This is a post-peer-review, pre-copyedit version of an article published in Plant Journal. The final authenticated version is available online at: https://doi.org/10.1111/tpj.15108

Rewired phenolic metabolism and improved saccharification efficiency of a *Zea mays cinnamyl alcohol dehydrogenase 2* (*zmcad2*) mutant

Xinyu Liu[1,2], Rebecca Van Acker[1,2], Wannes Voorend[1,2], Andreas Pallidis[1,2], Geert Goeminne[1,3], Jacob Pollier[1,2,3], Kris Morreel[1,2], Hoon Kim[4,5], Hilde Muylle[1,2], Mickael Bosio[6], John Ralph[4,5], Ruben Vanholme[1,2,†] and Wout Boerjan[1,2,†,*]

[1] Ghent University, Department of Plant Biotechnology and Bioinformatics, Ghent, Belgium
[2] VIB Center for Plant Systems Biology, Ghent, Belgium
[3] VIB Metabolomics Core, Ghent, Belgium
[4] Department of Biochemistry, University of Wisconsin-Madison, Madison, Wisconsin 53706, USA
[5] Department of Energy's Great Lakes Bioenergy Research Center, Wisconsin Energy Institute, Madison, Wisconsin 53726, USA
[6] Biogemma, Site de la Garenne, Chappes, France
[†]These authors contributed equally to the article.
[*]Correspondence (Tel. +32 9 331 38 81; fax +32 9 331 38 09; email wout.boerjan@psb.vib-ugent.be)

ORCID IDs: 0000-0002-0201-4137 (X.L.); 0000-0002-0092-1155 (R.V.A.); 0000-0001-9254-1557 (W.V.); 0000-0001-7680-912X (A.P.); 0000-0002-0337-2999 (G.G.); 0000-0002-1134-9238 (J.P.); 0000-0002-3121-9705 (K.M.); 0000-0001-7425-7464 (H.K.); 0000-0001-7350-4179 (H.M.), 0000-0003-3245-1984 (M.B.); 0000-0002-6093-4521 (J.R.); 0000-0001-5848-3138 (R.V.); 0000-0003-1495-510X (W.B.)

Corresponding author:

Wout Boerjan
Center for Plant Systems Biology
VIB - Ghent University
Technologiepark 71
B-9052 Gent (Belgium)

Running title: Phenolic metabolism and saccharification of *zmcad2*

Keywords: Lignin, CAD, brown midrib, *bm1*, maize, field-grown, saccharification, pretreatment, metabolomics, phenolic profiling

SUMMARY

Lignocellulosic biomass is an abundant byproduct from cereal crops that can potentially be valorized as a feedstock to produce biomaterials. *Zea mays CINNAMYL ALCOHOL DEHYDROGENASE* 2 (*ZmCAD2*) is involved in lignification and is a promising target to improve the cellulose-to-glucose conversion of maize stover. Here, we analyzed a field-grown *zmcad2 Mutator* transposon insertional mutant. Z*mcad2* mutant plants had an 18% lower Klason lignin content whereas their cellulose content was similar to that of control lines. The lignin in *zmcad2* mutants contained increased levels of hydroxycinnamaldehydes, i.e., the substrates of ZmCAD2, ferulic acid, and tricin. Ferulates decorating hemicelluloses were not altered. Phenolic profiling further revealed that hydroxycinnamaldehydes are partly converted into (dihydro)ferulic acid and sinapic acid and their derivatives in *zmcad2* mutants. Syringyl lactic acid hexoside, a metabolic sink in CAD-deficient dicot trees, appeared not to be a sink in *zmcad2* maize. The enzymatic cellulose-to-glucose conversion efficiency was determined after ten different thermochemical pretreatments. *Zmcad2* yielded significantly higher conversions compared to controls for almost every pretreatment. However, the relative increase in glucose yields after alkaline pretreatment was not higher than the relative increase when no pretreatment was applied, suggesting that the positive effect of the incorporation of hydroxycinnamaldehydes was leveled off by the negative effect of reduced *p*-coumarate levels in the cell wall. Taken together, our results reveal how phenolic metabolism is affected in *CAD*-deficient maize and further support mutating *CAD* genes in cereal crops as a promising strategy to improve lignocellulosic biomass for sugar-platform biorefineries.

INTRODUCTION

The consumption of finite fossil resources and its consequences on global climate change highlight the need for alternative renewable feedstocks, such as plant biomass (van der Weijde *et al.*, 2013; Vanholme *et al.*, 2013a). Cereal grasses (e.g., maize, wheat, rice, and sorghum) are particularly interesting as they can be cultivated as dual-purpose crops, on the one hand for the production of cereals as food, and on the other hand for the production of straw or stover as feedstock for cellulosic glucose production (Vermerris *et al.*, 2007). A major challenge in developing a cost-effective cellulosic glucose production platform is the improvement of the saccharification efficiency of lignocellulosic biomass (Carroll and Somerville, 2009; Torres *et al.*, 2016). Lignin, an aromatic polymer of plant cell walls, is a prime component responsible for biomass recalcitrance to enzymatic hydrolysis. In grasses, lignin can represent 20% or more of the biomass (Halpin, 2019). Because lignin covers the cellulose microfibrils, it hinders the saccharification enzymes (cellulases) from accessing the cellulose surface. In addition, saccharification efficiency is negatively influenced by the unspecific absorption of cellulases onto the lignin polymer (Mansfield et al., 1999; Jørgensen *et al.*, 2007). Consequently, the saccharification efficiency can be increased by lowering the lignin amount or by changing its structure (Vanholme *et al.*, 2012a; Vanholme *et al.*, 2008; Zeng *et al.*, 2014).

Maize is a staple crop in many regions of the world. In addition, maize is used as a model system to study biological processes (Hake and Ross-Ibarra, 2015). Lignin biosynthesis is relatively well-studied in maize, and involves the biosynthesis of three main canonical monolignols, *p*-coumaryl, coniferyl, and sinapyl alcohol. Portions of these alcohols are further esterified to either acetate or *p*-coumarate, resulting in mainly coniferyl acetate and sinapyl *p*-coumarate (Hatfield *et al.*, 2009; Lan *et al.*, 2015; Lu and Ralph, 1999; Marita *et al.*, 2014). In addition, the flavone tricin is a natural lignin monomer in maize (Eloy *et al.*, 2017; Fornalé *et*

al., 2017; Lan *et al.*, 2015; Lan *et al.*, 2016). After the biosynthesis of these monomers in the cytoplasm and their translocation to the apoplast, putatively through diffusion (Vermaas *et al.*, 2019), the monomers are oxidized by peroxidases and/or laccases, resulting in monolignol radicals that combinatorially couple to each other and to the growing lignin polymer. Hereby, *p*-coumaryl alcohol gives rise to *p*-hydroxyphenyl (H) units, coniferyl alcohol and its esters to guaiacyl (G) units, and sinapyl alcohol and its esters to syringyl (S) units (Boerjan *et al.*, 2003; Freudenberg, 1959; Vanholme *et al.*, 2013b; Vanholme *et al.*, 2010a). The incorporation of tricin results in tricin (T) units (del Río *et al.*, 2012; Eloy *et al.*, 2017; Fornalé *et al.*, 2017; Lan *et al.*, 2015).

The last step in the biosynthesis of *p*-coumaryl, coniferyl, and sinapyl alcohols is catalyzed by CINNAMYL ALCOHOL DEHYDROGENASE (CAD), which reduces hydroxycinnamaldehydes into their corresponding alcohols (Mansell *et al.*, 1974; Morrison *et al.*, 1994). Considering the importance of CAD for lignin formation in plant cell walls, mutants and transgenic plants with reduced CAD activity have been thoroughly investigated in pine (Ralph *et al.*, 1997; Wu *et al.*, 1999) and many dicot species such as Arabidopsis (Sibout *et al.*, 2005), tobacco (Bernard Vailhé *et al.*, 1998; Chabannes *et al.*, 2001; Halpin *et al.*, 1994), alfalfa (Baucher *et al.*, 1999), poplar (Baucher *et al.*, 1996; Ralph *et al.*, 2001; Van Acker *et al.*, 2017), eucalyptus (Valério *et al.*, 2003), and mulberry (Yamamoto *et al.*, 2020), and monocot species such as maize (Fornalé *et al.*, 2012; Halpin *et al.*, 1998), sorghum (Porter *et al.*, 1978; Pillonel *et al.*, 1991; Saballos *et al.*, 2008; 2009; Sattler *et al.*, 2009; Scully *et al.*, 2016), switchgrass (Fu *et al.*, 2011; Saathoff *et al.*, 2011), rice (Koshiba *et al.*, 2013; Martin *et al.*, 2019; Zhang *et al.*, 2006), tall fescue (Chen *et al.*, 2003), and Brachypodium (Bouvier d'Yvoire *et al.*, 2013; Trabucco *et al.*, 2013).

In maize, *ZmCAD2* is one of the *CAD* family members involved in lignification. Mutants in *ZmCAD2* were first discovered by their brown leaf midribs, after which the mutants were

named *brown midrib 1* (*bm1*) (Eyster, 1926; Kuc and Nelson, 1964; Chen *et al.*, 2012). Intriguingly, the incorporation of hydroxycinnamaldehydes into the lignin polymer, a hallmark for CAD deficiency in many plants, has either not been unambiguously proven to be increased in *ZmCAD2*-deficient maize (Halpin *et al.*, 1998; Provan *et al.*, 1997), or appeared at much lower levels in *ZmCAD2*-deficient maize than in *CAD*-deficient plants of other species (Barrière *et al.*, 2013; Marita *et al.*, 2003; Provan *et al.*, 1997). We therefore questioned whether the hydroxycinnamaldehydes, i.e. the CAD substrates, have another metabolic fate in *ZmCAD2*-deficient maize.

In general, the perturbation of CAD results in moderate reductions in lignin amount and causes the incorporation of hydroxycinnamaldehydes into the lignin, leading to a lignin polymer with an altered composition and structure (Ralph *et al.*, 2001) typified by an increased frequency of free phenolic units (Lapierre *et al.*, 1999). Depending on the plant species and pretreatment used, a reduced CAD activity results in an improved cell wall saccharification yield (Anderson *et al.*, 2015; Fornalé *et al.*, 2012; Fu *et al.*, 2011; Van Acker *et al.*, 2017). RNAi-mediated, *CAD*-downregulated maize had an 8% higher cellulosic ethanol production as compared to controls (Fornalé *et al.*, 2012) and preliminary enzymatic saccharification experiments showed a 40% increase in glucose yield of *bm1* (in the inbred A619 background) as compared to its control (Vermerris et al., 2007). However, the exact cause for this increased conversion efficiency is not known. Lignin amount is known to negatively correlate with enzymatic cellulose hydrolysis (Mechin *et al.*, 2000; Chen and Dixon, 2007; Van Acker *et al.*, 2013; van der Weijde *et al.*, 2016; Ostos Garrido *et al.*, 2018). In addition, structural elements in the cell wall that interfere with alkaline pretreatment could further explain the observed increase in cellulosic hydrolysis (Fornalé *et al.*, 2012). For instance, the incorporation of hydroxycinnamaldehydes introduces conjugated carbonyl functions into the lignin, which facilitates the cleavage of 8–*O*–4 ether structures under alkaline conditions (Van Acker *et al.*,

2017). An alkaline pretreatment also hydrolyzes ferulates esterified to arabinoxylans of which a portion is cross-linked to lignins, thereby loosening the cell wall (Jung, 2003; MacAdam and Grabber, 2002). In addition, the *p*-coumarates decorating the maize lignin that make up to 20% (by weight) of the lignin are hydrolyzed in alkaline pretreatments (Hatfield *et al.*, 2008; Lapierre, 2010; Li *et al.*, 2015; Ralph, 2010; Ralph *et al.*, 1994). Also a high frequency of free phenolic ends in the lignin has been shown to improve saccharification efficiency after alkaline pretreatment (Elumalai *et al.*, 2012; Eudes *et al.*, 2014; Grabber *et al.*, 2008; Grabber *et al.*, 2010; Lapierre, 2010; Zhang *et al.*, 2012). To obtain insights into the parameters determining improved cellulose conversion, we tested the saccharification ability of *zmcad2* lignocellulosic biomass after a series of neutral, acid and alkaline pretreatments.

In our study, all analyses were performed on the field-grown Ev2210bm1 Z*mCAD2*-deficient *Mutator* transposon insertional mutant *Zmcad2-m2210::Mu*, hereafter called *zmcad2*, and its control (Barrière *et al.*, 2013). *Zmcad2* mutants have a transposon insertion in the *ZmCAD2*-coding sequence, resulting in an inactive, truncated ZmCAD2 protein (Barrière *et al.*, 2013). *Zmcad2* lines displayed the hallmarks of *bm1* mutants, including a brown midrib phenotype (Barrière *et al.*, 2013).

RESULTS

Altered lignin composition in *zmcad2* mutants

Field-grown *zmcad2* mutants and controls were harvested at silage stage. After removing the cobs, the remaining lignocellulosic biomass (stems and leaves) was analyzed for lignin and cellulose content. The cell wall residue (CWR) fraction, obtained after extraction of the solubles from the biomass, was slightly reduced for *zmcad2* compared to control samples (Table 1). The total amount of cellulose (α-cellulose) in the CWR and the fraction of crystalline cellulose in *zmcad2* were similar to those in controls (Table 1). The *zmcad2* mutant had an 18% lower

Klason lignin content (Table 1), which was in the same range as the previously reported decreased portion of lignin in the internodes of the *bm1* mutant (-20%) (Halpin *et al.*, 1998). On the other hand, *zmcad2* was previously reported to have only a modest Klason lignin reduction (-6%) in whole plants (Barrière *et al.*, 2013). This discrepancy is most likely due to the fact that the plants were grown in independent experiments, in different growth years.

The lignin composition was analyzed via thioacidolysis, which cleaves 8–O–4 ether bonds and some ester bonds, thereby releasing phenolic moieties that are not linked via carbon-carbon inter-unit linkages (Figure 1). The released moieties were derivatized via trimethylsilylation prior to GC-MS analysis. The amounts of H, G, and S units released via thioacidolysis and expressed per unit of lignin, were all decreased to a similar extent in *zmcad2* mutants as compared to control plants (-36%, -35%, and -37%, respectively; Table 1). Different from the results reported by Barrière *et al.* (2013), where the S/G ratio increased from about 1.13 in control to 1.26 in *zmcad2*, the S/G ratio was found to be about 1.2 for both genotypes in this independent growth. Also in *bm1* mutants, the thioacidolysis-based S/G ratio was not different from its control (Halpin *et al.*, 1998).

Thioacidolysis also released markers for the incorporation of coniferaldehyde, sinapaldehyde, vanillin, and syringaldehyde. For coniferaldehyde, G dithioketal and G indene structures were detected (Figure 1). G dithioketal structures are derived from 4–O coupled coniferaldehyde units, i.e., in which coniferaldehyde acted as starting units for polymerization. These structures were increased by 151% in *zmcad2* mutants as compared to the control (Table 1). G indene structures are markers for 8–O-coupled coniferaldehyde, i.e., in which coniferaldehyde is linked via its 8-position to the 4–O-position of a growing lignin polymer (Kim *et al.*, 2002). In line with the results of Barrière *et al.* (2013), G indenes were below the detection limit in control samples, but were readily detected in the *zmcad2* samples. The S dithioketal structures, which could theoretically be formed from 4–O-coupled sinapaldehyde

units, remained below the detection limit in both *zmcad2* mutants and control samples. However, S indenes, markers for the incorporation of 8–O-coupled sinapaldehyde, increased substantially (about 20-fold) in *zmcad2*. Moreover, vanillin dithioketal and syringaldehyde dithioketal, markers for incorporation of 4–O-coupled vanillin and syringaldehyde, were increased by 212% and 233%, respectively (Figure 1, Table 1).

Finally, thioacidolysis also released markers for the incorporation of *p*-coumaric acid and ferulic acid, and their respective esters. *p*-Coumaric acid is mainly present in the maize cell wall as *p*-coumarate esters decorating the lignin, and to a lesser extent also decorating arabinoxylan (Lapierre *et al.*, 2018). The abundance of *p*-coumarate in the cell wall could be determined via two products released via thioacidolysis (Figure 1). The first product is *p*-coumaric acid itself (*p*CA-I), which was significantly reduced by 27% in *zmcad2* mutants as compared to the control. The second product is the ethanethiol addition product of *p*-coumaric acid (*p*CA-II), which showed a trend to be lower in *zmcad2* compared to the control (*p*-value=0.095). In maize, ferulic acid is mainly found as esters decorating arabinoxylans and to a lower extent as conjugate esters in the lignin polymer, as part of coniferyl and sinapyl ferulate derived units (Ishii, 1991; Karlen *et al.*, 2016; Ralph *et al.*, 2019; Terrett and Dupree, 2018). In addition, low levels of ferulic acid may also act as a genuine lignin monomer (Ralph *et al.*, 2008). The amount of ferulate esters could be determined via two products released via thioacidolysis; as ferulic acid (FA-I) and as the ethanethiol addition product of ferulic acid (FA-II). The abundance of neither of these products was changed in *zmcad2* as compared to the amount in control samples. The incorporation of ferulic acid as a monomer in the lignin polymer could be determined via the release of 7,8,8-trithioethyl ethylguaiacol (i.e., A_G, also known as the CCR marker) that derives from *bis*-8–*O*–4 coupled ferulic acid (Leplé *et al.*, 2007; Mir Derikvand *et al.*, 2008; Ralph *et al.*, 2008). Its abundance increased about 5-fold in *zmcad2* (Table 1). We also searched targeted on thioacidolysis-derived products from sinapic acid, but these remained under the

detection limit. In summary, thioacidolysis showed that the mutation of *ZmCAD2* resulted in changes in lignin composition beyond the incorporation of hydroxycinnamaldehydes.

To further investigate changes in the chemical structure of lignin, 2D ^{1}H–^{13}C correlation heteronuclear single-quantum coherence (HSQC) NMR analysis was performed on the enzyme lignins (ELs) isolated from *zmcad2* and control samples (Figure 2, Table S1). The signals from H units co-appear with those of phenylalanine (Kim *et al.*, 2017). After correction for the presence of phenylalanine, H units were found to be present only at trace levels in both control and *zmcad2* samples. The relative amount of coniferaldehyde endgroups (X2G′) increased by 74% in *zmcad2* samples, whereas sinapaldehyde endgroups (X2S′) were below detection limit. The increase in X2G′ and absence of X2S′ is in line with the results obtained via thioacidolysis. Notably, the relative amount of S* units (Figure 2) in *zmcad2* samples increased by 121%. Because these alpha-keto functionalities are thought to be derived from S units during the ball-milling process (Holtman *et al.*, 2006), their increase could indicate that *zmcad2* lignin is more prone to oxidation than wild-type lignin. The S/G ratio (including aldehydes) slightly decreased from 0.97 in the control to 0.91 in *zmcad2*.

The aromatic region of the HSQC spectra also showed signals of *p*-coumarate and tricin. Their signal intensities were expressed relative to the sum of those of S, S*, G and X2G′. The relative abundance of *p*-coumarate increased slightly (~5%) in *zmcad2* compared to control samples. This indicates that a higher proportion of the units derived from coniferyl and sinapyl alcohol are decorated with *p*-coumarate. Nevertheless, as the total amount of lignin is reduced by about 18% in *zmcad2*, the absolute level of *p*-coumarate in the cell wall is reduced in these mutants. In line with the fact that most ferulates decorate arabinoxylans, ferulates were not detected in ELs in which the hemicellulose fraction is largely removed. The *bis*-8–O–4 coupled ferulic acid (De Meester *et al.*, 2018; Ralph *et al.*, 2008) was below the detection limit. Finally,

the relative abundance of tricin was increased by 40% in *zmcad2* as compared to control samples.

The oxygenated-aliphatic region provides structural details on the distribution of interunit linkage types present in the ELs (Figure 2B). The relative abundance of monolignol-derived β-aryl ether (β–O–4, A), phenylcoumaran (8–5, B) and resinol (8–8, C) structures were about 83%, 7% and 10%, respectively in both *zmcad2* and control samples. NMR indicated that cinnamyl alcohol end groups, including those part of *p*-coumarate and acetate esters (X1+X1′) were slightly reduced (~7%) in *zmcad2*. The reduction of these end-groups, which are lignin polymer initiation sites, is in line with the observations that the relative abundance of coniferaldehyde (X2G′) and tricin initiation sites increased.

The aldehyde region of the HSQC spectra shows the aldehyde functionalities in various inter-monomeric and end-unit linkage types in the lignin (Figure 2C). Coniferaldehyde and sinapaldehyde both coupled via their 8-position onto the 4–O position of S units to form G′(8–O–4)S and S′(8–O–4)S structures, respectively. In addition, sinapaldehyde coupled via its 8-position onto the 4–O position of G units to form S′(8–O–4)G structures. Signals from these three structures were clearly observed in *zmcad2* ELs, whereas they were not detected in control samples. It has been proven that coniferaldehyde does not 8–O–4-cross-couple with G units (Ralph *et al.*, 2019). In line with this, no G′(8–O–4)G structures could be detected. In addition, the 8–8-cross-coupled homodimeric or heterodimeric coniferaldehydes and sinapaldehydes (G′(8–8)G′, S′(8–8)G′, G′(8–8)S′ and S′(8–8)S′) were not detected in either *zmcad2* or control samples. The signals from vanillin (V) and syringaldehyde (SA) end-groups were difficult to untangle. We therefore integrated the combined signal that corresponds to V+SA. The relative contribution of V+SA decreased by 51% in *zmcad2* compared to the control.

Zmcad2 mutants accumulate feruloyl and sinapoyl hexose conjugates

The total amount of lignin was reduced in *zmcad2*, and only minor amounts of hydroxycinnamaldehydes were incorporated. If the aldehydes are not all incorporated into the lignin of *zmcad2*, into what are they metabolized? To answer this question, we investigated the impact of perturbation of *ZmCAD2* on the phenolic metabolism of the ear internode via LC-MS-based metabolic profiling. A total of 5948 peaks was detected in the samples. To identify those metabolites that were most affected by the *zmcad2* mutation, we performed an untargeted analysis. Therefore, we selected peaks that were highly abundant, highly differential (at least 10-fold), and for which the difference in abundance between *zmcad2* and control was highly significant. The stringent selection criteria resulted in a total of 12 and 51 peaks with a lower and higher intensity, respectively, in *zmcad2*. The 12 peaks with a lower intensity could be assigned to 11 compounds of which seven could be structurally annotated (Table 2, Figure 3 and Figure S1). Compounds **1** and **2** were both trilignols, small lignin precursors, and their reduced abundance in *zmcad2* is in line with the observed reduced lignin amount in these mutants. In addition, compounds **3**, **4,** and **5** were hexosylated coupling products of coniferyl alcohol and sinapyl alcohol. Because coniferyl and sinapyl alcohol are direct products of CAD, the reduced abundance of their coupling products **1**-**5** in *zmcad2* mutants is most likely a direct consequence of reduced CAD activity. In addition, also vanilloyl pentosyl hexose **6** and tricin decorated with hexuronic acid, pentose and syringic acid **7** were reduced in abundance.

The 51 peaks with increased intensity in *zmcad2* could be assigned to 43 compounds of which 29 could be (partially) structurally annotated (Table 2, Figure 3 and Figure S1). Five compounds with an increased abundance in *zmcad2* had a feruloyl hexose moiety (**8**-**11** and **33**) and 22 compounds had a sinapoyl hexose moiety (**12**-**33**). In general, compounds with a ferulic acid and sinapic acid moiety point to new metabolic sinks by re-routing the metabolic flux through the pathway. Moreover, a vanillic acid conjugate **34**, hydroxybenzoyl hexose **35** and a tricin derivative **36** were annotated among the compounds with increased abundance.

Oligolignols composed of coniferyl and sinapyl alcohols were reduced, whereas those with one sinapaldehyde unit were increased in abundance in *zmcad2* mutants

The untargeted metabolite analysis revealed that the abundances of trilignol **1** and **2** were strongly reduced in *zmcad2* samples. These trilignols are only two members of the large oligolignol metabolic class (Morreel *et al.*, 2010a; Morreel *et al.*, 2010b). To gain insight in how the *zmcad2* mutation influenced the abundance of other oligolignols, we performed a targeted metabolite analysis (Table 3, Figure 3). Based on *m/z*, retention time and MS/MS fragmentation, we characterized 17 oligolignols that were coupling products of (acylated) coniferyl alcohol and (*p*-coumaroylated) sinapyl alcohol (**37**-**49**). In addition, we found oligolignols in which tricin was linked to one or two (acetyl) coniferyl alcohol units (**52**-**59**). In line with the reduced flux towards coniferyl alcohol and sinapyl alcohol in *zmcad2* mutants, all these oligolignols were significantly reduced in abundance to levels between 77% and 1% of the level in the control samples. Also two coupling products of tricin and *p*-coumaryl alcohol (**50** and **51**) were found. Their abundance was not significantly different between *zmcad2* and control samples (*p*-value=0.41 and 0.70, respectively).

Next, we searched for oligolignols with hydroxycinnamaldehyde units. The MS/MS spectra of oligolignols in which coniferaldehyde is linked via its O–4 position into an 8–O–4 linkage, are typified by *m/z* 177 and *m/z* 162 fragments with a high intensity (Figure S2). Based on this characteristic, we could annotate three trilignols that contained coniferaldehyde (**60**-**62**). Similarly, oligolignols with an 8–O–4-linked sinapaldehyde unit are typified by *m/z* 192 and *m/z* 207 fragments. As such, four sinapaldehyde containing oligolignols were found (**63**-**66**). In addition, we could also detect S'(8-8)S' **67** based on *m/z* and retention time similarity with data from an earlier report (Van Acker *et al.*, 2017). No oligolignols in which *p*-coumaraldehyde linked via its O–4 position into an 8-O–4 linked unit were found, based on a search for MS/MS spectra with *m/z* 147 fragments. Although one of the three coniferaldehyde-containing trilignols

(**60**) was not significantly different in abundance (*p*-value=0.41), the two others (**61** and **62**) were reduced in abundance to 13% and 19% of the levels in the control (*p*-value<0.01). A decreased abundance of **61** and **62** can partly be explained because the total pool of oligolignols that contain coniferyl alcohol and sinapyl alcohol and their acetate and *p*-coumarate esters is reduced (as in **37**-**49** and **52**-**59**). Three oligolignols with a sinapaldehyde unit were increased 2.8- to 6.7-fold, as compared to control (*p*-value<0.01) (**63**-**65**). Oligolignol **66** also contained one sinapaldehyde unit and was increase 16-fold in *zmcad2*, but at a lower significance level (*p*-value=0.016). Moreover, the dilignol composed out of two sinapaldehyde units, S'(8-8)S' **67** was below the detection limit in control, but was detected in all *zmcad2* samples. Taken together, the oligolignol-targeted metabolomics showed that the abundances of oligolignols solely made up from coniferyl alcohol and sinapyl alcohol (and their acylated forms) were reduced, those that contained coniferaldehyde were either reduced or did not change, and those that contained sinapaldehyde were increased in *zmcad2* internodes.

Hexosylated dihydroferulic acid accumulated in *zmcad2*, whereas hexosylated syringyl lactic acid did not

Several metabolites with a ferulic acid and sinapic acid moiety (**8**-**33**) accumulated in the *zmcad2* mutants. Metabolites of these compound classes also accumulated in CAD-deficient tobacco (*Nicotiana tabacum*), Arabidopsis, poplar and mulberry (Dauwe *et al.*, 2007; Thévenin *et al.*, 2011; Van Acker *et al.*, 2017; Vanholme *et al.*, 2012b; Yamamoto *et al.*, 2020). In CAD-deficient poplar and mulberry, hexose-linked dihydroferulic acid and syringyl lactic acid were also found to accumulate (Van Acker *et al.*, 2017; Yamamoto *et al.*, 2020). To investigate if these metabolites also accumulated in *zmcad2*, we performed a targeted metabolite search. Hexosylated dihydroferulic acid increased over 5-fold in *zmcad2* as compared to control (Table 3). For syringyl lactic acid hexoside, a low signal was found at 3.14 min with *m/z* 403.1238,

which differed 2.00 ppm from the calculated *m/z* 403.1246. There was no MS/MS spectrum available, so its characterization remains highly uncertain. The intensity of the signal was increased by 42% (*p*-value=0.02). For comparison, in CAD-deficient poplar and mulberry, syringyl lactic acid hexoside was increased 8,650 and 27-fold, respectively (Van Acker *et al.*, 2017; Yamamoto *et al.*, 2020).

CAD downregulation improves saccharification efficiency

To estimate the effect of the *zmcad2* mutation on the enzymatic cellulose-to-glucose conversion efficiency, we subjected *zmcad2* biomass to saccharification assays. Furthermore, to test the influence of the altered lignin structure in *zmcad2* mutants on cell wall reactivity and solubility, we subjected the biomass to different thermochemical pretreatments. The biomass was subjected to either hot water (98°C), sulfuric acid (H_2SO_4; 0.4M), hydrochloric acid (HCl; 1M), ammonium hydroxide (NH_4OH; 1M) or sodium hydroxide (NaOH; 6.25, 12.5, 25, 50, 62.5, or 150mM) pretreatment. For both lines, the residual biomass after pretreatment was highest for the samples that were not pretreated, and lowest for samples pretreated with acids (Figure 4). When applying pretreatment with NaOH, an inverse trend of relationship between residual biomass recovery and pretreatment severity was observed for both lines. However, there was no significant difference between *zmcad2* and control. By applying an acid pretreatment, 10% more biomass was removed from *zmcad2* relative to the control. These data indicate that the intrinsic structure of *zmcad2* biomass is modified in such a way that it becomes easier to degrade in acidic, but not in the alkaline conditions that were used.

After pretreatment, the residual biomass was incubated with an enzyme mixture. By taking into account the initial amount of cellulose in the CWR, the cellulose-to-glucose conversion was calculated (Figure 4 and Table S2). Biomass from *zmcad2* yielded significantly higher cellulose conversions compared to control for almost every pretreatment condition. The

highest relative difference between control and *zmcad2* was observed for 6.25mM NaOH (+24%), but even without any pretreatment, *zmcad2* had a 20% higher cellulose-to-glucose conversion compared to control. Based on the concentration range of NaOH, increasing the severity of the pretreatment resulted in higher cellulose-to-glucose conversions for both control and *zmcad2*. However, the difference between *zmcad2* and control lines in cellulose-to-glucose conversion decreased with increasing pretreatment severity and was even not significant anymore for the two highest concentrations of NaOH used (62.5 and 150mM). Because *zmcad2* mutants had no significant difference in cellulose content (Table 1) and only a 4% decrease in CWR as compared to control plants, the trends in glucose yield are largely similar when expressed per gram dry biomass instead of per gram CWR (Table S2).

DISCUSSION

The effects of mutating *ZmCAD2* on lignin composition

Klason lignin amount was reduced by 18% in *zmcad2* mutants as compared to control plants. Based on thioacidolysis, the reduction in total lignin in *zmcad2* mutants is caused by an equal reduction of the traditional H, G, and S units (including their acetate and *p*-coumarate esters). The traditional units are not reduced to zero in the lignin of *zmcad2* mutants, thus additional enzymes with CAD activity must be involved in lignification. In the maize *zmcad2* mutant, ZmCAD1 and ZmCAD4 are potentially involved in the biosynthesis of the traditional lignin units, because their corresponding genes were shown to be significantly up-regulated in the internodes of this mutant (Barrière et al., 2013).

In line with previous results, we found that *zmcad2* mutants had a lower yield of thioacidolysis-released H, G, and S units, expressed per lignin (Halpin *et al.*, 1998; Van Acker *et al.*, 2017). Such a lower thioacidolysis yield is typically interpreted to be caused by a relative increase of condensed linkages (8–8 and 8–5) and a relative decrease of ether linkages (8–O–4) in the lignin

(Halpin *et al.*, 1998). Here, this explanation is refuted by the NMR results, which showed no shifts in the relative abundance of the condensed linkages and ether linkages between *zmcad2* mutants and the control. Instead, the lower thioacidolysis yield could be caused by the observed increased incorporation of the alternative monomers coniferaldehyde, sinapaldehyde, tricin and ferulic acid in *zmcad2* lignin. Neither of these are part of the H+G+S-thioacidolysis yield. Alternatively, the shift in lignin structure changed its reactivity upon thioacidolysis, resulting in a lower yield. In line with this, we also found an increase in alpha-keto S units (S*) in *zmcad2* mutants. Even though these units are considered as an artifact generated during ball milling, their relatively higher proportion in lignin of *zmcad2* mutants, hints at a different reactivity of this lignin type as compared to the control.

Thioacidolysis and NMR both showed that the fraction of 8–O–4 and 4–O-coupled coniferaldehyde and 8–O–4-coupled sinapaldehyde were higher in the lignin of *zmcad2* mutants as compared to the control. 4–O-coupled sinapaldehyde was not detected by NMR in the lignin of *zmcad2* mutants and also the S dithioketal compounds derived from these structures upon thioacidolysis remained under the detection limit. Also, 8–8-coupled coniferaldehydes and sinapaldehydes remained under the detection limit via NMR. On the contrary, oligolignols with 4–O- and 8–8-coupled sinapaldehyde (**63**-**67**) significantly increased in abundance in *zmcad2* plants, whereas oligolignols with 4–O-coupled coniferaldehyde did not (**60**-**62**; Table 3). Because oligolignols are a pool of metabolites with the potential to become larger lignin structures, shifts in lignin are often reflected in the oligolignol pool (Vanholme *et al.*, 2012b). The apparent inconsistency between lignin units observed via thioacidolysis, NMR, and in the oligolignols, can be explained by the fact that thioacidolysis, NMR analysis, and phenolic profiling are complementary techniques that only provide partial insight into lignin composition. Based on the combined results of thioacidolysis, NMR analysis and phenolic profiling, we can conclude that the lignin of *zmcad2* is characterized by an increased

incorporation of coniferaldehyde, sinapaldehyde, ferulic acid, and tricin monomers. At the same time, the incorporation of the total pool of sinapyl alcohol, sinapyl *p*-coumarate, coniferyl alcohol, and potentially also coniferyl acetate is decreased in *zmcad2* lignin.

Phenolic profiling reveals metabolic sinks in *zmcad2* mutants

Pathway perturbations designed for lignin engineering typically also result in severe shifts in the soluble phenolic metabolism (Coleman *et al.*, 2008; Morreel *et al.*, 2014; Vanholme *et al.*, 2013b; Vanholme *et al.*, 2012a). Instead of accumulating the direct substrates of ZmCAD2, *zmcad2* internodes mainly accumulated feruloyl hexose and sinapoyl hexose and further conjugates (**8**-**33**, Table 2). Hexosylation of accumulating phenylpropanoids has been suggested as a strategy of the plant to detoxify potentially harmful metabolites (Vanholme *et al.*, 2019; Vanholme *et al.*, 2012a; Vanholme *et al.*, 2010b). More specifically, ferulic acid and sinapic acid conjugates accumulated in Arabidopsis *cad6* single mutants and *cad-c cad-d* (*cad2 cad6*) double mutants, in the naturally occurring mulberry *cad1* mutant (Sekizaisou), in poplar downregulated for *CAD1* and in tobacco downregulated for *CAD2* (Dauwe *et al.*, 2007; Thévenin *et al.*, 2011; Van Acker *et al.*, 2017; Vanholme *et al.*, 2012b; Yamamoto *et al.*, 2020). In Arabidopsis and *Brassica napus*, coniferaldehyde and sinapaldehyde are metabolized into ferulic and sinapic acid, respectively, by the enzymatic activity of HYDROXYCINNAMALDEHYDE DEHYDROGENASE (HCALDH, also known as REDUCED EPIDERMAL FLUORESCENCE 1, REF1) (Nair *et al.*, 2004; Mittasch *et al.*, 2013; Emrani *et al.*, 2015). The accumulation of feruloyl hexose and sinapoyl hexose and their conjugates in *zmcad2* mutants, suggests the presence of a similar bio-conversion step in maize. In line with this hypothesis, *HCALDH* homologs are present in grasses and the enzymes encoded by the four closest *HCALDH* homologs in maize (*RF2C*, *RF2D*, *RF2E* and *RF2F*) catalyze *in vitro* the conversion of coniferaldehyde and sinapaldehyde into ferulic acid and

sinapic acid, respectively (Končitíková *et al.*, 2015). In addition, *RF2C* and *RF2D* are relatively highly expressed in wild-type maize stem (Končitíková *et al.*, 2015). This observation further supports the existence of the HCALDH-conversion in maize internodes. Besides the possible route via HCALDH, a second route to ferulic acid might occur in *zmcad2* mutants. In poplar, CAD1 and CCR2 physically interact, and the suppression of CAD1 causes a reduction in CCR activity in cells that biosynthesize monolignols (Yan *et al.*, 2019). The reduced CCR activity may in turn lead to the accumulation of feruloyl-CoA, and further to the de-esterification of feruloyl-CoA into ferulic acid and CoA (Leplé *et al.*, 2007; Van Acker *et al.*, 2014). Further research is needed to investigate the role of the proposed HCALDH candidates, whether ZmCAD2 interacts with CCR enzymes in maize and whether a ZmCAD2 deficiency would, similar to the situation in poplar, result in reduced CCR activity in maize.

In addition, hexosylated dihydroferulic acid (**68**) was also found to accumulate in *zmcad2* maize, and was previously also found to accumulate in mulberry *cad1* mutant and *CAD1*-downregulated poplar (Van Acker *et al.*, 2017; Yamamoto *et al.*, 2020). This compound is the glycosylated form of dihydroferulic acid, which might be made via a DOUBLE BOND REDUCTASE (DBR) activity on either feruloyl-CoA or ferulic acid as substrate (Ibdah *et al.*, 2014). Apparently, a metabolic route towards dihydroferulic acid exists both in dicots (mulberry and poplar) and monocots (maize). However, unlike the mulberry *cad1* mutant and *CAD1*-downregulated poplar, syringyl lactic acid hexoside did not accumulate in *zmcad2* maize. The enzymes involved in the biosynthesis of syringyl lactic acid are not known, but potentially a DBR and a hydroxylase are involved (Van Acker *et al.*, 2017; Yamamoto *et al.*, 2020). The combined findings hint that the pathway towards syringyl lactic acid is either not present or not activated in *zmcad2* maize and, thus, showing that the detoxification route differs between dicots and monocots.

The metabolic flux of *zmcad2* mutants shifted into ferulic acid and sinapic acid and their derivatives. Thioacidolysis showed that part of the ferulic-acid pool was incorporated in the lignin of *zmcad2* mutants, which is consistent with the previous observations with *zmcad2* mutants and the Arabidopsis *cad-c cad-d* double mutant (Barrière *et al.*, 2013; Thévenin *et al.*, 2011). Notably, the fraction of ferulates decorating the hemicellulose did not increase in abundance (Table 1). This is similar to previous findings in which either no differences or even a small decrease in cell-wall bound ferulate was detected in CAD-deficient maize (Barrière *et al.*, 2004; Marita *et al.*, 2003; Provan *et al.*, 1997). This observation also suggests that a rate limiting step controls the biosynthetic route from the overproduced ferulic acid to hemicellulose ferulate (Terrett and Dupree, 2018).

Effect of pretreatments on saccharification efficiency of *zmcad2* lignocellulosic biomass

Lignin amount and lignin composition are two factors that influence cell-wall recalcitrance towards saccharification (Chen and Dixon, 2007; Van Acker *et al.*, 2013). The cellulose conversion without any pretreatment was increased by 20% in *zmcad2*, as compared to control samples. Because there was no significant difference in cellulose content in the *zmcad2* lines compared with the control, the improved saccharification efficiency is most likely caused by the lower lignin amount (Chen and Dixon, 2007; Van Acker *et al.*, 2017). In addition, we can not exclude the possibility that the altered lignin structure in the *zmcad2* line affects the adsorption properties of cellulase onto the lignin polymer, which would further influence the saccharification efficiency in the mutant line (Mansfield et al., 1999; Jørgensen *et al.*, 2007). The hot water pretreatment (98°C) only had a minor effect on saccharification. Notably, we used lower temperature than those typically used in hot water pretreatments (160-230°C), and which are required to dissolve the hemicelluloses into the hydrolysate to render the cellulose

more accessible for subsequent enzymatic saccharification (Li *et al.*, 2014; Mosier *et al.*, 2005a).

An acid pretreatment mainly hydrolyzes hemicellulosic glycosidic bonds, dissolving the hemicellulose and partially pre-hydrolyzing the cellulose (Gómez *et al.*, 2014; Shuai *et al.*, 2010). Both acid pretreatments resulted in a severe reduction in the residual biomass, leaving only about 60% of the CWR of control (Figure 4). The residual biomass after acid pretreatment dropped even more for *zmcad2*, to about 54% of the CWR of the non-pretreated samples. The difference between control and *zmcad2* biomass is indicative for the fact that *zmcad2* has a higher fraction of acid-cleavable linkages in the cell wall. Possibly, this can be attributed to the fraction of acetal interunit bonds derived from the incorporation of ferulic acid into the lignin polymer (Figure 1) (Ralph *et al.*, 2008; Vanholme *et al.*, 2012a), which is higher in the *zmcad2*. Also after the acid pretreatments, the saccharification efficiency was higher for *zmcad2* compared to the control. Notably, the relative difference in saccharification efficiency between control and *zmcad2* after acid pretreatment was in the same range (+21% for 1M HCl) or lower (+12% for 0.4M H_2SO_4) as compared to no pretreatment (+20%). This hints that the additional structures that were removed during acid pretreatment from *zmcad2* and not from control samples, were either not hindering the saccharification process, or also contained a part of the cellulose.

Alkaline pretreatments mainly hydrolyze ester bonds, releasing (di)ferulates that acylate the hemicellulose and *p*-coumarate, ferulate, and acetate that acylate lignin (Ralph, 2010), thereby loosening the linkages between hemicelluloses and lignin and causing partial delignification of the biomass. Moreover, alkaline pretreatments will potentially also cleave 4–O-coupled coniferaldehyde, sinapaldehyde, syringaldehyde, and vanillin structures in the lignin because the (conjugated) carbonyl functionality increases the reactivity of the 8–O–4 aryl ether linkages towards hydrolysis (Oyarce *et al.*, 2019; Van Acker *et al.*, 2017). Previous studies have

shown that CAD-deficient plants have increased susceptibility towards alkaline pretreatment in several dicots (Anderson *et al.*, 2015; Baucher *et al.*, 1996; Ralph *et al.*, 2001; Van Acker *et al.*, 2017). In *CAD1*-downregulated poplars, only upon alkaline pretreatments, cellulose conversions were higher compared with the wild type (Van Acker *et al.*, 2017). In control maize, the influence of the minute amounts of aldehyde monomers on the delignification in alkaline conditions is probably negligible as compared to the influence of the hydrolysis of the abundant ester bonds. Lignin in *zmcad2* mutants has two compositional shifts with opposite effects on alkaline pretreatment. The increased amount of coniferaldehyde and sinapaldehyde units could potentially increase the delignification and solubilization after alkaline pretreatment, whereas the reduced amount of esterified *p*-coumaric acid and acetic acid in the cell wall would rather reduce the delignification and solubilization. The saccharification yields hint that these two effects level off, as the relative increase in saccharification after alkaline pretreatment was not higher than the increase in saccharification when no pretreatment was applied. The reason for the overall increased saccharification efficiency observed after alkaline pretreatments is likely attributable to the lower amount of lignin, not to the shifts in its composition.

Although pretreatments improve the efficiency of the process, they are expensive and energy demanding, making them less attractive for the biorefinery. Indeed, pretreatments can account for up to 20% of the total capital investment needed (Aden *et al.*, 2002). It is therefore highly beneficial to produce easily degradable feedstock and decrease the requirement for pretreatments. Notably, *zmcad2* mutants displayed a higher saccharification efficiency when no pretreatment was used, and also after eight out of the ten thermochemical pretreatments tested.

The actual parameters that will be used in future biorefineries are yet unknown, but will probably differ at several levels. First, the *zmcad2* mutant and its control studied here are inbred lines, whereas maize hybrids are typically cultivated because of their overall higher yield and

performance. Second, plants were harvested at silage stage, which is a commonly used harvesting stage to determine genotypic effects on cell wall composition and digestibility (see e.g., Barrière *et al.*, 2009; 2013; Torres *et al.*, 2016). It is a relevant harvesting stage for maize when used as cattle feed, but not necessarily the stage of choice when maize is will be used as a dual-purpose feedstock where grains are used as food and stover for cellulosic glucose production. Third, we used lower temperatures and longer incubation times for the pretreatment than those that will likely be used in industry (Mosier et al., 2005b). In addition, we used small-scale limited-saccharification conditions to allow a determination of the relative ease of saccharification; the parameters will differ from those that will be used in future industrial processes to maximize fermentable sugar yield. As with all models, it remains to be shown to what extent the maize *bm1/cad2* mutants will be useful as feedstock when grown as hybrids and processed in an industrial setting.

EXPERIMENTAL PROCEDURES

Plant material

As described previously (Barrière *et al.*, 2013), *zmcad2* (*Zmcad2-m2210::Mu*, Ev2210bm1) and control (Ev2210) were isolated by Biogemma. The plants were field-grown in 2009 at Clermont-Ferrand (France), eight plots per genotype, each plot containing 20–25 plants. Six plots were used for metabolite profiling: ear internodes (i.e., the internode below the cob) were harvested fourteen days after the silking stage. Hence, six biological replicates were made per genotype, each a pool of five internodes from different plants in a single plot. For two other plots, whole plants without the cob were harvested at silage stage (i.e., the kernel milk stage, ~70% moisture content), and five individuals per genotype were randomly picked (spread over

these two plots). The plant material was oven-dried, milled and the resulting powder was used for cell-wall and saccharification analyses.

Cell-wall analyses

Klason lignin was determined as described previously (de Lyra Soriano Saleme *et al.*, 2017). The lignin composition was analyzed with thioacidolysis as previously described (Van Acker *et al.*, 2013). Crystalline cellulose content was measured using a colorimetric method as described previously (Voorend *et al.*, 2016). The α-cellulose was determined as described previously (de Lyra Soriano Saleme *et al.*, 2017).

NMR

The ELs were prepared and analyzed via NMR samples in dimethyl sulfoxide-d_6/pyridine-d_5 (4:1, v/v) on a Bruker Biospin AVANCE 700 MHz spectrometer fitted with a cryogenically cooled 5-mm quadruple-resonance $^{1}H/^{31}P/^{13}C/^{15}N$ QCI gradient probe with inverse geometry (proton coils closest to the sample) as described previously (Kim and Ralph, 2010; Kim *et al.*, 2008; Mansfield *et al.*, 2012; Van Acker *et al.*, 2017). Volume integration of contours in HSQC plots was performed on data processed without linear prediction and used Bruker's TopSpin 4.0.7 (Mac version) software. Note that integrals are to be used for relative comparisons: they do not represent absolute quantification and, in particular, end-groups are substantially overestimated (Mansfield *et al.*, 2012). The peaks characteristic of lignin structures were estimated after the contour level adjustment to achieve optimal peak separation. N=5 for control and *zmcad2*.

Saccharification assay

Milled biomass (10 mg) samples were used as starting material for saccharification. Pretreatments were applied before saccharification. The biomass was either non-pretreated or pretreated for 2 h with 1 mL of either hot water (98°C), 1M HCl (80°C), 0.4M H_2SO_4 (80°C), 1M NH_4OH (90°C) or NaOH (6.25, 12.5, 25, 50, 62.5 or 150mM, 90°C). The residual biomass of the pretreated samples was subsequently washed three times with 1 mL water to obtain neutral pH. All samples, including those that were not subjected to a pretreatment, were subsequently incubated overnight at 55°C with 1 mL 70% (v/v) ethanol while shaking at 750 rpm. Next, all samples were sequentially washed three times with 1 mL 70% ethanol and once with 1 mL acetone. Pellets obtained after evaporation of the solvent were weighed to calculate the amount of CWR (for non-pretreated samples) and residual biomass (for pretreated samples). The saccharification assay was performed as (Van Acker *et al.*, 2016; Van Acker *et al.*, 2013).

Metabolic profiling

For the *zmcad2* mutant and the control, six biological replicates of the ear internode at fourteen days after silking were used for metabolic profiling. The samples were collected in liquid nitrogen and subsequently kept at -70°C. The metabolic profiling was performed as described previously (Eloy *et al.*, 2017). The most significantly increased and decreased metabolites in *zmcad2* compared to control consisted of *m/z* values fulfilling the following criteria: (1) detected in all samples of at least one genotype, (2) average normalized abundance higher than 5000 counts, (3) ten-fold increased/decreased abundance in *zmcad2* versus control and (4) $P<0.001$. Annotation of compounds was based on accurate *m/z* (+/- 0.02 Da), retention time, isotope distribution and MS/MS similarities.

Statistical analysis

All statistics for cell-wall and saccharification analyses were performed using SAS Enterprise Guide 6® (SAS Institute Inc., USA). One-way ANOVA determinations followed by post-hoc Dunnett's tests (2-sided) were applied to test for differences between *zmcad2* and control. For metabolic profiling, the normalized peak area was first transformed (arcsinh) and then subjected to a two-tailed student *t*-test, performed in Microsoft Excel (Microsoft, Redmond, Washington, USA).

ACKNOWLEDGEMENTS

This work was supported by the European Community's Seventh Framework Programme (FP7/2009) under grant agreement 251132 (SUNLIBB) and by the Hercules program of Ghent University for the Synapt Q-Tof (grant no. AUGE/014). H.K. and J.R. were funded by the DOE Great Lakes Bioenergy Research Center (DOE BER Office of Science DE- SC0018409), X.L. by the Chinese Scholarship Council (CSC) and Bijzonder Onderzoeksfonds (BOF), and A.P. and R.V. by the Research Foundation Flanders (Project G0C1914N).

AUTHOR CONTRIBUTIONS

M.B. and W.B. conceived and designed the analysis. X.L., R.V.A., W.V., A.P., G.G., H.K., and M.B. conducted the experiments. X.L., R.V.A., W.V., A.P., G.G., J.P., K.M., H.K., J.R., R.V., and W.B. analyzed and interpreted the data. X.L., J.R., R.V., and W.B. wrote the manuscript, with contributions of all authors. M.B., J.R. and W.B. funded the research. All authors read and approved the final manuscript.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the Supporting Information tab of this article:

Figure S1. MS/MS-based annotation of compounds with a higher or lower abundance in *zmcad2* plants compared to control plants.

Figure S2. MS/MS-based annotation of oligolignols and dihydroferulic acid + hexose.

Table S1. Enzyme lignin NMR data. NMR (2D HSQC) integral data for the lignin structures identified in the cell wall samples as shown in Figure 2.

Table S2. Saccharification results for control and *zmcad2* plants

REFERENCES

Aden, A., Ruth, M., Ibsen, K., Jechura, J., Neeves, K., Sheehan, J., Wallace, B., Montague, L., Slayton, A. and Lukas, J. (2002) *Lignocellulosic biomass to ethanol process design and economics utilizing co-current dilute acid prehydrolysis and enzymatic hydrolysis for corn stover. National Renewable Energy Laboratory Technical Report NREL/TP-510-32438 (http://www.nrel.gov/docs/fy02osti/32438.pdf)*.

Anderson, N.A., Tobimatsu, Y., Ciesielski, P.N., Ximenes, E., Ralph, J., Donohoe, B.S., Ladisch, M. and Chapple, C. (2015) Manipulation of guaiacyl and syringyl monomer biosynthesis in an Arabidopsis cinnamyl alcohol dehydrogenase mutant results in atypical lignin biosynthesis and modified cell wall structure. *Plant Cell* **27**, 2195-2209.

Barrière, Y., Chavigneau, H., Delaunay, S., Courtial, A., Bosio, M., Lassagne, H., Derory, J., Lapierre, C., Méchin, V. and Tatout, C. (2013) Different mutations in the *ZmCAD2* gene underlie the maize brown-midrib1 (bm1) phenotype with similar effects on lignin characteristics and have potential interest for bioenergy production. *Maydica* **58**, 6-20.

Barrière, Y., Méchin, V., Lafarguette, F., Manicacci, D., Guillon, F., Wang, H., Lauressergues, D., Pichon, M., Bosio, M. and Tatout, C. (2009) Toward the discovery of maize cell wall genes involved in silage maize quality and capacity to biofuel production. Maydica **54**, 161-198.

Barrière, Y., Ralph, J., Méchin, V., Guillaumie, S., Grabber, J.H., Argillier, O., Chabbert, B. and Lapierre, C. (2004) Genetic and molecular basis of grass cell wall biosynthesis and degradability. II. Lessons from brown-midrib mutants. *C. R. Biol.* **327**, 847-860.

Baucher, M., Bernard-Vailhé, M.A., Chabbert, B., Besle, J.-M., Opsomer, C., Van Montagu, M. and Botterman, J. (1999) Down-regulation of cinnamyl alcohol dehydrogenase in transgenic alfalfa (*Medicago sativa* L.) and the effect on lignin composition and digestibility. *Plant Mol. Biol.* **39**, 437-447.

Baucher, M., Chabbert, B., Pilate, G., Van Doorsselaere, J., Tollier, M.-T., Petit-Conil, M., Cornu, D., Monties, B., Van Montagu, M., Inzé, D., Jouanin, L. and Boerjan, W. (1996) Red xylem and higher lignin extractability by down-regulating a cinnamyl alcohol dehydrogenase in poplar. *Plant Physiol.* **112**, 1479-1490.

Bernard Vailhé, M.A., Besle, J.M., Maillot, M.P., Cornu, A., Halpin, C. and Knight, M. (1998) Effect of down‐regulation of cinnamyl alcohol dehydrogenase on cell wall composition and on degradability of tobacco stems. *J. Sci. Food Agric.* **76**, 505-514.

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

Bosch, M., Mayer, C.-D., Cookson, A. and Donnison, I.S. (2011) Identification of genes involved in cell wall biogenesis in grasses by differential gene expression profiling of elongating and non-elongating maize internodes. *J. Exp. Bot.* **62**, 3545-3561.

Bouvier d'Yvoire, M., Bouchabke-Coussa, O., Voorend, W., Antelme, S., Cézard, L., Legée, F., Lebris, P., Legay, S., Whitehead, C., McQueen-Mason, S.J., Gomez, L.D., Jouanin, L., Lapierre, C. and Sibout, R. (2013) Disrupting the *cinnamyl alcohol dehydrogenase 1* gene (*BdCAD1*) leads to altered lignification and improved saccharification in *Brachypodium distachyon*. *Plant J.* **73**, 496-508.

Carroll, A. and Somerville, C. (2009) Cellulosic biofuels. *Annu. Rev. Plant Biol.* **60**, 165-182.

Chabannes, M., Barakate, A., Lapierre, C., Marita, J.M., Ralph, J., Pean, M., Danoun, S., Halpin, C., Grima-Pettenati, J. and Boudet, A.M. (2001) Strong decrease in lignin content without significant alteration of plant development is induced by simultaneous down-regulation of cinnamoyl CoA reductase (CCR) and cinnamyl alcohol dehydrogenase (CAD) in tobacco plants. *Plant J.* **28**, 257-270.

Chen, F. and Dixon, R.A. (2007) Lignin modification improves fermentable sugar yields for biofuel production. *Nat. Biotechnol.* **25**, 759-761.

Chen, L., Auh, C.K., Dowling, P., Bell, J., Chen, F., Hopkins, A., Dixon, R.A. and Wang, Z.Y. (2003) Improved forage digestibility of tall fescue (*Festuca arundinacea*) by transgenic down-regulation of cinnamyl alcohol dehydrogenase. *Plant Biotechnol. J.* **1**, 437-449.

Chen, W., VanOpdorp, N., Fitzl, D., Tewari, J., Friedemann, P., Greene, T., Thompson, S., Kumpatla, S., Zheng, P. (2012) Transposon insertion in a cinnamyl alcohol dehydrogenase gene is responsible for a brown midrib1 mutation in maize. *Plant Mol Biol.* **80**, 289-297.

Coleman, H.D., Samuels, A.L., Guy, R.D. and Mansfield, S.D. (2008) Perturbed lignification impacts tree growth in hybrid poplar—A function of sink strength, vascular integrity, and photosynthetic assimilation. *Plant Physiol.* **148**, 1229-1237.

Dauwe, R., Morreel, K., Goeminne, G., Gielen, B., Rohde, A., Van Beeumen, J., Ralph, J., Boudet, A.-M., Kopka, J., Rochange, S.F., Halpin, C., Messens, E. and Boerjan, W. (2007) Molecular phenotyping of lignin-modified tobacco reveals associated changes in cell-wall metabolism, primary metabolism, stress metabolism and photorespiration. *Plant J.* **52**, 263-285.

de Lyra Soriano Saleme, M., Cesarino, I., Vargas, L., Kim, H., Vanholme, R., Goeminne, G., Van Acker, R., Campos de Assis Fonseca, F., Pallidis, A., Voorend, W., Nicomedes, J.J., Padmakshan, D., Van Doorsseleare, J., Ralph, J. and Boerjan, W. (2017) Silencing *CAFFEOYL SHIKIMATE ESTERASE* affects lignification and improves saccharification in poplar. *Plant Physiol.* **175**, 1040-1057.

De Meester, B., de Vries, L., Özparpucu, M., Gierlinger, N., Corneillie, S., Pallidis, A., Goeminne, G., Morreel, K., De Bruyne, M., De Rycke, R., Vanholme, R. and Boerjan, W. (2018) Vessel-specific reintroduction of CINNAMOYL-COA REDUCTASE1 (CCR1) in dwarfed *ccr1* mutants restores vessel and xylary fiber integrity and increases biomass. *Plant Physiol.* **176**, 611-633.

del Río, J.C., Rencoret, J., Prinsen, P., Martínez, A.n.T., Ralph, J. and Gutiérrez, A. (2012) Structural characterization of wheat straw lignin as revealed by analytical pyrolysis, 2D-NMR, and reductive cleavage methods. *J. Agric. Food Chem.* **60**, 5922-5935.

Eloy, N.B., Voorend, W., Lan, W., Saleme, M.D.S., Cesarino, I., Vanholme, R., Smith, R.A., Goeminne, G., Pallidis, A., Morreel, K., Nicomedes, J., Ralph, J. and Boerjan, W. (2017) Silencing *CHALCONE SYNTHASE* in maize impedes the incorporation of tricin into lignin and increases lignin content. *Plant Physiol.* **173**, 998-1016.

Elumalai, S., Tobimatsu, Y., Grabber, J.H., Pan, X. and Ralph, J. (2012) Epigallocatechin gallate incorporation into lignin enhances the alkaline delignification and enzymatic saccharification of cell walls. *Biotechnol. Biofuels* **5**, 59.

Emrani, N., Harloff, H.-J., Gudi, O., Kopisch-Obuch, F., and Jung C. (2015) Reduction in sinapine content in rapeseed (*Brassica napus L.*) by induced mutations in sinapine biosynthesis genes. *Mol. Breeding* **35**, 37.

Eudes, A., Liang, Y., Mitra, P. and Loqué, D. (2014) Lignin bioengineering. *Curr. Opin. Biotechnol.* **26**, 189-198.

Eyster, W.H. (1926) chromosome VIII in maize. *Science* **64**, 22.

Fornalé, S., Capellades, M., Encina, A., Wang, K., Irar, S., Lapierre, C., Ruel, K., Joseleau, J.-P., Berenguer, J., Puigdomènech, P., Rigau, J. and Caparrós-Ruiz, D. (2012) Altered lignin biosynthesis improves cellulosic bioethanol production in transgenic maize plants down-regulated for cinnamyl alcohol dehydrogenase. *Mol. Plant* **5**, 817-830.

Fornalé, S., Rencoret, J., García-Calvo, L., Encina, A., Rigau, J., Gutiérrez, A., del Río, J.C. and Caparros-Ruiz, D. (2017) Changes in cell wall polymers and degradability in maize mutants lacking 3´-and 5´-*O*-methyltransferases involved in lignin biosynthesis. *Plant Cell Physiol.* **58**, 240-255.

Freudenberg, K. (1959) Biosynthesis and constitution of lignin. *Nature* **183**, 1152-1155.

Fu, C., Xiao, X., Xi, Y., Ge, Y., Chen, F., Bouton, J., Dixon, R.A. and Wang, Z.-Y. (2011) Downregulation of cinnamyl alcohol dehydrogenase (CAD) leads to improved saccharification efficiency in switchgrass. *BioEnergy Res.* **4**, 153-164.

Gómez, L.D., Vanholme, R., Bird, S., Goeminne, G., Trindade, L.M., Polikarpov, I., Simister, R., Morreel, K., Boerjan, W. and McQueen-Mason, S.J. (2014) Side by side comparison of chemical compounds generated by aqueous pretreatments of maize stover, Miscanthus and sugarcane bagasse. *BioEnergy Res.* **7**, 1466-1480.

Grabber, J.H., Hatfield, R.D., Lu, F. and Ralph, J. (2008) Coniferyl ferulate incorporation into lignin enhances the alkaline delignification and enzymatic degradation of cell walls. *Biomacromolecules* **9**, 2510-2516.

Grabber, J.H., Schatz, P.F., Kim, H., Lu, F. and Ralph, J. (2010) Identifying new lignin bioengineering targets: 1. Monolignol-substitute impacts on lignin formation and cell wall fermentability. *BMC Plant Biol.* **10**, 114.

Hake, S. and Ross-Ibarra, J. (2015) The natural history of model organisms: genetic, evolutionary and plant breeding insights from the domestication of maize. *eLife* **4**, e05861.

Halpin, C. (2019) Lignin engineering to improve saccharification and digestibility in grasses. *Curr. Opin. Biotechnol.* **56**, 223-229.

Halpin, C., Holt, K., Chojecki, J., Oliver, D., Chabbert, B., Monties, B., Edwards, K., Barakate, A. and Foxon, G.A. (1998) *Brown-midrib* maize (*bm1*) – a mutation affecting the cinnamyl alcohol dehydrogenase gene. *Plant J.* **14**, 545-553.

Halpin, C., Knight, M.E., Foxon, G.A., Campbell, M.M., Boudet, A.M., Boon, J.J., Chabbert, B., Tollier, M.-T. and Schuch, W. (1994) Manipulation of lignin quality by down-regulation of cinnamyl alcohol-dehydrogenase. *Plant J.* **6**, 339-350.

Hatfield, R., Ralph, J. and Grabber, J.H. (2008) A potential role for sinapyl *p*-coumarate as a radical transfer mechanism in grass lignin formation. *Planta* **228**, 919-928.

Hatfield, R.D., Marita, J.M., Frost, K., Grabber, J., Ralph, J., Lu, F. and Kim, H. (2009) Grass lignin acylation: *p*-coumaroyl transferase activity and cell wall characteristics of C3 and C4 grasses. *Planta* **229**, 1253-1267.

Holtman, K.M., Chang, H.-m., Jameel, H. and Kadla, J.F. (2006) Quantitative ^{13}C NMR characterization of milled wood lignins isolated by different milling techniques. *J. Wood Chem. Technol.* **26**, 21-34.

Ibdah, M., Berim, A., Martens, S., Valderrama, A.L.H., Palmieri, L., Lewinsohn, E. and Gang, D.R. (2014) Identification and cloning of an NADPH-dependent hydroxycinnamoyl-CoA double bond reductase involved in dihydrochalcone formation in *Malus* × *domestica* Borkh. *Phytochemistry* **107**, 24-31.

Ishii, T. (1991) Isolation and characterization of a diferuloyl arabinoxylan hexasaccharide from bamboo shoot cell-walls. *Carbohydr. Res.* **219**, 15-22.

Jørgensen, H., Kristensen, J.B. and Felby, C. (2007) Enzymatic conversion of lignocellulose into fermentable sugars: challenges and opportunities. *Biofuels Bioprod. Biorefin.* **1**, 119–134.

Jung, H.-J.G. (2003) Maize stem tissues: ferulate deposition in developing internode cell walls. *Phytochemistry* **63**, 543-549.

Karlen, S.D., Zhang, C., Peck, M.L., Smith, R.A., Padmakshan, D., Helmich, K.E., Free, H.C.A., Lee, S., Smith, B.G., Lu, F., Sedbrook, J.C., Sibout, R., Grabber, J.H., Runge, T.M., Mysore, K.S., Harris, P.J., Bartley, L.E. and Ralph, J. (2016) Monolignol ferulate conjugates are naturally incorporated into plant lignins. *Sci. Adv.* **2**, e1600393.

Kim, H., Padmakshan, D., Li, Y.D., Rencoret, J., Hatfield, R.D. and Ralph, J. (2017) Characterization and elimination of undesirable protein residues in plant cell wall materials for enhancing lignin analysis by solution-state nuclear magnetic resonance spectroscopy. *Biomacromolecules* **18**, 4184-4195.

Kim, H. and Ralph, J. (2010) Solution-state 2D NMR of ball-milled plant cell wall gels in DMSO-d_6/pyridine-d_5. *Org. Biomol. Chem.* **8**, 576-591.

Kim, H., Ralph, J. and Akiyama, T. (2008) Solution-state 2D NMR of ball-milled plant cell wall gels in DMSO-d_6. *BioEnergy Res.* **1**, 56-66.

Kim, H., Ralph, J., Lu, F., Pilate, G., Leplé, J.-C., Pollet, B. and Lapierre, C. (2002) Identification of the structure and origin of thioacidolysis marker compounds for cinnamyl alcohol dehydrogenase deficiency in angiosperms. *J. Biol. Chem.* **277**, 47412-47419.

Končitíková, R., Vigouroux, A., Kopečná, M., Andree, T., Bartoš, J., Šebela, M., Moréra, S. and Kopečný, D. (2015) Role and structural characterization of plant aldehyde dehydrogenases from family 2 and family 7. *Biochem J.* **468**, 109–123.

Koshiba, T., Murakami, S., Hattori, T., Mukai, M., Takahashi, A., Miyao, A., Hirochika, H., Suzuki, S., Sakamoto, M. and Umezawa, T. (2013) *CAD2* deficiency causes both *brown midrib* and *gold hull and internode* phenotypes in *Oryza sativa* L. cv. Nipponbare. *Plant Biotechnol.* **30**, 365-373.

Kuc, J. and Nelson, O.E., (1964) The abnormal lignins produced by the *brown-midrib* mutants of maize: I. the *brown-midrib-1* mutant. *Arch. Biochem. Biophys.* **105**, 103-113.

Lan, W., Lu, F., Regner, M., Zhu, Y., Rencoret, J., Ralph, S.A., Zakai, U.I., Morreel, K., Boerjan, W. and Ralph, J. (2015) Tricin, a flavonoid monomer in monocot lignification. *Plant Physiol.* **167**, 1284-1295.

Lan, W., Morreel, K., Lu, F., Rencoret, J., del Río, J.C., Voorend, W., Vermerris, W., Boerjan, W. and Ralph, J. (2016) Maize tricin-oligolignol metabolites and their implications for monocot lignification. *Plant Physiol.* **171**, 810-820.

Lapierre, C. (2010) Determing lignin structure by chemical degradations. In: *Lignin and Lignans* (Heitner, C., Dimmel, D.R. and Schmidt, J.A. eds), pp. 11-48. Boca Raton, FL: CRC Press.

Lapierre, C., Pilate, G., Pollet, B., Mila, I., Leplé, J.-C., Jouanin, L., Kim, H. and Ralph, J. (2004) Signatures of cinnamyl alcohol dehydrogenase deficiency in poplar lignins. *Phytochemistry* **65**, 313-321.

Lapierre, C., Pollet, B., Petit-Conil, M., Toval, G., Romero, J., Pilate, G., Leplé, J.-C., Boerjan, W., Ferret, V., De Nadai, V. and Jouanin, L. (1999) Structural alterations of lignins in transgenic poplars with depressed cinnamyl alcohol dehydrogenase or caffeic acid *O*-methyltransferase activity have an opposite impact on the efficiency of industrial kraft pulping. *Plant Physiol.* **119**, 153-164.

Lapierre, C., Voxeur, A., Karlen, S.D., Helm, R.F. and Ralph, J. (2018) Evaluation of feruloylated and *p*-coumaroylated arabinosyl units in grass arabinoxylans by acidolysis in dioxane/methanol. *J. Agric. Food Chem.* **66**, 5418-5424.

Leplé, J.-C., Dauwe, R., Morreel, K., Storme, V., Lapierre, C., Pollet, B., Naumann, A., Kang, K.-Y., Kim, H., Ruel, K., Lefèbvre, A., Joseleau, J.-P., Grima-Pettenati, J., De Rycke, R., Andersson-Gunnerås, S., Erban, A., Fehrle, I., Petit-Conil, M., Kopka, J., Polle, A., Messens, E., Sundberg, B., Mansfield, S.D., Ralph, J., Pilate, G. and Boerjan, W. (2007) Downregulation of cinnamoyl-coenzyme A reductase in poplar: multiple-level phenotyping reveals effects on cell wall polymer metabolism and structure. *Plant Cell* **19**, 3669-3691.

Li, H.-Q., Jiang, W., Jia, J.-X. and Xu, J. (2014) pH pre-corrected liquid hot water pretreatment on corn stover with high hemicellulose recovery and low inhibitors formation. *Bioresour. Technol.* **153**, 292-299.

Li, M., Heckwolf, M., Crowe, J.D., Williams, D.L., Magee, T.D., Kaeppler, S.M., de Leon, N. and Hodge, D.B. (2015) Cell-wall properties contributing to improved deconstruction by alkaline pre-treatment and enzymatic hydrolysis in diverse maize (*Zea mays* L.) lines. *J. Exp. Bot.* **66**, 4305-4315.

Lu, F. and Ralph, J. (1999) Detection and determination of *p*-coumaroylated units in lignins. *J. Agric. Food Chem.* **47**, 1988-1992.

MacAdam, J.W. and Grabber, J.H. (2002) Relationship of growth cessation with the formation of diferulate cross-links and *p*-coumaroylated lignins in tall fescue leaf blades. *Planta* **215**, 785-793.

Mansell, R.L., Gross, G.G., Stöckigt, J., Franke, H. and Zenk, M.H. (1974) Purification and properties of cinnamyl alcohol dehydrogenase from higher plants involved in lignin biosynthesis. *Phytochemistry* **13**, 2427-2435.

Mansfield, S.D., Kim, H., Lu, F. and Ralph, J. (2012) Whole plant cell wall characterization using solution-state 2D NMR. *Nat. Protoc.* **7**, 1579-1589.

Mansfield, S.D., Mooney, C., Saddler, J.N. (1999) Substrate and enzyme characteristics that limit cellulose hydrolysis. *Biotechnol. Prog.* **15**, 804–816.

Marita, J.M., Hatfield, R.D., Rancour, D.M. and Frost, K.E. (2014) Identification and suppression of the *p*-Coumaroyl CoA:hydroxycinnamyl alcohol transferase in *Zea mays* L. *Plant J.* **78**, 850-864.

Marita, J.M., Vermerris, W., Ralph, J. and Hatfield, R.D. (2003) Variations in the cell wall composition of maize *brown midrib* mutants. *J. Agric. Food Chem.* **51**, 1313-1321.

Martin, A.F., Tobimatsu, Y., Kusumi, R., Matsumoto, N., Miyamoto, T., Lam, P.Y., Yamamura, M., Koshiba, T., Sakamoto, M. and Umezawa, T. (2019) Altered lignocellulose chemical structure and molecular assembly in *CINNAMYL ALCOHOL DEHYDROGENASE*-deficient rice. *Sci. Rep.* **9**, 17153.

Mechin, V., Argillier, O., Menanteau, V., Barriere, Y., Mila, I., Pollet, B. and Lapierre C. (2000) Relationship of cell wall composition to *in vitro* cell wall digestibility of maize inbred line stems. *J. Sci. Food Agric.*, **80**, 574-580.

Mir Derikvand, M., Berrio Sierra, J., Ruel, K., Pollet, B., Do, C.-T., Thévenin, J., Buffard, D., Jouanin, L. and Lapierre, C. (2008) Redirection of the phenylpropanoid pathway to feruloyl malate in *Arabidopsis* mutants deficient for cinnamoyl-CoA reductase 1. *Planta* **227**, 943-956.

Mittasch, J., Böttcher, C., Frolov, A., Strack, D. and Milkowski, C. (2013) Reprogramming the phenylpropanoid metabolism in seeds of oilseed rape by suppressing the orthologs of REDUCED EPIDERMAL FLUORESCENCE. *Plant Physiol.* **161**, 1656–1669.

Morreel, K., Dima, O., Kim, H., Lu, F., Niculaes, C., Vanholme, R., Dauwe, R., Goeminne, G., Inzé, D., Messens, E., Ralph, J. and Boerjan, W. (2010a) Mass spectrometry-based sequencing of lignin oligomers. *Plant Physiol.* **153**, 1464-1478.

Morreel, K., Kim, H., Lu, F., Dima, O., Akiyama, T., Vanholme, R., Niculaes, C., Goeminne, G., Inzé, D., Messens, E., Ralph, J. and Boerjan, W. (2010b) Mass spectrometry-based fragmentation as an identification tool in lignomics. *Anal. Chem.* **82**, 8095-8105.

Morreel, K., Saeys, Y., Dima, O., Lu, F., Van de Peer, Y., Vanholme, R., Ralph, J., Vanholme, B. and Boerjan, W. (2014) Systematic structural characterization of metabolites in *Arabidopsis* via candidate substrate-product pair networks. *Plant Cell* **26**, 929-945.

Morrison, T.A., Kessler, J.R., Hatfield, R.D. and Buxton, D.R. (1994) Activity of two lignin biosynthesis enzymes during development of a maize internode. *J. Sci. Food Agric.* **65**, 133-139.

Mosier, N., Hendrickson, R., Brewer, M., Ho, N., Sedlak, M., Dreshel, R., Welch, G., Dien, B., Aden, A. and Ladisch, M. (2005a). Industrial scale-up of pH-controlled liquid hot water pretreatment of corn fiber for fuel ethanol production. *Appl. Biochem. Biotechnol.*, **125**, 77-97.

Mosier, N., Wyman, C., Dale, B., Elander, R., Lee, Y.Y., Holtzapple, M. and Ladisch, M., (2005b) Features of promising technologies for pretreatment of lignocellulosic biomass. *Bioresource Technology* **96**, 673-686.

Nair, R.B., Bastress, K.L., Ruegger, M.O., Denault, J.W. and Chapple, C. (2004) The *Arabidopsis thaliana REDUCED EPIDERMAL FLUORESCENCE1* gene encodes an aldehyde dehydrogenase involved in ferulic acid and sinapic acid biosynthesis. *Plant Cell* **16**, 544-554.

F.J. Ostos Garrido, F.J., Pistón, F., Gómez, L.D., McQueen-Mason, S.J. (2018) Biomass recalcitrance in barley, wheat and triticale straw: correlation of biomass quality with classic agronomical traits. *PLoS One*, 13(11): e0205880.

Oyarce, P., De Meester, B., Fonseca, F., de Vries, L., Goeminne, G., Pallidis, A., De Rycke, R., Tsuji, Y., Li, Y., Van den Bosch, S., Sels, B., Ralph, J., Vanholme, R. and Boerjan, W. (2019) Introducing curcumin biosynthesis in *Arabidopsis* enhances lignocellulosic biomass processing. *Nat. Plants* **5**, 225-237.

Porter, K. S., Axtell, J. D., Lechtenberg, V. L. and Colenbrander, V. F. (1978) Phenotype, fiber composition, and *in-vitro* dry-matter disappearance of chemically-induced *brown midrib* (*bmr*) mutants of sorghum. *Crop Sci.* **18**, 205-208.

Provan, G.J., Scobbie, L. and Chesson, A. (1997) Characterisation of lignin from CAD and OMT deficient *Bm* mutants of maize. *J. Sci. Food Agric.* **73**, 133-142.

Pillonel, C., Mulder, M. M., Boon, J. J., Forster, B. and Binder, A. (1991) Involvement of cinnamyl-alcohol dehydrogenase in the control of lignin formation in *Sorghum bicolor* L. Moench. *Planta* **185**, 538–544.

Ralph, J. (2010) Hydroxycinnamates in lignification. *Phytochem. Rev.* **9**, 65-83.

Ralph, J., Hatfield, R.D., Quideau, S., Helm, R.F., Grabber, J.H. and Jung, H.-J.G. (1994) Pathway of *p*-coumaric acid incorporation into maize lignin as revealed by NMR. *J. Am. Chem. Soc.* **116**, 9448-9456.

Ralph, J., Kim, H., Lu, F., Grabber, J.H., Leplé, J.-C., Berrio-Sierra, J., Mir Derikvand, M., Jouanin, L., Boerjan, W. and Lapierre, C. (2008) Identification of the structure and origin of a thioacidolysis marker compound for ferulic acid incorporation into angiosperm lignins (and an indicator for cinnamoyl CoA reductase deficiency). *Plant J.* **53**, 368-379.

Ralph, J., Lapierre, C. and Boerjan, W. (2019) Lignin structure and its engineering. *Curr. Opin. Biotechnol.* **56**, 240-249.

Ralph, J., Lapierre, C., Marita, J.M., Kim, H., Lu, F., Hatfield, R.D., Ralph, S., Chapple, C., Franke, R., Hemm, M.R., Van Doorsselaere, J., Sederoff, R.R., O'Malley, D.M., Scott, J.T., MacKay, J.J., Yahiaoui, N., Boudet, A.-M., Pean, M., Pilate, G., Jouanin, L. and Boerjan, W. (2001) Elucidation of new structures in lignins of CAD- and COMT-deficient plants by NMR. *Phytochemistry* **57**, 993-1003.

Ralph, J., MacKay, J.J., Hatfield, R.D., O'Malley, D.M., Whetten, R.W. and Sederoff, R.R. (1997) Abnormal lignin in a loblolly pine mutant. *Science* **277**, 235-239.

Saathoff, A.J., Sarath, G., Chow, E.K., Dien, B.S. and Tobias, C.M. (2011) Downregulation of cinnamyl alcohol dehydrogenase in switchgrass by RNA silencing results in enhanced glucose release after cellulase treatment. *PLoS ONE* **6**, e16416.

Saballos, A., Vermerris, W., Rivera, L. and Ejeta, G. (2008) Allelic association, chemical characterization and saccharification properties of brown midrib mutants of sorghum (*Sorghum bicolor* (L.) Moench). *Bioenergy Res.* **2**, 198-204.

Saballos, A., Ejeta, G., Sanchez, E., Kang, C. and Vermerris, W. (2009) A genomewide analysis of the cinnamyl alcohol dehydrogenase family in sorghum (*Sorghum bicolor* (L.) Moench) identifies *SbCAD2* as the *Brown midrib6* gene. *Genetics* **181**, 783-795.

Sattler, S.E., Saathoff, A.J., Haas, E.J., Palmer, N.A., Funnell-Harris, D.L., Sarath, G. and Pedersen, J.F. (2009) A nonsense mutation in a cinnamyl alcohol dehydrogenase gene is responsible for the sorghum *brown midrib6* phenotype. *Plant Physiol.* **150**, 584-595.

Scully, E.D., Gries, T., Funnell-Harris, D.L., Xin, Z., Kovacs, F.A., Vermerris, W. and Sattler S.E. (2016) Characterization of novel *Brown midrib 6* mutations affecting lignin biosynthesis in sorghum. *J. Integr. Plant Biol.* **58**, 136-149.

Shuai, L., Yang, Q., Zhu, J.Y., Lu, F.C., Weimer, P.J., Ralph, J. and Pan, X.J. (2010) Comparative study of SPORL and dilute-acid pretreatments of spruce for cellulosic ethanol production. *Bioresour. Technol.* **101**, 3106-3114.

Sibout, R., Eudes, A., Mouille, G., Pollet, B., Lapierre, C., Jouanin, L. and Séguin, A. (2005) *CINNAMYL ALCOHOL DEHYDROGENASE-C* and *-D* are the primary genes involved in lignin biosynthesis in the floral stem of Arabidopsis. *Plant Cell* **17**, 2059-2076.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* **28**, 2731-2739.

Terrett, O.M. and Dupree, P. (2018) Covalent interactions between lignin and hemicelluloses in plant secondary cell walls. *Curr. Opin. Biotechnol.* **56**, 97-104.

Thévenin, J., Pollet, B., Letarnec, B., Saulnier, L., Gissot, L., Maia-Grondard, A., Lapierre, C. and Jouanin, L. (2011) The simultaneous repression of CCR and CAD, two enzymes of the lignin biosynthetic pathway, results in sterility and dwarfism in *Arabidopsis thaliana*. *Mol. Plant* **4**, 70-82.

Torres, A.F., Slegers, P.M., Noordam-Boot, C.M.M., Dolstra, O., Vlaswinkel, L., van Boxtel, A.J.B., Visser, R.G.F. and Trindade, L.M. (2016) Maize feedstocks with improved digestibility reduce the costs and environmental impacts of biomass pretreatment and saccharification. *Biotechnol. Biofuels* **9**, 63.

Trabucco, G.M., Matos, D.A., Lee, S.J., Saathoff, A.J., Priest, H.D., Mockler, T.C., Sarath, G. and Hazen, S.P. (2013) Functional characterization of cinnamyl alcohol dehydrogenase and caffeic acid *O*-methyltransferase in *Brachypodium distachyon*. *BMC Biotechnol.* **13**, 61.

Valério, L., Carter, D., Rodrigues, J.C., Tournier, V., Gominho, J., Marque, C., Boudet, A.-M., Maunders, M., Pereira, H. and Teulières, C. (2003) Down regulation of cinnamyl alcohol dehydrogenase, a lignification enzyme, in *Eucalyptus camaldulensis*. *Mol. Breeding* **12**, 157-167.

Van Acker, R., Déjardin, A., Desmet, S., Hoengenaert, L., Vanholme, R., Morreel, K., Laurans, F., Kim, H., Santoro, N., Foster, C., Goeminne, G., Légée, F., Lapierre, C., Pilate, G., Ralph, J. and Boerjan, W. (2017) Different routes for conifer- and sinapaldehyde and higher saccharification upon deficiency in the dehydrogenase CAD1. *Plant Physiol.* **175**, 1018-1039.

Van Acker, R., Leplé, J.-C., Aerts, D., Storme, V., Goeminne, G., Ivens, B., Légée, F., Lapierre, C., Piens, K., Van Montagu, M.C.E., Santoro, N., Foster, C.E., Ralph, J., Soetaert, W., Pilate, G. and Boerjan, W. (2014) Improved saccharification and ethanol yield from field-grown transgenic poplar deficient in cinnamoyl-CoA reductase. *Proc. Natl. Acad. Sci. USA* **111**, 845-850.

Van Acker, R., Vanholme, R., Piens, K. and Boerjan, W. (2016) Saccharification protocol for small-scale lignocellulosic biomass samples to test processing of cellulose into glucose. *Bio-Protocol* **6**, e1701 (http://www.bio-protocol.org/e1701).

Van Acker, R., Vanholme, R., Storme, V., Mortimer, J.C., Dupree, P. and Boerjan, W. (2013) Lignin biosynthesis perturbations affect secondary cell wall composition and saccharification yield in *Arabidopsis thaliana*. *Biotechnol. Biofuels* **6**, 46.

van der Weijde, T., Lessa Alvim Kamei, C., Torres, A.F., Vermerris, W., Dolstra, O., Visser, R.G.F. and Trindade, L.M. (2013) The potential of C4 grasses for cellulosic biofuel production. *Front. Plant Sci.* **4**, 107.

van der Weijde, T., Torres, A.F., Dolstra, O., Dechesne, A., Visser, R.G.F. and Trindade, L.M. (2016) Impact of different lignin fractions on saccharification efficiency in diverse species of the bioenergy crop Miscanthus. *Bioenerg. Res.* **9**, 146-156.

Vanholme, B., Desmet, T., Ronsse, F., Rabaey, K., Van Breusegem, F., De Mey, M., Soetaert, W. and Boerjan, W. (2013a) Towards a carbon-negative sustainable bio-based economy. *Front. Plant Sci.* **4**, 174.

Vanholme, R., Cesarino, I., Rataj, K., Xiao, Y., Sundin, L., Goeminne, G., Kim, H., Cross, J., Morreel, K., Araujo, P., Welsh, L., Haustraete, J., McClellan, C., Vanholme, B., Ralph, J., Simpson, G.G., Halpin, C. and Boerjan, W. (2013b) Caffeoyl shikimate esterase (CSE) is an enzyme in the lignin biosynthetic pathway in *Arabidopsis*. *Science* **341**, 1103-1106.

Vanholme, R., De Meester, B., Ralph, J. and Boerjan, W. (2019) Lignin biosynthesis and its integration into metabolism. *Curr. Opin. Biotechnol.* **56**, 230-239.

Vanholme, R., Demedts, B., Morreel, K., Ralph, J. and Boerjan, W. (2010a) Lignin biosynthesis and structure. *Plant Physiol.* **153**, 895-905.

Vanholme, R., Morreel, K., Darrah, C., Oyarce, P., Grabber, J.H., Ralph, J. and Boerjan, W. (2012a) Metabolic engineering of novel lignin in biomass crops. *New Phytol.* **196**, 978-1000.

Vanholme, R., Morreel, K., Ralph, J. and Boerjan, W. (2008) Lignin engineering. *Curr. Opin. Plant Biol.* **11**, 278-285.

Vanholme, R., Ralph, J., Akiyama, T., Lu, F., Pazo, J.R., Kim, H., Christensen, J.H., Van Reusel, B., Storme, V., De Rycke, R., Rohde, A., Morreel, K. and Boerjan, W. (2010b) Engineering traditional monolignols out of lignin by concomitant up-regulation of *F5H1* and down-regulation of *COMT* in Arabidopsis. *Plant J.* **64**, 885-897.

Vanholme, R., Storme, V., Vanholme, B., Sundin, L., Christensen, J.H., Goeminne, G., Halpin, C., Rohde, A., Morreel, K. and Boerjan, W. (2012b) A systems biology view of responses to lignin biosynthesis perturbations in *Arabidopsis*. *Plant Cell* **24**, 3506-3529.

Vermaas, J.V., Dixon, R.A., Chen, F., Mansfield, S.D., Boerjan, W., Ralph, J., Crowley, M.F. and Beckham, G.T. (2019) Passive membrane transport of lignin-related compounds. *Proc. Natl. Acad. Sci. USA* **116**, 23117-23123.

Vermerris, W., Saballos, A., Ejeta, G., Mosier, N.S., Ladisch, M.R. and Carpita, N.C. (2007) Molecular breeding to enhance ethanol production from corn and sorghum stover. *Crop Sci.* **47**, S142-S153.

Voorend, W., Nelissen, H., Vanholme, R., De Vliegher, A., Van Breusegem, F., Boerjan, W., Roldán-Ruiz, I., Muylle, H. and Inzé, D. (2016) Overexpression of *GA20-OXIDASE1* impacts plant height, biomass allocation and saccharification efficiency in maize. *Plant Biotechnol. J.* **14**, 997–1007.

Wu, R.L., Remington, D.L., MacKay, J.J., McKeand, S.E. and O'Malley, D.M. (1999) Average effect of a mutation in lignin biosynthesis in loblolly pine. *Theor. Appl. Genet.* **99**, 705-710.

Yamamoto, M., Tomiyama, H., Koyama, A., Okuizumi, H., Liu, S., Vanholme, R., Goeminne, G., Hirai, Y., Shi, H., Nuoendagula, N., Takata, N., Ikeda, T., Uesugi, M., Kim, H., Sakamoto, S., Mitsuda, N., Boerjan, W., Ralph, J. and Kajita, S. (2020) A century-old mystery unveiled: Sekizaisou is a natural lignin mutant. *Plant Physiol.*, **182**, 1821-1828.

Yan, X., Liu, J., Kim, H., Liu, B., Huang, X., Yang, Z., Lin, Y.-C.J., Chen, H., Yang, C., Wang, J.P., Muddiman, D.C., Ralph, J., Sederoff, R.R., Li, Q. and Chiang, V.L. (2019) CAD1 and CCR2 protein complex formation in monolignol biosynthesis in *Populus trichocarpa*. *New Phytol.* **222**, 244-260.

Zeng, Y., Zhao, S., Yang, S. and Ding, S.-Y. (2014) Lignin plays a negative role in the biochemical process for producing lignocellulosic biofuels. *Curr. Opin. Biotechnol.* **27**, 38-45.

Zhang, K., Bhuiya, M.-W., Rencoret Pazo, J., Miao, Y., Kim, H., Ralph, J. and Liu, C.-J. (2012) An engineered monolignol 4-*O*-methyltransferase depresses lignin biosynthesis and confers novel metabolic capability in *Arabidopsis*. *Plant Cell* **24**, 3135-3152.

Zhang, K., Qian, Q., Huang, Z., Wang, Y., Li, M., Hong, L., Zeng, D., Gu, M., Chu, C. and Cheng, Z. (2006) *GOLD HULL AND INTERNODE2* encodes a primarily multifunctional cinnamyl-alcohol dehydrogenase in rice. *Plant Physiol.* **140**, 972-983.

FIGURES AND TABLES

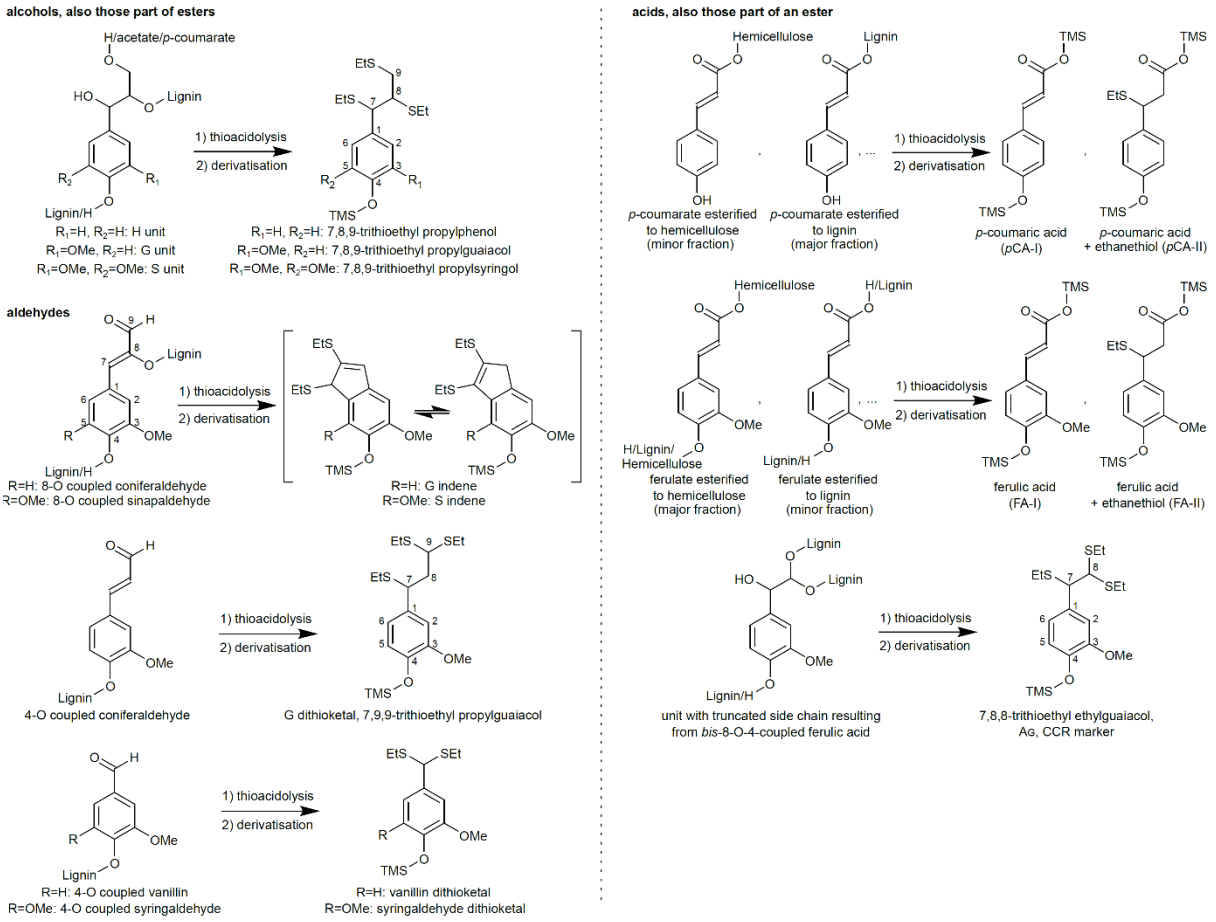

Figure 1. Products derived via thioacidolysis followed by trimethylsilyl (TMS) derivatization. Structures of the main thioacidolysis-derived compounds found in *zmcad2*, including the conventional thioacidolysis monomers that are derived from cleaving β-aryl ether structures. G indene and S indene are markers for the incorporation of 8–O–4-coupled coniferaldehyde and sinapaldehyde, respectively, whereas G dithioketal is derived from 8–O–4-coupled coniferaldehyde units (Kim *et al.*, 2002; Lapierre *et al.*, 2004). Vanillin dithioketal and syringaldehyde dithioketal are markers for incorporation of 4–O-coupled vanillin and syringaldehyde, respectively (Lapierre *et al.*, 2004; Ralph *et al.*, 2008). *p*CA-I and *p*CA-II are markers for the presence of *p*-coumarates in the cell wall, and FA-I and FA-II are markers for the presence of ferulates in the cell wall (Ralph *et al.*, 2008). Ferulic acid as a genuine lignin monomer is indicated by the A_G compound (Ralph *et al.*, 2008).

Figure 2. $^{1}H–^{13}C$ HSQC NMR spectra from enzyme lignins of control and *zmcad2* lines at silage stage. The genotype is indicated in the top left corner of each plot. (A) Lignin aromatic and double bond region. The relative signals of the S, S*, G, and G' units sum to 100%. (B) Lignin oxygenated-aliphatic and polysaccharide region. (C) Aldehyde region. For A, B and C, the colors of the contours correspond with the structures drawn. The average values for both control (n=5) and *zmcad2* (n=5) are given, and a representative spectrum is shown. Level data are from uncorrected integrals only.

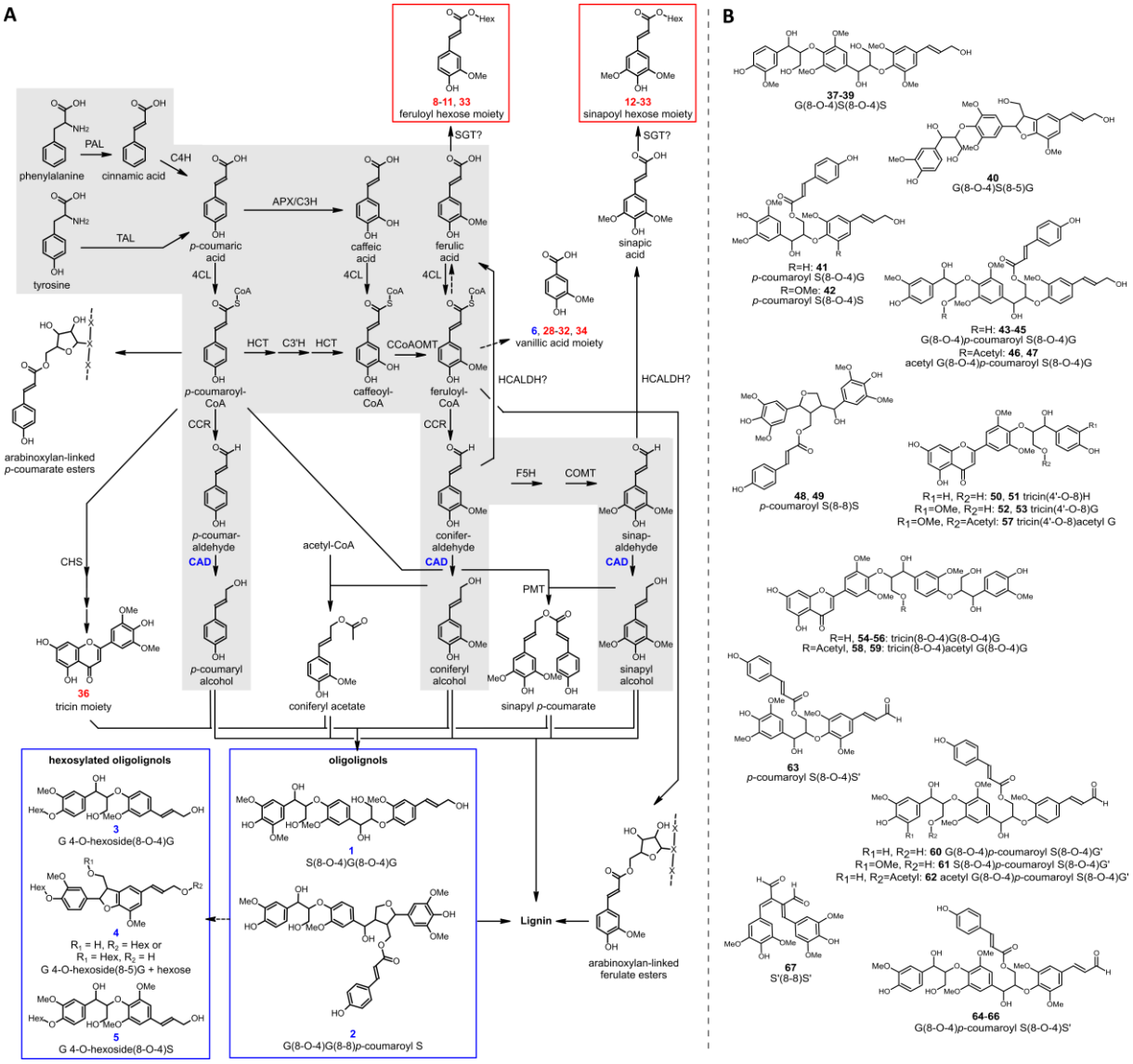

Figure 3. Metabolic pathway and structurally characterized oligolignols. (A) The main pathway involved in monolignol biosynthesis is shown with a grey background. The relative increase and decrease in metabolite abundances in *zmcad2* compared to control are mapped on the pathway. Red and blue boxes represent metabolic classes with an increased and decreased abundance, respectively. Solid arrows represent enzymatic conversions for which experimental evidence is available, and dashed arrows represent suggested conversions. Two successive arrows represent two or more metabolic conversions. PAL, phenylalanine ammonia-lyase; TAL, tyrosine ammonia-lyase; APX/C3H, ascorbate peroxidase / *p*-coumarate 3-hydroxylase C4H, cinnamate 4-hydroxylase; 4CL, 4-coumaric acid:CoA ligase; HCT, hydroxycinnamoyl-

CoA shikimate hydroxycinnamoyl transferase; C3'H, *p*-coumaroyl shikimate 3'-hydroxylase; CCoAOMT, caffeoyl-CoA *O*-methyltransferase; CCR, cinnamoyl-CoA reductase; F5H, ferulate 5-hydroxylase; COMT, caffeic acid *O*-methyltransferase; CAD, cinnamyl alcohol dehydrogenase; HCALDH, hydroxycinnamaldehyde dehydrogenase; SGT, UDP-glucose:sinapate glucosyltransferase; CHS, chalcone synthase. (B) Proposed molecular structures of oligolignols identified via UPLC-MS (Table 5). For the structural elucidation of the compounds, see Figure S2.

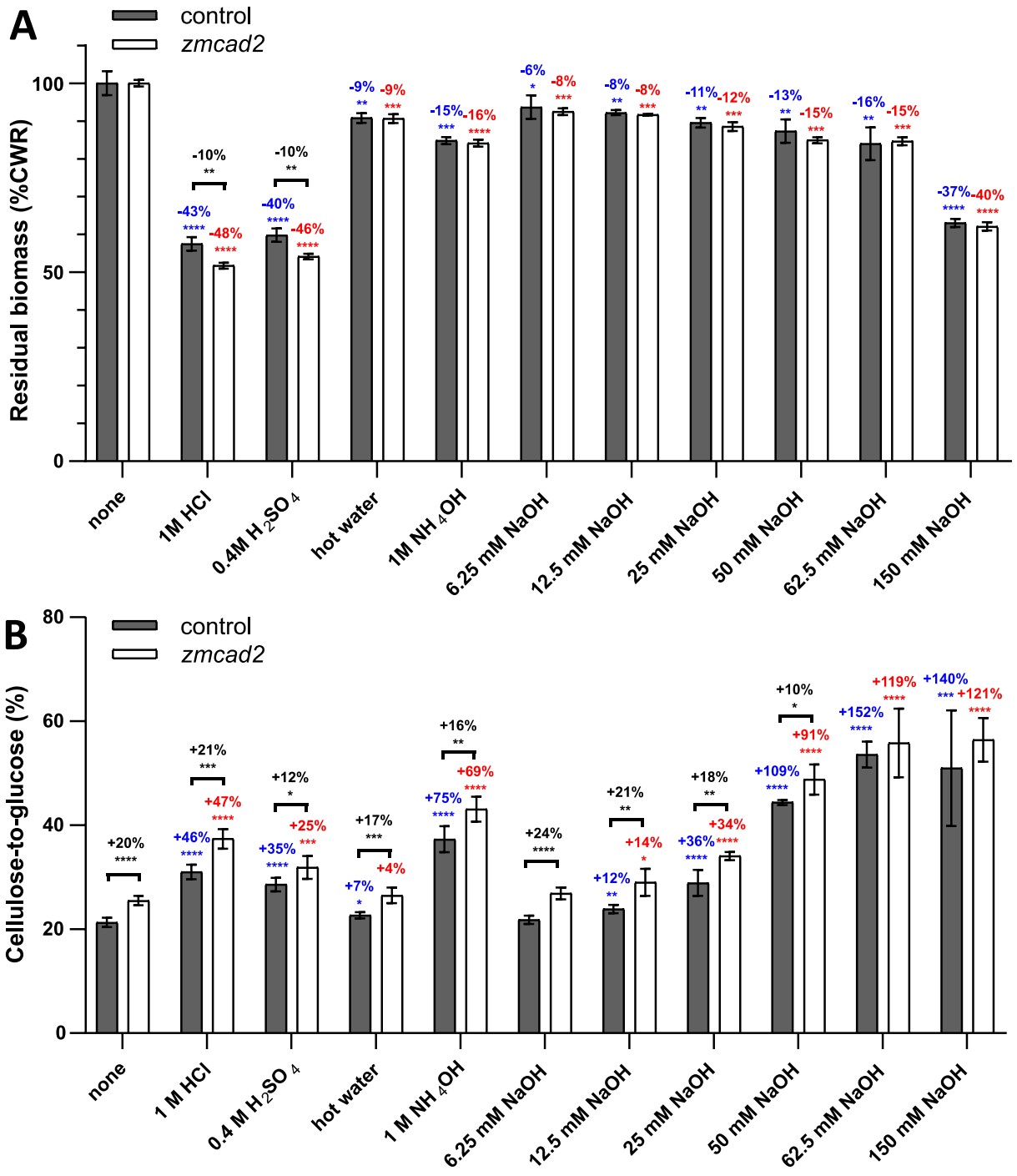

Figure 4. Residual biomass after pretreatment and cellulose-to-glucose conversion. (A) The average amount of residual biomass, expressed per amount of cell wall residue, that is left over after pretreatment and washing with ethanol for control (grey bars) and *zmcad2* (white bars) biomass (n=5). (B) The cellulose-to-glucose conversion rate is calculated based on the amount of glucose released after 48 h of saccharification and the total amount of cellulose (alpha cellulose) in the non-pretreated CWR, expressed as percentage. For A and B: The pretreatments

are arranged according to increasing pH. Error bars represent standard deviations. The relative difference between control and *zmcad2*, for a particular pretreatment, is given as percentage in black. The relative difference between a particular pretreatment and no treatment (none) for control and *zmcad2* is given as percentage in blue and red, respectively. Statistical significances were determined using the Dunnett adjusted *t*-test (*P < 0.05, **P < 0.01, ***<0.001, ****P < 0.0001).

Table 1. Cell wall and lignin composition of *zmcad2* mutants. The data represent means ± SD (n=5). Lignin composition was analyzed via thioacidolysis. The nomenclature is explained in Figure 1. S/G is a ratio, all other values are expressed in μmol per gram Klason lignin. Boldface or underlined values indicate significantly increased or decreased values, respectively, as compared with those of control. For the parameters that are significantly different between control and *zmcad2*, the percentage of difference is indicated (Dunnett adjusted *t*-test, P<0.05). b.d.l., below detection limit; inf., infinite (or just very large, from division by zero).

	control	***zmcad2***	**% difference**
CWR, % dry weight	66.0 ± 2.1	<u>63.4 ± 1.4</u>	−4 %
Klason lignin, % CWR	12.0 ± 1.0	<u>9.8 ± 0.4</u>	−18 %
alpha cellulose, % CWR	41.7 ± 0.2	42.4 ± 1.2	
crystalline cellulose, % CWR	21.7 ± 6.0	18.0 ± 4.1	
H unit	16.7 ± 1.2	<u>10.7 ± 0.2</u>	−36%
G unit	387 ± 39	<u>252 ± 9</u>	−35%
S unit	477 ± 58	<u>300 ± 14</u>	−37%
H + G + S units	881 ± 58	<u>562 ± 34</u>	−36%
S/G ratio	1.22 ± 0.04	1.19 ± 0.01	
8–O–4-coupled coniferaldehyde (G indene)	b.d.l.	**2.87 ± 0.13**	inf.
8–O–4-coupled sinapaldehyde (S indene)	0.16 ± 0.03	**3.35± 0.05**	+1994%
4–O-coupled coniferaldehyde (G dithioketal)	1.19 ± 0.27	**2.99 ± 0.01**	+151%
p-coumaric acid (*p*CA-I)	88.9 ± 5.6	<u>65.2 ± 2.3</u>	−27%
p-coumaric acid + ethanethiol (*p*CA-II)	33.4 ± 4.9	23.8 ± 1.2	
ferulic acid (FA-I)	27.7 ± 1.3	31.8 ± 2.7	
ferulic acid + ethanethiol (FA-II)	21.1 ± 3.1	23.1 ± 1.5	
bis 8–O–4 coupled ferulic acid (A_G)	0.35 ± 0.11	**1.64 ± 0.48**	+369%
syringaldehyde dithioketal	1.22 ± 0.30	**3.81 ± 0.91**	+212%
vanillin dithioketal	3.64± 0.79	**12.14 ± 2.33**	+233%

Table 2. Phenolic profiling of *zmcad2*. For each compound, its unique number (No.), peak area (average ± SD, n=6) and ratio of the peak area in *zmcad2* as compared to the control are given. [1]Compounds detected as formic acid adduct. [2]Vanilloyl pentosyl hexose was detected as heterodimer with a co-eluting compound. For the structural elucidation of the compounds, see Figure 3 and Figure S1.

				control			*zmcad2*			
No.	**RT (min)**	**$m/z_{experimental}$**	**name**	**average**	**±**	**s.d.**	**average**	**±**	**s.d.**	***zmcad2*/control**
down										
1	11.62	601.2277	S(8-O-4)G(8-O-4)G	9978	±	672	336	±	208	0.034
2	19.73	777.2753	G(8-O-4)S(8-8)*p*-coumaroyl S	14531	±	2700	965	±	563	0.066
3	7.41	583.2002	G 4-O-hexoside(8–O–4)G [1]	17726	±	2228	1196	±	431	0.067
4	7.43	727.2447	G 4–O–hexoside(8–5)G + hexose [1]	11695	±	1131	282	±	201	0.024
5	7.59	613.2108	G 4–O-hexoside (8–O–4)S [1]	10030	±	1657	646	±	283	0.064
6	2.35	803.2462	vanilloyl pentosyl hexose [2]	10915	±	2827	1031	±	321	0.094
7	11.65	817.1838	tricin + hexuronic acid + pentose + syringic acid	10062	±	714	710	±	501	0.071
up										
8	5.26	355.1017	feruloyl hexose	2573	±	3036	61049	±	12710	24
9	9.21	563.1399	feruloyl hexose + 208 Da	2	±	6	26293	±	6081	>100
10	7.84	581.1871	feruloyl hexose + 226 Da	1333	±	1512	29756	±	4276	22
11	8.15	639.1924	feruloyl hexose + hexose + 122 Da	220	±	240	12534	±	2773	57
12	5.47	385.1130	sinapoyl hexose 1	16378	±	2279	370493	±	44073	23
13	6.04	385.1128	sinapoyl hexose 2	1027	±	240	42439	±	6397	41
14	6.28	547.1666	sinapoyl hexose + hexose	475	±	237	37917	±	5745	80
15	10.92	537.1605	sinapoyl hexose + 152 Da 1	2110	±	438	47561	±	7372	23
16	11.81	537.1611	sinapoyl hexose + 152 Da 2	679	±	105	19807	±	3319	29
17	12.63	547.1816	sinapoyl hexose + 162 Da	4	±	6	21874	±	4441	>100
18	13.12	563.1761	sinapoyl hexose + 178Da	40	±	47	24212	±	5285	>100

19	8.62	593.1510	sinapoyl hexose + 208 Da	457	±	206	18679	±	3556	41
20	11.37	789.2247	sinapoyl hexose + 198 Da + 206 Da	0	±	0	14908	±	2511	>100
21	13.22	591.1717	disinapoyl hexose	0	±	0	26976	±	6029	>100
22	10.95	591.1718	disinapoyl hexose	1	±	2	19012	±	3869	>100
23	9.16	753.2251	disinapoyl hexose + hexose	3	±	5	34855	±	6359	>100
24	11.25	725.2298	sinapoyl hexose + hexose + 178Da 1	0	±	0	20433	±	4118	>100
25	11.54	725.2298	sinapoyl hexose + hexose + 178Da 2	0	±	0	13574	±	2624	>100
26	11.07	887.2841	sinapoyl hexose + hexose + hexose + 178Da	0	±	0	12249	±	3026	>100
27	13.15	905.2745	sinapoyl hexose + sinapic acid + hexose + 152 Da	62	±	81	42007	±	7583	>100
28	11.64	535.1460	sinapoyl hexose + vanillic acid	3683	±	667	44722	±	8313	12
29	9.47	697.1954	sinapoyl hexose + vanillic acid + hexose	254	±	115	57465	±	9200	>100
30	12.01	847.2319	sinapoyl hexose + vanilloyl hexose + vanillic acid	0	±	0	10918	±	2430	>100
31	11.50	819.2346	sinapoyl hexose + vanilloyl hexose + 122 Da	3364	±	812	80672	±	13703	24
32	11.71	849.2463	sinapoyl hexose + vanilloyl hexose + 152 Da	83	±	52	35422	±	6545	>100
33	10.21	723.2141	sinapoyl feruloyl hexose + hexose	117	±	54	20651	±	4354	>100
34	9.69	653.2059	vanilloyl coniferyl hexoside + hexose	517	±	246	13066	±	2561	25
35	2.63	299.0759	hydroxybenzoyl hexose	752	±	128	32093	±	4675	43
36	15.85	389.0871	tricin + 60 Da	10	±	14	18813	±	2539	>100

Table 3. Targeted metabolites. For each compound, its unique number (No.), peak area (average ± SD, n=6) and ratio of the peak area in *zmcad2* as compared to the control are given. For the structural elucidation of the compounds, see Figure S2.

				control			*zmcad2*				
No.	**RT (min)**	***m/z*experimental**	**name**	**average**	**±**	**s.d.**	**average**	**±**	**s.d.**	***zmcad2*/control**	***p*-value**
37	12.96	631.2390	G(8–O–4)S(8–O–4)S 1	10616	±	858	1290	±	333	0.12	<0.001
38	13.99	631.2390	G(8–O–4)S(8–O–4)S 3	18729	±	988	2868	±	602	0.15	<0.001
39	13.49	631.2390	G(8–O–4)S(8–O–4)S 2	39124	±	2772	5937	±	1020	0.15	<0.001
40	14.76	583.2132	G(8–O–4)S(8–5)G 1	50780	±	2744	8984	±	1239	0.18	<0.001
41	18.05	551.1921	*p*-coumaroyl S(8–O–4)G	10050	±	1310	7790	±	918	0.78	0.0054
42	19.03	581.2025	*p*-coumaroyl S(8–O–4)S	9561	±	1130	2536	±	644	0.27	<0.001
43	19.42	747.2656	G(8–O–4)*p*-coumaroyl S(8–O–4)G 1	16448	±	2324	8050	±	551	0.49	<0.001
44	18.35	747.2654	G(8–O–4)*p*-coumaroyl S(8–O–4)G 2	17058	±	2332	7198	±	611	0.42	<0.001
45	18.93	747.2657	G(8–O–4)*p*-coumaroyl S(8–O–4)G 3	23377	±	3162	11146	±	555	0.48	<0.001
46	21.58	789.2765	acetyl G(8–O–4)*p*-coumaroyl S(8–O–4)G 1	3245	±	1317	23	±	29	0.01	0.001
47	21.24	789.2767	acetyl G(8–O–4)*p*-coumaroyl S(8–O–4)G 2	597	±	514	4	±	9	0.01	<0.001
48	16.87	581.2015	*p*-coumaroyl S(8–8)S 1	62132	±	7056	15223	±	2273	0.25	<0.001
49	17.28	581.2018	*p*-coumaroyl S(8–8)S 2	14481	±	1452	4701	±	809	0.32	<0.001
50	17.57	495.1294	tricin(4'–O–8)H 1	12602	±	3944	10645	±	2026	0.84	0.408
51	18.31	495.1293	tricin(4'–O–8)H 2	9193	±	3280	8215	±	1826	0.89	0.698
52	17.96	525.1405	tricin(4'–O–8)G 1	112073	±	14089	82949	±	7507	0.74	0.001
53	18.70	525.1401	tricin(4'–O–8)G 2	82995	±	12218	61481	±	6041	0.74	0.003
54	16.11	721.2123	tricin(4'–O–8)G(4–O–8)G 3	14380	±	2481	3635	±	769	0.25	<0.001
55	16.37	721.2136	tricin(4'–O–8)G(4–O–8)G 1	10671	±	2147	2745	±	532	0.26	<0.001
56	16.81	721.2134	tricin(4'–O–8)G(4–O–8)G 2	30414	±	4657	9245	±	1160	0.30	<0.001
57	22.09	567.1501	tricin(4'–O–8)acetyl G	19056	±	5127	4907	±	712	0.26	<0.001

58	20.25	763.2244	tricin(4'–O–8)acetyl G(4–O–8)G 1	11199	±	1444	113	±	71	0.01	0.002
59	20.99	763.2244	tricin(4'–O–8)acetyl G(4–O–8)G 2	2041	±	359	39	±	32	0.02	0.004
60	20.81	745.2490	G(8–O–4)*p*-coumaroyl S(8–O–4)G'	3290	±	1336	3699	±	790	1.12	0.411
61	20.97	775.2604	S(8–O–4)*p*-coumaroyl S(8–O–4)G'	1291	±	451	164	±	80	0.13	<0.001
62	23.01	787.2603	acetyl G(8–O–4)*p*-coumaroyl S(8–O–4)G'	1245	±	676	241	±	247	0.19	0.007
63	20.75	579.1867	*p*-coumaroyl S(8–O–4)S'	861	±	171	5778	±	1563	6.71	<0.001
64	21.28	775.2614	G(8–O–4)*p*-coumaroyl S(8–O–4)S'	1543	±	717	4401	±	780	2.85	0.003
65	21.53	775.2604	G(8–O–4)*p*-coumaroyl S(8–O–4)S'	1307	±	422	5720	±	1145	4.38	<0.001
66	21.81	775.2605	G(8–O–4)*p*-coumaroyl S(8–O–4)S'	234	±	173	3754	±	1035	16.02	0.016
67	12.25	413.1222	S'(8–8)S'	0	±	0	448	±	374	>100	-
68	4.65	357.1179	Dihydroferulic acid + hexose	1349	±	248	7317	±	671	5.43	<0.001