



Metabolic engineering of *p*-hydroxybenzoate in poplar lignin

Yaseen Mottiar^{1,2}  , Steven D. Karlen² , Robyn E. Goacher³ , John Ralph^{2,4}  and Shawn D. Mansfield^{1,2,5,*} 

¹Department of Wood Science, University of British Columbia, Vancouver, BC, Canada

²Department of Energy Great Lakes Bioenergy Research Center, Madison, WI, USA

³Department of Biochemistry, Chemistry, and Physics, Niagara University, North Tonawanda, NY, USA

⁴Department of Biochemistry, University of Wisconsin, Madison, WI, USA

⁵Department of Botany, University of British Columbia, Vancouver, BC, Canada

Received 14 January 2022;
revised 10 August 2022;
accepted 20 September 2022.

*Correspondence (Tel +1-604-822-0196;
fax +1-604-822-9159; email
shawn.mansfield@ubc.ca)

Keywords: cell-wall-bound phenolics, designer lignins, ester-linked pendent groups, 4-hydroxybenzoic acid, lignin engineering.

Summary

Ester-linked *p*-hydroxybenzoate occurs naturally in poplar lignin as pendent groups that can be released by mild alkaline hydrolysis. These ‘clip-off’ phenolics can be separated from biomass and upgraded into diverse high-value bioproducts. We introduced a bacterial chorismate pyruvate lyase gene into transgenic poplar trees with the aim of producing more *p*-hydroxybenzoate from chorismate, itself a metabolic precursor to lignin. By driving heterologous expression specifically in the plastids of cells undergoing secondary wall formation, this strategy achieved a 50% increase in cell-wall-bound *p*-hydroxybenzoate in mature wood and nearly 10 times more in developing xylem relative to control trees. Comparable amounts also remained as soluble *p*-hydroxybenzoate-containing xylem metabolites, pointing to even greater engineering potential. Mass spectrometry imaging showed that the elevated *p*-hydroxybenzoylation was largely restricted to the cell walls of fibres. Finally, transgenic lines outperformed control trees in assays of saccharification potential. This study highlights the biotech potential of cell-wall-bound phenolate esters and demonstrates the importance of substrate supply in lignin engineering.

Introduction

Lignin is a complex phenolic biopolymer found primarily in the secondary cell walls of vascular plants. The development of lignin was a milestone in plant evolution, and the lignified plant vasculature is a defining feature of tracheophytes (Weng and Chapple, 2010). Lignin is a critical component of water conduction and plant defence systems, and it contributes greatly to the compressive strength of stems and branches (Dixon and Barros, 2019). Although the chemical complexity and recalcitrance of lignin presents a major challenge for industrial biomass processing, it also paves the way for diverse biochemicals and bioproducts to be made from lignin (Rinaldi *et al.*, 2016). To that end, there is considerable interest in developing designer biomass feedstocks with engineered lignin polymers (Mottiar *et al.*, 2016).

Lignin is assembled primarily from three monolignol precursors—*p*-coumaryl, coniferyl, and sinapyl alcohol (Freudenberg and Neish, 1968). Once incorporated into a lignin polymer, these become *p*-hydroxyphenyl (H), guaiacyl (G), and syringyl (S) lignin units, respectively. The monolignols are synthesized in the cytosol prior to their export to the cell wall where laccase and peroxidase enzymes generate monolignol radicals that undergo coupling and cross-coupling reactions (Figure 1; Ralph *et al.*, 2004). This process of dehydrogenative polymerization proceeds in a combinatorial manner such that the structure of the resulting polymer reflects the complement of monomers delivered to the site of lignification.

Consequently, a wide variety of lignin composition and structure exists in Nature. The flexibility of lignification is perhaps

best exemplified by loss/gain-of-function experiments. For example, downregulation of *p*-coumaroyl-CoA 3'-hydroxylase (C3'H) in transgenic poplar led to dramatically increased levels of H units (Coleman *et al.*, 2008), whereas overexpression of ferulate 5-hydroxylase (F5H) resulted in a polymer built predominantly with S units (Huntley *et al.*, 2003; Stewart *et al.*, 2009).

The inherent plasticity of lignification extends beyond the three primary monolignols, as demonstrated by the growing number of non-canonical monomers that have been reported (Del Río *et al.*, 2022). Among these, the γ -linked monolignol conjugates are the most prevalent and best studied. For instance, the lignin found in kenaf bast fibres is naturally acetylated due to the incorporation of pre-acetylated monolignols (Lu and Ralph, 2008; Ralph, 1996). Similarly, the lignin of commelinid monocots is decorated with *p*-coumarate groups derived from monolignol-*p*-coumarate conjugates produced via *p*-coumaroyl-CoA:monolignol transferases, members of the BAHD superfamily of acyltransferases (Hatfield *et al.*, 2008; Ralph *et al.*, 1994; Withers *et al.*, 2012).

Analogously in poplars and willows, *p*-hydroxybenzoate groups (henceforth denoted as *p*HB for the ester-linked pendent form, and *p*HBA for the free acid) comprise up to 10% of the lignin (Goacher *et al.*, 2021; Smith, 1955). These too derive from monolignol conjugates and are bound to lignin via the γ position (Lu *et al.*, 2004; Morreel *et al.*, 2004). Cell-wall-bound *p*HB has also been detected throughout the family Arecaceae (palms), in Japanese spikenard (*Aralia cordata*), in the roots of carrot (*Daucus carota*), in the stems of purple mountain saxifrage (*Saxifraga oppositifolia*), and at especially high levels in the Mediterranean

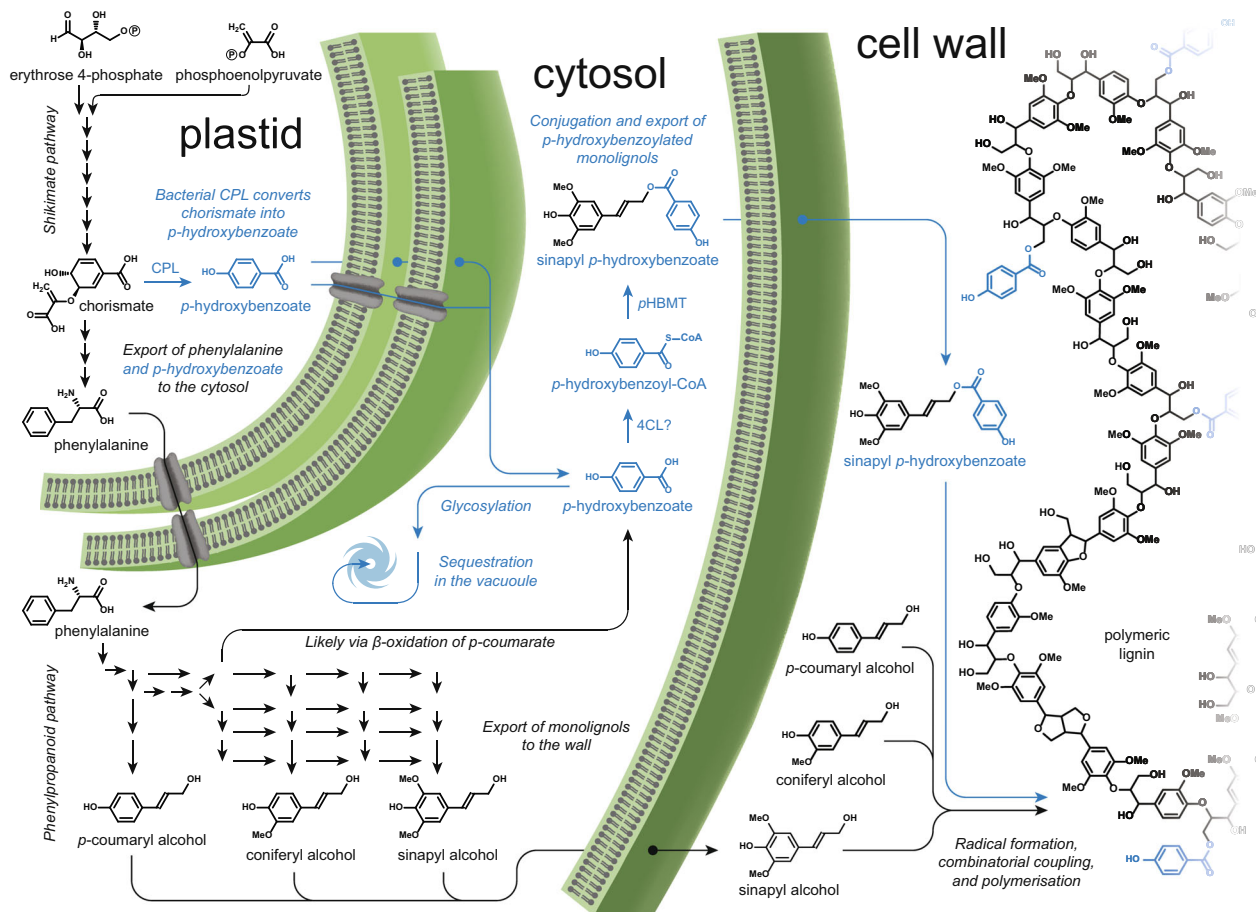


Figure 1 Metabolic map depicting the shikimate and general phenylpropanoid pathways in the plastid and cytosol, as well as the export of lignin monomers and polymerization in poplar cell walls. The novel route from chorismate catalysed by bacterial chorismate pyruvate lyase (CPL) that leads to the incorporation of more *p*-hydroxybenzoate pendent groups in lignin is shown in blue.

seagrass *Posidonia oceanica* (Faleva *et al.*, 2020; Hibino *et al.*, 1994; Parr *et al.*, 1997; Pearl *et al.*, 1959; Rencoret *et al.*, 2020).

As *p*HBA groups are ester-linked to lignin polymers, they can be released by mild alkaline hydrolysis. Once separated from the biomass, these ‘clip-off’ phenolics may be used directly as platform chemicals or upgraded into a variety of other biochemicals and bioproducts (Becker and Wittmann, 2019; Rinaldi *et al.*, 2016; Wang, Bilal, *et al.*, 2018). For example, derivatives of *p*HBA are used in cosmetics (as paraben preservatives), in polyester plastics (as liquid crystal copolymers), and as precursors to a wide range of pharmaceutical compounds (Aalto *et al.*, 1953; Jackson Jr and Kuhfuss, 1976; Manuja *et al.*, 2013). Alternatively, engineered microbes have been developed that catabolize and funnel phenolics including *p*HBA into metabolic pathways leading to high-value compounds (Beckham *et al.*, 2016; Kamimura *et al.*, 2017). However, to economize the use of clip-offs for such applications, the innate levels in feedstock plants must be substantially improved (Karlen *et al.*, 2020).

Herein, we report a novel strategy to increase the levels of *p*HBA groups by bolstering the supply of *p*HBA precursors. It has been established that *p*HBA derives from phenylpropanoid biosynthesis (Figure 1; El-Basyouni *et al.*, 1964; Zenk and Müller, 1964); however, the endogenous pathway has not been fully elucidated

and remains an inaccessible engineering target. Fortunately, an alternate route has evolved in bacteria: chorismate pyruvate lyase (CPL) cleaves chorismate, a product of the shikimate pathway, to produce *p*HBA and pyruvate (Siebert *et al.*, 1994). Previous studies on the heterologous expression of this enzyme *in planta* using different promoter sequences reported hyperaccumulation of various soluble *p*HBA-containing metabolites in leaves and plant cell cultures (Li *et al.*, 1997; Siebert *et al.*, 1996). Although this historical work demonstrated the utility of CPL in metabolic engineering, heterogeneous product streams can be a challenge in industrial processing. As poplar can deploy *p*HBA to the cell wall in a stable form where it can later be released by alkaline hydrolysis, we asked whether expression of bacterial CPL would lead to increased *p*-hydroxybenzoylation of lignin in poplar wood and thereby provide a readily accessible and homogeneous source of *p*HBA.

Results and discussion

Transgenic poplars expressing bacterial CPL

By expressing bacterial CPL in plastids, we sought to divert carbon away from chorismate and the metabolic pathway leading to lignin. In this way, the addition of more pendent groups could simultaneously result in reduced lignin content. Although recent

work has identified a cytosolic supply of chorismate (Qian *et al.*, 2019), it is generally thought that chorismate which is destined for lignin predominates in plastids. Moreover, plastidic expression of CPL is known to be more effective (Sommer and Heide, 1998; Viitanen *et al.*, 2004). As chorismate is also an essential precursor to the biosynthesis of proteins and a wide range of other metabolites, a cellulose synthase promoter, *PtCesA8pro*, was selected in order to drive expression only in cells undergoing secondary cell wall formation.

Transgenic hybrid poplar lines were generated by *Agrobacterium*-mediated transformation to evaluate the effects of heterologous expression of bacterial CPL (Figure S1, Table S1). Five lines chosen for in-depth analysis were grown in a randomized greenhouse trial alongside non-transformed wild-type (WT) control trees for 18 weeks. Other than a small penalty in growth and the occurrence of sylleptic branching, which is characteristic of fast-growing hybrid poplar and could be a response to metabolic stress (Ceulemans *et al.*, 1990), no major developmental defects were observed (Figure 2a). In comparison with control trees, transgenic lines were as much as 10% shorter and 12% thinner in stem diameter (Figure 2b). RT-qPCR was used to measure the transgene expression in developing xylem tissue and revealed that line 5 had the highest expression (Figure 2c).

Effects on lignin content, composition and structure

The xylem cell wall composition was analysed using extractive-free wood powder. The total Klason lignin content was slightly, but significantly, reduced in all lines compared with the WT control (Figure 3a). Of these, line 5 had the lowest lignin content with a 5% reduction in acid-insoluble lignin and a 15% reduction in acid-soluble lignin. Thioacidolysis was then used to assess whether there was a change in lignin composition. Analysis of liberated monomers revealed a significant albeit small decrease in S lignin units in all transgenic lines except line 1 (Figure 3a).

Two-dimensional ^1H - ^{13}C heteronuclear single-quantum coherence (HSQC) NMR spectroscopy was used to further examine changes in the composition and structure of lignin. Spectra of enzyme-lignin preparations confirmed the small shift in monomer composition (Figure 3b). Moreover, this analysis revealed an increase in cell-wall-bound *p*HB groups. Volume integrations showed a 50% increase in the amount of *p*HB in the mature xylem of line 5 compared with WT. Furthermore, the proportion of β -aryl ethers was slightly increased, primarily at the expense of resinols (Figure S2). Resinol structures arise from the dimerisation of monolignol radicals when the γ -OH traps the quinone methide intermediate (Ralph *et al.*, 2004). However, the occurrence of γ -linked *p*HB groups precludes this mechanism such that other structures prevail in highly γ -acylated lignin (Lu and Ralph, 2008).

Next, gel-permeation chromatography was used to ascertain whether increases in *p*-hydroxybenzoylation affected lignin molecular weight (Figure 3c). This analysis showed that there were no statistically significant differences on either a weight- (M_w) or number-average basis (M_n).

To corroborate the NMR results, the amount of *p*HBA released from extractive-free cell wall material was quantified following mild alkaline hydrolysis (Figure 4a). As *p*HB groups are ester linked, they can be liberated by saponification reactions. The transgenic poplars all contained significantly more *p*HB than WT control trees, with line 5 again being the most altered. In

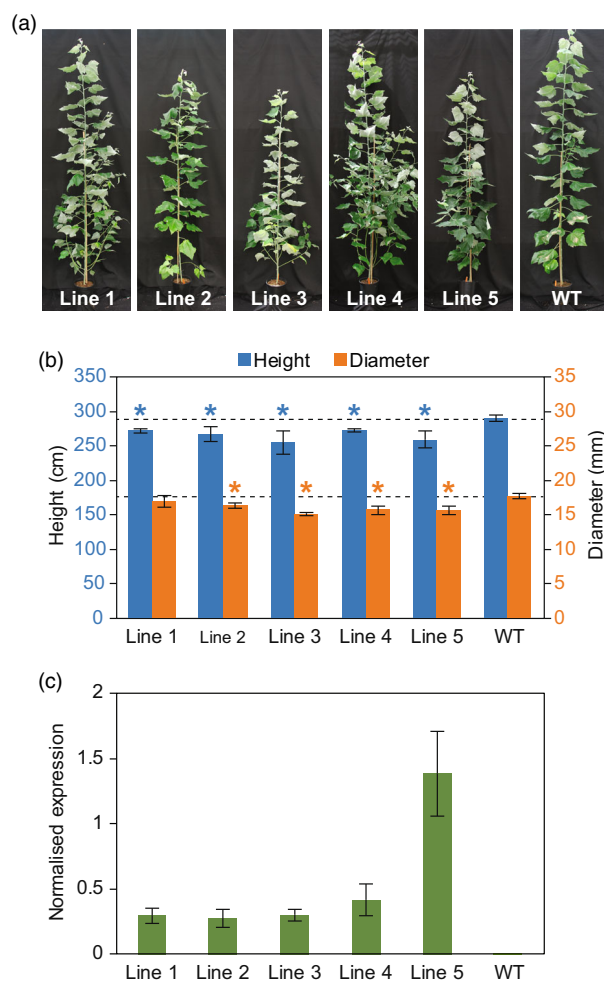


Figure 2 Expression of bacterial chorismate pyruvate lyase in transgenic poplar. (a) Representative photos of transgenic and wild-type (WT) control trees. (b) Heights and stem diameters measured at the time of harvest are shown in blue and orange on the left and right axes, respectively. Those values marked with an asterisk are significantly different from the WT control, depicted with horizontal dashed lines (one-way ANOVA with Dunnett's test, $n = 5$ for each line with technical triplicates, P -value < 0.05). (c) Relative expression of the *ubiC* gene, shown in green, was measured by RT-qPCR. No expression was detected for the WT control. Five biological replicates were analysed for each line using technical triplicates.

developing xylem, line 5 had at least nine times greater levels of *p*HB groups than the control trees. However, the differences were less pronounced in mature xylem which is comprised mostly of cells that have undergone programmed cell death. In mature xylem, line 5 had 50% more *p*HB than WT, a value that conforms with the NMR spectra of the same samples.

Next, the derivatization followed by reductive cleavage (DFRC) method was used to validate the findings from alkaline hydrolysis and to confirm that *p*HB groups originated from monolignol-*p*HB conjugates. As reductive cleavage releases monomers bound via β -aryl ether linkages but does not cleave ester bonds, DFRC yields a diagnostic product of *S-p*HB conjugates, namely 4-acetoxysinapyl *p*-acetoxybenzoate. Analysis of mature xylem showed significantly higher levels of *S-p*HB conjugates in all transgenic lines except line 1, with line 5 having 70% more than the WT control (Figure 4b).

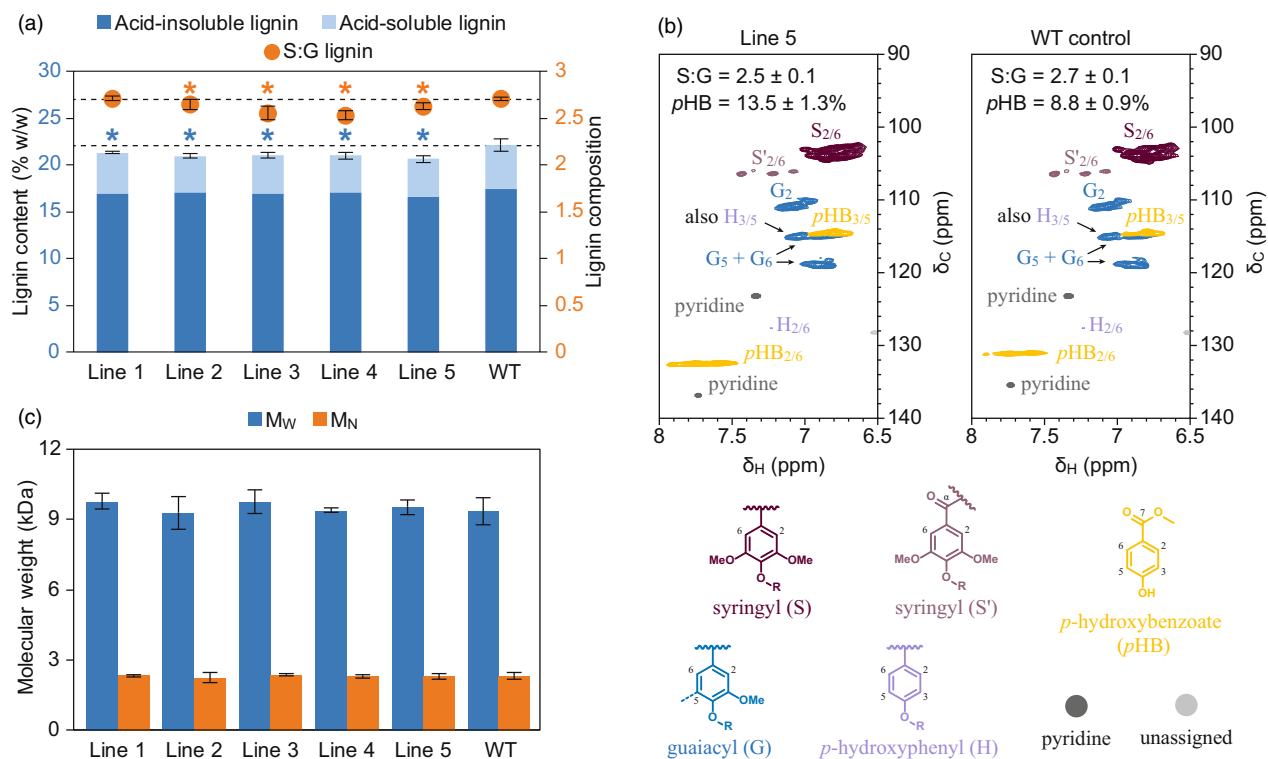


Figure 3 Lignin content, composition, and structure were altered in transgenic poplars due to the incorporation of *p*-hydroxybenzoate. (a) The Klason lignin content is plotted on the left axis with the acid-insoluble fraction shown in dark blue and the acid-soluble fraction in light blue, and the ratio of syringyl to guaiacyl lignin units (S:G) is plotted on the right axis in orange. Those values marked with an asterisk are significantly different from the wild-type (WT) control, depicted with horizontal dashed lines (one-way ANOVA with Dunnett's test, $n = 5$ for each line with technical triplicates, P -value < 0.05). (b) Two-dimensional ^1H - ^{13}C HSQC NMR spectra for enzyme-lignin samples of line 5 and the WT control, as labelled, showing the aromatics region. The colour-coding and peak annotations for *p*-hydroxyphenyl (H, light purple), guaiacyl (G, blue), and syringyl units (S and S', dark and light purple), as well as *p*-hydroxybenzoate pendent groups (yellow, *p*HB) are elaborated with the structures shown below. Proportions based on integrated peak volumes are provided as the mean \pm standard error for three biological replicates. (c) The average molecular weight is plotted on a weight-average basis (M_w) in blue and on a number-average basis (M_n) in orange.

Accumulation of *p*-hydroxybenzoate-containing metabolites

As CPL produces *p*HBA, the key precursor to *p*HB groups, it was reasonable to anticipate elevated levels in transgenic trees. However, to be deployed to the cell wall, *p*HBA must be first converted into *p*-hydroxybenzoyl-CoA, then conjugated with a monolignol, and finally exported to the site of lignification in the apoplast (see Figure 1). Accordingly, the levels of free *p*HBA were measured next. The developing xylem of all transgenic lines contained significantly more than control trees (Figure 5a). At the same time, far more *p*HBA occurred as conjugates—in both alkali-labile and acid-labile forms. Compared with control trees, line 5 had 20 times more total *p*HBA as metabolites in developing xylem.

UHPLC coupled with high-resolution mass spectrometry (HRMS) was then applied to methanolic extracts of line 5 and WT control samples to separate and identify the various *p*HBA-containing metabolites. This analysis revealed numerous compounds that were over-represented in line 5 compared with control trees (Figure 5b, Table S2). Many of these metabolites appeared to contain *p*-hydroxybenzoate moieties, as shown by the presence of a fragment ion at m/z 137.0246. Although a m/z 137.0246 fragment can also be derived from salicylate (*ortho*-hydroxybenzoate, *o*HBA), the greater abundance of glycosylated

*p*HB (peaks 1 and 4) compared with glycosylated salicylic acid (peaks 2 and 5), combined with the presence of peaks generating the m/z 137.0246 fragment in CPL poplar that are greatly reduced or entirely absent in the WT supports their designation as *p*HB-containing metabolites. The most abundant peak in line 5 samples (peak 1) was identified as the phenolic glucoside of *p*HBA, whereas the second-most abundant peak (peak 4) was determined to be the *p*HB acid glucoside (Figure 5c,d). Much weaker signals were identified by MS as glucosides of salicylic acid (m/z 299.0787, peak 2 and 5) and vanillic acid (m/z 329.0895, peak 3). These assignments were validated by matching retention times and MS fragmentation patterns with authentic standards. Furthermore, these glucosides of benzoic acid derivatives have previously been detected in the extracts of a variety of plant taxa (Herrmann, 1978; Klick and Herrmann, 1988).

For comparative purposes, methanolic extracts of line 5 and WT leaf tissues were also analysed (Figure S3, Table S3). The two glucosides of *p*HBA and various other peaks that generate m/z 137.0246 fragmentation ions were important features in the chromatograms, indicating that the effects of CPL expression were not merely restricted to xylem tissue.

Diverting carbon flux with CPL

This study is the first to exploit the inherent plasticity of lignification in order to deliberately engineer *p*HB levels in poplar.

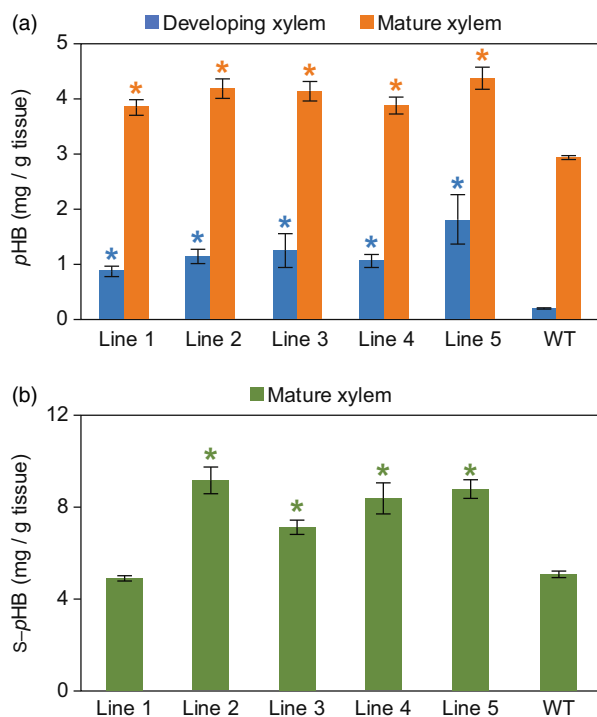


Figure 4 Transgenic poplars hyperaccumulate ester-linked monolignol conjugates of *p*-hydroxybenzoate. (a) The amount of *p*-hydroxybenzoate released by mild alkaline hydrolysis is shown for developing xylem (blue) and mature xylem (orange). (b) The amount of sinapyl *p*-hydroxybenzoate conjugates released by DFRC (sum of *cis* and *trans* isomers of 4-acetoxy sinapyl *p*-acetoxybenzoate, S-pHB) is shown in green for mature xylem. No conjugates of coniferyl *p*-hydroxybenzoate were detected. Those values marked with an asterisk are significantly different from the wild-type (WT) control (one-way ANOVA with Dunnett's test, $n = 5$ for each line with technical triplicates, P -value < 0.05).

Previously, it has been shown that *p*HBA is synthesized via *p*-coumarate and a metabolic branch of the phenylpropanoid pathway, and recent evidence favours a β -oxidation route (Terashima et al., 1975; Widhalm and Dudareva, 2015; Yazaki et al., 1991). In this context, it is not surprising that metabolic flux through the lignin pathway can affect *p*HBA content. For example, elevated *p*HBA levels were observed in transgenic poplar deficient in the lignin biosynthetic enzymes caffeoyl-CoA *O*-methyltransferase, cinnamoyl-CoA reductase, C3'H, *p*-hydroxycinnamoyl-CoA:shikimate *p*-hydroxycinnamoyl transferase, cinnamate 4-hydroxylase (C4H), and 4-coumarate:CoA ligase (Coleman et al., 2008; Leplé et al., 2007; Meyermans et al., 2000; Peng et al., 2014; Tsai et al., 2020; Wang, Matthews, et al., 2018; Zhong et al., 2000). Conversely, *p*HBA was reduced in F5H-overexpressing poplar (Stewart et al., 2009).

Despite the increases in *p*-hydroxybenzoylation observed in CPL poplar, a large pool of *p*HBA remained unbound to the cell wall and accumulated as soluble xylem metabolites. Plants possess endogenous glycosyltransferase enzymes that can act on free *p*HBA to produce both the acid glucoside and the phenolic glucoside (Katsumata et al., 1989; Lim et al., 2002; Nishizaki et al., 2014). These same compounds accumulated when CPL was expressed in tobacco, potato and sugarcane (Köhle et al., 2003; McQualter et al., 2005; Siebert et al., 1996; Viitanen et al., 2004). Furthermore, it has been reported that

glucosylation of *p*HBA occurs cytosolically in *Lithospermum* and that these glucosides are sequestered in the vacuole (Yazaki et al., 1995).

As conjugation of *p*HBA with monolignols is also believed to occur in the cytosol, the implication is that glucosylation potentially competes with monolignol acylation. This could be problematic because metabolites that have been transported into the vacuole are presumably inaccessible to β -glucosidase enzymes which could release the phenolic aglycones. However, when radiolabelled *p*HBA was provided to tobacco cell cultures, both the phenolic glucoside and the acid glucoside initially accumulated before a portion of the latter was subsequently incorporated into the cell wall (Li et al., 1997). Although a fulsome cell wall analysis was not provided in that study nor in any of the earlier work with heterologous CPL expression in plants, this observation suggests that the acid glucoside of *p*HBA could potentially be cleaved and deployed to the apoplast, perhaps as cells mature, undergo cell death, and release their vacuolar contents.

Glucosylation is not the only fate for surplus *p*HBA in the developing xylem of poplar. Various derivatives of the glucoside salicin occur naturally in poplar and willow (Boeckler et al., 2011). These metabolites contain phenolics, typically benzoate or cinnamate, that acylate either the sugar, as in populin and tremuloidin, or the salicyl alcohol group, as in populoside (Braconnot, 1830; Erickson et al., 1970; Pearl and Darling, 1959). Although *p*HBA-containing versions have been synthesized (Stepanova et al., 2014), these apparently have not been reported as naturally occurring. By examining the mass fragmentation patterns of metabolites that were differentially abundant in CPL poplar xylem, we have detected peaks that may correspond to *p*HBA-containing derivatives of salicin in the developing xylem of CPL poplar (Figure 5b,d; parent ions producing m/z 137.0246 and m/z 123.0445 product ions are coloured in magenta). Six of the metabolites with an apparent parent ion of m/z 405.1191 elute throughout the chromatogram (at 10.1, 24.1, 24.5, 25.0, 25.1, and 26.4 min), and these could be *p*HBA-containing homologues of populin, tremuloidin, populoside, salicyloylsalicin, deltoidin, chaenomeloidin, trichocarposide or a fragment of larger metabolites such as *p*HBA-containing homologues vigaureoside A, and symconoside A.

Glucosides of gentisyl alcohol also occur in poplar, and these too can be acylated with phenolic acids as in salireposide and nigracin (Thieme and Benecke, 1967; Wattiez, 1931). One compound in the pool of line 5 xylem metabolites (m/z 421.1091 at 21.8 min) might be a *p*-hydroxybenzoylated homologue of these. Another one of the metabolites (m/z 451.1246 at 26.5 min) could be vanilloylcalleryanin. Finally, two of the *p*HBA-containing metabolites (m/z 397.1503 at 8.4 and 23.9 min) could be homologues of grandidentatin or purpurein, a glucoside of 1,2-cyclohexanediol acylated with *p*-coumarate which has been identified in the bark of *Populus grandidentata* and in the xylem of RNAi-suppressed C3'H transgenic poplar (Coleman et al., 2008; Pearl and Darling, 1962). The occurrence of *p*HBA groups in these metabolites in place of other phenolics underscores the effectiveness of bacterial CPL *in planta*, and reveals substantial plasticity in the biosynthesis of these compounds in poplar xylem. It should be emphasized that, other than the glucosides of *p*HBA, salicylic acid, and vanillic acid presented in Figure 5c that were validated using reference standards, the remaining chemical designations are merely putative and provided here only for reference. Definitive identification of these

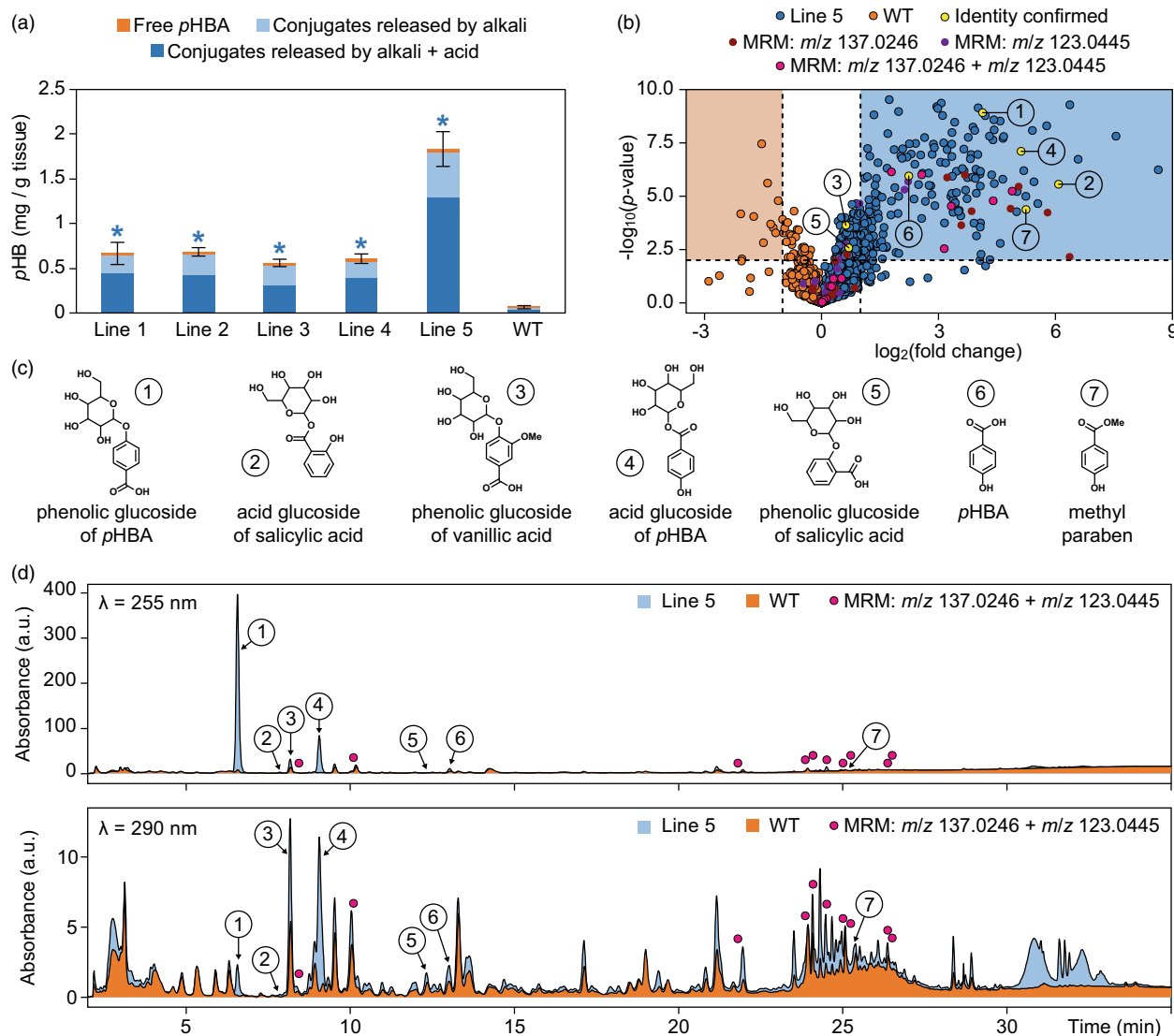


Figure 5 Metabolites of *p*-hydroxybenzoate hyperaccumulate in the developing xylem of transgenic poplars. (a) The amounts of free *p*-hydroxybenzoate (orange) as well as *p*-hydroxybenzoate released by alkali treatment (light blue) and *p*-hydroxybenzoate released by alkali and acid treatments (dark blue) are shown for methanolic extracts of developing xylem. Those values marked with an asterisk are significantly different from the wild-type (WT) control (one-way ANOVA with Dunnett's test, $n = 5$ for each line, P -value < 0.05). (b) Volcano plot showing xylem metabolites in methanolic extracts with an abundance significantly different in line 5 compared to the WT control. The metabolites in the shaded areas have a two-fold change in signal intensity with a P -value < 0.01 (Student's t -test, $n = 12$ for each line). Parent molecular ions that fragment to m/z 137.0246 (e.g. *p*HBA or salicylate) are coloured red, those that fragment to m/z 123.0445 (e.g. salicyl alcohol) are coloured purple, and those that fragment to both m/z 137.0246 and m/z 123.0445 (possibly *p*HBA-containing homologues of salicinoids) are coloured magenta. Key metabolites described in the text are indicated by compound identification numbers (1–7). (c) Chemical structures of the phenolic and acid glucosides of *p*HBA (1, 4) and salicylic acid (5, 2), the phenolic glucoside of vanillic acid (3), *p*HBA (6), and methyl paraben (7), which may be an artefact arising from esterification of *p*HBA with methanol during the extraction. (d) UHPLC-PDA traces for methanolic extracts of line 5 and the WT control at 255 and 290 nm with the peaks corresponding to confirmed compounds marked with compound identification numbers (1–7) and peaks with parent ions that fragment to both m/z 137.0246 and m/z 123.0445 and which may be *p*HBA-containing homologues of salicinoids marked with magenta circles.

metabolites would require isolation of each compound followed by structural determination by NMR and ideally comparison with chemical standards.

Distribution of cell-wall-bound *p*-hydroxybenzoate

Microscopic analysis of stem cross-sections with phloroglucinol-HCl and toluidine blue showed no obvious differences in xylem

cell morphology nor lignin distribution (Figure S4). Time-of-flight secondary-ion mass spectrometry (ToF-SIMS) was then used to evaluate the distribution of *p*HB groups in xylem cross-sections.

The peak at m/z 121.03, which has previously been attributed to cell-wall-bound *p*HB (Goacher *et al.*, 2021), was used for mapping purposes and to prepare false-colour overlay

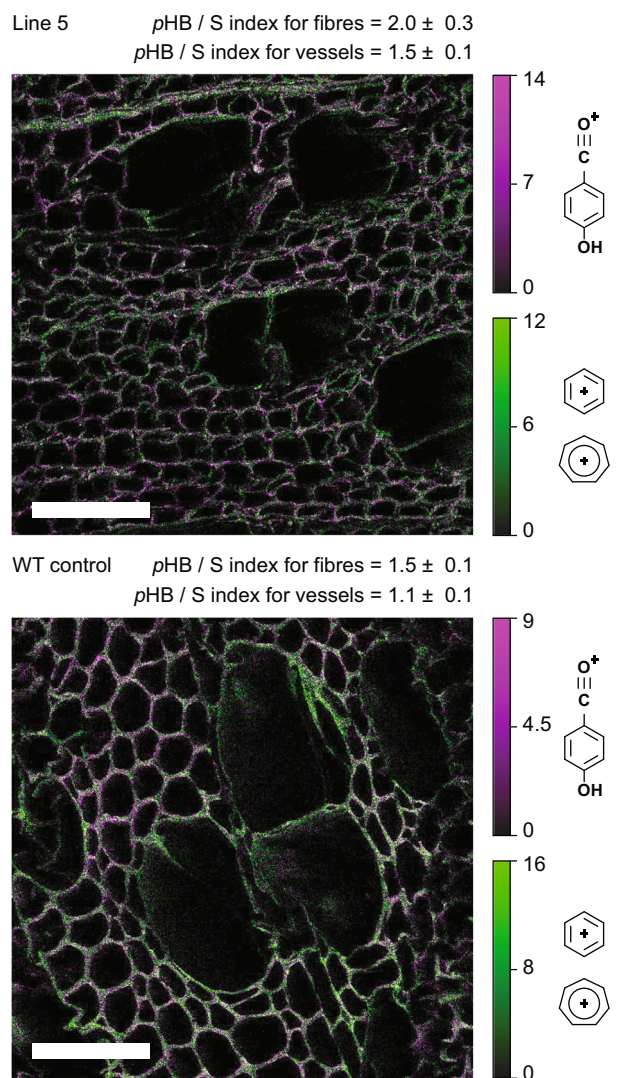


Figure 6 Distribution of cell-wall-bound *p*-hydroxybenzoate groups examined by positive-ion ToF-SIMS imaging. False-colour ion images with *p*-hydroxybenzoate groups shown in magenta (*m/z* 121.03) and lignin shown in green (sum of *m/z* 77.04 and *m/z* 91.05) for extractive-free transversal xylem cross-sections of line 5 and the wild-type (WT) control, as labelled. The gradient scales on the right show ion intensities, and the white scale bars represent 50 μm . The *p*HB/S index was calculated as a ratio of ion counts for *p*-hydroxybenzoate at *m/z* 121.03 divided by the sum of S-lignin at *m/z* 167.07 and *m/z* 181.05, and the values shown are the mean \pm standard error ($n = 5$ for line 5, $n = 3$ for WT).

images (Figure 6). This analysis revealed that *p*HB groups were predominantly limited to the cell walls of fibres in the xylem of both line 5 and WT. An index calculated for *p*HB ion counts normalized to those of S-lignin structures confirmed that the fibres of line 5 contained significantly more *p*HB groups compared with the WT control. When calculated for vessel cell walls, the index was again higher for line 5, but the values were lower due to the weak signal intensity of *p*HB in vessels.

We have previously shown that *p*HB groups occur predominantly in the cell walls of xylem fibres in poplar wood (Goacher et al., 2021). Although there was a small but proportionate

increase in cell-wall-bound *p*HB groups in the vessels of CPL poplar, the levels remained much lower than in fibres. The selection of the cellulose synthase promoter likely ensured an ample supply of *p*HBA for all xylem cell types, but monolignol-*p*HB conjugates were apparently still not efficiently produced in vessels. This could be the result of limited expression of *p*-hydroxybenzoyl-CoA:monolignol transferases, or perhaps related to the poor availability of sinapyl alcohol in vessels as these acyltransferases are known to prefer sinapyl alcohol as an acyl acceptor (de Vries et al., 2022; Zhao et al., 2021a). In order to achieve *p*-hydroxybenzoylation in the G-rich vessel cell walls of poplar, future engineering work may need to contemplate sinapyl alcohol availability or use acyltransferases that accept coniferyl alcohol as a substrate.

Strategies to maximize *p*-hydroxybenzoylation of poplar lignin

Realizing even greater *p*HB levels may ultimately require a multipronged approach. Recently, the *p*-hydroxybenzoyl-CoA:monolignol transferase responsible for *p*-hydroxybenzoylation of poplar lignin was identified (de Vries et al., 2022; Zhao et al., 2021a). A priority for future work will be to overexpress this gene in CPL poplar, but additional strategies should also be contemplated since monolignol acyltransferases alone may not achieve the desired titres of *p*HBA. For example, heterologous expression of *p*-coumaroyl-CoA:monolignol transferases in poplar and *Arabidopsis* achieved at most 0.35 and 1.3% w/w cell-wall-bound *p*-coumarate, respectively (Sibout et al., 2016; Smith et al., 2015).

One of the most impactful interventions would be blocking the formation of *p*HBA-containing metabolites. UDP-glycosyltransferases play a central role in detoxifying and modulating the accumulation of phenolics (Le Roy et al., 2016). Unfortunately, the specific enzymes responsible for glucosylation of *p*HBA in poplar are currently unknown. Although the biosynthesis of salicinoids also remains largely obscured (Babst et al., 2010), recent work has uncovered some of the steps involved in the formation of phenolic glycosides in poplar so these may soon be accessible engineering targets (Fellenberg et al., 2020).

Phenolic metabolites in the xylem of poplar could also represent an untapped fount of valuable compounds (Devappa et al., 2015). Many salicinoid metabolites are active as defence compounds in poplar (Boeckler et al., 2011), and some of these may also have pharmaceutical value. Furthermore, as many of the *p*HBA-containing compounds in the xylem of CPL poplar were ester-linked conjugates, treatment with alkali could be used to liberate the *p*HBA aglycones. In many ways, the accumulation of readily extractable *p*HBA as xylem metabolites in CPL poplar could be highly advantageous.

As chorismate is an important precursor to a wide range of biosynthetic pathways in plants, the introduction of CPL could conceivably have pleiotropic effects. Chorismate lies upstream of protein biosynthesis, via the aromatic amino acid and shikimate pathways, and upstream of several plant hormones and signalling molecules such as indoles and salicylic acid (Wang et al., 2022). In this study, the selection of a secondary cell wall-specific promoter should have mitigated the impacts on the metabolome. However, to preclude any potential adverse effects, it could be worth trialling other promoter sequences, including those specific to lignin biosynthesis and those restricted to specific tissues and developmental stages.

Impacts of ester-linked monolignol conjugates

Although the biological role of ester-linked pendent groups remains unclear, it has been suggested that one result may be increased polymerization rates for S-rich lignin (Takahama *et al.*, 1996). However, *p*HB groups are evidently negatively correlated with S units in poplar (Mottiar and Mansfield, 2022). There is also some evidence that *p*HB groups may promote plant defence as carrot cells respond to a fungal elicitor by incorporating cell-wall-bound *p*HB (Schnitzler and Seitz, 1989). Although plant growth was slightly impaired in CPL poplar, no irregular xylem phenotypes were observed and there was no indication that changes in lignin structure had other discernible effects on development. A recent study showed that gravitropism is altered in transgenic poplars with reduced *p*HB levels (Zhao *et al.*, 2021b). Consequently, it is possible that the minor effects on growth observed in CPL poplar could perhaps be directly related to increased *p*-hydroxybenzoylation.

Aside from any potential biological implications, elevated *p*HB levels led to improved biomass digestibility in assays of saccharification potential following various pretreatments. Transgenic poplar yielded greater amounts of glucose and xylose after 72 h of enzymatic hydrolysis following either acidic or alkaline pretreatments (Figure 7), commensurate with the observed increases in *p*-hydroxybenzoylation. For both glucose and xylose release, alkaline pretreatments performed the best. Compared to the WT control, line 5 released 10% more glucose and 50% more xylose following alkaline pretreatment. As the total amount of cell wall polysaccharides was actually slightly lower in line 5 compared with the control samples (Table S4), this increase in sugar release represents a *bone fide* improvement in saccharification. Previous lignin engineering efforts have achieved improved saccharification by significantly altering the lignin content, composition, and/or molecular weight (Mottiar *et al.*, 2016). In one of the few reports on the effects of pendent groups, overexpression of a *p*-coumaroyl-CoA:monolignol transferase in *Arabidopsis* led to novel *p*-coumarate groups and significantly improved saccharification, although this was also accompanied by reductions in lignin content (Sibout *et al.*, 2016).

By contrast, CPL-expressing poplar had only modest reductions in lignin (<5%), a small increase in β -aryl ether units, and no change in average molecular weight. These observations suggest

that the gains realized in saccharification could be specifically related to elevated *p*HB levels. It may be that these pendent groups alter the interactions between cell wall components or otherwise improve substrate access during biomass pretreatment and/or enzymatic hydrolysis. It has also been suggested that released *p*HBA acts as a blocking agent to prevent lignin condensation reactions and effectively enhance depolymerisation rates (Chua and Wayman, 1979).

Beyond improvements in biomass processing, cell-wall-bound *p*HB groups are themselves potentially valuable as phenolic compounds. However, the industrial use of such clip-offs will only be economical if the recovery processes are cost effective (Karlen *et al.*, 2020). In this light, CPL poplar trees with even higher titres of *p*HB could be an ideal feedstock candidate for biorefineries as alkaline hydrolysis would provide a pure stream of phenolics without the need for further separations unlike, for example, bioenergy grasses which would release heterogenous mixtures of *p*-coumarate and ferulate.

In general terms, the lignin engineering strategy promoted herein could be applied to the manipulation of other clip-offs as well. For example, increased *p*-coumaroylation and feruloylation could be facilitated by enriching the pools of those respective phenolic acyl groups. As support for this approach, the levels of cell-wall-bound benzoate, which was only recently found to be compatible with lignification in poplar, were elevated in plants deficient in C4H and C3'H (Kim *et al.*, 2020).

In summary, expression of bacterial CPL in transgenic poplar led to increased *p*-hydroxybenzoylation of lignin and enhanced saccharification potential of the wood, but no change in the distribution of *p*HB groups in the xylem. This work demonstrates the importance of substrate supply in engineering cell-wall-bound phenolics. However, abundant *p*HBA-containing metabolites in these transgenic poplars also represent untapped potential. Accordingly, *p*HB groups warrant further research consideration in the context of lignin valorisation as these ester-linked pendent groups can be clipped-off for use in the production of a wide variety of biochemicals, biomaterials, and bioproducts.

Experimental procedures

Detailed methodology is provided in the [Supporting Information](#) for cloning of the bacterial *ubiC* gene encoding CPL,

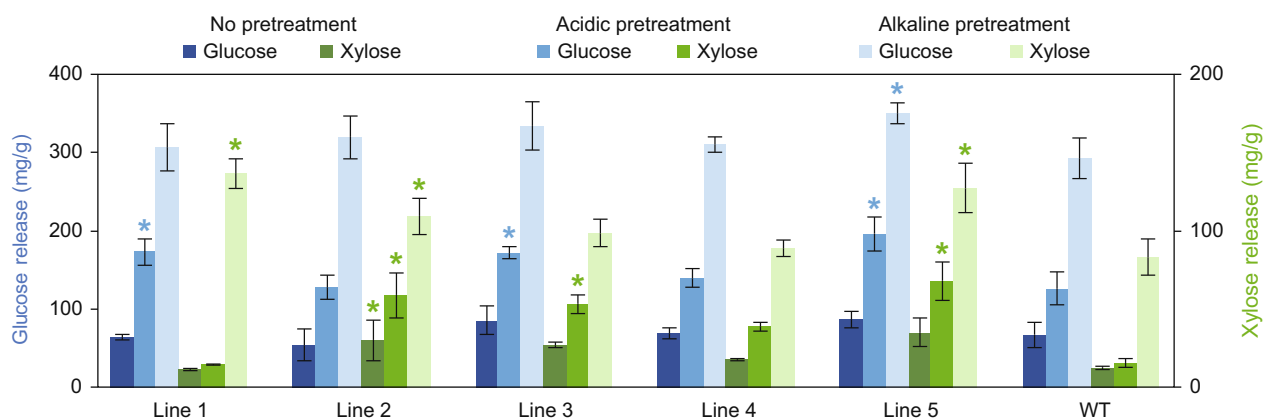


Figure 7 Saccharification potential of mature wood was improved in transgenic lines. The release of glucose (shades of blue) and xylose (shades of green) is shown following various pretreatments (alkaline, acidic, and no pretreatment, as labelled) and 72 h of enzymatic hydrolysis. Those values marked with an asterisk are significantly different from the wild-type (WT) control (one-way ANOVA with Dunnett's test, $n = 3$ for each line, P -value < 0.05).

establishment of transgenic hybrid poplar trees, measurement of transgene expression, xylem histology, ToF-SIMS imaging, gel-permeation chromatography, quantification of structural polysaccharides, and evaluation of saccharification potential.

Lignin content and composition

The Klason lignin method was used to evaluate lignin content as the sum of acid-insoluble and acid-soluble fractions. Wood powder, which had been processed in a Wiley mill to pass a 40-mesh sieve, was exhaustively Soxhlet-extracted for 24 h using hot acetone. The extractive-free wood powder was dried overnight at 50 °C and subsamples of 200 mg were accurately weighed into 125-mL glass serum bottles. Following the addition of 3 mL of cold 72% sulphuric acid, each sample was stirred with a glass rod every 10 min for 2 h prior to the addition of 112 mL of water. Once capped and sealed, the bottles were autoclaved for 1 h at 121 °C and 15 psi. Preweighed and oven-dried medium-coarseness sintered glass filtering crucibles were used to collect the remaining residues which were then rinsed with 150 mL of water and dried overnight at 105 °C before the acid-insoluble lignin was determined gravimetrically. The acid-soluble lignin component was evaluated by UV spectrophotometry of acid-hydrolysis filtrates using the Beer–Lambert law with an extinction coefficient of 110 L g⁻¹ cm⁻¹ at 205 nm.

The lignin composition was evaluated by quantifying lignin monomers released with thioacidolysis (Robinson and Mansfield, 2009). Briefly, 10-mg samples of extractive-free wood powder was subjected to thioacidolysis using 1 mL of freshly distilled dioxane containing 2.5% boron trifluoride etherate and 10% ethanethiol for 4 h at 100 °C with periodic mixing. The reactions were then incubated at -20 °C for 5 min and quenched with 0.3 mL of 0.4 M sodium bicarbonate. As an internal standard, 0.2 mL of 5 mg mL⁻¹ tetracosane was added. The reaction products were collected following phase separation using 1 mL dichloromethane and 2 mL deionized water. The organic phase was then dried by passing through a loosely packed column of anhydrous sodium sulphate, and the solvent was evaporated using a vacuum centrifuge. Samples were resuspended in 0.7 mL dichloromethane and derivatised by adding 20 µL pyridine and 100 µL *N,O*-bis(trimethylsilyl) acetamide to a 20-µL volume of sample. Derivatised lignin monomers were then separated using a Trace 1310 GC apparatus (Thermo Scientific, Waltham, MA, USA) equipped with a TraceGOLD 5MS column (30 m × 0.32 mm, 0.25 µm thickness) and a flame-ionization detector. Helium was used as the carrier gas with a constant-velocity flow rate of 1 mL min⁻¹, the injection port was maintained at 250 °C, and the column oven temperature was controlled as follows: hold at 130 °C for 3 min, then ramp at 6.5 °C min⁻¹ to 230 °C and hold for 6 min, and then finally ramp at 2.5 °C min⁻¹ to 250 °C and hold for 10 min. Peaks corresponding to the main diastereoisomers of trithioethylated G and S-lignin monomers were identified with reference to synthetic standards and were integrated using the Chromeleon 7 software package (Thermo Scientific). Finally, peak areas were corrected with previously determined response factors and then used to calculate the monomer ratio.

NMR analysis

Two-dimensional ¹H-¹³C HSQC NMR spectra were collected using enzyme-lignin samples prepared by enzymatic digestion of ball-milled wood tissue (Kim and Ralph, 2010). Wood powder which had been processed in a Wiley mill to pass a 40-mesh sieve

was subjected to sequential rounds of solvent extraction in an ultrasonic water bath as follows: 3 × 20 min with water, 3 × 20 min with 80% ethanol, 3 × 20 min with acetone. After drying in an oven at 50 °C, extractive-free wood powder was processed in a Fritsch Pulverisette 7 planetary ball mill using 20-mL agate jars containing 600 mg of wood powder and ten 10-mm agate balls at 600 rpm with a standardized procedure consisting of 30 cycles of 10 min grinding and 5 min rest. After ball-milling, the wood powder was quantitatively transferred into 50-mL Falcon tubes using 45 mL of 50 mM sodium acetate buffer (pH 5.0) and treated with 30 mg of Cellulysin cellulase (from *Trichoderma viride*; Calbiochem, MilliporeSigma, Burlington, MA, USA). After 3 days of incubation at 37 °C on a rotary shaker, the powder was rinsed twice with acetate buffer and fresh enzyme was added. After another 3 days of incubation, the powder was rinsed twice with distilled water and then freeze-dried. Samples were processed randomly, and WT controls were treated in the same manner as transgenic wood samples.

Approximately 10 mg of enzyme-lignin powder was dissolved in 0.5 mL of 4:1 dimethyl sulphoxide-d₆:pyridine-d₅ in a 5-mm outer diameter NMR tube. Spectra were collected using a standard adiabatic-pulse programme (hsqcetgpsisp2.2) on a Biospin AVANCE 700 MHz spectrometer (Bruker Corp., Billerica, MA, USA) equipped with a 5-mm QCI ¹H/³¹P/¹³C/¹⁵N cryoprobe with the proton coils closest to the sample. The dimethyl sulphoxide peak at δ_H 2.5 ppm and δ_C 39.5 ppm was used for calibration, and volume integrations for the peaks of interest were performed using TopSpin 4.08 (Bruker Corp.). Proportions were expressed on a basis where % pHB = 1/2 pHB_{2/6}/100% aromatics, 100% aromatics = 1/2 (H_{2/6}) + G₂ + 1/2 (S_{2/6} + S'_{2/6}), and 100% aliphatics = A_α + B_α + 1/2 (C_α) + 1/2 (C'_α). Peak annotations followed from previous reports (Kim and Ralph, 2010).

Metabolites and cell-wall-bound *p*-hydroxybenzoate

Developing xylem powder that had been ground with a mortar and pestle and stored at -80 °C was freeze-dried and used for the quantification of soluble and cell-wall-bound pHBA. 10-mg portions of wood powder were weighed into 2-mL screw-cap vials and 950 µL of 80% methanol was added along with 50 µL of 1 mg mL⁻¹ *o*-anisic acid as an internal standard. Once firmly sealed, the vials were incubated for 3 h at 50 °C in a hybridiser-incubator rotating at 40 rpm. After cooling for 5 min at -20 °C, the phenolic extracts were centrifuged for 5 min at 16 000 *g* and filtered using 0.45-µm nylon syringe filters. The remaining wood powder was set aside for further analysis of cell-wall-bound pHB, whereas 200-µL aliquots of the methanolic extracts were used for the analysis of free pHBA.

Saponification was then used to cleave ester-linked conjugates of pHBA in the methanolic extracts. 200-µL aliquots of the extracts were transferred into new vials, 10 µL of 0.2 M NaOH was added to reduce the volatility by converting to sodium salts, and the methanol was evaporated using a vacuum centrifuge. To the dried residue, 300 µL of 2 M NaOH was added and the vials were incubated for 24 h at 30 °C in a hybridiser-incubator. 180-µL aliquots of the saponified extracts were used for quantification of ester-linked pHBA metabolites after acidifying with 18 µL of 72% H₂SO₄.

To cleave ether-linked conjugates, the saponified methanolic extracts were subjected to acid hydrolysis. 120-µL aliquots of the saponified extracts were transferred into new vials and 120 µL of 3 M HCl was added. The vials were incubated at 95 °C for 3 h

with mixing every 15 min, then cooled at 20 °C for 5 min and used for the quantification of ether-linked *p*HBA metabolites.

The wood powder remaining after extraction with methanol was used for the quantification of cell-wall-bound *p*HB in developing xylem. To ensure that all soluble metabolites were removed, three additional extractions were performed with 1 mL acetone each for 24 h in a hybridiser-incubator at 30 °C. Residual acetone was removed by evaporation in a vacuum centrifuge and the dried wood powder was accurately weighed into fresh vials. To release cell-wall-bound *p*HB groups, 500 µL of 2 M NaOH was added which contained 50 µg mL⁻¹ of 3,4,5-trimethoxycinnamic acid as an internal standard. After incubating for 24 h at 30 °C in a hybridiser-incubator, the saponification reactions were acidified with 50 µL of 72% H₂SO₄, and the remaining wood powder was removed by centrifugation at 16 000 *g* for 5 min. The amount of cell-wall-bound *p*HB in mature xylem was determined in the same manner except that acetone-extracted 40-mesh wood powder was used directly.

Quantification of *p*HBA was performed using an Agilent Infinity II 1290 UHPLC apparatus (Agilent Technologies Inc., Santa Clara, CA, USA) equipped with a Zorbax Eclipse Plus C-18 column (2.1 × 50 mm, 1.8 µm particle size; Agilent) and a diode array detector. Good peak separation was achieved using an injection volume of 1 µL, a flow rate of 0.5 mL min⁻¹, and a binary gradient of 5%–25% of acetonitrile containing 0.01% trifluoroacetic acid (eluent A) with water containing 0.01% trifluoroacetic acid (eluent B) over 6 min. Integration of *p*HBA, *o*-anisic acid, and 3,4,5-trimethoxycinnamic acid peaks was performed at the UV maxima of 255, 296, and 304 nm, respectively. Subtraction of the amount of free *p*HBA in methanolic extracts from the total following saponification yielded the amount of alkali-labile *p*HBA. Similarly, subtraction of the amounts of free and alkali-labile *p*HBA from the total following acid hydrolysis yielded the amount of *p*HBA released by alkali and acid. A five-point calibration curve prepared with authentic standards enabled quantification.

Identification of metabolites was performed by HRMS using a Nexara X2 UHPLC apparatus (Shimadzu Corp., Kyoto, Japan) equipped with a diode array detector and an Impact II Ultra-High Resolution Qq-Time-of-Flight mass spectrometer (Bruker Corp.). Methanolic extracts were prepared using freeze-dried samples of developing xylem, as described above.

First, 5-µL samples were injected onto a Symmetry-C18 column (4.6 × 250 mm, 5-µm particle size, Waters Corp., Milford, MA, USA). The mobile phase was a binary gradient of 0.1% formic acid in water (eluent A) and 0.1% formic acid in 3:1 acetonitrile/methanol (eluent B) with a flow rate of 1.0 mL min⁻¹. The gradient programme was applied over 60 min as follows: 5% eluent B for 1 min, then ramp to 25% eluent B at 20 min, continue ramping to 99% eluent B at 42 min, hold until 50 min, then return to 5% eluent B at 55 min and hold until the end of the programme at 60 min to recondition the column. UV detection was performed from 250–600 nm with 1-nm resolution and 1.5625-Hz sampling rate. HRMS chromatograms were acquired from *m/z* 50–1000 with the electrospray-ionization source operating in negative-ion mode with a sampling rate of 2 Hz. The MS configuration was as follows: capillary voltage, +3.5 kV; nebulising gas, nitrogen at 4 bar; temperature, 210 °C; drying gas, nitrogen at 8 L min⁻¹; collision energy, 20 eV. The HRMS chromatograms were analysed using MetaboScape 2021 (Bruker Corp.). Calibration of the MS detector was performed independently for each sample using 10 mM sodium formate in

50:50 isopropanol/water, which was injected into the ionization source at 1.7–1.75 min as the mobile phase was switched from bypassing the MS source to pumping into the electrospray ionization source. The bucket table selection criteria were as follows: intensity threshold, 3000 counts; minimum peak length, 8 spectra; time range, 5–35 min; mass range, *m/z* 50–1000. The peaks were validated by assessing the extracted ion chromatograms, the MS–MS fragments, and the average mass spectra over the elution peak. Chemical formulae were calculated using the SmartFormula algorithm for compounds containing CHNOPSNa with the maximum limit set to C_xH_yN₂P₁S₁Na₁ and a scoring tolerance of *mSigma* <50. Putative metabolite identities were deduced from these chemical formulae using CompoundCrawler to search the PubChem and Chemical Entities of Biological Interest (ChEBI) databases. Finally, the putative identities were further scrutinized using MetFrag to examine the predicted fragment ions.

The retention times and MS–MS fragmentation patterns were measured for standards of the phenolic glucosides: 4-(β-D-glucosyloxy)-benzoic acid (compound 1, 6.7 min), 4-(β-D-glucosyloxy)-salicylic acid (compound 5, 12.6 min), 4-(β-D-glucosyloxy)-vanillic acid (compound 3, 8.4 min), 4-(β-D-glucosyloxy)-isovanillic acid (10.9 min, not found in extracts), for the acid glucoside 1-*O*-(4-hydroxybenzoyl)-β-D-glucopyranoside (compound 4, 9.19 min), for free *p*HBA (compound 6, 13.2 min), and for methyl paraben (compound 7, 25.3 min). The *p*HBA was purchased from MilliporeSigma, whereas the glucosides were synthesized following previously published protocols (Klick and Herrmann, 1988).

Monolignol-*p*-hydroxybenzoate conjugates

DFRC was used to identify ester-linked conjugate groups in lignin (Regner *et al.*, 2018). To a 50-mg sample of extractive-free 40-mesh wood powder in a small glass vial, 3 mL of brominating solution was added containing 20% acetyl bromide in acetic acid. After incubating at 50 °C for 2 h with stirring, the acetyl bromide was evaporated in a vacuum centrifuge. A 0.5-mL portion of ethanol was added to quench any residual acetyl bromide, and this too was then evaporated. Next, 100–150 mg of nanopowdered zinc was added along with 5 mL of a 5:4:1 mixture of 1,4-dioxane/acetic acid/water and the vial was incubated for 1 h in an ultrasonic water bath. After spiking with an internal standard solution containing deuterated synthetic standards, the reaction products were extracted in a separatory funnel charged with 10 mL of saturated ammonium chloride solution using 3 × 10 mL of dichloromethane. The combined organic phases were dried with anhydrous sodium sulphate, filtered into a round-bottom flask through filter paper, and evaporated using a rotary evaporator apparatus. The dried residue was acetylated by adding 1 mL each of acetic anhydride and pyridine, which was then also removed by evaporation. The residue was dissolved in 5 mL of dichloromethane and quantitatively transferred to a Supelclean LC-SI filter (500 mg silicon dioxide; Supelco Analytical, MilliporeSigma) and then eluted using 10 mL of a 1:1 mixture of ethyl acetate/hexanes. After drying with a rotary evaporator, the residue was dissolved in 1 mL of dichloromethane and the reaction products were analysed using a TQ8030 GC–MS (Shimadzu Corp.) equipped with an RXi-5Sil MS column (Restek Corp., Bellefonte, PA, USA) with helium as the carrier gas. Ions corresponding to *cis* and *trans* 4-acetoxysinapyl *p*-acetoxybenzoate were detected using multiple reaction monitoring mass detection as described previously (Regner *et al.*, 2018). No peaks were

detected for the G analogue, 4-acetoxyconiferyl *p*-acetoxybenzoate.

Replicates and statistical analysis

For the measurement of transgene expression, tree height, stem diameter, lignin content, lignin composition, cell-wall-bound *p*HB, soluble *p*HBA metabolites, and structural polysaccharides, five biological replicates were analysed per line and triplicate samples were processed for each. The DFRC reactions and the saccharification analysis used three biological replicates per line measured in triplicate. Three biological replicates were analysed using 2D-NMR and gel-permeation chromatography. Although all lines were analysed by 2D-NMR, representative spectra are only provided for line 5. HRMS of metabolites was performed using five biological replicates each of line 5 and the WT control. And finally, individual representative images were selected for xylem histology and ToF-SIMS imaging. For the HRMS analysis, Student's *t*-test was performed in Microsoft Excel using a two-tail, heteroscedastic model with any missing data ("0") imputed as 213.5 counts (Tables S2 and S3). Metabolite differences with a *P*-value of 0.01 and a two-fold change in signal intensity were deemed to be statistically significant. Student's *t*-test with equal variances was used for the ToF-SIMS image analysis. For all other analyses, statistically significant differences between transgenic lines and the WT control were evaluated using a one-way ANOVA with a *post hoc* Dunnett's test performed using SPSS Statistics 27 (IBM, Armonk, NY, USA).

Acknowledgements

This work was supported by the Great Lakes Bioenergy Research Center, U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under award DE-SC0018409.

Conflicts of interest

The authors report no conflicts of interest.

Author contributions

YM devised the study with help from SDM and JR. YM, SDK, and REG collected the data and analysed the results. YM prepared the manuscript and all authors contributed to the final version.

References

- Aalto, T.R., Firman, M.C. and Rigler, N.E. (1953) *p*-Hydroxybenzoic acid esters as preservatives: I. Uses, antibacterial and antifungal studies, properties and determination. *J. Am. Pharm. Assoc., Sci. Ed.* **42**, 449–457.
- Babst, B.A., Harding, S.A. and Tsai, C.-J. (2010) Biosynthesis of phenolic glycosides from phenylpropanoid and benzenoid precursors in *Populus*. *J. Chem. Ecol.* **36**, 286–297.
- Becker, J. and Wittmann, C. (2019) A field of dreams: lignin valorization into chemicals, materials, fuels, and health-care products. *Biotechnol. Adv.* **37**, 107360.
- Beckham, G.T., Johnson, C.W., Karp, E.M., Salvachúa, D. and Vardon, D.R. (2016) Opportunities and challenges in biological lignin valorization. *Curr. Opin. Biotechnol.* **42**, 40–53.
- Boeckler, G.A., Gershenzon, J. and Unsicker, S.B. (2011) Phenolic glycosides of the Salicaceae and their role as anti-herbivore defences. *Phytochemistry*, **72**, 1497–1509.
- Braconnot, H. (1830) Examen chimique de l'écorce de tremble. De la présence d'une quantité remarquable de salicine dans plusieurs espèces de peupliers. Nouveau principe immédiat. (la populine) [Chemical examination of aspen bark. The presence of a remarkable amount of salicin in several species of poplars. New principal intermediate (populin)]. *Ann. Chim. Phys.* **44**, 296–314 [In French].
- Ceulemans, R., Stettler, R.F., Hincley, T.M., Isebrands, J.G. and Heilman, P.E. (1990) Crown architecture of *Populus* clones as determined by branch orientation and branch characteristics. *Tree Physiol.* **7**, 157–167.
- Chua, M.G.S. and Wayman, M. (1979) Characterization of autohydrolysis aspen (*P. tremuloides*) lignins. Part 1. Composition and molecular weight distribution of extracted autohydrolysis lignin. *Can. J. Chem.* **57**, 1141–1149.
- Coleman, H.D., Park, J.Y., Nair, R., Chapple, C. and Mansfield, S.D. (2008) RNAi-mediated suppression of *p*-coumaroyl-CoA 3'-hydroxylase in hybrid poplar impacts lignin deposition and soluble secondary metabolism. *Proc. Natl. Acad. Sci. USA*, **105**, 4501–4506.
- de Vries, L., MacKay, H.A., Smith, R.A., Mottiar, Y., Karlen, S.D., Unda, F., Muirragui, E. et al. (2022) *p*HBMT1, a BAHD-family monolignol acyltransferase, mediates lignin acylation in poplar. *Plant Physiol.* **188**, 1014–1027.
- Del Río, J.C., Rencoret, J., Gutiérrez, A., Kim, H. and Ralph, J. (2022) Unconventional lignin monomers—extension of the lignin paradigm. *Adv. Bot. Res.* **104**, 1–39.
- Devappa, R.K., Rakshit, S.K. and Dekker, R.F.H. (2015) Forest biorefinery: potential of poplar phytochemicals as value-added co-products. *Biotechnol. Adv.* **33**, 681–716.
- Dixon, R.A. and Barros, J. (2019) Lignin biosynthesis: old roads revisited and new roads explored. *Open Biol.* **9**, 190215.
- El-Basyouni, S.Z., Chen, D., Ibrahim, R.K., Neish, A.C. and Towers, G.H.N. (1964) The biosynthesis of hydroxybenzoic acids in higher plants. *Phytochemistry*, **3**, 485–492.
- Erickson, R.L., Pearl, I.A. and Darling, S.F. (1970) Populoside and grandidentoside from the bark of *Populus grandidentata*. *Phytochemistry*, **9**, 857–863.
- Faleva, A.V., Kozhevnikov, A.Y., Pokryshkin, S.A., Belesov, A.V. and Pikovskoi, I.I. (2020) Structural characterization of the lignin from *Saxifraga oppositifolia* (L.) stems. *Int. J. Biol. Macromol.* **155**, 656–665.
- Fellenberg, C., Corea, O., Yan, L.-H., Archinuk, F., Piirtola, E.-M., Gordon, H., Reichelt, M. et al. (2020) Discovery of salicyl benzoate UDP-glycosyltransferase, a central enzyme in poplar salicinoid phenolic glycoside biosynthesis. *Plant J.* **102**, 99–115.
- Freudenberg, K. and Neish, A.C. (1968) *Constitution and Biosynthesis of Lignin*. Berlin–Heidelberg; New York, NY, USA: Springer-Verlag.
- Goacher, R.E., Mottiar, Y. and Mansfield, S.D. (2021) ToF-SIMS imaging reveals that *p*-hydroxybenzoate groups specifically decorate the lignin of fibres in the xylem of poplar and willow. *Holzforschung*, **75**, 452–462.
- Hatfield, R.D., Marita, J.M. and Frost, K. (2008) Characterization of *p*-coumarate accumulation, *p*-coumaroyl transferase, and cell wall changes during the development of corn stems. *J. Sci. Food Agric.* **88**, 2529–2537.
- Herrmann, K. (1978) Hydroxymethylsäuren und Hydroxybenzoesäuren enthaltende Naturstoffe in Pflanzen [Hydroxycinnamic acid and hydroxybenzoic acid-containing natural substances in plants]. In *Fortschritte der Chemie Organischer Naturstoffe [Progress in the Chemistry of Organic Natural Products]* (Herz, W., Grisebach, H. and Kirby, G.W., eds), pp. 73–132. Vienna, Austria: Springer-Verlag. [In German].
- Hibino, T., Shibata, D., Ito, T., Tsuchiya, D., Higuchi, T., Pollet, B. and Lapierre, C. (1994) Chemical properties of lignin from *Aralia cordata*. *Phytochemistry* **37**, 445–448.
- Huntley, S.K., Ellis, D., Gilbert, M., Chapple, C. and Mansfield, S.D. (2003) Significant increases in pulping efficiency in C4H-F5H-transformed poplars: improved chemical savings and reduced environmental toxins. *J. Agric. Food Chem.* **51**, 6178–6183.
- Jackson, W.J., Jr. and Kuhfuss, H.F. (1976) Liquid crystal polymers. I. Preparation and properties of *p*-hydroxybenzoic acid copolyesters. *J. Polym. Sci., Part A: Polym. Chem.* **14**, 2043–2058.
- Kamimura, N., Takahashi, K., Mori, K., Araki, T., Fujita, M., Higuchi, Y. and Masai, E. (2017) Bacterial catabolism of lignin-derived aromatics: new findings in a recent decade. *Environ. Microbiol. Rep.* **9**, 679–705.

- Karlen, S.D., Fasahati, P., Mazaheri, M., Serate, J., Smith, R.A., Sirobhusanam, S., Chen, M. et al. (2020) Assessing the viability of recovery of hydroxycinnamic acid from lignocellulosic biorefinery alkaline pretreatment waste streams. *ChemSusChem*, **13**, 2012–2024.
- Katsumata, T., Shige, H. and Ejiri, S.-I. (1989) UDP glucose: 4-(β -D-glucopyranosyloxy) benzoic acid glucosyltransferase from the pollen of *Pinus densiflora*. *Phytochemistry*, **28**, 359–362.
- Kim, H. and Ralph, J. (2010) Solution-state 2D NMR of ball-milled plant cell wall gels in DMSO- d_6 /pyridine- d_5 . *Org. Biomol. Chem.* **8**, 576–591.
- Kim, H., Li, Q., Karlen, S.D., Smith, R.A., Shi, R., Liu, J., Yang, C. et al. (2020) Monolignol benzoates incorporate into the lignin of transgenic *Populus trichocarpa* depleted in C3H and C4H. *ACS Sustainable Chem. Eng.* **8**, 3644–3654.
- Klick, S. and Herrmann, K. (1988) Glucosides and glucose esters of hydroxybenzoic acids in plants. *Phytochemistry*, **27**, 2177–2180.
- Köhle, A., Sommer, S., Li, S.-M., Schilde-Rentschler, L., Ninnemann, H. and Heide, L. (2003) Secondary metabolites in transgenic tobacco and potato: high accumulation of 4-hydroxybenzoic acid glucosides results from high expression of the bacterial gene *ubiC*. *Mol. Breed.* **11**, 15–24.
- Le Roy, J., Huss, B., Creach, A., Hawkins, S. and Neutelings, G. (2016) Glycosylation is a major regulator of phenylpropanoid availability and biological activity in plants. *Front. Plant Sci.* **7**, 735.
- Leplé, J.C., Dauwe, R., Morreel, K., Storme, V., Lapierre, C., Pollet, B., Naumann, A. et al. (2007) Downregulation of cinnamoyl-coenzyme A reductase in poplar: multiple-level phenotyping reveals effects on cell wall polymer metabolism and structure. *Plant Cell*, **19**, 3669–3691.
- Li, S.-M., Wang, Z.-X., Wemakor, E. and Heide, L. (1997) Metabolization of the artificial secondary metabolite 4-hydroxybenzoate in *ubiC*-transformed tobacco. *Plant Cell Physiol.* **38**, 844–850.
- Lim, E.-K., Doucet, C.J., Li, Y., Elias, L., Worrall, D., Spencer, S.P., Ross, J. et al. (2002) The activity of *Arabidopsis* glycosyltransferases toward salicylic acid, 4-hydroxybenzoic acid, and other benzoates. *J. Biol. Chem.* **277**, 586–592.
- Lu, F. and Ralph, J. (2008) Novel tetrahydrofuran structures derived from β - β -coupling reactions involving sinapyl acetate in Kenaf lignins. *Org. Biomol. Chem.* **6**, 3681–3694.
- Lu, F., Ralph, J., Morreel, K., Messens, E. and Boerjan, W. (2004) Preparation and relevance of a cross-coupling product between sinapyl alcohol and sinapyl *p*-hydroxybenzoate. *Org. Biomol. Chem.* **2**, 2888–2890.
- Manuja, R., Sachdeva, S., Jain, A. and Chaudhary, J. (2013) A comprehensive review on biological activities of *p*-hydroxy benzoic acid and its derivatives. *Int. J. Pharm. Sci. Rev. Res.* **22**, 109–115.
- McQualter, R.B., Chong, B.F., Meyer, K., Van Dyk, D.E., O'Shea, M.G., Walton, N.J., Viitanen, P.V. et al. (2005) Initial evaluation of sugarcane as a production platform for *p*-hydroxybenzoic acid. *Plant Biotech. J.* **3**, 29–41.
- Meyermans, H., Morreel, K., Lapierre, C., Pollet, B., De Bruyn, A., Busson, R., Herdewijn, P. et al. (2000) Modifications in lignin and accumulation of phenolic glucosides in poplar xylem upon down-regulation of caffeoyl-coenzyme A *O*-methyltransferase, an enzyme involved in lignin biosynthesis. *J. Biol. Chem.* **275**, 36899–36909.
- Morreel, K., Ralph, J., Kim, H., Lu, F., Goeminne, G., Ralph, S., Messens, E. et al. (2004) Profiling of oligolignols reveals monolignol coupling conditions in lignifying poplar xylem. *Plant Physiol.* **136**, 3537–3549.
- Mottiar, Y. and Mansfield, S.D. (2022) Lignin *p*-hydroxybenzoylation is negatively correlated with syringyl units in poplar. *Front. Plant Sci.* **13**, 938083.
- Mottiar, Y., Vanholme, R., Boerjan, W., Ralph, J. and Mansfield, S.D. (2016) Designer lignins: harnessing the plasticity of lignification. *Curr. Opin. Biotechnol.* **37**, 190–200.
- Nishizaki, Y., Sasaki, N., Yasunaga, M., Miyahara, T., Okamoto, E., Okamoto, M., Hirose, Y. et al. (2014) Identification of the glucosyltransferase gene that supplies the *p*-hydroxybenzoyl-glucose for 7-polyacylation of anthocyanin in delphinium. *J. Exp. Bot.* **65**, 2495–2506.
- Parr, A.J., Ng, A. and Waldron, K.W. (1997) Ester-linked phenolic components of carrot cell walls. *J. Agric. Food Chem.* **45**, 2468–2471.
- Pearl, I.A. and Darling, S.F. (1959) Studies on the barks of the family Salicaceae. I. Tremuloidin, a new glucoside from the bark of *Populus tremuloides*. *J. Org. Chem.* **24**, 731–735.
- Pearl, I.A. and Darling, S.F. (1962) Studies on the barks of the family Salicaceae. V. Grandidentatin, a new glucoside from the bark of *Populus grandidentata*. *J. Org. Chem.* **27**, 1806–1809.
- Pearl, I.A., Beyer, D.L. and Laskowski, D. (1959) Alkaline hydrolysis of representative palms. *TAPPI J.* **42**, 779–782.
- Peng, X.-P., Sun, S.-L., Wen, J.-L., Yin, W.-L. and Sun, R.-C. (2014) Structural characterization of lignins from hydroxycinnamoyl transferase (HCT) down-regulated transgenic poplars. *Fuel*, **134**, 485–492.
- Qian, Y., Lynch, J.H., Guo, L., Rhodes, D., Morgan, J.A. and Dudareva, N. (2019) Completion of the cytosolic post-chorismate phenylalanine biosynthetic pathway in plants. *Nat. Commun.* **10**, 15.
- Ralph, J. (1996) An unusual lignin from Kenaf. *J. Nat. Prod.* **59**, 341–342.
- Ralph, J., Hatfield, R.D., Quideau, S., Helm, R.F., Grabber, J.H. and Jung, H.-J.G. (1994) Pathway of *p*-coumaric acid incorporation into maize lignin as revealed by NMR. *J. Am. Chem. Soc.* **116**, 9448–9456.
- Ralph, J., Lundquist, K., Brunow, G., Lu, F., Kim, H., Schatz, P.F., Marita, J.M. et al. (2004) Lignins: natural polymers from oxidative coupling of 4-hydroxyphenylpropanoids. *Phytochem. Rev.* **3**, 29–60.
- Regner, M., Bartuce, A., Padmakshan, D., Ralph, J. and Karlen, S.D. (2018) Reductive cleavage method for quantitation of monolignols and low-abundance monolignol conjugates. *ChemSusChem* **11**, 1600–1605.
- Rencoret, J., Marques, G., Serrano, O., Kaal, J., Martínez, A.T., del Río, J.C. and Gutiérrez, A. (2020) Deciphering the unique structure and acylation pattern of *Posidonia oceanica* lignin. *ACS Sustainable Chem. Eng.* **8**, 12521–12533.
- Rinaldi, R., Jastrzebski, R., Clough, M.T., Ralph, J., Kennema, M., Bruijninx, P.C.A. and Weckhuysen, B.M. (2016) Paving the way for lignin valorisation: recent advances in biorefining, biorefining and catalysis. *Angew. Chem. Int. Ed.* **55**, 8164–8215.
- Robinson, A.R. and Mansfield, S.D. (2009) Rapid analysis of poplar lignin monomer composition by a streamlined thioacidolysis procedure and near-infrared reflectance-based prediction modeling. *Plant J.* **58**, 706–714.
- Schnitzler, J.-P. and Seitz, H.U. (1989) Rapid responses of cultured carrot cells and protoplasts to an elicitor from the cell wall of *Pythium aphanidermatum* (Edson) Fitzp. *Z. Naturforsch.* **44c**, 1020–1028.
- Sibout, R., Le Bris, P., Legée, F., Cézard, L., Renault, H. and 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.
- Siebert, M., Severin, K. and Heide, L. (1994) Formation of 4-hydroxybenzoate in *Escherichia coli*: characterization of the *ubiC* gene and its encoded enzyme chorismate pyruvate-lyase. *Microbiology*, **140**, 897–904.
- Siebert, M., Sommer, S., Li, S.-M., Wang, Z.-X., Severin, K. and Heide, L. (1996) Genetic engineering of plant secondary metabolism: accumulation of 4-hydroxybenzoate glucosides as a result of the expression of the bacterial *ubiC* gene in tobacco. *Plant Physiol.* **112**, 811–819.
- Smith, D.C.C. (1955) *p*-Hydroxybenzoate groups in the lignin of aspen (*Populus tremula*). *J. Chem. Soc.* **3**, 2347–2351.
- Smith, R.A., Gonzales-Vigil, E., Karlen, S.D., Park, J.-Y., Lu, F., Wilkerson, C.G., Samuels, L. et al. (2015) Engineering monolignol *p*-coumarate conjugates into poplar and *Arabidopsis* lignins. *Plant Physiol.* **169**, 2992–3001.
- Sommer, S. and Heide, L. (1998) Expression of bacterial chorismate pyruvate-lyase in tobacco: Evidence for the presence of chorismate in the plant cytosol. *Plant Cell Physiol.* **39**, 1240–1244.
- Stepanova, E.V., Belyanin, M.L. and Filimonov, V.D. (2014) Synthesis of acyl derivatives of salicin, salirepin, and arbutin. *Carbohydr. Res.* **388**, 105–111.
- Stewart, J.J., Akiyama, T., Chapple, C., Ralph, J. and Mansfield, S.D. (2009) The effects on lignin structure of overexpression of ferulate 5-hydroxylase in hybrid poplar. *Plant Physiol.* **150**, 621–635.
- Takahama, U., Oniki, T. and Shimokawa, H. (1996) A possible mechanism for the oxidation of sinapyl alcohol by peroxidase-dependent reactions in the apoplast: enhancement of the oxidation by hydroxycinnamic acids and components of the apoplast. *Plant Cell Physiol.* **37**, 499–504.
- Terashima, N., Mori, I. and Kanda, T. (1975) Biosynthesis of *p*-hydroxybenzoic acid in poplar lignin. *Phytochemistry*, **14**, 1991–1992.
- Thieme, H. and Benecke, R. (1967) Isolierung eines neuen Phenolglykosids aus *Populus nigra* L. [Isolation of a new phenylglycoside from *Populus nigra* L.]. *Pharmazie*, **22**, 59–60 [In German].

- Tsai, C.-J., Xu, P., Xue, L.-J., Hu, H., Nyamdari, B., Naran, R., Zhou, X. et al. (2020) Compensatory guaiacyl lignin biosynthesis at the expense of syringyl lignin in *4CL1*-knockout poplar. *Plant Physiol.* **183**, 123–136.
- Viitanen, P.V., Devine, A.L., Khan, M.S., Deuel, D.L., Van Dyk, D.E. and Daniell, H. (2004) Metabolic engineering of the chloroplast genome using the *Escherichia coli ubiC* gene reveals that chorismate is a readily abundant plant precursor for *p*-hydroxybenzoic acid biosynthesis. *Plant Physiol.* **136**, 4048–4060.
- Wang, J.P., Matthews, M.L., Williams, C.M., Shi, R., Yang, C., Tunlaya-Anukit, S., Chen, H.C. et al. (2018) Improving wood properties for wood utilization through multi-omics integration in lignin biosynthesis. *Nat. Commun.* **9**, 1579.
- Wang, S., Bilal, M., Hu, H., Wang, W. and Zhang, X. (2018) 4-Hydroxybenzoic acid – a versatile platform intermediate for value-added compounds. *Appl. Microbiol. Biotechnol.* **102**, 3561–3571.
- Wang, S., Li, Y., He, L., Yang, J., Fernie, A.R. and Luo, J. (2022) Natural variance at the interface of plant primary and specialized metabolism. *Curr. Opin. Plant Biol.* **67**, 102201.
- Wattiez, N. (1931) Le <<Saliréposide>>, hétéroside nouveau retiré des écorces de *Salix repens* L. (Salicacées) [Salireposide, a new glycoside from the bark of *Salix repens* L. (Salicaceae)]. *Bull. Soc. Chim. Biol.* **13**, 658 [In French].
- Weng, J.K. and Chapple, C. (2010) The origin and evolution of lignin biosynthesis. *New Phytol.* **187**, 273–285.
- Widhalm, J.R. and Dudareva, N. (2015) A familiar ring to it: biosynthesis of plant benzoic acids. *Mol. Plant* **8**, 83–97.
- Withers, S., Lu, F., Kim, H., Zhu, Y., Ralph, J. and Wilkerson, C.G. (2012) Identification of grass-specific enzyme that acylates monolignols with *p*-coumarate. *J. Biol. Chem.* **287**, 8347–8355.
- Yazaki, K., Heide, L. and Tabata, M. (1991) Formation of *p*-hydroxybenzoic acid from *p*-coumaric acid by cell free extract of *Lithospermum erythrorhizon* cell cultures. *Phytochemistry*, **30**, 2233–2236.
- Yazaki, K., Inushima, K., Kataoka, M. and Tabata, M. (1995) Intracellular localization of UDPG: *p*-hydroxybenzoate glucosyltransferase and its reaction product in *Lithospermum* cell cultures. *Phytochemistry*, **38**, 1127–1130.
- Zenk, M.H. and Müller, G. (1964) Biosynthese von *p*-hydroxybenzoesäure und anderer Benzoesäuren in höheren Pflanzen [Biosynthesis of *p*-hydroxybenzoic acid and other benzoic acids in higher plants]. *Z. Naturforsch.* **19b**, 398–405 [In German].
- Zhao, Y., Yu, X., Lam, P.-Y., Zhang, K., Tobimatsu, Y. and Liu, C.-J. (2021a) Monolignol acyltransferase for lignin *p*-hydroxybenzoylation in *Populus*. *Nat. Plants*, **7**, 1288–1300.
- Zhao, Y., Yu, X.-H. and Liu, C.-J. (2021b) The inducible accumulation of cell wall-bound *p*-hydroxybenzoates is involved in the regulation of gravitropic response of poplar. *Front. Plant Sci.* **12**, 755576.
- Zhong, R., Morrison, W.H., 3rd, Himmelsbach, D.S., Poole, F.L., 2nd and Ye, Z.H. (2000) Essential role of caffeoyl coenzyme A *O*-methyltransferase in lignin biosynthesis in woody poplar plants. *Plant Physiol.* **124**, 563–577.

Supporting information

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

Appendix S1 Additional Experimental Procedures.

Figure S1 DNA sequences.

Figure S2 Aliphatics region of the ^1H - ^{13}C HSQC NMR spectra.

Figure S3 Analysis of leaf metabolites.

Figure S4 Xylem cell wall histology.

Table S1 PCR primers and reaction conditions.

Table S2 Data table for developing xylem metabolites.

Table S3 Data table for leaf metabolites.

Table S4 Composition of structural polysaccharides in the wood.