

Biosynthesis and incorporation of side-chain-truncated lignin monomers to reduce lignin polymerization and enhance saccharification

Aymerick Eudes^{1,2}, Anthe George^{1,3}, Purba Mukerjee⁴, Jin S. Kim^{1,2}, Brigitte Pollet⁵, Peter I. Benke^{1,2}, Fan Yang^{1,2}, Prajakta Mitra^{1,2}, Lan Sun^{1,3}, Özgül P. Çetinkol^{1,3,†}, Salem Chabout⁵, Grégory Mouille⁵, Ludivine Soubigou-Taconnat⁶, Sandrine Balzergue⁶, Seema Singh^{1,3}, Bradley M. Holmes^{1,3}, Aindrila Mukhopadhyay^{1,2}, Jay D. Keasling^{1,2,7}, Blake A. Simmons^{1,3}, Catherine Lapierre⁵, John Ralph⁴ and Dominique Loque^{1,2,*}

¹Joint BioEnergy Institute, EmeryStation East, Emeryville, CA, USA

²Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA

³Sandia National Laboratory, Livermore, CA, USA

⁴Department of Biochemistry, The Wisconsin Bioenergy Initiative, and the DOE Great Lakes Bioenergy Research Center, University of Madison, Madison, WI, USA

⁵Institut Jean Pierre Bourgin, INRA-AgroParis Tech, UMR1318, Versailles, France

⁶UMR INRA 1165 CNRS 8114, Recherche en Génomique Végétale, UEVE, Evry Cedex, France

⁷Departments of Bioengineering, Chemical & Biomolecular Engineering, University of California, Berkeley, CA, USA

Received 22 December 2011;

revised 9 February 2012;

accepted 14 February 2012.

*Correspondence (Tel +1 510 486 7332;

fax +1 510 486 4252;

email dloque@lbl.gov)

†Present address: Chemistry Division, Doğuş University, Istanbul, Turkey.

Summary

Lignocellulosic biomass is utilized as a renewable feedstock in various agro-industrial activities. Lignin is an aromatic, hydrophobic and mildly branched polymer integrally associated with polysaccharides within the biomass, which negatively affects their extraction and hydrolysis during industrial processing. Engineering the monomer composition of lignins offers an attractive option towards new lignins with reduced recalcitrance. The presented work describes a new strategy developed in *Arabidopsis* for the overproduction of rare lignin monomers to reduce lignin polymerization degree (DP). Biosynthesis of these ‘DP reducers’ is achieved by expressing a bacterial hydroxycinnamoyl-CoA hydratase-lyase (HCHL) in lignifying tissues of *Arabidopsis* inflorescence stems. HCHL cleaves the propanoid side-chain of hydroxycinnamoyl-CoA lignin precursors to produce the corresponding hydroxybenzaldehydes so that plant stems expressing HCHL accumulate in their cell wall higher amounts of hydroxybenzaldehyde and hydroxybenzoate derivatives. Engineered plants with intermediate HCHL activity levels show no reduction in total lignin, sugar content or biomass yield compared with wild-type plants. However, cell wall characterization of extract-free stems by thioacidolysis and by 2D-NMR revealed an increased amount of unusual C₆C₁ lignin monomers most likely linked with lignin as end-groups. Moreover the analysis of lignin isolated from these plants using size-exclusion chromatography revealed a reduced molecular weight. Furthermore, these engineered lines show saccharification improvement of pretreated stem cell walls. Therefore, we conclude that enhancing the biosynthesis and incorporation of C₆C₁ monomers (‘DP reducers’) into lignin polymers represents a promising strategy to reduce lignin DP and to decrease cell wall recalcitrance to enzymatic hydrolysis.

Keywords: cell wall, lignin, hydroxycinnamoyl-CoA hydratase-lyase, saccharification, polymerization degree, bioenergy.

Introduction

Plant cell walls are composed of a polysaccharide-rich network of cellulose microfibrils and hemicelluloses embedded in aromatic polymers known as lignins. Lignins are mainly derived from polymerization of three hydroxyphenylpropanoid monomers (i.e. monolignols), namely *p*-coumaryl, coniferyl and sinapyl alcohols that give rise to the *p*-hydroxyphenyl (H), guaiacyl (G) and syringyl (S) lignin units (Boerjan *et al.*, 2003). Monolignols have a C₆C₃ carbon skeleton that consists of a phenyl ring (C₆) and a propanoid (C₃) side-chain. Lignin is essential for higher plants as it provides mechanical strength for upright growth, confers hydrophobicity to vessels that transport water and nutrients and acts as a physical barrier against pathogens that degrade cell walls (Boudet, 2007).

Lignocellulosic biomass is widely used as a raw material for the industrial production of pulp and paper and as a ruminant livestock feed. Plant feedstocks also represent a source of fermentable sugars for the production of ethanol as well as pharmaceuticals and advanced fuels using engineered micro-organisms (Keasling, 2010). However, lignin confers recalcitrance to the processing of plant cell walls, and negative correlations exist between lignin content in plant biomass and pulp yield, forage digestibility or polysaccharide saccharifiability (Baucher *et al.*, 2003; Chen and Dixon, 2007; Jung and Allen, 1995; Studer *et al.*, 2011; Taboada *et al.*, 2010). Consequently, reducing lignin content from plant feedstocks or facilitating its removal is of major interest in the field of lignocellulosic biofuels and pulping (Simmons *et al.*, 2010).

Lignin biosynthesis has been extensively studied and was shown to be well conserved across land plants (Halpin, 2004; Umezawa, 2010; Weng and Chapple, 2010). Genetic modifications such as transcript reduction and allelic variation of specific genes involved in particular steps of this pathway have been employed to reduce lignin content in various plant species (Chen and Dixon, 2007; Leplé *et al.*, 2007; Li *et al.*, 2008; Xu *et al.*, 2011). However, these approaches often result in undesired phenotypes such as dwarfism, sterility and increased susceptibility to environmental stresses owing to loss of cell wall integrity or the constitutive activation of defence responses (Bonawitz and Chapple, 2010; Chen and Dixon, 2007; Gallego-Giraldo *et al.*, 2011a,b; Voelker *et al.*, 2011).

Alternatively, changing the recalcitrant structure of lignin can be achieved by modifying its monomer composition and its physico-chemical properties. For example, incorporation of coniferyl ferulate into lignin improves enzymatic degradation of cell wall polysaccharides (Grabber *et al.*, 2008). Moreover, it has been demonstrated recently that enrichment in either 5-hydroxy-G- or S-units in lignin contributes to enhanced saccharification or pulping efficiencies without significantly affecting biomass yields and lignin content (Dien *et al.*, 2011; Fu *et al.*, 2011; Li *et al.*, 2010; Stewart *et al.*, 2009).

In this study, as an alternative strategy to reduce lignin recalcitrance, we developed a dominant approach that diverts precursors from the lignin pathway and enhances production of C₆C₁ aromatics that are known as non-conventional lignin monomers after export to the apoplast. Compared with regular C₆C₃ monolignols, these C₆C₁ monomers have reduced polymerization properties as they lack propanoid side-chain and its conjugated double bond, disabling them from undergoing condensation at their β-position. Such C₆C₁ aromatics are usually found in trace amounts in most lignins and frequently form lignin end-groups (Kim and Ralph, 2010; Kim *et al.*, 2003; Morreel *et al.*, 2004; Ralph *et al.*, 1997, 2001, 2008). This dominant approach is designed to increase the amount of C₆C₁ monomers in lignin to reduce its DP and to be easily transferable from model plant species to crops as it does not require any particular genetic background. For this purpose, we expressed a hydroxycinnamoyl-CoA hydratase-lyase (HCHL; EC 4.2.2.101/EC 4.1.2.41) from *Pseudomonas fluorescens* in stems of Arabidopsis. HCHL is an enzyme that catalyses the hydration of the double bond of the lignin precursor thioesters: *p*-coumaroyl-CoA, caffeoyl-CoA and feruloyl-CoA, followed by a retroaldol cleavage reaction that produces the corresponding C₆C₁ hydroxybenzaldehydes and acetyl-CoA (Figure S1; Mitra *et al.*, 1999). The promoter of a secondary cell wall cellulose synthase gene (*CesA4* / *IRX5*) was used to restrict HCHL expression to lignifying tissues of the stem and to prevent depletion of hydroxycinnamoyl-CoAs in other tissues where they serve as precursors for hydroxycinnamate derivatives involved in plant defence and development (Buer *et al.*, 2010; Gou *et al.*, 2009; Naoumkina *et al.*, 2010). Such secondary cell wall-specific promoter was preferred to those of lignin biosynthetic genes that are known to be active in non-lignified tissues for the synthesis of phenylpropanoid-derived metabolites (Vogt, 2010). Similarly, constitutively active promoters such as pCaMV35S were not selected on the basis that high HCHL expression would be detrimental to the development and growth of non-lignified tissues (Merali *et al.*, 2012). We show that most of the plants expressing the HCHL gene under the control of the *IRX5* promoter display no significant changes in lignin content, plant development

or biomass yield. We also demonstrate that C₆C₁ monomers accumulate as end-groups in the lignin of HCHL transgenics, resulting in reduced lignin DP and cell walls less recalcitrant to enzymatic hydrolysis.

Results

Expression of HCHL in Arabidopsis stems

The tissue-specific activity of the *IRX5* promoter used to express HCHL was investigated using beta-glucuronidase (GUS) as a reporter gene (Figure S2). GUS activity was essentially detected in the xylem vessels of the stem. Prolonged incubation also revealed strong GUS activity in stem interfascicular fibres, and more moderate staining was observed in the vascular system of young seedlings, siliques, rosette and cauline leaves. No activity was detected in other organs or tissues except for the style and anthers. A codon-optimized DNA sequence encoding for HCHL from *P. fluorescens* was cloned downstream the *IRX5* promoter for preferential expression in lignifying tissues of Arabidopsis stems. The presence of HCHL transcripts in the main stem of five independent transformants was verified by RT-PCR in the T1 generation (Figure 1a). T2 plants homozygous for the *IRX5:HCHL* construct were used to analyse HCHL protein expression and activity. The presence of HCHL protein in stem extracts of the five selected lines was confirmed by Western blot analysis (Figure 1a), and HCHL activity in these lines ranged from 0.025 to 0.16 pkat vanillin/μg protein using feruloyl-CoA as substrate, whereas no activity was detected in wild-type plants (Table 1). Based on these results, four transgenic lines were selected for detailed analysis: two lines exhibiting both extremes of HCHL activity and two others with intermediate activity levels.

Growth characteristics and tissue anatomy of *IRX5:HCHL* lines

IRX5:HCHL plants had growth and development characteristics visually similar to those of the wild type from the early rosette stage and until senescence (Figure 1b). However, senesced mature stems from lines *IRX5:HCHL*(4) and *IRX5:HCHL*(5) were slightly shorter (22% and 13% reduction) and had lower dry weight yield (30% and 16% reduction) compared with control plants, whereas those from lines *IRX5:HCHL*(1) and *IRX5:HCHL*(2) were not significantly different (Table 2). Stem tissues of *IRX5:HCHL* plants were inspected using light microscopy, and transverse cross-sections stained with lignin-specific reagents such as toluidine blue O, Mäule (S-units) and phloroglucinol-HCl (hydroxycinnamaldehyde units) showed similar patterns to those of wild-type plants except for the *IRX5:HCHL*(4) line in which occasional sections showed a few collapsed xylem structures (Figure 1c). Overall, these data could not demonstrate drastic changes in lignin content and composition in *IRX5:HCHL* plants, although the observation of abnormal xylem vessels for one of the transgenic lines suggested that some modifications of cell wall integrity could sometimes occur.

IRX5:HCHL lines show enrichment in cell wall-bound C₆C₁ aromatics

Analysis of methanol-soluble fractions extracted from stems of *IRX5:HCHL* plants showed higher amount of various C₆C₁ aromatics compared with wild type (Tables S1–S2). Therefore, presence of these aromatics was investigated in extract-free cell wall residues (CWR) obtained from stems. Mild alkaline hydrolysis

Figure 1 Comparative analysis of *IRX5:HCHL* lines. (a) HCHL expression in *IRX5:HCHL* lines. Top panel: Detection by RT-PCR of *HCHL* transcripts using stem mRNA from five independent 5-week-old T1 transformants. cDNA synthesized from stem mRNA of wild-type plants were used as a negative control. *Tub8*-specific primers were used to assess cDNA quality for each sample; Bottom panel: Detection by western blot of HCHL tagged with the AttB2 peptide (approximate size 32 kDa) using the 'universal antibody' and stem proteins from five independent 5-week-old *IRX5:HCHL* T2 transformants. A stem protein extract from wild type was used as a negative control, and a Ponceau staining of Rubisco large subunit (RBCL) is shown as a loading control. (b) Comparison of the growth and development of *IRX5:HCHL* and wild-type plants at different stages. Top panel: 3-week-old rosette; middle panel: 5-week-old flowering stage; Bottom panel: 8-week-old senescing stage. (c) Histochemical stainings of stem sections from 5-week-old Arabidopsis plants. Stem cross-sections of wild-type and *IRX5:HCHL* lines were analysed by Mäule (top panel), phloroglucinol-HCl (middle panel) and toluidine blue O stainings (bottom panel). i, interfascicular fibres; x, xylem. Bars represent 50 μm (top and middle panels) and 20 μm (bottom panel). Note the collapsed xylem vessels (yellow arrows) observed for line *IRX5:HCHL*(4).

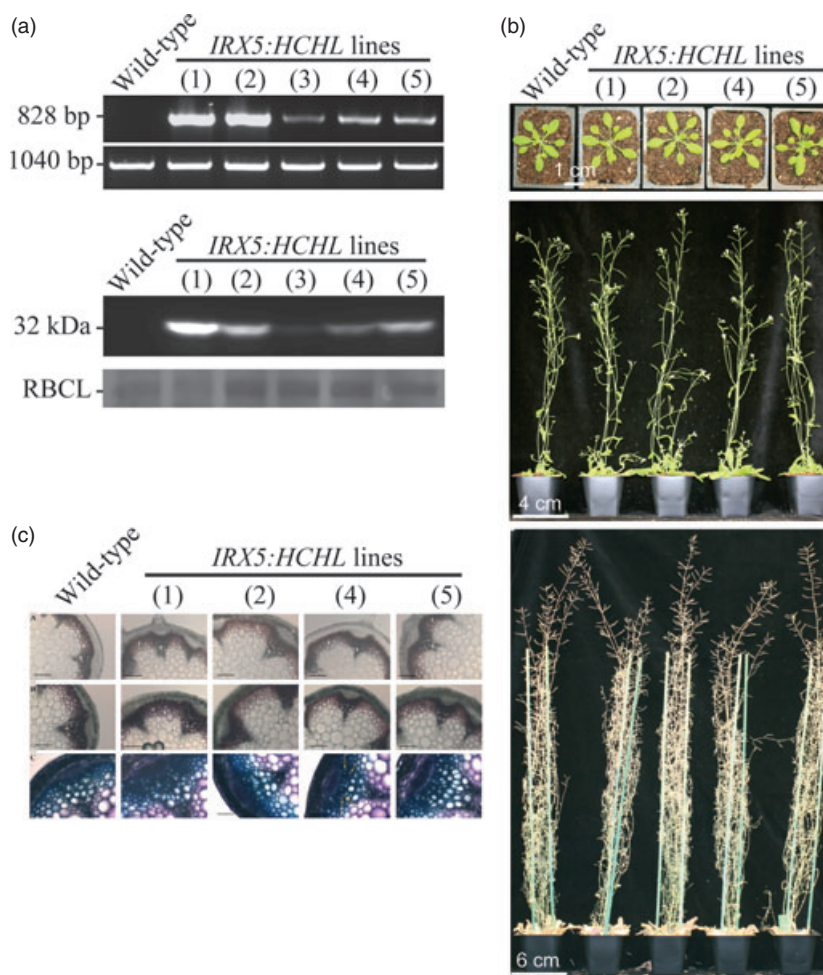


Table 1 *In vitro* HCHL enzyme activities in stems of 5-week-old wild-type (WT) and *IRX5:HCHL* plants. Values are means of three biological replicates ($n = 3$)

Plant line	Enzyme activity \pm SE (pkat vanillin/ μg protein)
WT	nd
<i>IRX5:HCHL</i> (1)	0.112 \pm 0.026
<i>IRX5:HCHL</i> (2)	0.075 \pm 0.022
<i>IRX5:HCHL</i> (3)	0.042 \pm 0.006
<i>IRX5:HCHL</i> (4)	0.160 \pm 0.038
<i>IRX5:HCHL</i> (5)	0.025 \pm 0.002

nd, not detected.

was performed on stem CWR to release loosely bound aromatics and revealed the presence of 4-hydroxybenzaldehyde (HBAlD), 3,4-dihydroxybenzaldehyde (3,4-HBAlD), vanillin (Van), 5-hydroxyvanillin (5OH-Van), syringaldehyde (SyrAlD), 4-hydroxybenzoic acid (HBA), vanillic acid (VA) and syringic acid (SyrA). 5OH-Van was only detected in *IRX5:HCHL* samples, and HBAlD, SyrAlD, HBA, VA and SyrA were increased by approximately 2-, 6-, 68-, 2- and 5-fold in cell walls of the different *IRX5:HCHL* lines compared with wild type (Table 3). These results indicate that larger amounts of C_6C_1 aromatics are

Table 2 Height and dry weight of the main inflorescence stem from senesced mature wild-type (WT) and *IRX5:HCHL* plants. n , number of plants analysed. Asterisks indicate significant differences from the wild-type using ANOVA followed by Scheffe *post hoc* test (* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$)

Plant line	Height (cm) Mean \pm SE	Dry weight (mg) Mean \pm SE	n
WT	62.4 \pm 4.6	477.7 \pm 51.3	16
<i>IRX5:HCHL</i> (1)	60.3 \pm 5.0	501.6 \pm 62.8	14
<i>IRX5:HCHL</i> (2)	56.0 \pm 4.6	435.3 \pm 62.5	12
<i>IRX5:HCHL</i> (4)	48.3 \pm 4.4***	335.7 \pm 63.4***	15
<i>IRX5:HCHL</i> (5)	54.1 \pm 7.6**	399.1 \pm 61.1*	16

loosely bound to cell walls in *IRX5:HCHL* plants. Moreover, similar amounts of ferulate and *p*-coumarate were quantified in cell walls of transgenics and wild type (data not shown).

Thioacidolysis reveals that lignin composition and structure are modified in *IRX5:HCHL* plants

Line *IRX5:HCHL*(2) was selected for further analyses as it showed no defective xylem structures and had biomass yields similar to wild-type plants. A preliminary screen performed on these plants using Fourier transform spectroscopy suggested no

Table 3 Quantitative analysis of cell wall-bound aromatics in stems from extract-free senesced mature wild-type (WT) and *IRX5:HCHL* plants. Values are means of four biological replicates ($n = 4$)

Plant line	Mean \pm SE ($\mu\text{g/g}$ dry weight)							
	HBAld	3,4-DHBAld	Van	5OH-Van	SyrAld	HBA	VA	SyrA
WT	5.8 \pm 0.6	1.1 \pm 0.0	59.4 \pm 6.5	nd	17.8 \pm 1.0	6.2 \pm 0.9	24.2 \pm 2.0	10.6 \pm 0.3
<i>IRX5:HCHL</i> (1)	11.1 \pm 0.4	0.6 \pm 0.0	36.9 \pm 2.7	0.8 \pm 0.1	107.8 \pm 6.4	486.4 \pm 28.2	42.2 \pm 3.2	47.5 \pm 2.3
<i>IRX5:HCHL</i> (2)	8.9 \pm 0.4	0.6 \pm 0.0	25.7 \pm 5.9	0.6 \pm 0.1	99.9 \pm 4.6	427.9 \pm 49.3	39.6 \pm 1.9	43.3 \pm 0.9
<i>IRX5:HCHL</i> (4)	9.1 \pm 0.9	0.7 \pm 0.0	29.9 \pm 2.7	0.8 \pm 0.1	122.2 \pm 14.8	421.8 \pm 28.2	36.8 \pm 1.4	54.1 \pm 6.1
<i>IRX5:HCHL</i> (5)	9.1 \pm 0.7	0.7 \pm 0.0	45.6 \pm 6.2	0.7 \pm 0.0	122.4 \pm 5.9	349.6 \pm 27.6	47.7 \pm 3.0	59.3 \pm 3.1

HBAld, 4-hydroxybenzaldehyde; 3,4-DHBAld, 3,4-dihydroxybenzaldehyde; Van, vanillin; 5OH-Van, 5-hydroxyvanillin; SyrAld, syringaldehyde; HBA, 4-hydroxybenzoic acid; VA, vanillic acid; SyrA, syringic acid; nd, not detected.

Table 4 Lignin content and main H, G and S lignin-derived monomers obtained by thioacidolysis of extract-free senesced mature stems from wild-type (WT) and *IRX5:HCHL*(2) plants. Acido-soluble lignin (ASL) and Klason lignin (KL) values are expressed as a relative percentage of the cell wall residue (CWR). Values are means \pm SE from duplicate analyses ($n = 2$)

Plant line		ASL% of CWR	Klason lignin KL % of CWR	Total yield (H+G+S) $\mu\text{mol/g}$ KL	% H	% G	% S	S/G
	<i>IRX5:HCHL</i>	1.32 \pm 0.05	20.12 \pm 0.15	1014 \pm 5	1.48 \pm 0.04	73.7 \pm 0.5	25.2 \pm 0.3	0.34 \pm 0.01
Culture 2	WT	1.50 \pm 0.15	20.32 \pm 0.25	1238 \pm 13	1.09 \pm 0.00	73.8 \pm 0.3	25.2 \pm 0.3	0.34 \pm 0.01
	<i>IRX5:HCHL</i>	1.25 \pm 0.07	21.29 \pm 0.14	1041 \pm 7	1.47 \pm 0.00	72.7 \pm 0.1	25.9 \pm 0.1	0.36 \pm 0.00

Table 5 Minor monomers obtained by thioacidolysis of extract-free senesced mature stems from WT and *IRX5:HCHL*(2) plants. Values are means \pm SE of duplicate analyses ($n = 2$). Values are expressed in $\mu\text{mol/g}$ KL and as a relative percentage of the total main H, G and S monomers released by thioacidolysis

Plant line		Vanalc $\mu\text{mol/g}$ KL	Syralc $\mu\text{mol/g}$ KL	Van $\mu\text{mol/g}$ KL	Syrald $\mu\text{mol/g}$ KL	Cald $\mu\text{mol/g}$ KL	VA $\mu\text{mol/g}$ KL	SyrA $\mu\text{mol/g}$ KL
		(%H+G+S)	(%H+G+S)	(%H+G+S)	(%H+G+S)	(%H+G+S)	(%H+G+S)	(%H+G+S)
Culture 1	WT	nd	nd	4.3 \pm 1 (0.31)	0.9 \pm 0.3 (0.06)	7.2 \pm 0.6 (0.53)	6.7 \pm 0.2 (0.49)	1.4 \pm 0.0 (0.10)
	<i>IRX5:HCHL</i>	5.0 \pm 0.1 (0.49)	2.6 \pm 0.2 (0.25)	6.5 \pm 1.4 (0.64)	18.7 \pm 3.5 (1.84)	7.9 \pm 0.3 (0.77)	6.8 \pm 0.2 (0.67)	2.2 \pm 0.0 (0.21)
Culture 2	WT	nd	nd	4.6 \pm 0.7 (0.37)	0.8 \pm 0.3 (0.06)	6.9 \pm 0.1 (0.55)	6.2 \pm 0.2 (0.50)	1.2 \pm 0.0 (0.09)
	<i>IRX5:HCHL</i>	5.3 \pm 0.1 (0.50)	2.9 \pm 0.1 (0.28)	6.3 \pm 0.7 (0.60)	16.7 \pm 1.9 (1.60)	6.8 \pm 0.1 (0.66)	7.0 \pm 0.0 (0.65)	2.1 \pm 0.0 (0.20)

Vanalc, vanillyl alcohol; Syralc, syringyl alcohol; Van, vanillin; SyrAld, syringaldehyde; Cald, coniferaldehyde; VA, vanillic acid; SyrA, syringic acid; nd, not detected.

difference in lignin content but significant changes in lignin structure (Figure S3). Consequently, lignin content and structure in stems from wild-type and *IRX5:HCHL*(2) plants were determined on CWR. Analysis of stem cell wall from two independent cultures revealed that Klason lignin (KL) and acido-soluble lignin (ASL) were identical in both genotypes, accounting for about 20% and 1.2%–1.5% of the CWR, respectively (Table 4). Lignin structure was evaluated by thioacidolysis, a chemical degradative method that specifically generates diagnostic thioethylated monomers from lignin units involved only in labile β -O-4-bonds, but also from cleaved ester-linked monomers. The total yield of conventional H, G and S monomers released from lignin after thioacidolysis was reduced by 25% and 16% in the two cultures of *IRX5:HCHL* plants compared with the wild type (Table 4). These lower yields indicate higher frequency of thioacidolysis-resistant bonds in the lignin of transgenics and/or a reduced amount of conventional monomers. In contrast, the

relative amount of G and S thioacidolysis monomers recovered from lignin was unchanged; both wild-type and transgenic samples showing an S/G ratio ranging between 0.34 and 0.36. However, the molar frequency of H monomers was significantly higher in *IRX5:HCHL* plants (Table 4).

In addition to the lignin-derived conventional monomers, thioacidolysis released more unconventional C_6C_1 units such as vanillin (Van) and syringaldehyde (Syrald) recovered as their dithioether derivatives, as well as vanillate (VA) and syringate (SyrA). Syrald and SyrA products released by thioacidolysis showed on average a 20.8- and 1.65-fold increase in *IRX5:HCHL* plants compared with wild-type plants (Table 5). In contrast, the contents of coniferaldehyde (Cald), Van and VA were not significantly different between the wild-type and transgenic samples (Table 5). Interestingly, two new lignin C_6C_1 units were released and identified as the thioethylated derivatives of vanillyl alcohol (Vanalc) and syringyl alcohol (Syralc) in the transgenic plants

(Table 5). The bonding mode of the Vanalc and Syralc parent structures that are cleaved during thioacidolysis is most likely as 4-O-ethers. Moreover, when expressed in the same units (e.g. $\mu\text{g/g}$ CWR), the amounts of Van, VA, Syrald and SyrA products released by thioacidolysis substantially exceed those obtained by mild alkaline hydrolysis (Table 3). Overall, these thioacidolysis data suggest that the lignin from *IRX5:HCHL* plants compared with wild type is enriched in resistant inter-unit bonds and in C_6C_1 monomers linked with lignin only by an ether bond at their C4-OH.

Analysis of lignin modifications by NMR

NMR (2D ^{13}C - ^1H -correlated, HSQC) spectra of whole cell wall material from *IRX5:HCHL(2)* stems had several new correlations compared with wild type, indicating the presence of 'new' components in the cell wall that are even more readily seen in the difference spectrum. These included a new oxidized S-unit (SA), an oxidized H-unit (HA) and correlations resembling those of an oxidized G-unit (Figure 2). However, upon pretreatment of the CWR with crude cellulases to remove most of the polysaccharides and enrich the lignin fraction in a so-called cellulolytic enzyme

lignin (CEL), the guaiacyl-type correlations disappeared from the spectra (Figure 2d), and therefore, the identification of this component has not been pursued. The chemical shifts of the SA component are consistent with either syringate esters or syringaldehyde groups. Although the SA correlations resemble those of benzylic-oxidized syringyl β -ether units that are often observed in dicot spectra (Ralph and Landucci, 2010), they cannot be assigned to such structures as the diagnostic β -C/H correlation (at $\sim 5.10/82.8$ ppm) is not present. In absolute agreement with the above-mentioned thioacidolysis observation of a 20.8-fold increase in syringaldehyde-derived units from the *IRX5:HCHL* plants, an enhanced aldehyde C/H correlation is also present with approximately the correct volume integral compared with that from the S2/6-correlations (Figure 2). Additionally, because of the overlap between syringate ester and syringaldehyde correlations, NMR analysis cannot rule out the presence of smaller fractions of syringate esters in the lignin of transgenics. The HA component's chemical shifts suggest that they are esters or, less likely, free acids, but do not match those of typical side-chain γ -*p*-hydroxybenzoates found in *Salix*, *Populus* or *Palmae* (Ralph and Landucci, 2010). Moreover, spec-

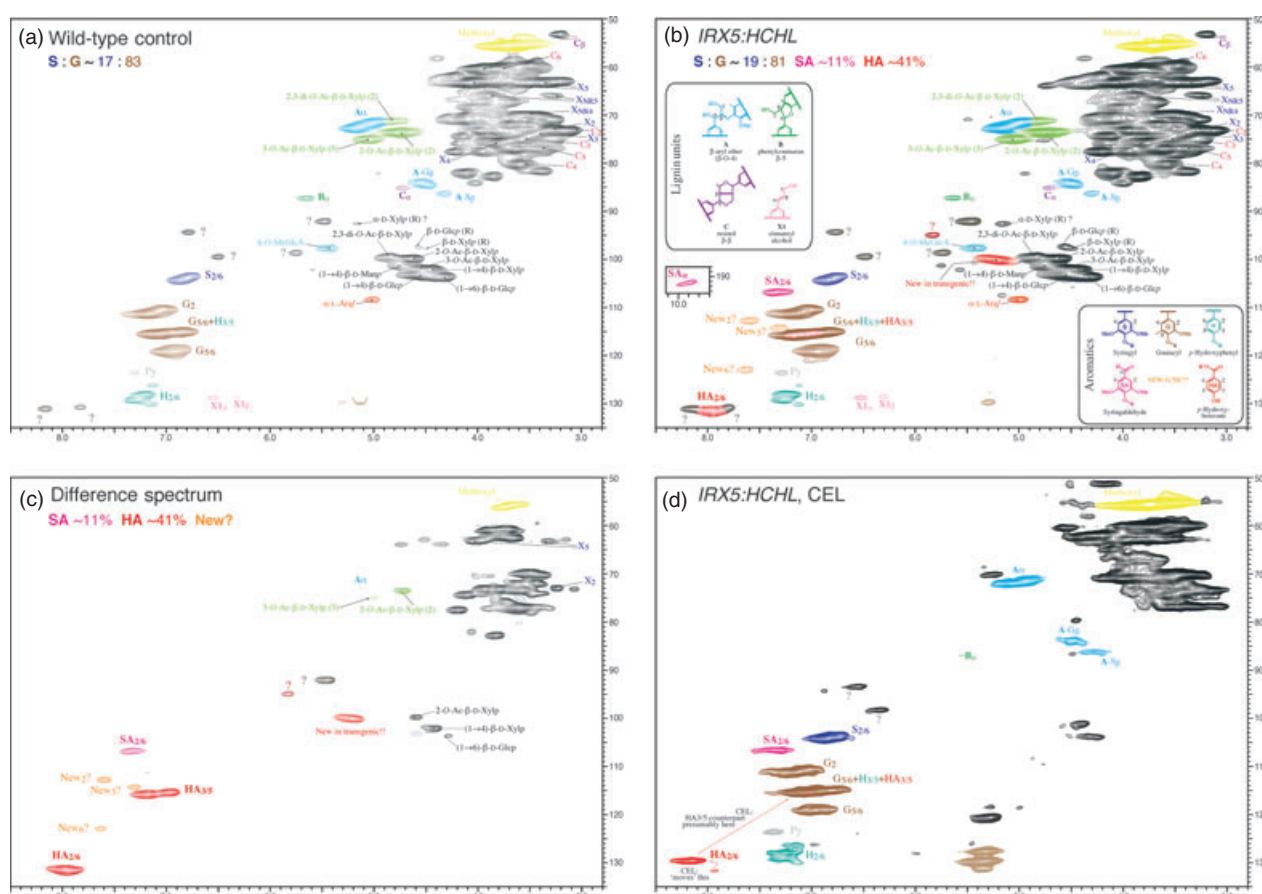


Figure 2 Partial 2D HSQC NMR spectra (aromatic and aliphatic regions) of cell wall material from wild-type and *IRX5:HCHL(2)* plants. (a) Spectrum from whole cell wall material of wild type showing normal S-G lignin with some H components, along with assigned polysaccharides and, in particular, the acetylated xylans. (b) Spectrum from whole cell wall material of *IRX5:HCHL(2)* showing the new SA and HA components. Inset: Aldehyde region attributable to the SA component. Shown in orange is an apparent oxidized G component. (c) Difference spectrum between transgenic and wild type by nulling the S2/6 correlation. (d) Spectrum of cellulolytic enzyme lignin (CEL) from *IRX5:HCHL(2)* confirming the association of the new HA and SA components with lignin. The apparent oxidized G component (orange peaks in b and c) is no longer detected and considered not to be associated with lignin. Similarly, migration of the HA2/6 correlation compared with that observed in b and c suggests association of this component with polysaccharides.

tra after cellulase treatment revealed that the HA unit converted into another species, which possibly assigns it as a glucoside ester (Figure 2d). Acetylation of the wall produces complex shifts that are predictable for free phenolic groups because the C/H-3/5 positions move significantly. Therefore, if HA units are linked with lignin, they must be esters and not free acids, but what monolignols they acylate remains to be determined. Interestingly, the core lignin relative S/G levels do not change significantly—when the S2/6 correlation is nulled in the difference spectrum, the G2 correlation is also essentially nulled. This agrees with the thioacidolysis data described earlier, although the S/G values calculated from NMR were 0.21 for the wild-type and 0.23 for the *IRX5:HCHL* plants (vs. 0.34–0.36 from thioacidolysis); such a difference is explained by the fact that NMR measures S/G on the entire lignin, whereas thioacidolysis only release a fraction of the lignin monomers.

Lignins from *IRX5:HCHL* plants have a lower degree of polymerization

Cellulolytic enzyme lignin fractions were isolated from wild-type and *IRX5:HCHL(2)* stems with similar yields and used to determine their polydispersity using size-exclusion chromatography (SEC). Elution profiles acquired by monitoring UV-A absorbance (SEC UV-A₃₀₀) and UV-F fluorescence (SEC UV-F_{ex250/em450}) of the dissolved CEL revealed differences between wild-type and *IRX5:HCHL* plants (Figure 3a–b). First, the total area of the largest mass peak detected between 7 and 13.5 min was significantly reduced in transgenics owing to an important diminution of the largest lignin fragments that elute between 7 and 9 min. Similarly, smaller molecular mass material that elutes later in a second peak between 13.5 and 19.5 min was more abundant (increased by 27% and 16% using UV-A and UV-F detections respectively) in *IRX5:HCHL* samples (Figure 3a–b). Finally, the area corresponding to the smallest lignin fragments detected between 19.5 and 26.5 min using UV-F is increased by 55% in transgenics (Figure 3b). These results demonstrate that the DP of lignins purified from *IRX5:HCHL* plants is reduced compared with wild type.

IRX5:HCHL lines have similar sugar content and improved saccharification efficiency

Monosaccharide composition was determined in stems after hydrolysis of total cell wall polysaccharides. Although both genotypes released similar amounts of total monosaccharides, which is in agreement with the FT-Raman analysis (Figure S3a), transgenics showed a reduction in glucose (–12%) and increases in xylose (+22%) and arabinose (+16%) compared with wild type (Table 6). Moreover, hemicellulosic monosaccharides released from CWR using trifluoroacetic acid were 23% higher in *IRX5:HCHL* stems, which corresponds to higher xylose (+23%) and arabinose (+22%) contents compared with wild type (Table 6).

To examine the impact of lignin DP reduction on cell wall digestibility, we conducted saccharification assays on biomass derived from stems pretreated with hot water, dilute alkaline and dilute acid solutions. After a 72-h incubation with cellulase and glucosidase, pretreated biomass of *IRX5:HCHL* plants released more sugars compared with wild type, showing improvements of saccharification efficiency ranging between 34% and 77% after hot water, 43% and 71% after dilute alkaline, and 15% and 31% after dilute acid pretreatments (Figure 4). Considering the similar total sugar contents in stems

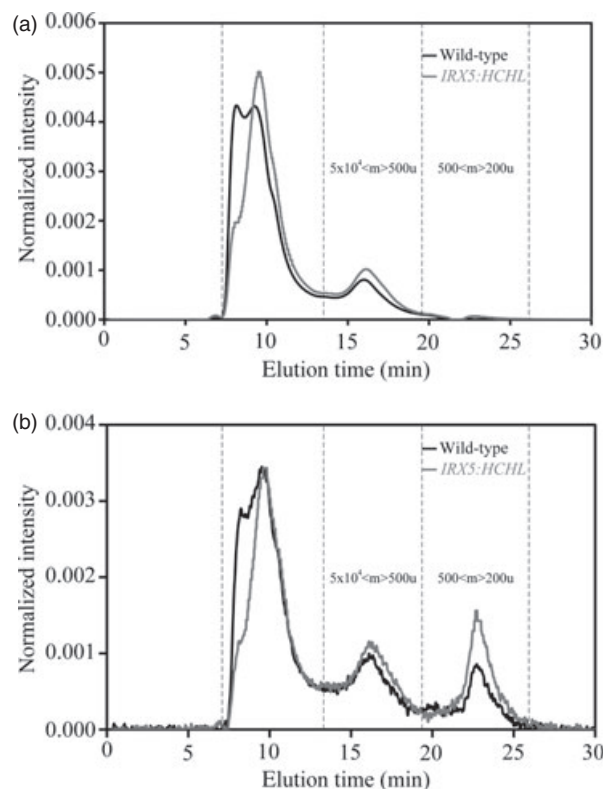


Figure 3 Polydispersity of cellulolytic enzyme lignin (CEL) from wild-type and *IRX5:HCHL(2)* plants. CEL lignins were purified from stems of wild-type (black line) and *IRX5:HCHL(2)* plants (grey line) and analysed for size polydispersity by size-exclusion chromatography (SEC). SEC chromatograms were obtained using (a) UV absorbance (λ_{300}) and (b) UV fluorescence (Ex_{250}/Em_{450}).

Table 6 Chemical composition of total and hemicellulosic cell wall sugars in senesced mature stems from wild-type (WT) and *IRX5:HCHL(2)* plants. Values are means \pm SE of three biological replicates ($n = 3$). Asterisks indicate significant differences from the WT ($P < 0.001$)

Sugar	Mean \pm SE (mg/g cell wall residues)			
	Total sugars		Hemicellulosic sugars	
	WT	<i>IRX5:HCHL</i>	WT	<i>IRX5:HCHL</i>
Fucose	2.2 \pm 0.1	2.2 \pm 0.1	1.4 \pm 0.0	1.5 \pm 0.1
Rhamnose	10.9 \pm 0.4	11.1 \pm 0.2	9.1 \pm 0.2	9.8 \pm 0.3
Arabinose	16.0 \pm 0.6	18.6 \pm 0.5*	10.1 \pm 0.3	12.4 \pm 0.4*
Galactose	23.0 \pm 0.7	22.7 \pm 0.8	15.3 \pm 0.3	16.5 \pm 0.5
Glucose	442.8 \pm 7.1	388.7 \pm 7.6*	10.1 \pm 0.3	11.2 \pm 0.3
Xylose	201.6 \pm 1.7	245.2 \pm 3.3*	114.4 \pm 1.0	141.2 \pm 4.2*
Galacturonic acid	93.7 \pm 2.6	99.9 \pm 1.5	37.1 \pm 1.9	40.6 \pm 1.1
Glucuronic acid	4.1 \pm 0.2	4.6 \pm 0.4	2.7 \pm 0.2	3.1 \pm 0.1
Total	794.3 \pm 13.4	793.0 \pm 14.4	200.2 \pm 4.2	236.3 \pm 7.0*

of *IRX5:HCHL* and wild-type plants, these data demonstrate that the cell wall of the transgenics is less recalcitrant to enzymatic hydrolysis.

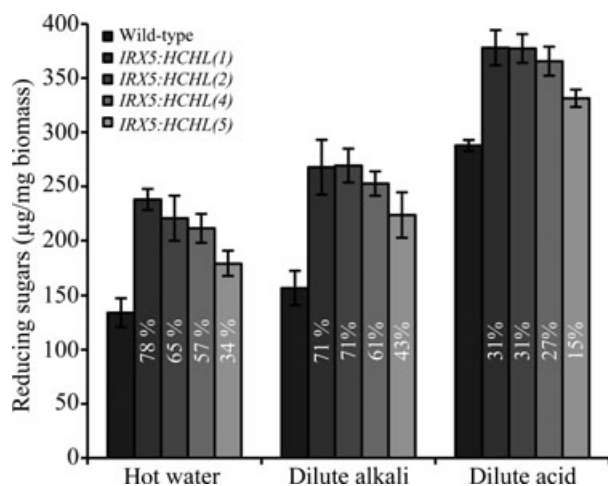


Figure 4 Saccharification of biomass from mature senesced stems of *IRX5:HCHL* and wild-type plants after various pretreatments. Amounts of sugars released from biomass after a 72-h enzymatic digestion are shown. Values are means \pm SE of four biological replicates ($n = 4$).

Discussion

Expression of HCHL in plants has originally been considered for *in planta* production of valuable and easily extractable compounds such as vanillin (Van) and *p*-hydroxybenzoic acid (HBA). However, owing to strong ectopic HCHL expression, adverse phenotypes such as chlorotic and senescing leaves, stunting, low pollen production, male sterility, collapsed xylem vessels and reduction in biomass were observed in transgenic tobacco and sugarcane (Mayer *et al.*, 2001; McQualter *et al.*, 2005; Merali *et al.*, 2007). In the present study, HCHL expression was restricted to the lignifying tissues using a secondary cell wall-specific promoter because our goal was to reduce the lignin DP and the risk of undesirable effects and also to enhance translation successes of this approach from Arabidopsis to various crops as the general secondary cell wall regulatory network and the lignin biosynthetic pathway are well conserved across vascular plants (Ruprecht *et al.*, 2010). Plants transformed with the *IRX5:HCHL* construct were neither dwarfed nor sterile and had unchanged lignin content in contrast to previous studies (Table 4; Mayer *et al.*, 2001; Merali *et al.*, 2007), and their young rosette leaves did not show reduced epidermal blue fluorescence under UV which is symptomatic of alteration in phenylpropanoid-derived sinapoyl-malate pools (data not shown; Bonawitz and Chapple, 2010; Vogt, 2010). The observation of unaffected lignin contents in the transgenics analysed in this work is in contrast with previous reports in tobacco that showed lignin diminution (Merali *et al.*, 2007). This difference could be explained by the fact that the non-tissue-specific, strong and constitutive promoter (pCaMV35S) was used to drive *HCHL* expression in all tissues in tobacco plants (including photosynthetic tissues), in contrast to the secondary cell wall cellulose synthase *IRX5* promoter which is active in tissues developing lignified secondary cell walls. *IRX5:HCHL* plants had similar total sugar content and a slight reduction in glucose/xylose ratio compared with wild type, which is in agreement with observations in tobacco plants expressing constitutively HCHL (Merali *et al.*, 2007). Although two *IRX5:HCHL* lines showed reduced

biomass, which for one of them might be caused a reduction in cell wall integrity creating occasional collapsed xylem vessels, other *IRX5:HCHL* lines were comparable with wild-type plants.

As expected, the transgenic lines showed increased amounts of soluble C_6C_1 aldehydes (HBAld, 3,4-DHBAld and Van) that are produced upon HCHL activity after cleavage of hydroxybenzoyl-CoA, 3,4-dihydroxybenzoyl-CoA and feruloyl-CoA (Tables S1–S2). HCHL has no activity against sinapoyl-CoA (Mitra *et al.*, 1999), suggesting that Syrald accumulated in transgenics is a conversion product of Van after successive hydroxylation and methylation at the C3-position of the phenyl ring, which is supported by the identification of the new intermediate 5OH-Van (Tables 3 and S2). Interestingly, several genes encoding monooxygenases were up-regulated in plants expressing HCHL, but none of the known or predicted *O*-methyltransferase showed altered expression (Table S3). Analysis of soluble aromatics in transgenics also showed that C_6C_1 aldehydes are oxidized to their respective C_6C_1 acid forms (Tables S1–S2). Notably, several genes encoding for dehydrogenases/reductases (SDR, AKR and ALDH) are up-regulated in plants expressing HCHL, with in particular, *AKR4C9* (*At3g37770*), which is known to metabolize a range of hydroxybenzaldehydes (Table S3; Simpson *et al.*, 2009). In addition, soluble C_6C_1 aromatics predominantly accumulate as conjugates in transgenics as we showed that glucose conjugates represent \sim 90% of the HBA soluble pool (Tables S1–S2), presumably for vacuolar storage as previously described for other C_6C_1 aromatics (Eudes *et al.*, 2008). Such accumulation of C_6C_1 acid glucosides is in agreement with previous studies performed in tobacco, sugar beet, datura and sugar cane plants over-expressing HCHL (Mayer *et al.*, 2001; McQualter *et al.*, 2005; Mitra *et al.*, 2002). Interestingly, transcript profiling of HCHL plants revealed seven up-regulated genes of the UDP-glucosyltransferase (UGT) family; among them, UGT75B1 and UGT73B4 were previously shown to catalyse glucose conjugation of benzoates (Table S3; Eudes *et al.*, 2008; Lim *et al.*, 2002).

Furthermore, we showed that some C_6C_1 aromatics are released from extract-free cell wall fractions of senesced stems upon mild alkaline hydrolysis. Higher amounts of HBAld, 5OH-Van, Syrald, HBA, VA and SyrA were measured in the 'loosely wall-bound' fraction of *IRX5:HCHL* lines compared with wild type, which is in accordance with observations in tobacco plants expressing constitutively HCHL (Merali *et al.*, 2007). Although the type(s) of linkages involved is unclear, loosely attached C_6C_1 aromatics have also been shown to be present cell walls of Arabidopsis leaves and roots (Forcat *et al.*, 2010; Tan *et al.*, 2004).

The analysis of lignin from plants expressing HCHL showed increased content of C_6C_1 monomers upon thioacidolysis, and the characterization of its monomeric composition identified two novel units (Vanalc and Syralc) in addition to higher amounts of syringaldehyde and syringate (Table 5). These data suggest that a part of the C_6C_1 aldehydes is converted into C_6C_1 alcohols, although the latest were not detected in the soluble and 'loosely wall-bound' fractions from *IRX5:HCHL* lines (data not shown). Thioacidolysis cleaves almost exclusively 4-O- β -ether bonds, implying that these C_6C_1 monomers are most likely 4-O-linked with lignins as end-groups. However, we cannot exclude that part of the C_6C_1 acids is esterified to lignins because thioacidolysis also cleaves ester bonds. In agreement with this data, NMR analysis identified higher amount of etherified syringaldehyde

units in lignins of transgenics, and revealed the presence of new structures resembling free phenolic hydroxybenzoate esters (Figure 2).

It is unclear whether C₆C₁ aromatics represent more nucleation sites ('starting points') for lignification in the walls creating more polymer chains of lower molecular weight or whether they actually incorporate growing lignin chains. In both cases, these side-chain-truncated monomers will be relegated to end-units and decrease the average lignin DP. Indeed, owing to the absence of phenylpropanoid tail, these new monomers can only initiate a new lignin chain leading to unidirectional polymerization starting from the hydroxyphenyl ring (at its 4-O- or non-methoxylated 5-positions) to which conventional monolignols couple via 5-β- or 4-O-β-coupling or terminate a growing lignin chain by 4-O-5-coupling or, in the case of the guaiacyl monomers only, 5-5- or 5-O-4-coupling.

We postulated that the lignin DP would be reduced upon enrichment in the wall of unconventional side-chain-truncated monomers that would increase the number of nucleation sites for lignification and allow only unidirectional polymerization. Analysis of the polydispersity of lignin in plants overproducing these monomers showed significant reduction in high-molecular mass oligomers and significant increase in low-molecular mass oligomers, supporting our hypothesis (Figure 3). These observations are relevant for understanding the higher susceptibility of the biomass from HCHL lines to polysaccharide enzymatic hydrolysis (Figure 4). Although saccharification efficiency of biomass is determined by several characteristics of cell walls, the improvements observed in the *IRX5:HCHL* lines after different pretreatments suggest that plants with less ramified or shorter lignin polymers would have their cell wall polysaccharides less cross-linked or less embedded in the lignin matrix, thus improving the accessibility of cellulose and hemicelluloses to hydrolytic enzymes.

We conclude that the overproduction of lignin 'DP reducers' can be used *in planta* to modify the lignin structure and reduce lignin DP to reduce biomass recalcitrance and enhance its conversion efficiency into fermentable sugars. As this approach is independent of any particular plant genetic background, it should be transferable to various energy crops and compatible with several traits. Restricting the biosynthesis of these lignin 'DP reducers' in supporting lignified tissues (i.e. sclerenchyma fibres), as well as avoiding strong production in conductive tissues (i.e. vessels), should limit the risk of adverse effects on plant development and biomass yield.

Experimental procedures

Plant material and growth conditions

Arabidopsis thaliana (ecotype Columbia, Col-0) seeds were germinated directly on soil. Growing conditions were 14 h of light per day at 100 μmol/m²/s, 22 °C and 55% humidity. Selection of T1 and T2 homozygote transgenic plants was made on Murashige and Skoog vitamin medium (PhytoTechnology Laboratories, Shawnee Mission, KS) supplemented with 1% sucrose, 1.5% agar and containing 50 μg/mL kanamycin.

Generation of a pTKan-*pIRX5:HCHL* plasmid and plant transformation

A Gateway cloning cassette (Life Technologies, Grand Island, NY) was inserted between *Xho*I and *Pst*I restriction sites of the

binary vector pTKan⁺ (Schaaf *et al.*, 2006) to produce a pTKan-GW vector. The nucleotide sequence of the *IRX5* promoter was amplified by PCR from *Arabidopsis* (ecotype Columbia, Col-0) genomic DNA using oligonucleotides 5'-CCCGGCGGCATG AAGCCATCCTCTACCTCGGAAA-3' and 5'-CCCGGCTAGCGGC GAGGTACTGAGCTCTCGGAA-3' (*Not*I and *Nhe*I restriction sites underlined) and inserted between the *Ap*I and *Spe*I restriction sites of pTKan-GW to produce a pTKan-*pIRX5*-GW expression vector. A *HCHL* codon-optimized nucleotide sequence from *P. fluorescens* AN103 (GenBank accession number CAA73502) was synthesized without stop codon (GenScript, Piscataway, NJ) and cloned into the Gateway pDONR221-f1 entry vector (Lalonde *et al.*, 2010). A sequence-verified *HCHL* entry clone was LR recombined with the pTKan-*pIRX5*-GW vector to generate the pTKan-*pIRX5:HCHL* construct that was introduced into wild-type *Arabidopsis* plants (ecotype Col0) via *Agrobacterium tumefaciens*-mediated transformation (Bechtold and Pelletier, 1998).

RNA extraction and RT-PCR

Total RNA (1 μg) was extracted using the Plant RNeasy extraction kit (Qiagen, Valencia, CA) and reverse-transcribed using the Transcriptor First Strand cDNA Synthesis Kit (Roche Applied Science, Indianapolis, IN). The obtained cDNA preparation was quality controlled for PCR using *tub8*-specific oligonucleotides (5'-GGGCTAAAGGACACTACTAG-3' and 5'-CCTCTGCACTT CCACTTCGTCTTC-3') and used for the detection of *HCHL* expression using oligonucleotides 5'-ATGTCTACTTACGAGGGA AGATGG-3' and 5'-TCTCTTGTAAGCCTGGAGTCC-3'.

Western blot analysis

Proteins were extracted using a buffer containing 100 mM Tris-HCl pH 6.5, 2% (w/v) polyvinylpyrrolidone, 2% (v/v) β-mercaptoethanol, 1% (w/v) SDS and quantified using the Bradford method (Bradford, 1976). Proteins (5 μg) were separated by SDS-PAGE, blotted and immunodetected using the universal antibody as described in Eudes *et al.* (2011).

HCHL activity

Proteins were extracted using a buffer (EB-100 mM Tris-HCl, pH 8.5, 20 mM DTT, 10 mM Na₂EDTA-) containing 25 mg of polyvinylpyrrolidone. Extracts were shaken at 1400 rpm for 15 min at 4 °C and centrifuged for 30 min at 20 000 *g* at 4 °C. Proteins were applied to PD10 columns (GE Healthcare, Piscataway, NJ), eluted with EB and quantified using the Bradford method. For HCHL activity, 5 μL of protein extract was incubated for 15 min at 30 °C with 150 μM feruloyl-CoA in 100 mM Tris-HCl pH 8.5 in a total volume of 50 μL. Total amounts of protein per reaction varied from 4 to 6.5 μg. Reactions were stopped with 50 μL of cold acidified methanol (12% glacial acetic acid / 88% methanol, v/v) and stored at -70 °C until LC-MS analysis.

Microscopy

Basal stem transverse sections (100-μm-thick) were obtained using a vibratome (Leica Microsystems Inc., Buffalo Grove, IL). For toluidine blue O (TBO) staining, sections were incubated in a 0.05% (w/v) solution of TBO (Sigma-Aldrich, St Louis, MO) for 30 s and rinsed with water. For Wiesner lignin staining, sections were incubated for 3 min in phloroglucinol-HCl reagent (VWR International, Brisbane, CA) and rinsed with water (Nakano *et al.*, 1992). For Mäule lignin staining, sections were incubated

in 4% KMnO₄ for 5 min, rinsed with water, incubated in 37% HCl for 2 min and observed after addition of a drop of aqueous ammonia. All sections were analysed using bright field light microscopy.

Cell wall-bound aromatics extraction

Senesced stems were ball-milled using a Mixer Mill MM 400 (Retsch Inc., Newtown, PA) and stainless steel balls for 2 min at 30 s. Extract-free CWR were obtained by sequentially washing 60 mg of ball-milled stems with 1 mL of 96% ethanol at 95 °C twice for 30 min, and vortexing with 1 mL of 70% ethanol twice for 30 s. The resulting CWR were dried *in vacuo* overnight at 30 °C. CWR (6 mg), were mixed with 500 µL of 2 M NaOH and shaken at 1400 rpm for 24 h at 30 °C. The mixture was acidified with 100 µL of concentrated HCl and subjected to three ethyl acetate partitioning steps. Ethyl acetate fractions were pooled, dried *in vacuo* and suspended in 50% (v/v) methanol–water prior to LC-MS analysis.

LC-MS

A Poroshell-120 column (150 mm length, 3 mm internal diameter, 2.7 µm particle size) and a 1200 Series HPLC system (Agilent Technologies Inc., Santa Clara, CA) were used. C₆C₁ aromatics were separated using a gradient elution with mobile phase composition of 0.1% formic acid in water as solvent A and 0.1% formic acid in acetonitrile–water (98 : 2, v/v) as solvent B. The elution gradient was 0–5 min 13% B, 5–7 min 50% B, 7–8 min 50% B and 8–11 min 13% B using a flow rate of 0.55 mL min⁻¹. The HPLC system was coupled to an Agilent 6210 time-of-flight (TOF) mass spectrometer (MS) via a 1 : 7 post-column split. Analyses were conducted using Electrospray ionization (ESI) in the positive ion mode. Detection of [M + H]⁺ ions was carried out in full scan mode at 0.85 spectra/s and a cycle time of 1.176 s/spectrum using the following parameters: capillary voltage 3500 V, fragmentor 165 V, skimmer 50 V and OCT RF 170 V, drying gas flow rate 9 L min⁻¹, nebulizer pressure 15 psig and drying gas temperature 325 °C. Quantification of aromatics was made by comparison with standard curves of authentic compounds (Sigma-Aldrich; Alfa Aesar, Ward Hill, MA).

HPAEC-PAD monosaccharide analysis

Cell wall residues of ball-milled senesced stems were treated with sulphuric acid as previously described for total sugar hydrolysis prior to chromatography (Sluiter *et al.*, 2008). For hemicellulose hydrolysis, CWR of ball-milled senesced stems (5 mg) were incubated in 1 mL of 2 M trifluoroacetic acid (TFA) for 1 h at 120 °C. TFA was dried *in vacuo* and the residue suspended in deionized water prior to chromatography. Monosaccharide separation was performed on a PA20 column (Dionex, Sunnyvale, CA) at a flow rate of 0.5 mL min⁻¹. The column was equilibrated with 10 mM NaOH for 10 min. The elution programme consisted of a linear gradient from 10 mM NaOH to 5 mM NaOH from 0 to 1.5 min, followed by isocratic elution with 5 mM NaOH from 1.5 to 20 min and a linear gradient up to 800 mM NaOH from 20 to 43 min (Harholt *et al.*, 2006; ØBro *et al.*, 2004). Monosaccharides were detected using a pulsed amperometric detector (gold electrode) set on waveform A according to manufacturer's instructions. A calibration curve of monosaccharide standards (Sigma-Aldrich) was run for verification of response factors. Statistical analysis was performed using ANOVA followed by Tukey's test.

Lignin analysis

Cell wall residues of ball-milled senesced stems were prepared using a Soxhlet apparatus by sequentially extracting the ground material with toluene/ethanol (2 : 1, v/v), ethanol and water (Sluiter *et al.*, 2008). The determination of lignin content using the standard Klason procedure (Dence *et al.*, 1992) and the thioacidolysis procedure (Lapierre *et al.*, 1995, 1999) were carried out on CWR. The lignin-derived monomers were identified by GC-MS as their trimethyl-silylated derivatives. All the lignin analyses were performed in duplicate.

NMR sample preparation

The whole plant cell wall gel-state samples for NMR experiments were prepared as previously described by Kim *et al.* (2008). The dried Arabidopsis stems were preground for 1 min in a Retsch MM400 mixer mill at 30 Hz, using zirconium dioxide (ZrO₂) vessels (10 mL) containing ZrO₂ ball bearings (2 × 10 mm). The preground cell walls were extracted with distilled water (ultrasonication, 1 h, three times) and 80% ethanol (ultrasonication, 1 h, three times). Isolated cell walls were dried and ball-milled using a Fritsch Planetary Micro *pulverisette 7* planetary ball mill at 800 rpm, using ZrO₂ vessels (12 mL) containing ZrO₂ ball bearings (30 × 5 mm). Each sample (100 mg) was ground for 2 h (interval: 30 min, break: 5 min, 4 cycles, total run time 2.25 h). The cell walls were collected directly into the NMR tubes and gels formed using DMSO-d₆/pyridine-d₅ (4 : 1).

2D ¹³C–¹H HSQC NMR spectroscopy

NMR experiments for the whole plant cell wall gel-state samples were also performed as previously described (Kim and Ralph, 2010; Kim *et al.*, 2008). NMR spectra were acquired on a Bruker Biospin (Billerica, MA) Avance 500 MHz spectrometer fitted with a cryogenically cooled 5-mm TCI gradient probe with inverse geometry (proton coils closest to the sample). The central DMSO solvent peak was used as internal reference (δ_C 39.5, δ_H 2.49 ppm). The ¹³C–¹H correlation experiment was an adiabatic HSQC experiment (Bruker standard pulse sequence 'hsqcetgpsisp.2') phase-sensitive gradient-edited-2D HSQC using adiabatic pulses for inversion and refocusing (Kupce and Freeman, 2007). HSQC experiments were carried out using the following parameters: acquired from 10 to 0 ppm in F2 (¹H) with 1000 data points (acquisition time 100 ms), 200 to 0 ppm in F1 (¹³C) with 400 increments (F1 acquisition time 8 ms) of 72 scans with a 500 ms interscan delay; the d₂₄ delay was set to 0.89 ms (1/8J, J = 145 Hz). The total acquisition time was 5 h. Processing used typical matched Gaussian apodization (GB = 0.001, LB = -0.1) in F2 and squared cosine-bell and one level of linear prediction (32 coefficients) in F1. Volume integration (uncorrected) of contours in HSQC plots used Bruker's TopSpin 3 (Mac version) software.

Isolation of cellulolytic enzyme lignin

One gram of ball-milled mature senesced stems was mixed with 50 mM NaCl (30 mL) and incubated overnight at 4 °C. After centrifugation (2800 g, 10 min), the biomass was extracted sequentially by sonication (20 min) with 80% ethanol (three times), acetone (one time), chloroform-methanol (1 : 1, v/v, one time) and acetone (one time). The obtained CWR were ball-milled for 3 h per 500 mg of sample (in 10 min on/10 min off cycles) using a PM100 ball mill (Retsch) vibrating at 600 rpm

in zirconium dioxide vessels (50 mL) containing ZrO₂ ball bearings (10 × 10 mm). Ball-milled walls (490 mg for wild type and 480 mg for *IRX5:HCHL*) were digested four times over 3 days at 30 °C with crude cellulases (Cellulysin; 60 mg/g of sample; EMD Biosciences Inc, San Diego, CA) in NaOAc pH 5.0 buffer (30 mL). The obtained CEL was washed three times with deionized water and lyophilized overnight. CEL recovered were 131 mg for wild type (27.3%) and 101 mg for *IRX5:HCHL* (20.6%). For NMR, the CEL (30 mg) was swelled in DMSO-d₆/pyridine-d₅ as described earlier. For SEC analysis, 1% (w/v) CEL lignin solutions were prepared in 1-methyl-2-pyrrolidinone (NMP)-DMSO (1 : 1, v/v) after sonication for 3 h at 40 °C.

Size-exclusion chromatography

Polydispersity of dissolved CEL was determined using analytical techniques SEC UV-F and SEC UV-A as previously described (George et al., 2011). An Agilent 1200 series binary LC system (G1312B) equipped with FL (G1321A) and DA (G1315D) detectors was used. Separation was achieved with a Mixed-D column [5 mm particle size, 300 mm × 7.5 mm i.d., linear molecular weight range of 200–400 000 g/mol (Agilent Technologies Inc.)] at 80 °C using a mobile phase of NMP at a flow rate of 0.5 mL min⁻¹. Absorbance of material eluting from the column was detected at 300 nm (UV-A). Excitation 250 nm and emission 450 nm were used for UV-F detection. Intensities were area normalized, and molecular mass estimates were determined after calibration of the system with polystyrene standards.

Cell wall pretreatments and saccharification

Ball-milled senesced stems (10 mg) were mixed with 340 µL of water, 340 µL of H₂SO₄ (1.2%, w/v) or 340 µL of NaOH (0.25%, w/v) for hot water, dilute acid or dilute alkaline pretreatments, respectively, incubated at 30 °C for 30 min and autoclaved at 120 °C for 1 h. After cooling down at room temperature, samples pretreated with dilute acid and dilute alkaline solutions were neutralized with 5 N NaOH (25 µL) and 1.25 N HCl (25 µL), respectively. Saccharification was initiated by adding 635 µL of 100 mM sodium citrate buffer pH 6.2 containing 80 µg/mL tetracycline, 5% w/w cellulase complex NS50013 and 0.5% w/w glucosidase NS50010 (Novozymes, Bagsværd, Denmark). After 72 h of incubation at 50 °C with shaking (800 rpm), samples were centrifuged (20 000 g, 3 min), and 10 µL of the supernatant was collected for reducing sugar measurement using the DNS assay and glucose solutions as standards (Miller, 1959).

Acknowledgements

The authors are grateful to Carsten Rautengarten for providing feruloyl-CoA, to Laurent Cézard and Frédéric Legée for thioacidolysis and Klason analysis, to Kazufumi Yazaki for providing 4-hydroxybenzoate glucose conjugates, to Andreia M. Smith for statistical analysis of the data, to Jorge Rencoret Pazo, Hoon Kim and Yuki Tobimatsu for help with the NMR analysis, and to Novozymes for providing cellulase and glucosidase. This work was part of the DOE Joint BioEnergy Institute (<http://www.jbei.org>) supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy. PM and JR were funded by the DOE Great

Lakes Bioenergy Research Center (DOE BER Office of Science DE-FC02-07ER64494).

Conflict of interest

J.D.K. has financial conflict of interest in Amyris, LS9 and Lygos.

References

- Baucher, M., Halpin, C., Petit-Conil, M. and Boerjan, W. (2003) Lignin: genetic engineering and impact on pulping. *Crit. Rev. Biochem. Mol. Biol.* **38**, 305–350.
- Bechtold, N. and Pelletier, G. (1998) In planta *Agrobacterium*-mediated transformation of adult *Arabidopsis thaliana* plants by vacuum infiltration. *Methods Mol. Biol.* **82**, 259–266.
- Boerjan, W., Ralph, J. and Baucher, M. (2003) Lignin biosynthesis. *Annu. Rev. Plant Biol.* **54**, 519–546.
- Bonawitz, N.D. and Chapple, C. (2010) The genetics of lignin biosynthesis: connecting genotype to phenotype. *Annu. Rev. Genet.* **337**–363.
- Boudet, A.-M. (2007) Evolution and current status of research in phenolic compounds. *Phytochemistry*, **68**, 2722–2735.
- Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248–254.
- Buer, C.S., Imin, N. and Djordjevic, M.A. (2010) Flavonoids: new roles for old molecules. *J. Integr. Plant Biol.* **52**, 98–111.
- Chen, F. and Dixon, R.A. (2007) Lignin modification improves fermentable sugar yields for biofuel production. *Nat. Biotechnol.* **25**, 759–761.
- Dence, C.W., Lin, S.Y. and Dence, C.W. (eds) (1992) *Lignin Determination*. Berlin: Springer-Verlag.
- Dien, B.S., Miller, D.J., Hector, R.E., Dixon, R.A., Chen, F., McCaslin, M., Reisen, P., Sarath, G. and Cotta, M.A. (2011) Enhancing alfalfa conversion efficiencies for sugar recovery and ethanol production by altering lignin composition. *Bioresour. Technol.* **102**, 6479–6486.
- Eudes, A., Bozzo, G.G., Waller, J.C., Naponelli, V., Lim, E.-K., Bowles, D.J., Gregory, J.F. and Hanson, A.D. (2008) Metabolism of the folate precursor p-aminobenzoate in plants: glucose ester formation and vacuolar storage. *J. Biol. Chem.* **283**, 15451–15459.
- Eudes, A., Baidoo, E.E.K., Yang, F., Burd, H., Hadi, M.Z., Collins, F.W., Keasling, J.D. and Loque, D. (2011) Production of tranilast [N-(3',4'-dimethoxycinnamoyl)-anthranilic acid] and its analogs in yeast *Saccharomyces cerevisiae*. *Appl. Microbiol. Biotechnol.* **89**, 989–1000.
- Forcat, S., Bennett, M., Grant, M. and Mansfield, J.W. (2010) Rapid linkage of indole carboxylic acid to the plant cell wall identified as a component of basal defence in *Arabidopsis* against hrp mutant bacteria. *Phytochemistry*, **71**, 870–876.
- Fu, C., Mielenz, J.R., Xiao, X., Ge, Y., Hamilton, C.Y., Rodriguez, M., Chen, F., Foston, M., Ragauskas, A., Bouton, J., Dixon, R.A. and Wang, Z.-Y. (2011) Genetic manipulation of lignin reduces recalcitrance and improves ethanol production from switchgrass. *Proc. Natl Acad. Sci. USA*, **108**, 3803–3808.
- Gallego-Giraldo, L., Escamilla-Trevino, L., Jackson, L.A. and Dixon, R.A. (2011a) Salicylic acid mediates the reduced growth of lignin down-regulated plants. *Proc. Natl Acad. Sci. USA*, **108**, 20814–20819.
- Gallego-Giraldo, L., Jikumaru, Y., Kamiya, Y., Tang, Y. and Dixon, R.A. (2011b) Selective lignin downregulation leads to constitutive defense response expression in alfalfa (*Medicago sativa* L.). *New Phytol.* **190**, 627–639.
- George, A., Tran, K., Morgan, T.J., Benke, P.I., Berruoco, C., Lorente, E., Wu, B.C., Keasling, J.D., Simmons, B.A. and Holmes, B.M. (2011) The effect of ionic liquid cation and anion combinations on the macromolecular structure of lignins. *Green Chem.* **13**, 3375.
- Gou, J.-Y., Yu, X.-H. and Liu, C.-J. (2009) A hydroxycinnamoyltransferase responsible for synthesizing suberin aromatics in *Arabidopsis*. *Proc. Natl Acad. Sci. USA*, **106**, 18855–18860.
- 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.

- Halpin, C. (2004) Investigating and manipulating lignin biosynthesis in the postgenomic era. *Adv. Bot. Res.* **41**, 63–106.
- Harholt, J., Jensen, J.K., Sørensen, S.O., Orfila, C., Pauly, M. and Scheller, H.V. (2006) Arabinan deficient 1 is a putative arabinosyltransferase involved in biosynthesis of pectic arabinan in Arabidopsis. *Plant Physiol.* **140**, 49–58.
- Jung, H.G. and Allen, M.S. (1995) Characteristics of plant cell walls affecting intake and digestibility of forages by ruminants. *J. Anim. Sci.* **73**, 2774–2790.
- Keasling, J.D. (2010) Manufacturing molecules through metabolic engineering. *Science*, **330**, 1355–1358.
- 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., Lu, F., Ralph, S., Boudet, A., MacKay, J., Sederoff, R., Ito, T., Kawai, S., Ohashi, H. and Higuchi, T. (2003) NMR analysis of lignins in CAD-deficient plants. Part 1. Incorporation of hydroxycinnamaldehydes and hydroxybenzaldehydes into lignins. *Org. Biomol. Chem.* **1**, 268–281.
- Kim, H., Ralph, J. and Akiyama, T. (2008) Solution-state 2D NMR of ball-milled plant cell wall gels in DMSO-d(6). *Bio. Res.* **1**, 56–66.
- Kupce, E. and Freeman, R. (2007) Compensated adiabatic inversion pulses: broadband INEPT and HSQC. *J. Magn. Reson.* **187**, 258–265.
- Lalonde, S., Sero, A., Pratelli, R., Pilot, G., Chen, J., Sardi, M.I., Parsa, S.A., Kim, D.-Y., Acharya, B.R., Stein, E.V., Hu, H.-C., Villiers, F., Takeda, K., Yang, Y., Han, Y.S., Schwacke, R., Chiang, W., Kato, N., Loque, D., Assmann, S.M., Kwak, J.M., Schroeder, J., Rhee, S.Y. and Frommer, W.B. (2010) Frontiers: a membrane protein/signaling protein interaction network for Arabidopsis version AMPv2. *Front. Physiol.* **1**, 24.
- Lapierre, C., Pollet, B. and Rolando, C. (1995) New insights into the molecular architecture of hardwood lignins by chemical degradative methods. *Res. Chem. Intermed.* **21**, 397–412.
- Lapierre, C., Pollet, B., Petit-Conil, M., Toval, G., Romero, J., Pilate, G., Leple, J., Boerjan, W., Ferret, V.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.
- 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., Fehrlé, 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, X., Weng, J.-K. and Chapple, C. (2008) Improvement of biomass through lignin modification. *Plant J.* **54**, 569–581.
- Li, X., Ximenes, E., Kim, Y., Slininger, M., Meilan, R., Ladisch, M. and Chapple, C. (2010) Lignin monomer composition affects Arabidopsis cell-wall degradability after liquid hot water pretreatment. *Biotechnol. Biofuels*, **3**, 27.
- Lim, E.-K., Doucet, C.J., Li, Y., Elias, L., Worrall, D., Spencer, S.P., Ross, J. and Bowles, D.J. (2002) The activity of Arabidopsis glycosyltransferases toward salicylic acid, 4-hydroxybenzoic acid, and other benzoates. *J. Biol. Chem.* **277**, 586–592.
- Mayer, M.J., Narbad, A., Parr, A.J., Parker, M.L., Walton, N.J., Mellon, F.A. and Michael, A.J. (2001) Rerouting the plant phenylpropanoid pathway by expression of a novel bacterial enoyl-CoA hydratase/lyase enzyme function. *Plant Cell*, **13**, 1669–1682.
- McQualter, R.B., Chong, B.F., Meyer, K., Van Dyk, D.E., O'Shea, M.G., Walton, N.J., Viitanen, P.V. and Brumbley, S.M. (2005) Initial evaluation of sugarcane as a production platform for p-hydroxybenzoic acid. *Plant Biotechnol. J.* **3**, 29–41.
- Merali, Z., Mayer, M.J., Parker, M.L., Michael, A.J., Smith, A.C. and Waldron, K.W. (2007) Metabolic diversion of the phenylpropanoid pathway causes cell wall and morphological changes in transgenic tobacco stems. *Planta*, **225**, 1165–1178.
- Merali, Z., Mayer, M.J., Parker, M.L., Michael, A.J., Smith, A.C. and Waldron, K.W. (2012) Expression of a bacterial, phenylpropanoid metabolizing enzyme in tobacco reveals essential roles of phenolic precursors in normal leaf development and growth. *Physiol. Plant.* PMID: 22276599.
- Miller, G. (1959) Use of dinitrosalicylic acid reagent for determination of reducing sugar. *Anal. Chem.* **31**, 426–428.
- Mitra, A., Kitamura, Y., Gasson, M.J., Narbad, A., Parr, A.J., Payne, J., Rhodes, M.J., Sewter, C. and Walton, N.J. (1999) 4-hydroxycinnamoyl-CoA hydratase/lyase (HCHL) – an enzyme of phenylpropanoid chain cleavage from Pseudomonas. *Arch. Biochem. Biophys.* **365**, 10–16.
- Mitra, A., Mayer, M.J., Mellon, F.A., Michael, A.J., Narbad, A., Parr, A.J., Waldron, K.W. and Walton, N.J. (2002) 4-Hydroxycinnamoyl-CoA hydratase/lyase, an enzyme of phenylpropanoid cleavage from Pseudomonas, causes formation of C(6)-C(1) acid and alcohol glucose conjugates when expressed in hairy roots of *Datura stramonium* L. *Planta*, **215**, 79–89.
- Morreel, K., Ralph, J., Lu, F., Goeminne, G., Busson, R., Herdewijn, P., Goeman, J.L., Van der Eycken, J., Boerjan, W. and Messens, E. (2004) Phenolic profiling of caffeic acid O-methyltransferase-deficient poplar reveals novel benzodioxane oligolignols. *Plant Physiol.* **136**, 4023–4036.
- Nakano, J., Meshitsuka, G., Lin, S. and Dence, C. (eds) (1992) *The Detection of Lignin Methods in Lignin Chemistry*. Berlin: Springer-Verlag.
- Naoumkina, M.A., Zhao, Q., Gallego-Giraldo, L., Dai, X., Zhao, P.X. and Dixon, R.A. (2010) Genome-wide analysis of phenylpropanoid defence pathways. *Mol. Plant Path.* **11**, 829–846.
- ØBro, J., Harholt, J., Scheller, H.V. and Orfila, C. (2004) Rhamnogalacturonan I in *Solanum tuberosum* tubers contains complex arabinogalactan structures. *Phytochemistry*, **65**, 1429–1438.
- Ralph, J. and Landucci, L. (2010) NMR of lignins. In *Lignin and Lignans: Advances in Chemistry* (Heitner, C., Dimmel, D. and Schmidt, J.A., eds), pp. 137–234. Boca Raton, FL: CRC Press.
- Ralph, J., MacKay, J., Hatfield, R., O'Malley, D., Whetten, R. and Sederoff, R. (1997) Abnormal lignin in a loblolly pine mutant. *Science*, **277**, 235–239.
- Ralph, J., Lapierre, C., Marita, J.M., Kim, H., Lu, F., Hatfield, R.D., Ralph, S., Chapple, C., Franke, R., Hemm, M.R., Van Doorselaere, J., Sederoff, R.R., O'Malley, D.M., Scott, J.T., MacKay, J.J., Yahiaoui, N., Boudet, A., 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., Kim, H., Lu, F., Grabber, J.H., Leplé, J.-C., Berrio-Sierra, J., Derikvand, M.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.
- Ruprecht, C., Mutwil, M., Saxe, F., Eder, M., Nikoloski, Z. and Persson, S. (2010) Large-scale co-expression approach to dissect secondary cell wall formation across plant species. *Front. Plant Physiol.* **2**, 23.
- Schaaf, G., Honsbein, A., Meda, A.R., Kirchner, S., Wipf, D. and von Wirén, N. (2006) AtIREG2 encodes a tonoplast transport protein involved in iron-dependent nickel detoxification in Arabidopsis thaliana roots. *J. Biol. Chem.* **281**, 25532–25540.
- Simmons, B.A., Loque, D. and Ralph, J. (2010) Advances in modifying lignin for enhanced biofuel production. *Curr. Opin. Plant Biol.* **13**, 313–320.
- Simpson, P.J., Tantitadapitak, C., Reed, A.M., Mather, O.C., Bunce, C.M., White, S.A. and Ride, J.P. (2009) Characterization of two novel aldo-keto reductases from Arabidopsis: expression patterns, broad substrate specificity, and an open active-site structure suggest a role in toxicant metabolism following stress. *J. Mol. Biol.* **392**, 465–480.
- Sluiter, A., Hames, B., Ruiz, R., Scarlata, C., Sluiter, J., Templeton, D. and Crocker, D. (2008) Determination of structural carbohydrates and lignin in biomass. In *Laboratory Analytical Procedure* (Technical Report, NREL/TP-510-42618), Golden, CO: National Renewable Energy Laboratory.
- Stewart, J.J., Akiyama, T., Chapple, C., Ralph, J. and Mansfield, S.D. (2009) The effects on lignin structure of overexpression of ferulate 5-hydroxylase in hybrid poplar. *Plant Physiol.* **150**, 621–635.
- Studer, M.H., DeMartini, J.D., Davis, M.F., Sykes, R.W., Davison, B., Keller, M., Tuskan, G.A. and Wyman, C.E. (2011) Lignin content in natural Populus variants affects sugar release. *Proc. Natl Acad. Sci. USA*, **108**, 6300–6305.

- Taboada, A., Novo-Uzal, E., Flores, G., Loureda, M., Ros Barceló, A., Masa, A. and Pomar, F. (2010) Digestibility of silages in relation to their hydroxycinnamic acid content and lignin composition. *J. Sci. Food Agric.* **90**, 1155–1162.
- Tan, J., Bednarek, P., Liu, J., Schneider, B., Svatos, A. and Hahlbrock, K. (2004) Universally occurring phenylpropanoid and species-specific indolic metabolites in infected and uninfected *Arabidopsis thaliana* roots and leaves. *Phytochemistry*, **65**, 691–699.
- Umezawa, T. (2010) The cinnamate/monolignol pathway. *Phytochem. Rev.* **9**, 1–17.
- Voelker, S.L., Lachenbruch, B., Meinzer, F.C., Kitin, P. and Strauss, S.H. (2011) Transgenic poplars with reduced lignin show impaired xylem conductivity, growth efficiency and survival. *Plant Cell Environ.* **34**, 655–668.
- Vogt, T. (2010) Phenylpropanoid Biosynthesis. *Mol. Plant*, **3**, 2–20.
- Weng, J.-K. and Chapple, C. (2010) The origin and evolution of lignin biosynthesis. *New Phytol.* **187**, 273–285.
- Xu, B., Escamilla-Treviño, L.L., Sathitsuksanoh, N., Shen, Z., Shen, H., Zhang, Y.-H.P., Dixon, R.A. and Zhao, B. (2011) Silencing of 4-coumarate:coenzyme A ligase in switchgrass leads to reduced lignin content and improved fermentable sugar yields for biofuel production. *New Phytol.* **192**, 611–625.

Supporting information

Additional Supporting information may be found in the online version of this article:

Figure S1 Synthesis of C₆C₁ monomers upon HCHL activity.

Figure S2 Organ- and tissue-specific activity of the *IRX5* promoter in *Arabidopsis*.

Figure S3 Spectral analysis of *IRX5:HCHL* and wild-type plants.

Table S1 Quantitative analysis of soluble aromatics in stems from 5-week-old wild-type (WT) and *IRX5:HCHL* plants.

Table S2 Quantitative analysis of acid-hydrolysed soluble aromatics in stems from 5-week-old wild-type (WT) and *IRX5:HCHL* plants.

Table S3 List of genes differentially expressed genes in the *IRX5:HCHL* stems.

Data S1 Supporting experimental procedures for supplemental data.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.