

# Silencing *CAFFEOYL SHIKIMATE ESTERASE* Affects Lignification and Improves Saccharification in Poplar<sup>1[OPEN]</sup>

Marina de Lyra Soriano Saleme,<sup>a,b,2</sup> Igor Cesarino,<sup>a,b,c,2</sup> Lívia Vargas,<sup>a,b</sup> Hoon Kim,<sup>d,e</sup> Ruben Vanholme,<sup>a,b</sup> Geert Goeminne,<sup>a,b</sup> Rebecca Van Acker,<sup>a,b</sup> Fernando Campos de Assis Fonseca,<sup>a,b</sup> Andreas Pallidis,<sup>a,b</sup> Wannes Voorend,<sup>a,b</sup> José Nicomedes Junior,<sup>a,b,f</sup> Dharshana Padmakshan,<sup>d,e</sup> Jan Van Doorselaere,<sup>g</sup> John Ralph,<sup>d,e</sup> and Wout Boerjan<sup>a,b,3</sup>

<sup>a</sup>Ghent University, Department of Plant Biotechnology and Bioinformatics, 9052 Ghent, Belgium

<sup>b</sup>VIB Center for Plant Systems Biology, 9052 Ghent, Belgium

<sup>c</sup>Department of Botany, Institute of Biosciences, University of São Paulo, 05508-090 Butanta, Sao Paulo, Brazil

<sup>d</sup>Department of Energy Great Lakes Bioenergy Research Center, Wisconsin Energy Institute, University of Wisconsin, Madison, Wisconsin 53726

<sup>e</sup>Department of Biochemistry, University of Wisconsin, Madison, Wisconsin 53706

<sup>f</sup>Petróleo Brasileiro S.A., Centro de Pesquisas Leopoldo Américo Miguez de Mello, Rio de Janeiro, 21941-598, Brazil

<sup>g</sup>VIVES, 8800 Roeselare, Belgium

ORCID IDs: 0000-0001-6526-0055 (M.d.L.S.S.); 0000-0002-6789-2432 (I.C.); 0000-0001-6334-8782 (L.V.); 0000-0001-7425-7464 (H.K.); 0000-0001-5848-3138 (R.V.); 0000-0002-0337-2999 (G.G.); 0000-0002-0092-1155 (R.V.A.); 0000-0001-6434-5954 (F.C.d.A.F.); 0000-0001-7680-912X (A.P.); 0000-0001-9254-1557 (W.V.); 0000-0003-0591-7021 (J.N.J.); 0000-0002-2916-0417 (D.P.); 0000-0002-3768-2434 (J.V.D.); 0000-0002-6093-4521 (J.R.); 0000-0003-1495-510X (W.B.).

Caffeoyl shikimate esterase (CSE) was recently shown to play an essential role in lignin biosynthesis in *Arabidopsis thaliana* and later in *Medicago truncatula*. However, the general function of this enzyme was recently questioned by the apparent lack of CSE activity in lignifying tissues of different plant species. Here, we show that down-regulation of CSE in hybrid poplar (*Populus tremula* × *Populus alba*) resulted in up to 25% reduced lignin deposition, increased levels of *p*-hydroxyphenyl units in the lignin polymer, and a relatively higher cellulose content. The transgenic trees were morphologically indistinguishable from the wild type. Ultra-high-performance liquid chromatography-mass spectrometry-based phenolic profiling revealed a reduced abundance of several oligolignols containing guaiacyl and syringyl units and their corresponding hydroxycinnamaldehyde units, in agreement with the reduced flux toward coniferyl and sinapyl alcohol. These trees accumulated the CSE substrate caffeoyl shikimate along with other compounds belonging to the metabolic classes of benzenoids and hydroxycinnamates. Furthermore, the reduced lignin amount combined with the relative increase in cellulose content in the CSE down-regulated lines resulted in up to 62% more glucose released per plant upon limited saccharification when no pretreatment was applied and by up to 86% and 91% when acid and alkaline pretreatments were used. Our results show that CSE is not only important for the lignification process in poplar but is also a promising target for the development of improved lignocellulosic biomass crops for sugar platform biorefineries.

Secondary cell walls account for the majority of plant lignocellulosic biomass, which constitutes an important renewable and sustainable feedstock for the production of fermentable sugars, biochemicals, and biomaterials (Vanholme et al., 2013b; Marriott et al., 2016). Bio-refining of plant biomass has attracted significant attention due to global climate change and the need for alternatives to fossil resources for fuels and materials. In the biorefinery, plant cell wall polysaccharides are depolymerized into simple monomeric sugars, a process called saccharification, that are subsequently fermented to ethanol or other compounds by optimized microorganisms (Zeng et al., 2014). However, the complex chemical composition and physical structure of plant cell walls hampers the efficient hydrolysis of lignocellulose, a fact known as plant biomass recalcitrance (Chen and Dixon, 2007; Cullis and Mansfield, 2010; Chundawat et al., 2011; Ding et al., 2012; Zhao et al., 2012a; Rinaldi et al., 2016). One of the major

factors contributing to biomass recalcitrance is the presence of lignin, an aromatic polymer that provides strength and hydrophobicity to the cell wall. Lignin hinders the efficient enzymatic depolymerization of cellulose and hemicellulose into fermentable sugars by immobilizing the hydrolytic enzymes and physically limiting the access to their polysaccharide substrates (Chang and Holtzapfel, 2000; Berlin et al., 2006; Nakagame et al., 2010; Guo et al., 2014). Although a number of pretreatments have been developed to remove lignin and consequently lower biomass recalcitrance, the pretreatment is still a relatively expensive step in the conversion process (Zhao et al., 2012b; Vanholme et al., 2013b). In this regard, lignin bioengineering (e.g. engineering plants that either accumulate less lignin or produce lignin polymers more amenable to chemical degradation) holds promise to tailor plants with reduced biomass recalcitrance (Chen and Dixon, 2007; Mansfield et al., 2012a; Eudes et al., 2014; Wilkerson et al.,

2014; Mottiar et al., 2016). The rational engineering of lignin for the development of crops with improved bio-conversion properties requires a fundamental knowledge of the phenolic metabolism, explaining the large interest in lignin research.

Lignin is a complex aromatic biopolymer deposited mainly in thickened secondary walls of specialized cell types, where it provides mechanical reinforcement for the plant to stand upright, and imperviousness to the cell wall, allowing long-distance water transport (Boerjan et al., 2003; Ralph et al., 2004; Bonawitz and Chapple, 2010). Lignin is derived from monomers biosynthesized in the phenylpropanoid pathway, whose primary precursor is the aromatic amino acid Phe. Following the deamination of Phe by the entry-point enzyme PHENYLALANINE AMMONIA LYASE (PAL), the resulting product *trans*-cinnamic acid undergoes a series of aromatic ring and propene tail modifications to yield three hydroxycinnamyl alcohols, *p*-coumaryl, coniferyl, and sinapyl alcohol, differing in their degrees of methoxylation (Cesarino et al., 2012; Liu, 2012). Once incorporated into the polymer, these monolignols produce *p*-hydroxyphenyl (H), guaiacyl (G), and syringyl (S) units, respectively. In addition to the canonical monolignols, other less abundant units might be incorporated at varying levels, such as hydroxycinnamaldehydes, hydroxycinnamyl *p*-coumarates, and hydroxycinnamate esters (Ralph et al., 2004; Vanholme et al., 2012a). These monomers are first

synthesized in the cytoplasm before being transported to the apoplast, where they are oxidized by peroxidases and/or laccases for subsequent polymerization through combinatorial radical-radical coupling (Ralph et al., 2004; Wang et al., 2013). Coupling between two electron-delocalized monolignol radicals and, more importantly, between a monolignol radical and the growing polymer (radical) results in various linkage types, mainly  $\beta$ -aryl ether ( $\beta$ -O-4), phenylcoumaran ( $\beta$ -5), and resinol ( $\beta$ - $\beta$ ) bonding types (Freudenberg and Neish, 1968; Ralph et al., 2004; Morreel et al., 2010). Because lignin polymerization is combinatorial and not controlled by proteins, coupling reactions are understood to be solely under chemical control, being affected by typical physical parameters such as pH, ionic strength, and substrate supply, among others (Ralph et al., 2004, 2008; Vanholme et al., 2008; Dima et al., 2015). Finally, monolignol biosynthesis is not only developmentally regulated but also responds to specific environmental conditions, such as biotic and abiotic stresses (Moura et al., 2010; Barros et al., 2015).

The biochemical pathway leading to the production of the canonical monolignols was thought to be completely elucidated over a decade ago. When the gene encoding *p*-COUMARATE 3'-HYDROXYLASE (C3H) was finally identified in *Arabidopsis* (*Arabidopsis thaliana*) and its enzymatic activity properly characterized, it was found that free *p*-coumaric acid or its CoA esters were not metabolized by the recombinant enzyme, whereas its 5-*O*-shikimate and 5-*O*-quininate esters were actively converted into their corresponding caffeate conjugates (Fig. 1; Schoch et al., 2001). This observation suggested that the accepted version of the pathway at that time was likely to be incorrect and would require the activity of a shikimate/quininate acyltransferase to produce the *p*-coumarate esters. Finally, in 2003, this acyltransferase was identified in tobacco (*Nicotiana tabacum*) and shown to catalyze the production of *p*-coumarate esters of both shikimic and quinic acids from *p*-coumaroyl-CoA, and it was named *p*-HYDROXYCINNAMOYL-COENZYME A:QUINATE/SHIKIMATE *p*-HYDROXYCINNAMOYLTRANSFERASE (HCT; Hoffmann et al., 2003). Moreover, based on *in vitro* assays using the recombinant enzyme and caffeoyl quinate (chlorogenic acid) as substrate, HCT also was suggested to catalyze a second reaction in the pathway, the conversion of caffeate esters, which are products of C3H, into caffeoyl-CoA. No confirmation of this proposed role was ever made by means of reverse genetics, but the so-called second/reverse reaction of HCT was widely accepted and thought to be conserved among plant taxa. Recently, it was demonstrated that an enzyme, named CAFFEOYL SHIKIMATE ESTERASE (CSE), is able to catalyze the conversion of caffeoyl shikimate into caffeate in *Arabidopsis* (Vanholme et al., 2013c). The combined activity of CSE with that of *p*-coumaroyl/caffeoyl-CoA ligase (4CL), producing caffeoyl-CoA, would bypass the second/reverse reaction of HCT (Fig. 1). Genetic evidence for an active role of CSE in the lignification process in *Arabidopsis* was gained with a loss-of-function mutant showing reduced lignin content and an increased relative

<sup>1</sup> This work has been supported by Petrobras, the European Commission's Directorate General for Research within the 7th Framework Program [Grant 270089 (MULTIBIOPRO)], the Multidisciplinary Research Partnership Biotechnology for a Sustainable Economy (01MRB510W) of Ghent University, through the IWT-SBO project BIOLEUM (grant no. 130039) and the IWT-FISCH-SBO project ARBOREF (Grant 140894). H.K. and J.R. were funded by the DOE Great Lakes Bioenergy Research Center (DOE BER Office of Science DE-FC02-07ER64494). R.V. is indebted to the Research Foundation Flanders for a postdoctoral fellowship. I.C. was funded by the Foundation for Research of the State of São Paulo (FAPESP) via the BIOEN Young Investigators Awards research fellowship (grant 2015/02527-1). L.V., F.C.d.A.F., and M.d.L.S.S. were funded by the Program Ciência Sem Fronteiras for postdoctoral (Grants 248535/2013-7 and 206329/2014-8) and predoctoral (Grant 206863/2014-4) fellowships, respectively.

<sup>2</sup> These authors contributed equally to the article.

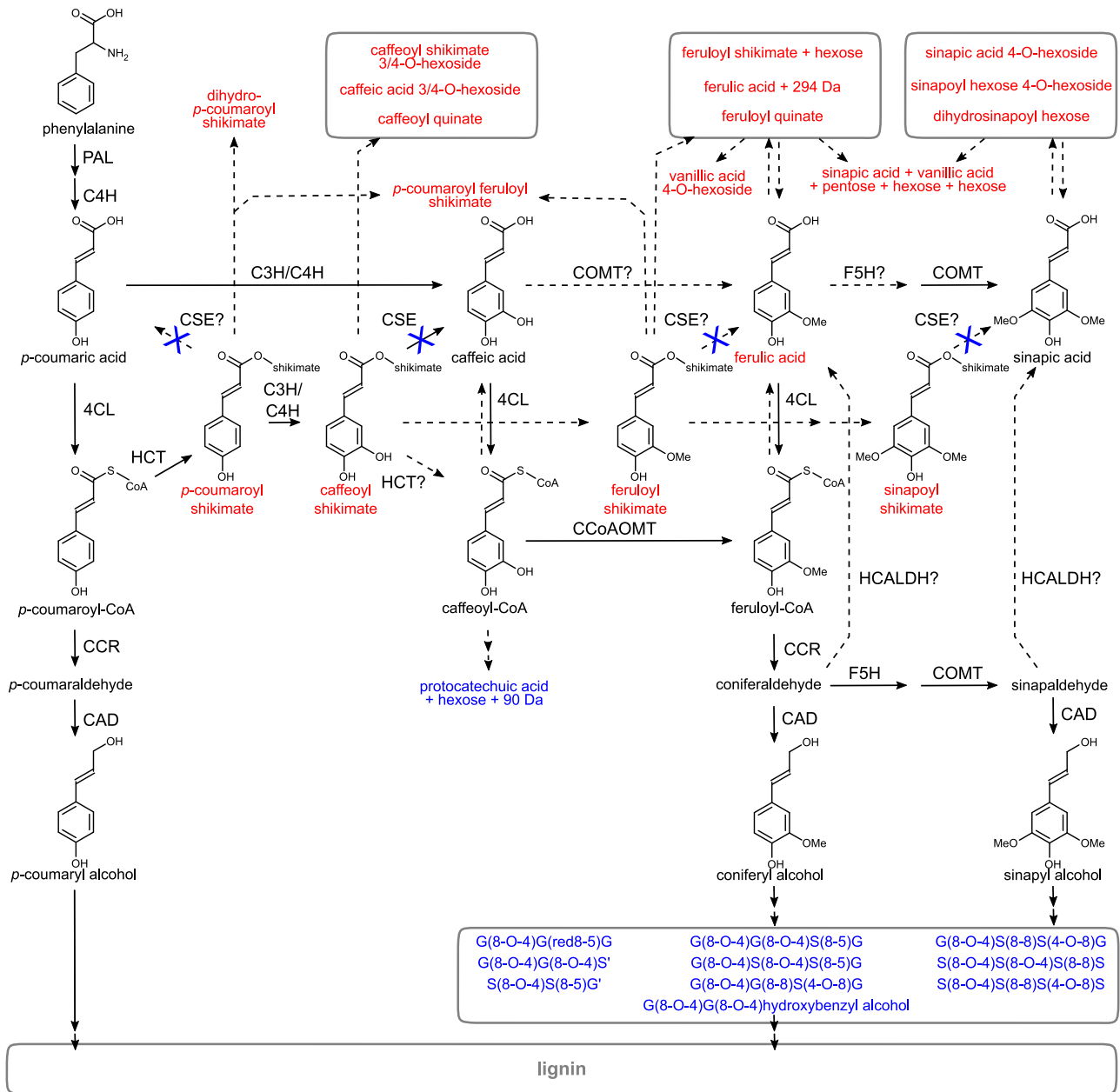
<sup>3</sup> Address correspondence to woboe@psb.vib-ugent.be.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors ([www.plantphysiol.org](http://www.plantphysiol.org)) is: Wout Boerjan (woboe@psb.vib-ugent.be).

I.C. performed experiments, analyzed data, and wrote the article with the contributions of all other authors; M.d.L.S.S. performed most of the experiments, analyzed data, and complemented the writing; L.V., R.V.A., F.C.d.A.F., A.P., W.V., J.N.J., D.P., and J.V.D. performed experiments and analyzed data; G.G. performed phenolic profiling and analyzed data; H.K. acquired and analyzed the NMR data; R.V. and J.R. analyzed data and complemented the writing; W.B. conceived the project and supervised and complemented the writing.

[OPEN] Articles can be viewed without a subscription.

[www.plantphysiol.org/cgi/doi/10.1104/pp.17.00920](http://www.plantphysiol.org/cgi/doi/10.1104/pp.17.00920)



**Figure 1.** Metabolic map of the general phenylpropanoid and monolignol-specific pathways showing the changes in phenolic metabolism upon *CSE* down-regulation in poplar. Metabolites found to have higher abundance in the *hpCSE* lines are highlighted in red, while those with decreased abundance are shown in blue. Metabolites belonging to the same class are framed with a box. Solid and dashed arrows represent enzymatic conversions validated by experimental evidence and suggested conversions, respectively (Van Acker et al., 2017). Two successive arrows represent two or more metabolic steps. PAL, PHENYLALANINE AMMONIA LYASE; C4H, CINNAMATE 4-HYDROXYLASE; 4CL, 4-COUMARATE:COENZYME A LIGASE; HCT, *p*-HYDROXYCINNAMOYL-COENZYME A:QUINATE/SHIKIMATE *p*-HYDROXYCINNAMOYLTRANSFERASE; C3H, *p*-COUMARATE 3'-HYDROXYLASE; CSE, CAFFEYOYL SHIKIMATE ESTERASE; CCoAOMT, CAFFEYOYL-COENZYME A O-METHYLTRANSFERASE; CCR, CINNAMOYL-COENZYME A REDUCTASE; F5H, FERULATE 5-HYDROXYLASE; COMT, CAFFEIC ACID O-METHYLTRANSFERASE; CAD, CINNAMYL ALCOHOL DEHYDROGENASE; HCALDH, HYDROXYCINNAMALDEHYDE DEHYDROGENASE.

proportion of H units in the lignin polymer (Vanholme et al., 2013c). Because *CSE* orthologs were found in the genomes of other plant species (Vanholme et al., 2013c; Carocha et al., 2015; Ha et al., 2016), it was first suggested that this enzymatic step was conserved within the plant

lineage. However, a subsequent study indicated that many unrelated species lack a bona fide *CSE* homolog in their genomes, suggesting that *CSE* might not be essential for lignification in all plants (Ha et al., 2016). Apart from *Arabidopsis*, the unambiguous confirmation of a role for

CSE in monolignol biosynthesis by means of reverse genetics was shown only in *Medicago truncatula*, for which CSE knockout resulted in strong negative effects on development, lower lignin levels, and preferential accumulation of H units in lignin (Ha et al., 2016).

A general role for CSE in lignin biosynthesis was recently questioned by a study reporting a predictive kinetic metabolic flux model that provides a mathematical description of the monolignol biosynthetic pathway in aspen/poplar (*Populus trichocarpa*) based on in vitro kinetics and absolute quantification of monolignol biosynthetic pathway proteins in the xylem (Wang et al., 2014). Although the *P. trichocarpa* genome harbors two bona fide CSE homologs, no CSE activity was detected in protein extracts from poplar secondary differentiating xylem, suggesting that caffeoyl shikimate is converted to caffeoyl-CoA by two isoforms of HCT in poplar (Wang et al., 2014). Moreover, CSE activity also was not detected in protein extracts of lignifying tissues of *Eucalyptus grandis*, switchgrass (*Panicum virgatum*), and rice (*Oryza sativa*), suggesting that CSE is not involved in the shikimate shunt in these species (Wang et al., 2014). Hence, to further assess a possible role for CSE in lignification in poplar (*Populus tremula* × *Populus alba*), we carried out a comprehensive study based on crude protein enzymatic activity assays, gene expression analysis, and multiple-level phenotyping of transgenic lines down-regulated in CSE. Moreover, we also evaluated the biotechnological potential of CSE by performing saccharification assays of wood samples derived from the CSE down-regulated lines.

## RESULTS

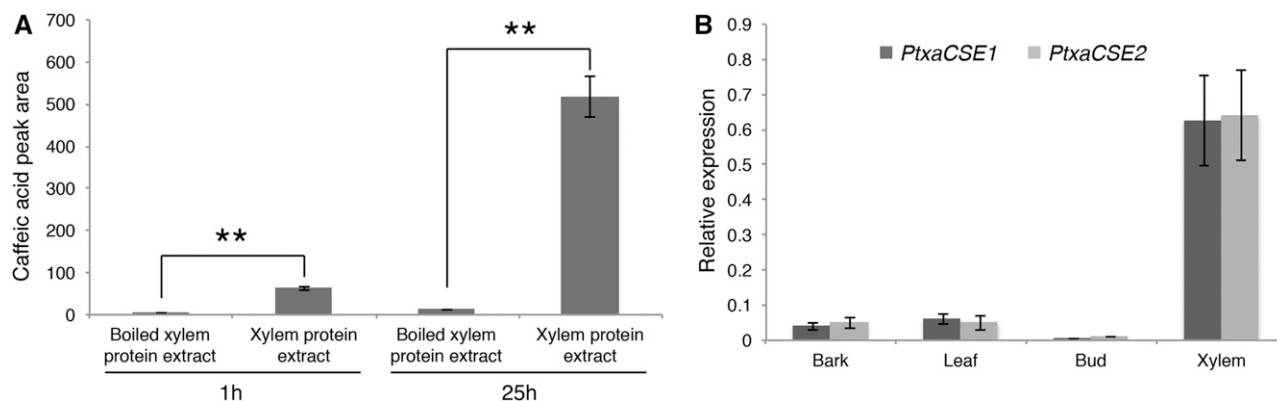
### Poplar Xylem Protein Extracts Show CSE Activity

The generality of a role for CSE in lignification was recently questioned based on the apparent lack of CSE activity in protein extracts of poplar xylem (Wang et al., 2014). Interestingly, with the exception of poplar, the

three other plant species analyzed previously (i.e. Arabidopsis, *M. truncatula*, and switchgrass) showed a positive correlation between the presence of bona fide CSE homologs in their genomes and CSE activity in protein extracts of their lignifying tissues (Vanholme et al., 2013c; Escamilla-Treviño et al., 2014; Wang et al., 2014; Ha et al., 2016). To further address this apparent dilemma, CSE activity was assayed by feeding crude protein extracts from scraped xylem tissue of wild-type *P. tremula* × *P. alba* plants with caffeoyl shikimate and analyzing the products by ultra-high-performance liquid chromatography-mass spectrometry (UHPLC-MS). Significant levels of caffeic acid were detected after 1 h of incubation, and an 8-fold increase in caffeic acid amount was detected after 25 h (Fig. 2A). Extremely low levels of caffeic acid were observed when incubations were performed with boiled xylem protein extracts, suggesting that the spontaneous hydrolysis of caffeoyl shikimate into caffeic acid is negligible. These results support the presence of CSE activity in the poplar xylem.

### The Poplar Genome Harbors Two Bona Fide CSE Homologs That Are Highly Expressed in the Xylem

A previous phylogenetic analysis of the CSE protein family in a set of different plant species has shown that these proteins group into two classes: class I, which potentially includes bona fide CSE proteins; and class II, which includes CSE homologs with low sequence identity and, therefore, are considered to be CSE-like proteins (Ha et al., 2016). In the genome of *P. trichocarpa*, two bona fide CSE homologs are found, *PttrCSE1* (Potri.001G175000) and *PttrCSE2* (Potri.003G059200), sharing 91% identity in amino acid sequence to each other and around 80% to AtCSE of Arabidopsis. Because not all family members for a given step of the monolignol biosynthetic pathway are necessarily involved in lignification, tissue-specific expression analysis was performed to evaluate whether the expression pattern of each CSE homolog in poplar



**Figure 2.** Poplar xylem CSE activity and CSE gene expression. A, CSE activity in crude poplar xylem protein extracts. Total poplar xylem protein was added to reaction buffer containing 100  $\mu\text{M}$  caffeoyl shikimate, and caffeate was measured via UHPLC-MS after 1 and 25 h. Boiled poplar xylem protein was used as a negative control. Error bars indicate se. Differences in caffeic acid abundance were assessed with Student's *t* test (\*\*,  $0.01 > P > 0.001$ ;  $n = 4$ ). B, Expression analysis of CSE homologs in different tissues of poplar as determined via RT-quantitative PCR (qPCR). Error bars indicate se ( $n = 4$ ).

correlates with developmental lignification. For such a purpose, the transcript abundance of both genes was determined in four *P. tremula* × *P. alba* tissues differing in their degree of lignification: bark, mature leaves, axillary buds, and developing xylem. Preferential expression in the xylem, a tissue that undergoes early and strong lignification, is expected for a gene involved in developmental lignification. Accordingly, both *PtxaCSE* genes were strongly and preferentially expressed in the xylem, whereas much lower transcript levels were detected in bark, leaves, and especially axillary buds (Fig. 2B). These results suggest a potential, and possibly redundant, role for the two *CSE* paralogues in poplar developmental lignification.

### Down-Regulation of *CSE* in Poplar Did Not Drastically Affect Plant Development

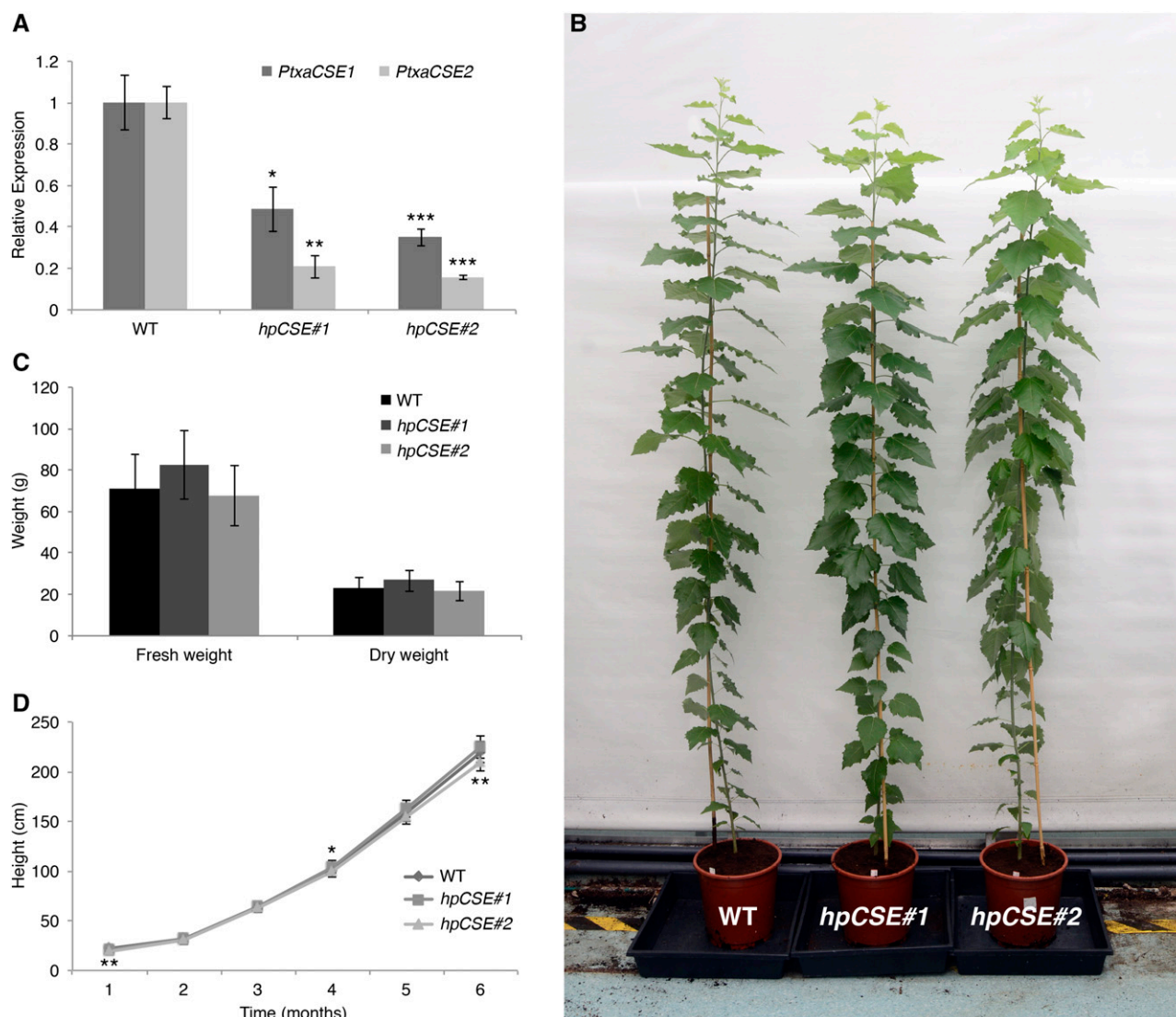
In order to gain genetic evidence for the involvement of *CSE* in the lignification process in poplar, constitutive down-regulation of both *CSE* genes was targeted by cloning a 120-bp fragment of *PtxaCSE2* (corresponding to Potri.003G059200) into the pK7GWIWG2II gene-silencing vector to produce cauliflower mosaic virus 35S promoter-driven *hpCSE* transgenic *P. tremula* × *P. alba* lines. The corresponding fragment in *PtxaCSE1* shared 90% identity with the cloned fragment of *PtxaCSE2*. Independent lines (137) were grown for 45 d in the greenhouse before screening. Of note, none of the transgenic lines exhibited the typical red xylem phenotype that is visible upon debarking the stems of poplars down-regulated for *CINNAMOYL-COENZYME A REDUCTASE (CCR)*, *CAFFEIC ACID O-METHYLTRANSFERASE (COMT)*, or *CINNAMYL ALCOHOL DEHYDROGENASE* (Van Doorselaere et al., 1995; Baucher et al., 1996; Tsai et al., 1998; Lep le et al., 2007). The *hpCSE* lines were screened for the accumulation of caffeoyl shikimate, the substrate for *CSE*, via UHPLC-MS after extraction of methanol-soluble phenolics from the stem. Among all analyzed lines, one particular line showed a 36-fold increase in the caffeoyl shikimate content relative to the wild type, whereas the second highest amount of caffeoyl shikimate was 20-fold higher than that of the wild type. In the other lines, the caffeoyl shikimate content decreased to as low as wild-type levels, suggesting lower silencing efficiency or complete absence of silencing. A significant increase in H unit content and a reduction in total lignin amount are two other hallmarks of *CSE* silencing in *Arabidopsis* and *M. truncatula*. Therefore, relative levels of H units were determined via thioacidolysis, and lignin amount was determined via acetyl bromide lignin measurements in the top 20 lines with the highest accumulation of caffeoyl shikimate. Although the relative H unit content in lignin and the lignin amount in the transgenic lines did not differ drastically from those of the wild type, these data were taken into account to select the two lines with the highest accumulation of both caffeoyl shikimate and H units and with a reduction in lignin (Supplemental Table S1). These selected lines were named *hpCSE#1* and *hpCSE#2* and were subjected to further multiple-level phenotyping.

Quantitative reverse transcription (RT)-PCR was used to determine the residual transcript levels of *PtxaCSE1* and *PtxaCSE2* in the developing xylem of both wild-type and *hpCSE* lines (Fig. 3A). This analysis showed that the two *CSE* paralogues were differentially down-regulated in both *hpCSE* lines. For *PtxaCSE2*, 21% and 15% of residual transcript levels were observed for *hpCSE#1* and *hpCSE#2* compared with the wild type, whereas a much weaker down-regulation was observed for *PtxaCSE1*, for which 49% and 35% of residual transcript levels were found in those lines (Fig. 3A). To evaluate whether down-regulation of *CSE* caused any effect on plant growth and development, the *hpCSE* lines were grown alongside the wild type in the greenhouse for a period of 6 months. Plant height was followed weekly, and by the end of the growth period, the trees were harvested and biomass parameters were determined. The *hpCSE* lines were morphologically indistinguishable from the wild type (Fig. 3B), and no differences in either fresh or dry weight were observed after 6 months of growth (Fig. 3C). Stem height was similar for both lines compared with the wild type during the whole growth period, but after 6 months, the *hpCSE#2* plants were slightly shorter (4%) than the controls (Fig. 3D). These results suggest that these *CSE* down-regulated poplars develop normally with no severe effects on biomass accumulation.

### Down-Regulation of *CSE* Affects Lignin Content and Structure in Poplar

Silencing of a gene encoding an enzyme central to the lignin biosynthetic pathway is expected to affect lignin content and/or composition. To evaluate the impact of *CSE* down-regulation on lignin deposition, crude cell wall residue (CWR) was prepared from the stem of 6-month-old *hpCSE* and wild-type poplars, and the lignin content was determined by the Klason method. Both *hpCSE* lines showed significant reduction in (acid-insoluble) Klason lignin content, with decreases of 19% and 25% for *hpCSE#1* and *hpCSE#2* compared with that of the wild type (Fig. 4A; Supplemental Table S2). These reductions were accompanied by mild increases in cellulose content of 8% and 13% in *hpCSE#1* and *hpCSE#2* (Fig. 4B; Supplemental Table S2). Next, lignin structural changes were first evaluated by thioacidolysis, an analytical method that quantifies only lignin units linked by labile  $\beta$ -O-4 bonds. Although the absolute amount of H units remained low, the relative proportion of H-derived monomers increased by 51% in *hpCSE#1* and by 113% in *hpCSE#2* (Fig. 4C; Supplemental Table S2). The relative frequencies of G- and S-derived monomers showed only mild shifts compared with control plants. The relative amount of thioacidolysis-released G units increased by 7% and 5%, whereas the relative amount of thioacidolysis-released S units decreased by 4% and 3%, in *hpCSE#1* and *hpCSE#2* (Fig. 4D; Supplemental Table S2). These changes led to a mild but significant decrease in the S/G ratio from 2.06 in the wild type to 1.85 and 1.9 in



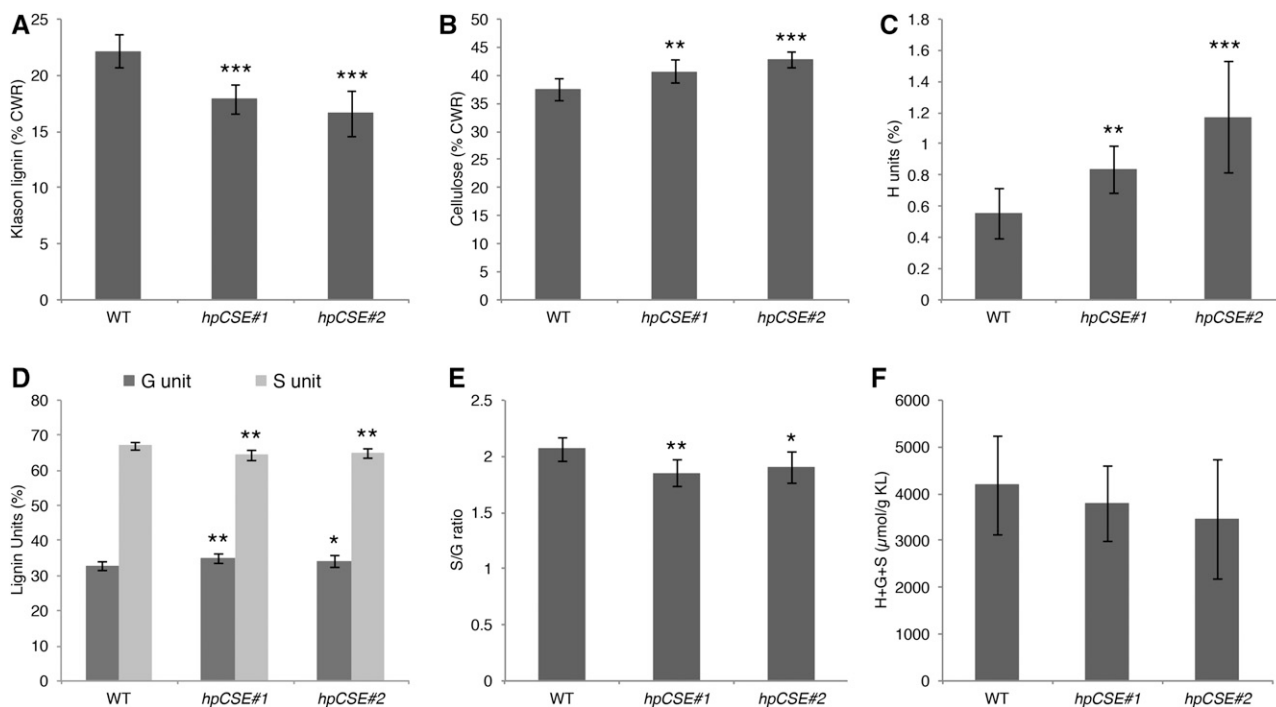


**Figure 3.** Residual expression of both *CSE* genes and growth analyses of poplar trees down-regulated in *CSE*. A, Residual expression levels of *PtxaCSE1* and *PtxaCSE2* in the xylem of wild-type (WT) poplar plants and the *hpCSE* lines determined by RT-qPCR. The expression of each *PtxaCSE* gene was normalized to that of the wild type. Error bars indicate SE. Differences in gene expression were assessed with Student's *t* test (\*,  $0.05 > P > 0.01$ ; \*\*,  $0.01 > P > 0.001$ ; and \*\*\*,  $P < 0.001$ ;  $n = 4$ ). B, Photograph of representative plants grown for 6 months in the greenhouse. C, Fresh and dry weights of stems harvested from wild-type and *hpCSE* plants at the end of the 6-month growth period. No statistical difference was found among the genotypes. Error bars indicate SD. D, Growth curves of the wild type and the *hpCSE* lines. Height was monitored every week for a period of 6 months. The average growth in each month was used to plot the graph. Error bars indicate SD. Differences in growth parameters between the wild type and the transgenic lines were assessed with Student's *t* test (\*,  $0.05 > P > 0.01$  and \*\*,  $0.01 > P > 0.001$ ; wild type,  $n = 25$ ; *hpCSE#1*,  $n = 23$ ; and *hpCSE#2*,  $n = 23$ ). The asterisk above the data point represents a statistical difference for line *hpCSE#1*, whereas asterisks below represent statistical differences for line *hpCSE#2*.

*hpCSE#1* and *hpCSE#2* (Fig. 4E; Supplemental Table S2). The degree of lignin condensation is inversely correlated with the thioacidolysis yield (i.e. the sum of released H, G, and S monomers). When expressed on a Klason lignin content basis, the thioacidolysis yield of the *hpCSE* lines was not significantly different from that of the wild type (Fig. 4F; Supplemental Table S2).

In order to gain further insight into the structural changes caused by *CSE* down-regulation, 2D NMR was

performed on enzyme lignins from the stems of the wild type and the *hpCSE* lines. By analyzing the aromatic regions of the two-dimensional  $^1\text{H}$ - $^{13}\text{C}$  heteronuclear single-quantum coherence (HSQC) spectra, it is possible to visualize differences in lignin monomeric composition irrespective of the interunit linkage distribution, in contrast to thioacidolysis, in which only units linked by ether bonds are determined. In general, the NMR data largely corroborated the structural



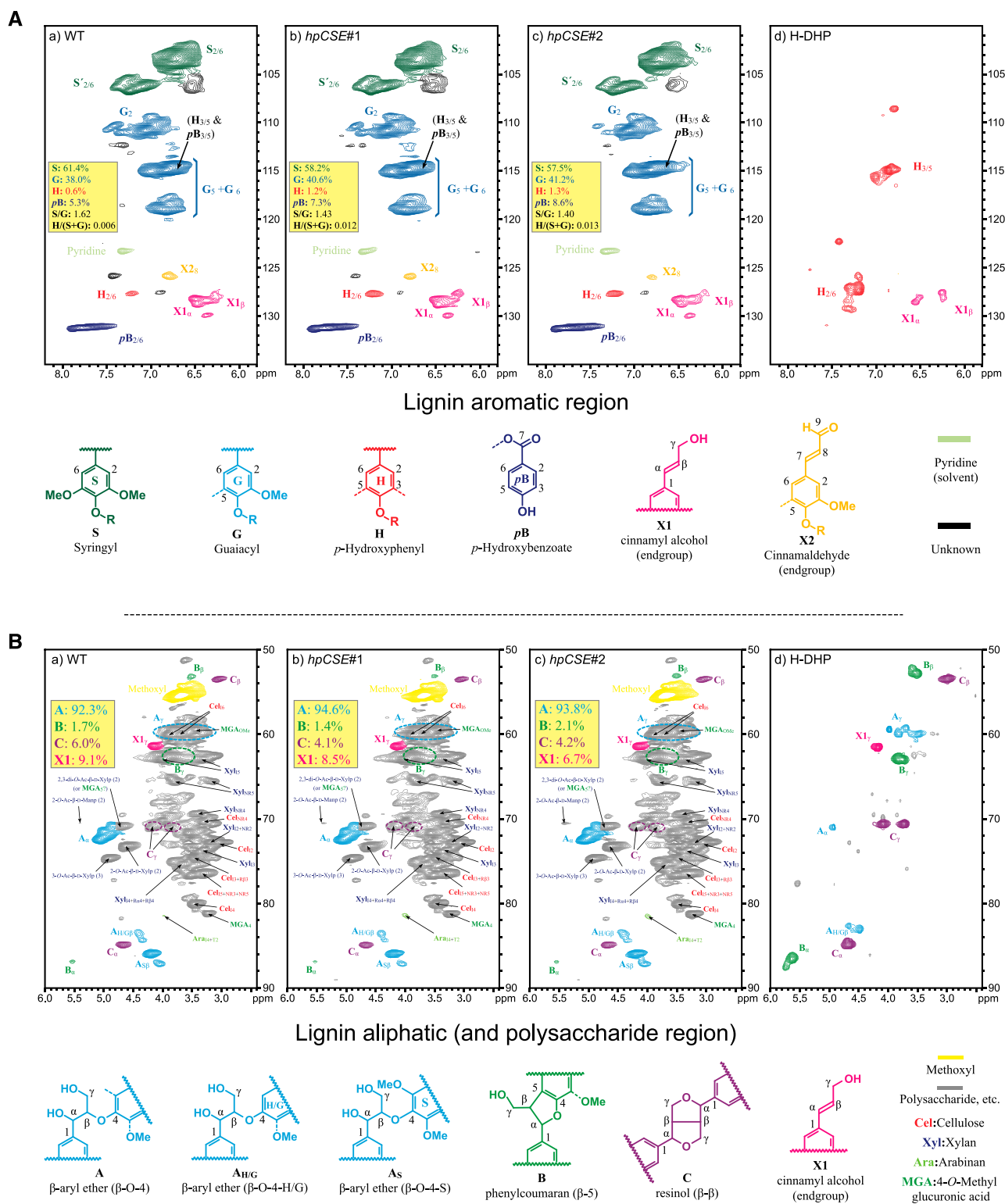
**Figure 4.** Lignin content and composition and cellulose content of stems from *hpCSE* lines and control plants. A, Total acid-insoluble lignin content determined by the Klason method. Data are expressed as percentage of CWR. B, Cellulose content was determined gravimetrically and is expressed as percentage of CWR. C, Relative levels of releasable H monomers determined by thioacidolysis. D, Relative levels of releasable G and S monomers determined by thioacidolysis. E, S/G ratio determined based on thioacidolysis data. F, Total thioacidolysis yield as determined from the sum of released H, G, and S monomers, expressed in  $\mu\text{mol g}^{-1}$  Klason lignin. For all analyses, data are means  $\pm$  SD ( $n = 10$ ), except for thioacidolysis yield, for which SE values are shown. Differences in cell wall parameters were assessed with Student's *t* test (\*,  $0.05 > P > 0.01$ ; \*\*,  $0.01 > P > 0.001$ ; and \*\*\*,  $P < 0.001$ ). WT, Wild type.

changes found with thioacidolysis, with both lines showing similar spectra. Compared with the wild type, the proportion of H units was increased by 100% and 110% in *hpCSE#1* and *hpCSE#2* (Fig. 5A). The relative fraction of G units was increased by 6.8% and 8.4%, whereas the relative frequency of S units was decreased by 5.2% and 6.3%, in each of these lines. Accordingly, the S/G ratio decreased significantly from 1.62 in the wild type to 1.43 and 1.40 in *hpCSE#1* and *hpCSE#2* (Fig. 5). In addition, the interunit linkage types can be deduced from the oxygenated aliphatic region of the HSQC spectrum (Fig. 5B). Minor shifts in the relative proportion of the major unit types were observed, with a slight increase in  $\beta$ -aryl ether ( $\beta$ -O-4) units A and a slight decrease in resinol ( $\beta$ - $\beta$ ) units C in both *hpCSE* lines (Fig. 5B), suggesting a lower proportion of carbon-carbon bonds in their lignin polymers. Altogether, these results show that down-regulation of *CSE* in poplar not only leads to reduced lignin levels but also affects lignin composition and structure, albeit mildly.

#### The Phenolic Profile Is Altered upon *CSE* Down-Regulation in Poplar

A shift in the phenolic metabolism is expected upon blocking a biosynthetic step central to the lignin

pathway. In order to assess the consequences of *CSE* down-regulation on phenolic metabolism, methanol-soluble metabolites were extracted from the developing xylem of 3.5-month-old *hpCSE* and wild-type plants and analyzed via UHPLC-MS. This procedure enables the detection and relative quantification of different classes of aromatic compounds, including intermediates and derivatives from the phenylpropanoid pathway, such as flavonoids, benzenoids, cinnamates, and products from monolignol coupling (oligolignols). A total of 2,443 peaks could be detected, aligned, and integrated in the chromatograms of wild-type, *hpCSE#1*, and *hpCSE#2* samples. Only peaks present in all replicates of at least one genotype and that were on average above a 500-count cutoff in at least one genotype were selected for statistical analysis. Some 618 peaks met these criteria. Principal component analysis (PCA) showed that the phenolic profile of the *hpCSE* lines clearly separated from that of the wild type, indicating major metabolic differences (Supplemental Fig. S1). To further select the metabolites for which the abundance was most affected by the down-regulation of *CSE*, filters were applied based on a significance level of 0.01 and a fold change of at least 2 between the *hpCSE* and wild-type samples. Accordingly, the 68 peaks with significantly higher abundance in both *hpCSE* lines could be assigned to 42 compounds, of which 26 could



**Figure 5.** Structural characterization of lignin by NMR. HSQC spectra are shown for the aromatic regions (A) and the oxygenated aliphatic regions (B) of whole cell walls from stems of the wild type (WT) and the two *hpCSE* lines, along with a synthetic lignin synthesized biomimetically from *p*-coumaryl alcohol to help validate the H unit assignments (H-DHP). Integrated values for each monomeric H, G, and S unit and the  $\alpha$ -C/H correlation peaks from the major lignin interunit structures A to C are provided. The colors of the substructures shown match those of the corresponding signals in the HSQC spectra (where they are resolved).



be structurally characterized (Fig. 1; Table I). Several of these compounds were observed only in the transgenic lines, being below the detection limit in wild-type samples. The identified compounds belonged to the metabolic classes of benzenoids and hydroxycinnamates, of which caffeate-, ferulate-, and sinapate-derived compounds were most abundant. As expected, one of the top-accumulating compounds was the CSE substrate caffeoyl shikimate (two free and two hexosylated forms), which is consistent with a reduced CSE expression. In addition, *p*-coumaroyl shikimate, caffeoyl quinate, feruloyl shikimate, feruloyl quinate, and sinapoyl shikimate also accumulated in the *hpCSE* lines. Among these 26 identified compounds with higher abundance in the *hpCSE* lines, a total of 12 metabolites were hexosylated, presumably to reduce their toxicity and allow for their sequestration in the vacuole (Dima et al., 2015).

The 26 peaks with significantly lower abundance (average relative abundance < 0.5) in the *hpCSE* lines compared with the wild type could be assigned to 26 compounds, of which 14 could be structurally characterized (Fig. 1; Table II). Interestingly, with the exception of hexosylated protocatechuic acid, all the metabolites with reduced abundance were oligolignols containing G and S units or their corresponding hydroxycinnamaldehyde analogs, in agreement with the reduced lignin deposition upon CSE down-regulation. Altogether, the metabolic alterations observed in the xylem of CSE down-regulated poplars (e.g. accumulation of caffeoyl shikimate and reduced amounts of oligolignols) suggest that the function of CSE in the lignin pathway is conserved in poplar.

### CSE Down-Regulation Improves Saccharification Efficiency of Poplar Stems

Previously, we showed that CSE loss of function in Arabidopsis resulted in a 4-fold increase in cellulose-to-Glc conversion of senesced stems when compared with the wild type (Vanholme et al., 2013c), one of the highest improvements in saccharification efficiency without biomass pretreatment ever reported. In order to evaluate whether the cell wall changes caused by CSE down-regulation also lead to improved saccharification efficiency in poplar, cellulose-to-Glc conversion of the wood from the *hpCSE* lines was compared in limited-saccharification assays with that of the wild type. Saccharification assays were performed with debarked stems without and with acid (1 M HCl, 80°C, 2 h) or alkaline (62.5 mM NaOH, 90°C, 3 h) pretreatment, with Glc release measured after 3, 6, 24, 30, and 48 h. The cellulose-to-Glc conversion was calculated based on the amount of Glc released upon saccharification, and the cellulose content was measured for each sample (Fig. 6; Supplemental Table S3). When no pretreatment was applied, the cellulose-to-Glc conversion of line *hpCSE#1* was significantly higher compared with that of the wild type already after 6 h of saccharification, whereas no significant differences were observed for *hpCSE#2* (Fig. 6A). After 48 h, the cellulose-to-Glc

conversion from the stems of line *hpCSE#1* was 31% higher than that of the wild type (Fig. 6A; Supplemental Table S3). Both lines performed better than the wild type in saccharification assays when the pretreatments were performed, and the differences were significant at all analyzed time points. The cellulose-to-Glc conversion after 48 h was increased by 48% and 46% in *hpCSE#1* and *hpCSE#2* lines when an acid pretreatment was applied (Fig. 6C; Supplemental Table S3), whereas alkaline pretreatment resulted in an increase of 51% and 40% compared with the wild type (Fig. 6E; Supplemental Table S3).

In Arabidopsis, the inflorescence stems of the *cse2* null mutant are 37% smaller and 42% lighter at senescence than those of the wild type, revealing a substantial yield penalty upon CSE loss of function. Because down-regulation of CSE did not affect biomass accumulation in poplar, and because the *hpCSE* lines have slightly more cellulose, we anticipated that the saccharification yield on a plant basis would be increased substantially in the *hpCSE* lines. Accordingly, we compared the total Glc release per plant among all genotypes with and without pretreatment. Similarly to the cellulose-to-Glc conversion results, only the *hpCSE#1* line showed increased Glc release on a plant basis without pretreatment, releasing 62% more than the wild type (Fig. 6B). Additionally, *hpCSE#1* and *hpCSE#2* lines released 86% and 57% more Glc than the wild type on a plant basis when acid pretreatment was applied (Fig. 6D), whereas the improvements were of 91% and 50% with alkaline pretreatment (Fig. 6F). These results show that recalcitrance is reduced and Glc release is improved in poplar stems down-regulated for CSE.

## DISCUSSION

### CSE Is Important for Lignification in Poplar

The identification of CSE as an enzyme central to the lignin biosynthetic pathway in Arabidopsis led to a major question on whether this catalytic step is conserved across plant lineages. The fact that many unrelated plant species lack a bona fide CSE homolog in their genomes suggests that CSE might not be essential for lignification in all plants. Interestingly, from the plant species analyzed to date, the presence of bona fide CSE homologs in the genome was positively correlated with the detection of CSE activity in protein extracts prepared from lignifying tissues. Weak esterase activity was detected in protein preparations from stems of *Brachypodium distachyon* and maize (*Zea mays*), whose genomes do not possess a bona fide CSE gene, whereas high CSE activity was found in Arabidopsis and switchgrass, which harbor CSE homologs (Vanholme et al., 2013c; Ha et al., 2016). The exception was *P. trichocarpa*, for which no CSE enzymatic activity could be detected in extracts prepared from developing secondary xylem (Wang et al., 2014), even though its genome possesses two bona fide CSE homologs. These results led those authors to conclude that there was no

**Table 1.** List of compounds for which the relative abundance was significantly higher in both hpCSE lines compared with the wild type

For each metabolite, a unique number (No.), mass-to-charge ratio (*m/z*), retention time (RT), peak area (as mean  $\pm$  sd), and ratio of the peak area in the hpCSE lines and the wild type are given. For the structural elucidation of the compounds, see Supplemental Figure S2.

No.	<i>m/z</i>	RT <i>min</i>	Name	Wild Type	hpCSE#1	Ratio hpCSE#1:Wild Type	hpCSE#2	Ratio hpCSE#2:Wild Type
1	319.0833	7.43	<i>p</i> -Coumaric acid derivatives	0 $\pm$ 0	482 $\pm$ 545	>100	805 $\pm$ 741	>100
2	321.0679	3.28	<i>p</i> -Coumaroyl shikimate	2,049 $\pm$ 513	6,785 $\pm$ 1,213	3.3	6,514 $\pm$ 817	3.2
3	495.1314	18.32	Dihydro- <i>p</i> -coumaroyl shikimate	0 $\pm$ 0	1,745 $\pm$ 1,207	>100	3,623 $\pm$ 3,388	>100
4	341.0895	4.10	<i>p</i> -Coumaroyl feruloyl shikimate	2,628 $\pm$ 334	6,561 $\pm$ 1,431	2.5	6,785 $\pm$ 1,193	2.6
5	341.0925	5.57	Caffeic acid 3/4- <i>O</i> -hexoside 1	248 $\pm$ 104	764 $\pm$ 221	3.1	778 $\pm$ 247	3.1
6	335.0772	6.23	Caffeic acid 3/4- <i>O</i> -hexoside 2	300 $\pm$ 206	24,163 $\pm$ 8,550	80	26,146 $\pm$ 17,480	87
7	335.0780	7.72	Caffeoyl shikimate 1	0 $\pm$ 0	875 $\pm$ 487	>100	1,152 $\pm$ 1,116	>100
8	497.1312	3.75	Caffeoyl shikimate 2	0 $\pm$ 0	831 $\pm$ 687	>100	1,330 $\pm$ 754	>100
9	497.1313	5.17	Caffeoyl shikimate 3/4- <i>O</i> -hexoside 1	0 $\pm$ 0	2,309 $\pm$ 890	>100	2,687 $\pm$ 821	>100
10	353.0887	4.28	Caffeoyl shikimate 3/4- <i>O</i> -hexoside 2	83 $\pm$ 115	1,475 $\pm$ 815	18	1,131 $\pm$ 779	14
11	329.0877	2.29	Caffeoyl quininate (chlorogenic acid)	7,748 $\pm$ 796	25,854 $\pm$ 3,977	3.3	25,260 $\pm$ 2,298	3.3
12	193.0506	8.40	Ferulic acid derivatives	0 $\pm$ 0	575 $\pm$ 729	>100	669 $\pm$ 669	>100
13	349.0941	8.03	Vanillic acid 4- <i>O</i> -hexoside	0 $\pm$ 0	737 $\pm$ 838	>100	964 $\pm$ 905	>100
14	349.0938	8.21	Feruloyl shikimate 1	6 $\pm$ 15	5,015 $\pm$ 4,305	>100	5,646 $\pm$ 3,729	>100
15	349.0928	9.17	Feruloyl shikimate 2	759 $\pm$ 691	30,521 $\pm$ 19,272	40	31,096 $\pm$ 20,848	41
16	349.0937	10.27	Feruloyl shikimate 3	118 $\pm$ 167	8,590 $\pm$ 5,677	73	10,284 $\pm$ 8,429	87
17	511.1467	4.64	Feruloyl shikimate 4	0 $\pm$ 0	2,243 $\pm$ 1,507	>100	4,763 $\pm$ 1,064	>100
18	511.1474	7.04	Feruloyl shikimate + hexose 1	0 $\pm$ 0	265 $\pm$ 478	>100	1,148 $\pm$ 336	>100
19	367.1048	6.80	Feruloyl shikimate + hexose 2	0 $\pm$ 0	1,110 $\pm$ 1,118	>100	922 $\pm$ 1,077	>100
20	487.1473	5.68	Feruloyl quininate	65 $\pm$ 111	696 $\pm$ 459	11	845 $\pm$ 614	13.1
21	385.1137	4.48	Ferulic acid + 294 D	1,064 $\pm$ 571	30,844 $\pm$ 4,722	29	35,447 $\pm$ 2,908	33
22	385.1145	6.18	Sinapic acid derivatives	190 $\pm$ 140	10,965 $\pm$ 2,100	58	15,412 $\pm$ 1,665	81
23	593.1739	3.10	Sinapic acid 4- <i>O</i> -hexoside 1	0 $\pm$ 0	1,354 $\pm$ 666	>100	2,196 $\pm$ 1,648	>100
24	815.2274	7.42	Sinapoyl hexose 4- <i>O</i> -hexoside (formic acid adduct)	0 $\pm$ 0	2,153 $\pm$ 760	>100	2,834 $\pm$ 966	>100
25	379.1044	9.11	Sinapic acid + vanillic acid + pentose + hexose + hexose	6 $\pm$ 15	4,335 $\pm$ 4,372	>100	5,620 $\pm$ 3,167	>100
26	387.1303	5.36	Sinapoyl shikimate	0 $\pm$ 0	625 $\pm$ 335	>100	978 $\pm$ 220	>100
			Dihydrosinapoyl hexose					

**Table II.** List of compounds for which the relative abundance was significantly lower in both *hpCSE* lines compared with the wild type

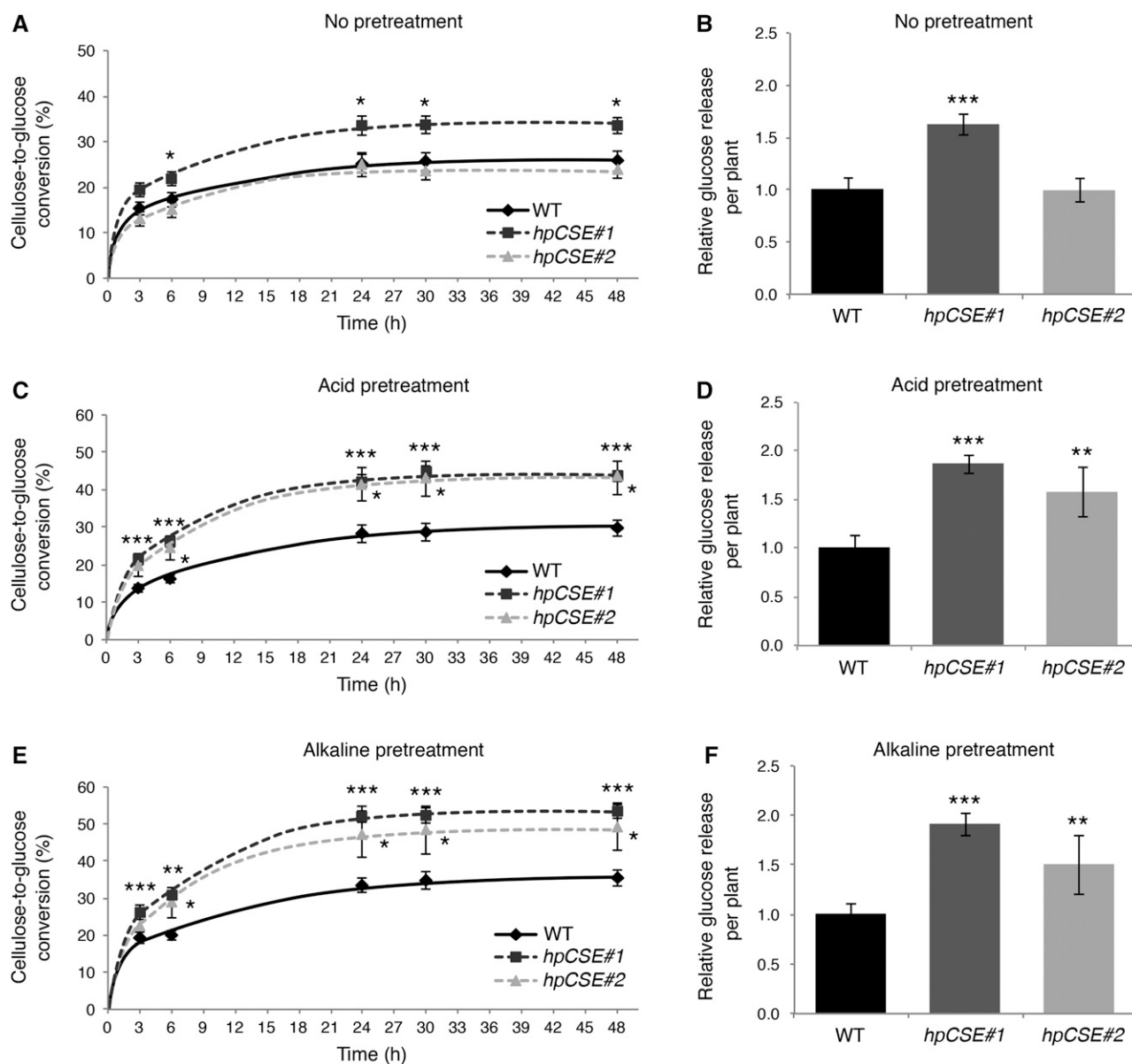
For each metabolite, a unique number (No.), mass-to-charge ratio (*m/z*), retention time (RT), peak area (as mean  $\pm$  sd), and ratio of the peak area in the *hpCSE* lines and the wild type are given. For the structural elucidation of the compounds, see Supplemental Figure S2.

No.	<i>m/z</i>	RT <i>min</i>	Name	Wild Type	<i>hpCSE</i> #1	Ratio <i>hpCSE</i> #1:Wild Type	<i>hpCSE</i> #2	Ratio <i>hpCSE</i> #2:Wild Type
27	405.1201	12.63	Protocatechuic acid + hexose + 90 D	732 $\pm$ 353	3 $\pm$ 10	0.005	2 $\pm$ 6	0.003
28	555.1898	15.41	G( $\beta$ -O-4)C(red $\beta$ -5)G	757 $\pm$ 157	103 $\pm$ 108	0.136	28 $\pm$ 38	0.037
29	599.2138	14.04	G( $\beta$ -O-4)G( $\beta$ -O-4)S' 1	5,568 $\pm$ 339	2,757 $\pm$ 760	0.495	1,389 $\pm$ 723	0.249
30	599.2142	14.86	G( $\beta$ -O-4)G( $\beta$ -O-4)S' 2	1,306 $\pm$ 395	608 $\pm$ 258	0.465	241 $\pm$ 211	0.185
31	515.1939	7.01	G( $\beta$ -O-4)C( $\beta$ -O-4)Hydroxybenzyl alcohol 1	1,778 $\pm$ 381	147 $\pm$ 147	0.083	28 $\pm$ 45	0.016
32	515.1940	7.47	G( $\beta$ -O-4)C( $\beta$ -O-4)Hydroxybenzyl alcohol 2	2,051 $\pm$ 319	340 $\pm$ 282	0.166	106 $\pm$ 138	0.052
33	611.2146	16.67	S( $\beta$ -O-4)S( $\beta$ -5)G'	5,672 $\pm$ 702	2,550 $\pm$ 328	0.450	2,385 $\pm$ 974	0.421
34	779.2940	15.19	G( $\beta$ -O-4)C( $\beta$ -O-4)S( $\beta$ -5)G 1	2,090 $\pm$ 881	273 $\pm$ 259	0.130	26 $\pm$ 48	0.012
35	779.2939	15.30	G( $\beta$ -O-4)C( $\beta$ -O-4)S( $\beta$ -5)G 2	601 $\pm$ 203	27 $\pm$ 66	0.045	8 $\pm$ 23	0.013
36	779.2933	17.26	G( $\beta$ -O-4)G( $\beta$ -O-4)S(4-O- $\beta$ )G	2,677 $\pm$ 770	976 $\pm$ 359	0.364	1,032 $\pm$ 896	0.385
37	809.3044	16.74	G( $\beta$ -O-4)S( $\beta$ -O-4)S( $\beta$ -5)G	2,365 $\pm$ 624	734 $\pm$ 630	0.310	10 $\pm$ 29	0.004
38	839.3155	18.64	G( $\beta$ -O-4)S( $\beta$ -O-4)S(4-O- $\beta$ )G	3,899 $\pm$ 353	1,561 $\pm$ 414	0.400	1,533 $\pm$ 614	0.393
39	869.3261	17.80	S( $\beta$ -O-4)S( $\beta$ -O-4)S( $\beta$ - $\beta$ )S	662 $\pm$ 265	29 $\pm$ 54	0.044	64 $\pm$ 121	0.097
40	869.3256	17.90	S( $\beta$ -O-4)S( $\beta$ - $\beta$ )S(4-O- $\beta$ )S	1,207 $\pm$ 326	81 $\pm$ 135	0.067	133 $\pm$ 227	0.111

direct evidence for a role for CSE in lignification in poplar and that caffeoyl shikimate is likely converted to caffeoyl-CoA by two xylem-specific HCT isoforms (Wang et al., 2014). However, the apparent lack of CSE activity in protein extracts from *P. trichocarpa* xylem is difficult to reconcile with the recent observation that both recombinant CSEs from *Populus deltoides*, which share 98% amino acid identity with their orthologs from *P. trichocarpa*, show high esterase activity in vitro (Ha et al., 2016). In addition, CSE activity in lignifying tissues was not detected in switchgrass, whereas a previous study had already shown that protein extracts prepared from switchgrass stems produce caffeic acid, instead of caffeoyl-CoA, when incubated with caffeoyl shikimate and CoA (Escamilla-Treviño et al., 2014). Subsequent work from the same group confirmed high CSE activity in crude extracts from switchgrass stems (Ha et al., 2016). Here, we showed not only that both poplar CSE homologs are preferentially expressed in developing xylem but also that protein preparations from this tissue show CSE activity, providing supporting evidence for a role for CSE in lignification in poplar.

The reactions of CSE need to be augmented by an additional 4CL step to regenerate the activated caffeoyl-CoA from the CSE product, caffeic acid. As an additional molecule of ATP is required, in contrast to the transesterification reaction catalyzed by HCT, the involvement of CSE in monolignol biosynthesis might seem counterintuitive. Nevertheless, this second/reverse reaction of HCT has not yet been supported by genetic evidence, because *HCT* loss of function also affects the conversion of *p*-coumaroyl-CoA to *p*-coumaroyl-shikimate by HCT upstream of the pathway. In addition, kinetic analyses of recombinant enzymes have shown that, although poplar xylem-specific HCT isoforms can convert caffeoyl shikimate into caffeoyl-CoA, the catalytic efficiency of this reaction is 20-fold lower than that of the forward acyl-transferase HCT reaction (i.e. the formation of caffeoyl shikimate from caffeoyl-CoA and shikimate; Wang et al., 2014). Similar results were found with recombinant HCT enzymes from switchgrass, which show preferences of around 15-fold for the forward compared with the reverse reaction (Escamilla-Treviño et al., 2014). When Hoffmann et al. (2003) first described and characterized HCT in tobacco, the conversion of caffeoyl-CoA to caffeoyl shikimate also was shown to be the most efficient reaction performed by the recombinant enzyme. Therefore, it seems counterintuitive that the reverse transesterification reaction of HCT has an important role in vivo, as the production of caffeoyl shikimate from caffeoyl-CoA would be highly favored. These observations suggest that caffeoyl-CoA is produced from caffeoyl shikimate by a two-step reaction catalyzed by CSE and 4CL instead of the single HCT-dependent transesterification reaction, at least in plant species harboring a bona fide CSE homolog.

Our data clearly demonstrate that CSE plays an important role in lignification in poplar. The *hpCSE* lines deposited up to 25% less lignin than control plants and



**Figure 6.** Saccharification efficiency of stem biomass from wild-type (WT) and *hpCSE* plants. A, C, and E, Cellulose-to-Glc conversion during limited saccharification of stems from the wild type (solid lines) and the *hpCSE* lines (dashed lines) without (A) and with acidic (C) and alkaline (E) pretreatments. Cellulose-to-Glc conversion was calculated based on the amounts of released Glc and the quantified amount of cellulose. B, D, and F, Relative Glc release per plant after 48 h of saccharification without (B) and with acid (D) and alkaline (F) pretreatments, normalized to the values of the wild type. Error bars indicate SE. Significant differences were assessed with Student's *t* test (\*,  $0.05 > P > 0.01$ ; \*\*,  $0.01 > P > 0.001$ ; and \*\*\*,  $P < 0.001$ ;  $n = 10$ ).

the relative level of H units was increased slightly, which is in line with a function for CSE after the branch where the biosynthesis of H units diverges from that of G and S units. These lignin changes were not associated with any drastic effect on plant growth and development. This mild phenotype is likely due to residual expression of both *PtxaCSE* paralogues observed in the *hpCSE* lines. Accordingly, although the Arabidopsis loss-of-function *cse-2* mutant exhibited a substantial developmental arrest, a 36% reduction in lignin amount, and a 30-fold increase in H unit content, the knockdown *cse-1* mutant,

with a T-DNA insertion in the promoter and 6.3% residual CSE expression, did not show any developmental abnormality, with total lignin reduced by 17% and only a 4-fold increase in H unit content (Vanholme et al., 2013c). Interestingly, the loss-of-function phenotype of *M. truncatula cse* mutants is even more severe than that of Arabidopsis, with lignin levels reduced by over 80% and H unit content increased by around 50-fold (Ha et al., 2016). Given that both are knockout mutants, the contrasting phenotypes are probably due to differential catalytic properties of HCT in converting caffeoyl shikimate into

caffeoyl-CoA in these two species. Although the kinetics of this reaction are unknown for *M. truncatula* HCT, Arabidopsis HCT has been shown to catalyze this conversion in vitro, albeit with a much lower affinity for caffeoyl shikimate than the Arabidopsis CSE (Vanholme et al., 2013c). Therefore, a lower catalytic efficiency of *M. truncatula* HCT might result in a lower flux toward G and S units from the accumulated caffeoyl shikimate pool caused by CSE loss of function, resulting in a more drastic phenotype when compared with Arabidopsis. In poplar, the efficiency of recombinant HCT1 and HCT6 in converting caffeoyl shikimate into caffeoyl-CoA was shown to be very low (Wang et al., 2014), which explains why even intermediate levels of CSE down-regulation in this species still result in the accumulation of caffeoyl shikimate and a significantly lower lignin content (this work). Altogether, these data suggest that only strong reductions in CSE mRNA levels cause developmental abnormalities, whereas intermediate suppression does not necessarily affect plant yield but could still lead to alterations in lignin deposition that might be beneficial for biomass processing. These results highlight the biotechnological potential of CSE engineering for the development of bioenergy feedstock with improved extractability.

#### CSE Down-Regulation in Poplar Results in a Shift in Phenolic Metabolism

Lignin loss-of-function mutants and down-regulated plants often accumulate the substrate of the corresponding deficient enzyme along with other pathway intermediates or derivatives, which results in conspicuous shifts in their phenolic profiles (Morreel et al., 2004; Rohde et al., 2004; Coleman et al., 2008; Vanholme et al., 2012b, 2013c). Therefore, phenolic profiling is a powerful tool with which to gain insight into the potential in vivo substrate of a given enzyme as well as to evaluate how the carbon flux through the phenylpropanoid pathway is redirected in response to a perturbation (i.e. upon blocking a particular step in the pathway). The majority of the compounds with significantly reduced abundance in *hpCSE* poplars were oligolignols composed mainly of G and/or S units and their corresponding cinnamaldehydes, which is consistent with the reduced flux toward the main monolignols if CSE is a pathway enzyme utilized in poplar lignification. Interestingly, upon CSE down-regulation, the carbon flux was not massively redirected to the biosynthesis of H units, and no differences in the relative amount of H-containing oligolignols were observed in the metabolome of *hpCSE* plants: the relative amount of H units in the lignin polymer doubled but remained very low overall. This is in contrast with the results found in Arabidopsis, in which three H-containing oligolignols accumulated to significant levels in both *cse-1* and *cse-2* mutants (Vanholme et al., 2013c). This discrepancy might again be explained by the higher residual CSE expression in comparison with the corresponding Arabidopsis mutants, but other explanations

also are possible. For example, the metabolite profiling revealed the accumulation of *p*-coumaroyl shikimate and caffeoyl shikimate in the *hpCSE* lines, and both compounds have been described to inhibit poplar 4CL activities (Lin et al., 2015). Such inhibition could cause the observed accumulation of ferulic acid and derivatives of caffeic acid (e.g. caffeic acid glucosides) in the *hpCSE* lines, but it also could be the reason for a lower flux toward H units in the *hpCSE* lines. Furthermore, the differences between the redirection of the carbon flux in Arabidopsis *cse-2* and poplar *hpCSE* might be due to intrinsic differences in the phenolic metabolism or different regulatory mechanisms to compensate for the lower amount of lignin. Accordingly, it has been shown that CINNAMATE 4-HYDROXYLASE (C4H) interacts physically with C3H in poplar xylem and that this protein complex is able to catalyze the 3-hydroxylation of *p*-coumaric acid into caffeic acid, providing an alternative pathway to the production of G and S units that bypasses the shikimate shunt in poplar (Chen et al., 2011). Such a pathway also might explain the lower redirection of carbon skeletons toward H units in the *hpCSE* lines as compared with Arabidopsis.

The observed accumulation of *p*-coumaroyl shikimate in the *hpCSE* lines might logically be explained by product inhibition of C3H activity. In addition to *p*-coumaroyl shikimate, feruloyl shikimate and sinapoyl shikimate also accumulated in the *hpCSE* lines. It is unlikely that feruloyl shikimate is made via 4CL and HCT activities from its corresponding acid, because the intermediate feruloyl-CoA is expected to be used by CCR and thereby shunted into the monolignol-specific pathway. More likely, feruloyl shikimate and sinapoyl shikimate are synthesized via hydroxylation and methylation of the accumulating caffeoyl shikimate. If confirmed, these reactions add another layer to the complex grid of phenylpropanoid derivatizations; besides hydroxylation and methylation at the free-acid level, the CoA-thioester level, and the aldehyde level, these reactions now also appear to occur at the shikimate ester level. However, the relevance of this route in wild-type poplars remains to be elucidated further. Notably, the accumulation of feruloyl shikimate and sinapoyl shikimate in the *hpCSE* lines suggests that these compounds also are in vivo CSE substrates.

Eight phenylpropanoid derivatives were detected in their hexosylated forms. The accumulation of these compounds hints at a detoxification route for potentially harmful pathway intermediates (phenolic acids) that accumulate upon CSE perturbation (Vanholme et al., 2010; Sundin et al., 2014; Dima et al., 2015).

#### Total Glc Release per Plant Is Increased Remarkably upon CSE Down-Regulation in Poplar

Despite its essential role in plant growth and development, lignin is widely recognized as a major limiting factor that hinders the efficient processing of plant biomass into fermentable sugars (Zhao et al., 2012a;



Cesarino et al., 2016). The recalcitrant nature of this phenolic polymer has stimulated biotechnological approaches to improve polysaccharide extraction and processing in bioenergy crops via genetic modification of lignin biosynthesis, either by reducing lignin amounts or altering lignin composition (Chen and Dixon, 2007; Vanholme et al., 2012a, 2013b; Van Acker et al., 2014; Mottiar et al., 2016; Sibout et al., 2016). Accordingly, a 4-fold increase in saccharification yield was observed for the *Arabidopsis cse-2* loss-of-function mutant, which shows a 36% reduction in lignin content (Vanholme et al., 2013c). However, the improvement in Glc release was accompanied by a substantial biomass yield penalty, with reductions of 37% and 42% in the inflorescence height and weight at senescence, respectively. Indeed, drastic reductions in the biosynthesis of lignin are typically followed by deleterious effects on plant development, which might significantly outweigh the gains in fermentable sugar yield (Gallego-Giraldo et al., 2011; Bonawitz and Chapple, 2013; Yang et al., 2013; Van Acker et al., 2014). Although the underlying molecular mechanisms that result in developmental arrest in lignin-deficient plants are largely unknown, it is likely that they depend on the particular step of the pathway that is blocked (Bonawitz and Chapple, 2013). Recently, we found evidence that vasculature collapse underlies the yield penalty found in the *Arabidopsis cse-2* mutant, as restoring CSE expression specifically in vessels also restored vasculature morphology and final stem weight, leading to an even higher total Glc release per plant (Vargas et al., 2016). In poplar, CSE down-regulation resulted in higher amounts of Glc released upon saccharification with and without pretreatment, which is likely associated with reduced lignin amounts combined with increased relative cellulose contents. Similar results (i.e. an inverse relation between lignin and cellulose) also have been found when other biosynthetic genes, such as *C4H*, *4CL*, *C3H*, *CCR*, and *COMT* (Jouanin et al., 2000; Leplé et al., 2007; Coleman et al., 2008; Bjurhager et al., 2010; Voelker et al., 2010), were down-regulated in poplar. However, increases in cellulose content were not found in a set of *Arabidopsis* mutants defective in lignin biosynthesis, in which lower deposition of lignin was compensated for by a relative increase in hemicellulose content (Van Acker et al., 2013). Also in the case of the *Arabidopsis cse-2* mutant, the lower lignin content was associated with a decrease in cellulose content (Vanholme et al., 2013c; i.e. the opposite to what we found for transgenic poplars down-regulated in CSE). Of particular note, the improved polysaccharide extractability of *hpCSE* plants was not accompanied by severe negative effects on plant yield when grown for 6 months under greenhouse conditions. Consequently, in limited-saccharification experiments, the total Glc released per plant was increased by up to 62% when no pretreatment was applied and by up to 86% and 91% when acid and alkaline pretreatments were used, compared with the controls. Overall, the observation that CSE silencing results in significantly improved

sugar yield without drastic effects on biomass production, at least under greenhouse conditions, opens the possibility for the exploitation of this gene as a target for genetic engineering of lignin biosynthesis in bioenergy crops.

The improvements in saccharification yield observed in the *hpCSE* lines were not as remarkable as those observed for the *Arabidopsis cse-2* mutant, which were among the highest saccharification efficiencies ever reported (Van Acker et al., 2013; Vanholme et al., 2013c). However, the *cse-2* plants exhibited not only more drastic reductions in lignin levels but the lignin polymer also was massively enriched in H units, which might help explain their higher performance in saccharification assays. Because *p*-coumaryl alcohol is devoid of methoxy groups that, as in *p*-coumarate, help stabilize the free radical, it favors radical transfer reactions (Hatfield et al., 2008; Ralph, 2010). Consequently, *p*-coumaryl alcohol frequently occurs as free-phenolic end groups when incorporated into lignin (Lapierre and Monties, 1988; Russell et al., 2006; Mottiar et al., 2016). Therefore, H-rich lignins have a lower degree of polymerization (Lapierre and Rolando, 2009; Ziebell et al., 2010). Presumably, shorter lignin chains would be less cross-linked, have potentially fewer connections to the cell wall ultrastructure, and have a greater solubility (Ziebell et al., 2010). Therefore, increasing the incorporation of H units may increase lignin extractability and improve biomass digestibility. Because the relative amount of H units in the *hpCSE* lines was increased only slightly, it is unlikely that this parameter has contributed to the improved saccharification efficiency found in these trees. Nevertheless, boosting H unit incorporation in order to produce shorter lignin polymers remains a promising strategy to increase cellulose-to-Glc conversion in dedicated crops (Vanholme et al., 2012a; Bonawitz et al., 2014; Mottiar et al., 2016). In this regard, the generation of CSE knockouts in poplar via CRISPR/Cas9 would likely produce plants with lower lignin amounts and much higher incorporation of H units into lignin.

#### CSE Activity Might Not Be Essential for Lignification in All Plant Species

The recent characterization of CSE in *M. truncatula* (Ha et al., 2016) and poplar (this work) is shedding some light on the generality of CSE function in monolignol biosynthesis. To date, reverse genetics has unambiguously confirmed an important role for CSE in lignification only in plants harboring bona fide class I homologs in their genomes. How caffeate esters are channeled into the production of G and S lignin units in plants that do not possess genes encoding class I CSE enzymes, which is the case for most grasses, therefore remains an open question. One possible explanation is that, in those plants, HCT has a higher efficiency in converting caffeoyl shikimate into caffeoyl-CoA than its counterparts in CSE-harboring species. Unfortunately, limited data on kinetic parameters for HCT are

available, especially for those species shown to lack a class I CSE. Hitherto, the only study characterizing the reaction mechanism of HCT in a plant that lacks a bona fide CSE focused exclusively on the forward reaction, with no information on the reverse transesterification reaction (Walker et al., 2013). Alternatively, the phylogenetically distant class II CSE enzymes might be catalytically able to fulfill this enzymatic step. Interestingly, switchgrass seems to employ an HCT-like enzyme for the biosynthesis of chlorogenic acids (CGAs; Escamilla-Treviño et al., 2014). In plants that accumulate CGAs, such as tobacco and coffee (*Coffea arabica*), the enzyme HYDROXYCINNAMOYL-COENZYME A:QUINATE HYDROXYCINNAMOYLTRANSFERASE (HQT) is normally responsible for the formation of CGA from caffeoyl-CoA and quinic acid, but close HQT orthologs were not found in switchgrass. In contrast, an HCT-like enzyme phylogenetically distant from HCTs or HQTs was discovered from transcriptome analysis of lignifying switchgrass cell suspensions and elongating internodes (Shen et al., 2013) and shown to exhibit HQT activity, preferring quinate instead of shikimate as its acyl acceptor (Escamilla-Treviño et al., 2014). A similar scenario may occur for class II CSEs in lignin biosynthesis. In support of this hypothesis, it was shown recently that the class II CSEs from maize and sorghum (*Sorghum bicolor*) are targeted by the lignin repressors MYB31 and MYB42 in chromatin immunoprecipitation experiments (Agarwal et al., 2016), suggesting a role for these genes in lignification. Conversely, the only Arabidopsis class II CSE gene was characterized recently and shown to encode a monoacylglycerol lipase that catalyzes the last step of triacylglycerol breakdown, the hydrolysis of monoacylglycerol to fatty acid and glycerol (Kim et al., 2016). Therefore, it is likely that the lack of class I CSEs contributed to the process of subfunctionalization or neofunctionalization of class II CSEs, allowing these enzymes to be part of the lignin toolbox in certain plants only. Further genetic analysis is necessary to determine how the shikimate shunt works in plants lacking class I CSE genes.

## MATERIALS AND METHODS

### Plant Material and Generation of Transgenic *hpCSE* Lines

The first 120 bp of the *PtxaCSE2* coding sequence from *Populus tremula* × *Populus alba* (INRA 717-1B4), corresponding to *Populus trichocarpa* Potri.003G059200, was PCR amplified from cDNA obtained from *P. tremula* × *P. alba* stems using iProof High Fidelity DNA Polymerase and cloned into the pDONR221 vector using BP Clonase (Invitrogen). Sequence identity was confirmed by sequencing. The fragment was introduced subsequently via LR Clonase (Invitrogen) into the pK7GWIWG2II vector suited for cauliflower mosaic virus 35S-driven intron-spliced hairpin RNA-mediated gene silencing. Restriction analysis with *Xba*I (New England Biolabs) was performed to confirm that intron flip, which may reduce silencing efficiency, did not occur. The recombinant vector was transferred into *Agrobacterium tumefaciens* strain C58C1 PMP90 by electroporation. *A. tumefaciens*-mediated transformation of *P. tremula* × *P. alba* was performed according to Leplé et al. (1992).

In total, 137 transgenic lines and their corresponding wild type were individually transferred to soil in pots of 5.5 cm diameter, placed in a tray filled with water, and covered with a cage liner (Tecniplast APET disposable cage liner for cage body 1291H) for acclimatization. After 2 weeks, one side of the cage liner was lifted and kept accordingly for 1 d, after which the other side also was lifted. The next day, the

acclimatized plants were transferred to bigger pots (24 cm diameter) and grown for 45 d in the greenhouse. A 5-cm debarked stem fragment was collected at the base of each plant for the first screening (relative increase of caffeoylshikimate). The debarked stem fragment was scraped, and the scrapings were extracted subsequently with 1 mL of methanol at 70°C for 15 min under 1,000 rpm shaking. After centrifugation, 800  $\mu$ L of the liquid phase was dried in a SpeedVac, and the resulting pellet was resuspended in a mixture of water:cyclohexane (100  $\mu$ L:100  $\mu$ L). The tubes were vortexed and centrifuged at 20,000g for 10 min, after which a 15- $\mu$ L aliquot of the lower water phase was analyzed via UHPLC-MS. Mass spectrometric analysis and caffeoyl shikimate relative quantification were performed as described previously (Vanholme et al., 2013c). The top 20 lines with highest accumulation of caffeoyl shikimate were chosen for the second screening, which was performed with a debarked 5-cm stem fragment harvested immediately above the fragment used for the first screening. The stem fragment was dried at room temperature for 2 weeks, ground in 2-mL Eppendorf tubes using a Retsch MM300 mill (20 Hz, 4-mm bead), and subjected to sequential extraction to obtain the CWR (Van Acker et al., 2013). Lignin content was determined using the acetyl bromide method adapted to small samples (Johnson et al., 1961; Van Acker et al., 2013), whereas thioacidolysis (Robinson and Mansfield, 2009) was performed to determine the relative level of H units in each sample. The best two lines with highest accumulation of caffeoyl shikimate, and the lowest lignin content with the highest level of H units, were chosen for further multiple-level phenotyping.

Two CSE down-regulated lines (*hpCSE#1* and *hpCSE#2*) were selected and, together with the corresponding wild type, propagated simultaneously and grown randomly in the greenhouse to obtain enough biological replicates of each line. For the analysis of biomass parameters, wild-type ( $n = 25$ ), *hpCSE#1* ( $n = 23$ ), and *hpCSE#2* ( $n = 23$ ) plants were grown under a 16-h-light/8-h-dark photoperiod at  $\pm 21^\circ\text{C}$  in 10-L pots filled with a Saniflor commercial soil (Van Israel). The height of the trees was measured weekly for a period of 6 months. At the end of the growth period, the stems were harvested to determine their final fresh and dry weights.

For phenolic profiling, stems from seven replicates of each genotype grown for approximately 6 months under the same greenhouse conditions were cut 10 cm above the soil. A basal 10-cm stem fragment was debarked, frozen in liquid nitrogen, and stored at  $-70^\circ\text{C}$  until use. For cell wall analyses ( $n = 10$  per line), 2D NMR ( $n = 3$  per line), and saccharification assays ( $n = 10$  per line), the stems of another set of 6-month-old plants grown under the same conditions were cut 10 cm above the soil. The stems were debarked, air dried, ground in a ball mill, and sieved to pass a mesh of 0.5 mm, after which the CWR was prepared for each sample as described previously (Van Acker et al., 2013).

### CSE Activity Assays of Poplar Xylem

The preparation of crude protein extracts of xylem from 3.5-month-old *P. tremula* × *P. alba* wild-type plants was performed as described previously for Arabidopsis (*Arabidopsis thaliana*; Vanholme et al., 2013a). The reaction mixture contained 50 mM Tris-HCl buffer, pH 7, 1 mM DTT, 100  $\mu$ M caffeoyl shikimate, and 10  $\mu$ g of soluble xylem proteins in a total volume of 40  $\mu$ L. Samples were incubated at 25°C for 1 or 25 h, and the reaction was terminated by boiling for 5 min. Boiled xylem protein extracts were used as a negative control. Quantification of caffeic acid was performed as described previously (Vanholme et al., 2013c).

### Gene Expression Analysis via RT-qPCR

Tissue-specific expression analysis was performed in bark, mature leaves, axillary buds, and developing xylem of 6-month-old *P. tremula* × *P. alba* wild-type plants. A 10-cm stem fragment was harvested in a region 40 cm below the apex of the plants, debarked, and immediately frozen in liquid nitrogen. The bark, leaves, and axillary buds were harvested and also frozen in liquid nitrogen. All plant materials were ground in 2-mL Eppendorf tubes using a Retsch MM300 mill (20 Hz, 4-mm bead). Total RNA was extracted using the RNeasy Plant Mini Kit (Qiagen) and treated with Ambion DNA-free (Life Technologies) to remove contaminating genomic DNA. Total RNA (1  $\mu$ g) was used as a template for the synthesis of cDNA using the iScript cDNA Synthesis Kit (Bio-Rad). cDNAs from two plants were pooled to constitute a biological replicate, and a total of four biological replicates for each line were used in the analysis. The transcript levels of *PtxaCSE1* (corresponding to Potri.001G175000) and *PtxaCSE2* (corresponding to Potri.003G059200) were determined via quantitative RT-PCR as described previously (Vargas et al., 2016). Poplar 18S RIBOSOMAL RNA (AF206999), POLYUBIQUITIN (BU879229), and LEAFY/FLORICAULA (Potri.015G106900) were used as reference genes (Brunner et al., 2004; Li et al., 2009). In addition, transcript levels of *PtxaCSE1* and *PtxaCSE2*

were determined in the *hpCSE* lines in order to evaluate the level of CSE down-regulation. This analysis was performed essentially as described above, but total RNA was extracted from scraped xylem of a stem fragment harvested 60 cm below the top of 6-month-old plants. All primers used in this study are listed in Supplemental Table S4.

## Cell Wall Analysis and Saccharification Assays

Lignin content was measured via the acetyl bromide lignin method to screen the transformed lines and via the Klason method to characterize *hpCSE#1* and *hpCSE#2*. Acetyl bromide lignin quantification was performed as described by Van Acker et al. (2014). The Klason method was performed via a protocol modified from Van den Bosch et al. (2015). First, 1.5 mL of 72% H<sub>2</sub>SO<sub>4</sub> (v/v) was added to 100 mg of purified CWR in 15-mL glass vials and incubated at room temperature for 2 h under magnetic stirring. Next, the content was transferred to 100-mL bottles and diluted to 3% H<sub>2</sub>SO<sub>4</sub> (v/v) using 36 mL of water. The diluted solution was then autoclaved for 1 h (1 bar, 121°C), filtered, and washed with hot water through a preweighed filter paper (Sartorius AG) using a Büchner filter system (Merck Millipore). The filter paper containing the extracted lignin was placed in a petri dish and oven dried at 80°C overnight. After cooling for 1 h at room temperature, Klason lignin was determined gravimetrically. Cellulose was measured according to a protocol modified from Brendel et al. (2000). In 5-mL autoclavable borosilicate vials with polytetrafluoroethylene-coated screw caps (Supelco/Sigma-Aldrich), 2 mL of 80% acetic acid (v/v) and 0.2 mL of 69% nitric acid (v/v) were added to 100 mg of milled stem, and the mixture was autoclaved for 20 min (1 bar, 121°C). After cooling at room temperature, 2.5 mL of absolute ethanol was added, and the samples were filtered through a preweighed filter paper (Sartorius AG) using a Büchner filter system (Merck Millipore). The pellet was washed sequentially as follows: (1) twice with 5 mL of absolute ethanol; (2) twice with 5 mL of water; (3) once with 5 mL of absolute ethanol; and (4) once with 4 mL of acetone. The filter paper with extracted cellulose was placed in a petri dish and oven dried at 50°C overnight. After cooling for 1 h at room temperature, cellulose was determined gravimetrically. Saccharification assays were performed as described previously (Van Acker et al., 2016). The total amount of Glc release on a plant basis was calculated with the following formula: (Glc release/dry weight)\*average dry weight of the stem.

## 2D NMR

Preparation of whole-plant CW gel-state NMR samples in dimethyl sulfoxide-*d*<sub>6</sub>/pyridine-*d*<sub>5</sub> (4:1) and NMR experiments were performed as described previously (Kim et al., 2008). NMR spectra were recorded on a Bruker Biospin AVANCE 700-MHz spectrometer fitted with a cryogenically cooled 5-mm quadrupole-resonance <sup>1</sup>H/<sup>31</sup>P/<sup>13</sup>C/<sup>15</sup>N QCI gradient probe with inverse geometry (proton coils closest to the sample). Bruker's TopSpin 3.2 (Mac version) software was used to process the spectra, and volume integration of contours in HSQC plots was performed on data processed without linear prediction. The peaks characteristic of the lignin monomers (aromatic distributions) and linkages (side chain distributions) were selected and estimated after the contour-level adjustment to achieve optimal peak separation. As end groups, in particular, are substantially overestimated (Mansfield et al., 2012b), the integrals were used in relative comparisons and do not represent true absolute quantification. This relative quantification was performed for all three genotypes using three biological replicates.

## Phenolic Profiling

The debarked 10-cm basal stem fragments from 3.5-month-old *P. tremula* × *P. alba* wild-type and *hpCSE* plants were scraped, and the resulting developing xylem was ground using mortar and pestle. Soluble phenolic compounds were extracted from approximately 100 mg of scraped and ground xylem with 1 mL of methanol at 70°C for 15 min under 1,000 rpm shaking. After centrifugation at room temperature and maximum speed, 800 μL of the supernatants was dried under vacuum, and the pellet was redissolved in 200 μL of milliQ water: cyclohexane (1/1, v/v). The tubes were vortexed and centrifuged at 14,000 rpm (20,000g) for 10 min, after which a 15-μL aliquot of the aqueous phase was injected on a UHPLC system (Waters Acquity UPLC) equipped with a BEH C18 column (2.1 × 150 mm, 1.7 μm; Waters) and hyphenated to a time-of-flight mass spectrometer (Synapt Q-ToF; Waters) using gradient elution. Gradient elution information, negative mode mass spectrometry setting, chromatogram integration, and alignment via Progenesis QI software (Waters) were performed as described previously (Eloy et al., 2017). Peak abundances were normalized to the dry weight (mg) of the pellet remaining after methanol extraction and dried

in the SpeedVac. The criteria to select peak/metabolites with significantly higher or lower relative abundance in *hpCSE* lines compared with the wild type were as follows: (1) present in all samples of at least one of the three genotypes; (2) average normalized abundance higher than 500 counts in at least one of the three genotypes; (3) at least 2-fold increased or 50% decreased peak area in *hpCSE* versus the wild type; and (4) Student's *t* test *P* < 0.01. PCA was performed in R software via the *prcomp* command. Annotation of compounds matching these criteria was based on accurate *m/z*, isotope distribution, and tandem mass spectrometry (MS/MS) similarities. The MS/MS spectra were structurally elucidated based on their similarity with commercially available standards and previously identified metabolites that have already been described in the literature. For targeted analysis of the oligolignols, compounds were identified via *m/z*, retention time, and MS/MS similarity.

## Supplemental Data

The following supplemental materials are provided.

**Supplemental Figure S1.** PCA plot.

**Supplemental Figure S2.** MS/MS spectra of the characterized metabolites.

**Supplemental Table S1.** Caffeoyl shikimate abundance, thioacidolysis monomeric composition, and lignin quantification used to screen the transgenic *hpCSE* lines.

**Supplemental Table S2.** Cell wall parameters measured in both *hpCSE* lines and the wild type.

**Supplemental Table S3.** Saccharification results from limited-saccharification tests.

**Supplemental Table S4.** List of primers used for cloning and quantitative RT-PCR.

Received July 7, 2017; accepted September 3, 2017; published September 6, 2017.

## LITERATURE CITED

- Agarwal T, Grotewold E, Doseff AI, Gray J (2016) MYB31/MYB42 syntelogs exhibit divergent regulation of phenylpropanoid genes in maize, sorghum and rice. *Sci Rep* 6: 28502
- Barros J, Serk H, Granlund I, Pesquet E (2015) The cell biology of lignification in higher plants. *Ann Bot (Lond)* 115: 1053–1074
- Baucher M, Chabbert B, Pilate G, Van Doorselaere J, Tollier MT, Petit-Conil M, Cornu D, Monties B, Van Montagu M, Inzé D, et al (1996) Red xylem and higher lignin extractability by down-regulating a Cinnamyl Alcohol Dehydrogenase in poplar. *Plant Physiol* 112: 1479–1490
- Berlin A, Balakshin M, Gilkes N, Kadla J, Maximenko V, Kubo S, Sadtler J (2006) Inhibition of cellulase, xylanase and β-glucosidase activities by softwood lignin preparations. *J Biotechnol* 125: 198–209
- Bjurhager I, Olsson AM, Zhang B, Gerber L, Kumar M, Berglund LA, Burgert I, Sundberg B, Salmén L (2010) Ultrastructure and mechanical properties of *Populus* wood with reduced lignin content caused by transgenic down-regulation of cinnamate 4-hydroxylase. *Biomacromolecules* 11: 2359–2365
- Boerjan W, Ralph J, Baucher M (2003) Lignin biosynthesis. *Annu Rev Plant Biol* 54: 519–546
- Bonawitz ND, Chapple C (2010) The genetics of lignin biosynthesis: connecting genotype to phenotype. *Annu Rev Genet* 44: 337–363
- Bonawitz ND, Chapple C (2013) Can genetic engineering of lignin deposition be accomplished without an unacceptable yield penalty? *Curr Opin Biotechnol* 24: 336–343
- Bonawitz ND, Kim JL, Tobimatsu Y, Ciesielski PN, Anderson NA, Ximenes E, Maeda J, Ralph J, Donohoe BS, Ladisch M, et al (2014) Disruption of Mediator rescues the stunted growth of a lignin-deficient Arabidopsis mutant. *Nature* 509: 376–380
- Brendel O, Iannetta PPM, Stewart D (2000) A rapid and simple method to isolate pure alpha-cellulose. *Phytochem Anal* 11: 7–10
- Brunner AM, Yakovlev IA, Strauss SH (2004) Validating internal controls for quantitative plant gene expression studies. *BMC Plant Biol* 4: 14
- Carocha V, Soler M, Hefer C, Cassan-Wang H, Fevereiro P, Myburg AA, Paiva JAP, Grima-Pettenati J (2015) Genome-wide analysis of the lignin toolbox of *Eucalyptus grandis*. *New Phytol* 206: 1297–1313

- Cesarino I, Araujo P, Domingues AP Jr, Mazzafera P (2012) An overview of lignin metabolism and its effect on biomass recalcitrance. *Braz J Bot* 35: 303–311
- Cesarino I, Simões MS, Brito MdS, Fanelli A, Silva TdF, Romanel E (2016) Building the wall: recent advances in understanding lignin metabolism in grasses. *Acta Physiol Plant* 38: 269
- Chang VS, Holtzapfel MT (2000) Fundamental factors affecting biomass enzymatic reactivity. *Appl Biochem Biotechnol* 84–86: 5–37
- Chen F, Dixon RA (2007) Lignin modification improves fermentable sugar yields for biofuel production. *Nat Biotechnol* 25: 759–761
- Chen HC, Li Q, Shuford CM, Liu J, Muddiman DC, Sederoff RR, Chiang VL (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
- Chundawat SPS, Donohoe BS, Sousa LdC, Elder T, Agarwal UP, Lu F, Ralph J, Himmel ME, Balan V, Dale BE (2011) Multi-scale visualization and characterization of lignocellulosic plant cell wall deconstruction during thermochemical pretreatment. *Energy Environ Sci* 4: 973–984
- Coleman HD, Park JY, Nair R, Chapple C, Mansfield SD (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
- Cullis IF, Mansfield SD (2010) Optimized delignification of wood-derived lignocellulosics for improved enzymatic hydrolysis. *Biotechnol Bioeng* 106: 884–893
- Dima O, Morreel K, Vanholme B, Kim H, Ralph J, Boerjan W (2015) Small glycosylated lignin oligomers are stored in *Arabidopsis* leaf vacuoles. *Plant Cell* 27: 695–710
- Ding SY, Liu YS, Zeng Y, Himmel ME, Baker JO, Bayer EA (2012) How does plant cell wall nanoscale architecture correlate with enzymatic digestibility? *Science* 338: 1055–1060
- Eloy N, Voorend W, Lan W, Saleme MdLS, Cesarino I, Vanholme R, Smith RA, Goeminne G, Pallidis A, Morreel K, et al (2017) Silencing *CHALCONE SYNTHASE* impedes the incorporation of tricetin in lignin and increases lignin content. *Plant Physiol* 173: 998–1016
- Escamilla-Treviño LL, Shen H, Hernandez T, Yin Y, Xu Y, Dixon RA (2014) Early lignin pathway enzymes and routes to chlorogenic acid in switchgrass (*Panicum virgatum* L.). *Plant Mol Biol* 84: 565–576
- Eudes A, Liang Y, Mitra P, Loqué D (2014) Lignin bioengineering. *Curr Opin Biotechnol* 26: 189–198
- Freudenberg K, Neish AC (1968) Constitution and Biosynthesis of Lignin. Springer-Verlag, Berlin
- Gallego-Giraldo L, Escamilla-Treviño L, Jackson LA, Dixon RA (2011) Salicylic acid mediates the reduced growth of lignin down-regulated plants. *Proc Natl Acad Sci USA* 108: 20814–20819
- Guo F, Shi W, Sun W, Li X, Wang F, Zhao J, Qu Y (2014) Differences in the adsorption of enzymes onto lignins from diverse types of lignocellulosic biomass and the underlying mechanism. *Biotechnol Biofuels* 7: 38
- Ha CM, Escamilla-Treviño L, Yance JCS, Kim H, Ralph J, Chen F, Dixon RA (2016) An essential role of caffeoyl shikimate esterase in monolignol biosynthesis in *Medicago truncatula*. *Plant J* 86: 363–375
- Hatfield R, Ralph J, Grabber JH (2008) A potential role for sinapyl *p*-coumarate as a radical transfer mechanism in grass lignin formation. *Planta* 228: 919–928
- Hoffmann L, Maury S, Martz F, Geoffroy P, 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
- Johnson DB, Moore WE, Zank LC (1961) The spectrophotometric determination of lignin in small wood samples. *Tappi* 44: 793–798
- Jouanin L, Goujon T, de Nadaï V, Martin MT, Mila I, Vallet C, Pollet B, Yoshinaga A, Chabbert B, Petit-Conil M, et al (2000) Lignification in transgenic poplars with extremely reduced caffeic acid *O*-methyltransferase activity. *Plant Physiol* 123: 1363–1374
- Kim H, Ralph J, Akiyama T (2008) Solution-state 2D NMR of ball-milled plant cell wall gels in DMSO-*d*<sub>6</sub>. *BioEnergy Res* 1: 56–66
- Kim RJ, Kim HJ, Shim D, Suh MC (2016) Molecular and biochemical characterizations of the monoacylglycerol lipase gene family of *Arabidopsis thaliana*. *Plant J* 85: 758–771
- Lapierre C, Monties B (1988) Thioacidolyses of diazomethane-methylated pine compression wood and wheat straw in situ lignins. *Holzforschung* 42: 409–411
- Lapierre C, Rolando C (2009) Thioacidolyses of pre-methylated lignin samples from pine compression and poplar woods. *Holzforschung* 42: 1–4
- Leplé JC, Brasileiro ACM, Michel MF, Delmotte F, Jouanin L (1992) Transgenic poplars: expression of chimeric genes using four different constructs. *Plant Cell Rep* 11: 137–141
- Leplé JC, Dauwe R, Morreel K, Storme V, Lapierre C, Pollet B, Naumann A, Kang KY, Kim H, Ruel K, 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 J, Brunner AM, Meilan R, Strauss SH (2009) Stability of transgenes in trees: expression of two reporter genes in poplar over three field seasons. *Tree Physiol* 29: 299–312
- Lin CY, Wang JP, Li Q, Chen HC, Liu J, Loziuk P, Song J, Williams C, Muddiman DC, Sederoff RR, et al (2015) 4-Coumaroyl and caffeoyl shikimic acids inhibit 4-coumaric acid:coenzyme A ligases and modulate metabolic flux for 3-hydroxylation in monolignol biosynthesis of *Populus trichocarpa*. *Mol Plant* 8: 176–187
- Liu CJ (2012) Deciphering the enigma of lignification: precursor transport, oxidation, and the topochemistry of lignin assembly. *Mol Plant* 5: 304–317
- Mansfield SD, Kang KY, Chapple C (2012a) Designed for deconstruction: poplar trees altered in cell wall lignification improve the efficacy of bio-ethanol production. *New Phytol* 194: 91–101
- Mansfield SD, Kim H, Lu F, Ralph J (2012b) Whole plant cell wall characterization using solution-state 2D NMR. *Nat Protoc* 7: 1579–1589
- Marriott PE, Gómez LD, McQueen-Mason SJ (2016) Unlocking the potential of lignocellulosic biomass through plant science. *New Phytol* 209: 1366–1381
- Morreel K, Dima O, Kim H, Lu F, Nicolaes C, Vanholme R, Dauwe R, Goeminne G, Inzé D, Messens E, et al (2010) Mass spectrometry-based sequencing of lignin oligomers. *Plant Physiol* 153: 1464–1478
- Morreel K, Ralph J, Lu F, Goeminne G, Busson R, Herdewijn P, Goeman JL, Van der Eycken J, Boerjan W, Messens E (2004) Phenolic profiling of caffeic acid *O*-methyltransferase-deficient poplar reveals novel benzodioxane oligolignols. *Plant Physiol* 136: 4023–4036
- Mottiar Y, Vanholme R, Boerjan W, Ralph J, Mansfield SD (2016) Designer lignins: harnessing the plasticity of lignification. *Curr Opin Biotechnol* 37: 190–200
- Moura JCMS, Bonine CAV, de Oliveira Fernandes Viana J, Dornelas MC, Mazzafera P (2010) Abiotic and biotic stresses and changes in the lignin content and composition in plants. *J Integr Plant Biol* 52: 360–376
- Nakagame S, Chandra RP, Saddler JN (2010) The effect of isolated lignins, obtained from a range of pretreated lignocellulosic substrates, on enzymatic hydrolysis. *Biotechnol Bioeng* 105: 871–879
- Ralph J (2010) Hydroxycinnamates in lignification. *Phytochem Rev* 9: 65–83
- Ralph J, Brunow G, Harris PJ, Dixon RA, Schatz PF, Boerjan W (2008) Lignification: are lignins biosynthesized via simple combinatorial chemistry or via proteinaceous control and template replication? *In* F Daayf, V Lattanzio, eds, Recent Advances in Polyphenol Research. Wiley-Blackwell, Oxford, UK, pp 36–66
- Ralph J, Lundquist K, Brunow G, Lu F, Kim H, Schatz PF, Marita JM, Hatfield RD, Ralph SA, Christensen JH, Boerjan W (2004) Lignins: natural polymers from oxidative coupling of 4-hydroxyphenyl-propanoids. *Phytochem Rev* 3: 29–60
- Rinaldi R, Jastrzebski R, Clough MT, Ralph J, Kennema M, Bruijninx PCA, Weckhuysen BM (2016) Paving the way for lignin valorisation: recent advances in bioengineering, biorefining and catalysis. *Angew Chem Int Ed Engl* 55: 8164–8215
- Robinson AR, Mansfield SD (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
- Rohde A, Morreel K, Ralph J, Goeminne G, Hostyn V, De Rycke R, Kushnir S, Van Doorselaere J, Joseleau JP, Vuylsteke M, et al (2004) Molecular phenotyping of the *pal1* and *pal2* mutants of *Arabidopsis thaliana* reveals far-reaching consequences on phenylpropanoid, amino acid, and carbohydrate metabolism. *Plant Cell* 16: 2749–2771
- Russell WR, Burkitt MJ, Scobbie L, Chesson A (2006) EPR investigation into the effects of substrate structure on peroxidase-catalyzed phenylpropanoid oxidation. *Biomacromolecules* 7: 268–273
- Schoch G, Goepfert S, Morant M, Hehn A, Meyer D, Ullmann P, 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
- Shen H, Mazarei M, Hisano H, Escamilla-Treviño L, Fu C, Pu Y, Rudis MR, Tang Y, Xiao X, Jackson L, et al (2013) A genomics approach to deciphering lignin biosynthesis in switchgrass. *Plant Cell* 25: 4342–4361

- Sibout R, Bris PL, Légée F, Cézard L, Renault H, Lapierre C (2016) Structural redesigning Arabidopsis lignins into alkali-soluble lignins through the expression of *p*-coumaroyl-CoA:monolignol transferase (PMT). *Plant Physiol* **170**: 1358–1366
- Sundin L, Vanholme R, Geerinck J, Goeminne G, Höfer R, Kim H, Ralph J, Boerjan W (2014) Mutation of the inducible *ARABIDOPSIS THALIANA* *CYTOCHROME P450 REDUCTASE2* alters lignin composition and improves saccharification. *Plant Physiol* **166**: 1956–1971
- Tsai CJ, Popko JL, Mielke MR, Hu WJ, Podila GK, Chiang VL (1998) Suppression of *O*-methyltransferase gene by homologous sense transgene in quaking aspen causes red-brown wood phenotypes. *Plant Physiol* **117**: 101–112
- Van Acker R, Déjardin A, Desmet S, Hoengenaert L, Vanholme R, Morreel K, Laurans F, Kim H, Santoro N, Foster C, Goeminne G, Légée F, Lapierre C, Pilate G, Ralph J, Boerjan W (2017) Different metabolic routes for coniferaldehyde and sinapaldehyde with CINNAMYL ALCOHOL DEHYDROGENASE1 deficiency. *Plant Physiol* **173**: ■■■-■■■
- Van Acker R, Leplé JC, Aerts D, Storme V, Goeminne G, Ivens B, Légée F, Lapierre C, Piens K, Van Montagu MCE, 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, Piens K, Boerjan W (2016) Saccharification protocol for small-scale lignocellulosic biomass samples to test processing of cellulose into glucose. *Bio Protoc* **6**: e1701
- Van Acker R, Vanholme R, Storme V, Mortimer JC, Dupree P, Boerjan W (2013) Lignin biosynthesis perturbations affect secondary cell wall composition and saccharification yield in *Arabidopsis thaliana*. *Biotechnol Biofuels* **6**: 46
- Van den Bosch S, Schutyser W, Vanholme R, Driessen T, Koelewijn SF, Renders T, Meester BD, Huijgen WJJ, Dehaen W, Courtin CM, et al (2015) Reductive lignocellulose fractionation into soluble lignin-derived phenolic monomers and dimers and processable carbohydrate pulps. *Energy Environ Sci* **8**: 1748–1763
- Van Doorselaere J, Baucher M, Chognot E, Chabbert B, Tollier M-T, Petit-Conil M, Leplé J-C, Pilate G, Cornu D, Monties B, et al (1995) A novel lignin in poplar trees with a reduced caffeic acid/5-hydroxyferulic acid *O*-methyltransferase activity. *Plant J* **8**: 855–864
- Vanholme B, Cesarino I, Goeminne G, Kim H, Marroni F, Van Acker R, Vanholme R, Morreel K, Ivens B, Pinosio S, 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, Boerjan W (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, Cross J, Morreel K, Araujo P, et al (2013c) Caffeoyl shikimate esterase (CSE) is an enzyme in the lignin biosynthetic pathway in Arabidopsis. *Science* **341**: 1103–1106
- Vanholme R, Morreel K, Darrach C, Oyarce P, Grabber JH, Ralph J, Boerjan W (2012a) Metabolic engineering of novel lignin in biomass crops. *New Phytol* **196**: 978–1000
- Vanholme R, Morreel K, Ralph J, Boerjan W (2008) Lignin engineering. *Curr Opin Plant Biol* **11**: 278–285
- Vanholme R, Ralph J, Akiyama T, Lu F, Pazo JR, Kim H, Christensen JH, Van Reusel B, Storme V, De Rycke R, et al (2010) Engineering traditional monolignols out of lignin by concomitant up-regulation of *F5H1* and down-regulation of *COMT* in Arabidopsis. *Plant J* **64**: 885–897
- Vanholme R, Storme V, Vanholme B, Sundin L, Christensen JH, Goeminne G, Halpin C, Rohde A, Morreel K, Boerjan W (2012b) A systems biology view of responses to lignin biosynthesis perturbations in *Arabidopsis*. *Plant Cell* **24**: 3506–3529
- Vargas L, Cesarino I, Vanholme R, Voorend W, de Lyra Soriano Saleme M, Morreel K, 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 SL, Lachenbruch B, Meinzer FC, Jourdes M, Ki C, Patten AM, Davin LB, Lewis NG, Tuskan GA, Gunter L, 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
- Walker AM, Hayes RP, Youn B, Vermerris W, Sattler SE, Kang C (2013) Elucidation of the structure and reaction mechanism of sorghum hydroxycinnamoyltransferase and its structural relationship to other coenzyme A-dependent transferases and synthases. *Plant Physiol* **162**: 640–651
- Wang JP, Naik PP, Chen HC, Shi R, Lin CY, Liu J, Shuford CM, Li Q, Sun YH, Tunlaya-Anukit S, 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
- Wang Y, Chantreau M, Sibout R, Hawkins S (2013) Plant cell wall lignification and monolignol metabolism. *Front Plant Sci* **4**: 220
- Wilkerson CG, Mansfield SD, Lu F, Withers S, Park JY, Karlen SD, Gonzales-Vigil E, Padmakshan D, Unda F, Rencoret J, et al (2014) Monolignol ferulate transferase introduces chemically labile linkages into the lignin backbone. *Science* **344**: 90–93
- Yang F, Mitra P, Zhang L, Prak L, Verhertbruggen Y, Kim JS, Sun L, Zheng K, Tang K, Auer M, et al (2013) Engineering secondary cell wall deposition in plants. *Plant Biotechnol J* **11**: 325–335
- Zeng Y, Zhao S, Yang S, Ding SY (2014) Lignin plays a negative role in the biochemical process for producing lignocellulosic biofuels. *Curr Opin Biotechnol* **27**: 38–45
- Zhao X, Zhang L, Liu D (2012a) Biomass recalcitrance. Part I. The chemical compositions and physical structures affecting the enzymatic hydrolysis of lignocellulose. *Biofuels Bioprod Biorefining* **6**: 465–482
- Zhao X, Zhang L, Liu D (2012b) Biomass recalcitrance. Part II. Fundamentals of different pre-treatments to increase the enzymatic digestibility of lignocellulose. *Biofuels Bioprod Biorefining* **6**: 561–579
- Ziebell A, Gracom K, Katahira R, Chen F, Pu Y, Ragauskas A, Dixon RA, Davis M (2010) Increase in 4-coumaryl alcohol units during lignification in alfalfa (*Medicago sativa*) alters the extractability and molecular weight of lignin. *J Biol Chem* **285**: 38961–38968