nucleation site of this polymer. Furthermore, FA is able to form links between proteins and polysaccharides through cysteine and tyrosine residues, and can form dimers (DFAs) and even oligomers through enzymatic or non-enzymatic oxidative coupling reactions, cross-linking this way hemicellulose polysaccharides. This net confers greater mechanical resistance to the cell wall [6–8].

The relationship between cell wall cross-linking and insect resistance was first suggested by Fry [1]. Since then, many studies have pointed out the role of hydroxycinnamate content in plant resistance to pests and diseases [9–12]. Greater amount of hydroxycinnamates monomers and diverse diferulic isomers, have been found in the pith tissues of borer resistant genotypes compared to susceptible ones. In addition, Barros-Rios et al. [11] found negative correlations between stem tunneling by borers and total DFAs, DFA-8-5' and PCA esters, or negative correlations between larval weight in feeding bioassays and total DFAs content. At first sight, modifications of their concentration would result in the improvement of borer resistance, however, a successful divergent selection program for changing diferulate ester concentration resulted in changes in the cell wall features [13]. When DFAs increased, glucose concentration and total cell walls concentration decreased, affecting properties related to animal digestibility [14]. Likewise, both, xylan-to-xylan ferulate linking [15] and ferulate-to-lignin cross-links [16, 17] limit the enzymatic depolymerization of cell wall polysaccharides increasing cell wall recalcitrance to deconstruction. This cell wall recalcitrance impacts negatively in digestibility, decreasing cell wall deconstruction by microorganisms and enzymes, thence limiting the energy intake by the cattle [18–20].

Finally, strategies that could reduce the incidence of ferulate cross-links in the cell wall have the potential to improve cell wall degradability properties relevant to cellulosic ethanol production [16]. Cell wall polysaccharides are the main substrate for ethanol fermentation, and cell wall recalcitrance increases the energy requirements, the cost and complexity of bio refinery operations and/or reduces the recovery of biomass carbon into desired products [21].

Some maize genomic regions mediating hydroxycinnamate accumulation have been detected throughout the genome [22–24]. Fontaine et al. [25], in a subset of 135 maize RILs, found seven QTL associated with FA and PCA. In addition, several QTL for hydroxycinnamate content, including PCA and FA, were identified by Barrière et al. [22] in a bi-parental population obtained from the cross F286 x F838, one of which co-localized with a QTL previously described by Fontaine et al. [25]. In the IBM population, Lorenzana et al. [26] found eight QTL associated with PCA and eight QTL associated with FA, whereas analyzing pith samples of second internode below the main ear in a biparental population

Santiago et al. [27] found two QTL associated with FA, seven with PCA and seven with total diferulates. Moreover, in one of the first high resolution association mapping analyses made to detect QTL associated to hydroxycinnamate content in the maize pith using a diversity panel, López-Malvar [24] found 22 QTL associated with cell wall bound hydroxycinnamates.

However, in the mentioned studies, mainly biparental populations have been used. In these populations you have limited variability, and although they are useful to identify QTL, the region where the QTL is located is usually quite wide. The use of inbred panels resolves both problems: they are diverse and enable high-resolution mapping of QTL to narrow genomic regions, where the searching for genes contributing to trait variability is more feasible. In this sense, as previously mentioned, a diversity panel of 272 inbreds was used to study cell wall bound hydroxycinnamates by our group [24]. Association studies using diversity panels still showing some disadvantages, as limited power to detect QTLs due to the small effect and/or low frequency (rare alleles) of some genetic variants, therefore we could have many undetected rare alleles that could be on the other hand valuable for breeding purposes when they show major effects [28–30]. Results from QTL mapping in Multi-Parent Advanced Generation Inter-Cross (MAGIC) populations could be complementary to those obtained by linkage mapping in biparental populations or association mapping using inbred panels. In this case several alleles can be simultaneously studied and none of them would be in low frequency. Besides, MAGIC populations present a known underlying structure that prevents from false positive associations [28]. Finally, in contrast to maize inbred panels constructed by inbreds of diverse origins, MAGIC populations, in general, are better adapted to a particular environment, so adaptation differences would not impair the phenotyping.

Therefore, a complete study of the maize functional genetic variability for cell wall hydroxycinnamates and final-use-related traits could be practical in order to identify relevant genetic combinations that may be incorporated into selection programs. The final objective will be to breed maize varieties for multiple purposes, without compromising each other.

To reach this goal we use cell wall bound hydroxycinnamates data and reports from three previous studies involving the same MAGIC population (López-Malvar et al. 2020 a,b submitted; [31]). We described genes in high resolution genomic regions..

## Methods

### Plant Materials

The MAGIC population was developed by the Maize Genetics and Breeding group at Mision Biologica de Galicia-CSIC. The eight parents of the MAGIC population were chosen because they show partial resistance to Mediterranean Corn Borer attack and high specific combining ability with the Reid germplasm group. Six of them come from European germplasm (EP17, EP43, EP53, EP86, PB130 and F473) and two from American germplasm (A509, EP125). Procedures used to release the 672 recombinant inbred lines of the population have been previously reported [31, 32]. The pedigree of each founder line is shown in Table 1.

### Experimental Design

A subset of 408 RILs of the MAGIC population together with the eight founders were tested in a single augmented design with 10 blocks in Pontevedra, Spain (42° 24'N, 8° 38'W and 20 m above sea level) in 2016 and 2017. Forty-two non replicated RILs plus the eight parents of the MAGIC (PB130 was replaced by EC212 in 2017 and EP43 by EP80 in both years due to lack of seed availability) were randomly assigned to each block. Only 30 RILs were evaluated in block 10. Each experimental plot consisted of a single row with 13 single-kernel hills planted manually, spacing between consecutive hills in a row being 0.18 m and 0.8 m between rows, obtaining a final density of ~ 70,000 plants ha<sup>-1</sup>. Local agronomical practices were fulfilled.

### Phenotypic Data

Biochemical analyses were performed in stover samples that were harvested approximately 55 days after silking (days from planting until half of the plants in the plot showed visible silks). Each sample was composed of tissue from at least two plants per plot. Samples were chopped, pre-dried at 35 °C in a forced air camera and dried at 60 °C in a stove. Last of all, dry stover samples from each plot were grounded in a Wiley mill with a 0.75 mm screen to be used in subsequent biochemical analysis.

A recently optimized protocol was used for hydroxycinnamate quantification [33]. Briefly, subsamples of 500 mg were extracted in 30 ml of 80% methanol and mixed using a Polytron mixer. Samples were extracted for 1 h and then centrifuged for 10 min at 1000 g. The remaining pellet containing the cell wall-bound material was shaken in 20 ml of 2 N NaOH under a nitrogen flow for 4 h. Digested samples were neutralized with 6 N HCl, and the pH was adjusted to 2.0. After centrifugation, the supernatant was collected and the pellet washed twice with distilled water (10 ml each). Supernatants were pooled and then extracted twice with ethyl acetate (40 ml each). Collected organic fractions were combined and dried using a SpeedVac for 6 h at a medium setting without a radiant cover. The final extract was dissolved in 1.5 ml of

HPLC grade methanol and stored at  $-20^{\circ}\text{C}$  prior to HPLC analysis. Standards and samples were filtered through a  $22\ \mu\text{m}$  tetrafluoroethylene filter before being analysed.

Chromatographic quantification was completed using a 2690 Waters Separations Module equipped with a model 996 photodiode array detector and a Waters YMC ODS-AM narrow bore column ( $100 \times 2\ \text{mm}$  internal diameter;  $3\ \mu\text{m}$  particle size). The solvent system consisted of acetonitrile (solvent A) and trifluoroacetic acid (0.05 %) in water (solvent B) as follows: initial conditions of 10:90 A:B, changing to 30:70 over 3.5 min, 32:68 over 6.5 min, 100:0 over 4 min, isocratic elution at 100:0 for 4.5 min, and finally returning to the initial conditions (10:90) over 3 min. The mobile phase flow rate was 0.3 ml/min, and the total analysis time was 17.5 min. The sample injection volume was  $4\ \mu\text{l}$ . Phenolic standards ferulic acid (FA) and *p*-coumaric acid (PCA) were purchased from Sigma-Aldrich Quimica SL, Madrid, Spain. The identities of FA dimers were confirmed by a comparison with the authentic 5-5 standard or published retention times and UV spectra. The total diferulate content (DFAT) was calculated as the sum of the following three identified and quantified DFA isomers: DFA 8-0-4, DFA 5-5, and DFA 8-5. The DFA 8-5 concentrations were calculated as the sum of 8-5-cyclic (or benzofuran)-DFA and 8-5-noncyclic (or open).

## Statistical Analysis

Inbred lines were previously genotyped with 955,690 SNPs using a genotyping-by-sequencing (GBS) method. Genotype matrix was filtered, i.e. SNPs with more than 50% missing data and a minor allele frequency less than 5% were omitted. Heterozygous genotypes were considered missing data. After filtering, 215,131 SNPs distributed across the maize genome were retained.

For each phenotypic trait, data from individual as well as from combined trials were analyzed according to the mixed model procedure (PROC MIXED) of the SAS program (version 9.4) [34] and the best linear unbiased estimator (BLUE) of each RIL was calculated based on the combined data for the 2-year analysis. The BLUES constituted the phenotype matrix. The comparison of means was carried out using the Fisher's protected least significant difference (LSD). Heritabilities ( $h^2$ ) were estimated for each trait on a family mean basis as previously described by Holland et al. [35]. The genetic ( $r_g$ ) and phenotypic ( $r_p$ ) correlation coefficients between each pair of traits were calculated using REML estimates according to a published SAS mixed model procedure [36].

We used the current field trials and population in previous studies in order to quantify the average saccharification efficiency, the digestibility of the organic matter (DOM) (López-Malvar et al. 2020 a,b submitted) and the tunnel length borers resistance [31]. In order to understand the interrelation among hydroxycinnamates and those final-use-related traits, multiple linear regression models using the stepwise method of the SAS PROC REG procedure. Saccharification efficiency, DOM and tunnel length were considered as dependent variables. Variables with a *p* value less than 0.15 were retained or excluded in the regression model.

The genome-wide association analysis was completed with Tassel 5 [37] based on a mixed linear model using a genotype-phenotype matrix and a kinship matrix using the centered IBS method [38]. Among the mixed linear model options, we used the optimum compression level and P3D to estimate the variance components.

## SNPs, QTL and Candidate Gene Selection

Two approaches were used to calculate the comparison-wise threshold for declaring significant an association between a trait and a SNP: (i) A modification of the classic Bonferroni approach where, first the number of independent tests was estimated by the Haploview program using the option four gamete rules, resulting in 12397 independent comparisons [39–41]. Then, the comparison-wise threshold was the coefficient between the experiment-wise threshold established (0.3) and the number of independent tests (12397); (ii) it was determined as the point where the observed and expected F test statistics deviated in the Q-Q plot of the model, performed with Tassel 5 [37] (Supplementary Figure 1).

We considered a  $\pm 700\ \text{kb}$  confident interval region around each significant SNP following previous association studies using the same mapping population [42]. In case confidence intervals of two SNPs overlapped they were assigned to a single QTL. The two described genes that delimit the  $\pm 700\ \text{kb}$  region around the SNP in the reference genome assembly version 2 were positioned in version 4 of the reference genome, and all genes contained in the region delimited by those genes were then identified and characterized based on the maize B73 reference genome assembly (version 4) available on the MaizeGDB browser [43]. Those genes are listed in Supplementary Table 1.

# Results

## Means, Analysis of Variance and Heritabilities

Means and ranks are shown in Table 2. Significant differences were found among the founders means and among RILs means for all the quantified hydroxycinnamates. *p*-coumaric and ferulic acid were the most abundant cell wall bound hydroxycinnamates. We found significant

variation for all traits; and also noted that the founder lines of the MAGIC population showed high variation for hydroxycinnamic acids as in previous evaluations [10]. Monomers presented moderate heritability values (0.59 for PCA and 0.60 for FA), whereas, dimers and total diferulate amount presented lower values (ranging from 0.32-0.42).

### Genotypic and Phenotypic Correlations

Correlation coefficients are detailed in Table 3. We found high positive genotypic and phenotypic correlation ( $r_{g,p} < 0.66$ ) between FA, total diferulates and individual dimers. On the other hand, very low correlation values were found between PCA and any other trait.

### Multiple Linear Regression Analysis

We looked for the multiple linear regression models that better fit data on saccharification efficiency, digestibility of the organic matter, and tunnel length damage using cell wall bound hydroxycinnamates as independent variables. For each model, the percentage of the partial variance for the dependent trait explained by each independent variable, the percentage of accumulated variance explained after incorporating that particular independent variable in the model, as well as estimates for regression coefficients are shown in Table 4.

The best model for saccharification efficiency only explained less than 1 % of the variance, mainly by PCA. Variation among RILs for DOM was partially determined by PCA (11 %) and DFA 8-5-I (2 %). By last, only 0.8 % of the variation for tunnel length was explained by cell wall bound hydroxycinnamates.

### Association Analysis

Following the modification of Bonferroni approach a marker was considered significantly associated with a trait at p values less than  $2.42 \times 10^{-5}$  ( $p$ -value = 4.6). For the Q-Q plots we stabilised that a marker was significantly associated with a trait with values less than  $1.00 \times 10^{-4}$  ( $p$ -value = 4.0). We followed the results of the Bonferroni modification approach for every trait with the exception of DFAT and DFA 5-5. We considered a +/-700kbp region around each significant SNP as the SNP confidence interval and two SNPs were clustered in the same QTL when their confidence intervals overlapped (Table 5).

A total of 24 SNPs, corresponding with 15 QTL, were significantly associated with cell wall bound hydroxycinnamates (Table 5). Each significant SNP explained phenotypic variances ranging from 5% to 8%; significant SNPs were distributed in bins 1.04, 1.07, 2.08, 3.09, 5.01, 5.06, 7.01, 7.02, 7.04, and 10.04. Minor and major frequency alleles contributed almost equally to increased DFA 5-5 and DFA 8-0-4 levels, while minor frequency alleles generally increased cell wall-bound DFA 8-5 and total diferulates concentrations; finally, major frequency alleles increased *p*-coumaric and ferulic acid concentrations.

### Candidate Gene Selection

The genes containing or physically close to SNPs significantly associated with traits were identified and characterized according to the maize B73 reference genome assembly (version 4). Analyses of +/- 700kbp regions surrounding significant SNPs resulted in the identification of the genes listed in Supplementary Table 1.

We highlight the following candidate genes within each QTL supporting interval for cell wall bound hydroxycinnamates: associated to PCA *MYB tf 41* (Zm00001d044117),  *$\omega$ -hydroxypalmitate O-feruloyl transferase* (Zm00001d024864) and *UDP-glycosyltransferase* (Zm00001d025037); associated to both FA and DFA 8-5, *Peroxidase* (Zm00001d017696) and *xyloglucan endotransglucosylase/hydrolase* (Zm00001d017699). As the candidates genes most probable for DFA 5-5 we found three *peroxidases*, we found associated both with DFA 5-5 and DFAT *UDP-glycosyltransferase* (Zm00001d020337). By last, within the supporting interval of QTLs for DFA 8-0-4 we found *UDP-glycosyltransferase* (Zm00001d020337), *Gibberellin-20 oxidase-2* (Zm00001d032223), *Peroxidase* (Zm00001d021965) and *Arabinosyltransferase RRA3* (Zm00001d021974) (Table 6).

## Discussion

First of all, we have observed genetic variation for cell wall bound hydroxycinnamates in the maize MAGIC population evaluated, for both the RILs and the founders agreeing with former evaluations [10]. The correlations between cell wall components traits followed the trends previously reported in the literature [2, 24, 44]. FA, particular dimers and DFAT showed co-variation. This means that if the target for improvement is FA, individual and total dimers would be also modified. On the other hand, this would not happen for PCA, as this trait was not correlated with any other trait.

Furthermore, moderate to high heritability for hydroxycinnamate contents agreed with the results obtained in previous studies [2, 24]. From those heritability values, we would expect a good response to selection, since additive effects are more important than additive  $\times$  environment

interaction effects. Although phenotypic selection could be effective based on those heritability estimates, a genomic selection approach could be implemented to speed up selection, in contrast to a time consuming phenotyping method.

### QTL Co-localization

With respect to other studies, novel genomic regions involved in hydroxycinnamate content were found, such as those in bins 10.04 (PCA), 5.06 and 7.01 (for FA), 2.08 and 7.02 (for DFA 5-5), 7.04 (for DFA 8-0-4), 5.06 (for DFA 8-5) and 5.01 and 7.02 (for DFAT). Nine of the ten SNPs associated with PCA, that correspond to three QTL, are located in bin 10.04. These QTL appear to be good candidates for selection targeting PCA. We found overlapping QTLs for FA and DFA 8-5 in the bin 5.06 and for DFA 5-5 and DFAT in bin 7.02. Those co-localizations have sense because genotypic correlations between those traits are high.

In addition, SNPs significantly associated with a trait in the current study co-localized with QTL for the same trait found by other authors. This is the case of the marker associated with PCA in bin 1.02, that co-localized in the same bin with one QTL associated with the same trait portrayed by García-Lara et al. [45], studying the cell wall phenolic composition of maize pericarp tissues in 163 F2:3 families; or the case of DFA 8-0-4 in bin 1.07 that co-localized with QTLs previously noted by López-Malvar et al. [24] using a maize diversity panel. Most interesting, in the same regions that we found QTL associated with cell wall bound hydroxycinnamates, QTLs for resistance to pests, animal digestibility, biofuel production and other cell wall components have already been reported. The aforementioned co-localizations helps to support the fact that the cell wall influences the final usage of maize.

However, despite the evidence that an indirect selection targeting cell wall hydroxycinnamates may be a good strategy for the improvement of final use of maize related traits, is not the case from the observed in the current MAGIC population. The low percentage of the variance of those traits explained by hydroxycinnamates, added to the lack of co-localizations among SNPs associated with hydroxycinnamates and SNPs associated with final use of maize related traits evaluated in previous research using the same MAGIC population ([31], López-Malvar et al. 2020a,b submitted), suggest the convenience of a direct and particular breeding strategy for the improvement of saccharification efficiency, digestibility of the organic matter or tunnel length damage resistance.

### Candidate Genes

Among genes proposed as candidates for cell-wall bound hydroxycinnamates we highlight peroxidases involved in oxidative coupling of FA to form dimers, genes that are responsible for transcriptional control of the phenylpropanoid pathway, implicated in xyloglucan and arabinoxylan synthesis, and involved in polysaccharides synthesis and modification. Finally, we also spot genes involved in gibberellin and suberin biosynthesis.

Cellulose and glucurono-arabinoxylans are the main constituents of lignified secondary walls; among the genes involved in the upstream parts of cell wall carbohydrate biosynthesis, we found several UDP-glycosyltransferases within the supporting intervals of QTL for PCA (qPCA\_10\_3), DFA 5-5 (qDFA5-5\_7\_1) and 8-0-4 (qDFA8-0-4\_1\_1), and total diferulates (qDFAT\_7\_1). UDP-glycotransferases are enzymes involved in the elongation of carbohydrate chains using nucleotide sugar as substrates and thereby causing variation in the cell wall polysaccharides structure [46]. For example, in rice, the gene underlying the brittle-culm-14 mutants, was a nucleotide sugar transporter that causes reduced mechanical strength by decreasing cellulose content and altering wall structure, including higher xylan extractability. In addition, we spotlight a *reduced residual arabinose 3* gene, which encodes an arabinosyltransferase that adds the second arabinose residue in a  $\beta$ -1,2 linkage in arabinoxylan chains, as candidate for the QTL qDFA8-0-4\_7\_1 [47]. Besides, FA is ester bond to arabinose residues of arabinoxylans chains, and FA dimers crosslink hemicellulose chains binding specifically to arabinose [48, 49]. Modifying arabinosyl transferase activities could be a promising strategy for modulating ferulate cross-linkages in the walls [50, 51], so we consider Zm00001d021974 a good candidate for DFA 8-0-4 content. Similarly, in the confidence interval of DFA 5-5, we found a glycosyl hydrolase (Zm00001d006940). Glycosyl hydrolases are mainly cellulases and xylanases; that modify the phenolic composition of cell walls as they cleave phenolic ester linkages. Thus, we considered Zm00001d006940, which encodes a glycosyl hydrolase, a good candidate for qDFA5-5\_2\_1. Besides, FA could be also bound to  $\alpha$ -xylosil of xyloglucan (XyG) [18] and diferulate crosslinking could also anchor the xyloglucan chains. XyG participates in the formation of cellulose- XyG network, allowing the attachment of cellulose to other wall polymers and playing an important role in cell wall extension during plant growth [52]. In this sense, we spot a *xyloglucan endotransglucosylase/hydrolase* gene that codifies for an enzyme involved in modifications of the xyloglucan for cell wall elongation as candidate for the overlapping QTL for FA and DFA 8-5.

Similarly, it has been hypothesized that GA could influence phenolic cross-linking via an effect on peroxidases [53]. As we mentioned throughout this work, phenolic groups can be oxidatively coupled by peroxidase + H<sub>2</sub>O<sub>2</sub> [54] and/or by oxidases (laccases) + O<sub>2</sub> [55] to form dimers. The conditions that increase the number of diferuoyl crosslinking included, among others, low gibberellin supply [56]. In this context,

within the supporting intervals of qDFA8-O-4\_1\_1, we highlight a *gibberlin 20-oxidase 2*, as candidate for DFA 8-O-4 content. In the same way, we also spotlight several peroxidase genes as candidates for FA and individual dimers in QTL qFA\_5\_1, qDFA8-5\_5\_1 and qDFA5-5\_2\_1.

Among the genes found in the interval of qPCA\_3\_1, we spot MYB tf 41 which has been classified by Du et al. [57] as involved in the “Phenylpropanoid Pathway”. Barrière et al. [58] proposed ZmMYB041 as the probable gene underlying a QTL for lignin content, and this could be also involved in PCA biosynthesis because increased lignification has been commonly associated to higher PCA concentration [59].

Finally, we found, in the QTL interval for PCA in chromosome 10, a gene encoding a  $\omega$ -hydroxypalmitate O-feruloyl transferase (Zm00001d024864) which takes part in the esterified suberin biosynthesis pathway. Suberin is a polymeric constituent of plant cell wall, which consists of two domains that are cross-linked. Concerning the aromatic fraction of suberin, hydroxycinnamates esters, such as PCA, fortify the crosslinking between arabinoxilans and suberin fatty acids; besides PCA deposition has been associated with highly suberized tissues. Even though there is some of the esterified PCA esterified to polysaccharides, as previously mentioned, most of the PCA in grasses is ester bond to S units of lignin [60, 61].

## Conclusions

To sum up, we pointed out new genomic regions associated to cell wall bound hydroxycinnamates in maize stover that could have an impact on their content across different genetic backgrounds. Based on their annotated functions the putative candidate genes for the traits under study are involved in polysaccharides synthesis and oxidative coupling, however, the genes that are not yet characterized, or the ones encoding a “hypothetical protein” that fell within QTLs intervals, could have an influence on the final hydroxycinnamates content. However, the effects of individual SNPs significantly associated with hydroxycinnamate content were low, and each SNP explained a low percentage of total genetic variability. Based on these results, and on the moderate heritability estimates observed, we suggest that the best breeding strategy to improve hydroxycinnamates content would be a genomic selection. On the other hand, we determine that the variation for cell wall bound hydroxycinnamates will not vary tunnel length damage, saccharification efficiency or forage digestibility, so we recommend carried out a direct selection program independently for each final use related trait.

## Declarations

### Ethics approval and consent to participate

Not applicable in this study.

### Consent for publication

Not applicable in this study.

### Availability of data and materials

The data sets used and/or analysed during the current study will be available upon reasonable request to the corresponding author.

### Competing interests

The authors declare that they have no competing interests.

### Funding

This research has been developed in the frame of the “Agri-Food Research and Transfer Centre of the Water Campus (CITACA)” at the University of Vigo (Spain), which is economically supported by the Galician Government, and in the Misión Biológica de Galicia (CSIC). It was funded by the “Plan Estatal de Ciencia y Tecnología de España” (projects RTI2018–096776-B-C21, and RTI2018–096776-B-C22 co-financed with European Union funds under the FEDER program). A. López-Malvar’s scholarship for the PhD fulfilment has been granted by University of Vigo and by a contract charged to the project RTI2018–096776-B-C22. The funding body played no role in study design, data analysis and manuscript preparation.

### Acknowledgements

We are grateful to Ana Carballada and David Figueroa for the performance of the cell wall bound hydroxycinnamates determinations.

### Author information

<sup>1</sup>Facultad de Biología, Departamento de Biología Vegetal y Ciencias del Suelo, Universidad de Vigo, As Lagoas Marcosende, Vigo 36310, Spain. Agrobiología Ambiental, Calidad de Suelos y Plantas (UVIGO), Unidad Asociada a la MBG (CSIC);

<sup>2</sup>Misión Biológica de Galicia (CSIC), Pazo de Salcedo, Carballeira 8, 36143, Spain

<sup>3</sup>E.E. Forestales, Dpto. Ingeniería Recursos Naturales y Medio Ambiente, Pontevedra 36005, Spain

### Corresponding author

\*Email: [alopezmalvar@uvigo.es](mailto:alopezmalvar@uvigo.es)

### Authors' contributions

RAM, RS conceived and designed the study. RAM, RS, AL, PR, and AB carried out the field trial and participated in sample collection; AL, RAM, RS performed data analysis; AL wrote the manuscript. XCS, contributed to the discussion of the manuscript. All authors reviewed and approved the final manuscript.

## References

1. Fry SC. Cross-Linking of Matrix Polymers in the Growing Cell Walls of Angiosperms. *Annu Rev Plant Physiol.* 1986;37:165–86.
2. Barrière Y, Méchin V, Riboulet C, Guillaumie S, Thomas J, Bosio M, et al. Genetic and genomic approaches for improving biofuel production from maize. *Euphytica.* 2009;170:183–202.
3. Santiago R, Barros-Rios J, Malvar RA. Impact of cell wall composition on maize resistance to pests and diseases. *Int J Mol Sci.* 2013;14:6960–80.
4. Lorenz AJ, Coors JG, Hansey CN, Kaeppler SM, de Leon N. Genetic analysis of cell wall traits relevant to cellulosic ethanol production in maize (*Zea mays* L.). *Crop Sci.* 2010;50:842–52.
5. Lu F, Ralph J. Detection and determination of p-coumaroylated units in lignins. *J Agric Food Chem.* 1999;47:1988–92.
6. Hartley RD, Harris PJ. Phenolic Constituents of the cell walls of dicotyledons. *Biochem Syst Ecol.* 1981;9:189–203.
7. Ralph J, Grabber JH, Hatfield RD. Lignin-ferulate cross-links in grasses: active incorporation of ferulate polysaccharide esters into ryegrass lignins. *Carbohydr Res.* 1995;275:167–78.
8. Grabber JH, Ralph J, Hatfield RD. Cross-linking of maize walls by ferulate dimerization and incorporation into lignin. *J Agric Food Chem.* 2000;48:6106–13.
9. Bergvinson DJ, Arnason JT. Phytochemical changes during recurrent selection for resistance to the European corn borer. *Crop Sci.* 1997;3.
10. Santiago R, Butron A, Arnason JT, Reid LM, Souto XC, Malvar RA. Putative role of pith cell wall phenylpropanoids in *Sesamia nonagrioides* (Lepidoptera: Noctuidae) resistance. *J Agric Food Chem.* 2006;54:2274–9.
11. Barros-Rios J, Malvar RA, Jung HJG, Santiago R. Cell wall composition as a maize defense mechanism against corn borers. *Phytochemistry.* 2011;72:365–71. doi:10.1016/j.phytochem.2011.01.004.
12. Assabgui RA. Correlation of Kernel (E)-Ferulic Acid Content of Maize with Resistance to *Fusarium graminearum*. *Phytopathology.* 1993;83:949.
13. Barros-Rios J, Malvar RA, Jung HJG, Bunzel M, Santiago R. Divergent selection for ester-linked diferulates in maize pith stalk tissues. Effects on cell wall composition and degradability. *Phytochemistry.* 2012;83:43–50. doi:10.1016/j.phytochem.2012.07.026.
14. Barros-Rios J, Malvar RA, Jung HJG, Bunzel M, Santiago R. Divergent selection for ester-linked diferulates in maize pith stalk tissues. Effects on cell wall composition and degradability. *Phytochemistry.* 2012;83:43–50. doi:10.1016/j.phytochem.2012.07.026.
15. Grabber JH, Ralph J, Hatfield RD. Ferulate Cross-Links Limit the Enzymatic Degradation of Synthetically Lignified Primary Walls of Maize. *J Agric Food Chem.* 1998;46:2609–14. doi:10.1021/JF9800099.
16. Grabber JH, Mertens DR, Kim H, Funk C, Lu F, Ralph J. Cell wall fermentation kinetics are impacted more by lignin content and ferulate cross-linking than by lignin composition. *J Sci Food Agric.* 2009;89:122–9.
17. Grabber JH, Ralph J, Hatfield RD. Severe inhibition of maize wall degradation by synthetic lignins formed with coniferaldehyde. *J Sci Food Agric.* 1998;78:81–7.
18. Ishii T, Hiroi T, Thomas JR. Feruloylated xyloglucan and p-coumaroyl arabinoxylan oligosaccharides from bamboo shoot cell-walls. *Phytochemistry.* 1990;29:1999–2003.

19. Jung HG, Ralph J. Phenolic-carbohydrate complexes in plant cell walls and their effect on lignocellulose utilization. *Microb plant Oppor to Improv Lignocellul Util by ruminants*. 1990;:173–82.
20. Besle J -M, Cornu A, Jouany J -P. Roles of structural phenylpropanoids in forage cell wall digestion. *J Sci Food Agric*. 1994;64:171–90.
21. Vermerris W, Saballos A, Ejeta G, Mosier NS, Ladisch MR, Carpita NC. Molecular breeding to enhance ethanol production from corn and sorghum stover. *Crop Sci*. 2007;47 SUPPL. DEC.
22. Barrière Y, Thomas J, Denoue D. QTL mapping for lignin content, lignin monomeric composition, p-hydroxycinnamate content, and cell wall digestibility in the maize recombinant inbred line progeny F838 × F286. *Plant Sci*. 2008;175:585–95.
23. Riboulet C, Fabre F, Dénoue D, Martinantä JP, Lefèvre B, Barrière Y. QTL mapping and candidate gene research for lignin content and cell wall digestibility in a top-cross of a flint maize recombinant inbred line progeny harvested at silage stage. *Maydica*. 2008;53:1–9.
24. López-Malvar A, Butrón A, Samayoa LF, Figueroa-Garrido DJ, Malvar RA, Santiago R. Genome-wide association analysis for maize stem Cell Wall-bound Hydroxycinnamates. *BMC Plant Biol*. 2019;19:1–12.
25. Fontaine AS, Briand M, Barrière Y. Genetic variation and QTL mapping of para-coumaric and ferulic acid contents in maize stover at silage harvest. *Maydica*. 2003;48:75–84.
26. Lorenzana RE, Lewis MF, Jung HJG, Bernardo R. Quantitative trait loci and trait correlations for maize stover cell wall composition and glucose release for cellulosic ethanol. *Crop Sci*. 2010;50:541–55.
27. Santiago R, Malvar RA, Barros-Rios J, Samayoa LF, Butrón A. Hydroxycinnamate Synthesis and Association with Mediterranean Corn Borer Resistance. *J Agric Food Chem*. 2016;64:539–51.
28. Myles S, Peiffer J, Brown PJ, Ersoz ES, Zhang Z, Costich DE, et al. Association mapping: Critical considerations shift from genotyping to experimental design. *Plant Cell*. 2009;21:2194–202.
29. Yan J, Warburton M, Crouch J. Association mapping for enhancing maize (*Zea mays* L.) genetic improvement. *Crop Sci*. 2011;51:433–49.
30. Bernardo R. Bandwagons I, too, have known. *Theor Appl Genet*. 2016;129:2323–32.
31. Jiménez-Galindo JC, Malvar RA, Butrón A, Santiago R, Samayoa LF, Caicedo M, et al. Mapping of resistance to corn borers in a MAGIC population of maize. *BMC Plant Biol*. 2019;19:1–17.
32. Butrón A, Santiago R, Cao A, Samayoa L, Malvar R. QTLs for Resistance to Fusarium Ear Rot in a Multiparent Advanced Generation Intercross (MAGIC) Maize Population. *Plant Dis*. 2019;103:897–904.
33. Santiago R, López-Malvar A, Souto C, Barros-Ríos J. Methods for Determining Cell Wall-Bound Phenolics in Maize Stem Tissues. *J Agric Food Chem*. 2018;66:1279–84.
34. SAS/STAT; SAS Institute Inc.: Cary, NC, 2007.
35. Holland JB, Nyquist WE, Cervantes-Martínez CT. Estimated and interpreting heritability for plant breeding: An update. *Plant Breed Rev*. 2003;22:9–122.
36. Holland JB. Estimating genotypic correlation and their standard errors using multivariate restricted maximum likelihood estimation with SAS Prco MIXED. *Crop Sci*. 2006;46:642–56.
37. Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES. TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics*. 2007;23:2633–5.
38. Endelman JB, Jannink JL. Shrinkage estimation of the realized relationship matrix. *G3 Genes, Genomes, Genet*. 2012;2:1405–13.
39. Li J. Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. *Heredity (Edinb)*. 2005;95:221–7.
40. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: Analysis and visualization of LD and haplotype maps. *Bioinformatics*. 2005;21:263–5.
41. Wang N, Akey JM, Zhang K, Chakraborty R, Jin L. Distribution of recombination crossovers and the origin of haplotype blocks: The interplay of population history, recombination, and mutation. *Am J Hum Genet*. 2002;71:1227–34.
42. Yi Q, Malvar RA, Álvarez-Iglesias L, Ordás B, Revilla P. Dissecting the genetics of cold tolerance in a multiparental maize population. *Theor Appl Genet*. 2020;133:503–16.
43. Andorf C, Cannon E, Portwood J, Gardiner J, Harper L, Schaeffer M, et al. MaizeGDB update: new tools, data and interface for the maize model organism database. *Nucleic acids research*. 2015; D1: 44, D1195-D1201.
44. Barros-Rios J, Santiago R, Malvar RA, Jung HJG. Chemical composition and cell wall polysaccharide degradability of pith and rind tissues from mature maize internodes. *Anim Feed Sci Technol*. 2012;172:226–36. doi:10.1016/j.anifeedsci.2012.01.005.
45. García-Lara S, Burt AJ, Arnason JT, Bergvinson DJ. QTL mapping of tropical maize grain components associated with maize weevil resistance. *Crop Sci*. 2010;50:815–25.

46. Kawakita M, Ishida N, Miura N, Sun-Wada GH, Yoshioka S. Nucleotide sugar transporters: Elucidation of their molecular identity and its implication for future studies. *J Biochem.* 1998;123:777–85.
47. Egelund J, Obel N, Ulvskov P, Geshi N, Pauly M, Bacic A, et al. Molecular characterization of two *Arabidopsis thaliana* glycosyltransferase mutants, *rra1* and *rra2*, which have a reduced residual arabinose content in a polymer tightly associated with the cellulosic wall residue. *Plant Mol Biol.* 2007;64:439–51.
48. Bunzel M, Ralph J, Steinhart H. Phenolic compounds as cross-links of plant derived polysaccharides. *Czech J Food Sci.* 2004;22:39–42. <http://www.dfrc.ars.usda.gov/DFRCWebPDFs/2004-Bunzel-CzechJFS-22-64.pdf>.
49. Hatfield RD, Marita JM. Enzymatic processes involved in the incorporation of hydroxycinnamates into grass cell walls. *Phytochem Rev.* 2010;9:35–45.
50. Chiniquy D, Sharma V, Schultink A, Baidoo EE, Rautengarten C, Cheng K, et al. XAX1 from glycosyltransferase family 61 mediates xylosyltransfer to rice xylan. *Proc Natl Acad Sci U S A.* 2012;109:17117–22.
51. Courtial A, Soler M, Chateigner-Boutin A-L, Reymond M, Mechin V, Wang H, et al. Breeding grasses for capacity to biofuel production or silage feeding value: an updated list of genes involved in maize secondary cell wall biosynthesis and assembly. *Maydica.* 2013;58:67–102. <http://cra-journals.cineca.it/index.php/maydica/article/view/918>.
52. Pauly M, Keegstra K. Cell-wall carbohydrates and their modification as a resource for biofuels. *Plant J.* 2008;54:559–68.
53. Gaspar T. Peroxidases 1970–1980. 1970.
54. Geissmann T V., Neukom H. Vernetzung von Phenolcarbonsäureestern von Polysacchariden durch oxydative phenolische Kupplung. *Helv Chim Acta.* 1971;54:1108–12.
55. Ranocha P, McDougall G, Hawkins S, Sterjiades R, Borderies G, Stewart D, et al. Biochemical characterization, molecular cloning and expression of laccases - A divergent gene family - In poplar. *Eur J Biochem.* 1999;259:485–95.
56. Fry SC. Gibberellin-controlled pectinic acid and protein secretion in growing cells. 1980;19:735–40.
57. Du H, Feng BR, Yang SS, Huang YB, Tang YX. The R2R3-MYB transcription factor gene family in maize. *PLoS One.* 2012;7:1–12.
58. Barrière Y, Courtial A, Soler M, Grima-Pettenati J. Toward the identification of genes underlying maize QTLs for lignin content, focusing on colocalizations with lignin biosynthetic genes and their regulatory MYB and NAC transcription factors. *Mol Breed.* 2015;35.
59. Méchin V, Argillier O, Menanteau V, Barrière Y, Mila I, Pollet B, et al. Relationship of cell wall composition to in vitro cell wall digestibility of maize inbred line stems. *J Sci Food Agric.* 2001;80 November 1999:574–80.
60. Gou JY, Yu XH, Liu CJ. A hydroxycinnamoyltransferase responsible for synthesizing suberin aromatics in *Arabidopsis*. *Proc Natl Acad Sci U S A.* 2009;106:18855–60.
61. Figueiredo R, Araújo P, Llerena JPP, Mazzafera P. Suberin and hemicellulose in sugarcane cell wall architecture and crop digestibility: A biotechnological perspective. *Food Energy Secur.* 2019;8:1–24.

## Tables

Table 1

Characteristics of the maize inbred lines used as founders of the MAGIC population, along with the checks EC212 and EP80.

Inbred lines	Grain color	Pedigree	Type of grain
<b>Founders</b>			
A509	Yellow	A78 × A109	Dent
EP17	Yellow	Selection from A1267	Flint
EP43	Yellow	Parderrubias (landrace from Atlantic Spain)	Flint
EP53	Yellow	Laro (landrace from Atlantic Spain)	Flint
EP86	Yellow	Nostrano dell'Isola (landrace from Italy)	Flint
EP125	Yellow	Selection from CO125	Dent
F473	White	Doré de Gomer (landrace from France)	Flint
PB130	Yellow	Rojo Vinoso de Aragón (landrace from Mediterranean Spain)	Flint
<b>Inbred lines used as checks</b>			
EC212	Yellow	EC203B×EC18	Flint
EP80	Yellow	Selection from EA2087, released from Azpeitia (landrace from North of Spain)	Flint

Table 2

Means, heritabilities and rank observed for cell wall component traits in the RILs of a maize MAGIC population evaluated during two years. Means for the checks are also included.

	PCA (mg/g)	FA (mg/g)	DFA 5-5 (mg/g)	DFA 8-0-4 (mg/g)	DFA 8-5 (mg/g)	DFAT (mg/g)
<b>RILs</b>						
Means	7.66	3.35	0.24	0.30	0.27	0.82
$h^2 \pm SE^1$	0.59 ± 0.04	0.60 ± 0.04	0.36 ± 0.06	0.42 ± 0.06	0.32 ± 0.06	0.40 ± 0.06
Rank	3.97-10.90	1.96-5.13	0.06-0.42	0.08-0.79	0.06-0.73	0.22-1.43
<b>LSD<sup>2</sup></b>	1.98	1.19	0.13	0.16	0.14	0.39
<b>Checks</b>						
A509	7.97	3.16	0.28	0.34	0.33	0.96
EP125	7.67	3.63	0.25	0.32	0.31	0.89
EP17	5.86	4.47	0.31	0.39	0.37	1.07
EP53	5.73	2.65	0.21	0.28	0.25	0.74
EP86	9.24	2.64	0.21	0.23	0.19	0.64
F473	7.10	3.65	0.29	0.37	0.26	0.84
EP80	9.26	4.23	0.29	0.33	0.34	0.96
EC212	7.76	3.51	0.18	0.25	0.24	0.66
<b>LSD</b>	1.58	0.43	0.05	0.06	0.10	0.14
PCA: <i>p</i> -coumaric acid; FA: ferulic acid; DFA 5-5: 5-5 diferulic acid; DFA 8-0-4: 8-0-4 diferulic acid; DFA 8-5: sum (DFA 8-5 I + DFA 8-5b); DFAT: total diferulates sum (DFA 5-5 + DFA 8-0-4 + DFA 8-5)						
<sup>1</sup> Heritability ± Standard errors						
<sup>2</sup> Least Significant Difference $p < 0.05$						

Table 3  
Genotypic (above diagonal) and phenotypic (below diagonal) correlation coefficient estimates for cell wall components.

	PCA	FA	DFA 5-5	DFA 8-0-4	DFA 8-5	DFAT
PCA		0.12 ± 0.09	-0.05 ± 0.41	-0.04 ± 0.33	-0.06 ± 0.14	-0.01 ± 0.13
FA	0.21 ± 0.04		0.84 ± 0.04	0.79 ± 0.05	0.78 ± 0.07	0.83 ± 0.04
DFA 5-5	0.13 ± 0.41	0.76 ± 0.02		0.97 ± 0.05	0.83 ± 0.06	0.97 ± 0.02
DFA 8-0-4	-0.01 ± 0.10	0.66 ± 0.02	0.76 ± 0.02		0.87 ± 0.08	0.98 ± 0.02
DFA 8-5	0.07 ± 0.14	0.66 ± 0.02	0.75 ± 0.02	0.63 ± 0.02		0.94 ± 0.03
DFAT	0.06 ± 0.04	0.76 ± 0.02	0.92 ± 0.01	0.88 ± 0.01	0.88 ± 0.01	

PCA: *p*-coumaric acid; FA: ferulic acid; DFA 5-5: 5-5 diferulic acid; DFA 8-0-4: 8-0-4 diferulic acid; DFA 8-5: sum (DFA 8-5 I + DFA 8-5b); DFAT: total diferulates sum (DFA 5-5 + DFA 8-0-4 + DFA 8-5).

Table 4

Multiple linear regression models (using stepwise selection) with saccharification efficiency, digestibility of organic matter and tunnel length on cell wall composition as dependent variables, and hydroxycinnamate-related traits in a maize MAGIC population.

Step Wise Selection			
Saccharification Efficiency (nmol mg <sup>-1</sup> material <sup>-1</sup> hour <sup>-1</sup> )			
Step	Variable introduced in the Model	R <sup>2</sup> Partial	Cumulated R <sup>2</sup>
1	<i>p</i> -Coumaric Acid (PCA)	0.0082	0.0082
<b>Model</b>	<b>SACCHARIFICATION</b> = 132.34 + 0.77 × PCA		
Digestibility of Organic Matter (%)			
1	<i>p</i> -Coumaric Acid (PCA)	0.1112	0.1112
2	DFA 8-5-I	0.0239	0.1351
<b>Model</b>	<b>DOM</b> = 62.06 - 0.65 × PCA + 14.51 × G DFA 8-5-I		
Tunnel length (cm)			
1	DFA 8-5-b	0.008	0.008
<b>Model</b>	<b>Tunnel Length</b> = 29.51 - 17.91 × DFA 8-5-b		

R<sup>2</sup> partial: percentage of the variance explained by each independent variable; R<sup>2</sup>: Total percentage of the variance explained by the model.

Table 5

SNPs and QTL significantly associated with cell wall component traits: including SNP's chromosome, bin and position within chromosome, allelic variants and additive effect for the SNP, proportion of total variance explained by the SNPs and P-value for the association between the SNP and the phenotype

Trait <sup>a</sup>	QTL <sup>b</sup>	Marker <sup>c</sup>	Chr <sup>d</sup>	Bin <sup>e</sup>	Alleles <sup>f</sup>	(No) <sup>g</sup>	Additive Effect <sup>h</sup>	P-value	R <sup>2i</sup>
PCA	qPCA_10_1	S10_90244975	10	10.04	T/A	292/54	0.42	1.81E-05	0.06
PCA	qPCA_10_2	S10_95123857	10	10.04	C/T	135/118	0.37	1.37E-05	0.07
PCA	qPCA_10_2	S10_94591480	10	10.04	A/G	109/148	0.34	1.89E-05	0.07
PCA	qPCA_10_2	S10_94591493	10	10.04	C/G	109/148	0.34	1.89E-05	0.07
PCA	qPCA_10_3	S10_98363769	10	10.04	G/C	258/61	0.39	2.10E-05	0.06
PCA	qPCA_10_3	S10_98799601	10	10.04	C/T	252/74	0.39	7.72E-06	0.07
PCA	qPCA_10_3	S10_99111128	10	10.04	G/A	277/42	0.48	1.86E-05	0.06
PCA	qPCA_10_3	S10_99204380	10	10.04	A/G	316/45	0.45	2.32E-05	0.05
PCA	qPCA_10_4	S10_100028254	10	10.04	G/A	295/50	0.45	1.63E-05	0.06
PCA	qPCA_3_1	S3_216794873	3	3.09	G/T	271/72	0.40	2.72E-06	0.06
FA	qFA_5_1	S5_198777365	5	5.06	A/C	240/49	0.21	6.39E-06	0.08
FA	qFA_5_1	S5_198777364	5	5.06	C/G	230/250	0.20	1.24E-05	0.07
FA	qFA_7_1	S7_9255010	7	7.01	C/T	146/85	0.16	2.16E-05	0.08
DFA 5-5	qDFA5-5_2_1	S2_212641077	2	2.08	G/T	65/158	0.02	1.00E-04	0.08
DFA 5-5	qDFA5-5_5_1	S5_4588557	5	5.01	C/T	17/273	0.03	1.01E-05	0.07
DFA 5-5	qDFA5-5_7_1	S7_104880233	7	7.02	A/C	134/117	0.02	5.29E-05	0.07
DFA 8-0-4	qDFA8-0-4_1_1	S1_214725368	1	1.07	C/G	71/257	0.02	1.19E-05	0.06
DFA 8-0-4	qDFA8-0-4_7_1	S7_162045550	7	7.04	A/G	113/152	0.02	2.32E-05	0.07
DFA 8 - 5	qDFA8-5_5_1	S5_198780500	5	5.06	A/G	112/132	0.02	6.76E-06	0.08
DFA 8 - 5	qDFA8-5_5_1	S5_198780501	5	5.06	T/A	113/132	0.02	1.03E-05	0.08
DFAT	qDFAT_1_1	S1_71604087	1	1.04	A/G	102/163	0.05	4.58E-05	0.07
DFAT	qDFAT_1_1	S1_71606983	1	1.04	G/A	101/173	0.05	7.58E-05	0.06
DFAT	qDFAT_5_1	S5_4588557	5	5.01	C/T	17/272	0.09	2.74E-05	0.06
DFAT	qDFAT_7_1	S7_104880233	7	7.02	A/C	134/117	0.05	2.68E-05	0.08

Trait <sup>a</sup>	QTL <sup>b</sup>	Marker <sup>c</sup>	Chr <sup>d</sup>	Bin <sup>e</sup>	Alleles <sup>f</sup>	(No) <sup>g</sup>	Additive Effect <sup>h</sup>	P-value	R <sup>2i</sup>
<p>a: PCA: <i>p</i>-coumaric acid; FA: ferulic acid; DFA 5-5: 5-5 diferulic acid; DFA 8-0-4: 8-0-4diferulic acid; DFA 8-5: sum (DFA 8-5 I + DFA 8-5b); DFAT: total diferulates sum (DFA 5-5 + DFA 8-0-4 + DFA 8-5).</p> <p>b: The number before the underscores indicates the chromosome and the number after the underscores indicates the QTL within the chromosome.</p> <p>c: The number before the underscores indicates the chromosome number and the number after the underscore indicates the physical position in bp within the chromosome.</p> <p>d: Chromosome.</p> <p>e: A bin is the interval that includes all loci from the leftmost or top Core Marker to the next Core Marker. The genetic maps are divided into 100 segments of approximately 20 centiMorgans designated with the chromosome number followed by a two-digit decimal.</p> <p>f: The letter before the diagonal is the nucleotide with the larger value; and the letter after the diagonal is the nucleotide with the smaller value.</p> <p>g: No = Number of inbred lines homozygous for a determined allelic variant. the number before the diagonal represents the number of homozygous with the largest mean value; and the number after the diagonal the number of homozygous with the smaller mean value.</p> <p>h: Additive effect: the additive effect was calculated as half the difference between the mean of the homozygous for the allele with the largest value and the mean of the homozygous for the allele with the smallest value.</p> <p>i: Phenotypic variance explained by each marker.</p>									

Table 6: Candidates genes associated with cell wall component traits identified in a maize MAGIC population.							
Trait <sup>a</sup>	QTL <sup>b</sup>	SNP <sup>c</sup>	Position <sup>d</sup>	Chr <sup>e</sup>	Bin <sup>f</sup>	Gene <sup>g</sup>	Gene Function <sup>h</sup>
PCA	qPCA_3_1	S3_216794873	216794873	3	3,09	Zm00001d044117	MYB-related-transcription factor 41
PCA	qPCA_10_1	S10_90244975	90244975	10	10,04	Zm00001d024864	$\omega$ -hydroxypalmitate O-feruloyl transferase
PCA	qPCA_10_3	S10_100028254	100028254	10	10,04	Zm00001d025037	UDP-glycosyltransferase 79
FA	qFA_5_1	S5_198777364 S5_198777365 S5_198780500 S5_198780501	198777364	5	5,06	Zm00001d017696	Peroxidase
DFA 8-5	qDFA8-5_5_1		198777365			Zm00001d017699	xyloglucan endotransglucosylase/hydrolase protein
			198780500				
			198780501				
DFA 5-5	qDFA5-5_2_1	S2_212641077	212641077	2	2,08	Zm00001d006933	Peroxidase
						Zm00001d006936	Peroxidase
						Zm00001d006937	Peroxidase
DFA 5-5	qDFA5-5_7_1	S7_104880233	104880233	7	7,02	Zm00001d020337	UDP-glycosyltransferase 90A1
DFAT	qDFAT_7_1						
DFA 8-O-4	qDFA8-O-4_1_1	S1_214725368	214725368	1	1,07	Zm00001d032222	UDP-glycosyltransferase 85A
						Zm00001d032223	Gibberellin 20 oxidase 2
DFA 8-O-4	qDFA8-O-4_7_1	S7_162045550	162045550	7	7.04	Zm00001d021965	Peroxidase 52
						Zm00001d021974	Arabinosyltransferase RRA3

a: PCA: *p*-coumaric acid; FA: ferulic acid; DFA 5-5: 5-5 diferulic acid; DFA 8-O-4: 8-O-4diferulic acid; DFA 8-5: sum (DFA 8-5 I + DFA 8-5b); DFAT: total diferulates sum (DFA 5-5 + DFA 8-O-4 + DFA 8-5).

b: The number before the underscores indicates the chromosome and the number after the underscores indicates the QTL within the chromosome.

c: The number before the underscores indicates the chromosome and the number after the underscores indicates the marker position.

d: Physical position of the marker in the B73 Reference Genome version 2.

e: Chr: Chromosome.

f: A bin is the interval that includes all loci from the leftmost or top Core Marker to the next Core Marker. The genetic maps are divided into 100 segments of approximately 20 centiMorgans designated with the chromosome number followed by a two-digit decimal.

g: Name of the gene in B73 Reference Genome version 4.

h: Gene function according to Zm-B73 reference form Gramene.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SuppelmentaryFigure1.docx](#)
- [SuppelmentaryFigure1.docx](#)
- [SuppelmentaryTable1.docx](#)
- [SuppelmentaryTable1.docx](#)