A Key Role for Apoplastic H$_2$O$_2$ in Norway Spruce Phenolic Metabolism

Teresa Laitinen, Kris Morreel, Nicolas Delhomme, Adrien Gauthier, Bastian Schiiffthaler, Kaloian Nickolov, Günter Brader, Jean-Lim, Teemu H. Teeri, Nathaniel R. Street, Wout Boerjan, and Anna Kärkönen

Department of Agricultural Sciences, Viikki Plant Science Centre, University of Helsinki, 00144 Helsinki, Finland
Ghent University, Department of Plant Biotechnology and Bioinformatics, 9052 Ghent, Belgium
VIB Center for Plant Systems Biology, 9052 Ghent, Belgium
Umeå Plant Science Centre, Department of Forest Genetics and Plant Physiology, Swedish University of Agricultural Sciences, 90183 Umeå, Sweden
Umeå Plant Science Centre, Department of Plant Physiology, Umeå University, 90187 Umeå, Sweden
Department of Biology, University of Oulu, 90014 Oulu, Finland
Department of Biosciences, University of Helsinki, 00144 Helsinki, Finland
Natural Resources Institute Finland (Luke), Green Technology, 00790 Helsinki, Finland

ORCID IDs: 0000-0002-3121-9705 (K.M.); 0000-0002-3053-0796 (N.D.); 0000-0002-1351-353X (A.G.); 0000-0002-9771-467X (B.S.); 0000-0002-5135-5691 (K.N.); 0000-0002-6682-7110 (G.B.); 0000-0003-2147-0215 (K.-J.L.); 0000-0002-3812-7213 (T.H.T.); 0000-0001-6031-005X (N.R.S.); 0000-0003-1495-510X (W.B.); 0000-0001-8870-3250 (A.K.).

Apoplastic events such as monolignol oxidation and lignin polymerization are difficult to study in intact trees. To investigate the role of apoplastic hydrogen peroxide (H$_2$O$_2$) in gymnosperm phenolic metabolism, an extracellular lignin-forming cell culture of Norway spruce (Picea abies) was used as a research model. Scavenging of apoplastic H$_2$O$_2$ by potassium iodide repressed lignin formation, in line with peroxidases activating monolignols for lignin polymerization. Time-course analyses coupled to candidate substrate-product pair network propagation revealed differential accumulation of low-molecular-weight phenolics, including flavonoids, and proanthocyanidins, in lignin-forming and H$_2$O$_2$-scavenging cultures and supported that monolignols are oxidatively coupled not only in the cell wall but also in the cytoplasm, where they are coupled to other monolignols and proanthocyanidins. Dilignol glycoconjugates with reduced structures were found in the culture medium, suggesting that cells are able to transport glycosylated dilignols to the apoplast. Transcriptomic analyses revealed that scavenging of apoplastic H$_2$O$_2$ resulted in remodulation of the transcriptome, with reduced carbon flux into the shikimate pathway propagating down to monolignol biosynthesis. Aggregated coexpression network analysis identified candidate enzymes and transcription factors for monolignol oxidation and apoplastic H$_2$O$_2$ production in addition to potential H$_2$O$_2$ receptors. The results presented indicate that the redox state of the apoplast has a profound influence on cellular metabolism.

Lignin, a phenolic polymer present in cell walls of water-conducting vessels and tracheids and in support-giving sclerenchyma cells, constitutes 20% to 35% of the dry weight of wood, making it the second most abundant terrestrial biopolymer after cellulose. Lignin has a high economic impact, as it hinders the industrial processing of lignocellulosic biomass into pulp and fermentable sugars, leading to large economic costs to extract it from the biomass. On the other hand, lignin is increasingly considered as a potentially valuable product, and new applications and profit revenues for this biopolymer are being identified (Ragauskas et al., 2014; Van den Bosch et al., 2015). In Norway spruce (Picea abies), an economically important conifer in Europe, lignin is composed mainly of coniferyl alcohol giving rise to G units, with a small proportion of p-coumaryl alcohol-derived H units (Boerjan et al., 2003). Although there is some degree of variation among species in the architecture of the monolignol biosynthetic pathway, it is largely consistent across different plant groups (Mottiar et al., 2016). In Arabidopsis (Arabidopsis thaliana), both laccases and peroxidases are required for monolignol oxidation in lignin polymerization, possibly acting sequentially, and/or in different tissues (Berthet et al., 2011; Lee et al., 2013; Novo-Uzal et al., 2013; Zhao et al., 2013; Barros et al., 2015; Shigeto and Tsutsumi, 2016). In Norway spruce, numerous peroxidase and laccase isoenzymes have been reported in the cell walls of developing xylem and in the culture medium of the lignin-forming cell culture utilized in this study (Kärkönen et al., 2002; Fagerstedt et al., 2010; Koutaniemi et al., 2015). Isolated peroxidases and laccases are able to mediate high-molecular-weight dehydrogenation polymer formation from coniferyl alcohol in vitro (Kärkönen et al., 2002; Koutaniemi et al., 2005; Warinowski et al., 2016). However, the relative contribution of these oxidative enzymes to lignin synthesis in vivo and the specific isoenzymes that
oxidize monolignols for lignin biosynthesis are not known.

The high number of putative peroxidase and laccase genes (over 200 and 100, respectively) in the Norway spruce genome (Nystedt et al., 2013) has hindered the identification of specific genes encoding lignification-related isoenzymes. The origin of apoplastic hydrogen peroxide (H2O2) required by peroxidases as an oxidant in cell wall cross-linking also is unclear, as is the contribution of the apoplastic redox state in controlling lignin formation. Several plant cell wall- and plasma membrane-located sources for apoplastic reactive oxygen species (ROS; superoxide, H2O2, and hydroxyl radical) exist (Kärkönen and Kuchitsu, 2015), including apoplastic peroxidases, various oxidases, and plasma membrane-located enzymes such as respiratory burst oxidase homologs (RBOHs; also called NADPH oxidases) and quinone reductases. There are indications that RBOHs produce ROS in zinnia (Zinnia elegans) during vascular xylem lignification (Ros Barceló, 1998) and in in vitro traechey element (TE) formation (Karlsson et al., 2005; Pesquet et al., 2013), with some participation of peroxidases (Karlsson et al., 2005). In addition, in Arabidopsis, ROS produced by AtRBOHF are essential for the formation of lignin-composed Casparian strips in the root endodermis (Lee et al., 2013). In the tissue-cultured spruce cells, at least two H2O2 generation systems are present in the apoplast: one sensitive to diphenylene iodonium, an inhibitor of flavin-containing enzymes such as RBOHs, and the other sensitive to a low concentration of azide, an inhibitor of heme-containing enzymes such as peroxidases (Kärkönen et al., 2009).

The aim of this work was to investigate the effects of apoplastic H2O2 level modulation on phenolic metabolism and lignification in Norway spruce in order to identify genes encoding isoenzymes (e.g., oxidative and ROS-producing enzymes as well as transcription factors) specific for lignin biosynthesis. Studying the role of the apoplastic redox state in lignin biosynthesis in tree stems during wood development is technically challenging, as differentiating xylem cells constitute a thin layer at the inner side of the cambium, with sampling-induced wounding inducing apoplastic ROS production. Therefore, the extracellular lignin-forming cell culture of Norway spruce used here represents a model system for such studies (Simola et al., 1992; Kärkönen and Koutaniemi, 2010). The cultured cells produce extracellular lignin when transferred from solid medium into liquid cultures (Simola et al., 1992; Kärkönen et al., 2002). Results of a comparative real-time reverse transcription-PCR analysis of the cultured cells and several lignin-forming tissues of spruce indicated that the same genes encoding enzymes in the phenylpropanoid and monolignol biosynthesis pathways are induced in the culture system as in developing xylem (Koutaniemi et al., 2007). In addition, the linkage structure of the extracellular lignin resembled that of native wood lignin (Brunow et al., 1993; Koutaniemi et al., 2005). The availability of the Norway spruce genome sequence (Nystedt et al., 2013), in addition to those of other conifer genomes (white spruce [Picea glauca] and loblolly pine [Pinus taeda]; De La Torre et al., 2014), now enables more comprehensive genomic, transcriptomic, and proteomic analyses to investigate the process of lignin biosynthesis in conifers. Here, the lignin-forming cell culture of Norway spruce was used in a systems biology approach combining liquid chromatography-mass spectrometry-based phenolic and RNA sequencing (RNA-Seq)-based transcriptome profiling over the course of lignin biosynthesis to investigate whether the scavenging of apoplastic ROS affects the coupling of monolignols into lignin and intracellular aromatic metabolism. Differential expression and gene coexpression network analyses identified candidate enzymes for monolignol oxidation and apoplastic H2O2 production, novel putative transcription factors, and potential H2O2 receptors. The results suggest that the apoplastic redox state is an important regulatory factor for apoplastic reactions as well as whole cellular metabolism.

RESULTS

Scavenging of H2O2 Inhibits Extracellular Lignin Production

A cell culture of Norway spruce that produces extracellular lignin in the culture medium was used to investigate the effects of modulating apoplastic H2O2.
levels on phenolic metabolism using phenolic and transcriptomic analyses (the experimental setup is presented in Fig. 1 and Supplemental Fig. S1). Spruce callus cells maintained on a solid culture medium were transferred into liquid cultures with a concomitant decrease in 2,4-dichlorophenoxyacetic acid concentration (by 20-fold), and the nutrient medium was supplemented with potassium iodide (KI; 5 mM), an H$_2$O$_2$ scavenger (Nose et al., 1995). Control cells treated with KCl (5 mM) or an equivalent volume of water (to ascertain that KI effects were due to iodide and not potassium) produced extracellular lignin after 5 d in liquid culture, visible as a fine, white precipitate in the culture medium (Figs. 1 and 2A). Cell weight did not increase during this period in any of the treatments (Fig. 2B). The addition of KI substantially decreased H$_2$O$_2$ concentration in the culture medium (Fig. 2C) and prevented extracellular lignin production (H$_2$O$_2$-scavenging conditions; Fig. 2A). The KI-treated cells also started to divide after 3 weeks of culturing, observed as an increase in cell fresh weight (Fig. 2B). KI was washed away after 20 d, and the cells were further cultured in the standard nutrient medium. Four to 12 d after KI removal, extracellular lignin subsequently appeared in the culture medium (lignin reformation; Fig. 1; Supplemental Fig. S1), showing that the prevention of lignin formation was reversible and that KI treatment did not impact cell viability. This is in contrast to the situation in tobacco (Nicotiana tabacum) BY-2 cells, which do not survive in the presence of elevated levels of KI (Väisänen et al., 2015). These results suggest an important role for apoplastic H$_2$O$_2$ in extracellular lignin production. Moreover, as iodide does not inhibit laccase activity in vitro (Ihssen et al., 2014), the results suggest that peroxidases are crucial for monolignol oxidation during extracellular lignin formation.

**H$_2$O$_2$ Scavenging Affects Both the Intracellular and Apoplastic Phenolic Pools**

To investigate the effects of apoplastic H$_2$O$_2$ scavenging on phenolic metabolism, a time-course sampling of cells and culture medium from the various types of culture (Supplemental Fig. S1A) was analyzed for methanol-soluble phenolic compounds using ultra-HPLC-mass spectrometry (Morreel et al., 2004, 2010a, 2010b, 2014). In cells and culture medium, 505 and 278 compounds were profiled, of which only 31 compounds were in common, providing strong evidence that cross-contamination between the medium and cell extracts was low or even absent. Subsequently, a candidate substrate-product pair (CSPP) network (Morreel et al., 2014) was constructed using all 752 (505 + 278 − 31) compounds (Fig. 3). In the CSPP network, nodes represent compounds while edges reflect mass differences corresponding with well-known metabolic conversions (for the included types of metabolic conversions, see Supplemental Table S1). Hence, this method aids structural characterization and provides insight into the various profiled biochemical compound classes. In total, 71 and 20 compounds were structurally elucidated in cells and culture medium, respectively (Table I; Supplemental Table S2). In the constructed CSPP network, the largest subnetworks represented compounds that were either almost exclusively present in the cells (circles, Fig. 3) or in

![Figure 1](https://www.plantphysiol.org/)

**Figure 1.** Setup for the experiments using the extracellular lignin-forming Norway spruce cell culture. Spruce cells were transferred at time 0 from the solid maintenance medium to liquid nutrient medium supplemented either with an H$_2$O$_2$ scavenger, KI (5 mM), or the corresponding volume of water (H$_2$O). After 5 d (H$_2$O$_{5d}$), extracellular lignin became visible in the culture medium of water-treated samples (lignin-forming conditions). In contrast, KI-treated cells did not form any visible extracellular lignin in the culture medium at 5 d (KI$_{5d}$) up to 20 d (KI$_{20d}$; H$_2$O$_2$-scavenging conditions). After 20 d of culturing, KI was removed by washing the cells three times with fresh nutrient medium (~KI) and cultivated further in the fresh nutrient medium without extra KI supplementation (KI$_{20+nd}$; with n, number of days in the culture after the medium-change step; lignin-reforming conditions).
the culture medium (squares, Fig. 3), with very few nodes representing compounds that were detected in both cells and the medium (triangles, Fig. 3). Subnetworks for the compounds in the culture medium represented mainly lignin oligomers, termed oligolignols (Morreel et al., 2004), whereas those for the compounds in the cells were associated predominantly with hexosylated phenylpropanoids, hexosylated flavonoids, and proanthocyanidins (condensed tannins). Especially the latter compounds were much more prevalent in the cell-specific CSPP subnetworks. Upon a targeted search for the presence of any of the monolignols, either as aglycones or glycosylated, only coniferin was observed in the chromatograms and was connected as a cellular compound in the subnetwork of the hexosylated oligolignols (Fig. 3).

The time course-associated abundance changes of metabolites were modeled by piece-wise regression and the metabolites with highly correlated abundances were grouped into cliques, which were further combined into clique clusters whenever they shared more than a defined number of metabolites (see “Materials and Methods”; Fig. 4; Supplemental Fig. S2). In agreement with the absence of any other detectable monolignol, coniferin could not be further grouped into a clique (Fig. 4). Coniferin was detected during the entire time course in the cells of cultures where H$_2$O$_2$ was initially scavenged (KI-treated and lignin-reforming cultures) but was never observed in the lignin-forming cultures. Likely, the coniferin formed in the H$_2$O$_2$-scavenging cultures remained stored in the cells after KI removal at day 20. When exploring the connection between CSPP network nodes and clique cluster membership, CSPP subnetwork nodes associated with cellular phenolic compounds were found to belong almost exclusively to clique cluster 1 (filled blue circles and triangles, Figs. 3 and 4), whereas those associated with phenolic compounds in the culture medium belonged mainly to either clique cluster 1 (squares and triangles with blue border, Figs. 3 and 4) or clique cluster 4 (squares and triangles with yellow border, Figs. 3 and 4). In addition, two and one nodes related to the culture medium (squares, Fig. 3), with very few nodes representing compounds that were detected in both cells and the medium (triangles, Fig. 3). Subnetworks for the compounds in the culture medium represented mainly lignin oligomers, termed oligolignols (Morreel et al., 2004), whereas those for the compounds in the cells were associated predominantly with hexosylated phenylpropanoids, hexosylated flavonoids, and proanthocyanidins (condensed tannins). Especially the latter compounds were much more prevalent in the cell-specific CSPP subnetworks. Upon a targeted search for the presence of any of the monolignols, either as aglycones or glycosylated, only coniferin was observed in the chromatograms and was connected as a cellular compound in the subnetwork of the hexosylated oligolignols (Fig. 3).

The time course-associated abundance changes of metabolites were modeled by piece-wise regression and the metabolites with highly correlated abundances were grouped into cliques, which were further combined into clique clusters whenever they shared more than a defined number of metabolites (see “Materials and Methods”; Fig. 4; Supplemental Fig. S2). In agreement with the absence of any other detectable monolignol, coniferin could not be further grouped into a clique (Fig. 4). Coniferin was detected during the entire time course in the cells of cultures where H$_2$O$_2$ was initially scavenged (KI-treated and lignin-reforming cultures) but was never observed in the lignin-forming cultures. Likely, the coniferin formed in the H$_2$O$_2$-scavenging cultures remained stored in the cells after KI removal at day 20. When exploring the connection between CSPP network nodes and clique cluster membership, CSPP subnetwork nodes associated with cellular phenolic compounds were found to belong almost exclusively to clique cluster 1 (filled blue circles and triangles, Figs. 3 and 4), whereas those associated with phenolic compounds in the culture medium belonged mainly to either clique cluster 1 (squares and triangles with blue border, Figs. 3 and 4) or clique cluster 4 (squares and triangles with yellow border, Figs. 3 and 4). In addition, two and one nodes related to

Figure 2. Scavenging of H$_2$O$_2$ by KI inhibits lignin formation in the Norway spruce cell culture. A, Amount of extracellular lignin formed after culturing of Norway spruce cells in the culture medium supplemented with water (H$_2$O), KCl (5 mM), or KI (5 mM). Culture medium was separated from cells by filtering through Miracloth. Extracellular lignin was pelleted by centrifugation, washed two times with water, lyophilized, and weighed. The experiment was repeated several times with a similar trend in results. Representative results are shown. Mean ± se; n = 3 (treatment replicates). There was a statistically significant difference between the groups (a, b, c, and d) as determined by one-way ANOVA [F(5,12) = 92.758, P = 3.6524E-9, Tukey’s honestly significant difference; SPSS software]. B, Cell growth in the medium supplemented with water or KI (5 mM). After 21 d, KI was removed and cell cultivation was continued in the regular nutrient medium without extra KI supplementation (KI$_{21+}$; with n, number of days of culture after the medium-change step). The values show representative results for experiments with 25 mL of cell culture per 100-mL flask. Mean ± se; n = 3 to 7 (treatment replicates). There was a statistically significant difference between the groups (a, b, c, and d) as determined by one-way ANOVA [F(6,27) = 39.917, P = 3.4759E-12, Tukey’s honestly significant difference; SPSS software]. FW, Cell fresh weight. C, H$_2$O$_2$ concentration (μM) in the culture medium supplemented with water or KI (5 mM). In this culturing experiment, extracellular lignin was clearly visible in the culture medium of water-treated cultures at day 14. After 20 d, KI was removed and cell cultivation was continued in the regular nutrient medium without extra KI supplementation until extracellular lignin was visible at KI$_{20+}$; with n, number of days of culture after the medium-change step). The values show representative results for
Figure 3. CSPP network. In the CSPP network, the nodes representing mass-to-charge ratio (m/z) features specific for cells and culture medium are shown as circles and squares, whereas triangles represent those found in both. The clique cluster (see Fig. 4) to which the m/z feature (node) belongs is indicated by a color. The node itself is colored whenever the clique cluster is derived from the cellular metabolites, whereas the node border color represents clique clusters derived from the culture medium-based metabolites. Only a few representative molecules for the various biochemical classes are shown, except for the CSPP subnetwork.
cellular phenolics of clique cluster 10 (filled yellow squares and triangles, Figs. 3 and 4) and to an apoplastic phenolic compound of clique cluster 20 (squares with orange border, Figs. 3 and 4) were observed. In the culture medium and cells, 114 and 318 compounds, of which 10 and 44 were structurally elucidated, respectively, showed significantly different (P < 0.001) abundances between the lignin-forming and H$_2$O$_2$-scavenging cultures.

In agreement with the known structure of gymnosperm lignin, the oligolignols in the medium of the lignin-forming cell cultures contained units derived from the incorporation of coniferyl alcohol (G units), p-coumaryl alcohol (H units), coniferaldehyde (G’ units), dihydro-p-coumaryl alcohol (DHH units), and dihydroconiferyl alcohol (DHG units) that are connected via 8-8, 8-5, and 8-O-4 linkages (Boerjan et al., 2003). Many of these oligolignols were also observed in the medium of H$_2$O$_2$-scavenging cultures, albeit often at much lower abundances (clique clusters 1, 4, 16, 28, 36, 57, and 58; Supplemental Fig. S2A; exemplified for clique clusters 1 and 4 in Fig. 4). Upon KI removal, higher abundances of these compounds were observed in the medium of lignin-reforming cultures (especially in clique clusters 4, 36, and 58), although these compounds generally did not accumulate to the levels observed in the medium of water-treated, lignin-forming cultures. A few oligolignols were more abundant in the medium of H$_2$O$_2$-scavenging cultures compared with that of lignin-forming cultures (clique clusters 20, 50, 56, and 74; Supplemental Fig. S2A; exemplified for clique cluster 20 in Fig. 4), which could be explained for clique clusters 20 and 56 by their consumption in further oligomerization reactions during lignin formation in the lignin-forming cultures [compare the accumulation pattern of G(8-O-4)DHG in clique cluster 20 with that of G(8-O-4)G(8-O-4)DHG in clique cluster 16 and that of G(8-O-4)G’ in clique cluster 56 with that of G(8-O-4)G(8-O-4)G’ in clique cluster 57; Supplemental Fig. S2A]. Remarkably, two hexosylated dilignols, isodihydrodihydrodiconiferyl alcohol (IDDDC) hexoside and pinoresinol hexoside, were observed in the medium. As IDDDC hexoside was the most abundant structurally characterized phenol in the cells (Table 1; Supplemental Table S3), showing a similar accumulation pattern in the medium (clique cluster 1; Fig. 4) and in the cells (clique cluster 1; Fig. 4), we cannot exclude that its presence in the culture medium might have resulted from a minor cellular contamination during isolation. In contrast, the accumulation patterns of pinoresinol hexoside in the medium (clique cluster 74; Supplemental Fig. S2A) and in the cells (clique cluster 10; Fig. 4) were entirely different, providing evidence for distinct pools of pinoresinol hexoside in the cells and the culture medium. In addition to IDDDC hexoside and pinoresinol hexoside that were present in the cells and the medium, several other dilignol hexosides were detected exclusively in the cells of all cultures, such as G(8-8)H hexoside, lariciresinol hexoside, G(8-5)G hexoside, G(8-O-4)G hexoside, and G(8-O-4)vanilloyl hexose (mainly clique clusters 10 and 69; Fig. 4; Supplemental Fig. S2B). Whereas the former two are 8-8-linked lignan-type dilignols, the latter three are either 8-5- or 8-O-4-linked neolignan-type dilignols. Furthermore, the latter three showed equal abundances in water-treated, lignin-forming cultures compared with H$_2$O$_2$-scavenging cultures, with an increase observed during lignin reformation.

Several types of flavonoids (e.g. dihydroflavonols and flavonols), condensed tannins (proanthocyanidins with few subunits), stilbenes, and hydroxycinnamic acid glycosides also were identified in the cells but not in the medium (clique clusters 1, 10, 23, 39, 52, and 69; Fig. 3; Supplemental Fig. S2B). Many of these compounds appeared at higher abundance in the cells that produced extracellular lignin (especially clique clusters 1, 23, 39, and 52; Fig. 4; Supplemental Fig. S2B). Among them were many flavonoids with antioxidative capacity (e.g. taxifolin, quercetin, and myricetin) as well as condensed tannins. The variety of condensed tannins was partitioned mainly between clique clusters 1 and 10, with those of clique cluster 1 more frequently comprising units derived from (epi)catechin compared with those of clique cluster 10.

**Figure 3.** (Continued.)

comprising monolignols, oligolignols, and their glycosides. The latter subnetwork includes both cell- and culture medium-specific nodes. Nevertheless, both node types are clearly separated, as the left and the central/right part of the subnetwork represent the hexosides (present in cells) and the aglycones (present in medium) of the monolignols and oligolignols, respectively.
Table 1. Abundances of phenolic compounds in the cells and in the culture medium of Norway spruce cell culture

Column A shows lignin forming at 15 d, and column B shows H2O2 scavenging at 15 d. The mass spectrometry (MS) response represents the image current generated in the ion cyclotron resonance cell and could be arbitrarily expressed in charge units. See Supplemental Table S3 to compare with the abundances of a particular compound at other time points. Note that extrapolating concentration differences between compounds from their relative MS response is flawed due to the highly varying ionization and transfer efficiencies of compounds by MS. NA, Not available.

<table>
<thead>
<tr>
<th>Trivial Name</th>
<th>Cells</th>
<th>Culture Medium</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>p-Coumaroyl hexose</td>
<td>2.1E+05</td>
<td>2.3E+04</td>
</tr>
<tr>
<td>Caffeic acid hexoside</td>
<td>3.5E+05</td>
<td>1.8E+05</td>
</tr>
<tr>
<td>Benzoyl hexose</td>
<td>7.7E+05</td>
<td>4.7E+05</td>
</tr>
<tr>
<td>Coniferin</td>
<td>NA</td>
<td>7.1E+05</td>
</tr>
<tr>
<td>Naringenin O-hexoside</td>
<td>2.4E+05</td>
<td>NA</td>
</tr>
<tr>
<td>Naringenin-8-C-hexoside</td>
<td>NA</td>
<td>2.0E+05</td>
</tr>
<tr>
<td>Aromadendrin hexoside</td>
<td>2.2E+05</td>
<td>1.2E+05</td>
</tr>
<tr>
<td>Aromadendrin hexoside</td>
<td>2.5E+05</td>
<td>1.2E+05</td>
</tr>
<tr>
<td>p-Coumaroyl quinoyl hexose</td>
<td>1.4E+05</td>
<td>3.0E+05</td>
</tr>
<tr>
<td>Methylmyricetin hexoside</td>
<td>1.4E+05</td>
<td>2.4E+04</td>
</tr>
<tr>
<td>Syringetin hexoside</td>
<td>NA</td>
<td>6.9E+03</td>
</tr>
<tr>
<td>G(8-O-4)Vanillyl hexose</td>
<td>1.1E+05</td>
<td>2.1E+05</td>
</tr>
<tr>
<td>G(8-O-4)G hexoside</td>
<td>NA</td>
<td>1.7E+05</td>
</tr>
<tr>
<td>G(8-O-4)IDDDC</td>
<td>3.0E+05</td>
<td>1.7E+05</td>
</tr>
<tr>
<td>G(8-O-4)IDDDC</td>
<td>1.0E+06</td>
<td>1.5E+05</td>
</tr>
<tr>
<td>Myricetin dihexoside</td>
<td>2.4E+05</td>
<td>NA</td>
</tr>
<tr>
<td>(Epigallocatechin)</td>
<td>4.9E+05</td>
<td>2.9E+05</td>
</tr>
<tr>
<td>Caffeoyl shikimic acid</td>
<td>2.8E+05</td>
<td>NA</td>
</tr>
<tr>
<td>Cinnamoyl hexose</td>
<td>7.1E+05</td>
<td>3.3E+05</td>
</tr>
<tr>
<td>Vanillyl alcohol hexoside</td>
<td>4.6E+04</td>
<td>9.7E+04</td>
</tr>
<tr>
<td>Transpiceatannol hexoside</td>
<td>1.9E+06</td>
<td>4.0E+05</td>
</tr>
<tr>
<td>Naringenin-6-C-hexoside</td>
<td>NA</td>
<td>1.8E+05</td>
</tr>
<tr>
<td>Kaempherol hexoside</td>
<td>8.1E+05</td>
<td>2.7E+05</td>
</tr>
<tr>
<td>Myricetin deoxyhexoside</td>
<td>3.5E+05</td>
<td>NA</td>
</tr>
<tr>
<td>Quercetin hexoside</td>
<td>1.8E+06</td>
<td>1.6E+05</td>
</tr>
<tr>
<td>Taxifolin hexoside</td>
<td>2.5E+05</td>
<td>NA</td>
</tr>
<tr>
<td>Myricetin hexoside</td>
<td>1.2E+06</td>
<td>6.7E+04</td>
</tr>
<tr>
<td>Myricetin hexoside</td>
<td>6.2E+05</td>
<td>3.7E+04</td>
</tr>
<tr>
<td>G(8-BH) hexoside</td>
<td>1.2E+05</td>
<td>NA</td>
</tr>
<tr>
<td>Methylmyricetin hexoside</td>
<td>1.9E+05</td>
<td>NA</td>
</tr>
<tr>
<td>Pinoresinol hexoside</td>
<td>5.1E+06</td>
<td>4.6E+06</td>
</tr>
<tr>
<td>Lariciresinol hexoside</td>
<td>3.8E+05</td>
<td>5.3E+05</td>
</tr>
<tr>
<td>(Epi)catechin-4,8'/7',2-(epi)catechin</td>
<td>1.1E+06</td>
<td>5.4E+05</td>
</tr>
<tr>
<td>(Epi)catechin-4,8'/7',2-(epi)catechin</td>
<td>4.1E+05</td>
<td>2.5E+05</td>
</tr>
<tr>
<td>(Epi)catechin-4,8'-(epi)catechin</td>
<td>6.4E+06</td>
<td>1.5E+06</td>
</tr>
<tr>
<td>(Epi)catechin-4,8'-(epi)catechin</td>
<td>4.3E+06</td>
<td>9.9E+05</td>
</tr>
<tr>
<td>(Epi)catechin-4,8'-(epi)catechin</td>
<td>2.8E+06</td>
<td>NA</td>
</tr>
<tr>
<td>(Epi)catechin</td>
<td>9.2E+06</td>
<td>2.7E+06</td>
</tr>
<tr>
<td>G(8-5)C hexoside</td>
<td>9.1E+05</td>
<td>1.4E+06</td>
</tr>
<tr>
<td>IDDDC hexoside</td>
<td>3.7E+07</td>
<td>1.4E+07</td>
</tr>
<tr>
<td>(Epi)gallolocatechin-4,8'/7',2-(epi)catechin</td>
<td>4.5E+05</td>
<td>3.8E+05</td>
</tr>
<tr>
<td>(Epi)gallolocatechin-4,8'/7',2-(epi)catechin</td>
<td>7.0E+05</td>
<td>3.4E+05</td>
</tr>
<tr>
<td>(Epi)gallolocatechin-4,8'/7',2-(epi)catechin</td>
<td>8.9E+05</td>
<td>4.9E+05</td>
</tr>
<tr>
<td>(Epi)gallolocatechin-4,8'/7',2-(epi)catechin</td>
<td>9.1E+05</td>
<td>3.0E+05</td>
</tr>
<tr>
<td>(Epi)gallolocatechin-4,8'/7',2-(epi)catechin</td>
<td>3.3E+05</td>
<td>1.0E+05</td>
</tr>
<tr>
<td>(Epi)gallolocatechin-4,8'/7',2-(epi)catechin</td>
<td>1.2E+06</td>
<td>8.7E+05</td>
</tr>
<tr>
<td>(Epi)gallolocatechin-4,8'/7',2-(epi)catechin</td>
<td>2.0E+06</td>
<td>8.9E+05</td>
</tr>
<tr>
<td>(Epi)gallolocatechin-4,8'/7',2-(epi)catechin</td>
<td>8.2E+05</td>
<td>3.2E+05</td>
</tr>
<tr>
<td>Hydroxybenzoic acid hexoside</td>
<td>2.0E+06</td>
<td>3.4E+05</td>
</tr>
<tr>
<td>(Epi)gallolocatechin-4,8'/7',2-(epi)catechin</td>
<td>2.2E+05</td>
<td>2.7E+05</td>
</tr>
<tr>
<td>(Epi)gallolocatechin-4,8'/7',2-(epi)catechin</td>
<td>8.2E+05</td>
<td>3.6E+05</td>
</tr>
<tr>
<td>Quercetin dihexoside</td>
<td>6.9E+05</td>
<td>2.8E+05</td>
</tr>
<tr>
<td>(Epi)catechin-4,8'-(epi)catechin-4',8''/7'',2'-(epi)catechin</td>
<td>1.9E+06</td>
<td>3.6E+05</td>
</tr>
<tr>
<td>(Epi)catechin-4,8'-(epi)catechin-4',8''-epicatechin</td>
<td>1.8E+06</td>
<td>5.0E+05</td>
</tr>
<tr>
<td>(Epi)catechin-4,8'-(epi)catechin-4',8''-(epi)catechin</td>
<td>6.9E+06</td>
<td>7.6E+05</td>
</tr>
<tr>
<td>(Epi)catechin-4,8'-(epi)catechin-4',8''-(epi)catechin</td>
<td>3.0E+06</td>
<td>3.0E+05</td>
</tr>
</tbody>
</table>

(Table continues on following page.)
and non-lignin-forming conditions \((P_{adj} < 0.001;\) Supplemental Table S4). As KI was used as an \(H_2O_2\) scavenger, it is possible that part of the identified differential expression resulted from KI-induced effects rather than from the difference between lignin formation and not. To address this concern, we performed a range of comparisons the results of which indicate that KI did not induce a significant transcriptional response (Fig. 6; Supplemental Table S5; Supplemental Fig. S3; Supplemental Protocol S1).

In the comparison of lignin-forming versus non-lignin-forming conditions, only 5.8% (619) of the genes were affected by KI, and as a caution, these were removed from subsequent analyses. Gene Ontology (GO) category enrichment testing of the differentially expressed genes revealed that large differences in cellular metabolism were seen in lignin-forming cells as compared with non-lignin-forming cells (Supplemental Fig. S4). In lignin-forming conditions, genes involved in carbohydrate, amine, and lipid metabolism and response to stress were up-regulated, for example (Supplemental Fig. S4A). Also, genes encoding enzymes in the shikimate pathway all the way down to monolignol biosynthesis were strongly induced (see below). In non-lignin-forming conditions, genes involved in cell division, DNA replication, and response to various stimuli were enriched (e.g. all putative cyclin genes except one were up-regulated in non-lignin-forming conditions \([P_{adj} < 0.001];\) Supplemental Fig. S4B). Together with microscopic evaluation, these observations suggest that the observed increase in cell fresh weight (Fig. 2B) was due to cell division rather than cell enlargement. These GO over-enrichment results were in agreement with other findings from the study.

### Shikimate, Phe, and Phenylpropanoid Pathways Are Induced in Lignin-Forming, But Not in \(H_2O_2\)-Scavenging Conditions

To understand the role of apoplastic \(H_2O_2\) in monolignol biosynthesis, the expression of genes putatively encoding enzymes in the shikimate, aromatic amino acid, phenylpropanoid, and monolignol biosynthesis pathways was investigated. Examination of the set of annotated gene models (high, medium, and low confidence; Nystedt et al., 2013) identified gene models for all enzymes of these pathways. \(p\)-Coumaroyl shikimate 3-hydroxylase is a P450 enzyme the annotation of which is problematic, as enzyme activity cannot be inferred from the gene sequence. Additionally, there are a considerable number of P450s present in the genome annotation, and these all share high sequence similarity yet will represent a broad potential range of activities.

<table>
<thead>
<tr>
<th>Trivial Name</th>
<th>Cells</th>
<th>Culture Medium</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>(Epi)catechin-4,8'-(epi)catechin-4',8''-(epi)catechin</td>
<td>2.6E+06</td>
<td>9.9E+05</td>
</tr>
<tr>
<td>(Epi)catechin-4,8'-(epi)gallocatechin-4',8''-(epi)catechin</td>
<td>5.9E+05</td>
<td>4.2E+05</td>
</tr>
<tr>
<td>(Epi)gallocatechin-4,8/7,2-(epi)catechin-4',8''-(epi)catechin</td>
<td>2.0E+06</td>
<td>3.4E+05</td>
</tr>
<tr>
<td>(Epi)gallocatechin-4,8'-(epi)catechin-4',8''-(epi)catechin</td>
<td>8.9E+05</td>
<td>3.5E+05</td>
</tr>
<tr>
<td>(Epi)gallocatechin-4,8'/7',2-(epi)gallocatechin-4',8''-(epi)catechin</td>
<td>3.9E+05</td>
<td>1.4E+05</td>
</tr>
<tr>
<td>(Epi)gallocatechin-4,8'/7',2-(epi)gallocatechin-4',8''-(epi)catechin</td>
<td>1.4E+06</td>
<td>5.2E+05</td>
</tr>
<tr>
<td>(Epi)gallocatechin-4,8'/7',2-(epi)gallocatechin-4',8''-(epi)catechin</td>
<td>1.9E+06</td>
<td>3.1E+05</td>
</tr>
<tr>
<td>(Epi)gallocatechin-4,8'/7',2-(epi)gallocatechin-4',8''-(epi)catechin</td>
<td>1.3E+05</td>
<td>1.4E+05</td>
</tr>
<tr>
<td>(Epi)gallocatechin-4,8'/7',2-(epi)gallocatechin-4',8''-(epi)catechin</td>
<td>1.1E+06</td>
<td>3.2E+05</td>
</tr>
<tr>
<td>(Epi)gallocatechin-4,8'/7',2-(epi)gallocatechin-4',8''-(epi)catechin</td>
<td>1.9E+05</td>
<td>1.4E+05</td>
</tr>
<tr>
<td>(Epi)gallocatechin-4,8'/7',2-(epi)gallocatechin-4',8''-(epi)catechin</td>
<td>1.5E+05</td>
<td>1.2E+05</td>
</tr>
<tr>
<td>(Epi)gallocatechin-4,8'/7',2-(epi)gallocatechin-4',8''-(epi)catechin</td>
<td>1.7E+05</td>
<td>1.1E+06</td>
</tr>
<tr>
<td>(Epi)gallocatechin-4,8'/7',2-(epi)gallocatechin-4',8''-(epi)catechin</td>
<td>1.5E+06</td>
<td>5.9E+04</td>
</tr>
<tr>
<td>(Epi)gallocatechin-4,8'/7',2-(epi)gallocatechin-4',8''-(epi)catechin</td>
<td>5.1E+05</td>
<td>5.3E+05</td>
</tr>
<tr>
<td>(Epi)gallocatechin-4,8'/7',2-(epi)gallocatechin-4',8''-(epi)catechin</td>
<td>1.2E+05</td>
<td>5.1E+04</td>
</tr>
<tr>
<td>(Epi)gallocatechin-4,8'/7',2-(epi)gallocatechin-4',8''-(epi)catechin</td>
<td>1.2E+05</td>
<td>1.6E+05</td>
</tr>
<tr>
<td>(Epi)gallocatechin-4,8'/7',2-(epi)gallocatechin-4',8''-(epi)catechin</td>
<td>1.7E+06</td>
<td>1.6E+05</td>
</tr>
<tr>
<td>(Epi)gallocatechin-4,8'/7',2-(epi)gallocatechin-4',8''-(epi)catechin</td>
<td>1.8E+06</td>
<td>2.3E+06</td>
</tr>
<tr>
<td>(Epi)gallocatechin-4,8'/7',2-(epi)gallocatechin-4',8''-(epi)catechin</td>
<td>1.9E+06</td>
<td>3.1E+05</td>
</tr>
<tr>
<td>(Epi)gallocatechin-4,8'/7',2-(epi)gallocatechin-4',8''-(epi)catechin</td>
<td>2.3E+06</td>
<td>1.5E+06</td>
</tr>
<tr>
<td>(Epi)gallocatechin-4,8'/7',2-(epi)gallocatechin-4',8''-(epi)catechin</td>
<td>1.5E+06</td>
<td>6.3E+05</td>
</tr>
<tr>
<td>(Epi)gallocatechin-4,8'/7',2-(epi)gallocatechin-4',8''-(epi)catechin</td>
<td>1.4E+06</td>
<td>2.3E+06</td>
</tr>
<tr>
<td>(Epi)gallocatechin-4,8'/7',2-(epi)gallocatechin-4',8''-(epi)catechin</td>
<td>4.2E+05</td>
<td>5.2E+05</td>
</tr>
<tr>
<td>(Epi)gallocatechin-4,8'/7',2-(epi)gallocatechin-4',8''-(epi)catechin</td>
<td>5.1E+05</td>
<td>1.2E+05</td>
</tr>
<tr>
<td>(Epi)gallocatechin-4,8'/7',2-(epi)gallocatechin-4',8''-(epi)catechin</td>
<td>1.7E+06</td>
<td>1.1E+06</td>
</tr>
<tr>
<td>(Epi)gallocatechin-4,8'/7',2-(epi)gallocatechin-4',8''-(epi)catechin</td>
<td>1.5E+06</td>
<td>4.7E+06</td>
</tr>
<tr>
<td>(Epi)gallocatechin-4,8'/7',2-(epi)gallocatechin-4',8''-(epi)catechin</td>
<td>1.9E+06</td>
<td>8.9E+05</td>
</tr>
<tr>
<td>(Epi)gallocatechin-4,8'/7',2-(epi)gallocatechin-4',8''-(epi)catechin</td>
<td>1.2E+06</td>
<td>1.6E+05</td>
</tr>
<tr>
<td>(Epi)gallocatechin-4,8'/7',2-(epi)gallocatechin-4',8''-(epi)catechin</td>
<td>2.3E+05</td>
<td>3.8E+05</td>
</tr>
<tr>
<td>(Epi)gallocatechin-4,8'/7',2-(epi)gallocatechin-4',8''-(epi)catechin</td>
<td>1.5E+06</td>
<td>2.9E+06</td>
</tr>
<tr>
<td>(Epi)gallocatechin-4,8'/7',2-(epi)gallocatechin-4',8''-(epi)catechin</td>
<td>1.7E+05</td>
<td>5.4E+05</td>
</tr>
<tr>
<td>(Epi)gallocatechin-4,8'/7',2-(epi)gallocatechin-4',8''-(epi)catechin</td>
<td>1.3E+06</td>
<td>6.8E+05</td>
</tr>
</tbody>
</table>

| (Epi)gallocatechin-4,8'/7',2-(epi)gallocatechin-4',8''-(epi)catechin | 1.7E+06 | 8.7E+04 | NA | NA |

Table I. (Continued from previous page.)
To constrain the analysis, only candidate genes ($P_{adj} < 0.001$) with at least a 2-fold expression difference between lignin-forming and non-lignin-forming conditions were considered. Out of 127 putative genes encoding enzymes in the shikimate and aromatic amino acid pathways, 26 were up-regulated in lignin-forming conditions (Fig. 7; Supplemental Table S6). The four most highly induced genes putatively encode arogenate dehydratase/prephenate dehydratase, enzymes catalyzing the final steps in Phe biosynthesis. Only four genes in the pathway from chorismate to Trp had a significant increase in expression, although it was of a small magnitude (1.3- to 1.5-fold; $P_{adj}, 0.001$). This observation suggests that H$_2$O$_2$ scavenging did not have a major effect on the pathway leading to Trp but that it strongly affected that leading to Phe (Fig. 7).

Subsequently, the effects of H$_2$O$_2$ scavenging on the phenylpropanoid and monolignol biosynthesis pathways were evaluated. In lignin-forming conditions, several genes encoding enzymes belonging to these pathways were up-regulated (Fig. 7; Supplemental Table S6). Typically, one or few members of each gene family were induced. Interestingly, several genes annotated as CAFFEATE/5-HYDROXYCONIFERALDEHYDE O-METHYLTRANSFERASE (COMT) were up-regulated, even though sinapyl alcohol is not present in spruce wood lignin or in this cell culture. Several genes with similarity to the recently discovered CAFFEOLSHIKIMATE ESTERASE (Vanholme et al., 2013) were identified, many of which had a moderate expression in the tissue-cultured cells. However, none of these were induced over 2-fold in lignin-forming conditions ($P_{adj} < 0.001$). Transcripts for many enzymes of the flavonoid pathway also were strongly induced during lignin formation (Fig. 7; Supplemental Table S6). These results support those obtained from the phenolic analysis, where elevated abundances of various flavonoids and proanthocyanidins were observed in lignin-forming cells (Fig. 3; Supplemental Fig. S2B).

Several LACCASE, PEROXIDASE, and DIRIGENT Genes Were Differentially Expressed between Lignin-Forming and Non-Lignin-Forming Conditions

Next, the expression of genes encoding oxidative enzymes (class III peroxidases and laccases) that are specifically involved in monolignol oxidation was...
examined. Thirty-nine of 281 PEROXIDASE genes (including full-length and partial sequences) were induced in lignin-forming conditions (Fig. 8A; Supplemental Table S7), whereas 18 were up-regulated in non-lignin-forming conditions. Phylogenetic sequence comparison of these 57 genes did not reveal any condition-specific sequence clade (data not shown).

LACCASE genes behaved similarly to PEROXIDASEs, with nine and eight full-length LACCASE genes induced in lignin-forming and non-lignin-forming cells, respectively (Fig. 8B; Supplemental Table S7).

In some species, such as weeping forsythia (Forsythia suspensa), dirigent proteins guide a stereoselective coupling of two coniferyl alcohol radicals to an optically active lignan, (+)-pinoresinol (Davin et al., 1997). The participation of dirigent proteins in lignin biosynthesis is controversial. In Norway spruce, there are 78 annotated putative DIRIGENT genes, of which 12 were strongly induced (up to 72-fold) in lignin-forming conditions and four in non-lignin-forming conditions (Fig. 8B; Supplemental Table S7). Two LACCASE and two DIRIGENT genes were constitutively expressed in all conditions.

Pinoresinol Formed in the Medium Was a Racemic Mixture

To ascertain whether dirigent proteins are functional in dilignol formation in the spruce cell culture, pinoresinol was isolated from the culture medium of KI-treated cultures, and its optical activity was determined. The optical rotation $[\alpha]_{22}^D$ was less than +3°, indicating that pinoresinol was a racemic mixture of both (+) and (−)-forms in almost equal amounts, as the optically active enantiomer (−)-pinoresinol shows an optical rotation $[\alpha]_{23}^D$ of −42.1° (Ishii et al., 1983).

Applastic H$_2$O$_2$ Production and Signaling

The inhibition of extracellular lignin formation under H$_2$O$_2$ scavenging suggests that peroxidases oxidize monolignols for extracellular lignin formation in the cell culture, requiring sources of apoplastic ROS, such as RBOHs, to be present. RBOHs are plasma membrane-located enzymes that produce superoxide in the apoplast (Ros Barceló, 1998; Suzuki et al., 2011; Kärkönen and Kuchitsu, 2015). Remarkably, five putative RBOH genes were significantly up-regulated in non-lignin-forming conditions but none in conditions of lignin formation (Fig. 8; Supplemental Table S7). Other sources of apoplastic H$_2$O$_2$ have been reported, including a number of cell wall-located oxidases, such as polyamine oxidases, copper-containing amine oxidases, and germins with oxalate oxidase activity (Bernier and Berna, 2001; Cona et al., 2006; Kärkönen and Kuchitsu, 2015). Plasma membrane-located quinone reductases also may contribute to superoxide production into the apoplast (Lüthje et al., 2013). In line with these findings, the expression of several genes encoding amine oxidases, germin-like proteins, and quinone reductases was strongly induced during lignin formation in the spruce cell culture (Fig. 8C; Supplemental Table S7).

Cys-rich receptor-like kinases (CRKs) have been shown to be involved in signaling during biotic and abiotic stresses that generate apoplastic ROS, and CRKs are potential candidates for apoplastic ROS perception and signaling (Bourdais et al., 2015). Among genes coding for CRKs, 47 and nine were up-regulated in lignin-forming conditions and when H$_2$O$_2$ was scavenged, respectively.
respectively (Fig. 8D; Supplemental Table S7), suggesting a potential role in $\text{H}_2\text{O}_2$ signaling.

The Expression of Antioxidant and Related Genes Suggests That Spruce Cells Experience an Oxidative Stress during Lignin-Forming Conditions

In order to evaluate the effectiveness of KI as an $\text{H}_2\text{O}_2$ scavenger, the expression of antioxidant protein-encoding genes and genes of related metabolic functions was analyzed. In conditions of $\text{H}_2\text{O}_2$ scavenging (KI_5d and KI_20d), several genes annotated as ascorbate oxidases were up-regulated (Fig. 9A; Supplemental Table S8). In contrast, in lignin-forming cultures, several genes encoding putative monodehydroascorbate reductases, catalases, and glutathione $\text{S}$-transferases were induced, suggesting that the antioxidant system was activated and that the synthesis of glutathione conjugates was increased during lignin formation (Fig. 9B; Supplemental Table S8).

A Broad Array of Transcription Factor Families Were Induced during Lignin Formation

To gain insight into the transcriptional regulation of extracellular lignin biosynthesis, the expression of putative transcription factor genes was assessed. In lignin-forming samples, the expression of 167 transcription factor/regulator genes was up-regulated by 2-fold or greater ($P_{\text{adj}} < 0.001$), while that of 138 transcription factor genes was more abundant in non-lignin-forming conditions. For gene identifiers, see Supplemental Table S6.
samples (Supplemental Table S9). MYB and WRKY transcription factor families were enriched among the differentially expressed genes (Supplemental Table S9). NAC (NAM, ATAF1/2, and CUC2) transcription factors are important for secondary cell wall development in Arabidopsis (Nakano et al., 2015; Zhong and Ye, 2015) and are likely also involved in the regulation of secondary cell wall biosynthesis in conifers (Duval et al., 2014; Raherison et al., 2015; Lamara et al., 2016). Ten NAC genes were up-regulated during lignin formation, whereas six NAC genes were down-regulated (Supplemental Table S9). Similar observations were made for the R2R3-MYB transcription factor family, members of which act as second-level master regulators and/or direct regulators of secondary cell wall biosynthesis (Duval et al., 2014; Raherison et al., 2015; Lamara et al., 2016). Ten NAC genes were up-regulated during lignin formation, whereas six NAC genes were down-regulated (Supplemental Table S9).

Gene Network Inference Analysis

A gene coexpression network was inferred, an analysis that has acknowledged limitations due to the almost binomial state of the experimental setup (i.e. lignin-forming versus non-lignin-forming conditions) but that was deemed to be more comprehensive than a more naive clustering based on expression patterns. This inference analysis helped circumvent the paucity of functionally annotated genes in the initial release of the genome by using a guilt-by-association approach to infer the role of genes of unknown function based on coexpression with genes having functional annotations. After stringent filtering (see “Materials and Methods”), the network consisted of 4,485 genes (Supplemental Table S10), with a main subcluster containing 3,472 nodes (representing genes) that, as expected, showed a strong polarity separating the lignin-forming and non-lignin-forming conditions (Fig. 10A). The network included 13 out of 196 gene models of the pathway starting from phosphoenolpyruvate and D-erythrose-4-phosphate and leading to Phe, 10 out of 111 general phenylpropanoid, and 11 out of 117 monolignol-specific pathway genes. This set of genes was used as a bait to gather all of their first and second degree neighbors, which are likely to be involved in similar biological processes. The network obtained (Supplemental Fig. S5) was further reduced of the excess-KI-affected genes (Supplemental Table S5). The resulting 399 nodes (Fig. 10, B and C) were enriched in the GO categories secondary metabolism, including phenylpropanoid metabolism, transmembrane receptor kinase signaling pathway, and protein modification (Supplemental Fig. S6; Supplemental Table S11). Of these, 203 genes also were differentially expressed in the lignin-forming versus non-lignin-forming comparison ($P_{adj} < 0.001$). This subnetwork contained many genes with a putative link to lignin biosynthesis (Table II). Additionally, six flavonoid biosynthesis genes were present. Of the 26 genes encoding transcription factors, the MYB family
was most represented with seven members (Tables II and III).

Gas Chromatography-Mass Spectrometry

The results from the phenolic profiling and transcriptome analyses suggested oxidative stress in water-treated, lignin-forming cells. Further insight was gained by analyzing primary metabolism with gas chromatography-mass spectrometry (GC-MS). The profiling of cellular extracts from 5-d-old lignin-forming and H₂O₂-scavenged cultures yielded 67 metabolites, including various organic acids, amino acids, sugars, amines, and lipids (Supplemental Table S12). Comparative analysis via PCA revealed the largest proportion of explained variation in the data set (33%) was associated with the difference between the lignin-forming and the H₂O₂-scavenged cell cultures (Supplemental Fig. S7, left plot). To reveal metabolites that contributed most to the separation between lignin-forming and non-lignin-forming cell cultures, a loading plot was constructed (Supplemental Fig. S7, right plot). Metabolites that show up on the left and on the right side of the loading plot were more abundant in lignin-forming (H₂O₂_5d) and in non-lignin-forming (KI_5d) cells, respectively. To verify whether these trends also were apparent in univariate analyses, Student’s t test was performed independently on the metabolite abundances. Metabolites in orange and red indicate those that were associated with values of $P < 0.1$ and $P < 0.05$, respectively. Among the most discriminating metabolites were putrescine, oxalic acid, ribonic acid, 2-keto-i-gluconic acid, glycerol-3-phosphate, palmitic acid, malic acid, Glu, pyro-Glu, and tartric acid, which were all higher in abundance in lignin-forming cells compared with non-lignin-forming cells, whereas discriminating compounds with a lower abundance in lignin-forming cells were represented by GalUA, 2-O-glycerol-β-D-galactoside, threonic acid, and hydroxylamine (Supplemental Fig. S7; Supplemental Table S12).

DISCUSSION

The Antioxidant System Is Induced during Lignin Formation in Cultured Spruce Cells

In this study, the effects of modulated apoplastic ROS levels on phenolic metabolism and its transcriptional regulation in a cell culture of Norway spruce that produces extracellular lignin were studied. In this cell culture, extracellular lignin becomes visible a few days after the transfer of cells from the solid maintenance medium into liquid cultures (Simola et al., 1992; Kärkönen et al., 2002). The reduction in auxin concentration at this stage was the factor proposed to stimulate lignin production in liquid conditions (Simola et al.,...
Even though the majority of cultured spruce cells (\( \sim 97\% \)) remain undifferentiated (Kärkönen and Koutaniemi, 2010), the structure of the extracellular lignin is more similar to that of native wood lignin than that of any other currently available, artificially produced lignin (Brunow et al., 1993; Koutaniemi et al., 2005; Warinowski et al., 2016). Furthermore, the same gene family members of the phenylpropanoid and monolignol biosynthesis pathways are induced in the tissue-cultured cells during lignin formation as in developing xylem (Koutaniemi et al., 2007). During culturing in liquid medium, micromolar levels of \( \text{H}_2\text{O}_2 \) are detectable in the culture medium (Kärkönen and Fry, 2006; Kärkönen et al., 2009). The substantial scavenging of \( \text{H}_2\text{O}_2 \) from the medium by KI (Huwiler et al., 1985; Nose et al., 1995; Fig. 2C) at a concentration of 5 mM prevented extracellular lignin formation (Fig. 2A; Supplemental Fig. S1). The inhibition was similar to that detected in a loblolly pine cell suspension culture where extracellular lignin formation had been induced by a high-osmoticum treatment (8% Suc; Nose et al., 1995). This observation suggests that peroxidases and not laccases are essential for the oxidative monolignol activation during lignin biosynthesis in the spruce cell culture.

Although we cannot exclude the possibility that some of the genes we identified as differentially expressed represent off-target KI effects, our analyses suggest that any such effects were minimal (Fig. 6; Supplemental Protocol S1). In support of this, KI repressed lignin formation, and KI-treated cells started to divide actively without showing any symptoms of toxicity (Fig. 2). Furthermore, the results are in accordance with those obtained in transgenic tobacco with a modified apoplastic redox state (Pignocchi et al., 2006) and are supported by the phenolic analyses presented here. The observations corroborate that the detected effects were due to \( \text{H}_2\text{O}_2 \) scavenging. The induced expression of genes encoding antioxidant and related

---

**Figure 10.** Inference analysis networks. A, Inference analysis with stringent filtering resulted in one big network and several small clusters (at right). The big network shows high polarity. Red squares indicate positive and blue circles indicate negative fold change. Fold-change values are positive and negative when the gene expression is up-regulated in lignin-forming conditions and in non-lignin-forming conditions, respectively. Yellow triangles indicate the shikimate and phenylpropanoid pathway genes (no cutoff), and gray hexagons indicate nonsignificant (\( P_{\text{adj}} > 0.05 \)) or excess-KI-affected genes. B, Subnetworks of the stringent network created by selecting nodes to the second degree with shikimate pathway genes leading to Phe and with phenylpropanoid pathway genes. Differentially expressed pathway genes are named. Symbols are the same as in A. For abbreviations of the pathway genes, see the legend for Figure 7. C, Subnetworks highlighted for the transcriptional regulators (purple diamonds; \( P_{\text{adj}} < 0.05 \)). The pathway genes (triangles) are colored for fold change: red, positive; blue, negative; yellow, nonsignificant. For gene identifiers, see Table II and Supplemental Tables S10 and S11.
Table II. Gene identifications and annotations of discussed genes of the subnetworks with $P_{adj} < 0.05$

The fold-change value is positive and negative when the gene expression is up-regulated in lignin-forming conditions and in non-lignin-forming conditions, respectively.

<table>
<thead>
<tr>
<th>Annotation</th>
<th>Gene Identifier</th>
<th>Lignin-Forming versus Non-Lignin-Forming</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Fold Change</td>
</tr>
<tr>
<td><strong>Shikimate pathway</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-Deoxy-D-arabino-heptulosonate 7-phosphate synthase</td>
<td>MA_6189912g0010</td>
<td>2.7</td>
</tr>
<tr>
<td>3-Deoxy-D-arabino-heptulosonate 7-phosphate synthase</td>
<td>MA_10142969g0010</td>
<td>1.4</td>
</tr>
<tr>
<td>3-Dehydroquininate dehydratase/shikimate dehydrogenase</td>
<td>MA_10436080g0010</td>
<td>2.3</td>
</tr>
<tr>
<td>Chorismate synthase</td>
<td>MA_2539599g0010</td>
<td>2.4</td>
</tr>
<tr>
<td>Putative prephenate aminotransferase</td>
<td>MA_17674g0010</td>
<td>1.5</td>
</tr>
<tr>
<td>Arogenate dehydratase/prephenate dehydratase</td>
<td>MA_7514745g0010</td>
<td>39.3</td>
</tr>
<tr>
<td>Arogenate dehydratase/prephenate dehydratase</td>
<td>MA_864553g0010</td>
<td>27.2</td>
</tr>
<tr>
<td>Arogenate dehydratase/prephenate dehydratase</td>
<td>MA_387748g0010</td>
<td>1.8</td>
</tr>
<tr>
<td>Arogenate dehydrogenase</td>
<td>MA_864553g0020</td>
<td>32.4</td>
</tr>
<tr>
<td>Arogenate dehydrogenase</td>
<td>MA_42232g0010</td>
<td>6.8</td>
</tr>
<tr>
<td><strong>Phenylpropanoid pathway</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phe ammonia lyase</td>
<td>MA_5862513g0010</td>
<td>4.8</td>
</tr>
<tr>
<td>Cinnamate 4-hydroxylase</td>
<td>MA_10159692g0010</td>
<td>5.0</td>
</tr>
<tr>
<td>Cinnamate 4-hydroxylase</td>
<td>MA_10435536g0010</td>
<td>61.7</td>
</tr>
<tr>
<td>4-Coumarate-CoA ligase-like</td>
<td>MA_166573g0010</td>
<td>2.9</td>
</tr>
<tr>
<td>Cinnamoyl-CoA reductase</td>
<td>MA_137109g0010</td>
<td>2.7</td>
</tr>
<tr>
<td>Caffeate/5-hydroxyconiferaldehyde O-methyltransferase</td>
<td>MA_10430251g0010</td>
<td>−4.1</td>
</tr>
<tr>
<td>Caffeate/5-hydroxyconiferaldehyde O-methyltransferase</td>
<td>MA_10375504g0010</td>
<td>4.7</td>
</tr>
<tr>
<td>Caffeate/5-hydroxyconiferaldehyde O-methyltransferase</td>
<td>MA_630428g0010</td>
<td>−5.1</td>
</tr>
<tr>
<td>Caffeate/5-hydroxyconiferaldehyde O-methyltransferase</td>
<td>MA_10426580g0010</td>
<td>17.0</td>
</tr>
<tr>
<td>Caffeate/5-hydroxyconiferaldehyde O-methyltransferase</td>
<td>MA_86438g0010</td>
<td>9.1</td>
</tr>
<tr>
<td><strong>Flavonoid pathway</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flavonoid 3’-hydroxylase</td>
<td>MA_10434848g0010</td>
<td>12.8</td>
</tr>
<tr>
<td>Flavonoid 3’-hydroxylase</td>
<td>MA_158072g0010</td>
<td>4.0</td>
</tr>
<tr>
<td>Dihydroflavonol-4-reductase</td>
<td>MA_110462g0010</td>
<td>3.2</td>
</tr>
<tr>
<td>Leucoanthocyanidin dioxygenase</td>
<td>MA_169203g0010</td>
<td>33.5</td>
</tr>
<tr>
<td>Anthocyanidin 3-O-glucosyltransferase</td>
<td>MA_156209g0010</td>
<td>2.5</td>
</tr>
<tr>
<td><strong>Transcription factors</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abscisic stress-ripening protein, ASR</td>
<td>MA_60383g0010</td>
<td>2.4</td>
</tr>
<tr>
<td>Abscisic stress-ripening protein, ASRa</td>
<td>MA_773806g0010</td>
<td>3.8</td>
</tr>
<tr>
<td>Abscisic stress-ripening protein, ASRa</td>
<td>MA_89246g0010</td>
<td>7.7</td>
</tr>
<tr>
<td>bHLH transcription factor</td>
<td>MA_89575g0010</td>
<td>2.6</td>
</tr>
<tr>
<td>C2H2 zinc finger transcription factor</td>
<td>MA_756968g0010</td>
<td>8.2</td>
</tr>
<tr>
<td>Ethylene-responsive transcription factor</td>
<td>MA_4072g0020</td>
<td>−2.7</td>
</tr>
<tr>
<td>Ethylene-responsive transcription factor</td>
<td>MA_98464g0010</td>
<td>8.2</td>
</tr>
<tr>
<td>GRAS transcriptional regulator</td>
<td>MA_99004g0010</td>
<td>−2.1</td>
</tr>
<tr>
<td>MYB family transcription factor</td>
<td>MA_139448g0010</td>
<td>20.3</td>
</tr>
<tr>
<td>MYB family transcription factor</td>
<td>MA_937875g0010</td>
<td>9.3</td>
</tr>
<tr>
<td>MYB family transcription factor</td>
<td>MA_10225049g0010</td>
<td>8.3</td>
</tr>
<tr>
<td>MYB family transcription factor</td>
<td>MA_51173g0010</td>
<td>2.5</td>
</tr>
<tr>
<td>MYB family transcription factor</td>
<td>MA_37058g0010</td>
<td>2.7</td>
</tr>
<tr>
<td>MYB family transcription factor</td>
<td>MA_316475g0010</td>
<td>34.1</td>
</tr>
<tr>
<td>NAC transcription factor</td>
<td>MA_139896g0010</td>
<td>1.9</td>
</tr>
<tr>
<td>NAC transcription factor</td>
<td>MA_18958g0010</td>
<td>6.5</td>
</tr>
<tr>
<td>NAC transcription factor</td>
<td>MA_16595g0010</td>
<td>1.9</td>
</tr>
<tr>
<td>Probable WRKY transcription factor</td>
<td>MA_76002g0010</td>
<td>17.1</td>
</tr>
<tr>
<td>TIFY family transcription factor</td>
<td>MA_10427283g0030</td>
<td>11.7</td>
</tr>
<tr>
<td>TIFY family transcription factor</td>
<td>MA_10436587g0020</td>
<td>4.7</td>
</tr>
<tr>
<td>Two-component response regulator ARR</td>
<td>MA_8059983g0010</td>
<td>54.9</td>
</tr>
<tr>
<td><strong>Other</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ABC transporter G family member</td>
<td>MA_18507g0010</td>
<td>−3.2</td>
</tr>
<tr>
<td>Cys-rich receptor-like protein kinase</td>
<td>MA_87408g0010</td>
<td>−2.1</td>
</tr>
<tr>
<td>Cys-rich receptor-like protein kinase</td>
<td>MA_911707g0010</td>
<td>3.9</td>
</tr>
<tr>
<td>Gal oxidase</td>
<td>MA_812690g0010</td>
<td>5.6</td>
</tr>
<tr>
<td>Dirigent protein</td>
<td>MA_10428143g0020</td>
<td>19.5</td>
</tr>
<tr>
<td>Dirigent protein</td>
<td>MA_638586g0010</td>
<td>10.5</td>
</tr>
</tbody>
</table>

(Table continues on following page.)
proteins indicated that the cells were subjected to oxidative stress during extracellular lignin formation and that this oxidative condition was alleviated by KI treatment (Fig. 9). Likewise, various flavonoids and tannins accumulated preferentially in lignin-forming cells (e.g. clique clusters 1, 10, 39, and 52; Supplemental Fig. S2B), with some, such as taxifolin, quercetin, and myricetin, being efficient radical scavengers and/or antioxidants for lipid peroxidation, albeit sugar moieties can influence their antioxidant activity (Hopia and Heinonen, 1999; Willför et al., 2003). The strong up-regulation of genes encoding glutathione S-transferases (Fig. 9B) and UDP-Glc/GlcA transferases in lignin-forming conditions suggests that the synthesis of glutathione and sugar conjugates was increased during lignin formation, further supporting that cells under lignin-forming conditions were subjected to oxidative stress. The up-regulation of similar transferase genes together with those of cytochrome P450s has been shown to be typical for a catalase-deficient Arabidopsis mutant that suffers from oxidative stress (Queval et al., 2012; Noctor et al., 2015). Glutathione S-transferases, for example, catalyze the conjugation of oxidized lipid compounds to glutathione (Farmer and Mueller, 2013; Noctor et al., 2015). In conditions of H2O2 scavenging, several genes encoding putative ascorbate oxidases were induced (Fig. 9A). In the less oxidative conditions under KI treatment, ascorbate would preferentially exist in a reduced form. Ascorbate oxidases regulate the redox state of apoplastic ascorbate by oxidizing ascorbate to monodehydroascorbate without generating ROS (Pignocchi and Foyer, 2003). However, not all proteins encoded by genes annotated as ascorbate oxidase have ascorbate oxidase activity (Ueda et al., 2015), and further studies are needed to resolve the function of the proteins encoded by these genes in spruce. GC-MS analysis additionally indicated altered ascorbate metabolism, with tartaric acid (a precursor for ascorbate biosynthesis) as well as oxalic acid and threonic acid (ascorbate breakdown products) belonging to the most discriminating metabolites between lignin-forming and non-lignin-forming cell cultures (Supplemental Fig. S7; Supplemental Table S12).

Various Low-Molecular-Weight Phenolic Compounds Were Present in Cells and in the Culture Medium in Both Lignin-Forming and Non-Lignin-Forming Conditions

The most apparent change in the cultures where H2O2 was scavenged was the abolition of extracellular lignin formation (Fig. 2A; Supplemental Fig. S1). In the lignin-forming and H2O2-scavenging cultures, a large

Table II. (Continued from previous page.)

<table>
<thead>
<tr>
<th>Annotation</th>
<th>Gene Identifier</th>
<th>Fold Change</th>
<th>P adj</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dirigent protein</td>
<td>MA_10428143g0010</td>
<td>8.5</td>
<td>1.4E-05</td>
</tr>
<tr>
<td>Dirigent protein</td>
<td>MA_362458g0010</td>
<td>30.9</td>
<td>6.1E-21</td>
</tr>
<tr>
<td>Laccase</td>
<td>MA_414479g0010</td>
<td>11.2</td>
<td>3.2E-12</td>
</tr>
<tr>
<td>Laccase</td>
<td>MA_67291g0010</td>
<td>10.4</td>
<td>1.1E-20</td>
</tr>
<tr>
<td>Peroxidase</td>
<td>MA_10432865g0020</td>
<td>8.1</td>
<td>2.2E-23</td>
</tr>
<tr>
<td>Peroxidase</td>
<td>MA_10433564g0010</td>
<td>30.8</td>
<td>4.2E-64</td>
</tr>
<tr>
<td>Peroxidase</td>
<td>MA_179122g0010</td>
<td>23.5</td>
<td>4.3E-17</td>
</tr>
<tr>
<td>Pinoresinol reductase homolog</td>
<td>MA_445017g0010</td>
<td>18.5</td>
<td>6.2E-16</td>
</tr>
<tr>
<td>Pinoresinol reductase homolog</td>
<td>MA_222809g0010</td>
<td>10.9</td>
<td>3.4E-09</td>
</tr>
</tbody>
</table>

*aSwitch or regulation position in the subnetwork.

Table III. Transcription factor families found in the inference subnetworks, and occurrence in cell wall studies

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>MYB (7)</td>
<td>Yes</td>
<td>–</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>ASR (5)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>NAC (3)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>AP2/EREBP (2)</td>
<td>–</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>TIFY (2)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>bZIP (1)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>C2H2 zinc finger (1)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>GRAS (1)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>ARR (1)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>WRKY (1)</td>
<td>Yes</td>
<td>–</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Downloaded from on October 15, 2017 - Published by www.plantphysiol.org
Copyright © 2017 American Society of Plant Biologists. All rights reserved.
Apoplastic Redox State Affects Phenolic Metabolism

variety of low-molecular-weight phenolic compounds accumulated within the cells and in the culture medium. Several glycosylated dilignols were present in cells during H$_2$O$_2$ scavenging (Supplemental Fig. S2B; clique cluster 69), with further accumulation after KI removal (lignin reformation; clique clusters 10 and 69). In addition, several types of flavonoids, phenolic alcohols, and hydroxycinnamic acids were identified intracellularly. Many of these compounds also were detected in cells of the lignin-forming cultures, but with contrasting accumulation patterns. The observation that spruce cells contain various dilignol glycosides is in accordance with the data of Dima et al. (2015), who detected glycosylated dilignols, triligons, and tetriligons in vacuoles of Arabidopsis leaves. It has been demonstrated that vacuole-located oligolignols are generated by protoplasmic combinatorial radical coupling of monolignols similar to that observed in cell walls, likely due to photooxidative stress (Dima et al., 2015). Also in lignifying tissues, intracellular oxidative cross-coupling of monolignols was suggested to occur based on the detection of cysteinylated 8-O-4-type dilignols in the xylem of transgenic poplar (Populus spp.) with a diminished phenylcoumaran benzylc ether reductase (PCBER) activity (Niculaes et al., 2014). The latter aryl ether dilignols are formed when Cys rather than water acts as a nucleophile in the rearrangement of quinone methides formed by the 8-O-4 cross-coupling of monolignol radicals. As Cys is not present in the apoplast, this observation implied that these quinone methides were formed inside the cells. Intracellular radical-radical cross-coupling of monolignols is further supported in this study, as the time-course profiles of pinoresinol hexoside inside the cells and in the culture medium were clearly different (Supplemental Fig. S2). In contrast to Arabidopsis leaf cells, in which the radical-radical cross-coupling-based scavenging mechanism yields only oligolignol glycosides (Dima et al., 2015), the condensed tannins within the spruce cells also are supposedly formed by such a mechanism. Indeed, condensed tannin polymerization has been suggested to proceed via the intermediate formation of quinone methides with the involvement of polyphenol oxidases and/or peroxidases (Pourcel et al., 2007). In agreement with such a mechanism, an increased abundance of condensed tannins was observed in the lignin-forming and lignin-reforming cultures (Supplemental Fig. S2B; clique clusters 1 and 10).

The presence of pinoresinol hexoside in the culture medium (Table I) suggests that glycosylated dimers are transported to the apoplast. Indeed, UDP-glycosyl transferases and activated sugars required for the glycosylation reaction are considered to locate intracellularly (Le Roy et al., 2016). An alternative hypothesis for the presence of pinoresinol hexoside in the culture medium is that some of the cells had ruptured and released their cytoplasmic content. This hypothesis is not supported by the data, as the accumulation profiles of pinoresinol hexoside inside and outside the protoplast were clearly different (Supplemental Fig. S2). The fact that flavonoids, some of which were abundant in cells, were not detected in the medium also indicates that the medium was not contaminated with cellular content.

In the spruce cell culture, the expression of several putative dirigent-encoding genes was induced significantly in lignin-forming cells and some in non-lignin-forming conditions (Fig. 8B). In order to ascertain whether dirigent proteins contributed to dilignol synthesis, we isolated pinoresinol from the culture medium of KI-treated cultures. The isolated pinoresinol was a racemic mixture of both (+)- and (−)-forms, suggesting that dirigent proteins did not participate in pinoresinol synthesis when H$_2$O$_2$ was scavenged. All putative DIRIGENT genes of spruce (at least all those for which a full-length sequence could be obtained) encode an N-terminal signal peptide, suggesting that they enter the secretory pathway. If a dirigent protein was responsible for pinoresinol synthesis, its concentration in the culture medium may have been too low for the protein to catch the conifer alcohol radicals. Hence, purification of pinoresinol from the apoplastic fluid of developing xylem is needed to evaluate the role of dirigent proteins in dilignol formation in the cell wall.

Usually, conditions leading to apoplastic ROS formation (e.g. wounding and pathogen infection) induce DIRIGENT expression (Ralph et al., 2006). Their exact role in defense has not yet been elucidated.

Reduced structures [IDDDC hexoside and G(8-O-4) IDDDC] were identified in the culture medium in both conditions (clique clusters 1 and 74; Supplemental Fig. S2A). IDDDC is formed from dehydroconiferyl alcohol [G(8-5)G] by a PCBER that, in poplar, is one of the most abundant proteins present in xylem (Niculaes et al., 2014). Since the PCBER is located in the cytoplasm (Gang et al., 1999; Vander Mijnsbrugge et al., 2000a, 2000b; Niculaes et al., 2014), the data suggest that these reduced dilignols are produced intracellularly and transported into the apoplast. The spruce genome contains 10 sequences of high similarity to Populus trichocarpa PfrPCBER, several of which have good expression in developing xylem (Nystedt et al., 2013). In the spruce cell culture, four putative PCBER-encoding genes had a considerable, constant expression across all conditions, whereas one PCBER gene was induced in lignin-forming and one in non-lignin-forming conditions (Fig. 8B).

Heterologous production of the proteins and biochemical enzyme activity assays are needed to investigate which of these genes encodes the enzyme contributing to the biosynthesis of IDDDC. The observation that IDDDC hexoside was exported whereas many other hexosylated phenolics remained intracellular suggests an as yet unknown biological role for IDDDC hexoside in the apoplast.

Lariciresinol hexoside, another reduced dilignol glycoconjugate, was detected with a good response by mass spectrometry in the extracts of spruce cells in both lignin-forming and H$_2$O$_2$-scavenging conditions, but unlike IDDDC hexoside, it was not detected in the medium (Table I; Supplemental Table S2). This compound is formed from pinoresinol via the action of...
pinoresinol reductase (Fujita et al., 1999). Two genes encoding putative pinoresinol reductase were induced in lignin-forming conditions, and one gene had a constant expression in both conditions (Fig. 8B). This enzyme catalyzes the reduction of pinoresinol to lariciresinol and, in some species, also the further reduction of lariciresinol to secoisolariciresinol (Fujita et al., 1999). Native Norway spruce lignin contains pinoresinol and secoisolariciresinol structures (Zhang et al., 2003). Secoisolariciresinol units can originate via dehydrogenative coupling of secoisolariciresinol into lignin polymer (Zhang et al., 2003). In this work, lariciresinol was not detected in the culture medium of lignin-forming cells, neither as such nor as part of an oligolignol.

The Apoplastic Redox State Influences Cellular Metabolism

The RNA-Seq data analysis revealed that the phenylpropanoid pathway was up-regulated in lignin-forming conditions compared with non-lignin-forming conditions (Fig. 7; Supplemental Table S6), which is consistent with the accumulation of extracellular lignin in water-treated and in lignin-reforming cultures after KI removal. The data support that a mechanism exists, possibly mediated by apoplastic H₂O₂, sensing a reduced need for monolignol biosynthesis under H₂O₂-scavenging conditions when monolignols cannot polymerize, whereas under more oxidative conditions, the synthesis of monolignols is up-regulated. Alternatively, the signal can be a phenolic compound (Bonawitz et al., 2014). The differential expression analysis revealed several CRK-encoding genes that were up-regulated in lignin-forming conditions (Fig. 8D). CRK expression has been shown to be modulated in response to ROS- and cellular redox state-related stresses (Lehti-Shiu et al., 2009; Wrzaczek et al., 2010; Bourdais et al., 2015). Moreover, CRKs are thought to be involved in cellular redox and apoplastic ROS sensing by the presence of conserved Cys residues in their apoplastic domain. Thus, CRKs are potential candidates for ROS signaling and regulation (Idänheimo et al., 2014; Bourdais et al., 2015).

The cellular redox state has been shown to affect cell proliferation in animals (Diaz-Vivancos et al., 2015), and this is likely also to be the case in plants (Schippers et al., 2016). Differences in redox buffering (i.e. antioxidative capacity) in cellular compartments, with the apoplast being more oxidizing compared with the highly redox-buffered cytosol, suggest that plant cells can utilize redox-sensitive signal transduction as a powerful mechanism to integrate apoplastic stimuli to adjust cellular metabolism (Foyer and Noctor, 2016). During plant defense responses, transient oxidative bursts have been shown to induce major changes in plant metabolism (Wrzaczek et al., 2013). Interestingly, in spruce cells, the modulation of the longer term redox status of the apoplast had a profound influence, not only on apoplastic metabolism but on whole cellular metabolism. Our results are in accordance with those obtained with transgenic tobacco with modified ascorbate oxidase levels in the apoplast: the oxidation state of the apoplast had an effect on gene expression and hormone signaling, affecting growth and susceptibility to pathogens (Pignocchi et al., 2006). It is tempting to speculate that H₂O₂ present at micromolar concentrations in the culture medium of lignin-forming cultures (Kärkönen and Fry, 2006; Kärkönen et al., 2009) is perceived by plasma membrane-located receptors (e.g. CRKs); the consequent signal transduction would lead to the up-regulation of pathways leading to extracellular lignin and flavonoid biosynthesis. Under H₂O₂-scavenging conditions (i.e. higher reducing condition in the apoplast), on the contrary, different metabolic pathways were operating, with cells dividing actively. In cell cultures of zinnia that differentiate into TEs, the lignification of TEs initiates only after TEs undergo programmed cell death, even if the monolignols are already present (Pesquet et al., 2013). The authors hypothesized that ROS released from cells by programmed cell death could have a role in triggering lignification.

During lignin formation, the most highly induced genes of the shikimate and aromatic amino acid pathways presumably encode arogenate dehydratase/prephenate dehydratase, enzymes catalyzing the final steps in Phe biosynthesis (Fig. 7; Supplemental Table S6). Certain isoenzymes of arogenate dehydratase have been shown to profoundly and differentially modulate carbon flux into the downstream lignin biosynthesis pathway in Arabidopsis (Corea et al., 2012a, 2012b). It seems likely that the induced genes in spruce encode isoenzymes responsible for the synthesis of Phe destined for monolignol and/or flavonoid biosynthesis, while some other isoenzymes are responsible for the formation of Phe utilized for protein synthesis.

Both Peroxidases and Laccases Are Likely to Have a Role in Phenolic Coupling in the Spruce Cell Culture

Both PEROXIDASEs and LACCASEs showed differential expression in lignin-forming and non-lignin-forming conditions: many genes were either up- or down-regulated (Fig. 8, A and B). In addition, the coexpression subnetwork contained three PEROXIDASE and two LACCASE genes that were expressed similarly to monolignol and shikimate pathway genes (Table II; Supplemental Table S11). All these genes were up-regulated in lignin-forming conditions (Supplemental Table S7), and all but one have some expression in early wood or in vegetative shoots in spruce (Nystedt et al., 2013), making them strong candidates for further focus. One PEROXIDASE in the subnetwork (MA_10432865g0020; Table II) encodes PaPX2, a cationic peroxidase shown to be expressed in differentiating tracheids of developing xylem (Marjamaa et al., 2006). One subnetwork LACCASE (MA_67291g0010), on the other hand, is PaLAC3c (Koutaniemi et al., 2015) and has considerable expression in early wood and stems of vegetative shoots (absolute expression values of 6.98 and 9.68, respectively; Nystedt et al., 2013).
An analogous situation to that in the spruce cell culture has been observed in flax (Linum usitatissimum) plants, whose outer stem tissues contain hypolignified bast fibers with increased levels of glycosylated oligolignols compared with the inner stem, which contains lignified xylem and a higher content of less polar oligolignols (Huis et al., 2012). The data obtained from flax mutants with lignified bast fibers suggest that, similar to that in the KI-treated spruce cell cultures (Fig. 2A), insufficient polymerization of monolignols leads to hypolignification in the wild-type bast fibers (Chantreau et al., 2014). In the flax lignified bast fiber1 mutants, which have ectopic lignin formation in bast fibers, PEROXIDASE expression is increased in comparison with the wild type, suggesting that wild-type flax stems have insufficient peroxidase action, leading to the hypolignification of bast fibers. According to the current opinion in the field, laccases function in the initial coupling of monolignols to oligolignols, while peroxidases contribute to lignin formation at later stages (Sterjádés et al., 1993; Zhao et al., 2013). Additionally, some isoenzymes of them probably participate in the polymerization of proanthocyanidins (condensed tannins) detected in cells (Fig. 3; Supplemental Fig. S2B; P ourcel et al., 2005, 2007). As peroxidase action is essential in the formation of high-molecular-weight lignin polymer in this cell culture and polymers were not formed in the presence of KI, the involvement of another type of oxidative enzyme is likely. This observation supports the role of laccases in the synthesis of dillignols and oligolignols (Fig. 3). However, H2O2 scavenging by KI seemed to be incomplete (Fig. 2C), allowing for some peroxidase action during the KI treatment. Thus, the possibility exists that there was less peroxidase action in the KI-supplemented cultures due to reduced H2O2 concentration; this would lead to shorter polymers. Hence, the data from this work cannot exclude that only peroxidases participate in monolignol oxidation in the spruce cell culture.

Apoplastic ROS Generation

Since peroxidases are involved in monolignol activation for lignin biosynthesis in the spruce cell culture, an aim of this study was to identify the source(s) of apoplastic H2O2. Candidates include isoenzymes of apoplastic peroxidases that can form H2O2 if there is a reductant present (Daudi et al., 2012); however, the native apoplastic reductant is unknown. Other candidates include plasma membrane-located RBOH enzymes (Ros Barceló, 1998; Suzuki et al., 2011; Kärkönen and Kuchitsu, 2015). RBOHs generate apoplastic superoxide that dismutates to H2O2 either nonenzymatically in the acidic pH present in the cell wall or enzymatically via a superoxide dismutase (SOD)-catalyzed reaction (Karlsson et al., 2005). Surprisingly, none of the putative RBOH genes were induced in lignin-forming conditions, but the expression of several RBOH genes was high in all non-lignin-forming conditions (also at time 0; Fig. 8C). Although these data do not indicate a role for RBOHs, it has to be remembered that the activity of RBOH enzymes is efficiently regulated posttranscriptionally by calcium and phosphorylation (Kärkönen and Kuchitsu, 2015).

Apoplastic isoenzymes of polyamine oxidases and copper-containing amine oxidases produce H2O2 in the cell wall using apoplastic diamines and polyamines as substrates (Cona et al., 2006). Oxalate oxidases (also called germins), on the other hand, use oxalic acid/oxalate as a substrate; alternatively, some have SOD activity (Bernier and Berna, 2001). In our experiment, several genes annotated as COPPER-CONTAINING AMINE OXIDASE or GERMIN were strongly up-regulated in lignin-forming conditions (Fig. 8C). However, externally added oxalate, putrescine, spermidine, or Arg (the precursor of the latter compounds) did not increase apoplastic H2O2 levels in the lignin-forming spruce cell culture, suggesting that these enzymes do not produce H2O2 into the culture medium (Kärkönen et al., 2009). Interestingly, higher amounts of putrescine and oxalate were detected in lignin-forming cells compared with non-lignin-forming cells (Supplemental Fig. S7; Supplemental Table S12). Levels of these metabolites in the culture medium, a site of apoplastic H2O2 formation, were not determined. Other possible apoplastic ROS-producing enzymes include plasma membrane-located quinone oxidoreductases (Schofer et al., 2008; Lüthje et al., 2013; Kärkönen et al., 2014). In the cell culture, two quinone oxidoreductase-encoding genes (FLAVOPROTEIN WrbA) were induced during lignin formation. Both of these genes also have strong expression (absolute expression values greater than 10.5) in developing xylem (Nystedt et al., 2013).

The inference analysis revealed novel enzyme candidates, Gal oxidases, for apoplastic H2O2 production (Table II; Supplemental Table S11). Furthermore, differential expression analysis revealed additional GAL OXIDASE genes that were induced in lignin-forming conditions or in non-lignin-forming conditions (Padj < 0.001; Fig. 8C). There have been reports suggesting the involvement of carbohydrate oxidases in defense-related H2O2 production in plants (Custers et al., 2004). The detection of GalUA as a main discriminating metabolite between the two types of cell cultures with higher abundance in non-lignin-forming cells (Supplemental Fig. S7; Supplemental Table S12) may support such a hypothesis, as galacturonaldheyde, the Gal oxidase product, should be readily oxidized to GalUA during GC-MS analysis. Alternatively, GalUA may have been released from pectins during cell wall modifications, followed by uptake into the cells for reactivation for cell wall biosynthesis. Clearly, more work is needed to elucidate the role of ROS-producing enzymes in plant development and, specifically, in lignin biosynthesis.

Transcriptional Regulation

Similar to that in Arabidopsis, in white spruce, a number of NAC domain transcription factors regulate...
the expression of secondary transcription factors that, in turn, regulate the expression of lignin and cell wall polysaccharide biosynthesis genes; other NACs directly regulate some secondary cell wall biosynthesis genes (Duval et al., 2014; Raherison et al., 2015; Lamara et al., 2016). This suggests conservation of the hierarchical transcriptional regulatory cascade network of secondary cell wall synthesis between angiosperms and gymnosperms. PgNAC7 is most likely a master switch regulating secondary cell wall formation in white spruce (Duval et al., 2014). It is also the most connected hub gene in the coexpression gene network of secondary cell wall-related genes (Lamara et al., 2016). PgNAC7 and its homolog in Norway spruce (MA_6777g0010) are highly expressed during wood formation in stems (Nystedt et al., 2013; Raherison et al., 2015). In the cultured spruce cells, the expression of MA_6777g0010 was highest in H2O2-scavenging conditions and at time 0 (cells on the solid maintenance medium), with expression down-regulated in conditions of lignin formation. In addition, genes annotated as MYB transcription factors (MA_139238g0010 and MA_62361g0010), which are putative orthologs of those regulating secondary cell wall and/or monolignol biosynthesis, PgMYB8/PtMYB8 and PgMYB1/PtMYB1 (Bedon et al., 2007; Bomal et al., 2008), respectively, had the highest expression at time 0, with a decrease in KI-supplemented conditions and an even lower expression in lignin-forming conditions. The majority of lignin-forming spruce cells remain undifferentiated. Therefore, up-regulation of the genes coding for transcription factors that regulate the entire process of secondary cell wall biosynthesis would not be expected when extracellular lignin is synthesized. However, some callus cells (~3%) differentiate to tracheids on the solid maintenance medium (Kärkönen and Koutaniemi, 2010), an observation that may explain the moderate levels of expression of secondary cell wall-regulating transcription factors at time 0. Interestingly, a recent transcriptomic study with coexpression network analysis over cambium and developing xylem up to mature xylem suggested that NAC transcription factors are not central regulators for secondary cell wall formation in Norway spruce (Jokipii-Lukkari et al., 2017).

In Arabidopsis, AtMYB58 and AtMYB63 directly activate lignin biosynthesis genes by binding to AC elements present in the promoters of most of the genes in the phenylpropanoid and monolignol pathways (Zhou et al., 2009). A few putative spruce MYB genes (MA_1043020g0010, MA_1201g0010, and MA_137934g0010) with similarity to these Arabidopsis MYB genes showed significant up-regulation during lignin formation. Furthermore, MYB family transcription factors were enriched within the differentially expressed genes (Supplemental Table S9), and MYB was the most abundant transcription factor family in the inference analysis subnetwork (Table III; Supplemental Table S11). A putative MYB gene (MA_139448g0010) was located in the subnetwork (Fig. 10C; Table II). This MYB was a first-degree neighbor of a putative 3-DEOXY-ß-ARABINO-HEPTULOSONATE 7-PHOSPHATE SYNTHASE gene, encoding the first enzyme in the shikimate pathway. Furthermore, the MYB gene was induced 20-fold in lignin-forming conditions (P adj < 0.001). These observations make this MYB an interesting candidate regulator of events leading to phenylpropanoid biosynthesis. Apart from monolignol biosynthesis, MYB transcription factors are important in regulating, for example, flavonoid biosynthesis, with different members important in distinct processes (Bomal et al., 2014).

In recent years, secondary cell wall and lignin biosynthesis studies have revealed new members of the transcriptional regulatory gene regulatory network (Cassan-Wang et al., 2013; Taylor-Teeples et al., 2015; Shi et al., 2017). Using a yeast one-hybrid approach on Arabidopsis root xylem-expressed transcription factors, Taylor-Teeples and colleagues (2015) revealed a highly interconnected gene regulatory network enabling subtle and highly focused regulation of cell wall biosynthesis. AP2/EREBP, bHLH, C2H2 zinc finger, and GRAS transcription factor families showed enrichment in their studies. In addition, transcription factors from bZIP and WRKY families coexpressed with cell wall biosynthesis genes in xylem of Populus trichocarpa (Shi et al., 2017). Previously, genes of the same transcription factor families were identified in a search for regulators of secondary cell wall biosynthesis in Arabidopsis (Cassan-Wang et al., 2013). Several genes belonging to these same transcription factor families had similar expression patterns to the shikimate and monolignol pathway genes (Table III) and also were differentially expressed (P adj < 0.001). In the present study, enrichment analysis (Taylor-Teeples et al., 2015) showed statistically significant overrepresentation of the members of AUX-IAA, homeobox, MYB, and WRKY transcription factor families within the differentially expressed genes (hypergeometric probability; Supplemental Table S9). Members of the MYB, Tify, and ABSCISIC STRESS-RIPENING PROTEIN (ASR) transcription factor families also were present in the subnetwork (Tables II and III; Supplemental Table S11), some of which have expression in developing xylem (Nystedt et al., 2013). A putative ASR1 gene was found in a switch position in the subnetwork (Fig. 10C; Table II). In tomato (Solanum lycopersicum), ASR1 has been found to act as a transcription factor also targeting cell wall-related genes (Ricardi et al., 2014). In addition, rice (Oryza sativa) ASR1 may directly scavenge H2O2 (Kim et al., 2012). Also, two putative TIFY genes were located in possible switch positions in the subnetwork (Fig. 10C; Table II). The TIFY transcription factor family has been connected to a variety of abiotic and biotic stresses and also to development (Zhang et al., 2015). In maize (Zea mays), ZML2, a Tify family member, and MYB11 together repress COMT expression (Vélez-Bermúdez et al., 2015).

The differential gene expression analysis identified AP2/EREBP transcription factors as the second most abundant group of transcription factors up-regulated during lignin formation (Supplemental Table S9). The
ethylene response transcription factors (ERFs) mediate the responses to ethylene, but some are involved in other hormonal (e.g. jasmonic acid and abscisic acid) signaling pathways (Vahala et al., 2013). In hybrid aspen (Populus tremula × Populus tremuloides), over-expression of certain ERFs affects cambial growth and wood chemistry (lignin and carbohydrate content; Vahala et al., 2013). During compression wood formation in maritime pine (Pinus pinaster), the ethylene-forming enzyme 1-aminocyclopropane-1-carboxylate oxidase accumulates strongly (Plomion et al., 2000). In addition to eccentric growth, compression wood of conifers contains higher levels of lignin compared with normal vertical wood. In Norway spruce and radiata pine (Pinus radiata), ethylene stimulates cambium activity (Barker, 1979; Eklund and Tiltu, 1999). However, no changes in lignin amount were detected in radiata pine wood formed under ethylene exposure, although the amount of water-soluble extractives increased (Barker, 1979).

CONCLUSION

The lignin-forming cell culture of spruce proved to be a valuable system for lignin biosynthesis studies, as it allows modification of the apoplastic redox state, an objective difficult to investigate in the developing wood of intact trees. The global analyses presented here revealed a major role for the apoplastic redox state in cellular metabolism. Cells with oxidative conditions in the apoplast supply large amounts of carbon into the synthesis of various phenylpropanoids (dilignols and oligolignols, lignin, and flavonoids), some of which have antioxidant properties. Under more reducing conditions in the apoplast, other metabolic pathways operate and cells proliferate actively. The isolation of apoplastic compounds separately from the cytoplasmic ones revealed that spruce cells are able to transport glycosylated dilignols into the apoplast. CSPP networks enabled the identification of several phenolic compounds based on their biochemical modifications. Gene inference network analysis combined with differential expression revealed candidate enzymes for monolignol oxidation and apoplastic H₂O₂ production, novel putative transcription factors and potential H₂O₂ receptors related to lignin biosynthesis. The results obtained are important for advancing understanding of the processes involved in lignin biosynthesis and have provided potential targets for future studies on the utilization of lignocellulosic materials.

MATERIALS AND METHODS

Plant Material

Norway spruce (Picea abies) cells (line A3/85) were maintained on a solid nutrient medium as described by Karkkonen et al. (2002). For experiments, cells (approximately 2.5 weeks after subculturing) were transferred into liquid medium 5’ (Simola and Santanen, 1990), and freshly prepared KI (final concentration, 5 mM), KCl (5 mM), or an equivalent volume of water was added to the medium. At the transfer from solid to liquid medium, the auxin (2,4-dichlorophenoxyacetic acid) concentration is reduced from 10 to 0.5 μM, whereas that of cytokinin (kinetin) remains unchanged (2.5 μM). In addition, some changes in macronutrients occur (e.g. total nitrogen concentration is reduced from 35 to 30 mM; Simola and Santanen, 1990). To ensure a uniform cell population, a large-volume cell suspension was initiated and divided into 100-mL aliquots (approximately 2 g of cells per 100 mL of medium in a 500-mL flask).

Time 0 samples were taken from the solid maintenance medium immediately before the cells were transferred into liquid cultures. Cell suspensions were cultivated at a shaker (100 rpm) at 20°C in an 18-h-light/6-h-dark cycle (30-50 μmol m⁻² s⁻¹; Osram warm white) and sampled at the indicated time points (Supplemental Fig. S1). After 20 d of culturing, the culture medium of KI-treated cultures was decanted away. To remove any residual KI from the apoplast of cells, 100 mL of fresh nutrient medium containing the normal amount of KI (20 μM as a micronutrient) was added onto the cells, and cells were incubated for 15 to 20 min on a shaker, after which the medium was decanted again. This was repeated three times, after which the culturing was continued as described above.

Oligolignol Profiling

Sample Preparation

To reduce variation between replicates, 100-mL suspension cultures in 500-mL flasks were used with four biological replicates; these were sampled throughout the time series (Supplemental Fig. S1A). The cultures with or without KI supplementation were initiated as described above. At each time point, 2-mL samples containing both cells and the culture medium were aseptically collected. Once the cells had settled down, culture medium (~1.5 mL) was transferred into a clean tube and centrifuged to remove any precipitates (5 min, 17,000g, and 4°C). Supernatant was transferred into a new tube and the air space was filled with nitrogen gas, after which the samples were frozen in liquid nitrogen. From the cells, the remaining medium was pipetted away, and the cells were washed once with 1.8 mL of water. As soon as the cells had settled again, the water was pipetted out, the air space was filled with nitrogen gas, and the cells were frozen in liquid nitrogen. Usually, it is impossible to cultivate cells of this spruce cell culture line in liquid cultures for longer periods, as the cells die a few days after extracellular lignin becomes visible in the culture medium. In this specific tissue culture experiment, water-treated cells remained alive even when they produced extracellular lignin; hence, the lignin-forming cultures were cultured with sampling up to 20 d. The samples were stored at −80°C before the analyses. Phenolic compounds were extracted following the procedure described by Morreel et al. (2014). Two replicate samples of cells and the medium of the lignin-forming and non-lignin-forming cultures were analyzed at time points 0, 5, 10, 15, and 20 d, and two replicates of the lignin-reforming cultures were analyzed at time points 20 + 4, 20 + 8, 20 + 12, 20 + 16, and 20 + 20 d, making a total of 60 samples.

Phenolic Profiling

Phenolic profiling was performed by analyzing 10 μL of the extract by reverse-phase ultra-HPLC (Accela UHPLC system; Thermo Electron)-electrospray ionization-Fourier transform ion cyclotron resonance-mass spectrometry (LTQ FT Ultra; Thermo Electron). Instrument conditions were as described previously (Morreel et al., 2014) with some minor modifications: the water content in the mobile phase was decreased from 99% (0 min) to 55% (30 min), and the electrospray ionization values for the spray voltage, capillary temperature, sheath gas, and auxiliary gas were set at −4 kV, 275°C, 20 arbitrary units, and 10 arbitrary units, respectively. Preprocessing of the chromatograms, including peak integration, alignment, and the grouping of peaks representing the same compound, was performed as described previously (Morreel et al., 2014). Profiling of the cells and the medium yielded 968 and 551 m/z features that could be reduced further to 430 and 225 m/z peak groups, respectively. However, 75 and 53 m/z features could not be grouped and remained as singletons (Broeckling et al., 2014; Morreel et al., 2014). Statistical analyses were performed on the pseudomolecular ion from each peak group and on the singleton ions (i.e. on 505 and 278 m/z features, respectively). Because 31 m/z features were in common between the 505 and 278 m/z features, a total of 751 unique m/z features were considered as putative nodes in the construction of a CISP network (Morreel et al., 2014). Metabolic conversion types on which the edges were based, and their corresponding number of CSPPs, are shown in Supplemental Table S1. Metabolic conversion types were selected from the list.
given by Morreel et al. (2014) based on their prevalence in C3P networks and presumed presence in gymnosperms. Structural elucidation of the multiple-stage mass spectrometry fragmentation MS^2 spectra was based on Eklund et al. (2008) and Morreel et al. (2004, 2010a, 2010b) for the (neo)lignins/oligolignols, on Cuyckens and Claeyss (2004), Fabre et al. (2001), Hughes et al. (2001), Morreel et al. (2006), and Hvattum and Ekeberg (2003) for the flavonoids, on Zhou et al. (2013) and Abu-Reidah et al. (2014) for the C-glycosylated flavonones, on Morreel et al. (2014) for the phenylpropanoids/benzenoids, and on Gu et al. (2003) and Jaiswal et al. (2012) for the condensed tannins. Shorthand names for the oligolignols are as explained by Morreel et al. (2004). The average abundances at each time point for the structurally characterized compounds are shown in Supplemental Table S3.

**Data Mining of Phenolic Profiles**

The metabolite data were modeled by piece-wise linear regression (lm function) and the final model was reanalyzed by piece-wise robust linear regression (r宋代 function, MASS library) in R version 3.1.3. Compared with classical regression, robust regression is much less biased by outliers. Therefore, the reliability of the classical regression coefficients can be judged by comparing them with those obtained via robust regression. Missing data were replaced by the mean at the particular time point whenever available or by the threshold value (100) in the whole phenolic profile data set. The full model contained knots at times 10, 15, 20, 20 + 4, 20 + 8, 20 + 12, 20 + 16, and 20 + 20 d. Model reduction occurred by iteratively removing the coefficient with the highest P value based on the Wald test and recalculation of the model. The final model contained only coefficients with a P value for culture medium and cell data, see Supplemental Tables S13 and S14, respectively. Significant regression models (P < 0.001) were obtained for 251 and 475 m/z features of the set of 278 and 505 m/z features present in the ultra-HPLC-mass spectrometry chromatograms of extracts from culture media and cells. Of these m/z features having a significant regression model, 114 and 318 m/z features differed significantly in abundance (P < 0.001) between lignin-forming and non-lignin-forming cultures at one or more time points during the experiment.

To cluster m/z features having covarying abundances over time, a correlation analysis was performed. For each m/z feature, the mean abundance at each time point was calculated and pairwise Pearson correlations (cor function) for all m/z features were computed in R version 3.1.3. Pairs of m/z features having correlations above 0.75 were selected (Supplemental Table S15) and a correlation network was constructed (graph edgelist function, igraph library). In this correlation network, groups (called cliques) were searched for (maximal cliques function, igraph library) in which the nodes (representing m/z features) were all mutually connected (i.e., having highly correlated abundances). However, a node can be a member of multiple cliques. Therefore, to allow further clustering, cliques that had more than 50% of their nodes in common were grouped (called clique clusters). All m/z features represented in the whole set of maximal clique clusters were retained in the clique clusters. Only clique clusters containing more than five m/z features are included in Supplemental Figure S2.

**Analysis of Primary Metabolites**

Cells grown for 5 d under either lignin-forming (H2O, 5d) or non-lignin-forming (KI, 5d) conditions (25 mL of cell suspension per 100-mL culture) were harvested and stored as described above for oligolignol profiling. Extraction, GC-MS profiling, and the subsequent processing of GC-MS data were conducted as described previously (Kaplan et al., 2004; Damiani et al., 2005; Desbrosses et al., 2005; Daune et al., 2007) using an HP6890 series mass spectrometry system coupled to a 5973 mass selective detector (Agilent Technologies). The identification of primary metabolites was based on the quadrupole mass spectral and retention time index library (i.e., Q-MSR; Schauer et al., 2005). Sixty-seven metabolites were annotated and their abundances analyzed by Student’s t test (t.test function in R version 3.1.3). In addition, the metabolite abundances were centered and unit variance was scaled prior to PCA (prcomp function in R version 3.1.3).

**Isolation of Pinosinol and Its Chirality Analysis**

Liquid cultures (4.2 L) containing 5 mS KI were initiated as described above. After culturing for 5 d, the culture medium was filtered through Miracloth and centrifuged (4500 g, 10 min, and +4°C). The medium was concentrated by freeze drying to roughly 1 L and extracted in a separation funnel portion wise (500 mL) three times with equal volumes of CHCl3. The combined CHCl3 fractions were evaporated to dryness with a rotavapor (40°C, water pump vacuum) and separated by column chromatography (Merck silica gel 60) by successive elution with hexane and ethyl acetate and finally purified by thin-layer chromatography (Merck silica gel 60, F254, 0.25 mm) using CHCl3-methanol (97.3 v/v), to yield 3.9 mg of pinosinol. The structure of pinosinol was identified by 1H- and 13C-NMR recorded on a Bruker Avance III 400 in CDCl3 (Merck Uvasol; 99.8% D), and optical rotation was measured on a polarimeter (Perkin Elmer 341).

**Spruce Transcriptomics**

For transcriptomic analysis, cell suspension cultures supplemented with water or KI were initiated as described above (100 mL of culture per 500-mL flask). A total of 12 replicate cultures were initiated for each treatment. Cells for RNA isolation were collected at time points 0 and 5 d, and KI-treated cells were additionally collected at time points 20, 20 + 4, and 20 + 12 d (Supplemental Fig. S1B). Approximately 7 mL of the culture was aseptically decanted into 15-mL tubes. After cell sedimentation (0.5–1 min), the supernatant was removed, and the cells were washed briefly with water (5 mL). After cell sedimentation, water was pipetted out and the cells were weighed and frozen in liquid nitrogen. To reduce variation between the biological replicates, cells from four separate flasks treated similarly were pooled at each time point for RNA isolation (hence, three biological replicates for each treatment were analyzed for each time point). Total RNA from approximately 1 g of cultured cells was isolated following the procedure of Chang et al. (1993). The integrity of the isolated total RNA was visualized on a 1.2% (w/v) agarose gel, and total RNA concentrations were determined on a GeneQuant 1300 spectrophotometer (GE Life Sciences). Genomic DNA was removed by DNase I treatment (Qiagen) according to the manufacturer’s instructions. Sample preparation for paired-end Illumina HiSeq sequencing was conducted in the Biomedical Functional Genomics Unit, University of Helsinki, using the NEBNext mRNA sample prep master mix set (version 3.0 6/12). The mRNA was purified using the NEBNext mRNA Magnetic Isolation Module (version 1.1.8/12) modified protocol. The sequencing was done at the Institute for Molecular Medicine Finland (Platform HiSeq 2000; number of lanes, two; paired-end sequencing, 2 × 100 bp; cycles, 101 + 7 + 101, TrueSeq version 3 [TrueSeq PE Cluster Kit version 3, TrueSeq SBS Kit version 3–4HS, 200 cycles]) with 20–3 million paired-end sequences per sample.

**Preprocessing of RNA-Seq Data and Differential Expression Analyses**

The quality of the raw sequence data was assessed, and reads were filtered to remove adapters and trimmed for quality according to the Biomedical Functional Genomics Unit sequencing facility standard procedure. Briefly, FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc) was used to assess the quality of the raw data, from which standard illumina adapters were removed using cutadapt (version 1.2rc2, with settings -m 15 -O 4; Martin, 2011), and reads were trimmed for quality using fastq_quality_trimmer (with settings –Q33 -t 16 +15) from the fastx_toolkit suite version 0.0.13. FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc) was used to assess the quality of the raw data, from which standard illumina adapters were removed using cutadapt (version 1.2rc2, with settings -m 15 -O 4; Martin, 2011), and reads were trimmed for quality using fastq_quality_trimmer (with settings –Q33 -t 16 +15) from the fastx_toolkit suite version 0.0.13. This GFF file is the only one transcript per gene model and contain all high-, medium-, and low-confidence gene models (Nystedt et al., 2013). Briefly, high-confidence gene models are well supported and predicted (greater than 70% coverage by either ESTs or UniProt proteins), medium-confidence gene models have less support (less than 70% and greater than 30% coverage), while low-confidence gene models have little support (less than 30%). This GFF file and the STAR read alignments were used as an input to the HTSeq (Anders et al., 2015) HTseq-count python utility to calculate exon-based read count values. The biseg-count utility takes only uniquely mapping reads into account. Statistical analysis of single-gene differential expression between conditions was performed in R (version 3.2.3; R Core Team, 2015) using the Bioconductor (version 3.2; Gentleman et al., 2004) DESeq2 package (version 1.10.1; Love et al., 2014). False discovery rate (FDR)-adjusted P values were used to assess significance; thresholds of 0.001, 0.01, and 0.05 were used for the analyses. For the data quality assessment and visualization, the read counts were normalized using variance-stabilizing transformation as implemented in DESeq2. The biological relevance of the data (e.g., biological replicate similarity) was assessed by PCA and other visualization tools (e.g., heat maps) using custom R scripts. For all subsequent expression analyses, which also were performed in R, the normalized read counts obtained from DESeq2 were used.
An overview of the data, including raw and post-quality assessment read counts and alignment rates, is given in Supplemental Table S16. Transcription factor families were compiled following the classification in plantfdb.cbi.pku.edu.cn and plantfdb.bio.uni-potsdam.de/v3.0/ by combining it with the UniProt and pfam annotation from ConGenIE.org. Gene family enrichment in the differentially expressed genes was determined using the hypergeometric distribution online tool (www.geneprof.org; Taylor-Teeples et al., 2015). Success is a number of genes differentially expressed in lignifying versus non-lignin-forming comparison ($P_{aush} < 0.001$) from the population of genes in the spruce genome. Sample is the size of the transcription factor family in the genome, and success in the sample is the number of family members in the differentially expressed genes.

**Extracting Excess-KI-Affected Genes**

Three independent differential expression analyses were conducted (H$_2$O$_5$d versus 0, KI$_5$d versus 0, and KI$_5$(5 and 20d) versus 0), and their results intersected (Supplemental Protocol S1). The intersection of KI$_5$(5 and 20d) versus 0, excluding H$_2$O$_5$d versus 0, contains genes that are differentially expressed in response to either the presence of KI (specific effect, bias) or the absence of H$_2$O$_5$ (biologically relevant). As these two effects could not be disentangled and in order to reduce the false-positive rate at the likely expense of increasing the false-negative rate, this gene set (Supplemental Table S5) was removed from further analysis.

**GO Enrichment**

Due to the low GO term coverage (approximately 40%; source, ConGenIE), GO enrichment was conducted using GO terms from the Arabidopsis (Arabidopsis thaliana) homologs (coverage of approximately 84%). agrigo software (Du et al., 2010) was used with Arabidopsis homologs obtained from ConGenIE.org (Complete GO, Fisher’s exact test with Yekutieli FDR under dependency correction, $q < 0.05$). GO treemaps were generated using REVIGO (Supek et al., 2011; FDR < 0.05, medium similarity) and custom R scripts.

**Gene Coexpression Network Inference**

Eight gene network inference methods were run: ANOVA (Küffner et al., 2012), CLR (Faith et al., 2007), GeneNet (Opgen-Rhein and Strimmer, 2007), GENIE3 (Huynh-Thu et al., 2010), NARROMI (Zhang et al., 2013), Pearson, Spearman, and a modified implementation of TIGRESS (Haury et al., 2012); their results were aggregated into a consensus network using an unweighted Borda counting method as suggested by Marbach et al. (2012). The resulting network was then analyzed using RmtGeneNet (Luo et al., 2007) to determine the weight threshold for which its structure deviated from a Poisson to a Gaussian distribution, as assessed by a $\chi^2$ goodness-of-fit test. This allowed for the selection of the core of the network. Similarly, an assessment of the scale-free property of the consensus network, fitting a heavy-tailed distribution using a log-log linear model, led to the selection of two additional cutoffs, one stringent, where the $r^2$ linear regression value was maximized, and one more lenient, where the weight was minimized, for an $r^2$ greater than or equal to 0.8. These two cutoffs allowed for the selection of the stringent and lenient networks. These networks were further visualized and processed using the Cytoscape (Shannon et al., 2003) tool. Subsets of the stringent network were created by selecting the first- and second-degree neighbor nodes of the shikimate pathway, the Phe pathway (Maeda and Dudareva, 2012), and the phenylpropanoid and monolignol pathway genes, except that the genes determined as excess-KI-affected genes by the Venn comparison (Fig. 6; Supplemental Table S5) were not used. The resulting subsets (Supplemental Fig. S5) were further reduced to subnetworks (Fig. 10B) by removing the KI-affected genes.

**Accession Numbers**

Sequence data from this article can be found in the European Nucleotide Archive/Short Read Archive (ENA/SRA) data libraries under the accession number ERP014615.

**Supplemental Data**

The following supplemental materials are available.

**Supplemental Figure S1.** Experimental setups for large-scale phenolic and transcriptomic analyses.

**Supplemental Figure S2.** Accumulation pattern of low-molecular-weight phenolic compounds in the Norway spruce cell culture during lignin formation and when lignin formation was inhibited by H$_2$O$_2$ scavenging.

**Supplemental Figure S3.** REVIGO treemap of the KI-affected genes from Figure 6.

**Supplemental Figure S4.** REVIGO treemaps of differentially expressed transcripts ($P_{aush} < 0.001$).

**Supplemental Figure S5.** Network nodes to the second degree pulled with the shikimate, Phe, phenylpropanoid, and monolignol pathway genes.

**Supplemental Figure S6.** REVIGO treemap of the subnetwork genes with the KI-affected genes removed.

**Supplemental Figure S7.** PCA of GC-MS-profiled compound abundances.

**Supplemental Table S1.** CSSP conversions.

**Supplemental Table S2.** Phenolic analysis and mass spectrometry data: culture medium and cells.

**Supplemental Table S3.** Time course-based raw data means of structurally characterized phenolic compounds.

**Supplemental Table S4.** Differentially expressed genes ($P_{aush} < 0.001$).

**Supplemental Table S5.** Venn diagram gene lists.

**Supplemental Table S6.** Biosynthesis pathway genes.

**Supplemental Table S7.** Genes encoding putative enzymes involved in monolignol oxidation, further modification of dilignols, apoplastic ROS production, and signaling.

**Supplemental Table S8.** Putative antioxidant and related protein genes.

**Supplemental Table S9.** Differentially expressed transcription factor/ regulator families.

**Supplemental Table S10.** Stringent network genes.

**Supplemental Table S11.** Subnetwork genes.

**Supplemental Table S12.** GC-MS data.

**Supplemental Table S13.** Culture medium, piece-wise regression, and clique cluster analysis.

**Supplemental Table S14.** Cultured cells, piece-wise regression, and clique cluster analysis.

**Supplemental Table S15.** Liquid chromatography-mass spectrometry-based clique analysis.

**Supplemental Table S16.** Summary of RNA-Seq sequencing reads and alignments.

**Supplemental Protocol S1.** Identifying and removing the putative KI effect on gene expression: rationale.

**ACKNOWLEDGMENTS**

We thank Tuja Niinimäki for subculturing spruce cells.

Received January 20, 2017; accepted May 16, 2017; published May 18, 2017.

**LITERATURE CITED**


Copyright © 2017 American Society of Plant Biologists. All rights reserved.


