Running title: Cytochromes P450 of Diterpenoid Conifer Defence

Author to whom correspondence should be sent:

Jörg Bohlmann
Michael Smith Laboratories
University of British Columbia
2185 East Mall
Vancouver, BC, V6T 1Z4, Canada
Phone: 1 604 822 0282
e-mail: (bohlmann@msl.ubc.ca)

Research Area: Biochemical Processes and Macromolecular Structures
Evolution of diterpene metabolism: Sitka spruce CYP720B4 catalyses multiple oxidations in resin acid biosynthesis of conifer defense against insects

Björn Hamberger²,³, Toshiyuki Ohnishi²,⁴, Britta Hamberger³, Armand Séguin and Jörg Bohlmann⁵

Michael Smith Laboratories, University of British Columbia, 2185 East Mall, Vancouver, BC, V6T 1Z4, Canada (BH, TO, BrH, JB); Canadian Forest Service, Laurentian Forestry Centre, Quebec City, Quebec, G1V 4C7, Canada (AS)
Footnotes

1 The work described in this paper was supported with funding to JB from the Natural Sciences and Engineering Research Council of Canada (EWR Steacie Memorial Fellowship and Discovery Grant); Genome British Columbia and Genome Canada for the Treenomix Conifer Forest Health Project (www.treenomix.ca). JB is supported in part by the University of British Columbia Distinguished University Scholars program.

2 These authors contributed equally to this work.

3 Present address: Department of Plant Biology & 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

5 Corresponding author: bohlmann@msl.ubc.ca
Abstract
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 multi-substrate, multi-functional enzyme by: Formation of oxygenated diterpenoids in metabolically engineered yeast, yeast in vivo transformation of diterpene substrates, in vitro assays with CYP720B4 protein produced in E. coli, and alteration of DRA profiles in RNAi 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 reveal differences in the evolution of specialized and general diterpene metabolism in a gymnosperm. The genomic and functional characterization of the gymnosperm CYP720B family highlight that evolution of specialized metabolism involves substantial diversification relative to conserved, general metabolism.
Introduction

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, online Dictionary of Natural Products 18.2). Diterpene resin acids (DRAs), together with mono- 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 which 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 Ophiostoma ips, a tree pathogenic fungal symbiont 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), 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 large-volume, renewable resource for industrial bio-products (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 increased production and accumulation of oleoresin (Zulak and Bohlmann, 2010). Formation of traumatic resin ducts and 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). 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 (Figure 1A). These diterpenoids are structurally similar to the tetracyclic ent-kaurane diterpenes, which include the ubiquitous gibberellin (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 (dTTPSs) and cytochrome P450 monooxygenases (P450s) into diterpene acids, i.e. ent-kaurenoic acid or DRAs. The dTTPSs of conifer DRA biosynthesis yield the tricyclic core structures (Keeling et al., 2008; Keeling et al., 2011a). Oxidation at the C-18 carbon of ring A by P450 activity leads to the corresponding diterpene alcohols, aldehydes, and acids (Figure 1A). The conifer dTTPS genes of specialized and general diterpenoid metabolism share common evolutionary ancestry (Keeling et al., 2010).

While dTTPS for DRA biosynthesis have been cloned and characterized in several species of the pine family (Stofer 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 into 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).
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 specialized DRA metabolism of conifers appears to be much larger compared to general ent-kaurenoic acid biosynthesis. In general, various pathways of specialized metabolism in plants evolved from conserved elements of general metabolism (e.g., Bohlmann et al., 1996; Pichersky and Gang, 2000; Ober 2010; de Kraker and Gershenzon, 2011; Shoji et al., 2011). Diversification of specialized metabolism can result from events of gene duplication, retention of multiple gene copies, as well as sub- 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; Keeling et al., 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 and 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 FLcDNAs members of the CYP720B family in Sitka spruce, we identified PsCYP720B4 with a role in DRA biosynthesis. The biochemical properties of the multi-substrate and multi-functional PsCYP720B4 enzyme highlight the molecular underpinnings of diversification of DRA biosynthesis in 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 the present work in a gymnosperm system are relevant for the broader understanding of specialized diterpene metabolism and the evolution of specialized metabolism in general.
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 non-oxidized 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 (Figure 1A; see figure legend for details) in the bark and phloem and in the wood and xylem of stems of five-year old Sitka spruce (Figure 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 (Figure 1B). Of the total amount of diterpenoids of 48.98 ± 4.30 μg mg⁻¹ stem DW in the outer stem tissues (Figure 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 ± 0.11 μg mg⁻¹ DW) 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 (Figure 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 analysis, 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, abietates; Figure 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. Pimarric 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 (Figure 1B). Therefore, the question arises as to how many of the possible 24 oxidations (8 homologous series of olefins, alcohols, aldehydes, and acids) involved in specialized DRA biosynthesis can be attributed to single or multiple P450 enzyme(s). To generate a comprehensive inventory of the CYP720B subfamily, we searched nearly one million spruce and pine ESTs and constructed a maximum likelihood phylogeny of the predicted CYP720B protein sequences from Sitka spruce, white spruce (P. glauca), interior spruce (P. glauca x engelmannii), loblolly pine, and lodgepole pine (P. contorta), rooted with Arabidopsis P450 CYP720A1 (At1g73340) as an outgroup (Figure 2). We identified 39 CYP720B members that form four distinct phylogenetic clades I - 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 orthologues 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 Figure 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 orthologues in the distantly related angiosperms (e.g. AtCYP85A2). These findings suggest an increased rate of sequence divergence within the CYP720B subfamily of specialized metabolism.

Full length (FL) cDNAs of Sitka spruce PsCYP720Bs. We cloned a representative set of 12 different Sitka spruce PsCYP720B members as FLcDNAs (Figure 2B). The translated proteins were predicted to cover the entire ORFs according to presence of starting methionines with stop codons upstream within the 5’ UTRs and patterns of ORF versus UTR sequence similarity within the CYP720B subfamily. We identified two pairs of putative allelic variants (≥98% amino acid identity) among the 12 PsCYP720B FLcDNA sequences (PsCYP720B5v1 and PsCYP720B5v2; PsCYP720B17v1 and PsCYP720B17v2, Figure 2B and Supplemental Figure 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 Figure S2 panel C).

Transcript profiles of PsCYP720Bs reveal differential gene expression across different organs and tissues. To assess correlations of PsCYP720B transcripts with oleoresin accumulation in Sitka spruce stems, we performed comparative and quantitative transcript analysis using RT-qPCR for the set of 12 distinct PsCYP720Bs across a range of tissues (Supplemental Figure S3): wood/xylem of the stem section of the first interwhorl from the top the tree (W), bark/phloem of the first (B1) and second (B2) interwhorl 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 - IV (Figure 2).
All PsCYP720Bs of clades I - 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, B2). These stem tissues accumulate substantial amounts of diterpene resin acids (Figure 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, MN) and young shoots which included needles (Supplemental Figure 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 below-ground tissues.

Transcript levels of PsCYP720B4 were the highest in B1 and B2, followed by YS which 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, 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, -7) showed MeJA-induced increase of transcript accumulation in bark tissue eight days after treatment of trees (Figure 4). This time point was chosen for correlation with 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 eight 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 (CRD) are the primary side 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, co-localize with diTPS transcripts to CRD epithelial cells (Figure 5). For comparative spatial analysis, we measured transcript levels of individual PsCYP720Bs in entire stem cross-sections dominated by the woody xylem (XSEC), in the cambial zone (CZ), and 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 XSEC (Figure 5). Low levels of PsLAS in CZ may be attributed to radial resin ducts. Conversely, PsCYP73A20 transcripts were most abundant in XSEC and CZ, but were detected at low levels in CRDs.

Matching the CRD localization of PsLAS transcripts, several PsCYP720Bs (PsCYP720B2, -4, -12,-15, -16, -17) showed highly enriched transcript levels in CRDs with much lower transcript levels in XSEC and CZ (Figure 5). Among clade III genes, transcripts of PsCYP720B4 were most highly enriched in CRDs. Consistent with results from transcript analysis in various samples across the entire plant (Figure 3), transcript levels of PsCYP720B10 and -14 were low or not detected in the laser microdissected stem tissues. Also consistent with lack of transcripts of PsCYP720B8 and PsCYP720B9 in complex stem samples (Figure 3), these transcripts were not detected in laser microdissected CZ or CRDs.
Functional characterisation of PsCYP720B4 using in vivo assays of diterpene resin acid formation in metabolically engineered yeast.

The combined patterns of gene expression profiles (Figures 3 - 5) and phylogenetic proximity with PtCYP720B1 (Figure 2) strongly support a role of PsCYP720B4 in DRA biosynthesis in Sitka spruce. We therefore selected PsCYP720B4 for a first comprehensive functional characterisation of a CYP720B enzyme, building in part on methods previously established for the initial characterisation 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 (ScGGPPS) 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\(^{-1}\) (Figure 6). The yeast strain expressing ScGGDPS/PaISO/AtCPR was newly developed for this study and produced isopimaradiene with yields of up to 1.6 mg L\(^{-1}\) (Figure 6). Although PaLAS can form multiple diterpene products in vitro (Keeling et al., 2008), only abietadiene accumulated in the yeast strain expressing ScGGDPS/PaLAS/AtCPR. Diterpenes were detected at approximately 80-fold higher levels in the yeast cell pellet compared with culture media suggesting either 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 Figure S2 panel B). When PsCYP720B4 was co-expressed with ScGGDPS/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 (Figure 6A). When PsCYP720B4 was co-expressed with ScGGDPS/PaLAS/AtCPR, levels of abietadiene were reduced and abietic acid accumulated up to 0.2 mg/L (Figure 6B). No oxidized diterpenoids were detected in strains lacking the PsCYP720B4 cDNA. None of the diterpene alcohol or aldehyde
intermediates of the conversion of isopimaradiene to isopimaric acid or abietadiene to
abietic acid accumulated to detectable levels. These results showed convincingly that
PsCYP720B4 expressed in yeast catalyses 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 substrate preference of CYP720B4 for isopimaradiene over
abietadiene. To further assess substrate preferences, we developed additional assay
systems for PsCYP720B4 expressed in yeast and *E. coli*.

Functional characterisation of PsCYP720B4 using *in vitro* assays with protein in *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, respectively abietadiene and isopimaradiene.
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 8 different diterpene skeletons
(abietadiene, palustradiene, levopimaradiene, neoabietadiene, pimaradiene,
isopimaradiene, sanadaracopimaradiene, dehydroabietadiene; Figure 1A). These
compounds represented all of the diterpenoid resin acids and intermediates detected in
Sitka spruce stem tissues (Figure 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 PsCYP720B4 protein produced in *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 presence of P450 protein by
SDS-PAGE analysis and CO difference spectroscopy (Supplemental Figure 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 either
dehydroabietadiene, dehydroabietadienol or dehydroabietadienal as substrate. All three substrates were converted by PsCPR/PsCYP720B4 (Figure 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 (Figure 7). When dehydroabietadienol was used as 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 for the first time with a CYP720B enzyme in 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 PsCYP720B4 protein in E. coli membrane fraction, to further substantiate the emerging pattern of PsCYP720B4 as a multi-substrate and multi-functional (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 (Figure 8). Assays with E. coli produced PsCYP720B4 protein were performed with limiting concentrations of PsCPR to restrict 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 (Figure 8A). Results of the in vitro assays and LC-MS analysis showed that the PsCYP720B4 enzyme is active with each of the 24 different substrates tested, confirming the multi-substrate and multi-functional 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 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 μM and the formation of the corresponding diterpene
resin acid was measured by GC-MS (Figure 8B and Supplemental Figure S5). With the
exception of two olefin substrates (neoabietadiene 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 to the
conversion of the olefins (Figure 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 (Figure 8A). This result may be due to
reduced olefin solubility or up-take into yeast cells, or it may reflect a lower rate of the
three-step oxidation of the olefins compared to the one- or two-step oxidations of the
aldehydes or alcohols, respectively.

In both types of assays, PsCYP720B4 consistently showed the highest activity
with the dehydroabietadienolate group of substrates (dehydroabietadiene,
dehydroabietadienol, and dehydroabietadienal) compared to the other seven sets of
diterpenoids tested (Figure 8). The dehydroabietadienolates are the only diterpenoids of the
Sitka spruce oleoresin with an aromatic C ring (Figure 1). Notably, the
dehydroabietadienolates were of low abundance in the diterpene mixture of the particular
clonal line of Sitka spruce investigated in this study (Figure 1), while across a range of
more than 100 genotypes of Sitka spruce dehydroabietic acid was associated with 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
multi-substrate and multi-functional 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 CO-difference assays (Supplemental Figure 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 (Figure 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 (Figure 8A and B). 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 E. coli membrane fraction with reactions restricted to a single oxidation step by limitation of PsCPR 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 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 to diterpene olefins. Thus, the apparent catalytic efficiencies ($k_{cat}/K_m$) of PsCYP720B4 for dehydroabietadienol were 1.5 and 3 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 (Figure 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 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 (Armand Séguin, unpublished results) and has not yet been reported to be successful in any other conifer except for radiata pine (P. radiata) (Wagner et al., 2007;...
Wagner et al., 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 one-year old RNAi lines (five biological replicates for each of four RNAi line) showed a significant decrease of PgCYP720B4 transcripts to as low as 25% relative to transcripts in stems of control trees transformed with empty pCambia vector (Figure 9; Supplemental Table SII).

We also tested the effect of cross-reactivity and co-suppression 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 the PgCYP720B4 target, but also some considerable suppression with other members of the clades III, II, and I in decreasing order of effect (Supplemental Table SII)

RNAi-mediated reduction of PgCYP720B transcript levels was paralleled by significant (independent T-test, p<0.05) change in the amounts of several (but not all) types of DRA (Figure 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 to 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 levopimaradienate was detected in any of the four RNAi lines. Pimaric acid was not detected in the white spruce samples.

Reduced levels of several DRA in the RNAi lines supports a role of CYP720B4, and perhaps other CYP720B genes, in DRA biosynthesis in planta. The strong effect on the accumulation of dehydroabietic acid, compared to the other diterpene resin acids, is in agreement with the results from enzyme assays and kinetic characterization of PsCYP720B4 which showed preferred enzyme activity in the formation of
dehydroabietic acid. The lack of obvious effect of PgCYP720B4 RNAi on several other
diterpene resin acids 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 multi-substrate nature of the PsCYP720B4
enzyme and may also be due to co-suppression 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 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 catalyze the oxidation of ent-kaurenoic acid to GA$^{12}$ (Helliwell
et al., 2001), and to CYP90, which catalyze C-22 and C-23 hydroxylations in the
biosynthesis of brassinosteroid hormones (Szekeres et al., 1996; Choe et al., 1998;
Ohnishi et al., 2006) (Figure 2). CYP701, CYP88 and CYP90 genes had not previously
been cloned from a gymnosperm. To validate 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 clade I -
IV, as well as for Sitka spruce PsCYP701, PsCYP88, and PsCYP90 (Figure 10). Shared
position of introns indicates 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 position of the
introns were highly conserved across the PsCYP720B, PsCYP88 and PsCYP90 genes
(Figure 10). In contrast, the PsCYP701 gene is organized in eight exons, and the size 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, \textit{i.e.}, DRAs or \textit{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 function in a context of diversification of conifer specialized metabolisms, while CYP701 functions in a context of a highly conserved general metabolism.

\textbf{Discussion}

Using several lines of evidence, we identified CYP720B4 as a multi-substrate and multi-functional P450 enzyme with an important role for structural diversity of specialized diterpenoid metabolism. A major product of CYP720B4 is dehydroabietic acid which is associated with resistance of Sitka spruce against its major insect pest, the spruce weevil (Robert et al., 2010). Beyond the general relevance of the present work for fundamental understanding of the evolution of specialized (\textit{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 biosynthesis and evolution of specialized and general (\textit{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 \textit{(Oryza sativa)} system (Peters, 2006; Swaminathan et al., 2009). Specialized diterpenoid metabolism of DRAs and the biosynthesis of \textit{ent}-kaurenoic acid, the ubiquitous precursor for GAs of general diterpene metabolism, are remarkably similar (Figure 11). These pathways share structurally similar diterpene olefins, alcohols, aldehydes and acids derived by 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 (Figure 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 (Figures 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-CPP synthase and ent-kaurene synthase, are monofunctional enzymes (Figure 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., 2010). On the genome level, conifer diTPSs of general metabolism seem to be encoded by single copy genes which 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 MYA 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; Keeling et al., 2011b), that set
them apart from diTPSs of general metabolism (Keeling et al., 2010), are retention and
neo-functionalization of duplicated genes which give rise to the initial diversification of
specialized metabolism. The products of the diTPSs of specialized metabolism are the
substrates for P450 enzymes of the conifer-specific CYP720B family (Figure 11), of
which PsCYP720B4 is the first member which has been characterized in detail with 24
different substrates as a multi-functional and multi-substrate enzyme. PsCYP720B4
substantially increases the spectrum of diterpenoids in specialized DRA metabolism.

While the recently discovered CYP720B genes exist with multi-gene families in
conifers (Figure 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 et al., 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 orthologues in the angiosperms. In contrast the conifer diTPSs and CYP720B
P450s of specialized metabolism do not have apparent orthologues in the angiosperms.

Apparently, the specific function of three-step consecutive oxidation at 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 biosynthesis of artemisinic acid (Ro et al., 2006) or
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 (Figure 1 and Figure 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 awaiting 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; Hartmann, 2007) in which (i) specialized and general metabolism share common origins, (ii) specialized metabolism evolved (and continues to evolve) with patterns of taxon-specific pathway diversification on the basis of gene duplication and neo-functionalization, while (iii) 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 multi-gene 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, five-year old) and white spruce (*P. glauca*, clone PG653, one-year old) were maintained as previously described (Miller *et al.*, 2005). Treatment with methyl jasmonate (0.1% in
0.1% Tween20 v/v) was as previously described (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, periderm), xylem/wood, mature and young needles (Supplemental Figure 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 (USA). Yeast strain WAT11 (MAT a; ade2-1; his3-11,-15; leu2-3,-112; ura3-1; canR; cyr+) carrying the Arabidopsis thaliana cytochrome P450 reductase (AtCPR) chromosomally integrated under control of the GAL promoter (Pompon et al., 1995) was provided by Dr. Werck-Reichhart (Institut de Biologie Moléculaire des Plantes, Centre National de la Recherche Scientifique, Strasbourg, France) and used as host for heterologous gene expression. The pCWori+ vector for heterologous expression of CYP720B4 in E.coli was provided by Dr. Mizutani (Graduate School of Agricultural Science, Kobe University, Japan) (Fujita et al., 2006).

Diterpene standards and substrates. Diterpene resin acids were purchased from Orchid Cellmark (New Westminster, BC, Canada). Diterpenoid olefins, alcohols and aldehydes were synthesized from the corresponding acids by Best West Laboratories (Salt Lake City, Utah, USA) following published methods (Ro et al., 2005). Purity of synthesized diterpenoids was assessed by 1H-NMR, 13C-NMR and GC-MS, and was equal or exceeded 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 (LMD). Details of sample preparation and LMD procedures are described in 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, Germany) and laser microdissected (DM 6000B/CTR 6500, Leica, Germany).
Extraction of diterpenoids. Methods for extraction and derivatisation 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, USA). Tissue (50-100 mg) was extracted by shaking overnight in 1.5 ml tert-butylmethylether with 200 μg ml⁻¹ dichlorodehydroabietic acid and 100 μg ml⁻¹ isobutylbenzene as internal standards. Extracted tissue was dried at 50°C for two days prior to weighing to determine the dry weight (DW). For derivatisation, 500 μl of the organic supernatant was mixed with 100 μl methanol and 120 μl trimethylsilyl-diazomethane (2.0 M in diethyl ether, Sigma-Aldrich) and incubated at room temperature for 20 min. Derivatised samples were concentrated under N₂ gas and adjusted to 500 μl with diethylether before GC-MS analysis.

Analysis of metabolites by GC-MS and LC-MS. Due to the complexity of 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 of 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 (30m x 0.25mm x 0.25μm, SGE) in SIM-scan mode (scan: m/z 40-400; SIM: m/z 270, 272, 286, 288, 314 and 316 (dwell time 50)). Volumes of 0.5μl sample were injected in pulsed splitless mode at 250°C with a column flow of 1 ml/min helium and 20 psi pulse pressure. The GC oven temperature was programmed to rise from initial 40°C at 3°C/min to 110°C, at 10°C/min to 180°C and at 15°C/min 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 (Agilent) on a Zorbax SB-C18 rapid resolution HT column (4.6 x 40mm x 1.8μm) in positive ESI mode (dry temperature 350°C, nebulizer 55 psi). 2.5μl samples were injected with a column flow of 1ml/min in isocratic mode (acetonitrile/water 85%/15% + 0.2% formic acid) with the diode array detector scanning from 190-400nm for a run time of 10 min. Diterpenes from Sitka spruce tissue extracts were analyzed by GC-MS with a AT1000 capillary column (30m x 0.25mm x 0.25μm, Alltech) in scan mode (scan: m/z
40-500). 0.5μl samples were injected in pulsed splitless mode at 250°C with a column
flow of 1ml He/min and 30psi pulse pressure. The GC oven temperature was
programmed to rise from initial 150°C at 1.5°C/min to 220°C, at 20°C/min 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. (2006) with minor
modifications. Extractions were scaled down to 50-100 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 in Miller et al. (2005) using Superscript III (Invitrogen,
Canada). For microdissected samples, RNA was isolated from 20 μm thick longitudinal
stem sections, epithelial cells of cortical resin ducts, and cambial zone tissue using the
RNAqueous-Micro Kit (Ambion, USA) as described in Abbott et al. (2010). Bioanalyzer
2100 RNA pico-chip assays (Agilent) were used to determine RNA integrity and
concentration.

**CYP720B gene discovery and FLcDNA cloning.** General procedures for mining of the
pine and spruce sequence databases for genes of the CYP85 clan including CYP720B
genes were previously described (Hamberger and Bohlmann, 2006). *In silico* generated
contigs for candidate CYP720B genes from three spruce [Sitka spruce (*P. sitchensis*),
white spruce (*P. glauca*), interior spruce (*P. glauca x engelmannii*), and two pine species
[loblolly pine (*P. taeda*), lodgepole pine (*P. 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, USA) was used to obtain FLcDNAs of partial Sitka spruce ESTs. PCR
amplification of FLcDNAs was performed over 40-45 cycles with Phusion polymerase
(Finnzymes, Finland) and gene-specific oligonucleotide primers (Integrated DNA
Technologies, USA). Amplicons were ligated into pJET (Fermentas, Canada), and inserts were sequenced completely.

**Real-time quantitative PCR (RT-qPCR) analysis.** cDNA was synthesized from equal amounts of total RNA using Superscript III reverse transcriptase (Invitrogen) with random hexamer oligonucleotides. Efficiency of reverse transcription reactions was assessed with internal calibrator reactions on four reference transcripts, elongation factor 1 (ELF1), translation initiation factor (EIF-5A)-like (TIF), polyubiquitin (UBI) and tubulin alpha chain homolog (TUA). For 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 (MJ Research, USA) with the Opticon3 software. HK-uracil N-glycosylase reaction (Epicentre Biotechnologies, Canada) preceded the quantitative PCR. Real-time PCR miner program (Zhao and Fernald, 2005) was used to determine efficiency and C_T values. Efficiency corrected ΔC_T values were calculated and ΔΔC_T were used to quantify relative differences in transcript accumulation. AmplifX was used for oligonucleotide optimization (Amplifx 1.5.4 Nicolas Jullien, [http://ifrjr.nord.univ-mrs.fr/AmplifX](http://ifrjr.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 Information (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/isoamylalcohol extraction, nucleic acids were pelleted with isopropanol, washed with 70% ethanol, resuspended in nuclease free water, and treated with RNaseA and proteinaseK 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 primers listed in Supplemental...
Information (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/ScGGDPS for co-expression of PaLAS and ScGGDPS in *S. cerevisiae* was previously described (Ro et al., 2005). Using the same approach, pESC-His::PaISO/ScGGDPS was generated for co-expression of ScGGDPS and PaISO lacking the N-terminal plastid-targeting sequence (Martin et al., 2004). Codon optimization of the 5’-region of PsCYP720B4 (Supplemental Figure S2) re-coded 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% (Marin et al., 2003) and rare yeast codons were replaced. Using eight partially overlapping longmer oligonucleotides, the optimized 561 bp 5’-region of PsCYP720B4 was synthesized by PCR and ligated to 891 bp of the native 3’-region. *SphI* 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 Information (Supplemental Text S1). To generate yeast strains for P450-dependent *de novo* diterpenoid formation, plasmids pESC-His::PaLAS/ScGGDPS or pESC-His::PaISO/ScGGDPS 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, media and culture conditions were previously described (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/ScGGDPS//pESC-Leu::PsCYP720B4 or pESC-His::PaISO/ScGGDPS//pESC-Leu::PsCYP720B4 are described below. Diterpenoids extracted from yeast cultures were analyzed by GC-MS after derivatisation.
**In vivo assays for PsCYP720 activity using yeast strains carrying the plasmid pESC-Leu::PsCYP720B4.** The basic procedures were as previously described in Ro et al. (2005). Heterologous protein expression in yeast cultures was initiated at an optical density of 0.6 OD\(_{600}\) by transfer into minimal selection medium containing 2% galactose. 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 in Ro et al. (2005). Diterpenoids were dissolved in either acetonitrile (alcohols), ethylacetate (alcohols, aldehydes) or diethyl ether (olefins) depending on the substrates polarity. While acetonitrile is 100% water soluble, ethyl acetate and to a lesser extend 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 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, the (i) olefins dehydroabietadiene, isopimaradiene, palustradiene, (ii) alcohols dehydroabietanienol, isopimaradienol, palustradienol or (iii) aldehydes dehydroabietadienal, isomiparadienal, palustradienal at 100 μM to 1 μM. For example, complete conversion of dehydroabietadienal to dehydroabietic acid was observed at 10 μM. To analyse diterpene resin acid formation in yeast, individual compounds were added to induced yeast cultures to a final concentration of 10 μM in 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/ScGGDPS//pESC-Leu::PsCYP720B4 or pESC-His::PaISO/ScGGDPS//pESC-Leu::PsCYP720B4.** For analysis of *de novo* diterpenoid formation in engineered yeast strains, cultures were grown to a density of 0.6 OD\(_{600}\) before transfer into minimal selection medium containing 2% galactose. After growth for 16 h, yeast cells were pelleted by centrifugation at 3,200 x g for 10 min. Cell pellets were extracted twice with 5 ml diethyl ether with 0.5 μM eicosene and 3 μM of dichlorodehydroabietic acid (Orchid Cellmark) as internal standard. Anhydrous sodium sulphate was added to remove residual water and the combined organic phases were
concentrated to 500 μl. To derivatise the diterpene resin acids, 100 μl methanol and 60 μl 
(Trimethylsilyl)diazo-methane (Aldrich, 2.0M 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 μl 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 
(MALLLAVF) of bovine CYP17a (Supplemental Figure S4) (Barnes et al., 1991; 
Kagawa et al., 2003). The modified PsCYP720B4 cDNA was ligated into Ndel and 
HindIII sites of the bacterial expression vector pCWori⁺. PsCPR was amplified from a 
Sitka spruce cDNA library constructed from flushing buds and ligated into NcoI and SalI 
sites of pET28 (Novagen) to generate the bacterial expression construct pET28-PsCPR. 
Oligonucleotides are given in Supplemental Information (Supplemental Text S1). PsCPR 
was heterologously expressed in the E. coli strain BL21(DE3) and JM109 was used for 
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 PsCYP720B4 in E. coli membrane faction 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 [3]. The E. coli strain JM109 was used for expression of recombinant 
CYP720B4 following Fujita et al. (2006). E. coli cells were harvested by centrifugation at 
25,000 x g 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 sucrose 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 3,000 x g for 15 min to remove cell debris. The supernatant was further centrifuged at 100,000 x g 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 either acetonitrile (alcohols), ethylacetate (alcohols, aldehydes) or diethyl ether (olefins) depending on the substrates polarity. PsCYP720B4 in E. coli membrane fractions was mixed with purified PsCPR consisting of 50 mM potassium phosphate (pH 7.25), 25 pmol/ml of PsCYP720B4, PsCPR (0.1 U for analysis of enzyme activities; 5 mU for analysis of relative activities and kinetic parameters), diterpene substrates (20 μM for analysis of enzyme activities and relative activities, and a range of concentration of substrate for kinetic assays of 0.625 to 40 μM). Reactions were initiated by 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 μl of ethyl acetate, 10 μl 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 μL 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 in 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 Information (Supplemental Text S1)] and cloned in sense and antisense orientation in the multicloning site of pTRAIN on either side of the intron as described in Levée et al. (2009). The RNAi cassette is positioned downstream of the Ubiquitin promoter-UTR region (Christensen et al., 1992). The resulting plasmid was HindIII digested and the excised cassette carrying the RNAi construct 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 followed 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 one 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 v7.0.9 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, JTT (Jones–Taylor–Thornton) substitution model, BioNJ starting tree and 100 bootstrap repetitions (Guindon and Gascuel, 2003)] and displayed as 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).

**Accession numbers.** The sequences reported in this manuscript have been deposited in the NCBI GenBank database under accession numbers HM245397-HM245410 and HQ888858.

**Author responsible for distribution of materials.** The author responsible for distribution of materials integral to the findings presented in this article in accordance with the Journal policy described in the Instructions for Authors (http://www.plantphysiol.org) is: Jörg Bohlmann (bohlmann@msl.ubc.ca). The contact address for distribution of materials is Michael Smith Laboratories, University of British Columbia, 2185 East Mall, Vancouver, BC, V6T 1Z4, Canada.

**Supplemental Data Files.**
Supplemental Text S1: Oligonucleotides used for PCR.
Supplemental Table S I: Identification of diterpenoids by LC-MS and GC-MS.
Supplemental Table S II: Co-Suppression effects of CYP720B4 RNAi
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. (A)
Sequence of the N-terminus of PsCYP720B4 modified for expression in E. coli. (B)
Carbon monoxide differential spectrum of PsCYP720B4 protein in E. coli membrane fraction.
Supplemental Figure S5: Transformation of diterpene olefins, alcohols and aldehydes added to yeast expressing PsCYP720B4 and PsCPR and formation of diterpene resin acids.

Acknowledgements
We thank Dr. Masaharu Mizutani (Graduate School of Agricultural Science, Kobe University, Japan) for providing the pCWori+ vector, Dr. David Nelson (Department of Molecular Sciences, TN, USA) for assigning P450 nomenclature, Dr. Krystyna Klimaszewska (Natural Resources Canada, Canadian Forest, Sainte-Foy, Québec) for white spruce seedlings, Dr. David Ellis (formerly CellFor Inc., Victoria, BC) for Sitka spruce seedlings, Ms. Lina Madilao (UBC) for excellent analytical support, Ms. Karen Reid (UBC) for superb lab management support, and Drs. Dawn Hall and Katherine Zulak for critical reading of the manuscript and helpful comments.
Literature Cited


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


Ohnishi T, Szatmari AM, Watanabe B, Fujita S, Bancos S, Koncz C, Lafos M,
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
Bioinformatics Chapter 6: Unit 6 2
Abietadiene synthase from grand fir (Abies grandis): characterization and
mechanism of action of the "pseudomature" recombinant enzyme. Biochemistry
39: 15592-15602
Pichersky E, Gang DR (2000) Genetics and biochemistry of secondary metabolites in
plants: an evolutionary perspective. Trends in Plant Science 5: 439-445
Science 4: 184-190
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. Environmental
Entomology 28: 148-154
abietadienol/abietadienal oxidase PtAO (CYP720B1) is a multifunctional,
multisubstrate cytochrome P450 monoxygenase. Proc Natl Acad Sci U S A 102:
8060-8065
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, Chang MCY, Withers ST, Shiba Y, Sarpong
R, Keasling JD (2006) Production of the antimalarial drug precursor artemisinic
acid in engineered yeast. Nature 440: 940-943
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
bifunctional geranyl and geranylgeranyl diphosphate synthase is involved in
Schmidt HA, von Haeseler A (2007) Maximum-likelihood analysis using TREE-
PUZZLE. Curr Protoc Bioinformatics Chapter 6: Unit 6 6


Shoji T, Hashimoto T (2011) Recruitment of a duplicated primary metabolism gene into the nicotine biosynthesis regulon in tobacco. Plant J 67: 949-959


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
**Figure legends**

**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 gibberellins plant hormones in general plant metabolism are shown. 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 dehydroabietadiene. The arrow indicates 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 to bottom. Traces: traces were identified but amounts were below limits for accurate quantification; n.d.: not detectable; error bar SE, n=4 independent biological replicates.

**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 bold 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 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 (Picea spp.) and pine (Pinus spp.). The maximum likelihood phylogeny was reconstructed using amino acid sequences deduced
from cDNAs of Sitka spruce (Ps, *P. sitchensis*), white spruce (Pg, *P. glauca*), interior spruce (Pge, *P. glauca x engelmannii*), loblolly pine (Pt, *P. taeda*), lodgepole pine (Pc, *P. contorta*). The CYP720B sequences cluster into four clades I-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 thaliana* AtCYP720A1 as outgroup. FL, full length cDNA clones; v, putative allelic variant.

**Figure 3. Relative transcript abundance of 12 different PsCYP720B in different parts of Sitka spruce.** Transcripts were measured in wood and xylem (W) of the first interwhorl from the top; bark and phloem of the first (B1) or second (B2) interwhorl; roots (R); young (YN) and mature (MN) needles; young shoots (YS) (Supplemental Figure S3). (A-D) Genes are grouped according to their phylogenetic association with clades I - IV as indicated with insets. Transcripts abundance was measured by RT-qPCR and normalized to the internal reference TIF. Error bars represent standard errors of biological replicates (SE, n=3). Statistically significant differences are indicated with different letters (*t* test, p<0.05). Transcript detected, but below threshold for quantification, t.d.; transcript not detectable, n.d..

**Figure 4. Relative transcript abundance of 12 different PsCYP720B and PsLAS in Sitka spruce bark and phloem tissue in response to treatment with MeJA.** (A-D) Genes are grouped according to their phylogenetic association with clades I - IV. Transcripts abundance was measured by RT-qPCR and normalized to the internal reference TIF. Error bars represent standard errors of biological replicates (SE, n=4). Statistical significant differences of the treated samples to control are indicated by asterisks (*, p<0.05; **, p<0.01). Transcript detected, but below threshold for quantification, t.d.; transcript not detectable, n.d..

**Figure 5. Transcript abundance of 12 different PsCYP720B, 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 cortical resin ducts (CRD) 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 numbers in parentheses indicate the clade. Error bars represent standard errors of biological replicates (SE, n=3). Statistically significant differences are indicated by letters (t test, p<0.05). Transcript detected, but below threshold for quantification, t.d.; transcript not detectable, n.d.

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

Figure 7. Oxydation of dehydroabietadiene, dehydroabietadienol and dehydroabietadienal in in vitro assays with PsCYP720B4 protein in E. coli membrane fraction. LC-MS selected ion chromatograms of oxygenated diterpene substrates and products from in vitro assays with recombinant PsCPR and PsCYP720B4 after 0 min or 30 min of incubation with dehydroabietadiene, dehydroabietadienol or dehydroabietadienal as substrates. Assays were performed with 1 mM NADPH and 0.1 U 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.
Dehydroabietic acid (6), dehydroabietadienol (7), dehydroabietadienal (8). See Supplemental Table S1 for additional details of metabolite identification.

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 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 mU 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 analysed with two technical replicates by GC-MS. No conversion detected, n.c.. 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.

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 months old white spruce plants of four independent RNAi lines and vector transformed controls. Error bars represent standard errors of biological replicates (SE, n=3). (B) Relative abundance of DRAs in stems of one-year old white spruce plants of four independent RNAi lines and vector transformed controls, scale 0 - 1. Asterisk indicates significant (t test, p<0.05) reduction compared with the empty vector control. Error bars represent standard errors of biological replicates (SE, n=5). (C) Relative abundance of diterpene resin acids in stems of one-year old white spruce plants of four independent RNAi lines and vector transformed controls, scale 0 - 6.
Figure 10. Intron-exon gene structures of representative Sitka spruce PsCYP720B1 genes of clades I - IV, PsCYP90A20, PsCYP88, and PsCYP701. The scale bar represents 1kb. Roman numerals indicate exons.

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 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 multi-gene 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 P450s of specialized DRA metabolism, accepts multiple diterpene substrates and catalyses three consecutive oxidations, thus contributing to the structural diversity of DRA in specialized metabolism.
## Tables

Table I. Apparent Kinetic Parameters of CYP720B4 \(^1,2\)

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

\(^1\) Assays were performed with 1 mM NADPH and 5 mU CPR.
\(^2\) Different physicochemical properties of olefins, alcohols and aldehydes (e.g., differences in solubility) may affect the apparent kinetic parameters.
Figure 1. Metabolite profiles of diterpene olefins, alcohols, aldehydes and acids in the outer and inner 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 gibberellins in general plant metabolism are shown. 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 dehydroabietadiene. The arrow indicates 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: traces were identified but amounts were below limits for accurate quantification; n.d.: not detectable; error bar SE, n=4 independent biological replicates.
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 bold 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 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 (Picea spp.) and pine (Pinus spp.). The maximum likelihood phylogeny was reconstructed using amino acid sequences deduced from cDNAs of Sitka spruce (Ps, P. sitchensis), white spruce (Pg, P. glauca), interior spruce (Pge, P. glauca x engelmannii), loblolly pine (Pt, P. taeda), lodgepole pine (Pc, P. contorta). The CYP720B sequences cluster into four clades I-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 thaliana AtCYP720A1 as outgroup. FL, full length cDNA clones; v, putative allelic variant.
Figure 3. Relative transcript abundance of 12 different PsCYP720B in different parts of Sitka spruce. Transcripts were measured in wood and xylem (W) of the first interwhorl from the top; bark and phloem of the first (B1) or second (B2) interwhorl; roots (R); young (YN) and mature (MN) needles; young shoots (YS) (Supplemental Figure S3). (A-D) Genes are grouped according to their phylogenetic association with clades I - IV as indicated with insets. Transcripts abundance was measured by RT-qPCR and normalized to the internal reference TIF. Error bars represent standard errors of biological replicates (SE, n=3). Statistically significant differences are indicated with different letters (t test, p<0.05). Transcript detected, but below threshold for quantification, t.d.; transcript not detectable, n.d.
Figure 4. Relative transcript abundance of 12 different PsCYP720B and PsLAS in Sitka spruce bark and phloem tissue in response to treatment with MeJA. (A-D) Genes are grouped according to their phylogenetic association with clades I - IV. Transcripts abundance was measured by qRT-PCR and normalized to the internal reference TIF. Error bars represent standard errors of biological replicates (SE, n=4); statistical significant differences of the treated samples to control are indicated by asterisks (*, p<0.05; **, p<0.01). Transcript detected, but below threshold for quantification, t.d.; transcript not detectable, n.d.
Figure 5. Transcript abundance of 12 different PsCYP720B, 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 cortical resin ducts (CRD) 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 numbers in parentheses indicate the clade. Error bars represent standard errors of biological replicates (SE, n=3). Statistically significant differences are indicated by letters (t test, p<0.05). Transcript detected, but below threshold for quantification, t.d.; transcript not detectable, n.d.
Figure 6. PsCYP720B4 dependent in vivo formation of DRAs isopimaric acid and abietic acid in engineered yeast. (A) GC-MS total ion chromatograms of extracts from engineered yeast strains expressing GGPPS, diTPS isopimaradiene synthase (PaISO) and CPR in the presence (lower trace) or absence (middle trace) of PsCYP720B4. (B) GC-MS total ion chromatograms of extracts from engineered yeast strains expressing GGPPS, diTPS abietadiene synthase PaLAS, and CPR in the presence (lower trace) or absence (middle trace) of PsCYP720B4. Retention times and mass spectra of compounds of compounds (1; isopimaradiene), (2; isopimaric acid methylester), (3; abietadiene) and (4; abietic acid methylester) match those of authentic standards. A minor peak labeled with an asterisk in (B) differs in retention from (4) and represents a phthalate derivative contamination. Upper traces in (A) and (B) are GC-MS total ion chromatograms of extracts from control engineered yeast strains transformed with empty vectors. See Supplemental Table S1 for additional details of metabolite identification.
Figure 7. Oxidation of dehydroabietadiene, dehydroabietadienol, and dehydroabietadienal in \textit{in vitro} assays with PsCYP720B4 in \textit{E. coli} membrane fraction. LC-MS selected ion chromatograms of oxygenated diterpene substrates and products from \textit{in vitro} assays with recombinant PsCPR and PsCYP720B4 after 0 min or 30 min of incubation with dehydroabietadiene, dehydroabietadienol or dehydroabietadienal as substrates. Assays were performed with 1 mM NADPH and 0.1 U CPR. Substrates are shown with their chemical structures. Dehydroabietadiene (top panel) is not detectable in this type of analysis (LC-MS). Diterpenoids were identified by comparison of their retention times and mass spectra with those of authentic standards. Dehydroabietic acid (6), dehydroabietadienol (7), dehydroabietadienal (8), with retention times extended from the authentic standards as dotted lines. See Supplemental Table S1 for additional details of metabolite identification.
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 in 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 mU 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 experiment was repeated three times, each time with two biological replicates, each of which was analysed with two technical replicates by GC-MS. No conversion detected, n.c.. 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.
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 months old white spruce plants of four independent RNAi lines and vector transformed controls. Error bars represent standard errors of biological replicates (SE, n=3). (B) Relative abundance of DRAs in stems of one-year old white spruce plants of four independent RNAi lines and vector transformed controls, scale 0 - 1. Asterisk indicates significant (T-test, p<0.05) reduction compared with the empty vector control. Error bars represent standard errors of biological replicates (SE, n=5). (C) Relative abundance of diterpene resin acids in stems of one-year old white spruce plants of four independent RNAi lines and vector transformed controls, scale 0 - 6.
Figure 10. Intron-exon gene structures of representative Sitka spruce PsCYP720B genes of clades I - IV, PsCYP88, PsCYP90, and PsCYP701. The scale bar represents 1kb. Roman numbers indicate exons, asterisks indicate expansion of three introns in PsCYP90A20.
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 general metabolism of ent-kaurenoic acid seem to be encoded by single copy genes. The diTPSs and P450s of general metabolism of ent-kaurenoic acid 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 multi-gene 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 P450s of specialized DRA metabolism, accepts multiple diterpene substrates and catalyzes three consecutive oxidations, thus contributing to the structural diversity of DRA in specialized metabolism.