


CRISPR-Cas9 editing of CAFFEYOYL SHIKIMATE ESTERASE 1 and 2 shows their importance and partial redundancy in lignification in *Populus tremula* × *P. alba*

Lisanne de Vries^{1,2,†}, Marlies Brouckaert^{1,2,†}, Alexandra Chanoca^{1,2}, Hoon Kim³, Matthew R. Regner³, Vitaliy I. Timokhin³, Yi Sun⁴, Barbara De Meester^{1,2}, Jan Van Doorselaere⁵, Geert Goeminne^{1,2,6}, Vincent L. Chiang^{4,7,8}, Jack P. Wang^{4,7}, John Ralph³, Kris Morreel^{1,2}, Ruben Vanholme^{1,2,†} and Wout Boerjan^{1,2,*,†} 

¹Department of Plant Biotechnology and Bioinformatics, Ghent University, Ghent, Belgium

²VIB Center for Plant Systems Biology, Ghent, Belgium

³Department of Biochemistry, and U.S. Department of Energy Great Lakes Bioenergy Research Center, Wisconsin Energy Institute, University of Wisconsin-Madison, Madison, WI, USA

⁴State Key Laboratory of Tree Genetics and Breeding, Northeast Forestry University, Harbin, China

⁵VIVES, Roeselare, Belgium

⁶VIB Metabolomics Core, Ghent, Belgium

⁷Forest Biotechnology Group, Department of Forestry and Environmental Resources, North Carolina State University, Raleigh, NC, USA

⁸Department of Forest Biomaterials, North Carolina State University, Raleigh, NC, USA

Received 10 April 2021;

revised 10 June 2021;

accepted 18 June 2021.

*correspondence (Tel +32 9 3313881; fax +32 9 3313809; email

wout.boerjan@psb.vib-ugent.be)

[†]These authors contributed equally to this work.

Summary

Lignins are cell wall-located aromatic polymers that provide strength and hydrophobicity to woody tissues. Lignin monomers are synthesized via the phenylpropanoid pathway, wherein CAFFEYOYL SHIKIMATE ESTERASE (CSE) converts caffeoyl shikimate into caffeic acid. Here, we explored the role of the two CSE homologs in poplar (*Populus tremula* × *P. alba*). Reporter lines showed that the expression conferred by both *CSE1* and *CSE2* promoters is similar. CRISPR-Cas9-generated *cse1* and *cse2* single mutants had a wild-type lignin level. Nevertheless, *CSE1* and *CSE2* are not completely redundant, as both single mutants accumulated caffeoyl shikimate. In contrast, the *cse1 cse2* double mutants had a 35% reduction in lignin and associated growth penalty. The reduced-lignin content translated into a fourfold increase in cellulose-to-glucose conversion upon limited saccharification. Phenolic profiling of the double mutants revealed large metabolic shifts, including an accumulation of *p*-coumaroyl, 5-hydroxyferuloyl, feruloyl and sinapoyl shikimate, in addition to caffeoyl shikimate. This indicates that the CSEs have a broad substrate specificity, which was confirmed by *in vitro* enzyme kinetics. Taken together, our results suggest an alternative path within the phenylpropanoid pathway at the level of the hydroxycinnamoyl-shikimates, and show that CSE is a promising target to improve plants for the biorefinery.

Keywords: poplar, CRISPR-Cas9, CSE, lignin, phenylpropanoids, metabolic engineering.

Introduction

Lignocellulosic biomass is a promising renewable feedstock for the production of bio-based chemicals and fermentable sugars (Marriott *et al.*, 2016; Van de Wouwer *et al.*, 2018). The cell wall polysaccharides can be hydrolysed to monosaccharides by saccharification, after which the monosaccharides can be fermented to ethanol or other products by specific microorganisms (Vanholme *et al.*, 2013b). However, efficient saccharification is hindered by the presence of lignin in the secondary-thickened cell walls, as lignin physically prevents the hydrolytic enzymes from accessing the cellulose surface and it also adsorbs the saccharification enzymes (Jørgensen *et al.*, 2007; Mansfield *et al.*, 1999). Biomass pretreatments are used to reduce the recalcitrance and therefore improve the saccharification efficiency. Nevertheless, these pretreatments are a costly step in the production of bio-based products (Aden *et al.*, 2002; Vanholme *et al.*, 2013b).

Alternatively, plants can be engineered to contain less lignin and/or lignin with a different composition to reduce biomass recalcitrance (Chanoca *et al.*, 2019; Chen and Dixon, 2007; De Meester *et al.*, 2018; de Vries *et al.*, 2018; Eudes *et al.*, 2014; Halpin, 2019; Mansfield *et al.*, 2012; Van Acker *et al.*, 2014; Van Acker *et al.*, 2013; Wilkerson *et al.*, 2014; Zhang *et al.*, 2012).

In angiosperms, lignin is mostly derived from the monolignols coniferyl alcohol and sinapyl alcohol, with traces of *p*-coumaryl alcohol. These monolignols are synthesized via the general phenylpropanoid and monolignol-specific pathways (Boerjan *et al.*, 2003; Bonawitz and Chapple, 2010; Freudenberg, 1965; Vanholme *et al.*, 2019; Vanholme *et al.*, 2010). After their biosynthesis, the monolignols are deposited in the cell wall, where they are oxidized to radicals by laccases and peroxidases, and subsequently polymerized through combinatorial radical-radical coupling to create the lignin polymer (Berthet *et al.*, 2011; Ralph *et al.*, 2019; Ralph *et al.*, 2004; Zhao *et al.*, 2013).

Upon incorporation into the lignin polymer, *p*-coumaryl alcohol, coniferyl alcohol and sinapyl alcohol give rise to *p*-hydroxyphenyl (H), guaiacyl (G) and syringyl (S) units, respectively. In addition to these three main monolignols, an increasing number of other aromatic compounds has been shown to couple into the lignin polymer (Mottiar *et al.*, 2016; Ralph *et al.*, 2019; Vanholme *et al.*, 2019). The biosynthesis of the two main monolignols coniferyl alcohol and sinapyl alcohol passes through caffeoyl-CoA. Caffeoyl-CoA can be produced from caffeoyl shikimate by *p*-HYDROXYCINNAMOYL-CoA:QUINATE/SHIKIMATE *p*-HYDROXYCINNAMOYLTRANSFERASE (HCT) (Franke *et al.*, 2002; Hoffmann *et al.*, 2004; Hoffmann *et al.*, 2003; Schoch *et al.*, 2001; Shadle *et al.*, 2007). Alternatively, it can be produced from *p*-coumaric acid via either 4-COUMARATE 3-HYDROXYLASE (C3H (Barros *et al.*, 2019)) or an enzyme complex involving *p*-COUMAROYL SHIKIMATE 3'-HYDROXYLASE/CINNAMATE 4-HYDROXYLASE (C3'H/C4H), followed by the action of 4-COUMARATE:CoA LIGASE (4CL) (Chen *et al.*, 2011). In addition to these routes, it was shown that the HCT-catalysed conversion of caffeoyl shikimate into caffeoyl-CoA could be bypassed by the combined activities of CAFFEYOYL SHIKIMATE ESTERASE (CSE) and 4CL in Arabidopsis (Vanholme *et al.*, 2013c). A loss-of-function mutant of CSE in Arabidopsis showed up to 36% reduced-lignin content and a 30-fold increase in the relative amount of H units in the lignin polymer (Vanholme *et al.*, 2013c). Stacking a C3H silencing construct with the *cse* loss-of-function mutation in Arabidopsis resulted in a further reduction in lignin content and an increase in the H unit frequency, as compared to the *cse* mutant (Barros *et al.*, 2019).

The essential role of CSE in monolignol biosynthesis has also been proven in *M. truncatula* and *P. tremula* × *P. alba* (Ha *et al.*, 2016; Saleme *et al.*, 2017). The loss-of-function mutation of CSE in *M. truncatula* caused more severe phenotypes than those caused by the loss-of-function mutation in Arabidopsis, such as severe dwarfism, reduction of lignin levels by 80% and levels of up to 84% H units in the lignin, as opposed to approximately 6% H units in the wild type (WT) (Ha *et al.*, 2016). Poplar has two CSE homologs, CSE1 (corresponding to Potri.001G17500) and CSE2 (corresponding to Potri.003G059200), that share 92% identity in amino acid sequence based on the genome sequence of *P. tremula* × *P. alba* 717-1B4 from AspenDB (Xue *et al.*, 2015; Zhou *et al.*, 2015). Simultaneous down-regulation of CSE1 and CSE2 by a hairpin approach in *P. tremula* × *P. alba* resulted in residual expression levels as low as 35% and 15%, respectively, a 25% reduction in lignin amount, and a twofold increase in H unit content (Saleme *et al.*, 2017). Although the saccharification efficiency of the *hpCSE* was 30% higher than that of the WT, this increase was much smaller than the fourfold improvement observed with the Arabidopsis *cse2* loss-of-function mutant (Vanholme *et al.*, 2013c). We hypothesized that the modest increase in saccharification efficiency and the small elevation in H unit content in comparison with the Arabidopsis *cse2* were caused by the residual CSE expression in the hairpin lines. Additionally, the hairpin strategy down-regulated both CSE genes simultaneously, hindering the understanding of potentially differential roles for each of the two genes individually.

Here, we investigated the role of CSE1 and CSE2 in *P. tremula* × *P. alba*, both independently and simultaneously, via the generation of knockout mutants through CRISPR-Cas9. Growth, lignin analysis and metabolic profiling of the mutants provided evidence that CSE1 and CSE2 are partially redundant in poplar. By comparative metabolic profiling of WT and double mutants and

in vitro enzyme kinetics, we obtained evidence for another metabolic layer in the grid-like phenylpropanoid pathway. Furthermore, the large increases in saccharification efficiency in *cse1 cse2* further supported the contention that CSE is a promising target for engineering trees for the biorefinery.

Results

Generation of *cse1* and *cse2* single mutants and *cse1 cse2* double mutants via CRISPR-Cas9 and GUS expression analyses

Based on the genome sequence of the *P. tremula* × *P. alba* 717-1B4 from AspenDB (Figure S1) (Xue *et al.*, 2015; Zhou *et al.*, 2015), the two CSE paralogs share 92% identity in amino acid sequence. In order to investigate the specific roles of CSE1 and CSE2 in poplar, we generated *cse1* and *cse2* single mutants, as well as *cse1 cse2* double mutants via CRISPR-Cas9. Four gRNAs were designed to target the DNA sequence upstream of the codon for the last amino acid of the catalytic triad so that a frameshift mutation would result in a complete loss of enzyme activity (Figure S2, Supplemental Experimental Procedures). After transformation and regeneration, we obtained independent poplar transformants with mutations caused by gRNA1 solely targeting CSE1 (Table S1), gRNA2 solely targeting CSE2 (Table S2) and gRNA4 targeting both CSE1 and CSE2 (Table S4). No biallelic double mutants targeting both CSE1 and CSE2 were observed with gRNA3 (Table S3). Only the lines that had biallelic frameshift mutations were analysed further.

Based on GUS expression analyses, CSE1 and CSE2 promoters (2 kb upstream sequence from start codon) conferred overlapping expression patterns (Supplemental Results 1.1, Figure S3).

Simultaneous mutation of CSE1 and CSE2 affects plant development

Perturbations of lignin biosynthesis often lead to a yield penalty in plants (Bonawitz and Chapple, 2013; Li *et al.*, 2010; Muro-Villanueva *et al.*, 2019). No differences were observed for the biomass of *cse1* and *cse2* as compared to WT plants, with respect to height, stem diameter, fresh and dry weight after 4 months of growth in the greenhouse (Table S5, Figure S4). In an independent growth experiment, *cse1 cse2* were grown alongside WT. In contrast to the single mutants, the height of the *cse1 cse2* was reduced by 35%, the stem diameter by 14%, stem fresh weight (not debarked) by 52% and stem dry weight (debarked) by 69% as compared to WT (Table S5, Figure S4). These data suggest redundancy of CSE1 and CSE2 in poplar.

Altered phenolic profile in *cse1* and *cse2* single, and *cse1 cse2* double mutant poplars

Next, we analysed the changes in the phenolic metabolism as a consequence of blocking CSE1 and/or CSE2 to gain insights into the individual roles of these genes. Methanol extracts of developing xylem and bark from 4-month-old WT and mutant lines were analysed via reverse-phase ultra-high-performance liquid chromatography–mass spectrometry (UHPLC-MS).

In a first experiment, phenolic profiles of xylem from *cse1* (*n* = 24) and *cse2* (*n* = 15) were compared with those from WT (*n* = 10), yielding a total of 6681 *m/z* peaks (1850 compounds estimated, Supplemental Experimental Procedures). The principal component analysis (PCA) on the 6681 peaks showed that the first two principal components (PCs) were not able to

differentiate the single mutants from WT (Figure S5a). Nevertheless, ANOVA revealed 20 differential peaks (11 compounds, estimated) with a minimum twofold change between the WT, and either *cse1* or *cse2*. For *cse1* and *cse2*, a total of 7 and 10 compounds, respectively, were higher in abundance compared with WT. All 7 compounds accumulating in *cse1* also accumulated in *cse2*, but none could be structurally characterized. For *cse2*, 2 compounds were characterized based on their MS/MS fragmentation spectra (Table 1, Dataset S1, Figure S6). Caffeoyl shikimate (**16**), the known substrate of CSE, increased 2.65-fold in *cse2* and 1.44-fold in *cse1* (below twofold threshold), as compared to WT. In addition, sinapoyl hexose (**10**) increased 5.30-fold in *cse2* in comparison with WT. One unknown compound was significantly decreased in *cse1*, compared with WT. No significantly decreased compounds were detected for *cse2*.

A comparative profiling of xylem extracts of *cse1 cse2* mutants ($n = 27$) versus WT ($n = 5$) yielded a total of 4270 m/z peaks (2356 compounds, estimated). PCA on the 4270 peaks showed that PC1 (counting for 11.5% of the variation) allowed separation between the *cse1 cse2* lines and the WTs (Figure S5b). Following ANOVA, 749 differential peaks (560 compounds, estimated) with a minimal 10-fold change were selected (see Supplemental Experimental Procedures for selection criteria). Of these, 316 compounds accumulated in the *cse1 cse2* mutant lines as compared to WT. Of the accumulating compounds, 39 (**1–25**, including 14 isomers) could be structurally characterized based on their MS/MS fragmentation spectra (Table 2, Figure 1, Dataset S2, Figure S6). Among the accumulating compounds were metabolites with a *p*-coumarate, caffeate, ferulate or sinapate moiety (**4–25**), or derivatives thereof (**1–3**). These included not only caffeoyl shikimate (**16**), the known substrate of CSE, but also *p*-coumaroyl shikimate (**15**), feruloyl shikimate (**18**) and sinapoyl shikimate (**22**). In total, 21 characterized metabolites (including isomers) were hexosylated (**1–14**, **17**, **20**, **21**, **23**), most likely to reduce their toxicity and for storage in the vacuole (Desmet *et al.*, 2021; Dima *et al.*, 2015; Le Roy *et al.*, 2016). There were 244 metabolites that were significantly decreased (with a minimal 10-fold change) in the mutants versus WTs, of which 36 could be characterized (**26–51**, including 13 isomers; Table 2, Figure 1, Dataset S2, Figure S6). In agreement with the reduction in lignin amount in the *cse1 cse2* lines (see below), all of these compounds were oligolignols. These were mainly composed of units derived from coniferyl alcohol, sinapyl alcohol, coniferaldehyde and sinapaldehyde, connected by 8-O-4, 8-5 and 8-8 linkages (**26–42**) (Morreel *et al.*, 2010a; Morreel *et al.*, 2010b). Aliphatic end-groups were sometimes oxidized (**43–45**), and units with 8-5 linkages were reduced (**46**). Also, units derived from sinapyl *p*-

hydroxybenzoate (*p*HBA) (**48–50**), typical for poplar lignin (Morreel *et al.*, 2010a; Morreel *et al.*, 2004), were reduced in abundance. Furthermore, the abundance of syringyl glycerol *p*-hydroxybenzoate ester (**48**) in *cse1 cse2* lines was less than 4% relative to levels in WT. This compound has not been reported before. In addition, one compound that was reduced in abundance was characterized as G(8-O-4)G(8-O-4)*p*-hydroxybenzyl alcohol (**51**).

In a similar set-up, both single and double mutant bark tissues were subjected to phenolic profiling. Analogously to the xylem samples, hydroxycinnamoyl-shikimates [caffeoyl shikimate (**16**), feruloyl shikimate (**18**) and 5-hydroxyferuloyl shikimate (**66**)] and several derivatives thereof accumulated in the *cse1 cse2* bark samples (**17**, **19**, **21**, **63–67**) (Figure 1, Supplemental Results 1.2).

In short, large metabolic shifts (~20% of the estimated number of compounds) were observed in the profiles of the *cse1 cse2* mutants, including the accumulation of various hydroxycinnamoyl-shikimic acid derivatives. This indicates that CSE is able to convert hydroxycinnamoyl-shikimates other than caffeoyl shikimate to their respective acid forms. The fact that caffeoyl shikimate (the known substrate of CSE) accumulates in each of the *cse1* and *cse2* mutants indicates that the CSE1 and CSE2 are not fully redundant at the metabolic level.

A broader role for CSE in the phenylpropanoid pathway

The phenolic profiling showed an accumulation in the *cse1 cse2* mutants of *p*-coumaroyl shikimate (**15**), caffeoyl shikimate (**17**), feruloyl shikimate (**18**), 5-hydroxyferuloyl shikimate (**66**) and sinapoyl shikimate (**22**), suggesting that CSE has a broad substrate specificity. Furthermore, *in silico* protein structure prediction through Phyre2 revealed a L189F substitution for CSE2 in the substrate binding site (Supplemental Results 1.3, Figure S7). To investigate whether CSE1 and CSE2 have different catalytic properties, enzyme kinetics were performed on CSE1 and CSE2 derived from *Populus trichocarpa* [PtrCSE1 has a 98% amino acid sequence identity compared with CSE1 encoded by both the *P. alba* and *P. tremula* alleles; PtrCSE2 has a 97% and 98% amino acid sequence identity compared with CSE2 encoded by the *P. alba* and *P. tremula* alleles, respectively]. The Michaelis–Menten kinetic parameter (K_m), the turnover number (k_{cat}) and the catalytic efficiency (k_{cat}/K_m) of PtrCSE1 and PtrCSE2 were determined by *in vitro* assays for four hydroxycinnamoyl-shikimates (**15**, **16**, **18** and **22**) at each of the enzyme's respective optimal pH (Figure S8a,b) and optimal temperature (Figure S8c,d), and with the protein concentrations mentioned (Figure S8e,f, Table S8). The four tested hydroxycinnamoyl-shikimates were all substrates of PtrCSE1 and PtrCSE2 (Figure S9). PtrCSE1 had the highest k_{cat}/K_m for caffeoyl shikimate, followed by feruloyl shikimate, *p*-coumaroyl

Table 1 List of characterized compounds in xylem extracts of *cse2* single mutants. Column 'No.': the number the compound was assigned in Figure 1 and the text. tR: retention time in minutes. Fold changes were calculated as the ratio of average normalized abundance in mutants/average normalized abundance in WT. The structures are given in Figure S6.

No.	tR	m/z	Name	Average normalized abundance in WT	Average normalized abundance in mutants	fold change
UP						
10	4.39	385.113	Sinapoyl hexose	826	4381	5.3
16	6.12	335.075	Caffeoyl shikimate	5331	14 134	2.7

Table 2 List of characterized compounds in xylem extracts of *cse1 cse2* double mutants. Column 'No.': the number the compound was assigned in Figure 1 and the text. tR: retention time in minutes. Fold changes were calculated as the ratio of average normalized abundance in mutants/average normalized abundance in WT. The structures are given in Figure S6. Superscripts: 1 = feature detected as ion-neutral complex, 2 = feature detected as formate adduct, 3 = in source fragment of unknown origin. b.d.l. = below detection limit, * = exact position of OH-groups could not be inferred from the fragmentation spectra, biochemical knowledge was used to support the characterization

No.	tR	m/z	Name	Average normalized abundance in WT	Average normalized abundance in mutants	Fold change
UP						
1	1.8	315.072	Protocatechoyl hexose*	299	5185	17
2	2.32	659.177	Vanillic acid hexoside ¹	2099	99 993	48
3	3.66	461.128	Vanillic acid pentosyl-hexoside	81	2488	31
4	14.89	475.124	Hydroxybenzoyl (feruloyl) hexoside	b.d.l.	6797	>100
5	11.97	477.139	Dihydroxybenzyl alcohol (feruloyl) hexoside	b.d.l.	11 459	>1000
6	5.45	355.103	Feruloyl hexose	9234	154 450	17
6	5.96	355.103	Feruloyl hexose	611	28 801	47
6	6.53	355.103	Feruloyl hexose	592	17 342	29
7	12.75	531.149	Diferuloyl hexose	b.d.l.	28 838	>1000
8	12.9	491.119	Protocatechuic acid ferulic acid hexoside*	256	16 260	63
9	2.26	343.104	Dihydrocaffeoyl hexoside	129	3711	29
10	5.66	385.113	Sinapoyl hexose	1746	78 843	45
10	6.84	385.114	Sinapoyl hexose	378	5137	14
11	11.8	535.144	Sinapoyl (vanilloyl) hexose	b.d.l.	6502	>100
12	11.97	537.160	Hydroxyvanillyl alcohol (sinapoyl) hexoside	b.d.l.	7646	>100
13	1.66	506.131	1,2,3,4,5-pentahydroxy-(cysteinyl-1-S)-hexanoyl sinapic acid	b.d.l.	10 669	>1000
14	2.91	593.166	4-O-hexosyl-sinapoyl hexose ²	b.d.l.	2032	>100
14	3.17	593.168	4-O-hexosyl-sinapoyl hexose ²	82	11 325	138
15	7.38	319.082	<i>p</i> -coumaroyl shikimate	372	36 405	98
16	2	335.079	Caffeoyl shikimate	11	1330	120
16	5.78	335.077	Caffeoyl shikimate	1871	48 164	26
16	6.2	335.077	Caffeoyl shikimate	20 360	308 659	15
16	7.71	335.078	Caffeoyl shikimate	703	11 941	17
17	3.78	995.257	Caffeoyl shikimate hexoside ¹	b.d.l.	6625	>100
17	4.96	497.128	Caffeoyl shikimate hexoside	b.d.l.	58 902	>1000
17	7.21	497.130	Caffeoyl shikimate hexoside	28	4675	166
18	4.69	349.093	Feruloyl shikimate	222	9958	45
18	5.61	349.093	Feruloyl shikimate	12	1676	135
18	8.19	349.093	Feruloyl shikimate	2265	40 966	18
18	8.38	349.093	Feruloyl shikimate	31	4885	158
18	9.13	349.093	Feruloyl shikimate	10 196	135 217	13
19	6	429.048	Sulfoferuloyl shikimate	32	2147	67
20	5.1	527.138	5-hydroxyferuloyl shikimate hexoside	b.d.l.	6816	682
21	6.38	511.143	Feruloyl shikimate hexoside	b.d.l.	10 136	1014
22	9.09	379.103	Sinapoyl shikimate	1736	36 183	21
23	5.14	541.154	Sinapoyl (shikimoyl) hexose	b.d.l.	14 722	1473
24	6.8	367.103	Feruloyl quinate	19	1684	88
25	2.45	470.110	Cysteinyl feruloyl quinate	b.d.l.	17 458	1746
25	2.8	470.110	Cysteinyl feruloyl quinate	b.d.l.	17 685	1769
DOWN						
26	15.15	779.288	G(8-O-4)G(8-O-4)S(8-5)G	4774	13	0.0028
26	15.28	779.288	G(8-O-4)G(8-O-4)S(8-5)G	839	b.d.l.	<0.1
27	19.54	809.301	G(8-O-4)G(8-O-4)S(8-8)S	5981	101	0.0170
28	17.96	1005.373	G(8-O-4)G(8-O-4)S(8-8)S(4-O-8)G	13 342	491	0.0368
29	18.16	809.300	G(8-O-4)S(8-8)S(4-O-8)G	152 444	13 896	0.0912
29	18.86	809.300	G(8-O-4)S(8-8)S(4-O-8)G	71 685	5293	0.0738
30	16.04	809.297	G(8-O-4)S(8-O-4)S(8-5)G	1588	b.d.l.	<0.01
30	16.31	809.299	G(8-O-4)S(8-O-4)S(8-5)G	6879	15	0.0021
31	15.71	583.215	G(e8-O-4)S(8-5)G	23 863	386	0.0162
32	15.28	553.206	G(t8-O-4)G(8-8)G ³	48 422	2654	0.0548
33	17.89	839.310	S(8-O-4)S(8-8)S(4-O-8)G	25 988	1918	0.0738

Table 2 Continued

No.	tR	m/z	Name	Average normalized abundance in WT	Average normalized abundance in mutants	Fold change
33	18.59	839.311	S(8-O-4)S(8-8)S(4-O-8)G ³	13 430	355	0.0264
34	17.88	869.320	S(8-O-4)S(8-O-4)S(8-8)S	5703	222	0.0390
34	18.32	869.320	S(8-O-4)S(8-O-4)S(8-8)S	7257	92	0.0126
35	14.67	613.224	S(t8-O-4)S(8-5)G	8700	124	0.0142
36	16.95	777.273	G(8-O-4)G(8-O-4)S(8-5)G'	10 751	179	0.0166
37	15.85	825.293	G(8-O-4)G(8-O-4)S(8-O-4)S'	38 278	441	0.0115
37	16.53	825.294	G(8-O-4)G(8-O-4)S(8-O-4)S'	15 730	360	0.0229
38	14.83	599.210	G(8-O-4)G(8-O-5)S'	6577	23	0.0036
39	15.58	629.220	G(e8-O-4)S(8-O-4)S'	5858	184	0.0314
40	16.93	581.201	G(t8-O-4)S(8-5)G'	164 401	8753	0.0532
41	14.84	629.221	G(t8-O-4)S(8-O-4)S'	38 762	1220	0.0315
42	14.7	659.231	S(8-O-4)S(8-O-4)S'	14 655	246	0.0168
43	10.66	813.290	G(8-O-4)G(8-O-4)S(8-5)G-glycerol	6216	77	0.0124
44	5.69	439.159	G(8-O-4)S-glycerol	6586	76	0.0116
45	12.45	599.210	G'(8-O-4)G(8-O-4)S-OH	10 721	59	0.0055
46	12.96	555.221	G(8-O-4)G(red8-5)G	61 815	6121	0.0990
46	13.26	555.221	G(8-O-4)G(red8-5)G	40 959	1281	0.0313
46	13.7	555.221	G(8-O-4)G(red8-5)G ³	5570	26	0.0047
47	15.46	737.241	G(8-O-4)G(8-O-4)alpha-oxoS-pHBA	15 724	418	0.0266
48	8.41	363.109	S-glycerol pHBA	11 431	452	0.0396
49	11.31	559.180	G(8-O-4)S-glycerol-pHBA	18 713	11	0.0006
49	11.96	559.180	G(8-O-4)S-glycerol-pHBA	12 764	19	0.0015
49	12.08	559.180	G(8-O-4)S-glycerol-pHBA	16 551	78	0.0047
50	11.95	589.189	S(8-O-4)S-glycerol-pHBA	8562	11	0.0013
51	7.51	515.190	G(8-O-4)G(8-O-4)hydroxybenzyl alcohol	21 700	147	0.0068

shikimate and sinapoyl shikimate (Table 3). PtrCSE2 had the highest catalytic efficiency for feruloyl shikimate followed by caffeoyl shikimate, *p*-coumaroyl shikimate and sinapoyl shikimate (Table 3). Together with the phenolic profiling data, these results suggest an extra metabolic layer and a role for the different hydroxycinnamoyl-shikimates in the phenylpropanoid pathway (Figure 1).

Reduced-lignin amount and altered lignin composition in *cse1 cse2* mutants

Next, we investigated the effect of mutating *cse1* and/or *cse2* on lignin amount and composition. First, the cell wall residue (CWR) of wood from the mutants was determined by removing soluble compounds in a sequential solvent extraction. No differences in % CWR were observed between *cse1*, *cse2*, *cse1 cse2* and WT (Figure 2a,b, Figure S11 for individual values). Whereas no differences in Klason (acid-insoluble) lignin were observed between *cse1* and *cse2* and WT poplars (Figure 2c), *cse1 cse2* poplars showed a significant reduction of 35% in Klason lignin compared with WT (Figure 2d).

Because the *cse1* and *cse2* mutants did not show a difference in lignin amount, these samples were not further investigated. The lignin composition of *cse1 cse2* mutants was determined via thioacidolysis (releasing only 8-O-4-linked monomers) and two-dimensional nuclear magnetic resonance (2D-NMR) spectroscopy. The relative amount of thioacidolysis-released G units did not differ significantly and was about 30% in both the *cse1 cse2* mutant and WT samples (Figure 2e, Figure S11 for individual values). However, there was a decrease in thioacidolysis-released S units from 70% in WT to 66% in the *cse1 cse2* mutant. The

thioacidolysis-released H units increased about ninefold, from 0.45% in the WT to 4.05% in the *cse1 cse2* mutant (Figure 2e, Figure S11 for individual values).

Next, NMR was performed on 3 pools of *cse1 cse2* mutants and 3 pools of corresponding WT samples (2 plants per pool). Minor differences were evident between the NMR spectra of enzyme lignin from *cse1 cse2* mutants and those of WT (Figure 2f,g, Figure S10). The relative intensities of signals from S units were lower in the *cse1 cse2* poplars compared with WT, whereas the relative intensities of signals from G units were higher; the S/G ratio was therefore lower in the double mutants. The signals assigned to H units were detected at a trace level (0.6%) in WT, as usual, but their relative contribution was considerably elevated to 4.3% (sevenfold) in the *cse1 cse2* poplars (Figure 2f). The relative intensity of the signal assigned to pHBA was similar in spectra of *cse1 cse2* mutant samples and those of the WT (Figure 2f, Figure S10). Due to the changed lignin composition with the change in H levels, the aliphatic regions also differed (Figure 2g, Figure S10c,d). There is a relative decrease in the 8-5 lignin interunit level, a relative increase in the end-groups and a slight relative increase in the 8-O-4 (β -O-4) interunit levels in the lignin of *cse1 cse2* mutant as compared to those of the WT.

Compromised vessels in *cse1 cse2* mutants

Microscopy was performed to investigate the cell anatomy and lignin distribution in the xylem. The bottom part of 4-month-old stems of *cse1 cse2* and WT was debarked, sectioned and investigated via phloroglucinol and Mäule staining and via fluorescence microscopy. The *cse1 cse2* mutant stems were

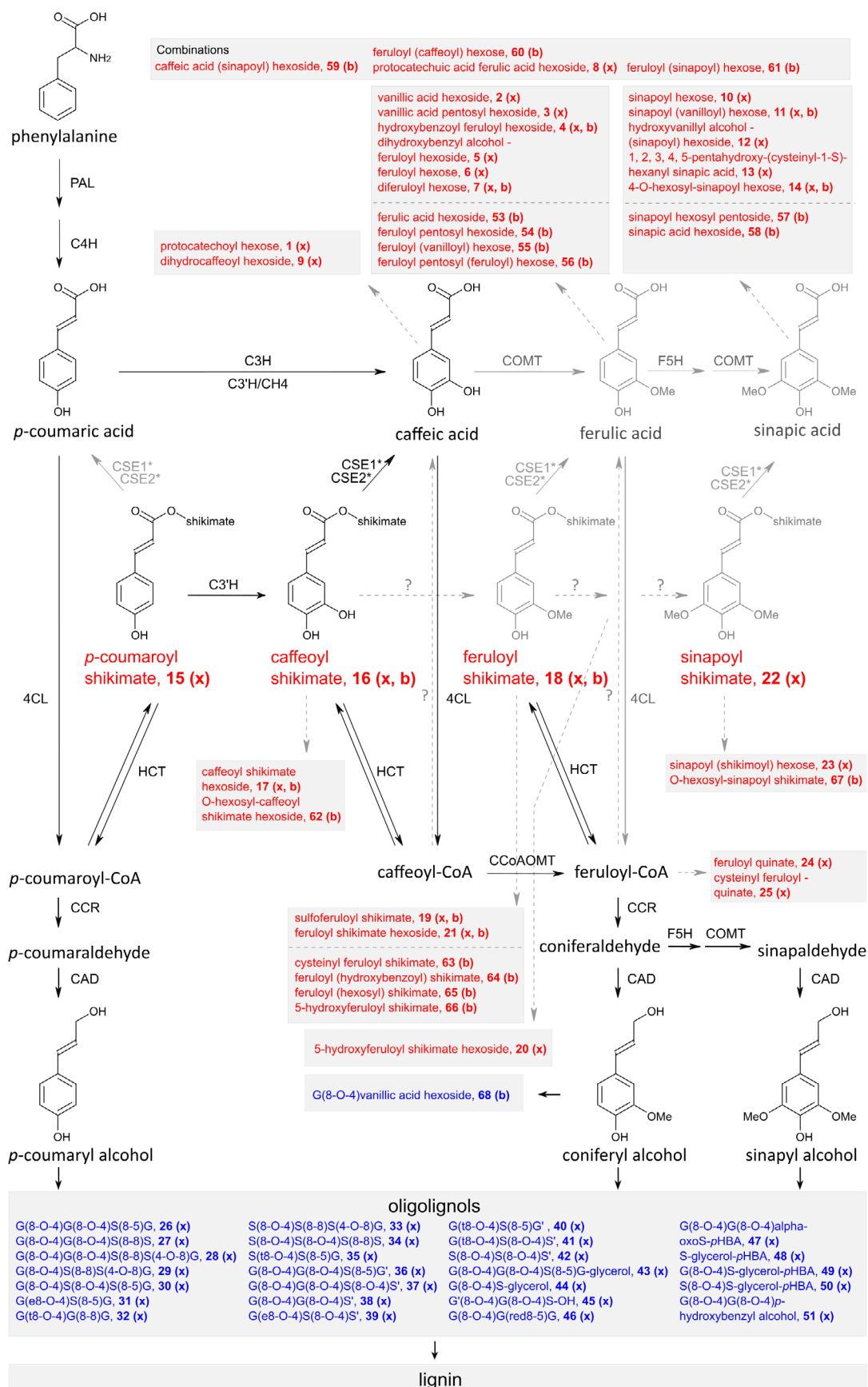


Figure 1 The general phenylpropanoid and monolignol-specific pathways, with the changes in phenolic metabolism indicated for *cse1 cse2* xylem and/or bark extracts compared with WT. Red and blue names indicate metabolites that are up or down, respectively, in *cse1 cse2* xylem and/or bark extracts as compared to WT. Characterized metabolites uniquely detected in bark or xylem profiles are indicated with a (b) or (x), respectively. Co-occurring metabolites are indicated by (x, b). Metabolites that are framed in a box belong to the same class. The flow through the pathway towards the conventional monolignols is in black in WT, whereas (putative) alternative flows are in grey. Successive arrows show two or more metabolic steps. Solid arrows show enzymatic conversions that are validated by experimental evidence, whereas dashed arrows show suggested conversions (Saleme *et al.*, 2017; Van Acker *et al.*, 2017). *The catalytic activity of CSE1 and CSE2 is suggested *in planta* and proven *in vitro* in this work. Each metabolite is indicated with a unique number, corresponding to the metabolites mentioned in the manuscript and Tables 1 and 2 and Tables S6 and S7. Their structures are given in Figure S6. CAD, CINNAMYL-ALCOHOL DEHYDROGENASE; CCoAOMT, CAFFEYOYL-CoA O-METHYLTRANSFERASE; CCR, CINNAMOYL-CoA REDUCTASE; CHS, CHALCONE SYNTHASE; COMT, CAFFEIC ACID O-METHYLTRANSFERASE; F5H, FERULATE 5-HYDROXYLASE; PAL, PHENYLALANINE AMMONIA-LYASE.

considerably softer and more pliable than those of WT; upon sectioning, the tissue often got crushed. Phloroglucinol staining (staining hydroxycinnamaldehyde end-groups is considered to reflect the total amount of lignin) was less intense in the *cse1 cse2* mutant (Figure 3a). Both in *cse1 cse2* and in WT, the vessel cell walls showed a more intense staining than the fibres. Vessel morphology appeared compromised in the double mutants. WT showed the expected Mäule staining pattern, with fibres (rich in S units) stained in red and vessels (rich in G units) in tan. The *cse1 cse2* mutant, however, showed a strong reduction in staining, with fibres and vessels staining in a light orange tone. Moreover, *cse1 cse2* mutant xylem showed uneven staining and stronger signal in ray parenchymal cells compared with other cells. Finally, imaging of lignin autofluorescence by confocal microscopy demonstrated that, in addition to irregular vessel walls, the structure and organization of the xylem fibres were also compromised in the *cse1 cse2* mutant. Accordingly, vessel circularity, determined on microscopy sections visualized via autofluorescence, was reduced (Figure 3b).

Increased cellulose-to-glucose conversion in *cse1 cse2* mutants

To evaluate whether the reduced-lignin amount in the *cse1 cse2* poplars leads to an increase in saccharification efficiency, as in *Arabidopsis cse* mutants and CSE-down-regulated poplars (Saleme *et al.*, 2017; Vanholme *et al.*, 2013c), the cellulose-to-glucose conversion of dried debarked wood of *cse1 cse2* mutants and WT was determined using limited saccharification assays. Cellulose content was measured for each sample via the Updegraff method (Figure 2h, Figure S11 for individual values). No significant difference in cellulose content between *cse1 cse2* and WT was observed. Saccharification assays were performed without pretreatment, and with acid and alkaline pretreatment (Figure 4, Figure S12 for individual values). To determine the cellulose-to-glucose conversion, glucose release was measured 3, 6, 24, 30, 48 and 72 h after addition of the saccharification enzymes. Whereas in WT samples, the cellulose-to-glucose conversion was 15%, 20% and 20% without pretreatment, with an acid and alkaline pretreatment, respectively, the conversion of *cse1 cse2* poplars was significantly higher at all time points and for all conditions. The cellulose-to-glucose conversion after 72 h was 63%, 88% and 84% without, with acid and with alkaline pretreatments, respectively, for the *cse1 cse2* mutants. This means that, independently of the pretreatment, the cellulose-to-glucose conversion increased approximately fourfold in the *cse1 cse2* mutants. However, because of the general biomass yield penalty, no increase was observed when the glucose yield was calculated on a per plant basis (Figure S13).

Discussion

Biosynthesis of phenylpropanoids and lignin in *cse1 cse2* mutant poplar

We have previously shown that down-regulation of CSE in poplar using a hairpin approach resulted in up to 25% reduced-lignin content, a 100%–110% increase in H units and a 62%–91% increase in saccharification efficiency on a plant basis, depending on the pretreatment (Saleme *et al.*, 2017). However, the relatively modest increases in H units and saccharification efficiency compared with the *Arabidopsis cse2* mutant suggested that a full knockout of CSE generated by CRISPR-Cas9 could result in stronger reductions in lignin content and associated improvements in saccharification efficiency (Vanholme *et al.*, 2013c). In addition, CRISPR-Cas9 generated *cse* knockout mutants could reveal the role of the individual CSE genes. Here, we show that lignin content is not affected in poplar *cse1* and *cse2* mutants, and is reduced in *cse1 cse2* mutants by about 35% in comparison with WT. The remaining 65% lignin in *cse1 cse2* mutants imply other routes towards the biosynthesis of lignin in the absence of functional CSE in poplar. One possible route towards coniferyl and sinapyl alcohol could be through HCT. The activity of recombinant poplar HCTs involved in lignification (HCT1 and HCT6) is very low towards caffeoyl shikimate (Wang *et al.*, 2014), but because caffeoyl shikimate increased to very high levels in the *cse1 cse2* mutants, HCT enzymes might work near their V_{\max} in these mutants. Despite that, accumulation of caffeoyl shikimate (**16**) and caffeoyl shikimate hexoside (**17**) in *cse1 cse2* mutants showed that HCT was unable to convert all caffeoyl shikimate to caffeoyl-CoA and that part of it was detoxified via glycosylation or converted to other hydroxycinnamoyl-shikimate derivatives. A second possible route towards coniferyl and sinapyl alcohol is the bypass via either the recently characterized ascorbate peroxidase C3H enzyme or the C3'H/C4H heteromeric complex, whereby *p*-coumaric acid is converted to caffeic acid (Barros *et al.*, 2019; Chen *et al.*, 2011). A third possibility is that a CSE-like protein has partially taken over the role of CSE1 and CSE2, as discussed previously (Ha *et al.*, 2016; Saleme *et al.*, 2017).

Although the frequency of H units had increased by sevenfold in *cse1 cse2* mutants (based on NMR data), the total incorporation of H units remained relatively low (approximately 4% of the total H+G+S units), whereas in *M. truncatula*, a knockout of CSE dramatically increased the frequency of H units in lignin from 6% up to 84% and in *Arabidopsis* from almost 0% up to 44% (Ha *et al.*, 2016; Vanholme *et al.*, 2013c). A possible reason for the relatively low frequency of H units in the lignin in *cse1 cse2* mutant poplars in comparison with the *M. truncatula* mutant could be that CCR2 in poplar is not efficient in catalysing the reaction from *p*-coumaroyl-CoA to *p*-coumaraldehyde [k_{cat}/K_m of

Table 3 Enzyme kinetics of PtrCSE1 and PtrCSE2. Corresponding K_m , k_{cat} and k_{cat}/K_m values with standard deviations for 3 to 5 technical replicates are shown for each tested substrate. CSE1 shows the highest k_{cat}/K_m for caffeoyl shikimate, whereas CSE2 shows highest k_{cat}/K_m for feruloyl shikimate.

Substrates	PtrCSE1			PtrCSE2		
	K_m (μM)	k_{cat}	k_{cat}/K_m	K_m (μM)	k_{cat}	k_{cat}/K_m
<i>p</i> -Coumaroyl shikimate	234 \pm 35	0.43 \pm 0.05	$1.78 \times 10^{-3} \pm 0.41 \times 10^{-3}$	49.40 \pm 7.90	1.27 \pm 0.06	$2.56 \times 10^{-2} \pm 0.20 \times 10^{-2}$
Caffeoyl shikimate	111 \pm 22	2.05 \pm 0.19	$1.87 \times 10^{-2} \pm 0.03 \times 10^{-2}$	30.00 \pm 3.90	1.70 \pm 0.03	$5.65 \times 10^{-2} \pm 0.65 \times 10^{-2}$
Feruloyl shikimate	13.0 \pm 0.9	$6.54 \times 10^{-2} \pm 0.02 \times 10^{-2}$	$5.06 \times 10^{-3} \pm 0.41 \times 10^{-3}$	0.85 \pm 0.11	$7.24 \times 10^{-1} \pm 0.80 \times 10^{-1}$	0.86 \pm 0.10
Sinapoyl shikimate	618 \pm 67	$3.84 \times 10^{-2} \pm 0.14 \times 10^{-2}$	$6.21 \times 10^{-5} \pm 0.07 \times 10^{-5}$	95.30 \pm 8.90	$1.16 \times 10^{-1} \pm 0.05 \times 10^{-1}$	$1.21 \times 10^{-3} \pm 0.15 \times 10^{-3}$

2.51 $\mu M/min$ for feruloyl-CoA vs 0.15 for *p*-coumaroyl-CoA (Wang et al., 2014)], whereas for *M. truncatula* CCR2, *p*-coumaroyl-CoA is actually the preferred substrate with the highest k_{cat}/K_m [0.90 versus 0.40, 0.37 and 0.49 $\mu M/min$ for feruloyl-CoA, sinapoyl-CoA and caffeoyl-CoA respectively (Zhou et al., 2010)]. Nevertheless, for Arabidopsis the preferred substrate is feruloyl-CoA [k_{cat}/K_m of 4.52 $\mu M/min$ vs 3.03 and 0.72 for sinapoyl-CoA and *p*-coumaroyl-CoA, respectively (Baltas et al., 2005)]. However, the k_{cat}/K_m is (relatively) still much higher for Arabidopsis CCR1 compared with poplar CCR2. These observations could suggest that CCR (and/or CAD) cannot efficiently catalyse the biosynthesis of *p*-coumaryl alcohol in poplar, resulting in a relatively low accumulation of H units in the *cse1 cse2* knockouts.

The *cse1 cse2* lines accumulate several hydroxycinnamoyl-shikimates in xylem and bark tissues, such as *p*-coumaroyl shikimate (15), caffeoyl shikimate (16), feruloyl shikimate (18), 5-hydroxyferuloyl shikimate (66), sinapoyl shikimate (22) and several derivatives thereof (17, 20, 21, 23, 63-65, 67). The accumulation of these hydroxycinnamoyl-shikimates suggests that they are also substrates of CSE. Indeed, our enzyme kinetic studies of PtrCSE1 and PtrCSE2 showed that both CSE1 and CSE2 are able to convert shikimates other than caffeoyl shikimate, such as *p*-coumaroyl, feruloyl and to a lesser extent also sinapoyl shikimate *in vitro* (Table 3, Figure S9). Together with the observed accumulation of the hydroxycinnamoyl-shikimates and their derivatives in *cse1 cse2* mutant lines, the enzyme kinetics hint at the existence of another metabolic layer in the

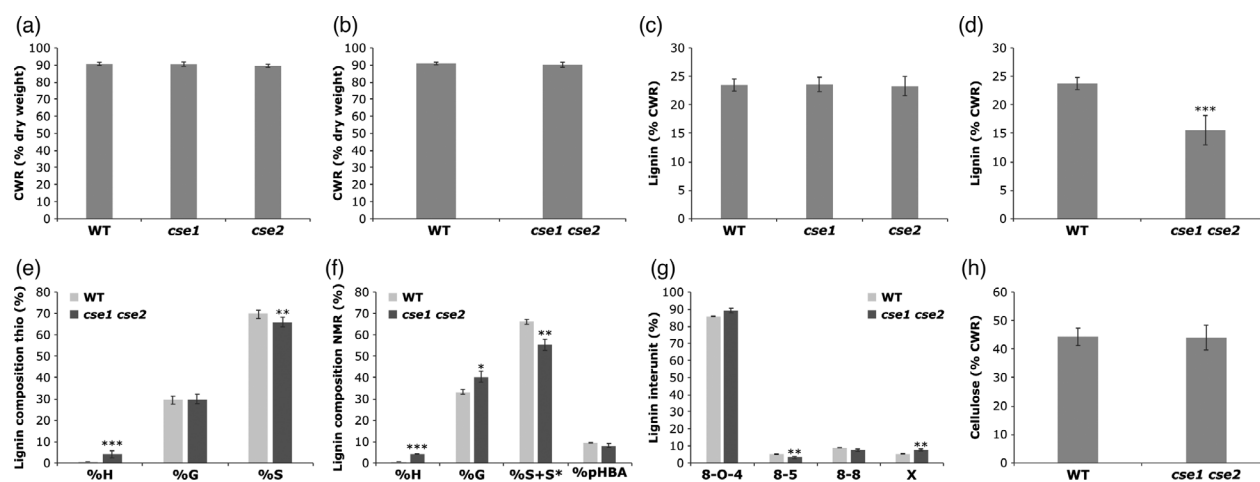


Figure 2 Cell wall characteristics. (a) CWR of WT, *cse1* and *cse2* and (b) CWR of WT and *cse1 cse2* were determined gravimetrically after extraction. (c) Acid-insoluble (Klason) lignin of WT, *cse1* and *cse2* and (d) WT and *cse1 cse2*. (e, f) Lignin composition analysed by thioacidolysis (e) and 2D-NMR (NMR) (f) of *cse1 cse2*, values are expressed in percentage of the respective released lignin unit relative to the sum of H + G + S. S* units are only observed via NMR. These are 5 units with α -keto functionalities that are thought to be derived from S units during the ball-milling process (see Figure 3). (g) Lignin aliphatic region analysed by NMR, values are expressed in percentage of the respective lignin interunit level relative to the sum of all the lignin interunits. 8-O-4: 8-O-aryl ether; 8-5: phenylcoumaran; 8-8: resinol; X: cinnamyl alcohol (end-group). (h) Cellulose content of WT and *cse1 cse2*, determined via the colorimetric Updegraff method. For *cse1* ($n = 15$), *cse2* ($n = 15$) and their corresponding WT ($n = 10$) statistical differences were assessed with ANOVA. For acid-insoluble lignin of *cse1 cse2* $n = 16$ [6 independent lines (line 3, 5.2, 5.3, 7.1, 26 and 28)] and its corresponding WT $n = 5$, for thioacidolysis and cellulose content: *cse1 cse2* $n = 16$ [6 independent lines (line 3, 5.2, 5.3, 7.1, 26 and 28)] and WT $n = 6$ and for NMR *cse1 cse2* $n = 3$ (each sample is a pool of wood derived from 3 poplars) and WT $n = 3$ (each sample is a pool of wood derived from 3 poplars), differences were assessed via Student's t-test. Error bars represent standard deviation. *, 0.05 > P > 0.01; **, 0.01 > P > 0.001; and ***, P < 0.001.

phenylpropanoid pathway, in which hydroxylation and methylation occur at the shikimate ester level, besides the free-acid level, the CoA-thioester level and the aldehyde level, as recently suggested (Saleme *et al.*, 2017). Moreover, cell lysates of yeast lines co-expressing *Populus nigra* HCT1 and *Arabidopsis thaliana* 4CL4 were able to convert caffeic and ferulic acid to the corresponding shikimate esters, illustrating that *Populus nigra* HCT1 is able to use caffeoyl-CoA and feruloyl-CoA (formed by 4CL4 from the acids) as substrates, albeit to a lesser extent than *p*-coumaroyl-CoA (Vanholme *et al.*, 2013a). These results suggest a route in the phenylpropanoid pathway from the acids, through the CoA-thioester intermediates, towards the shikimate esters that, in their turn, could be converted back to their acid forms by the CSE enzymes and in this way potentially form a substrate cycle. Such substrate cycles play an important role in cellular homeostasis, allowing for fluctuations in the cycle without directly affecting other fluxes in the metabolic network (Sridharan *et al.*, 2015), and in this particular case could be important in maintaining the levels of free CoA by regulating the accumulation of CoA-ester intermediates such as caffeoyl-CoA.

CSE1 and CSE2 are partially redundant in poplar

Both CSE promoters confer similar expression patterns (Figure S3a–l), and both CSE proteins were shown to have similar relative activities towards caffeoyl shikimate (Ha *et al.*, 2016; Figure S9), suggesting that CSE1 and CSE2 are redundant in poplar. Despite the normal development and lignin accumulation of the *cse1* and *cse2* mutants, they show metabolic shifts in xylem and bark, including the accumulation of the CSE substrate, caffeoyl shikimate. Furthermore, *cse2* mutants showed more differential compounds in both bark and xylem tissues in comparison with *cse1* mutants, demonstrating that the redundancy of CSE1 and CSE2 is incomplete at the metabolic level. These results, in combination with the difference in catalytic efficiency of PtrCSE1 and PtrCSE2 (Table 3), suggest that CSE1 and CSE2 might be undergoing subfunctionalization. Moreover, our data indicate that the conversion of caffeoyl shikimate to caffeic acid (and potentially the conversion of *p*-coumaroyl shikimate, feruloyl shikimate and sinapoyl shikimate to *p*-coumaric acid, ferulic acid and sinapic acid, respectively) is not the rate-limiting step in the pathway flux to lignin in WT plants, as its decrease in efficiency in the *cse1* and *cse2* mutants, substantiated by the product accumulation, does not result in notable changes in lignin amount.

CSE as a target for engineering low-lignin trees for the biorefinery

A ~fourfold improvement in cellulose-to-glucose conversion was observed in *cse1 cse2* poplars, independent of pretreatment. However, due to the significant biomass penalty this improvement does not persist when the glucose yield is expressed on a plant basis. Nevertheless, if the yield penalty of the plants could be overcome, CSE-edited trees would represent a significantly improved feedstock for the biorefinery. The vessel cell morphology is compromised in *cse1 cse2* lines, as observed in several other plants with perturbations in the lignin biosynthetic pathway (Coleman *et al.*, 2008; Jones *et al.*, 2001; Leplé *et al.*, 2007; Schillmiller *et al.*, 2009; Vanholme *et al.*, 2013c; Voelker *et al.*, 2010). This is a possible reason for the observed yield penalty, either because the function of the vessel cell wall is impaired or because its physico-chemical defects signal a stress response (Gallego-Giraldo *et al.*, 2020; Gallego-Giraldo *et al.*, 2018; Ha

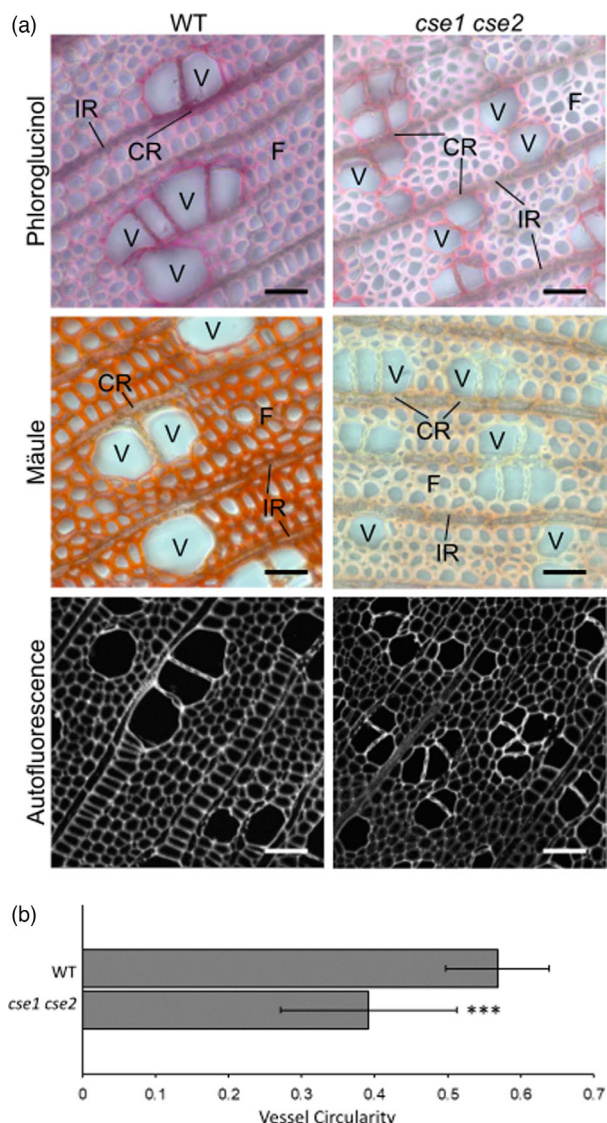


Figure 3 Cellular histology of *cse1 cse2* double mutants and WT. (a) Phloroglucinol, Mäule and autofluorescence pictures of a representative sample of WT and *cse1 cse2* show compromised vessels in the *cse1 cse2* double mutant. Sections stained with phloroglucinol and Mäule were analysed by a light microscope, autofluorescence was analysed by a confocal microscope. V: vessels; F: fibres; CR: contact ray parenchyma; IR: isolation ray parenchyma. (b) vessel circularity was calculated as circularity = $4\pi(\text{area}/\text{perimeter}^2)$. Thirty vessels were measured for each genotype, and differences were assessed via Student's t-test. Error bars represent standard deviation. ***, $P < 0.001$. Bar = 50 μm .

et al., 2021; Muro-Villanueva *et al.*, 2019); restoring lignification in the vessel cells of *Arabidopsis* lignin mutants has been successfully used to recover the vessel cell morphology and overall biomass, while still maintaining the enhanced saccharification efficiency of the mutant biomass (De Meester *et al.*, 2018; Vargas *et al.*, 2016; Yang *et al.*, 2013). However, implementing this strategy in trees might not be straightforward to rescue the biomass yield penalty. It has been shown that in poplar, the fibres adjacent to vessels have a similar lignin composition as the vessels, suggesting that monolignols diffuse from vessels to fibres (Gorzás *et al.*, 2011). Although it is not clear to what extent this

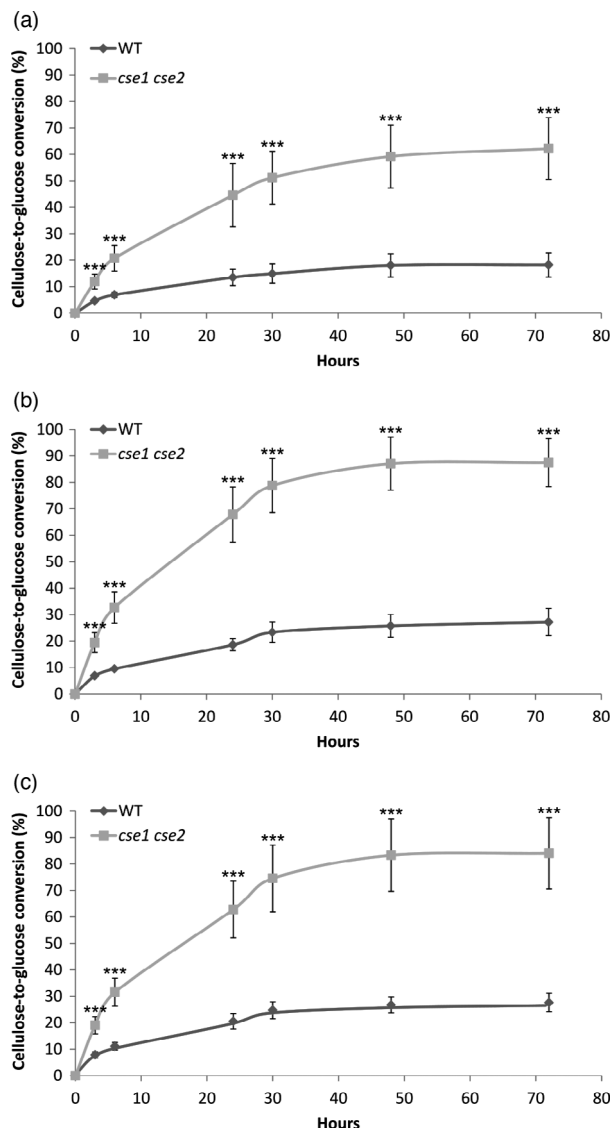


Figure 4 Saccharification efficiency of *cse1 cse2* wood. Wood from *cse1 cse2* ($n = 16$, 6 independent lines (line 3, 5.2, 5.3, 7.1, 26 and 28)) and WT ($n = 5$) was tested for saccharification efficiency at 3, 6, 24, 30, 48 and 72 h, either without pretreatment (a), with acid pretreatment (b) or with alkaline pretreatment (c). Error bars represent standard deviations. Significant differences were assessed with Student's *t*-test (***, $P < 0.001$).

diffusion could re-lignify poorly lignified fibres, such as in the *Arabidopsis* *cse* mutants, it should be plausible to screen for transgenic lines that have normal-shaped and lignified vessels but hypolignified fibres. Such biomass would still present advantages over wild-type biomass.

Down-regulation of *CSE1* and *CSE2* expression in poplar through RNA interference (*hpCSE*) did not result in a yield penalty, even though the lignin amount was reduced by 25% (Saleme *et al.*, 2017). This suggests that there are no phenotypic effects as long as the lignin amount stays above a certain threshold in *CSE*-down-regulated lines. The saccharification efficiency of the *hpCSE*-down-regulated poplars was increased by 31% when no pretreatment was used (Saleme *et al.*, 2017), whereas that of the *cse1 cse2*-mutated poplars had an increase of 320% when no pretreatment was applied. To achieve a high

saccharification efficiency while at the same time avoiding a yield penalty, instead of knocking out *CSE1* and *CSE2*, weak alleles may be generated via CRISPR-Cas9, as recently illustrated for the *CINNAMOYL-CoA REDUCTASE 2* gene by De Meester *et al.* (2020).

Experimental Procedures

Generation of *CSE* mutations via CRISPR-Cas9, poplar transformation and plant material

CSE1 and *CSE2* sequences of *P. tremula* × *P. alba* and a list of predefined gRNAs (used as basis for gRNA design) were obtained from AspenDB (Xue *et al.*, 2015; Zhou *et al.*, 2015). The gRNAs were further selected based on criteria outlined in Supplemental Experimental Procedures. For targeting *CSE1* and *CSE2* individually, one gRNA per gene fulfilled these criteria. To design gRNAs that target *CSE1* and *CSE2* simultaneously, an alignment of *CSE1* and *CSE2* was made in CLC Main Workbench. A search on the PAM-sequence (5'-NGG-3') was performed, resulting in a list of thirteen possible gRNAs that target both *CSE1* and *CSE2* simultaneously. Eight gRNAs fulfilled the selection criteria (Supplemental Experimental Procedures), of which two were chosen. The gRNAs were then cloned via Gibson Assembly into the p201N-Cas9 vector, as previously reported (Jacobs and Martin, 2016).

The different constructs were transformed via electroporation in the *Agrobacterium tumefaciens* strain C58C1. Transformation of *P. tremula* × *P. alba* cv 717-1B4 was performed as previously reported (Leplé *et al.*, 1992). The transformed poplars and their corresponding WT were transferred to soil in pots of 5.5 cm diameter and covered with a cage liner for acclimatization. DNA was extracted from a leaf via the Edwards procedure (Edwards *et al.*, 1991), and amplified and Sanger-sequenced around the different target sites (Table S9 for primer sequences). The Sanger sequences of the different transformed poplar lines were analysed for possible mutations via TIDE (Brinkman *et al.*, 2014).

The transformed poplars that carried mutations (caused by gRNA4) were first propagated to obtain enough biological replicates. After propagation, the poplars were transferred to bigger pots (24 cm diameter) and their heights were measured weekly over 12 weeks ($n = 30$ for *cse1 cse2* and $n = 6$ for WT, $n = 31$ for *cse1*, $n = 33$ for *cse2* and $n = 17$ for WT). After 4 months of growth, the poplars were harvested and their diameters (5 cm above the ground) and weights were measured.

Phenolic profiling

Bark and wood from 10-cm basal stem fragments from 4-month-old *P. tremula* × *P. alba* *cse1 cse2*, *cse1*, *cse2* and WT were subjected to metabolic profiling according to Saleme *et al.* 2017 with the following adaptations: 250 mg of ground tissue was used instead of 100 mg, and 600 µL of methanol supernatants was dried instead of 800 µL (Supplemental Experimental Procedures for adaptations and metabolite selection criteria).

Structural characterization was performed using MS/MS spectral matching against an in-house spectral database, spectral elucidation using published MS fragmentation pathways (Morreel *et al.*, 2014), and *de novo* spectral elucidation software [CSI: FingerID (Dührkop *et al.*, 2015); CFM-ID (Allen *et al.*, 2015)].

Lignin amount and thioacidolysis

Poplars stems were harvested and dried in a Memmert UF 750 oven at 30 °C for five days and subsequently ground into powder

(Fritsch Funnel P-19 for long and bulk solids). This powder was used for cell wall analysis and saccharification. To obtain purified CWR, the samples were subjected to sequential extraction as described in Saleme *et al.* (2017). Lignin quantification was performed via a modified Klason protocol following Saleme *et al.* 2017. Thioacidolysis was performed as previously described (Robinson and Mansfield, 2009). For quantification of the conventional lignin units H, G and S, response factors were used as reported earlier (Yue *et al.*, 2012). To determine the lignin composition regardless of the interunit linkages, 2D-NMR was used as previously described (Kim *et al.*, 2008). For 2D-NMR, three biological replicates were analysed from *cse1 cse2* and WT; each of the biological replicates comprised wood derived from two different lines.

Isolation of enzyme lignin (EL) and NMR

Ball-milled cell wall was prepared, and enzyme lignin (EL) was isolated as described previously (Kim *et al.*, 2017). NMR spectra were acquired according to Saleme *et al.* (2017) with slight adaptations (Supplemental Experimental Procedures).

Lignin staining and vessel circularity analyses

Lignin staining with Wiesner and Mäule reagents was performed as previously described (Pradhan Mitra and Loqué, 2014). Sections of 100 μm were obtained in a Leica Vibratome, mounted in 20 mM Tris, pH 7, and 50% glycerol and visualized using a Zeiss AxioScope. Vessel circularity was determined using the measure command in ImageJ, using 20 \times thresholded images. Circularity is calculated as $\text{circularity} = 4\pi(\text{area}/\text{perimeter}^2)$. Thirty vessels were measured for each genotype.

Cellulose quantification

Cellulose content was estimated via the colorimetric Updegraff method on CWR as previously described (Foster *et al.*, 2010). The absorbance was measured at room temperature at 625 nm.

Saccharification assays

Saccharification assays were performed as previously described (Van Acker *et al.*, 2016; Van Acker *et al.*, 2013), without pretreatment, acid pretreatment 1 M HCl at 80 °C for 2 h) and alkaline pretreatment (62.5 mM NaOH at 90 °C for 3 h). Glucose measurements were performed at 3, 6, 24, 48 and 72 h after the saccharification enzymes were added. The enzyme activity added to each aliquot was 0.015 FPU. To calculate the glucose yield per plant, the cellulose-to-glucose conversion at time point 72 h was multiplied by the average dry weight of debarked stems and the average cellulose content.

Recombinant protein expression and purification

Full-length coding regions of *PtrCSE1* (Potri.003G059200) and *PtrCSE2* (Potri.001G175000) were amplified from cDNA of *P. trichocarpa* (Nisqually-1) stem differentiating xylem and verified by Sanger sequencing. *PtrCSE1* and *PtrCSE2* coding regions were then cloned into the pET101/D-TOPO vector (Invitrogen, Carlsbad, CA) to express recombinant proteins fused with C-terminal 6 \times histidine tag. The assembled constructs (*PtrCSE1-pET101* and *PtrCSE2-pET101*) were transformed into *E. coli* Rosetta 2 (DE3) (EMD Millipore Corp., Billerica, MA), and induced for protein expression using 0.2 mM isopropyl β -D-1-thiogalactopyranoside (IPTG) at 22 °C for 16 h. Recombinant proteins were purified using the ProBond Purification Kit (Invitrogen, Carlsbad, CA) as

described previously (Shuford *et al.*, 2012) and stored at -80 °C before enzyme kinetic analysis.

Enzyme reaction conditions and HPLC analysis

Enzyme kinetic analysis using *p*-coumaroyl, caffeoyl, feruloyl and sinapoyl shikimates, synthesized following Timokhin *et al.* (in preparation), as substrates for *PtrCSE1* and *PtrCSE2*, was performed using purified recombinant proteins and the optimal reaction temperature and pH for each enzyme (Supplemental Experimental Procedures). The enzyme assays were performed as described for optimal temperature and pH determination, using different concentrations of each substrate ranging from 10 to 1000 μM , and reaction time was adjusted between 10 and 240 min to ensure initial reaction velocity was measured. Three to five technical replicates were performed for each enzyme–substrate combination. Kinetic parameters (K_m and K_{cat}) were obtained using Prism 8 (GraphPad, San Diego, CA) by applying the nonlinear Michaelis–Menten curve-fitting function.

Acknowledgements

We thank Chung-Jui (CJ) Tsai for advice on gRNA design and cloning strategy, Andreas Pallidis for the analysis of the thioacidolysis data, Robin De Buck for help with experiments, and Annick Bleys for preparing the manuscript for submission. The authors acknowledge funding from the IWT-FISCH-SBO project ARBOREF (grant number 140894), the IWT-SBO project BIOLEUM (grant number 130039) and the Energy Transition Fund AD-LIBIO. J.R., V.I.T., M.R.R. and H.K. were funded by the DOE Great Lakes Bioenergy Research Center (DOE Office of Science BER DE-SC0018409).

Conflict of interest

R.V. and W.B. are inventors of patent number 9834776.

Author contributions

L.d.V., M.B., A.C., R.V. and W.B. designed the research; L.d.V., M.B., A.C., H.K., M.R.R., V.I.T., Y.S., B.D.M., J.V.D. and G.G. performed the experiments; L.d.V., M.B., A.C., H.K., V.L.C., J.P.W., J.R., K.M. and R.V., performed data analysis; and L.d.V., M.B., A.C., R.V. and W.B. wrote the article.

References

- Van Acker, R., Lep  , J.-C., Aerts, D., Storme, V., Goeminne, G., Ivens, B., L  g  , F. *et al.* (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., 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 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., D  jardin, A., Desmet, S., Hoengenaert, L., Vanholme, R., Morreel, K., Laurans, F. *et al.* (2017) Different Routes for Conifer- and Sinapaldehyde and Higher Saccharification upon Deficiency in the Dehydrogenase CAD1. *Plant Physiol.*, **175**, 1018–1039. <https://doi.org/10.1104/pp.17.00834>.

- 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>).
- Allen, F., Greiner, R. and Wishart, D. (2015) Competitive fragmentation modeling of ESI-MS/MS spectra for putative metabolite identification. *Metabolomics*, **11**, 98–110.
- Baltas, M., Lapeyre, C., Bedos-Belval, F., Maturano, M., Saint-Aguet, P., Roussel, L., Duran, H. et al. (2005) Kinetic and inhibition studies of cinnamoyl-CoA reductase 1 from *Arabidopsis thaliana*. *Plant Physiol. Biochem.*, **43**, 746–753.
- Barros, J., Escamilla-Trevino, L., Song, L., Rao, X., Serrani-Yarce, J.C., Palacios, M.D., Engle, N. et al. (2019) 4-Coumarate 3-hydroxylase in the lignin biosynthesis pathway is a cytosolic ascorbate peroxidase. *Nat. Commun.*, **10**, 1994.
- Berthet, S., Demont-Caulet, N., Pollet, B., Bidzinski, P., Cézard, L., Le Bris, P., Borrega, N. et al. (2011) Disruption of *LACCASE4* and 17 results in tissue-specific alterations to lignification of *Arabidopsis thaliana* stems. *Plant Cell*, **23**, 1124–1137.
- 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.*, **44**, 337–363.
- Bonawitz, N.D. and Chapple, C. (2013) Can genetic engineering of lignin deposition be accomplished without an unacceptable yield penalty? *Curr. Opin. Biotechnol.*, **24**, 336–343.
- Brinkman, E.K., Chen, T., Amendola, M. and van Steensel, B. (2014) Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res.*, **42**, e168.
- Chanoca, A., de Vries, L. and Boerjan, W. (2019) Lignin engineering in forest trees. *Front. Plant Sci.*, **10**, 912.
- Chen, F. and Dixon, R.A. (2007) Lignin modification improves fermentable sugar yields for biofuel production. *Nat. Biotechnol.*, **25**, 759–761.
- Chen, H.-C., Li, Q., Shuford, C.M., Liu, J., Muddiman, D.C., Sederoff, R.R. and Chiang, V.L. (2011) Membrane protein complexes catalyze both 4- and 3-hydroxylation of cinnamic acid derivatives in monolignol biosynthesis. *Proc. Natl. Acad. Sci. USA*, **108**, 21253–21258.
- Coleman, H.D., Park, J.-Y., Nair, R., Chapple, C. and Mansfield, S.D. (2008) RNAi-mediated suppression of *p*-coumaroyl-CoA 3'-hydroxylase in hybrid poplar impacts lignin deposition and soluble secondary metabolism. *Proc. Natl. Acad. Sci. USA*, **105**, 4501–4506.
- Desmet, S., Saey, Y., Verstaen, K., Dauwe, R., Kim, H., Nicolaes, C., Fukushima, A. et al. (2021) Maize specialized metabolome networks reveal organ-preferential mixed glycosides. *Comp. Struct. Biotechnol. J.*, **19**, 1127–1144.
- Dima, O., Morreel, K., Vanholme, B., Kim, H., Ralph, J. and Boerjan, W. (2015) Small glycosylated lignin oligomers are stored in *Arabidopsis* leaf vacuoles. *Plant Cell*, **27**, 695–710.
- Dührkop, K., Shen, H.B., Meusel, M., Rousu, J. and Böcker, S. (2015) Searching molecular structure databases with tandem mass spectra using CSI:FingerID. *Proc. Natl. Acad. Sci. USA*, **112**, 12580–12585.
- Edwards, K., Johnstone, C. and Thompson, C. (1991) A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucleic Acids Res.*, **19**, 1349.
- Eudes, A., Liang, Y., Mitra, P. and Loqué, D. (2014) Lignin bioengineering. *Curr. Opin. Biotechnol.*, **26**, 189–198.
- Foster, C.E., Martin, T.M. and Pauly, M. (2010) Comprehensive compositional analysis of plant cell walls (lignocellulosic biomass). Part II: Carbohydrates. *J. Vis. Exp.*, **37**, e1837.
- Franke, R., Humphreys, J.M., Hemm, M.R., Denault, J.W., Ruegger, M.O., Cusumano, J.C. and Chapple, C. (2002) The *Arabidopsis* *REF8* gene encodes the 3-hydroxylase of phenylpropanoid metabolism. *Plant J.*, **30**, 33–45.
- Freudenberg, K. (1965) Lignin: its constitution and formation from *p*-hydroxycinnamyl alcohols. *Science*, **148**, 595–600.
- Gallego-Giraldo, L., Liu, C., Pose-Albacete, S., Pattathil, S., Peralta, A.G., Young, J., Westpheling, J. et al. (2020) ARABIDOPSIS DEHISCENCE ZONE POLYGALACTURONASE 1 (ADPG1) releases latent defense signals in stems with reduced lignin content. *Proc. Natl. Acad. Sci. USA*, **117**, 3281–3290.
- Gallego-Giraldo, L., Posé, S., Pattathil, S., Peralta, A.G., Hahn, M.G., Ayre, B.G., Sunuwar, J. et al. (2018) Elicitors and defense gene induction in plants with altered lignin compositions. *New Phytol.*, **219**, 1235–1251.
- Gorzás, A., Stenlund, H., Persson, P., Trygg, J. and Sundberg, B. (2011) Cell-specific chemotyping and multivariate imaging by combined FT-IR microspectroscopy and orthogonal projections to latent structures (OPLS) analysis reveals the chemical landscape of secondary xylem. *Plant J.*, **66**, 903–914.
- Ha, C.M., Escamilla-Trevino, L., Yarce, J.C.S., Kim, H., Ralph, J., Chen, F. and Dixon, R.A. (2016) An essential role of caffeoyl shikimate esterase in monolignol biosynthesis in *Medicago truncatula*. *Plant J.*, **86**, 363–375.
- Ha, C.M., Rao, X., Saxena, G. and Dixon, R.A. (2021) Growth-defense trade-offs and yield loss in plants with engineered cell walls. *New Phytol.*, **231**, 60–74.
- Halpin, C. (2019) Lignin engineering to improve saccharification and digestibility in grasses. *Curr. Opin. Biotechnol.*, **56**, 223–229.
- Hoffmann, L., Besseau, S., Geoffroy, P., Ritzenthaler, C., Meyer, D., Lapierre, C., Pollet, B. et al. (2004) Silencing of hydroxycinnamoyl-coenzyme A shikimate/quinate hydroxycinnamoyltransferase affects phenylpropanoid biosynthesis. *Plant Cell*, **16**, 1446–1465.
- Hoffmann, L., Maury, S., Martz, F., Geoffroy, P. and Legrand, M. (2003) Purification, cloning, and properties of an acyltransferase controlling shikimate and quinate ester intermediates in phenylpropanoid metabolism. *J. Biol. Chem.*, **278**, 95–103.
- Jacobs, T.B. and Martin, G.B. (2016) High-throughput CRISPR vector construction and characterization of DNA modifications by generation of tomato hairy roots. *J. Vis. Exp.*, **110**, e53843.
- Jones, L., Ennos, A.R. and Turner, S.R. (2001) Cloning and characterization of *irregular xylem4* (*irx4*): a severely lignin-deficient mutant of *Arabidopsis*. *Plant J.*, **26**, 205–216.
- Jørgensen, H., Kristensen, J.B. and Felby, C. (2007) Enzymatic conversion of lignocellulose into fermentable sugars: challenges and opportunities. *Biofuels Bioprod. Biorefining*, **1**, 119–134.
- 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. *Biomacromol.*, **18**, 4184–4195.
- 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.
- Leplé, J.C., Brasileiro, A.C.M., Michel, M.F., Delmotte, F. and Jouanin, L. (1992) Transgenic poplars: expression of chimeric genes using four different constructs. *Plant Cell Rep.*, **11**, 137–141.
- Leplé, J.-C., Dauwe, R., Morreel, K., Storme, V., Lapierre, C., Pollet, B., Naumann, A. et al. (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., Bonawitz, N.D., Weng, J.-K. and Chapple, C. (2010) The growth reduction associated with repressed lignin biosynthesis in *Arabidopsis thaliana* is independent of flavonoids. *Plant Cell*, **22**, 1620–1632.
- Mansfield, S.D., Kang, K.-Y. and Chapple, C. (2012) Designed for deconstruction - poplar trees altered in cell wall lignification improve the efficacy of bioethanol production. *New Phytol.*, **194**, 91–101.
- Mansfield, S.D., Mooney, C. and Saddler, J.N. (1999) Substrate and enzyme characteristics that limit cellulose hydrolysis. *Biotechnol. Prog.*, **15**, 804–816.
- Marriott, P.E., Gómez, L.D. and McQueen-Mason, S.J. (2016) Unlocking the potential of lignocellulosic biomass through plant science. *New Phytol.*, **209**, 1366–1381.
- De Meester, B., Madariaga Calderón, B., de Vries, L., Pollier, J. et al. (2020) Tailoring poplar lignin without yield penalty by combining a null and haploinsufficient CINNAMOYL-CoA REDUCTASE2 allele. *Nat. Commun.*, **11**, 5020 <https://doi.org/10.1038/s41467-020-18822-w>
- De Meester, B., de Vries, L., Özpapucu, M., Gierlinger, N., Corneille, S., Pallidis, A., Goeminne, G. et al. (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.

- Morreel, K., Dima, O., Kim, H., Lu, F., Nicolaes, C., Vanholme, R., Dauwe, R. et al. (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., Nicolaes, C. et al. (2010b) Mass spectrometry-based fragmentation as an identification tool in lignomics. *Anal. Chem.*, **82**, 8095–8105.
- Morreel, K., Ralph, J., Kim, H., Lu, F., Goeminne, G., Ralph, S., Messens, E. et al. (2004) Profiling of oligolignols reveals monolignol coupling conditions in lignifying poplar xylem. *Plant Physiol.*, **136**, 3537–3549.
- Morreel, K., Saey, Y., Dima, O., Lu, F., Van de Peer, Y., Vanholme, R., Ralph, J. et al. (2014) Systematic structural characterization of metabolites in *Arabidopsis* via candidate substrate-product pair networks. *Plant Cell*, **26**, 929–945.
- Mottiar, Y., Vanholme, R., Boerjan, W., Ralph, J. and Mansfield, S.D. (2016) Designer lignins: harnessing the plasticity of lignification. *Curr. Opin. Biotechnol.*, **37**, 190–200.
- Muro-Villanueva, F., Mao, X. and Chapple, C. (2019) Linking phenylpropanoid metabolism, lignin deposition, and plant growth inhibition. *Curr. Opin. Biotechnol.*, **56**, 202–208.
- Pradhan Mitra, P. and Loqué, D. (2014) Histochemical staining of *Arabidopsis thaliana* secondary cell wall elements. *J. Vis. Exp.*, **87**, e51381.
- Ralph, J., Lapierre, C. and Boerjan, W. (2019) Lignin structure and its engineering. *Curr. Opin. Biotechnol.*, **56**, 240–249.
- Ralph, J., Lundquist, K., Brunow, G., Lu, F., Kim, H., Schatz, P.F., Marita, J.M. et al. (2004) Lignins: natural polymers from oxidative coupling of 4-hydroxyphenyl-propanoids. *Phytochem. Rev.*, **3**, 29–60.
- Robinson, A.R. and Mansfield, S.D. (2009) Rapid analysis of poplar lignin monomer composition by a streamlined thioacidolysis procedure and near-infrared reflectance-based prediction modeling. *Plant J.*, **58**, 706–714.
- Le Roy, J., Huss, B., Creach, A., Hawkins, S. and Neutelings, G. (2016) Glycosylation is a major regulator of phenylpropanoid availability and biological activity in plants. *Front. Plant Sci.*, **7**, 735.
- Saleme, M.L.S., Cesarino, I., Vargas, L., Kim, H., Vanholme, R., Goeminne, G., Van Acker, R. et al. (2017) Silencing *CAFFEYOYL SHIKIMATE ESTERASE* affects lignification and improves saccharification in poplar. *Plant Physiol.*, **175**, 1040–1057.
- Schillmiller, A.L., Stout, J., Weng, J.-K., Humphreys, J., Ruegger, M.O. and Chapple, C. (2009) Mutations in the *cinnamate 4-hydroxylase* gene impact metabolism, growth and development in *Arabidopsis*. *Plant J.*, **60**, 771–782.
- Schoch, G., Goepfert, S., Morant, M., Hehn, A., Meyer, D., Ullmann, P. and Werck-Reichhart, D. (2001) CYP98A3 from *Arabidopsis thaliana* is a 3'-hydroxylase of phenolic esters, a missing link in the phenylpropanoid pathway. *J. Biol. Chem.*, **276**, 36566–36574.
- Shadle, G., Chen, F., Reddy, M.S.S., Jackson, L., Nakashima, J. and Dixon, R.A. (2007) Down-regulation of hydroxycinnamoyl CoA: shikimate hydroxycinnamoyl transferase in transgenic alfalfa affects lignification, development and forage quality. *Phytochemistry*, **68**, 1521–1529.
- Shuford, C.M., Li, Q., Sun, Y.-H., Chen, H.-C., Wang, J., Shi, R., Sederoff, R.R. et al. (2012) Comprehensive quantification of monolignol-pathway enzymes in *Populus trichocarpa* by protein cleavage isotope dilution mass spectrometry. *J. Proteome Res.*, **11**, 3390–3404.
- Sridharan, G.V., Ullah, E., Hassoun, S. and Lee, K. (2015) Discovery of substrate cycles in large scale metabolic networks using hierarchical modularity. *BMC Syst. Biol.*, **9**, 5.
- Vanholme, B., Cesarino, I., Goeminne, G., Kim, H., Marroni, F., Van Acker, R., Vanholme, R. et al. (2013a) Breeding with rare defective alleles (BRDA): a natural *Populus nigra* HCT mutant with modified lignin as a case study. *New Phytol.*, **198**, 765–776.
- Vanholme, B., Desmet, T., Ronsse, F., Rabaey, K., Van Breusegem, F., De Mey, M., Soetaert, W. et al. (2013b) 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. et al. (2013c) Caffeoyl shikimate esterase (CSE) is an enzyme in the lignin biosynthetic pathway in *Arabidopsis*. *Science*, **341**, 1103–1106.
- Vanholme, R., Demedts, B., Morreel, K., Ralph, J. and Boerjan, W. (2010) Lignin biosynthesis and structure. *Plant Physiol.*, **153**, 895–905.
- Vanholme, R., De Meester, B., Ralph, J. and Boerjan, W. (2019) Lignin biosynthesis and its integration into metabolism. *Curr. Opin. Biotechnol.*, **56**, 230–239.
- Vargas, L., Cesarino, I., Vanholme, R., Voorend, W., Saleme, M.L.S., Morreel, K. and Boerjan, W. (2016) Improving total saccharification yield of *Arabidopsis* plants by vessel-specific complementation of *caffeoyl shikimate esterase* (*cse*) mutants. *Biotechnol. Biofuels*, **9**, 139.
- Voelker, S.L., Lachenbruch, B., Meinzer, F.C., Jourdes, M., Ki, C., Patten, A.M., Davin, L.B. et al. (2010) Antisense down-regulation of *4CL* expression alters lignification, tree growth, and saccharification potential of field-grown poplar. *Plant Physiol.*, **154**, 874–886.
- de Vries, L., Vanholme, R., Van Acker, R., De Meester, B., Sundin, L. and Boerjan, W. (2018) Stacking of a low-lignin trait with an increased guaiacyl and 5-hydroxyguaiacyl unit trait leads to additive and synergistic effects on saccharification efficiency in *Arabidopsis thaliana*. *Biotechnol. Biofuels*, **11**, 257.
- Wang, J.P., Naik, P.P., Chen, H.-C., Shi, R., Lin, C.-Y., Liu, J., Shuford, C.M. et al. (2014) Complete proteomic-based enzyme reaction and inhibition kinetics reveal how monolignol biosynthetic enzyme families affect metabolic flux and lignin in *Populus trichocarpa*. *Plant Cell*, **26**, 894–914.
- Wilkerson, C.G., Mansfield, S.D., Lu, F., Withers, S., Park, J.-Y., Karlen, S.D., Gonzales-Vigil, E. et al. (2014) Monolignol ferulate transferase introduces chemically labile linkages into the lignin backbone. *Science*, **344**, 90–93.
- Van de Wouwer, D., Boerjan, W. and Vanholme, B. (2018) Plant cell wall sugars: sweeteners for a bio-based economy. *Physiol. Plant.*, **164**, 27–44.
- Xue, L.-J., Alabady, M.S., Mohebbi, M. and Tsai, C.-J. (2015) Exploiting genome variation to improve next-generation sequencing data analysis and genome editing efficiency in *Populus tremula* × *alba* 717–1B4. *Tree Genet. Genomes*, **11**, 82.
- Yang, F., Mitra, P., Zhang, L., Prak, L., Verhertbruggen, Y., Kim, J.-S., Sun, L. et al. (2013) Engineering secondary cell wall deposition in plants. *Plant Biotechnol. J.*, **11**, 325–335.
- Yue, F., Lu, F., Sun, R.-C. and Ralph, J. (2012) Syntheses of lignin-derived thioacidolysis monomers and their uses as quantitation standards. *J. Agric. Food Chem.*, **60**, 922–928.
- Zhang, K., Bhuiya, M.-W., Pazo, J.R., 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.
- Zhao, Q., Nakashima, J., Chen, F., Yin, Y., Fu, C., Yun, J., Shao, H. et al. (2013) *LACCASE* is necessary and nonredundant with *PEROXIDASE* for lignin polymerization during vascular development in *Arabidopsis*. *Plant Cell*, **25**, 3976–3987.
- Zhou, R., Jackson, L., Shadle, G., Nakashima, J., Temple, S., Chen, F. and Dixon, R.A. (2010) Distinct cinnamoyl CoA reductases involved in parallel routes to lignin in *Medicago truncatula*. *Proc. Natl. Acad. Sci. USA*, **107**, 17803–17808.
- Zhou, X.H., Jacobs, T.B., Xue, L.-J., Harding, S.A. and Tsai, C.-J. (2015) Exploiting SNPs for biallelic CRISPR mutations in the outcrossing woody perennial *Populus* reveals 4-coumarate:CoA ligase specificity and redundancy. *New Phytol.*, **208**, 298–301.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Data S1 List of structurally characterized metabolites with a different abundance in the *cse1* and *cse2* single mutant xylem extracts as compared to WT.

Data S2 List of structurally characterized metabolites with a different abundance in the *cse1 cse2* double mutant xylem extracts as compared to WT.

Data S3 List of structurally characterized metabolites with a different abundance in the *cse1* and *cse2* single mutant bark extracts as compared to WT.

Data S4 List of structurally characterized metabolites with a different abundance in the *cse1 cse2* double mutant bark extracts as compared to WT.

Figure S1 *P. tremula* x *P. alba* sequences of CSE1 and CSE2.

Figure S2 Schematic Representation of *P. tremula* x *P. alba* CSE1 and CSE2 with the Location of the possible gRNAs.

Figure S3 GUS Staining of *pCSE1:GUS* and *pCSE2:GUS* Plants.

Figure S4 Phenotype of representative poplars grown in the greenhouse for four months and individual values of height measurements

Figure S5 Principal Component Analysis (PCA) plots of phenolic profiling experiments

Figure S6 Molecular structures of characterized compounds that are higher or lower in abundance in the *cse1 cse2* double mutant, or in the *cse1* and *cse2* single mutant xylem and/or bark extracts compared to WT.

Figure S7 Protein modeling of the *P. tremula* and *P. alba* CSE1 and CSE2 proteins.

Figure S8 Optimization of pH and temperature for enzyme kinetics.

Figure S9 Enzyme kinetics of PtrCSE1 and PtrCSE2.

Figure S10 Lignin Composition Analyzed by 2D-NMR.

Figure S11 Individual values of cell wall characteristics for WT and *cse1 cse2*.

Figure S12 Individual values for saccharification efficiency at 72 h timepoint.

Figure S12 Released glucose per plant of WT (n=5) and *cse1 cse2* (n=16).

Table S1 Sequencing results of *P. tremula* x *P. alba* poplars around target site of gRNA1 targeting CSE1.

Table S2 Sequencing results of *P. tremula* x *P. alba* poplars around target site of gRNA2 targeting CSE2.

Table S3 Sequencing results of *P. tremula* x *P. alba* poplars around target site of gRNA3 targeting both CSE1 and CSE2.

Table S4 Sequencing results of *P. tremula* x *P. alba* poplars around target site of gRNA4 targeting both CSE1 and CSE2.

Table S5 Biomass Yield of *cse1* and *cse2* Single Mutants and *cse1 cse2* Double Mutants.

Table S6 List of characterized compounds in bark extracts of *cse1* and *cse2* single mutants.

Table S7 List of characterized compounds in bark extracts of *cse1 cse2* double mutants.

Table S8 Protein concentrations of CSE1 and CSE2 used for enzyme kinetic assays.

Table S9 Primers used for sequencing CSE1 and CSE2 to confirm mutations caused by CRISPR-Cas9 and primers used for confirming destination vector *pCSE1:GUS* and *pCSE2:GUS*.

Table S10 gRNA1 alignments with the 25 closest potential off-targets.

Table S11 gRNA2 alignments with the 25 closest potential off-targets.

Table S12 gRNA4 alignments with the 25 closest potential off-targets.