

Evolution of Diterpene Metabolism: Sitka Spruce CYP720B4 Catalyzes Multiple Oxidations in Resin Acid Biosynthesis of Conifer Defense against Insects^{1[C][W][OA]}

Björn Hamberger^{2,3}, Toshiyuki Ohnishi^{2,4}, Britta Hamberger³, Armand Séguin, and Jörg Bohlmann*

Michael Smith Laboratories, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z4 (Bj.H., T.O., Br.H., J.B.); and Canadian Forest Service, Laurentian Forestry Centre, Quebec City, Quebec, Canada G1V 4C7 (A.S.)

Diterpene resin acids (DRAs) are specialized (secondary) metabolites of the oleoresin defense of conifers produced by diterpene synthases and cytochrome P450s of the CYP720B family. The evolution of DRA metabolism shares common origins with the biosynthesis of *ent*-kaurenoic acid, which is highly conserved in general (primary) metabolism of gibberellin biosynthesis. Transcriptome mining in species of spruce (*Picea*) and pine (*Pinus*) revealed CYP720Bs of four distinct clades. We cloned a comprehensive set of 12 different Sitka spruce (*Picea sitchensis*) CYP720Bs as full-length cDNAs. Spatial expression profiles, methyl jasmonate induction, and transcript enrichment in terpenoid-producing resin ducts suggested a role of CYP720B4 in DRA biosynthesis. CYP720B4 was characterized as a multisubstrate, multifunctional enzyme by the formation of oxygenated diterpenoids in metabolically engineered yeast, yeast in vivo transformation of diterpene substrates, in vitro assays with CYP720B4 protein produced in *Escherichia coli*, and alteration of DRA profiles in RNA interference-suppressed spruce seedlings. CYP720B4 was active with 24 different diterpenoid substrates, catalyzing consecutive C-18 oxidations in the biosynthesis of an array of diterpene alcohols, aldehydes, and acids. CYP720B4 was most active in the formation of dehydroabietic acid, a compound associated with insect resistance of Sitka spruce. We identified patterns of convergent evolution of CYP720B4 in DRA metabolism and *ent*-kaurene oxidase CYP701 in gibberellin metabolism and revealed differences in the evolution of specialized and general diterpene metabolism in a gymnosperm. The genomic and functional characterization of the gymnosperm CYP720B family highlights that the evolution of specialized metabolism involves substantial diversification relative to conserved, general metabolism.

Nature produces an impressive assortment of more than 3,300 known diterpenoids, many of which are specialized (i.e. secondary) metabolites of plant origin (Buckingham, 2010). Diterpene resin acids (DRAs), to-

gether with monoterpenes and sesquiterpenes, are major components of the oleoresin defense system of the coniferous trees of the pine family (Pinaceae; Gymnospermae; Phillips and Croteau, 1999; Langenheim, 2003; Keeling and Bohlmann, 2006). Oleoresin terpenoids occur in conifers as complex mixtures of dozens of compounds that may protect the long-lived sessile trees against potentially faster evolving and highly mobile pests and pathogens such as weevils, bark beetles, and their associated fungi (Keeling and Bohlmann, 2006; Boone et al., 2011). For example, the DRAs abietic acid and isopimaric acid inhibit spore germination, and abietic acid also inhibits mycelial growth of the bark beetle *Ips pini* (Kopper et al., 2005). In larch (*Larix laricina*), the DRAs dehydroabietic acid, abietic acid, isopimaric acid, and neoabietic acid are associated with feeding deterrence of gypsy moth (*Lymantria dispar*) larvae (Powell and Raffa, 1999). The same DRAs also reduce feeding damage and growth rates of several sawfly species (Wagner et al., 1983; Schuh and Benjamin, 1984). In Sitka spruce (*Picea sitchensis*), the accumulation of dehydroabietic acid is associated with resistance against spruce shoot weevil (also referred to as white pine weevil [*Pissodes strobi*]; Robert et al., 2010). Beyond their roles in conifer defense, DRAs provide a

¹ This work was supported by the Natural Sciences and Engineering Research Council of Canada (EWR Steacie Memorial Fellowship and Discovery Grant to J.B.), by Genome British Columbia and Genome Canada for the Treenomix Conifer Forest Health Project, and by the University of British Columbia Distinguished University Scholars program (to J.B.).

² These authors contributed equally to the article.

³ Present address: Department of Plant Biology and Biotechnology, Faculty of Life Sciences, University of Copenhagen, Thorvaldsensvej 40, 1871 Frederiksberg C, Copenhagen, Denmark.

⁴ Present address: Division of Global Research Leaders, Shizuoka University, 836 Ohya, Suruga-ku, Shizuoka 422-8529, Japan.

* Corresponding author; e-mail bohlmann@msl.ubc.ca.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantphysiol.org) is: Jörg Bohlmann (bohlmann@msl.ubc.ca).

[C] Some figures in this article are displayed in color online but in black and white in the print edition.

[W] The online version of this article contains Web-only data.

[OA] Open Access articles can be viewed online without a subscription.

www.plantphysiol.org/cgi/doi/10.1104/pp.111.185843

large-volume, renewable resource for industrial bioproducts (Bohlmann and Keeling, 2008).

In spruce species (*Picea* spp.), large amounts of oleoresin diterpenoids accumulate constitutively in axial resin ducts in the phloem and cortex, in radial resin ducts that connect cortex, phloem, and xylem, and to a lesser degree in the scarce axial resin ducts of the xylem (Banan, 1936; Fahn, 1979). In addition, traumatic resin ducts are formed de novo in the cambial zone and in the differentiating xylem as an induced defense response against stem-boring insects or fungal inoculation, providing anatomical structures for the increased production and accumulation of oleoresin (Zulak and Bohlmann, 2010). The formation of traumatic resin ducts and the induced biosynthesis and accumulation of terpenoids in spruce xylem can be triggered by treatment of trees with methyl jasmonate (MeJA; Franceschi et al., 2002; Martin et al., 2002; Zulak et al., 2009). The biosynthesis of diterpenoids is localized to epithelial cells of cortical and traumatic resin ducts (Keeling and Bohlmann, 2006; Abbott et al., 2010; Schmidt et al., 2010; Zulak et al., 2010).

The diterpenoids of conifer oleoresin are largely members of three structural groups, the abietanes, pimaranes, and dehydroabietanes, all of which are characterized by tricyclic parent skeletons (Fig. 1A). These diterpenoids are structurally similar to the tetracyclic *ent*-kaurane diterpenes, which include the ubiquitous GA phytohormones. Both the oleoresin diterpenoids of specialized metabolism and the GAs of general (i.e. primary) metabolism are derived from the common diterpenoid precursor geranylgeranyl diphosphate (GGPP), which is converted by the sequential activity of diterpene synthases (*diTPS*s) and cytochrome P450 monooxygenases (P450s) into diterpene acids (i.e. *ent*-kaurenoic acid or DRAs). The *diTPS*s of conifer DRA biosynthesis yield the tricyclic core structures (Keeling et al., 2008, 2011a). Oxidation at the C-18 carbon of ring A by P450 activity leads to the corresponding diterpene alcohols, aldehydes, and acids (Fig. 1A). The conifer *diTPS* genes of specialized and general diterpenoid metabolism share common evolutionary ancestry (Keeling et al., 2010).

While *diTPS*s for DRA biosynthesis have been cloned and characterized in several species of the pine family (Vogel et al., 1996; Martin et al., 2004; Ro and Bohlmann, 2006; Keeling et al., 2011b), only a single P450 of DRA biosynthesis has been reported, with the cDNA cloning and functional characterization of PtCYP720B1 from loblolly pine (*Pinus taeda*; Ro et al., 2005). PtCYP720B1 was shown to catalyze two consecutive oxidations of abietadienol and abietadienal to abietic acid and also converts the alcohol and aldehyde forms of dehydroabietadiene, isopimaradiene, as well as levopimaradienol to the corresponding DRAs. PtCYP720B1 was characterized as the first member of the newly identified, conifer-specific CYP720B group. However, despite the emergence of additional members of the

CYP720B family in conifer EST sequences (Hamberger and Bohlmann, 2006) and their likely importance for the biochemical diversification of DRA biosynthesis, the complexity of the CYP720B family has not yet been dissected in detail in any species, and enzyme biochemical data have been lacking to support preliminary results that CYP720B enzymes may catalyze all three consecutive C-18 oxidations of DRA biosynthesis (Ro et al., 2005).

The specialized diterpenoid metabolism of DRAs and the biosynthesis of *ent*-kaurenoic acid, the ubiquitous precursor for GAs of general diterpenoid metabolism, are remarkably similar and share some common ancestry (Keeling et al., 2010). However, the diversity of diterpenoid metabolites in the specialized DRA metabolism of conifers appears to be much larger compared with general *ent*-kaurenoic acid biosynthesis. In general, various pathways of specialized metabolism in plants evolved from conserved elements of general metabolism (Bohlmann et al., 1996; Pichersky and Gang, 2000; Ober 2010; de Kraker and Gershenzon, 2011; Shoji and Hashimoto, 2011). The diversification of specialized metabolism can result from events of gene duplication, retention of multiple gene copies, as well as subfunctionalization and neofunctionalization. Apparently, diversification is much more restricted in general metabolism. This pattern is well illustrated with the plant TPS gene family of specialized and general metabolism (Chen et al., 2011) and the gymnosperm TPS-d family in particular (Keeling et al., 2010, 2011b). Likewise, P450s are well known to have multiplied in plant genomes and contribute much to the chemical diversity of plant specialized metabolism (Nelson and Werck-Reichhart, 2011).

To investigate the diversification of specialized diterpene metabolism and the contributing role of P450s in a gymnosperm system, and to compare features of diversification in specialized metabolism with features of conserved general metabolism, we performed a comprehensive gene discovery of members of the conifer-specific CYP720B family across several species of spruce and pine. Based on transcript profiling and phylogeny of 12 full-length (FL) cDNA members of the CYP720B family in Sitka spruce, we identified PsCYP720B4 with a role in DRA biosynthesis. The biochemical properties of the multisubstrate and multifunctional PsCYP720B4 enzyme highlight the molecular underpinnings of the diversification of DRA biosynthesis in the specialized metabolism of conifer defense. The activity of PsCYP720B4 resembles that of CYP701 in *ent*-kaurenoic acid biosynthesis, but their functions appear to have evolved independently in specialized and general metabolism.

Given the large number of thousands of diterpenoid structures in nature, the results of this work in a gymnosperm system are relevant for the broader understanding of specialized diterpene metabolism and the evolution of specialized metabolism in general.

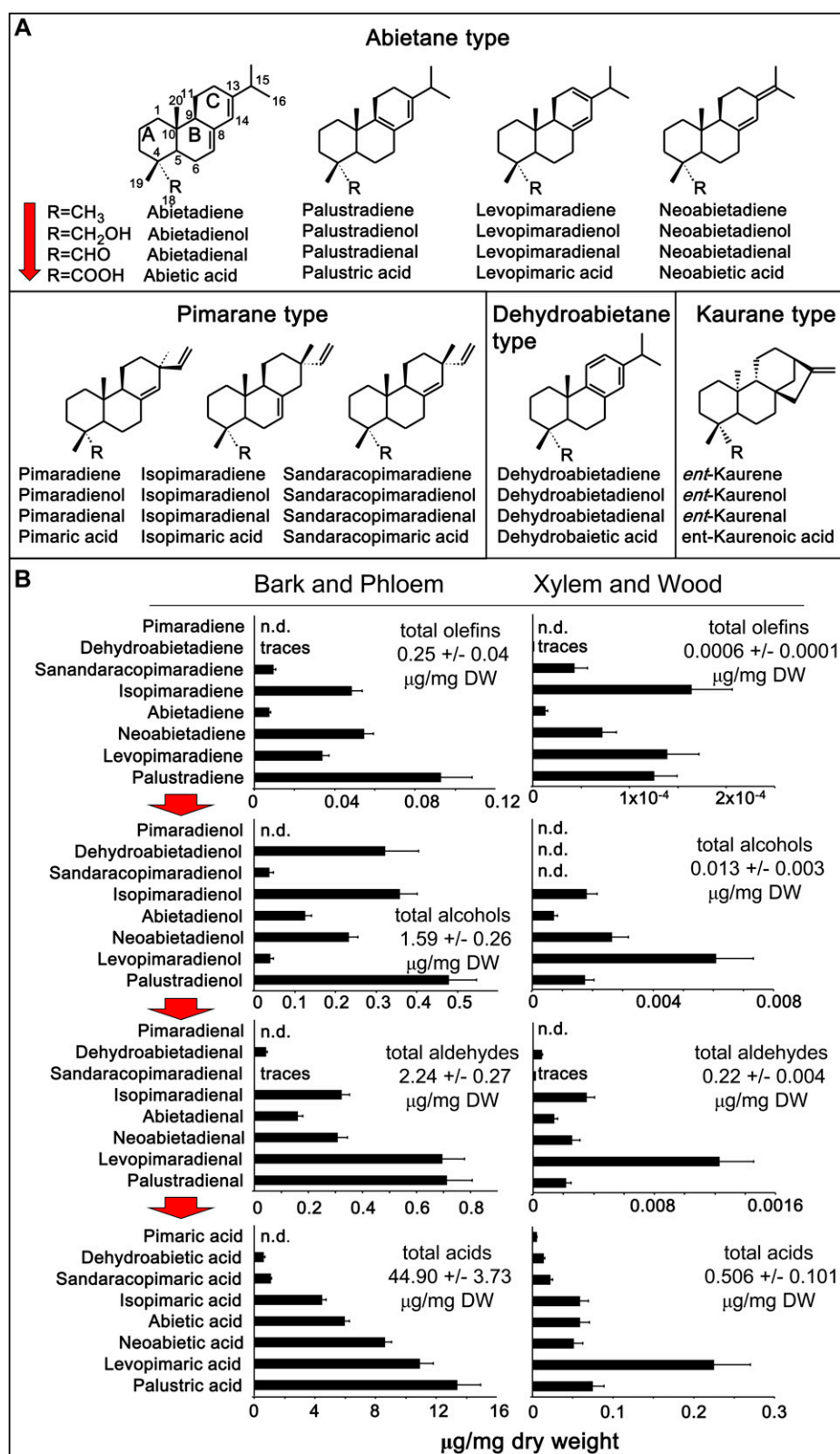


Figure 1. Metabolite profiles of diterpene olefins, alcohols, aldehydes, and acids in the outer (bark/phloem) and inner (wood/xylem) stem tissues of Sitka spruce. A, Structures of tricyclic abietane, pimarane, and dehydroabietane skeletons of diterpenoids of the specialized metabolism of conifer oleoresin and the structurally related tetracyclic *ent*-kaurane diterpene precursors of GA plant hormones in general plant metabolism. The abietane class of diterpenoids includes the various oxygenated forms (i.e. homologous series of olefins, alcohols, aldehydes, and acids) of abietadiene, palustradiene, levopimaradiene, and neoabietadiene. The pimarane class of diterpenoids includes the various oxygenated forms of pimaradiene, isopimaradiene, and sandaracopimaradiene. The dehydroabietane class of diterpenoids includes the various oxygenated forms of dehydroabietadiene. The arrows indicate the direction of modifications (ene → enol → enal → enoic acid). B, Qualitative and quantitative composition of diterpenoids of Sitka spruce outer and inner stem tissues. The arrows indicate the direction of modifications (ene → enol → enal → enoic acid) as well as the increase of concentrations of compounds from top to bottom. "Traces" indicates that traces were identified but amounts were below the limits for accurate quantification. DW, Dry weight; n.d., not detectable. Error bars indicate SE; *n* = 4 independent biological replicates. [See online article for color version of this figure.]

RESULTS

Sitka Spruce Oleoresin Contains a Complex Mixture of Diterpene Olefins, Alcohols, Aldehydes, and Acids

To assess the diversity of oleoresin diterpenoids and the levels of diterpene oxidation in Sitka spruce stems, we determined the quantitative and qualitative composition of oxidized and nonoxidized oleoresin diterpenoids in outer (bark/phloem) and inner (wood/xylem) stem tissues. We identified and quantified 28 different compounds of the abietane, pimarane, and dehydroabietane classes of diterpenoids (Fig. 1A) in the bark and phloem and in the wood and xylem of stems of 5-year-old Sitka spruce (Fig. 1B). In both types of samples, the various C-18 resin acids represented the largest fraction of diterpenoids, followed in decreasing order by much lower amounts of the C-18 aldehydes, alcohols, and olefins (Fig. 1B). Of the total amount of diterpenoids of $48.98 \pm 4.30 \mu\text{g mg}^{-1}$ stem dry weight in the outer stem tissues (Fig. 1B), 92% were resin acids, 5% were aldehydes, 3% were alcohols, and less than 1% were olefins. This result showed that most of the diterpene pool formed by diTPSs is readily converted, presumably by P450 activity, to the fully oxidized C-18 DRAs. Substantially lower amounts of diterpenoids ($0.54 \pm 0.11 \mu\text{g mg}^{-1}$ dry weight) were present in the inner stem tissues (approximately 1% of the amount found in outer stem tissues), which is probably due to the low abundance of terpenoid-producing axial resin ducts in xylem tissue of spruce stems (Zulak and Bohlmann, 2010). Recent work suggested that the levopimaradiene/abietadiene synthase (LAS) type of conifer diTPSs produces epimeric C-13 allylic alcohols as highly unstable, initial products (Keeling et al., 2011a), which dehydrate spontaneously to some of the olefins detected in Sitka spruce (Fig. 1). Although some of the diterpenoids detected in the oleoresin of Sitka spruce may also occur as C-13 alcohols, such compounds have not been found in our or previous analyses and would be difficult to identify with current methods (Keeling et al., 2011a).

The majority of diterpenoids accumulating in outer and inner stem tissues were of the abietane type (palustrates, levopimarates, neoabietates, and abietates; Fig. 1B). Among the pimarane-type diterpenoids, isopimarates were the most abundant. In trees of the clonally propagated Sitka spruce line FB3-425 used in this study, the dehydroabietanes were among the least abundant in both tissue types, with the exception that a relatively high amount of dehydroabietadienol was found in outer stem tissues. Although the same diterpene classes were found in outer and inner stem tissues, some distinct qualitative differences stood out, in particular the relatively high abundance of levopimaric acid and its lesser oxidized precursors in inner stem tissues. Pimaric acid was only found in inner stem tissues. It is important to note that the abundance of individual terpenoids can vary

across different genotypes of Sitka spruce (Robert et al., 2010).

Gene Discovery of the Conifer-Specific CYP720B Subfamily of the CYP85 Clan

The 28 diterpenoids identified in Sitka spruce stem tissues represent various degrees of C-18 oxidation of eight different parent structures and include all of the fully oxidized DRAs (Fig. 1B). Therefore, the question arises of how many of the possible 24 oxidations (eight homologous series of olefins, alcohols, aldehydes, and acids) involved in specialized DRA biosynthesis can be attributed to single or multiple P450 enzymes. To generate a comprehensive inventory of the CYP720B subfamily, we searched nearly 1 million spruce and pine ESTs and constructed a maximum likelihood phylogeny of the predicted CYP720B protein sequences from Sitka spruce, white spruce (*Picea glauca*), interior spruce (*P. glauca* × *Picea engelmannii*), loblolly pine, and lodgepole pine (*Pinus contorta*), rooted with *Arabidopsis thaliana* (*Arabidopsis thaliana*) P450 CYP720A1 (At1g73340) as an outgroup (Fig. 2). We identified 39 CYP720B members that form four distinct phylogenetic clades, clades I to IV. Twelve distinct CYP720Bs were found as Sitka spruce cDNAs, two of those with additional putative allelic variants; 11 distinct CYP720Bs were identified in white spruce, with one putative allelic variant; and 11 distinct CYP720Bs were identified in loblolly and lodgepole pine. Unique gene identifiers were assigned for each distinct CYP720B cDNA (<http://drnelson.uthsc.edu/CytochromeP450.html>; Nelson, 2009). Within each of the four clades (I–IV), groups of probable CYP720B orthologs were identified in Sitka spruce, white spruce, interior spruce, as well as loblolly pine. These phylogenetic patterns are indicative of complete or nearly complete coverage of the expressed complement of Sitka spruce CYP720B genes in the available EST sequences.

Across the four clades of the CYP720B subfamily, members show a high level of sequence divergence, with amino acid sequence identity being as low as 49% between clades I and IV (Supplemental Fig. S1). For comparison, conifer P450s of subfamilies representing general metabolism in the CYP85 clan (e.g. PsCYP85A17) share from 55% to 69% sequence identity with putative orthologs in the distantly related angiosperms (e.g. AtCYP85A2). These findings suggest an increased rate of sequence divergence within the CYP720B subfamily of specialized metabolism.

FLcDNAs of Sitka Spruce PsCYP720Bs

We cloned a representative set of 12 different Sitka spruce PsCYP720B members as FLcDNAs (Fig. 2B). The translated proteins were predicted to cover the entire open reading frames (ORFs), according to the presence of starting Met residues with stop codons upstream within the 5' untranslated regions (UTRs) and patterns of ORF versus UTR sequence similarity

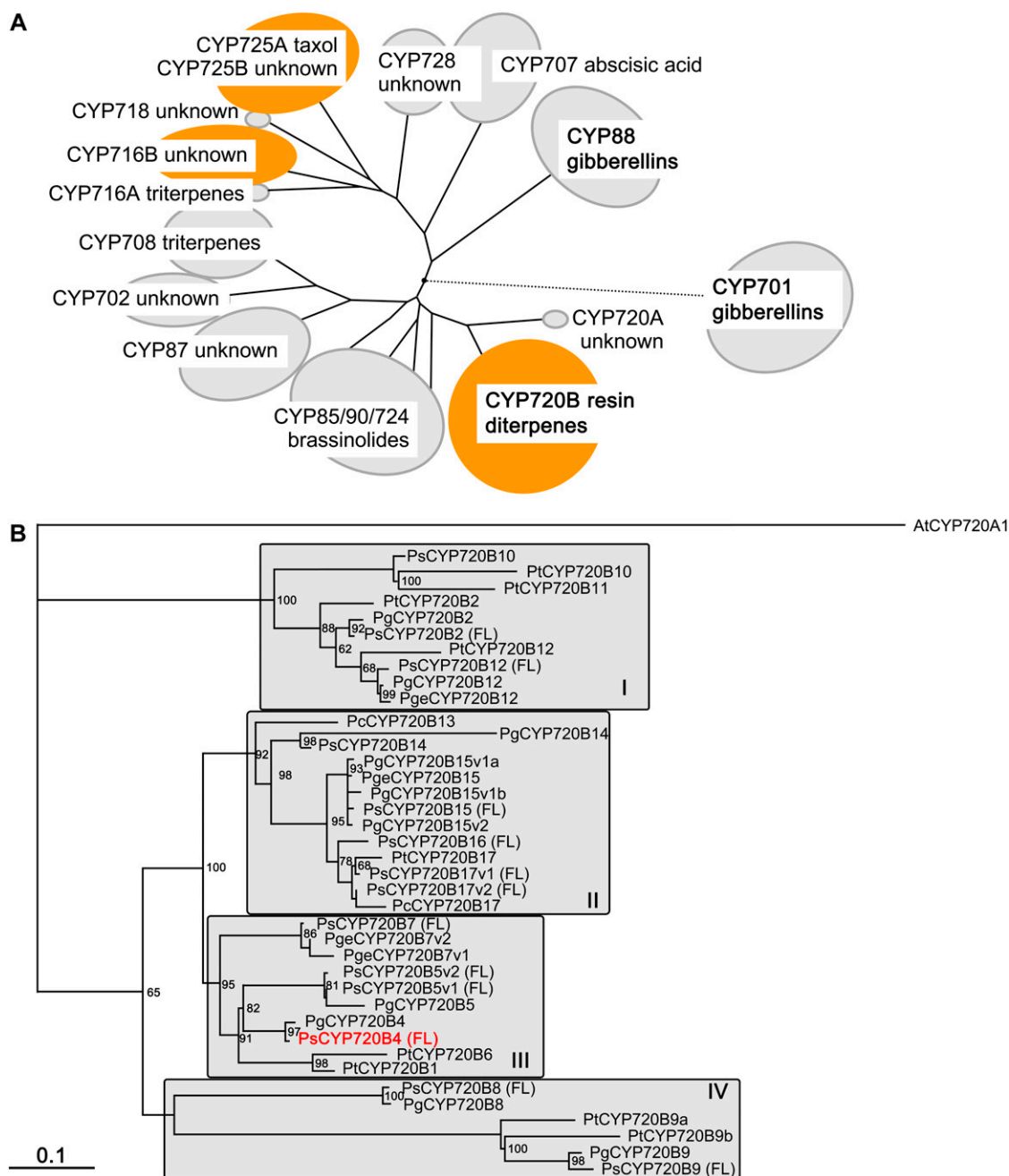


Figure 2. Phylogenetic relationships of the conifer-specific CYP720B genes. A, Schematic of the CYP720B subfamily relative to other P450 subfamilies of the CYP85 clan. Subfamilies are represented with their roles in terpenoid metabolism as known. Subfamilies highlighted in yellow appear to be conifer specific or may have substantially expanded in conifers. Shown in boldface are the CYP720B subfamily of conifer DRA biosynthesis and subfamilies CYP701 (*ent*-kaurene oxidase) and CYP88 (*ent*-kaurenoic acid oxidase) of GA biosynthesis. Note that subfamily CYP701 is not part of the CYP85 clan (dotted line), but it was included in the phylogeny because of the highly similar reactions catalyzed by CYP720B and CYP701. B, Phylogeny of 39 CYP720B deduced amino acid sequences from species of spruce and pine. The maximum likelihood phylogeny was reconstructed using amino acid sequences deduced from cDNAs of Sitka spruce (Ps), white spruce (Pg), loblolly pine (Pt), and lodgepole pine (Pc). The CYP720B sequences cluster into four clades, I to IV. PsCYP720B4 is highlighted in red. The scale bar represents 0.1 amino acid substitutions per site. Bootstrap confidence values over 60% are given. The tree is rooted using Arabidopsis AtCYP720A1 as an outgroup. v, Putative allelic variant. [See online article for color version of this figure.]

within the CYP720B subfamily. We identified two pairs of putative allelic variants (98% or greater amino acid identity) among the 12 PsCYP720B FLcDNA sequences (PsCYP720B5v1/PsCYP720B5v2 and PsCYP720B17v1/PsCYP720B17v2; Fig. 2B; Supplemental Fig. S1). The PsCYP720B FLcDNAs represent members of each of the four CYP720B clades. They also represent the complete set of groups of orthologous spruce and pine CYP720Bs. Only two PsCYP720Bs, PsCYP720B10 and PsCYP720B14, could not be cloned as FLcDNAs despite extensive efforts. Of the PsCYP720Bs, PsCYP720B4 is the closest related to the loblolly pine PtCYP720B1 (Ro et al., 2005), with 86% amino acid identity and 92% similarity (Supplemental Fig. S2C).

Transcript Profiles of PsCYP720Bs Reveal Differential Gene Expression across Different Organs and Tissues

To assess the correlations of PsCYP720B transcripts with oleoresin accumulation in Sitka spruce stems, we performed comparative and quantitative transcript analysis using real-time quantitative (RT-q) PCR for the set of 12 distinct PsCYP720Bs across a range of tissues (Supplemental Fig. S3): wood/xylem of the stem section of the first interwhorl from the top of the tree (W), bark/phloem of the first (B1) and second (B2) interwhorls from the top of the tree, roots (R), young (YN) and mature (MN) needles, and young shoots (YS). Transcript profiles are shown in Figure 3 organized by phylogenetic clades I to IV (Fig. 2).

All PsCYP720Bs of clades I to III, with the exception of PsCYP720B10 (clade I) and PsCYP720B14 (clade II), were expressed at relatively high levels in the three different types of stem samples (W, B1, and B2). These stem tissues accumulate substantial amounts of DRAs (Fig. 1). For two genes, PsCYP720B10 and PsCYP720B14, transcript levels were very low across all samples tested. PsCYP720B8 also showed low levels of transcript in stem tissues, while transcripts for CYP720B9 of the same clade IV were not detected in stem tissues, suggesting that these four genes are not substantially involved in the DRA biosynthesis of Sitka spruce stems. PsCYP720B8 was preferentially expressed in young and mature needles (YN and MN) and young shoots that included needles (Supplemental Fig. S3). PsCYP720B9 was unique in that transcripts were detected only in roots. Conifer roots are poorly characterized for oleoresin biosynthesis, and this finding warrants future studies into the role of PsCYP720B9 in belowground tissues.

Transcript levels of PsCYP720B4 were highest in B1 and B2, followed by YS that also contains bark and xylem, R, and W. This pattern of expression is consistent with a role in DRA biosynthesis in stems and possibly roots. Transcript levels of PsCYP720B4 were low in young and mature needles (YN and MN), in which levels of DRAs were also below detection limit (Miller et al., 2005).

MeJA-Induced Accumulation of PsCYP720B Transcripts of Clade III Is Consistent with a Role in Induced DRA Biosynthesis

Treatment with MeJA simulates the effect of insect attack on Sitka spruce, with increased accumulation of DRAs and diTPS transcripts in the bark and phloem (Miller et al., 2005). Consistent with induced DRA biosynthesis, the PsCYP720Bs of clade III (PsCYP720B4, -5, and -7) showed MeJA-induced increases of transcript accumulation in bark tissue 8 d after treatment of trees (Fig. 4). This time point was chosen for correlation with the MeJA-induced increase of diTPS (Zulak et al., 2009). In addition, a less pronounced induction was also observed with PsCYP720B8 of clade IV. The induction of clade III PsCYP720B transcripts is similar to that of diTPS (PsLAS) transcripts, in agreement with a role in MeJA-induced DRA biosynthesis. The transcript profiles for other PsCYP720Bs of clades I, II, and IV at day 8 after MeJA treatment do not indicate a role in MeJA-induced diterpene resin biosynthesis; however, induced expression with different temporal and/or spatial patterns of transcript accumulation cannot be excluded.

Localization of PsCYP720B Transcript in Laser-Microdissected Stem Tissues Shows Preferential Accumulation in Cortical Resin Ducts

Cortical resin ducts (CRDs) are the primary site of DRA accumulation in spruce stems, and diTPS proteins are localized to epithelial cells of CRDs (Zulak et al., 2010). Using recently established methods for transcript analysis in laser-microdissected CRD tissue of spruce stems (Abbott et al., 2010), we tested by RT-qPCR if PsCYP720B transcripts, and in particular those of clade III, colocalize with diTPS transcripts to CRD epithelial cells (Fig. 5). For comparative spatial analysis, we measured transcript levels of individual PsCYP720Bs in entire stem cross sections dominated by the woody xylem, in the cambial zone, and in epithelial cells of CRDs isolated by laser microdissection. We also measured diTPS (PsLAS) transcripts as a positive control for diterpenoid biosynthesis in CRD epithelial cells and PsCYP73A20, a P450 encoding cinnamate 4-hydroxylase associated with lignification and xylem development, as a negative control. PsLAS transcripts were detected at high levels in laser-microdissected CRDs and were substantially enriched relative to the stem cross sections (Fig. 5). Low levels of PsLAS in cambial zone may be attributed to radial resin ducts. Conversely, PsCYP73A20 transcripts were most abundant in stem cross sections and cambial zone but were detected at low levels in CRDs.

Matching the CRD localization of PsLAS transcripts, several PsCYP720Bs (PsCYP720B2, -4, -12, -15, -16, and -17) showed highly enriched transcript levels in CRDs, with much lower transcript levels in stem cross sections and cambial zone (Fig. 5). Among clade III genes, transcripts of PsCYP720B4 were most highly enriched in CRDs. Consistent with results from transcript anal-

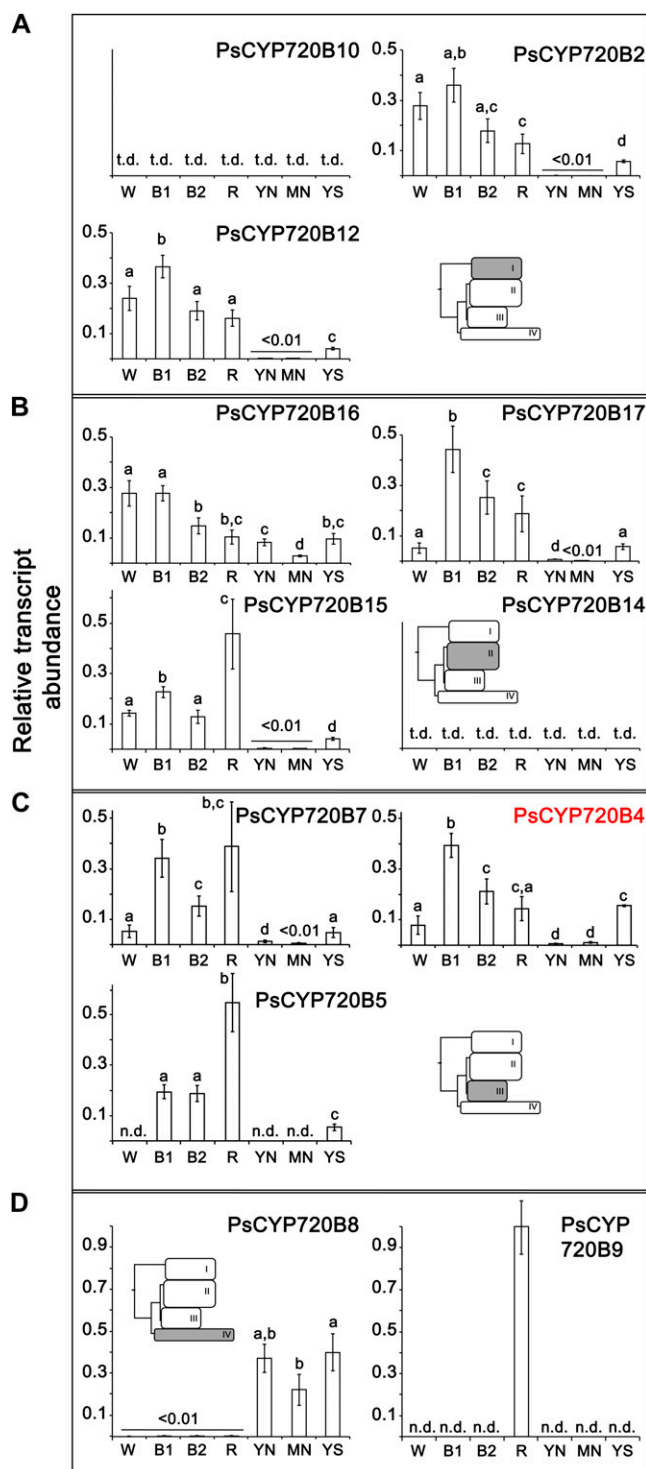


Figure 3. Relative transcript abundance of 12 different PsCYP720Bs in different parts of Sitka spruce. Transcripts were measured in wood and xylem (W) of the first interwhorl from the top; in bark and phloem of the first (B1) and second (B2) interwhorls; in roots (R); in young (YN) and mature (MN) needles; and in young shoots (YS; Supplemental Fig. S3). Genes are grouped according to their phylogenetic association with clades I to IV, as indicated with insets. Transcript abundance was measured by RT-qPCR and normalized to the internal reference TIF. Error bars represent SE values of biological replicates ($n = 3$). Statistically

ysis in various samples across the entire plant (Fig. 3), transcript levels of PsCYP720B10 and -14 were low or not detected in the laser-microdissected stem tissues. Also consistent with the lack of transcripts of PsCYP720B8 and PsCYP720B9 in complex stem samples (Fig. 3), these transcripts were not detected in laser-microdissected cambial zone or CRDs.

Functional Characterization of PsCYP720B4 Using in Vivo Assays of DRA Formation in Metabolically Engineered Yeast

The combined patterns of gene expression profiles (Figs. 3–5) and phylogenetic proximity with PtCYP720B1 (Fig. 2) strongly support a role of PsCYP720B4 in DRA biosynthesis in Sitka spruce. Therefore, we selected PsCYP720B4 for a first comprehensive functional characterization of a CYP720B enzyme, building in part on methods previously established for the initial characterization of PtCYP720B1 (Ro et al., 2005). To test for biochemical functions of PsCYP720B4, we generated two different diterpene olefin-producing yeast (*Saccharomyces cerevisiae*) strains as platforms for P450 in vivo assays. Each strain expressed the yeast GGPP synthase (GGPPS) together with one of two different spruce diTPSs, abietadiene synthase (PaLAS) or isopimaradiene synthase (PaISO; Martin et al., 2004), in the yeast WAT11 (with chromosomally integrated Arabidopsis NADPH-dependent P450 reductase ATR1; here referred to as AtCPR) background (Pompon et al., 1995). The induced strain produced abietadiene with yields of up to 6 mg L^{-1} (Fig. 6). The yeast strain expressing ScGGPPS/PaISO/AtCPR was newly developed for this study and produced isopimaradiene with yields of up to 1.6 mg L^{-1} (Fig. 6). Although PaLAS can form multiple diterpene products in vitro (Keeling et al., 2008), only abietadiene accumulated in the yeast strain expressing ScGGPPS/PaLAS/AtCPR. Diterpenes were detected at approximately 80-fold higher levels in the yeast cell pellet compared with culture medium, suggesting either a lack of efficient secretion and accumulation of diterpenes in yeast cells or their association with cell surfaces after secretion. No diterpenes were detected in the empty-vector control strains.

PsCYP720B4 cDNA was codon optimized for expression in yeast (Supplemental Fig. S2B). When PsCYP720B4 was coexpressed with ScGGPPS/PaISO/AtCPR, levels of the isopimaradiene olefin in yeast cells dropped below the detection limit, which was paralleled by the accumulation of isopimaric acid of up to 0.9 mg L^{-1} (Fig. 6A). When PsCYP720B4 was coexpressed with ScGGPPS/PaLAS/AtCPR, levels of abietadiene were reduced and abietic acid accumulated up to 0.2 mg L^{-1} (Fig. 6B). No oxidized diterpenoids were detected in strains lacking the PsCYP720B4

significant differences are indicated with different letters (t test, $P < 0.05$). n.d., Transcript not detectable; t.d., transcript detected but below the threshold for quantification. [See online article for color version of this figure.]

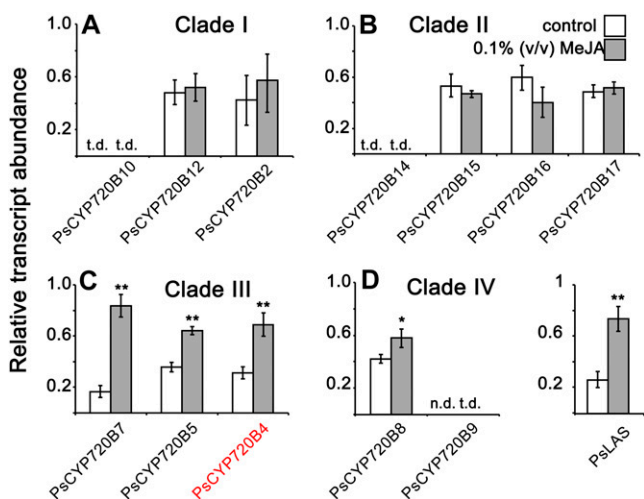


Figure 4. Relative transcript abundance of 12 different PsCYP720Bs and PsLAS in Sitka spruce bark and phloem tissue in response to treatment with MeJA. Genes are grouped according to their phylogenetic association with clades I to IV. Transcript abundance was measured by RT-qPCR and normalized to the internal reference TIF. Error bars represent SE values of biological replicates ($n = 4$). Statistically significant differences of the treated samples from the control are indicated by asterisks (* $P < 0.05$, ** $P < 0.01$). n.d., Transcript not detectable; t.d., transcript detected but below the threshold for quantification. [See online article for color version of this figure.]

cDNA. None of the diterpene alcohol or aldehyde intermediates of the conversion of isopimaradiene to isopimaric acid or of abietadiene to abietic acid accumulated to detectable levels. These results showed convincingly that PsCYP720B4 expressed in yeast catalyzes the complete three-step sequence of diterpene oxidation to form the DRAs isopimaric acid and abietic acid from the corresponding diterpene olefins. Differences in yields of DRAs in the metabolically engineered yeast cells may suggest a substrate preference of CYP720B4 for isopimaradiene over abietadiene. To further assess substrate preferences, we developed additional assay systems for PsCYP720B4 expressed in yeast and *Escherichia coli*.

Functional Characterization of PsCYP720B4 Using in Vitro Assays with Protein in an *E. coli* Membrane Fraction and Authentic Diterpenoid Substrates

Testing of PsCYP720B4 function in metabolically engineered yeast cells was limited to two diterpenes produced by PaLAS and PaISO, abietadiene and isopimaradiene, respectively. For further functional characterization, with a much larger set of synthesized authentic substrates and standards, we used 32 diterpene metabolites, including all of the olefins, alcohols, aldehydes, and acids corresponding to eight different diterpene skeletons (abietadiene, palustradiene, levopimaradiene, neoabietadiene, pimaradiene, isopimaradiene, sanadacopimaradiene, and dehydroabietadiene; Fig. 1A). These compounds represented all of the diterpenoid

resin acids and intermediates detected in Sitka spruce stem tissues (Fig. 1B). Putative C-13 alcohol derivatives of some of these compounds could be speculated as additional diterpenoids of the Sitka spruce resin, but these compounds have never been detected and could not be synthesized as substrates or standards, as they are highly unstable (Keeling et al., 2011a).

We used two strategies to test PsCYP720B4 against the panel of substrates: one approach involved feeding of substrates to yeast cell cultures expressing PsCYP720B4 (see below); the other approach used

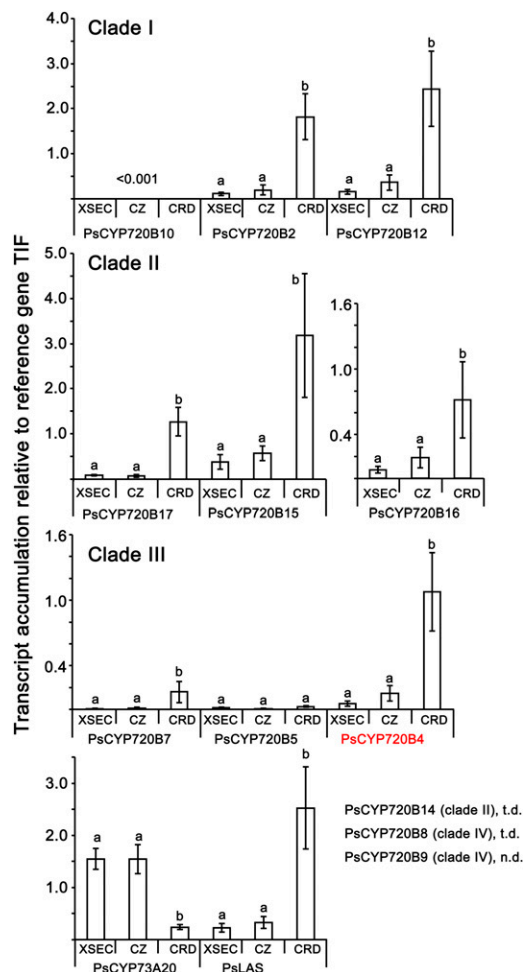


Figure 5. Transcript abundance of 12 different PsCYP720Bs, PsLAS, and PsCYP73A20 in whole cross sections and laser-microdissected cambial zone and cortical individual resin duct tissues of Sitka spruce stems. Transcript abundance was measured by RT-qPCR in whole cross sections (XSEC), cambial zone tissue (CZ), and epithelial tissue of CRDs isolated by laser microdissection from Sitka spruce stems. PsCYP73A20 is Sitka spruce cinnamate 4-hydroxylase; PsLAS is Sitka spruce levopimaradiene/abietadiene diterpene synthase. Transcript abundance was normalized to the internal reference gene TIF. Roman numerals in parentheses indicate the clade. Error bars represent SE values of biological replicates ($n = 3$). Statistically significant differences are indicated by letters (t test, $P < 0.05$). n.d., Transcript not detectable; t.d., transcript detected but below the threshold for quantification. [See online article for color version of this figure.]

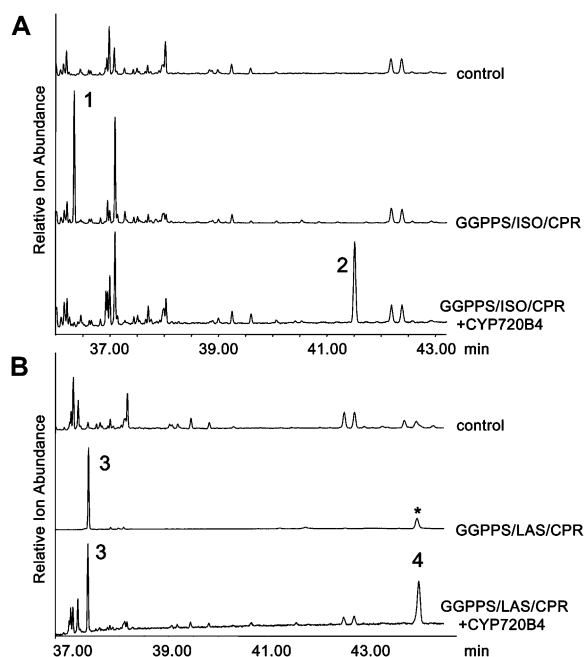


Figure 6. PsCYP720B4-dependent in vivo formation of the DRAs isopimaric acid and abietic acid in engineered yeast. A, GC-MS total ion chromatograms of extracts from engineered yeast strains expressing GGPPS, diTPS, PaISO, and CPR in the presence (bottom trace) or absence (middle trace) of PsCYP720B4. B, GC-MS total ion chromatograms of extracts from engineered yeast strains expressing GGPPS, diTPS, PaLAS, and CPR in the presence (bottom trace) or absence (middle trace) of PsCYP720B4. Retention times and mass spectra of compounds (1, isopimaradiene; 2, isopimaric acid methylester; 3, abietadiene; 4, abietic acid methylester) match those of authentic standards. A minor peak labeled with an asterisk in B differs in retention from peak 4 and represents a phthalate derivative contamination. Top traces in A and B are GC-MS total ion chromatograms of extracts from control engineered yeast strains transformed with empty vectors. For additional details of metabolite identification, see Supplemental Table S1.

PsCYP720B4 protein produced in an *E. coli* membrane fraction. For the latter approach, an N-terminally modified form of PsCYP720B4 was expressed in *E. coli*, and membrane fractions were tested for the presence of P450 protein by SDS-PAGE analysis and carbon monoxide difference spectroscopy (Supplemental Fig. S4). For initial tests of the activity of PsCYP720B4 produced in *E. coli*, we performed in vitro assays with Sitka spruce NADPH-dependent P450 reductase (PsCPR) and dehydroabietadiene, dehydroabietadienol, or dehydroabietadienal as a substrate. All three substrates were converted by PsCPR/PsCYP720B4 (Fig. 7). Following a 30-min incubation, the dehydroabietadiene substrate was converted primarily to the alcohol and the acid, with only traces of the aldehyde detectable (Fig. 7). When dehydroabietadienol was used as the substrate, dehydroabietadienal and dehydroabietic acid were formed as products; and dehydroabietadienal as a substrate was completely converted to dehydroabietic acid. These results with the dehydroabietadiene series of substrates showed, to our knowledge for the first

time with a CYP720B enzyme in an *E. coli* membrane fraction and in vitro assays, that PsCYP720B4 can catalyze three consecutive oxidations at C-18, confirming the in vivo assays with abietadiene and isopimaradiene in engineered yeast cells. Furthermore, these combined results showed that PsCYP720B4 acts on compounds of each of the three different structural classes of Sitka spruce oleoresin diterpenes, the abietane, pimarane, and dehydroabietane classes.

PsCYP720B4 Is Active with 24 Different Diterpenoid Substrates

We used both of the aforementioned assay systems, in vivo feeding of substrates to yeast cultures expressing the PsCYP720B4 cDNA and in vitro assays with

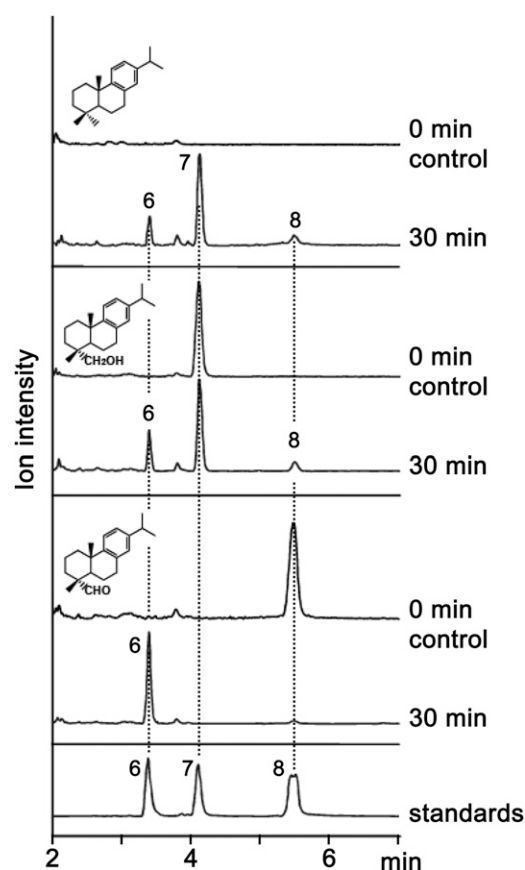


Figure 7. Oxidation of dehydroabietadiene, dehydroabietadienol, and dehydroabietadienal in in vitro assays with PsCYP720B4 protein in an *E. coli* membrane fraction. LC-MS selected ion chromatograms are shown for oxygenated diterpene substrates and products from in vitro assays with recombinant PsCPR and PsCYP720B4 after 0 or 30 min of incubation with dehydroabietadiene, dehydroabietadienol, or dehydroabietadienal as substrates. Assays were performed with 1 mM NADPH and 0.1 unit of CPR. Substrates are shown with their chemical structures. Diterpenoids were identified by comparison of their retention times and mass spectra with those of authentic standards. Numbered peaks are as follows: dehydroabietic acid (6), dehydroabietadienol (7), and dehydroabietadienal (8). For additional details of metabolite identification, see Supplemental Table S1.

PsCYP720B4 protein in an *E. coli* membrane fraction, to further substantiate the emerging pattern of PsCYP720B4 as a multisubstrate and multifunctional (i.e. catalyzing three consecutive oxidations) enzyme of DRA biosynthesis. We tested PsCYP720B4 in both assay systems with each of the 24 synthetic diterpene olefin, alcohol, and aldehyde substrates (Fig. 8). Assays with *E. coli* produced PsCYP720B4 protein and were performed with limiting concentrations of PsCPR to restrict the reactions of each assay to a single oxidation step. This approach allowed us to separately measure the formation of alcohols from olefin substrates, the formation of aldehydes from alcohol substrates, and the formation of acids from aldehyde substrates (Fig. 8A). Results of the *in vitro* assays and liquid chromatography-mass spectrometry (LC-MS) analysis showed that the PsCYP720B4 enzyme is active with each of the 24 different substrates tested, confirming the multisubstrate and multifunctional nature of this P450. These results also indicated a possible preference for the alcohol and aldehyde forms of the diterpene substrates and, in most cases, lower conversion of the olefins. This result may be influenced, in part, by the lower solubility of the olefins in the aqueous assay system, although measures were taken to maximize the solubility for each substrate without compromising the enzyme.

In *in vivo* assays with yeast cells expressing PsCYP720B4 and AtCPR (but not GGPPS and diTPS), induced cultures were supplemented with individual diterpene substrates at concentrations of 10 μ M and the formation of the corresponding DRA was measured by gas chromatography (GC)-MS (Fig. 8B; Supplemental Fig. S5). With the exception of two olefin substrates (neobietadiene and levopimaradiene), all of the substrates tested were converted. With the yeast assays, we detected a much higher rate of conversion when the aldehydes and alcohols were used as substrates compared with the conversion of the olefins (Fig. 8B). When we monitored single-step oxidations for each of the 24 substrates in *in vitro* assays with isolated PsCYP720B4 proteins, the olefins also showed consistently the lowest rate of conversion (Fig. 8A). This result may be due to reduced olefin solubility or to uptake into yeast cells, or it may reflect a lower rate of the three-step oxidation of the olefins compared with the one- or two-step oxidation of the aldehydes or alcohols, respectively.

In both types of assays, PsCYP720B4 consistently showed the highest activity with the dehydroabietadiene group of substrates (dehydroabietadiene, dehydroabietadienol, and dehydroabietadienal) compared with the other seven sets of diterpenoids tested (Fig. 8). The dehydroabietadienes are the only diterpenoids of the Sitka spruce oleoresin with an aromatic C ring (Fig. 1). Notably, the dehydroabietadienes were of low abundance in the diterpene mixture of the particular clonal line of Sitka spruce investigated in this study (Fig. 1), while across a range of more than 100 genotypes of Sitka spruce dehydroabietic acid was associated with

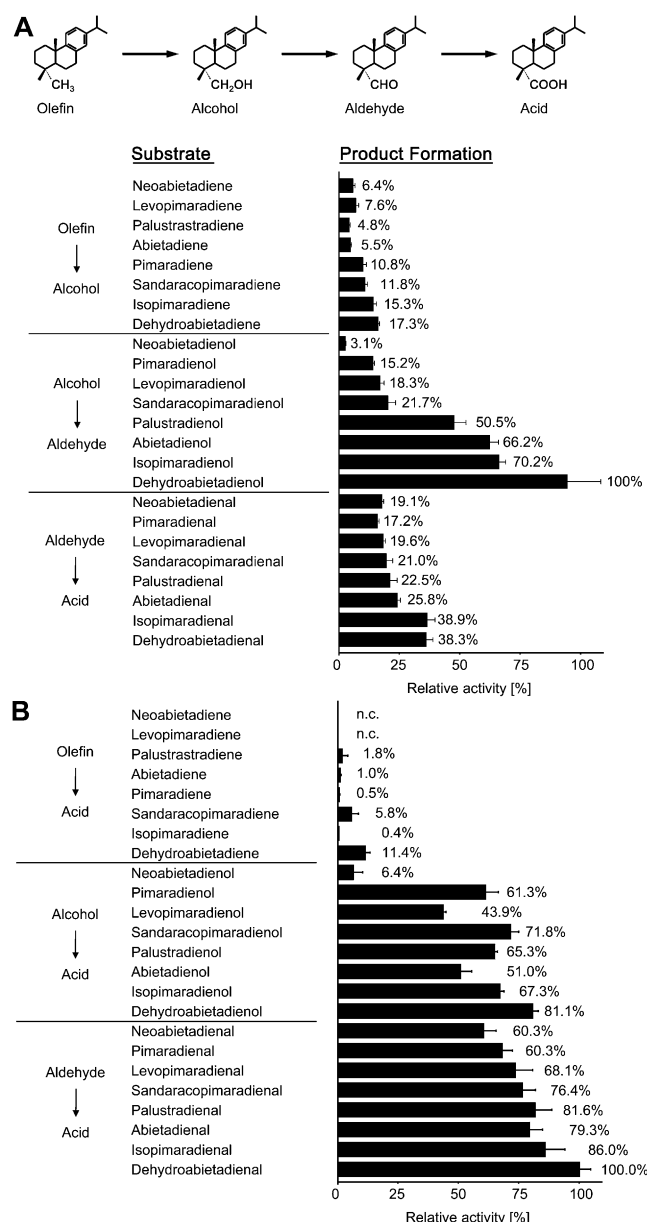


Figure 8. Relative activity of PsCYP720B4 with 24 different diterpene substrates representing eight olefins, eight alcohols, and eight aldehydes. A, Relative *in vitro* transformation of authentic diterpene substrates with PsCYP720B4 protein in an *E. coli* membrane fraction and PsCPR produced in *E. coli*. Transformations were restricted to single-step oxidation by limitation of PsCPR. Assays were performed with 1 mM NADPH and 5 milliunits of CPR. Product formation is given relative to activity with dehydroabietadienol (100%) and was determined with three independent biological replicates by LC-MS. B, Relative *in vivo* transformation of authentic diterpene substrates to the corresponding DRAs in cultures of yeast cells expressing PsCPR and PsCYP720B4. DRA formation with each substrate is shown relative to the formation of dehydroabietic acid from dehydroabietadienal (100%). The complete experiment was repeated three times, each time with two biological replicates, each of which was analyzed with two technical replicates by GC-MS. n.c., No conversion detected. For both A and B, it is important to note that different physicochemical properties of olefins, alcohols, and aldehydes (e.g. differences in solubility) may affect the apparent lower activity with the olefin substrates.

genetic resistance against spruce weevil (Robert et al., 2010).

Apparent Kinetic Properties of PsCYP720B4 Enzyme Show a Preference for the Formation of Dehydroabietic Acid

To substantiate the finding of PsCYP720B4 as a multisubstrate and multifunctional P450 enzyme, we determined basic kinetic properties with nine representative substrates. PsCYP720B4 protein expressed in *E. coli* was used for these experiments, as the isolated membrane fractions allowed for reliable and robust quantification of P450 via carbon monoxide difference assays (Supplemental Fig. S4). The selected substrates used for kinetic characterization included a homologous series of olefins, alcohols, and aldehydes from each of the three major classes of Sitka spruce oleoresin diterpenes, the abietane-, pimarane-, and dehydroabietane-type diterpenes (Fig. 1). Members of these classes were chosen to represent substrates that showed high to intermediate levels of relative activity in the yeast *in vivo* assays and in the *in vitro* assays with PsCYP720B4 produced in *E. coli* (Fig. 8). Specifically, we tested the enzyme kinetics of PsCYP720B4 with dehydroabietadiene, dehydroabietadienol, dehydroabietadienal, isopimaradiene, isopimaradienol, isomiparadienal, palustradiene, palustradienol, and palustradienal (Table I). The enzyme kinetic properties were measured in *in vitro* assays with PsCYP720B4 protein in an *E. coli* membrane fraction, with reactions restricted to a single oxidation step by limitation of PsCYP activity. It is important to note that different physicochemical properties of olefins, alcohols, and aldehydes (e.g. differences in solubility) may affect the apparent kinetic parameters.

PsCYP720B4 exhibited apparent dissociation constants (K_m) in the low micromolar range with all diterpene olefins, alcohols, and aldehydes tested. PsCYP720B4 exhibited the lowest apparent K_m values for two of the diterpene alcohols, palustradienol and dehydroabietadienol. Apparent turnover numbers (k_{cat}) were slightly higher for diterpene aldehydes and alcohols

compared with diterpene olefins. Thus, the apparent catalytic efficiencies (k_{cat}/K_m) of PsCYP720B4 for dehydroabietadienol were 1.5 and three times greater than those for dehydroabietadiene and dehydroabietadienal, respectively. These findings are in good agreement with the relative activity determined in *in vivo* and *in vitro* assays. Consistent with basic conversion rates assessed in assays with all 24 substrates (Fig. 8), PsCYP720B4 showed higher apparent k_{cat}/K_m with precursors of dehydroabietic acid and isopimaric acid than with precursors of palustric acid.

CYP720B4 RNA Interference Causes Changes in DRA Composition

To validate the *in vivo* functions of CYP720B genes in plants, we targeted the orthologous PgCYP720B4 sequence in white spruce (above 99% ORF identity at the nucleic acid level with PsCYP720B4) by RNA interference (RNAi). RNAi has recently been established for white spruce (A. Séguin, unpublished data) and has not yet been reported to be successful in any other conifer except for radiata pine (*Pinus radiata*; Wagner et al., 2007, 2009). Since Sitka spruce and white spruce are closely related species (Yeh and Arnott, 1986), we used white spruce as a suitable system for analysis of the effect of targeted RNAi on CYP720B transcript levels and diterpenoid metabolites.

We generated four independent PgCYP720B4 RNAi lines carrying a stably integrated 290-bp PgCYP720B4 hairpin loop RNAi element under the control of the white spruce ubiquitin promoter. RT-qPCR analysis in stem samples of 1-year-old RNAi lines (five biological replicates for each of the four RNAi lines) showed significant decreases of PgCYP720B4 transcripts to as low as 25% relative to transcripts in stems of control trees transformed with empty pCambia vector (Fig. 9; Supplemental Table S2). We also tested the effect of cross-reactivity and cosuppression with other members of the PgCYP720B family in white spruce. For the seven genes where RT-qPCR yielded detectable transcript levels in stems of white spruce controls and RNAi lines, we found the strongest RNAi effect with

Table I. Apparent kinetic parameters of CYP720B4

Assays were performed with 1 mM NADPH and 5 milliunits of CPR. Different physicochemical properties of olefins, alcohols, and aldehydes (e.g. differences in solubility) may affect the apparent kinetic parameters.

Substrate	K_m	k_{cat}	k_{cat}/K_m
	μM	min^{-1}	$min^{-1} \mu M^{-1}$
Palustradiene	11.6 ± 1.56	0.70 ± 0.07	0.06 ± 0.002
Palustradienol	3.42 ± 0.85	1.21 ± 0.03	0.35 ± 0.03
Palustradienal	10.1 ± 1.58	4.07 ± 0.63	0.40 ± 0.02
Isopimaradiene	13.9 ± 2.07	3.19 ± 0.03	0.23 ± 0.005
Isopimaradienol	10.9 ± 1.90	6.56 ± 0.92	0.60 ± 0.03
Isopimaradienal	8.34 ± 1.46	6.40 ± 0.63	0.77 ± 0.04
Dehydroabietadiene	8.54 ± 1.45	3.99 ± 0.15	0.47 ± 0.02
Dehydroabietadienol	4.92 ± 1.12	5.17 ± 0.29	1.05 ± 0.07
Dehydroabietadienal	10.7 ± 1.33	8.71 ± 0.41	0.81 ± 0.02

the PgCYP720B4 target but also some considerable suppression with other members of clades III, II, and I in decreasing order of effect (Supplemental Table S2).

RNAi-mediated reduction of PgCYP720B transcript levels was paralleled by significant (independent *t* test, $P < 0.05$) changes in the amounts of several (but not all) types of DRAs (Fig. 9). Levels of dehydroabietic acid, isopimaric acid, and sandaracopimaric acid were significantly reduced in at least three of the four independent RNAi lines. Levels of dehydroabietic acid were reduced by approximately 30% to 50% compared with the empty-vector transformed control. Isopimaric acid and sandaracopimaric acid were reduced in three of the four RNAi lines by approximately 35% to 65% of their respective levels in the controls. An RNAi effect on levels of abietic acid was found only in a single line, and no significant change of abundance of neoabietate, palustrate, and levopimaradiene was detected in any of the four RNAi lines. Pimaric acid was not detected in the white spruce samples.

Reduced levels of several DRAs in the RNAi lines support a role of CYP720B4, and perhaps other CYP720B genes, in DRA biosynthesis in planta. The strong effect on the accumulation of dehydroabietic acid, compared with the other DRAs, is in agreement with the results from enzyme assays and the kinetic characterization of PsCYP720B4, which showed preferred enzyme activity in the formation of dehydroabietic acid. The lack of an

obvious effect of PgCYP720B4 RNAi on several other DRAs, such as abietic acid, palustric acid, levopimaric acid, and neoabietic acid, suggests that other P450(s), presumably of the CYP720B subfamily, are required for the efficient oxidation of the abietane class of spruce oleoresin diterpenoids in planta, or that due to the incomplete silencing of CYP720B4 residual enzyme activity contributes to the observed profile. The fact that more than one type of DRA is affected by PgCYP720B4 RNAi is in agreement with the multisubstrate nature of the PsCYP720B4 enzyme and may also be due to the cosuppression effect with others CYP720Bs.

Convergent Evolution of Nearly Identical Enzyme Activities of CYP720B and CYP701 in Specialized and General Metabolism, Respectively

The activity of PsCYP720B4 in the three-step C-18 oxidation of tricyclic diterpenoids in conifer specialized metabolism resembles most closely the activity of the distantly related CYP701 in the three-step C-18 oxidation of the tetracyclic *ent*-kaurene to *ent*-kaurenoic acid in the general metabolism of GA biosynthesis (Helliwell et al., 1999). Phylogenetically, however, members of the conifer-specific CYP720B family are more closely related, within the CYP85 clan, for example to CYP88, which catalyzes the oxidation of *ent*-kaurenoic acid to GA₁₂ (Helliwell et al., 2001), and to CYP90, which catalyzes C-22 and C-23 hydroxylations in the biosynthesis of brassinosteroid hormones (Szekeres et al., 1996; Choe et al., 1998; Ohnishi et al., 2006; Fig. 2). CYP701, CYP88, and CYP90 genes had not previously been cloned from a gymnosperm. To validate the patterns of phylogeny with the homologous Sitka spruce set of genes, we identified the genomic sequences covering the full-length exonic coding regions and introns for four representative genes of PsCYP720B clades I to IV as well as for Sitka spruce PsCYP701, PsCYP88, and PsCYP90 (Fig. 10). Shared positions of introns indicate common evolutionary ancestry (Carmel et al., 2007), and analysis of the intron-exon structure can be used to infer evolutionary relationships between genes (Rogozin et al., 2005). The size of nine exons and the positions of the introns were highly conserved across the PsCYP720B, PsCYP88, and PsCYP90 genes (Fig. 10). In contrast, the PsCYP701 gene is organized in eight exons, and the sizes of exons and positions of introns were substantially different from the gene structure of PsCYP720B, PsCYP88, and PsCYP90. These results support the interpretation of convergent functional evolution of CYP720B and CYP701 for the principally identical three-step C-18 oxidation of diterpenes to diterpene acids (i.e. DRAs or *ent*-kaurenoic acid). An alternative interpretation would be that C-18 diterpene oxidation was an ancestral function that remained conserved in the distantly related CYP720B and CYP701; however, this is less likely if the CYP85 clan evolved from a sterol-metabolizing ancestor, as suggested by Nelson and Werck-Reichhart (2011). Notably, CYP720B functions in the

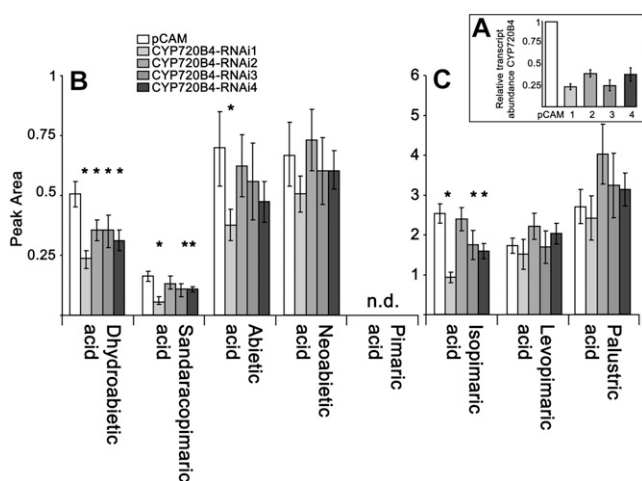


Figure 9. Effect of RNAi on PgCYP720B4 transcript abundance and composition of DRAs in white spruce stems. A, Relative transcript abundance of CYP720B4 in stems of 12-month-old white spruce plants of four independent RNAi lines and vector-transformed controls. Error bars represent SE values of biological replicates ($n = 3$). B, Relative abundance of DRAs in stems of 1-year-old white spruce plants of four independent RNAi lines and vector-transformed controls. Scale = 0 to 1. Asterisks indicate significant (*t* test, $P < 0.05$) reduction compared with the empty-vector control. Error bars represent SE values of biological replicates ($n = 5$). C, Relative abundance of DRAs in stems of 1-year-old white spruce plants of four independent RNAi lines and vector-transformed controls. Scale = 0 to 6.

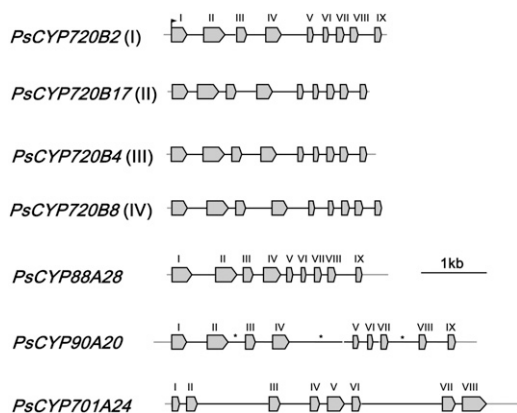


Figure 10. Intron-exon gene structures of representative Sitka spruce PsCYP720B genes of clades I to IV, PsCYP90A20, PsCYP88, and PsCYP701. Roman numerals indicate exons. The scale bar represents 1 kb.

context of the diversification of conifer specialized metabolisms, while CYP701 functions in the context of a highly conserved general metabolism.

DISCUSSION

Using several lines of evidence, we identified CYP720B4 as a multisubstrate and multifunctional P450 enzyme with an important role for the structural diversity of specialized diterpenoid metabolism. A major product of CYP720B4 is dehydroabietic acid, which is associated with the resistance of Sitka spruce against its major insect pest, the spruce weevil (Robert et al., 2010). Beyond the general relevance of this work for a fundamental understanding of the evolution of specialized (i.e. secondary) metabolism, as discussed below, the work described here also demonstrates the association of P450 activity of DRA biosynthesis with the specialized cells of terpenoid-accumulating resin ducts and provides new genes and proteins that can be developed into genetic markers or biomarkers for Sitka spruce resistance against weevils (Hall et al., 2011).

The abundance and structural diversity of diterpenoid oleoresin defenses in conifers make these gymnosperm trees a highly relevant system for the general understanding of the biosynthesis and evolution of specialized and general (i.e. primary) diterpenoid metabolism in plants. Knowledge of the evolution of specialized and general diterpenoid metabolism in gymnosperms complements similar work on diterpenoids in angiosperms, which is perhaps best established in the rice (*Oryza sativa*) system (Peters, 2006; Swaminathan et al., 2009). Specialized diterpenoid metabolism of DRAs and the biosynthesis of *ent*-kaurenoic acid, the ubiquitous precursor for GAs of general diterpene metabolism, are remarkably similar (Fig. 11). These pathways share structurally similar

diterpene olefins, alcohols, aldehydes, and acids derived by the activity of diTPS and P450 enzymes from the common GGPP precursor. The general reactions catalyzed by the diTPSs and P450s of the *ent*-kaurenoic acid and DRA pathways are the same in both pathways. However, a number of fundamental differences are emerging with regard to the evolution and diversification of gymnosperm genes, enzymes, and metabolites of related general and specialized diterpenoid biosynthesis. A major aspect of the conifer diterpenoid metabolism is the structural diversity of metabolites in specialized metabolism relative to comparable segments of general metabolism (Fig. 11). A diverse assortment of 32 different known compounds of eight homologous series of olefins, alcohols, aldehydes, and acids represents the specialized DRA metabolism of Sitka spruce (Figs. 1 and 11), and undoubtedly, additional specialized compounds exist in this species, such as simpler diterpenoid alcohols, contrasting *ent*-kaurenoic acid and its precursors in general metabolism.

Differences at the level of structural diversity in specialized and general metabolism are reflected in the corresponding diTPS (Keeling et al., 2010) and P450 genes and enzymes. In previous work, it was shown that conifer diTPSs of general and specialized metabolism share common ancestry, possibly resembling a bifunctional diTPS of the spike moss *Physcomitrella patens* (Hayashi et al., 2006; Keeling et al., 2010). Similar to the spike moss enzyme, bifunctional diTPSs of conifer specialized metabolism catalyze two consecutive cyclization reactions (Peters et al., 2000; Keeling et al., 2008). In contrast, conifer diTPSs of general *ent*-kaurenoic acid biosynthesis, namely *ent*-copalyl diphosphate synthase and *ent*-kaurene synthase, are monofunctional enzymes (Fig. 11; Keeling et al., 2010). The conifer diTPSs of specialized and general diterpene metabolisms arose by divergent evolution (Keeling et al., 2010; Chen et al., 2011), as is also supported by insights from the structural analysis of a gymnosperm diTPS (Köksal et al., 2011). At the genome level, conifer diTPSs of general metabolism seem to be encoded by single-copy genes that are functionally and structurally conserved across the angiosperms and gymnosperms (Keeling et al., 2010), suggesting that genes of general GA biosynthesis evolved prior to the separation of the angiosperms and gymnosperms more than 300 million years ago and remained largely unchanged with regard to structure and function. This pattern of conserved gene evolution is in agreement with a conserved role of *ent*-kaurenoic acid as an essential intermediate in the general metabolism of GA biosynthesis. In contrast, diTPSs of conifer specialized metabolism are substantially divergent from angiosperm diTPSs (Chen et al., 2011). Hallmarks of conifer diTPSs of specialized metabolism (Keeling et al., 2008, 2011b), which set them apart from diTPSs of general metabolism (Keeling et al., 2010), are retention and neofunctionalization of duplicated genes that give rise to the initial diversification of specialized metabolism. The products of the

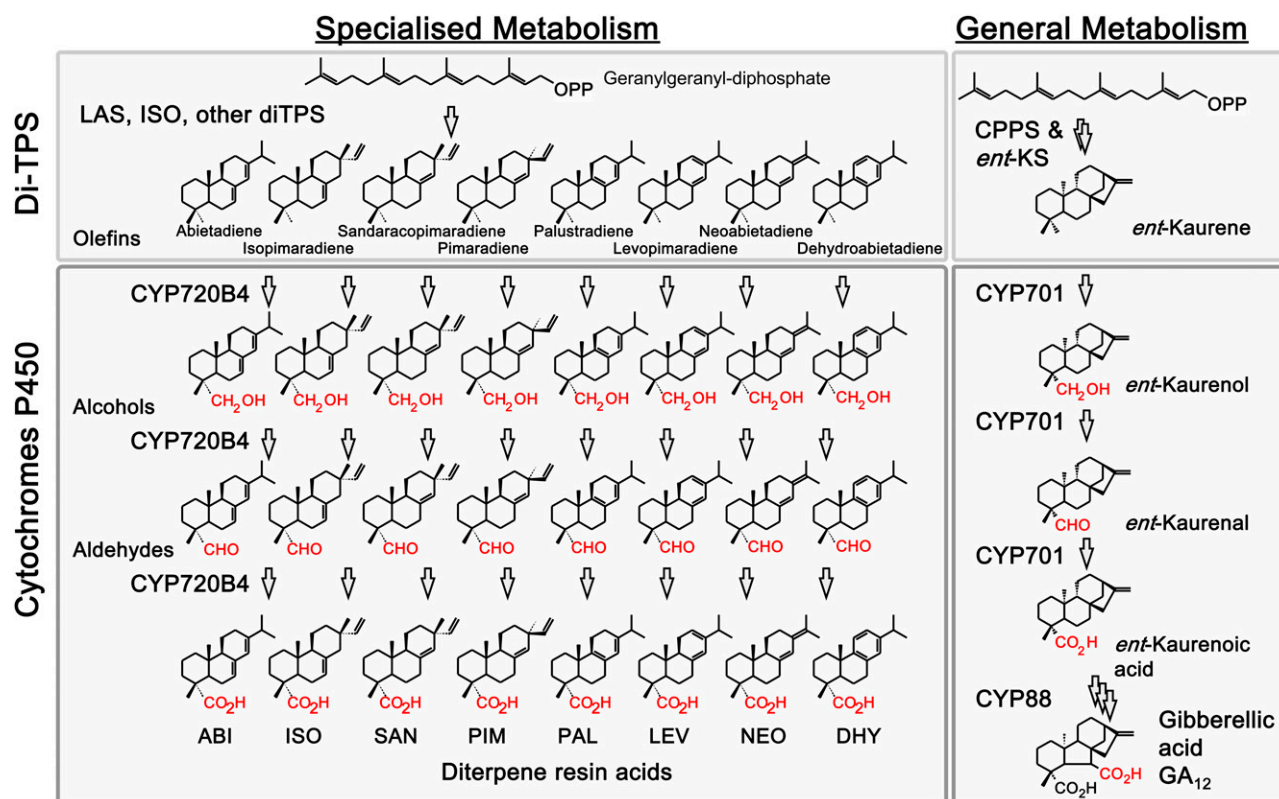


Figure 11. Comparison of specialized DRA biosynthesis and general GA biosynthesis in spruce. Similar diTPS and P450 enzymes catalyze the formation of diterpene acids in the form of a diverse assortment of DRAs in specialized metabolism and in the form of the ubiquitous *ent*-kaurenoic acid in general metabolism. Spruce diTPSs and P450s of the general metabolism of *ent*-kaurenoic acid seem to be encoded by single-copy genes. The diTPSs and P450s of general *ent*-kaurenoic acid metabolism are functionally conserved across the angiosperms and gymnosperms. In contrast, the diTPSs (TPS-d subfamily) and P450s (CYP720B subfamily) of specialized DRA biosynthesis are unique to gymnosperms and encoded by members of multigene families that resulted, most likely, from multiple events of gene duplication. Individual diTPSs of specialized DRA metabolism can produce multiple diterpene olefins. CYP720B4, as a representative P450 of specialized DRA metabolism, accepts multiple diterpene substrates and catalyzes three consecutive oxidations, thus contributing to the structural diversity of DRA in specialized metabolism. ABI, Abietic acid; ISO, isopimaric acid; SAN, sandaracopimaric acid; PAL, palustric acid; PIM, pimaric acid; LEV, levopimaric acid; NEO, neoabietic acid; DHY, dehydroabietic acid.

diTPSs of specialized metabolism are the substrates for P450 enzymes of the conifer-specific CYP720B family (Fig. 11), of which PsCYP720B4, to our knowledge, is the first member that has been characterized in detail, with 24 different substrates as a multifunctional and multisubstrate enzyme. PsCYP720B4 substantially increases the spectrum of diterpenoids in specialized DRA metabolism.

While the recently discovered CYP720B genes exist with multigene families in conifers (Fig. 2), no CYP720B member has been identified in the vast amount of angiosperm genome and transcriptome sequences. The CYP720Bs of specialized diterpene metabolism seem to be taxonomically restricted to conifers and perhaps other gymnosperms, where they diverged through events of repeated gene duplication. In contrast to CYP720Bs of specialized metabolism, mechanistically identical CYP701 P450s of *ent*-kaurenoic acid biosynthesis in general metabolism appear to be encoded by

single-copy functional genes in conifers (Hamberger and Bohlmann, 2006). Thus, with all known diTPSs and P450s of specialized and general diterpene metabolism of conifers, there is a consistent pattern of gene multiplication and diversification in specialized metabolism, in contrast to single-copy genes and conserved function in general metabolism. The conserved functions of conifer diTPSs and CYP701 P450s of general metabolism extend to their orthologs in the angiosperms. In contrast, the conifer diTPSs and CYP720B P450s of specialized metabolism do not have apparent orthologs in the angiosperms.

Apparently, the specific function of three-step consecutive oxidation at the C-18 position of cyclic diterpenes evolved convergently in the CYP701 and CYP720B P450s. It is important to note that there are other cytochrome P450 enzymes in different pathways that also catalyze consecutive oxidations into alcohols, aldehydes, and acids, such as CYP71AV1 in the bio-

synthesis of artemisinic acid (Ro et al., 2006) and CYP716A12 in the biosynthesis of triterpenoid saponins (Carelli et al., 2011). However, in the case of CYP701 and CYP720B4 from Sitka spruce or CYP720B1 from loblolly pine (Ro et al., 2005), the respective cytochrome P450 enzymes use very similar diterpene substrates and catalyze oxidations at the same C-18 positions of the A ring (Figs. 1 and 11), supporting the notion of convergent evolution of CYP701 and CYP720B in general and specialized metabolism, respectively. The level of functional diversification within the CYP720B family of any given species is currently not known and awaits future functional characterization of additional members of this gene family.

CONCLUSION

In summary, the results of the characterization of CYP720B4, along with the comparative analysis of P450 and diTPS genes of specialized and general gymnosperm diterpenoid metabolism, support models of plant metabolism (Hartmann, 1996, 2007) in which (1) specialized and general metabolism share common origins, (2) specialized metabolism evolved (and continues to evolve) with patterns of taxon-specific pathway diversification on the basis of gene duplication and neofunctionalization, while (3) general metabolism is conserved across distant plant lineages, such as the gymnosperms and angiosperms. The conserved general diterpenoid metabolism of *ent*-kaurenoic acid biosynthesis, with no obvious gene duplication in conifers, is also indicative of tight control of general plant hormone metabolism. Conversely, the specialized metabolism of defense-related DRA biosynthesis may not only tolerate relaxed metabolic diversification, but plants may indeed benefit from the diversity and dynamics of a broader assortment of specialized metabolites to protect long-lived, sessile plants against faster evolving pests and pathogens. Based on the study of specialized metabolism at the level of metabolites, enzyme functions, and transcriptome-wide capture of multigene families, it appears that specialized metabolism, in contrast to general metabolism, is permissive to redundancy and diversification, which may serve as an adaptive advantage, for example, in the chemical defense of sessile, long-lived conifer trees.

MATERIALS AND METHODS

Plant Materials

Clonally propagated Sitka spruce (*Picea sitchensis*; clone FB3-425; 5 years old) and white spruce (*Picea glauca*; clone PG653; 1 year old) were maintained as described previously (Miller et al., 2005). Treatment with MeJA (0.1% in 0.1% [v/v] Tween 20) was as described previously (Miller et al., 2005). From the upper two interwhorls of Sitka spruce, the following tissues were harvested for metabolite and/or RNA analysis: young shoots, bark/phloem (tissues outside the vascular cambium, including phloem, cortex, and periderm), xylem/wood, and mature and young needles (Supplemental Fig. S3). These samples as well as roots were flash frozen in liquid nitrogen and stored at -80°C .

Plasmids and Host Strains

Dual expression vectors pESC-His and pESC-Leu were from Stratagene. Yeast strain WAT11 (MATa; ade2-1; his3-11, -15; leu2-3, -112; ura3-1; canR; cyr+) carrying the Arabidopsis (*Arabidopsis thaliana*) AtCPR chromosomally integrated under the control of the GAL promoter (Pompon et al., 1995) was provided by Dr. D. Werck-Reichhart (Institut de Biologie Moléculaire des Plantes, Centre National de la Recherche Scientifique, Strasbourg, France) and used as a host for heterologous gene expression. The pCWori⁺ vector for heterologous expression of CYP720B4 in *Escherichia coli* was provided by Dr. M. Mizutani (Fujita et al., 2006).

Diterpene Standards and Substrates

DRAs were purchased from Orchid Cellmark. Diterpenoid olefins, alcohols, and aldehydes were synthesized from the corresponding acids by Best West Laboratories following published methods (Ro et al., 2005). The purity of synthesized diterpenoids was assessed by ^1H -NMR, ^{13}C -NMR, and GC-MS and was equal to or exceeded the purity of the starting materials. Abietadiene contained an unknown contamination, and levopimaradiene/levopimaradiene derivatives contained other diterpenoid isomers (up to 30%) that could not be removed.

Laser Microdissection

Details of sample preparation and laser microdissection procedures are described by Abbott et al. (2010). Stem segments of 1 cm length from the uppermost interwhorl were longitudinally cut into 20- μm sections on a cryostat (CM3050S; Leica) and laser microdissected (DM 6000B/CTR 6500; Leica).

Extraction of Diterpenoids

Methods for the extraction and derivatization of diterpenoids and GC-MS analysis were adapted from Lewinsohn et al. (1993) and performed with three technical replicates and at least three biological replicates for each sample type. All steps were carried out in 2-mL amber silanized vials (Agilent). Tissue (50–100 mg) was extracted by shaking overnight in 1.5 mL of *tert*-butylmethylether with 200 $\mu\text{g mL}^{-1}$ dichlorodehydroabietic acid and 100 $\mu\text{g mL}^{-1}$ isobutylbenzene as internal standards. Extracted tissue was dried at 50°C for 2 d prior to weighing to determine the dry weight. For derivatization, 500 μL of the organic supernatant was mixed with 100 μL of methanol and 120 μL of trimethylsilyl-diazomethane (2.0 M in diethyl ether; Sigma-Aldrich) and incubated at room temperature for 20 min. Derivatized samples were concentrated under N_2 gas and adjusted to 500 μL with diethyl ether before GC-MS analysis.

Analysis of Metabolites by GC-MS and LC-MS

Due to the complexity of the samples extracted from plant tissues and yeast in vivo assays, advantage was taken of the superior resolution and accuracy of GC-MS. For extracts from in vitro assays with heterologously expressed proteins, LC-MS was the superior method to measure the large number of samples from kinetic studies. Metabolite extracts from yeast in vivo assays were analyzed by GC-MS on a 6890A/5993N GC-MSD system (Agilent) with a SolGel Wax capillary column (30 m $\times$ 0.25 mm $\times$ 0.25 μm ; SGE Incorporated) in selected ion monitoring scan mode (scan, mass-to-charge ratio [m/z] 40–400; selected ion monitoring, m/z 270, 272, 286, 288, 314, and 316 [dwell time 50]). Volumes of 0.5 μL per sample were injected in pulsed splitless mode at 250°C with a column flow of 1 mL min^{-1} helium and 20 ψ pulse pressure. The GC oven temperature was programmed to rise from an initial 40°C at $3^{\circ}\text{C min}^{-1}$ to 110°C , at $10^{\circ}\text{C min}^{-1}$ to 180°C , and at $15^{\circ}\text{C min}^{-1}$ to a final temperature of 260°C for a run time of 60.7 min. Metabolites from in vitro assays with proteins produced in *E. coli* were identified and quantified by LC-MS on a LC-MSD-Trap-XCD_plus 1100 series device (Agilent) on a Zorbax SB-C18 rapid resolution HT column (4.6 $\times$ 40 mm $\times$ 1.8 μm) in positive electrospray ionization mode (dry temperature, 350°C ; nebulizer, 55 pounds per square inch). Samples of 2.5 μL were injected with a column flow of 1 mL min^{-1} in isocratic mode (acetonitrile:water, 85%:15% + 0.2% formic acid) with the diode array detector scanning from 190 to 400 nm for a run time of 10 min. Diterpenes from Sitka spruce tissue extracts were analyzed by GC-MS with an

AT1000 capillary column (30 m $\times$ 0.25 mm $\times$ 0.25 μ m; Alltech) in scan mode (scan, m/z 40–500). Samples of 0.5 μ L were injected in pulsed splitless mode at 250°C with a column flow of 1 mL helium min⁻¹ and 30 p.s.i. pulse pressure. The GC oven temperature was programmed to rise from an initial 150°C at 1.5°C min⁻¹ to 220°C and at 20°C min⁻¹ to 240°C for a run time of 63.7 min. Quantification was repeated with GC-flame ionization detection on an Agilent 6890 GC system under the same conditions. No significant differences were observed for the metabolites detected. Compounds were identified and quantified by comparison with the synthesized authentic standards.

RNA Isolation

RNA isolation followed the method of Kolosova et al. (2004) with minor modifications. Extractions were scaled down to 50 to 100 mg of tissue; the RNA pellet was washed with 70% ethanol after LiCl precipitation and resuspended in nuclease-free water; RNA integrity was assessed using the RNA 6000 Nano kit (Bioanalyzer 2100; Agilent). cDNA synthesis was as described by Miller et al. (2005) using SuperScript III (Invitrogen). For microdissected samples, RNA was isolated from 20- μ m-thick longitudinal stem sections, epithelial cells of CRDs, and cambial zone tissue using the RNAqueous-Micro Kit (Ambion) as described by Abbott et al. (2010). Bioanalyzer 2100 RNA pico-chip assays (Agilent) were used to determine RNA integrity and concentration.

CYP720B Gene Discovery and FcDNA Cloning

General procedures for mining of the pine and spruce sequence databases for genes of the CYP85 clan, including CYP720B genes, were described previously (Hamberger and Bohlmann, 2006). In silico-generated contigs for candidate CYP720B genes from three spruce species (Sitka spruce, white spruce [*Picea glauca*], and interior spruce [*P. glauca* $\times$ *Picea engelmannii*]) and two pine species (loblolly pine [*Pinus taeda*] and lodgepole pine [*Pinus contorta*]) were used to reconstruct a phylogeny of the CYP720B family by maximum likelihood analysis (Guindon and Gascuel, 2003) and for homology-based cloning of Sitka spruce CYP720B cDNAs. 5' RACE (Ambion) was used to obtain FcDNAs of partial Sitka spruce ESTs. PCR amplification of FcDNAs was performed over 40 to 45 cycles with Phusion polymerase (Finnzymes) and gene-specific oligonucleotide primers (Integrated DNA Technologies). Amplicons were ligated into pJET (Fermentas), and inserts were sequenced completely.

RT-qPCR Analysis

cDNA was synthesized from equal amounts of total RNA using SuperScript III reverse transcriptase (Invitrogen) with random hexamer oligonucleotides. The efficiency of RT reactions was assessed with internal calibrator reactions on four reference transcripts, elongation factor 1 (ELF1), translation initiation factor-like (TIF), polyubiquitin, and tubulin α -chain homolog. For the quantification of target transcripts, TIF and ELF were chosen as references, as they showed the lowest variation across all tissue types (Abbott et al., 2010). For each sample, RT-qPCR was performed with three technical replicates on three or four biological replicates. Detection of amplification was performed in real time with a final melting curve using the DyNAmo HS SYBR Green kit (Finnzymes) on a DNA Engine Opticon2 apparatus (MJ Research) with Opticon3 software. HK-uracil *N*-glycosylase reaction (EPICENTRE) preceded the qPCR. A real-time PCR miner program (Zhao and Fernald, 2005) was used to determine efficiency and cycle number over threshold (C_T) values. Efficiency-corrected ΔC_T values were calculated, and $\Delta\Delta C_T$ was used to quantify relative differences in transcript accumulation. AmpliX was used for oligonucleotide optimization (AmplifX 1.5.4; N. Jullien; <http://ifjr.nord.univ-mrs.fr/AmplifX>). Representative PCR products for all sample types and gene targets were sequenced to confirm target specificity. Sequences of oligonucleotides are given in Supplemental Text S1.

Isolation and Sequencing of Genomic CYP720B Clones

Isolation of genomic DNA from young shoots of Sitka spruce was performed according to Carlson et al. (1991) with the following modifications. After two rounds of chloroform/isoamyl alcohol extraction, nucleic acids were pelleted with isopropanol, washed with 70% ethanol, resuspended in nuclease-free water, and treated with RNase A and proteinase K before the addition of ice-cold isopropanol. DNA was hooked out of the solution, resuspended in nuclease-free water, and quantified spectrophotometrically.

Genomic clones covering the complete coding region were generated with 45 PCR cycles based on cDNA coding sequences using the primers listed in Supplemental Text S1 and Phusion polymerase (Finnzymes). PCR products were cloned into pJET (Fermentas), and inserts were sequenced completely.

Expression of Diterpene Synthases and PsCYP720B4 in Yeast

Plasmid pESC-His::PaLAS/ScGGPPS for the coexpression of PaLAS and ScGGPPS in *Saccharomyces cerevisiae* was described previously (Ro et al., 2005). Using the same approach, pESC-His::PaISO/ScGGPPS was generated for the coexpression of ScGGPPS and PaISO lacking the N-terminal plastid-targeting sequence (Martin et al., 2004). Codon optimization of the 5' region of PsCYP720B4 (Supplemental Fig. S2) recoded 561 bp (Batard et al., 2000). A transcription initiation enhancer element (Cavener and Ray, 1991) was introduced 5' of the initiation codon. GC content was adjusted from 47% in the 5' coding region to 33% (Marín et al., 2003), and rare yeast codons were replaced. Using eight partially overlapping long-mer oligonucleotides, the optimized 561-bp 5' region of PsCYP720B4 was synthesized by PCR and ligated to 891 bp of the native 3' region. *SpeI*-compatible restriction sites were introduced by PCR, and the purified amplicon was ligated into the yeast expression vector pESC-Leu (pESC-Leu::PsCYP720B4). Oligonucleotides are given in Supplemental Text S1. To generate yeast strains for P450-dependent de novo diterpenoid formation, plasmids pESC-His::PaLAS/ScGGPPS or pESC-His::PaISO/ScGGPPS were individually transformed into the WAT11 (AtCPR) background, and the resulting strains were transformed with pESC-Leu::PsCYP720B4. Engineered yeast strains to assay for P450-dependent oxidation of supplemented diterpenoid substrates were generated by transformation of pESC-Leu::PsCYP720B4 in the WAT11 (AtCPR) background. Details of yeast transformation, medium, and culture conditions were described previously (Pompon et al., 1995; Ro et al., 2005; Gietz and Schiestl, 2007). Details of heterologous protein expression and metabolite analysis in yeast strains carrying the plasmid pESC-Leu::PsCYP720B4 or the plasmid pairs pESC-His::PaLAS/ScGGPPS//pESC-Leu::PsCYP720B4 or pESC-His::PaISO/ScGGPPS//pESC-Leu::PsCYP720B4 are described below. Diterpenoids extracted from yeast cultures were analyzed by GC-MS after derivatization.

In Vivo Assays for PsCYP720 Activity Using Yeast Strains Carrying the Plasmid pESC-Leu::PsCYP720B4

The basic procedures were as described previously by Ro et al. (2005). Heterologous protein expression in yeast cultures was initiated at an optical density at 600 nm of 0.6 by transfer into minimal selection medium containing 2% Gal. PsCYP720B4 protein fused to the FLAG epitope was detected in crude cell lysates 16 h post induction by western analysis with an anti-FLAG antibody (Stratagene) as described by Ro et al. (2005). Diterpenoids were dissolved in acetonitrile (alcohols), ethyl acetate (alcohols and aldehydes), or diethyl ether (olefins), depending on the substrate polarity. While acetonitrile is 100% water soluble, ethyl acetate and to a lesser extent ethyl ether are less miscible with water. Due to their low polarity, olefins are the least soluble in the aqueous systems used for the assays. Hence, it cannot be excluded that the assays may have been influenced by the lower solubility of the olefins in water. The optimal concentration for diterpenoid substrates added to yeast was determined in a series of tests using nine terpenoids, selected based on initial high, medium, and low conversion, (1) the olefins dehydroabietadiene, isopimaradiene, and palustradiene, (2) the alcohols dehydroabietanienol, isopimaradienol, and palustradienol, or (3) the aldehydes dehydroabietadienal, isopimaradienal, and palustradienal, at 100 to 1 μ M. For example, complete conversion of dehydroabietadienal to dehydroabietic acid was observed at 10 μ M. To analyze DRA formation in yeast, individual compounds were added to induced yeast cultures to a final concentration of 10 μ M in a 20-mL culture volume and incubated for 1 h at 29°C.

In Vivo Assays for PsCYP720 Activity Using Yeast Strains Carrying the Plasmid Pairs pESC-His::PaLAS/ScGGPPS//pESC-Leu::PsCYP720B4 or pESC-His::PaISO/ScGGPPS//pESC-Leu::PsCYP720B4

For the analysis of de novo diterpenoid formation in engineered yeast strains, cultures were grown to an optical density at 600 nm of 0.6 before

transfer into minimal selection medium containing 2% Gal. After growth for 16 h, yeast cells were pelleted by centrifugation at 3,200g for 10 min. Cell pellets were extracted twice with 5 mL of diethyl ether with 0.5 μ M eicosene and 3 μ M dichlorodehydroabietic acid (Orchid Cellmark) as an internal standard. Anhydrous sodium sulfate was added to remove residual water, and the combined organic phases were concentrated to 500 μ L. To derivatize the DRAs, 100 μ L of methanol and 60 μ L of trimethylsilyl-diazomethane (Aldrich; 2.0 M in diethyl ether) were added, and the reactions were incubated at room temperature for 20 min. Organic solvent were removed under N₂ gas, and samples were adjusted to 200 μ L with diethyl ether before GC-MS analysis.

Expression of PsCPR and PsCYP720B4 in *E. coli*

Nine amino acids of the N-terminal membrane anchor region of PsCYP720B4 were replaced by the N-terminal region (MALLAVF) of bovine CYP17 α (Supplemental Fig. S4; Barnes et al., 1991; Kagawa et al., 2003). The modified PsCYP720B4 cDNA was ligated into *Nde*I and *Hind*III sites of the bacterial expression vector pCWori⁺. PsCPR was amplified from a Sitka spruce cDNA library constructed from flushing buds and ligated into *Nco*I and *Sall* sites of pET28 (Novagen) to generate the bacterial expression construct pET28-PsCPR. Oligonucleotides are given in Supplemental Text S1. PsCPR was heterologously expressed in the *E. coli* strain BL21(DE3), and JM109 was used for the expression of recombinant PsCYP720B4 as described (Fujita et al., 2006). PsCYP720B4 and PsCPR were expressed, purified, and quantified as detailed below. The reconstitution assay with recombinant PsCPR and PsCYP720B4 was performed following previously described methods (Ohnishi et al., 2006). Diterpenoid analysis by LC-MS is described above.

Production of PsCYP720B4 in an *E. coli* Membrane Fraction and in Vitro Enzyme Assays with PsCPR and PsCYP720B4 Produced in *E. coli*

The *E. coli* strain BL21 (DE3) was used for heterologous expression of PsCPR, and the enzyme was purified following Mizutani and Ohta (1998). PsCPR activity was assayed by the rate of cytochrome *c* reduction. The *E. coli* strain JM109 was used for the expression of recombinant CYP720B4 following Fujita et al. (2006). *E. coli* cells were harvested by centrifugation at 25,000g for 10 min and treated on ice for 30 min with 0.5 mg mL⁻¹ lysozyme in a buffer containing 50 mM Tris-HCl (pH 7.5), 250 mM Suc, and 0.5 mM EDTA. After centrifugation, spheroplasts were sonicated in buffer A containing 50 mM potassium phosphate (pH 7.3), 20% (w/v) glycerol, 1 mM EDTA, 0.1 mM dithiothreitol, and 1 mM phenylmethylsulfonyl fluoride and subjected to centrifugation at 3,000g for 15 min to remove cell debris. The supernatant was further centrifuged at 100,000g for 1 h, and the resulting pellet was homogenized with buffer A without phenylmethylsulfonyl fluoride to recover the membrane fractions. The concentration of active P450 was estimated from the carbon monoxide difference spectrum (Omura and Sato, 1964). For in vitro assays, substrates were added, dissolved in acetonitrile (alcohols), ethyl acetate (alcohols and aldehydes), or diethyl ether (olefins), depending on the substrate polarity. PsCYP720B4 in *E. coli* membrane fractions was mixed with purified PsCPR consisting of 50 mM potassium phosphate (pH 7.25), 25 pmol mL⁻¹ PsCYP720B4, PsCPR (0.1 unit for analysis of enzyme activities; 5 milliunits for analysis of relative activities and kinetic parameters), diterpene substrates (20 μ M for analysis of enzyme activities and relative activities; a range of substrate concentrations for kinetic assays [0.625–40 μ M]). Reactions were initiated by the addition of 1 mM NADPH and were carried out at 30°C for 30 min (analysis of enzyme activities and relative activities) or 20 min (analysis of kinetic parameters). After termination by adding 300 μ L of ethyl acetate, 10 μ L of 1 mM *O*-methyl podocarpic acid was added as an internal standard. The reaction products were extracted three times with an equal volume of ethyl acetate. The organic phase was collected and evaporated with N₂. The residue was dissolved in 100 μ L of acetonitrile before LC-MS analysis.

RNAi Suppression of CYP720B4: Binary Vector Construction, Spruce Transformation, and Somatic Seedling Production

RNAi-mediated suppression of CYP720B4 was performed in white spruce clone Pg653, where efficient transformation protocols are established (Klimaszewska et al., 2005), and followed the general strategies of RNAi-

mediated suppression described by Wesley et al. (2001). A 290-bp region of PgCYP720B4 with minimal similarity to paralogous CYP720B genes was amplified by PCR (oligonucleotides are listed in Supplemental Text S1) and cloned in the sense and antisense orientations in the multicloning site of pTRAIN on either side of the intron as described by Levée et al. (2009). The RNAi cassette is positioned downstream of the ubiquitin promoter-UTR region (Christensen et al., 1992). The resulting plasmid was *Hind*III digested, and the excised cassette carrying the RNAi construct was ligated into pCambia1305.2 (www.cambia.org). *Agrobacterium tumefaciens*-mediated stable transformation of the white spruce embryogenic line Pg653, embryo maturation, and somatic seedling production following the protocol of Klimaszewska et al. (2005). Ten independent kanamycin-resistant transgenic lines were obtained, and four were selected by transcript profiling of PgCYP720B4 by RT-qPCR on total RNA of whole seedlings. Plants were grown for 1 year prior to characterization, including analysis by RT-qPCR and metabolite analysis as described above.

Phylogenetic Analyses

Phylogenetic analyses were performed on aligned amino acid sequences (dialign2 [http://bioweb2.pasteur.fr/]; manually curated, Bioedit version 7.0.9 [http://www.mbio.ncsu.edu/BioEdit/bioedit.html]). The maximum likelihood tree was bootstrapped (PhyML [http://www.atgc-montpellier.fr/phyml/binaries.php]; four rate substitution categories, γ -shape parameter optimized, Jones-Taylor-Thornton substitution model, BioNJ starting tree, and 100 bootstrap repetitions; Guindon and Gascuel, 2003) and displayed as a phylogram using TreeView32 1.6.6 (http://taxonomy.zoology.gla.ac.uk/rod/treeview.html; Page, 2002). Tree topologies were supported using the independent maximum likelihood algorithm TREE-PUZZLE 5.2 (http://www.tree-puzzle.de; Schmidt and von Haeseler, 2007).

The sequences reported in this article have been deposited in GenBank under accession numbers HM245397 to HM245410 and HQ888858.

Supplemental Data

The following materials are available in the online version of this article.

Supplemental Figure S1. Multiple sequence alignment of PsCYP720Bs.

Supplemental Figure S2. Sequence optimization of PsCYP720B4 for expression in yeast.

Supplemental Figure S3. Organs and tissues of Sitka spruce used for transcript and metabolite profiling.

Supplemental Figure S4. Heterologous expression of PsCYP720B4 in *E. coli*.

Supplemental Figure S5. Transformation of diterpene olefins, alcohols, and aldehydes added to yeast expressing PsCYP720B4 and PsCPR, and formation of DRAs.

Supplemental Table S1. Identification of diterpenoids by LC-MS and GC-MS.

Supplemental Table S2. Cosuppression effects of CYP720B4 RNAi.

Supplemental Text S1. Oligonucleotides used for PCR.

ACKNOWLEDGMENTS

We thank Dr. Masaharu Mizutani (Graduate School of Agricultural Science, Kobe University) for providing the pCWori⁺ vector, Dr. David Nelson (Department of Molecular Sciences, University of Tennessee Health Sciences Center, Memphis) for assigning P450 nomenclature, Dr. Krystyna Klimaszewska (Natural Resources Canada) for white spruce seedlings, Dr. David Ellis (formerly CellFor, Inc.) for Sitka spruce seedlings, Ms. Lina Madilao (University of British Columbia) for excellent analytical support, Ms. Karen Reid (University of British Columbia) for superb laboratory management support, and Drs. Dawn Hall and Katherine Zulak, Michael Smith Laboratories, University of British Columbia, Vancouver) for critical reading of the manuscript and helpful comments.

Received August 19, 2011; accepted October 7, 2011; published October 12, 2011.

LITERATURE CITED

- Abbott E, Hall D, Hamberger B, Bohlmann J (2010) Laser microdissection of conifer stem tissues: isolation and analysis of high quality RNA, terpene synthase enzyme activity and terpenoid metabolites from resin ducts and cambial zone tissue of white spruce (*Picea glauca*). *BMC Plant Biol* 10: 106
- Banan MW (1936) Vertical resin ducts in the secondary wood of the Abietineae. *New Phytol* 35: 11–46
- Barnes HJ, Arlotto MP, Waterman MR (1991) Expression and enzymatic activity of recombinant cytochrome P450 17 α -hydroxylase in *Escherichia coli*. *Proc Natl Acad Sci USA* 88: 5597–5601
- Batard Y, Hehn A, Nedelkina S, Schalk M, Pallett K, Schaller H, Werck-Reichhart D (2000) Increasing expression of P450 and P450-reductase proteins from monocots in heterologous systems. *Arch Biochem Biophys* 379: 161–169
- Bohlmann J, Keeling CI (2008) Terpenoid biomaterials. *Plant J* 54: 656–669
- Bohlmann J, Lins T, Martin W, Eilert U (1996) Anthranilate synthase from *Ruta graveolens*: duplicated AS alpha genes encode tryptophan-sensitive and tryptophan-insensitive isoenzymes specific to amino acid and alkaloid biosynthesis. *Plant Physiol* 111: 507–514
- Boone CK, Aukema BH, Bohlmann J, Carroll AL, Raffa KF (2011) Efficacy of tree defense physiology varies with bark beetle population density: a basis for positive feedback in eruptive species. *Can J Res* 41: 1174–1188
- Buckingham J (2010) Dictionary of Natural Products. CRC Press Taylor & Francis Group. <http://dnp.chemnetbase.com/dictionary> (August 1, 2011)
- Carelli M, Biazzi E, Panara F, Tava A, Scaramelli L, Porceddu A, Graham N, Odoardi M, Piano E, Arcioni S, et al (2011) *Medicago truncatula* CYP716A12 is a multifunctional oxidase involved in the biosynthesis of hemolytic saponins. *Plant Cell* 23: 3070–3081
- Carlson JE, Tulsieram LK, Glaubitz JC, Luk VWK, Kauffeldt C, Rutledge R (1991) Segregation of random amplified DNA markers in F1 progeny of conifers. *Theor Appl Genet* 83: 194–200
- Carmel L, Rogozin IB, Wolf YI, Koonin EV (2007) Patterns of intron gain and conservation in eukaryotic genes. *BMC Evol Biol* 7: 192
- Cavener DR, Ray SC (1991) Eukaryotic start and stop translation sites. *Nucleic Acids Res* 19: 3185–3192
- Chen F, Tholl D, Bohlmann J, Pichersky E (2011) The family of terpene synthases in plants: a mid-size family of genes for specialized metabolism that is highly diversified throughout the kingdom. *Plant J* 66: 212–229
- Choe S, Dilkes BP, Fujioka S, Takatsuto S, Sakurai A, Feldmann KA (1998) The DWF4 gene of *Arabidopsis* encodes a cytochrome P450 that mediates multiple 22 α -hydroxylation steps in brassinosteroid biosynthesis. *Plant Cell* 10: 231–243
- Christensen AH, Sharrock RA, Quail PH (1992) Maize polyubiquitin genes: structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation. *Plant Mol Biol* 18: 675–689
- de Kraker J-W, Gershenzon J (2011) From amino acid to glucosinolate biosynthesis: protein sequence changes in the evolution of methylthioalkylmalate synthase in *Arabidopsis*. *Plant Cell* 23: 38–53
- Fahn A (1979) Secretory Tissues in Plants. Academic Press, New York
- Franceschi VR, Kreckling T, Christiansen E (2002) Application of methyl jasmonate on *Picea abies* (Pinaceae) stems induces defense-related responses in phloem and xylem. *Am J Bot* 89: 578–586
- Fujita S, Ohnishi T, Watanabe B, Yokota T, Takatsuto S, Fujioka S, Yoshida S, Sakata K, Mizutani M (2006) Arabidopsis CYP90B1 catalyzes the early C-22 hydroxylation of C27, C28 and C29 sterols. *Plant J* 45: 765–774
- Gietz RD, Schiestl RH (2007) Frozen competent yeast cells that can be transformed with high efficiency using the LiAc/SS carrier DNA/PEG method. *Nat Protoc* 2: 1–4
- Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* 52: 696–704
- Hall DE, Robert JA, Keeling CI, Domanski D, Quesada AL, Jancsik S, Kuzyk MA, Hamberger B, Borchers CH, Bohlmann J (2011) An integrated genomic, proteomic and biochemical analysis of (+)-3-carene biosynthesis in Sitka spruce (*Picea sitchensis*) genotypes that are resistant or susceptible to white pine weevil. *Plant J* 65: 936–948
- Hamberger B, Bohlmann J (2006) Cytochrome P450 mono-oxygenases in conifer genomes: discovery of members of the terpenoid oxygenase superfamily in spruce and pine. *Biochem Soc Trans* 34: 1209–1214
- Hartmann T (1996) Diversity and variability of plant secondary metabolism: a mechanistic view. *Entomol Exp Appl* 80: 177–188
- Hayashi K, Kawaide H, Notomi M, Sakigi Y, Matsuo A, Nozaki H (2006) Identification and functional analysis of bifunctional ent-kaurene synthase from the moss *Physcomitrella patens*. *FEBS Lett* 580: 6175–6181
- Helliwell CA, Chandler PM, Poole A, Dennis ES, Peacock WJ (2001) The CYP88A cytochrome P450, ent-kaurenoic acid oxidase, catalyzes three steps of the gibberellin biosynthesis pathway. *Proc Natl Acad Sci USA* 98: 2065–2070
- Helliwell CA, Poole A, Peacock WJ, Dennis ES (1999) Arabidopsis ent-kaurene oxidase catalyzes three steps of gibberellin biosynthesis. *Plant Physiol* 119: 507–510
- Kagawa N, Cao Q, Kusano K (2003) Expression of human aromatase (CYP19) in *Escherichia coli* by N-terminal replacement and induction of cold stress response. *Steroids* 68: 205–209
- Keeling CI, Bohlmann J (2006) Diterpene resin acids in conifers. *Phytochemistry* 67: 2415–2423
- Keeling CI, Dullat HK, Yuen M, Ralph SG, Jancsik S, Bohlmann J (2010) Identification and functional characterization of monofunctional ent-copalyl diphosphate and ent-kaurene synthases in white spruce reveal different patterns for diterpene synthase evolution for primary and secondary metabolism in gymnosperms. *Plant Physiol* 152: 1197–1208
- Keeling CI, Madilao LL, Zerbe P, Dullat HK, Bohlmann J (2011a) The primary diterpene synthase products of *Picea abies* levopimaradiene/abietadiene synthase (PaLAS) are epimers of a thermally unstable diterpenol. *J Biol Chem* 286: 21145–21153
- Keeling CI, Weisshaar S, Lin RP, Bohlmann J (2008) Functional plasticity of paralogous diterpene synthases involved in conifer defense. *Proc Natl Acad Sci USA* 105: 1085–1090
- Keeling CI, Weisshaar S, Ralph SG, Jancsik S, Hamberger B, Dullat HK, Bohlmann J (2011b) Transcriptome mining, functional characterization, and phylogeny of a large terpene synthase gene family in spruce (*Picea* spp.). *BMC Plant Biol* 11: 43
- Klimaszewska K, Rutledge RG, Séguin A (2005) Genetic transformation of conifers utilizing somatic embryogenesis. *Methods Mol Biol* 286: 151–164
- Köksal M, Jin Y, Coates RM, Croteau RB, Christianson DW (2011) Taxadiene synthase structure and evolution of modular architecture in terpene biosynthesis. *Nature* 469: 116–120
- Kolosova N, Miller B, Ralph S, Ellis BE, Douglas C, Ritland K, Bohlmann J (2004) Isolation of high-quality RNA from gymnosperm and angiosperm trees. *Biotechniques* 36: 821–824
- Kopper BJ, Illman BL, Kersten PJ, Klepzig KD, Raffa KF (2005) Effects of diterpene acids on components of a conifer bark beetle-fungal interaction: tolerance by *Ips pini* and sensitivity by its associate *Ophiostoma ips*. *Environ Entomol* 34: 486–493
- Langenheim JH (2003) Plant Resins: Chemistry, Evolution, Ecology and Ethnobotany. Timber Press, Portland, OR
- Levéé V, Major I, Levasseur C, Tremblay L, MacKay J, Séguin A (2009) Expression profiling and functional analysis of *Populus WRKY23* reveals a regulatory role in defense. *New Phytol* 184: 48–70
- Lewinsohn E, Savage T, Gijzen M, Croteau R (1993) Simultaneous analysis of monoterpenes and diterpenoids of conifer oleoresin. *Phytochem Anal* 4: 220–225
- Marín A, Gallardo M, Kato Y, Shirahige K, Gutiérrez G, Ohta K, Aguilera A (2003) Relationship between G+C content, ORF-length and mRNA concentration in *Saccharomyces cerevisiae*. *Yeast* 20: 703–711
- Martin D, Tholl D, Gershenzon J, Bohlmann J (2002) Methyl jasmonate induces traumatic resin ducts, terpenoid resin biosynthesis, and terpenoid accumulation in developing xylem of Norway spruce stems. *Plant Physiol* 129: 1003–1018
- Martin DM, Fäldt J, Bohlmann J (2004) Functional characterization of nine Norway spruce TPS genes and evolution of gymnosperm terpene synthases of the TPS-d subfamily. *Plant Physiol* 135: 1908–1927
- Miller B, Madilao LL, Ralph S, Bohlmann J (2005) Insect-induced conifer defense: white pine weevil and methyl jasmonate induce traumatic resinosis, de novo formed volatile emissions, and accumulation of terpenoid synthase and putative octadecanoid pathway transcripts in Sitka spruce. *Plant Physiol* 137: 369–382
- Mizutani M, Ohta D (1998) Two isoforms of NADPH:cytochrome P450 reductase in *Arabidopsis thaliana*: gene structure, heterologous expres-

- sion in insect cells, and differential regulation. *Plant Physiol* **116**: 357–367
- Nelson D, Werck-Reichhart D (2011) A P450-centric view of plant evolution. *Plant J* **66**: 194–211
- Nelson DR (2009) The cytochrome p450 homepage. *Hum Genomics* **4**: 59–65
- Ober D (2010) Gene duplications and the time thereafter: examples from plant secondary metabolism. *Plant Biol (Stuttg)* **12**: 570–577
- Ohnishi T, Szatmari AM, Watanabe B, Fujita S, Bancos S, Koncz C, Lafos M, Shibata K, Yokota T, Sakata K, et al (2006) C-23 hydroxylation by *Arabidopsis* CYP90C1 and CYP90D1 reveals a novel shortcut in brassinosteroid biosynthesis. *Plant Cell* **18**: 3275–3288
- Omura T, Sato R (1964) The carbon monoxide-binding pigment of liver microsomes. I. Evidence for its hemoprotein nature. *J Biol Chem* **239**: 2370–2378
- Page RD (2002) Visualizing phylogenetic trees using TreeView. *Curr Protoc Bioinformatics Chapter 6*: Unit 6.2
- Peters RJ (2006) Uncovering the complex metabolic network underlying diterpenoid phytoalexin biosynthesis in rice and other cereal crop plants. *Phytochemistry* **67**: 2307–2317
- Peters RJ, Flory JE, Jetter R, Ravn MM, Lee HJ, Coates RM, Croteau RB (2000) Abietadiene synthase from grand fir (*Abies grandis*): characterization and mechanism of action of the “pseudomature” recombinant enzyme. *Biochemistry* **39**: 15592–15602
- Phillips MA, Croteau RB (1999) Resin-based defenses in conifers. *Trends Plant Sci* **4**: 184–190
- Pichersky E, Gang DR (2000) Genetics and biochemistry of secondary metabolites in plants: an evolutionary perspective. *Trends Plant Sci* **5**: 439–445
- Pompon D, Perret A, Bellamine A, Laine R, Gautier JC, Urban P (1995) Genetically engineered yeast cells and their applications. *Toxicol Lett* **82–83**: 815–822
- Powell JS, Raffa KF (1999) Effects of selected *Larix laricina* terpenoids on *Lymantria dispar* (Lepidoptera: Lymantriidae) development and behavior. *Environ Entomol* **28**: 148–154
- Ro DK, Arimura G, Lau SY, Piers E, Bohlmann J (2005) Loblolly pine abietadienol/abietadienal oxidase PtAO (CYP720B1) is a multifunctional, multisubstrate cytochrome P450 monooxygenase. *Proc Natl Acad Sci USA* **102**: 8060–8065
- Ro DK, Bohlmann J (2006) Diterpene resin acid biosynthesis in loblolly pine (*Pinus taeda*): functional characterization of abietadiene/levopimaradiene synthase (PtTPS-LAS) cDNA and subcellular targeting of PtTPS-LAS and abietadienol/abietadienal oxidase (PtAO, CYP720B1). *Phytochemistry* **67**: 1572–1578
- Ro DK, Paradise EM, Ouellet M, Fisher KJ, Newman KL, Ndungu JM, Ho KA, Eachus RA, Ham TS, Kirby J, et al (2006) Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **440**: 940–943
- Robert JA, Madilao LL, White R, Yanchuk A, King J, Bohlmann J (2010) Terpenoid metabolite profiling in Sitka spruce identifies association of dehydroabietic acid, (+)-3-carene and terpinolene with resistance against white pine weevil. *Botany* **88**: 810–820
- Rogozin IB, Sverdlov AV, Babenko VN, Koonin EV (2005) Analysis of evolution of exon-intron structure of eukaryotic genes. *Brief Bioinform* **6**: 118–134
- Schmidt A, Wächter B, Temp U, Krekling T, Séguin A, Gershenzon J (2010) A bifunctional geranyl and geranylgeranyl diphosphate synthase is involved in terpene oleoresin formation in *Picea abies*. *Plant Physiol* **152**: 639–655
- Schmidt HA, von Haeseler A (2007) Maximum-likelihood analysis using TREE-PUZZLE. *Curr Protoc Bioinformatics Chapter 6*: Unit 6.6
- Schuh BA, Benjamin DM (1984) Evaluation of commercial resin acids as feeding deterrents against *Neodiprion dubiosus*, *N. lecontei*, and *N. rugifrons* (Hymenoptera: Diprionidae). *J Econ Entomol* **77**: 802–805
- Shoji T, Hashimoto T (2011) Recruitment of a duplicated primary metabolism gene into the nicotine biosynthesis regulon in tobacco. *Plant J* **67**: 949–959
- Swaminathan S, Morrone D, Wang Q, Fulton DB, Peters RJ (2009) CYP76M7 is an *ent*-cassadiene C11 α -hydroxylase defining a second multifunctional diterpenoid biosynthetic gene cluster in rice. *Plant Cell* **21**: 3315–3325
- Szekeress M, Németh K, Koncz-Kálmán Z, Mathur J, Kauschmann A, Altmann T, Rédei GP, Nagy F, Schell J, Koncz C (1996) Brassinosteroids rescue the deficiency of CYP90, a cytochrome P450, controlling cell elongation and de-etiolation in *Arabidopsis*. *Cell* **85**: 171–182
- Vogel BS, Wildung MR, Vogel G, Croteau R (1996) Abietadiene synthase from grand fir (*Abies grandis*): cDNA isolation, characterization, and bacterial expression of a bifunctional diterpene cyclase involved in resin acid biosynthesis. *J Biol Chem* **271**: 23262–23268
- Wagner A, Donaldson L, Kim H, Phillips L, Flint H, Steward D, Torr K, Koch G, Schmitt U, Ralph J (2009) Suppression of 4-coumarate-CoA ligase in the coniferous gymnosperm *Pinus radiata*. *Plant Physiol* **149**: 370–383
- Wagner A, Ralph J, Akiyama T, Flint H, Phillips L, Torr K, Nanayakkara B, Te Kiri L (2007) Exploring lignification in conifers by silencing hydroxycinnamoyl-CoA:shikimate hydroxycinnamoyltransferase in *Pinus radiata*. *Proc Natl Acad Sci USA* **104**: 11856–11861
- Wagner MR, Benjamin DM, Clancy KM, Schuh BA (1983) Influence of diterpene resin acids on feeding and growth of larch sawfly, *Pristiphora erichsonii* (Hartig). *J Chem Ecol* **9**: 119–127
- Wesley SV, Helliwell CA, Smith NA, Wang MB, Rouse DT, Liu Q, Gooding PS, Singh SP, Abbott D, Stoutjesdijk PA, et al (2001) Construct design for efficient, effective and high-throughput gene silencing in plants. *Plant J* **27**: 581–590
- Yeh FC, Arnott JT (1986) Electrophoretic and morphological differentiation of *Picea sitchensis*, *Picea glauca*, and their hybrids. *Can J Res* **16**: 791–798
- Zhao S, Fernald RD (2005) Comprehensive algorithm for quantitative real-time polymerase chain reaction. *J Comput Biol* **12**: 1047–1064
- Zulak KG, Bohlmann J (2010) Terpenoid biosynthesis and specialized vascular cells of conifer defense. *J Integr Plant Biol* **52**: 86–97
- Zulak KG, Dullat HK, Keeling CI, Lippert D, Bohlmann J (2010) Immunofluorescence localization of levopimaradiene/abietadiene synthase in methyl jasmonate treated stems of Sitka spruce (*Picea sitchensis*) shows activation of diterpenoid biosynthesis in cortical and developing traumatic resin ducts. *Phytochemistry* **71**: 1695–1699
- Zulak KG, Lippert DN, Kuzyk MA, Domanski D, Chou T, Borchers CH, Bohlmann J (2009) Targeted proteomics using selected reaction monitoring reveals the induction of specific terpene synthases in a multi-level study of methyl jasmonate-treated Norway spruce (*Picea abies*). *Plant J* **60**: 1015–1030

Supplemental Text S1 List of Oligonucleotides in 5' to 3' Direction

Cloning CYP720B4 for expression in E. coli

01 GGCATATGGCTCTGTTATTAGCAGTTTTTCACAGGTGGTGAATATCCAGAGAGG nucleotides modified are indicated in italics

02 AAGCTTAGTGATGGTGATGTTCACTCTACTCTACCATGAAGG four histidine codons are indicated in italics

Cloning CYP720B4 and re-formatting for expression in yeast

05 ATGGCTCCAATGGCTGATCAAATTTCTTTGTTATTGGTAGTTTTCACTGTTGCTGTTGCTTTGTTACATTTGATTCATAGATGGTGAATAT

06 AGTTTCACCAATCAATGGCCAACCAGTAGAACCAGGTGGCAAATGAACCTCTTTATTAGACATCTTTGGACCTCTTTGAATATTCCACCATCTATGAATC

07 GGCCATTGATTGGTGAACTTTCTCTTACTATAGATCTATGACTTCCAACCATCCAAGAAAGTTCATTGATGATAGAGAAAAAGATACGATTCTGATA

08 CATTTTGCAAAACAACTTGTTAAATTGTGGATCAGCAGAAACAACAGTTCTACCACCAACAAATGAGAAATGAAAATATCAGAATCGTATCTTTTTTC

09 ACAAGTTTGTTTTGCAAAAATGAAGGTAGATTTTTTCAAGCTCAATATCCAAAAGCTTTGAAGGCTTTGATTGGTAATTATGGTTTGTTATCTGTTTCATGG

10 TCCATGAAATCAACTTTCAATCTTTCAAATCTTAACAAATTAACAGCAATACCATGCAATTTTTCTTTGCAAATCACCATGAACAGATAACAAACCATAAT

11 ATGACATTCATTTTGCAAAGAAATTTCTTTTCATATCAGCCCATCTATCCAAAGTAGAATGAACCAAATTTTGAATTTCTCCATGAAATCAACTTTCAAT

12 ACTAGTAAAAATGGCTCCAATGGCTGATCAAATTTCT enhanced translation initiation leader is indicated in italics

13 ACTAGTTCATTCTCTACTCTACCATGAAGGC

14 CCATCTGATGACATTCAATTTTGCAAAGAAATT

15 CCAGCTGATGACATTCAATTTTGCAAAGAAATT

Cloning Sitka spruce CPR

16 ATGCAGTCTAATTTCGAACACCGTGG

17 TTACCATACATCTCGTAAATATCTTCCTTC

03 CCCCATGGAGTCTAATTCGAACACC

04 CGTCGACTTACCATACATCTCGTAAATATCTTCC

RT-qPCR

18 CATCCGCAAGAACGGCTACATC TIF translation initiation factor (EIF-5A)-like reference gene

19 GTAACATGAGGGACATCGCAG

20 GTTGCTGTAACAAGATGGATGC ELF elongation factor 1a-like reference gene

21 CCCTCAAAACCAGAGATAGGC

22 ATCAAGTTGAGAACCATCTGGTGACATTC CYP720B4

23 CTCGTCTCTTTGGAAGGAGATAAATCC

24 GATCCATGGCGATGGAACTTGATCAGG CYP720B9

25 ATGGTCATCTGCAAGCACATCCCATCTA

26 CTTTGCTGGCCAGACATCCTTCAGA CYP720B8

27 CAGCTGAGTTTGTGATGACCACCTTTC

28 CTATGCAAAGGGGCTTAAGGCGAGGA CYP720B5

29 CTCTAGAAGAAGTATCATGTCCAGCCAAGAGC

30 GCTATTGTGACTATACCAACGCACTTATG CYP720B7

31 GTCAAGTTTATTATGACCGCCCTTGGA

32 CTTAGATTATGCAACCTTGCACCAGGA CYP720B14

33 AAGAGACACCTCAAGTTTAGCTAGATGAAG

34 GGCAACTACGGGGTTCTCTCAGTG CYP720B16

35 TCTAGCCTTAATCCCTTTGCGTAGGTG

36 TTTCTTACTTATTGCCCTGAAGCAGTGG CYP720B15

37 CTTTCACGTAAATGTGTAGCCGCTAAAAC

38 TCAGTTCATGGCGATCTCCACAGG CYP720B17
39 CATCTCAGAAATCTTTCTTATGAGAAGC
40 CTTGAGGAACTGAAGGCTGAGCACGAT CYP720B2
41 GGCATAAACGACGGGTTCTTTGAGAGT
42 TATCAGGAGTTGAAGGCTGAGCATGAG CYP720B12
43 CATAAACGAGGGGTTCTTTGAGGGC
44 GGGAGCTGAGGGCGGAACACAAC CYP720B10
45 TCCAGATGGATGCCGTCATAAGCACCAAG

Cloning Sitka spruce FL genomic and cDNA sequences

46 GATGGAAGACGTTGAAGCAATGTTAAGG CYP701A24
47 TTAAGTGTTGGACATGGAGATATTGCTAGG
48 GGACACTTGTAAGGATGCTGAGTTTGG CYP88A28
49 GGTGTTGGCCTCATAACTTTCTGATCCT
50 GGAATAGCAATCATGGCGGAAACG CYP720B2
51 CTAGAGAGATTACTCGTCCTGTCC
52 CGCCAATGGCAGATCAAATATCATTAG CYP720B17
53 GCATGTCTTCATTGTCTACAATGAAGACG
54 CAACAATGGCAGACCAAATATCATTAC CYP720B4
55 TTATTCATTCTCTACTCTACCATGAAGGCT
56 AGTGGCGCTTGTTACCTTATCTACAGAG CYP720B8
57 GCCCTTGGTCAAGTGAGGTAATGGAAGATATG
58 TGCACGAAGAAGAATGCATGGTCTCC CYP90A20
59 TCATGTAAATTATATACATTCCTTCGAG

Supplemental Table S Ia: Identification of isopimaradiene, abietadiene and the corresponding resin acids by GC-MS

Yeast Strain	<i>In vivo</i> Product Authentic standard	Ret. Time [min]	Characteristic ions m/z [% rel. int.]
GGPPS/ISO/CPR	Isopimaradiene	36.7	272.0 (100), 257.0 (43), 229.0 (78)
		36.7	272.0 (100), 257.0 (47), 229.0 (82)
GGPPS/ISO/CPR/CYP720B4	Methyl isopimarate	41.25	316.0 (39), 257.0 (57), 241.0 (100)
		41.25	316.0 (44), 257.0 (60), 241.0 (100)
GGPPS/LAS/CPR	Abietadiene	37.2	272.0 (21), 257.0 (7), 135.0 (100)
		37.2	272.0 (25), 257.0 (15), 135.0 (100)
GGPPS/LAS/CPR/CYP720B4	Methyl abietate	43.5	316.0 (78), 256.0 (100), 241.0 (70)
		43.5	316.0 (85), 265.0 (100), 241.0 (72)

Supplemental Table S Ib: Identification of diterpenoids by LC-MS

Assay Substrate	Assay Product Authentic Standard	Ret. Time [min]	Characteristic ions m/z [% rel. int.]
Abietadiene	Abietadienol	10	289.0 (100), 271.0 (29.4)
		10	289 (100), 271.0 (8)
	Abietadienal	15.5	trace
		15.5	287.0(100), 269.0 (4), 243.0 (2.2)
	Abietic acid	7.8	303.0 (100), 123.0(10)
Isopimaradiene	Isopimaradienol	7.8	303.0 (100), 123.0 (8)
		11.5	289.0 (18), 271.0 (100), 174.9 (16)
	Isopimaradienal	11.8	289.0 (27), 271.0 (100), 174.9 (17)
		18	287.0(100), 269.0 (29), 242.9 (22)

		17.9	287.0(100), 269.0 (27), 242.9 (32)
	Isopimaric acid	7.9	303.0 (37), 257.0 (100)
		7.8	303.0 (46), 257.0 (100)
Sandaracopimaradiene	Sandaracopimaradienol	11.3	289.0 (7), 271.0 (100)
		11.4	289 (12), 271 (100), 174.9 (12)
	Sandaracopimaradienal	17.5	287.0(100), 269.0 (51), 242.9 (14)
		17	287.0(100), 269.0 (43), 242.9 (17)
	Sandaracopimaric acid	7.8	302.9 (8), 256.9 (100),
		7.9	303.0 (9), 257.0 (100)
Pimaradiene	Pimaradienol	10.2	289.0 (4), 271.0 (100)
		10.3	289.0 (4), 271.0 (100)
	Pimaradienal	17.4	287.0(100), 269.0 (39), 242.9 (14)
		17.6	287.0(100), 268.9 (48), 242.9 (20)
	Pimaric acid	7.8	303.0 (15), 257.0 (100), 174.9 (5)
		7.6	303.0 (13), 257.0 (100), 174.9 (6)
Neoabietadiene	Neoabietadienol	10.5	289.0 (100), 271.0 (75)
		10.5	289.0 (100), 271.0 (65)
	Neoabietadienal	16.5	287.0(100), 268.9 (3), 242.9 (2)
		16.4	287.0(100), 269.0 (14), 242.9 (13)
	Neoabietic acid	8	303.0 (100), 257.0 (12), 123.0 (7)
		8	303.0 (100), 257.0 (58), 123.0 (5)
Palustradiene	Palustradienol	10.7	289.0 (100), 271.0 (49)
		10.7	289.0 (100), 271.0 (23)
	Palustradienal	17.2	287.0 (100), 268.9 (2)
		16.9	287.0(100), 269.0 (4)
	Palustric acid	7.7	302.9 (100), 257.0 (38)
		7.5	303.0 (100), 257.0 (26)
Dehydroabietadiene	Dehydroabietanol	6.5	287.0 (3), 268.9 (100), 172.9 (90)
		6.5	287.0 (4), 268.9 (100), 172.9 (90)
	Dehydroabietanal	9.5	284.9 (45), 267.0 (33), 240.9 (100)
		9.5	284.9 (70), 267.0 (68), 240.9 (100)
	Dehydroabietic acid	5.1	300.9 (20), 257.0 (89), 172.9 (100)
		5.1	300.9 (15), 254.9 (84), 172.9 (100)
Levpimaradiene	Levpimaradienol	10.8	289.0 (100), 271 (91)

	10.4	289.0 (100), 271 (20)
Levopimaradienal	15.8	287.0(100), 269.0 (12)
	15.7	287.0 (100), 268.9 (17)
Levopimaric acid	8	303.0 (100), 254.0 (32)
	8	303.0 (100), 256.9 (19)

Supplemental Table S II: Co-suppression of transcripts of CYP720B family members in white spruce lines transformed for CYP720B4 RNAi

Family member	Transcript accumulation (% WT)	Nucleotide identity to target CYP720B4 RNAi sequence	21-mers with match to off-target, 20 out of 21 (21 out of 21)
CYP720B10 (class I)	* transcript detected in WT too low for quantification independent of RNAi	89%	92 (46)
CYP720B12 (class I)	32.1 ± 20.1	75%	14 (4)
CYP720B2 (class I)	38.0 ± 27.5	76%	17 (13)
CYP720B16 (class II)	14.4 ± 9.6	89%	92 (46)
CYP720B17 (class II)	* unspecific RT-qPCR in WT	89%	104 (43)
CYP720B15 (class II)	15.7 ± 6.7	89%	87 (45)
CYP720B14 (class II)	* transcript detected in WT too low for quantification independent of RNAi	**n.a. as sequence not FL	n.a.
CYP720B7 (class III)	18.5 ± 9.4	94%	163 (104)
CYP720B5 (class III)	20.5 ± 7.2	94%	195 (72)
CYP720B4 (class III)	***5.9 ± 1.0	100%	269 (269)
CYP720B8 (class IV)	* transcript not detected in WT independent of RNAi	82%	44 (22)
CYP720B9 (class IV)	* transcript not detected in WT independent of RNAi	78%	34 (27)

* These results are consistent with the constitutive low abundance of the orthologous gene transcripts in Sitka spruce stems (see Figure 3).

** n.a.: Could not be analyzed, as the sequence was not full length.

*** Please note that this number (5.9 % residual transcript level) is lower than the number (25% residual transcript level) obtained in a separate experiment shown in Figure 9A.

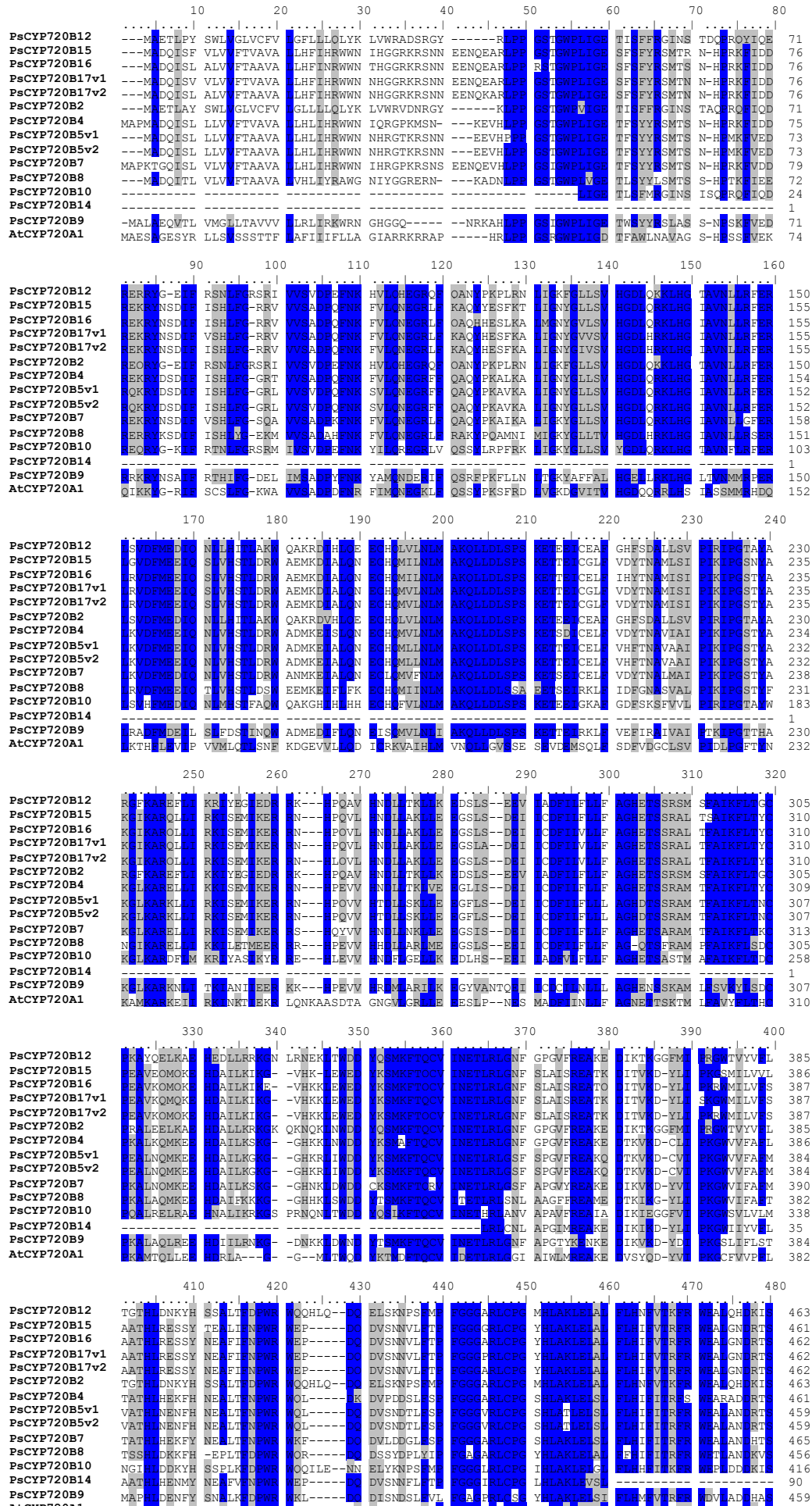
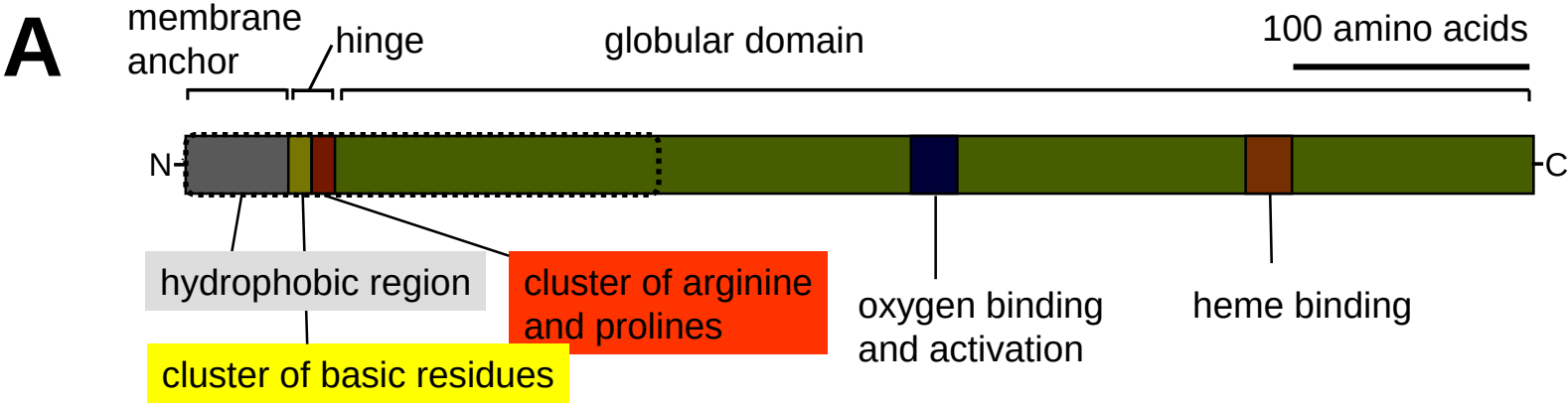


Figure S1: Multiple sequence alignment of Sitka spruce PsCYP720B and Arabidopsis AtCYP720A. Manually curated alignment of amino acid sequences generated by dialign¹ and visualised using bioedit². Highlighted in blue 80% identical positions, in grey similar amino acids (Blosum62). ¹B. Morgenstern (2004). DIALIGN: Multiple DNA and Protein Sequence Alignment at BiBiServ. Nucleic Acids Research 32, W33-W36. ²Hall, T.A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl. Acids. Symp. Ser. 41:95-98



B

M A P M A D Q I S L L L V V F T V A V A
ATGGCGCCCATGGCAGACCAAATATCATTACTGTTGGTGGTGTTCACGGTAGCGGTGGCG
ATGGCTCCAAATGGCTGATCAAATTTCTTTGTTATTGGTAGTTTTCACGTGTGCTGTGCT
L L H L I H R W W N I Q R G P K M S N K
CTCCTCCACCTTATTCACAGGTGGTGGGAATATCAGAGAGGCCCAAAATGAGTAATAAG
TTGTTACATTTGATTCA TAGATGGTGGGAATATCAAAGAGGTCCAAAGATGCTAATAAA
E V H L P P G S T G W P L I G E T F S Y
GAGGTTTCATCTGCCCTCTGGGTCGACTGGATGGCCGCTTATTGGCGAAACCTTCAGTTAT
GAAGTTTCATTTGCCACCTGGTTCTACTGGTTGGCCATTGATTGGTGAAACCTTCCTTAC
Y R S M T S N H P R K F I D D R E K R Y
TATCGCTCCATGACAGCAAATCATCCAGGAAATTCATCGACGACAGAGAGAAAAGATA
TATAGATCTATGACTTCCAAACCATCCAGAAAGTTTCATGATGATAGAGAAAAAGATAC
D S D I F I S H L F G G R T V V S A D P
GATTCGGAATTTTCATATCTCATCTATTGGAGGCCGGACGGTGTGATCAGCGGATCCC
GATTCGATATTTTCATTTCTCATTTGTTGGTGGTAGAACGTGTGTGTTCTGCTGATCCA
Q F N K F V L Q N E G R F F Q A Q Y P K
CAGTTCAACAAGTTTGTTCACAAAACGAGGGGAGATTCTTTCAAGCCCAATACCAAAG
CAATTTAACAAGTTTGTTCGCAAAATGAGGAGTAGATTCTTTCAAGCTCAATATCCAAA
A L K A L I G N Y G L L S V H G D L Q R
GCATGAAGGCTTTGATAGGCAACTACGGGCTGCTCTCTGTGCATGGAGATCTCCAGAGA
GCTTTGAAGGCTTTGATTGGTAAATATGTTGTTATCTGTTCATGGTGATTTGCAAGA
K L H G I A V N L L R F E R L K V D F M
AAGCTCCACGGGAATAGCTGTGAATTTGCTGAGGTTTGAGAGACTGAAAGTCGATTTTCATG
AAATTTGCAATGGTATGCTGTAAATTTGTTAAGATTTGAAGATTGAAAGTTGATTTTCATG
E E I Q N L V H S T L D R W A D M K E I
GAGGAGATACAGAATCTCGTGCACTCCACGTTGGATAGATGGGCAGATATGAAGGAAATT
GAAGAAATCAAAATTTGGTTCAATTCATCTTTGGATAGATGGGCTGATATGAAAGAAATT
S L Q N E C H Q M V L N L M A K Q L L D
TCTCTGCAGAAATGAATGTCAACAGATGGTTCTCAACTTGATGGCCAAACAACCTGCTGGAT
TCTTTGCAAAATGAATGTCAACAGATGGTTCTCAACTTGATGGCCAAACAACCTGCTGGAT

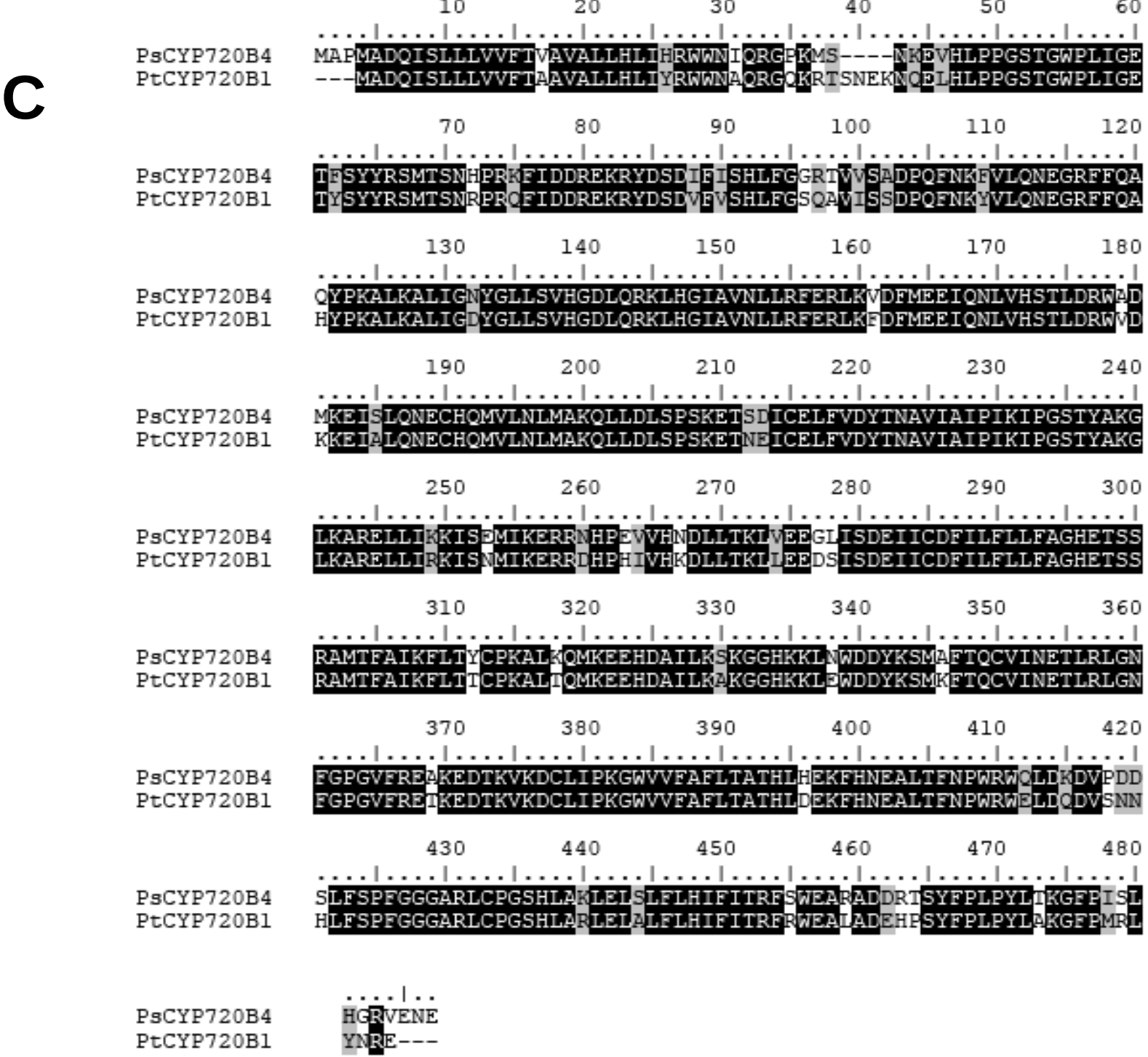
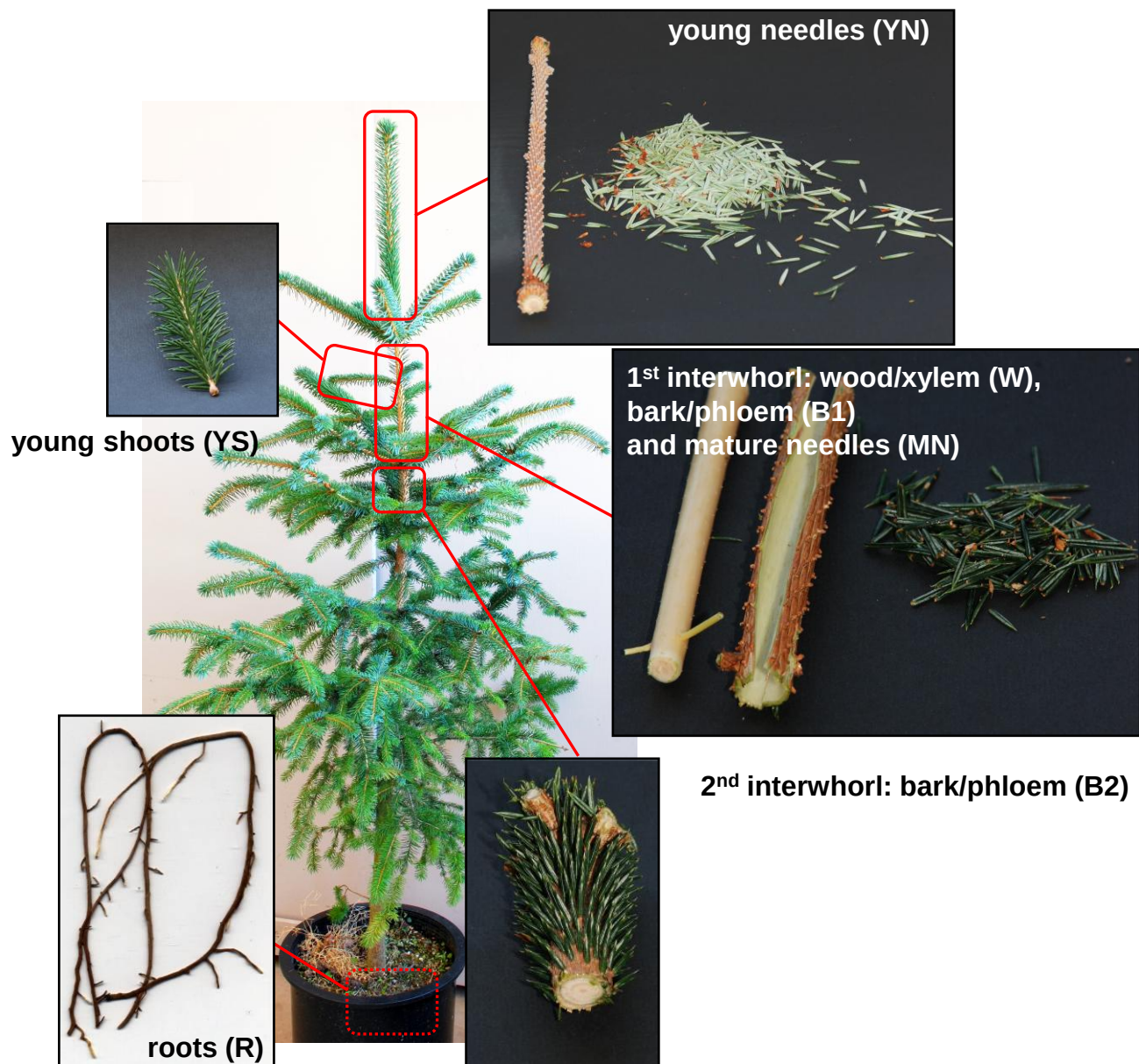


Figure S2: Optimization of PsCYP720B4 for expression in yeast. (A) Primary structure of the PsCYP720B4 protein. 600 bp encoding the N-terminal hydrophobic membrane anchor and the cluster of basic residues and arginines and prolines of the hinge region were optimized for heterologous expression in yeast, indicated by the dashed box. The oxygen binding domain and the heme binding domain are indicated in the globular domain of PsCYP720B4. The scale represents 100 amino acids. (B) High occurrence of rare yeast codons encoding amino acids of the basic cluster and clustering arginine and proline were adjusted together with the GC-content in the 600 bp 5'-coding region of CYP720B4 (the native PsCYP720B4 in the first, the optimized PsCYP720B4 in the second lane). (C) Amino acid alignment of PtCYP720B1 and PsCYP720B4. Highlighted in grey are 80% similar amino acids (Blosum80).



5-year old Sitka spruce trees of clonal line FB3-425, average height 150cm.

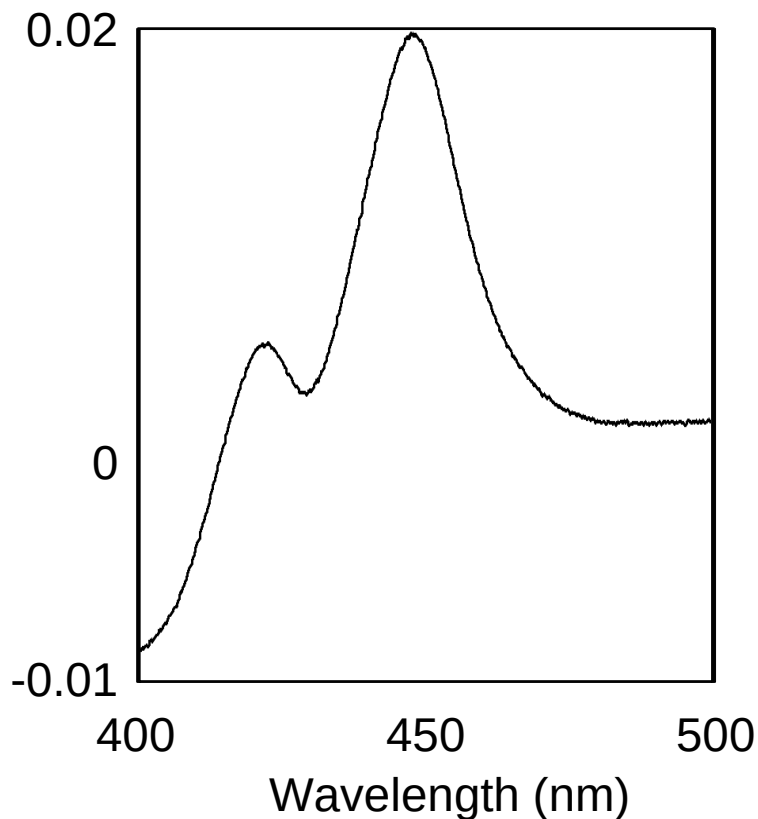
Figure S3: Organs and tissues of Sitka spruce used for transcript and metabolite profiling. YN, young needles; MN, mature needles; YS, young shoots; W, wood and xylem of the first interwhorl; B1, bark and phloem of the first interwhorl; B2, bark and phloem of the second interwhorl; R, roots.

A)

	1	11	21
CYP720B4 (FL cDNA)	M	A	P
CYP720B4 (modified)	M	A	~

CYP720B4 (FL cDNA) MAPMADQISLLLVVFTVAVALLHLLIHRW
CYP720B4 (modified) MA~~~~~LLLVFTVAVALLHLLIHRW

B)



82 nmol P450/L

Figure S4: Heterologous expression of CYP720B4 in *E.coli*. (A) Sequence of the N-terminus of PsCYP720B4 modified for expression in *E. coli*. (B) Carbon monoxide differential spectrum of PsCYP720B4 in *E. coli* membrane fraction.

