SPECIAL ISSUE REVIEWS AND SYNTHESES

Evolutionary Applications Open Access **WILEY**

Functional and morphological evolution in gymnosperms: A portrait of implicated gene families

Amanda R. De La Torre[1] iD | Anthony Piot[2,3,4] | Bobin Liu[1,5] | Benjamin Wilhite[1] | Matthew Weiss[1] | Ilga Porth[2,3,4]

[1]School of Forestry, Northern Arizona University, Flagstaff, AZ, USA

[2]Department of Wood and Forest Sciences, Laval University, Quebec City, Quebec, Canada

[3]Institute for System and Integrated Biology (IBIS), Laval University, Quebec City, Quebec, Canada

[4]Centre for Forest Research (CEF), Laval University, Quebec City, Quebec, Canada

[5]College of Forestry, Fujian Agricultural and Forestry University, Fuzhou, Fujian, China

Correspondence
Amanda R. De La Torre, School of Forestry, Northern Arizona University, Flagstaff, AZ, USA.
Email: Amanda.de-la-torre@nau.edu

Abstract

Gymnosperms diverged from their sister plant clade of flowering plants 300 Mya. Morphological and functional divergence between the two major seed plant clades involved significant changes in their reproductive biology, water-conducting systems, secondary metabolism, stress defense mechanisms, and small RNA-mediated epigenetic silencing. The relatively recent sequencing of several gymnosperm genomes and the development of new genomic resources have enabled whole-genome comparisons within gymnosperms, and between angiosperms and gymnosperms. In this paper, we aim to understand how genes and gene families have contributed to the major functional and morphological differences in gymnosperms, and how this information can be used for applied breeding and biotechnology. In addition, we have analyzed the angiosperm versus gymnosperm evolution of the *pleiotropic drug resistance* (*PDR*) gene family with a wide range of functionalities in plants' interaction with their environment including defense mechanisms. Some of the genes reviewed here are newly studied members of gene families that hold potential for biotechnological applications related to commercial and pharmacological value. Some members of conifer gene families can also be exploited for their potential in phytoremediation applications.

KEYWORDS

biotechnology, functional evolution, gene families, gymnosperms, PDR gene family, reproductive biology, stress

1 | INTRODUCTION

Gymnosperms are an ancient and widespread nonflowering plant lineage of great economic and ecological importance. With only 1,000 living species, gymnosperms represent four of the five seed plant lineages including conifers (Pinophyta), cycads (Cycadophyta), ginkgos (Ginkgophyta), and gnetophytes (Gnetophyta; Wang & Ran, 2014). Coniferous species make up 39% of the world's forests

and represent great value for forestry-dependent economies in Northern and Southern hemispheres (Armenise, Simeone, Piredda, & Schirone, 2012). Efforts to understand their biology, and genomic and functional evolution have been limited by their life-history characteristics and large genome sizes (De La Torre et al., 2014). Recent studies highlighted the underpinnings of the major morphological, genomic, and functional differences that shaped the evolutionary divergence among gymnosperms and flowering plants.

The most noteworthy differences between angiosperms and gymnosperms certainly occur at the morphological level. Flowers, the major functional innovation in angiosperms, are assumed to have evolved through the transformation of gymnosperms' separate male and female structures into an integrated hermaphrodite structure (Niu et al., 2016; Pires & Dolan, 2012). Similarly, angiosperms developed a more efficient method of water transport through vessels, while tracheids are present in gymnosperm species (with the exception of gnetales), but also in the basal angiosperm *Amborella trichopoda*. Gene families involved in secondary metabolism such as terpene biosynthesis or various alkaloid biosynthesis pathways evolved differently in gymnosperms and flowering plants (Chen, Tholl, Bohlmann, & Pichersky, 2011; Hall, Zerbe, et al., 2013b). In this review paper, we aim to understand how a subset of well-studied genes and gene families have contributed to the evolution of major morphological and functional differences between angiosperms and gymnosperms including their reproductive biology, water-conducting xylem tissues, secondary metabolism and stress, and noncoding and small RNAs. In addition, we analyzed the gene family evolution of the *pleiotropic drug resistance* (*PDR*) proteins, known to play important roles in plant–environment interactions in angiosperms. Some of the gymnosperm genes reviewed here are newly studied members of gene families such as PDR that hold potential for biotechnological applications with commercial and pharmacological value. Some members of conifer gene families have potential to be exploited for improved growth on marginal or disturbed soils, by increasing the detoxification potential of spruces in phytoremediation applications.

2 | GENOMIC EVOLUTIONARY DIFFERENCES BETWEEN ANGIOSPERMS AND GYMNOSPERMS

Before the extensive radiation of flowering plants during the late Cretaceous, gymnosperms dominated the world flora for almost 200 million years (Pennisi, 2009; Pires & Dolan, 2012). Extreme climatic shifts over the Cenozoic resulted in major extinctions in the gymnosperm lineage, which may account for the low diversity of extant gymnosperms in comparison with their sister seed plant clade (Crisp & Cook, 2011; Leslie et al., 2012). Extinctions were more pronounced in the Northern hemisphere in which older lineages were replaced by those better adapted to cooler and drier environmental conditions, resulting in higher species turnover rates in Pinaceae and Cupressaceae, compared to southern lineages (Leslie et al., 2012). More recently, climatic changes during the last Glaciation strongly shaped species distributions and patterns of speciation and adaptation for many Northern hemisphere gymnosperms which went through cycles of contraction and expansion from refugia (Shafer, Cullingham, Côté, & Coltman, 2010).

While angiosperm evolution has been shaped by whole-genome duplication (WGD) events leading to higher speciation rates and the development of key functional innovations, gymnosperm genomes have been less dynamic (Landis et al., 2018; Soltis & Soltis, 2016; Vanneste, Maere, & Van de Peer, 2014). The rarity of WGD, paucity of chromosomal rearrangements, and slow mutation rates have led to low levels of structural genomic and morphological variation among species, and low speciation rates in gymnosperms (De La Torre et al., 2014; De La Torre, Li, Van de Peer, & Ingvarsson, 2017; Jaramillo-Correa, Verdu, & Gonzalez-Martinez, 2010; Leitch & Leitch, 2012; Pavy et al., 2012). In the presence of polyploidy and retro-transposition, angiosperms have developed mechanisms to counteract the increase in genomic DNA by replication or recombination-based errors generating indels, and unequal recombination between sister chromosomes (Grover & Wendel, 2010; Leitch & Leitch, 2012).

Although polyploidy is largely absent in gymnosperms (exceptions are *Sequoia*, *Pseudolarix*, and *Ephedra*), a combination of a massive accumulation of long-terminal repeat retrotransposons (LTR-RTs), together with limited removal of transposable elements through unequal recombination, has resulted in very large genome sizes (mean 1C = 18.8 pg; De La Torre et al., 2014; Leitch & Leitch, 2012; Nystedt et al., 2013). Recent studies revealed transposable elements make up to 74%, 76.58%, 79%, and 85.9% of the genomes of *Pinus taeda*, *Ginkgo biloba*, *Pinus lambertiana*, and *Gnetum montanum*, respectively (Guan et al., 2016; Neale, Martínez-García, Torre, Montanari, & Wei, 2017; Neale et al., 2014; Wan et al., 2018; Wegrzyn et al., 2014). A comparative analysis among six diverged gymnosperms suggested the diversity and abundance of transposable elements is widely conserved among gymnosperm taxa (Nystedt et al., 2013). However, a more recent study focused on gnetophytes (*Gnetum*, *Welwitschia*, *Ephedra*) suggests higher frequencies of LTR-RT elimination due to recombination-based processes of genome downsizing may explain the smaller sizes of gnetophytes in comparison with other gymnosperm genomes (Wan et al., 2018).

Despite significant variations in noncoding regions of angiosperm and gymnosperm genomes, both plant lineages have comparable numbers of genes and gene families. Sequence similarities of expressed genes are 58%–61% between conifers and angiosperms, and 80% within Pinaceae (Prunier, Verta, & MacKay, 2016; Rigault et al., 2011). This suggests that functional differences observed between seed plant lineages may have evolved as a consequence of differences in rates of nucleotide substitution, frequency of copy number variant (CNV) formation (Prunier, Caron, & MacKay, 2017; Prunier, Caron, Lamothe, et al., 2017; for a discussion of poplar vs. spruce CNVs see Prunier et al., 2019), and/or differential gene family expansion or contraction (Zhou et al., 2019). A recent analysis of protein-coding genes across a broad phylogeny suggested slower rates of molecular evolution (number of synonymous substitutions dS and mutation rates), but higher substitution rate ratios (dN/dS) in gymnosperms than in angiosperms (De La Torre et al., 2017). Higher levels of dN/dS in gymnosperms suggest stronger and more effective selection pressures probably due to larger effective population sizes, especially in the Pinaceae (De La Torre et al., 2017). In addition, gymnosperms generally present high levels of within-population genetic diversity, while long-distance gene flow of wind-dispersed pollen between highly outcrossed populations leads to rapid decay of

TABLE 1 Gene families showing differential expansions in gymnosperms

Gene family	Function	Taxon	Reference
MYB	Defense response, vascular development	*Pinus taeda, Picea glauca*	Bedon et al. (2010); Patzlaff et al. (2003)
NB-LRR	Disease resistance	*Picea abies*	Fossdal et al. (2012)
sRNA's (SGS3, DCL1)	Epigenetics, transposable element silencing	*P. taeda, Pinus lambertiana, P. abies*	Gonzalez-Ibeas et al. (2016); Yakovlev, Carneros, Lee, Olsen, and Fossdal (2016)
Phenylalanine ammonia lyase (PAL)	Lignin biosynthesis	*P. taeda*	Bagal, Leebens-Mack, Lorenz, and Dean (2012)
Ty3/Gypsy	Transposable element	*P. abies, Pinus palustrus, Abies concolor, Podocarpus totara* and others	Nystedt et al. (2013); Friesen, Brandes, and Heslop-Harrison (2001)
Cytochrome P450	Monoterpenoid production	*Thuja plicata* / *P. glauca*	Gesell et al. (2015); Warren et al. (2015)
miR390-TAS3-ARF	Auxin signaling	*P. abies*	Xia, Xu, Arikit, and Meyers (2015)
CslE/J/G-like	Cellulose synthesis	*P. abies, Cryptomeria japonica, Pinus banksiana* and others	Yin et al. (2014)
Dehydrins	Drought, cold tolerance	*P. glauca, Abies balsamea, Larix laricina* and others	Stival Sena et al. (2018); Rigault et al. (2011)
Glucosinolate biosynthesis and α-bisabolene synthase	Terpene-mediated defense	*Ginkgo biloba*	Guan et al. (2016)
FSL2 and EFR	Bacterial infection defense	*G. biloba*	Guan et al. (2016)

linkage disequilibrium and low among-population genetic diversity (De La Torre et al., 2014, 2017; Porth & El-Kassaby, 2014). Higher gene turnover, which probably explains a higher species turnover, has been observed in Pinaceae. Although the cause of this is unknown, it is being suggested that this trend might be explained by an elevated frequency of gene CNVs, although rates of CNV formation in Pinaceae or any other gymnosperms are unknown (Casola & Koralewski, 2018).

Because gymnosperms predate angiosperms, most differential gene family expansions between angiosperms and gymnosperms seem to have occurred either by loss of genes in angiosperms (most likely scenario) or gain in gymnosperms (by neofunctionalization or subfunctionalization). Large expanded paralogous gene families such as leucine-rich repeats, cytochrome P450, MYB, and others (Table 1) have been observed in gymnosperms (De La Torre, Lin, Van de Peer, & Ingvarsson, 2015; Neale et al., 2014; Pavy et al., 2013; Porth, Hamberger, White, & Ritland, 2011; Warren et al., 2015). While comparing differentially expanded gene families using whole-genome data, our study found that *Picea abies*' larger gene ontologies, compared to those of *Arabidopsis thaliana*, are the consequence of the species' ability to respond to diverse stimuli (biotic and abiotic stress), transport mechanisms, and a variety of specific metabolic and biosynthetic processes (Figure 1).

Sequence variation leading to neofunctionalization in large paralogous families might have resulted in the evolution of lineage-specific expansions between angiosperms and gymnosperms or within gymnosperms lineages (e.g., *Picea* and *Pinus*; Gonzalez-Ibeas et al., 2016). Conifer-specific gene families have been observed in the Terpene synthase (TPS) subfamily (Hall, Zerbe, et al., 2013b; Shalev et al., 2018; Warren et al., 2015), in specific subfamilies of P450s such as the CYP720Bs subfamily (Geisler, Jensen, Yuen, Madilao, & Bohlmann, 2016; Warren et al., 2015) or the CYP750 family (Gesell et al., 2015; Porth et al., 2011), and in transcription factors such as mTERF involved in abiotic stress and plant growth (Gonzalez-Ibeas et al., 2016). F-box proteins that are subunits of the E3 ubiquitin ligase aggregations also known as the SCF quaternary complex (SKP1,
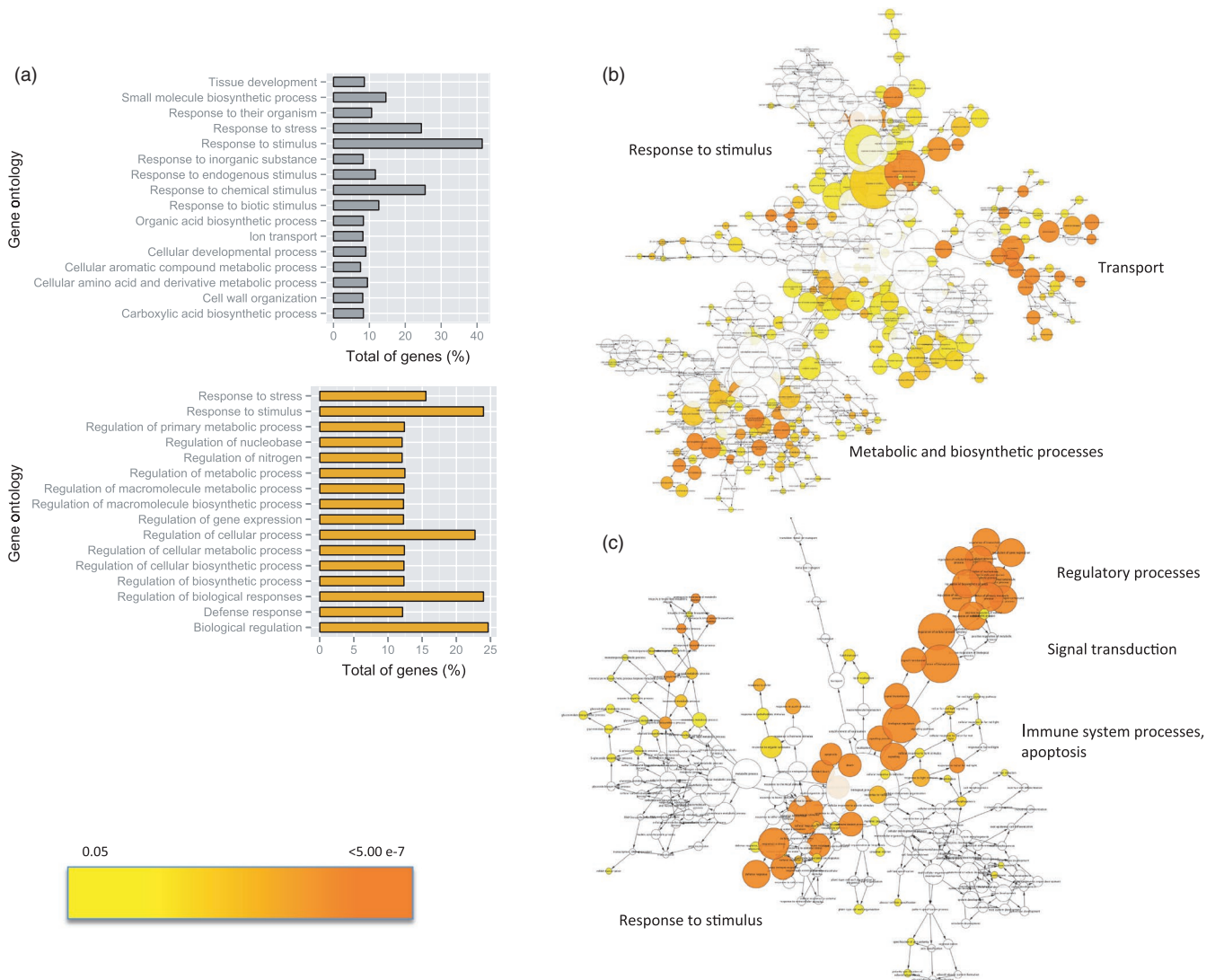
FIGURE 1 Functional enrichment of genes in gene families showing differential expansions between gymnosperms (*Picea abies*) and angiosperms (*A. thaliana*). (a) Barplot showing significant (*p* < 0.0001) enriched gene ontologies for gymnosperms (gray bars) and angiosperms (yellow bars); (b) Gene interaction network showing significant gene ontologies for gymnosperms species; (c) Gene interaction network showing significant gene ontologies for angiosperms species. *p*-Values vary from 0.05 (yellow) to <5.00e-7 (dark orange) according to the color scale in left bottom (A. De La Torre & P.K. Ingvarsson, unpublished). *P. abies* genome version 1.0 (Nystedt et al., 2013; http://congenie.org) was used for this analysis

Cullin1, F-box protein, and Rbx1) also reveal conifer-specific gene families (Gonzalez-Ibeas et al., 2016).

3 | REPRODUCTIVE BIOLOGY

The reproductive biology in gymnosperms is characterized by a largely outcrossing mating system, predominant anemophily (wind pollination) and wind-mediated seed dispersal. Other characteristics that differ between gymnosperms and angiosperms are the presence of uncovered seeds (lack of fruit), a *haploid* nourishing tissue (megagametophyte) surrounding the diploid embryo in the developing seed, and temporary polyembryony.

Genomic research on the reproductive biology of gymnosperms is not as advanced as in angiosperms (Cairney & Pullman, 2007; Rodrigues, De Vega, & Miguel, 2018). This is not surprising given the lack of genetic mutant lines required for discovering genes involved in plant development and reproduction (these resources are largely available for model plant species such as *A. thaliana*: the SALK lines, e.g.). Knowledge about the genetic underpinnings of reproductive and seed development is scarce in conifers, with *Pinus* being the most studied conifer genus to date, and *Picea* to a lesser extent. Recent studies in *Picea* suggest individual or multiple gene family members involved in reproductive development have gene expression-based relationships with tree growth and environmental stress (abiotic, biotic defenses); however, the exact molecular nature

of their involvement is unknown (Porth et al., 2012; Porth, White, Jaquish, & Ritland, 2018).

Some of the most thorough reproductive development studies on pines have been on trees growing under somewhat less natural conditions and therefore prone to artificial stress (such as botanical garden settings). This has resulted in a higher frequency of anomalies such as bisexual reproductive structures (hermaphroditism) than it is expected under natural conditions (Flores-Renteria et al., 2011; Niu et al., 2016). Nevertheless, teratology and hermaphroditism might have been common in early and now extinct gymnosperms (Rudall, Hilton, Vergara-Silva, & Bateman, 2011). While species members of the Pinaceae family are mostly monoecious (gymnosperms overall are mostly dioecious), their male and female reproductive structures are unisexual, and megasporophylls and microsporophylls occupy distinct locations on a tree's shoot. The female structures develop at the top while male structures form at the base of a shoot in normal tree development (Niu et al., 2016), making self-pollination difficult, and thus helping to maintain outcrossing rates and genetic diversity (Porth & El-Kassaby, 2014). Furthermore, incompatibility mechanisms (pollen discrimination) in conifers occur within the ovule and are mostly postzygotic (Owens, Takaso, & Runions, 1998).

In order to facilitate pollen release and dispersal through wind travel, gymnosperms' male reproductive structures (male cones; pollen grains) have evolved an impressive diversity of male cone positioning, and grain shapes (Lu et al., 2011). This is seen as a necessity to overcome the innate constraints from gymnosperm's heavy (but not always exclusive) reliance on anemophily. In contrast, angiosperms evolved flowers with attractive colors and fragrances as signals for pollination by insects and other animals, pollen and nectar rewards as food source for the pollinator, as well as fruits for their seeds' protection but also dispersal by animals. Ovular secretion is also crucial to reproduction in gymnosperms as it fosters pollen germination and pollen tube growth, eventually leading to fertilization of an egg cell within gymnosperms' archegonia. Yet, virtually nothing is known about the molecular genetic basis of ovular secretion, an important characteristic in gymnosperms' megagametophytes (Zhang & Zheng, 2016). A recent study on thirteen species representing all five main lineages of extant gymnosperms (Nepi et al., 2017) suggested that oval secretion chemical profiles showing higher levels of carbohydrates and certain amino acids can identify recent or hint at ancestral gymnosperm populations that are or were insect-pollinated, respectively. Pollination drops functioning as gymnosperm nectar have not been chemically analyzed (Von Aderkas, Prior, & Little, 2018), precluding a direct angiosperm–gymnosperm comparison at the moment. However, both gymnosperm ovular secretions and angiosperm floral nectaries have shown selection over time driven by insects and their nutritive needs related to a higher energy metabolism required for flight (Nepi et al., 2017), suggesting a similar adaptive mechanism. In any case, important insect pollinator–plant host shifts took place during angiosperm radiation in mid-Cretaceous, leaving few survivors of previously common gymnosperm–insect pollinators either because of extinctions or insects' transition to angiosperms' plant hosts (Peris et al., 2017). Anemophily could have

also been an ancient event in angiosperms due to its importance in pioneer habitats (Gottsberger, 1988).

Archegonia develop from initial cells within the female gametophyte of the ovule through subsequent rounds of divisions giving rise to (outward) neck cells, and the central cell. This later develops into the large egg cell and the small ventral canal cell, which degenerates as the egg cell matures. Despite its importance, the molecular regulation of archegonia development in the ovule of cone-bearing gymnosperms has not been extensively studied, and the role of archegonia in egg fertilization is largely unknown (Zhang & Zheng, 2016). The major challenge for such studies is the long duration of the pollination process (up to 13 months for pines) compared to the short period of time required for zygote formation. Archegonia were not found to produce pollen-specific signals, but neck cells might produce these (Zhang & Zheng, 2016). Evidence from lower archegoniatae such as ferns and bryophytes suggests auxin-responsive genes might be involved in reproductive organ morphogenesis, differentiation, and cell turnover related to archegonial development (Zhang & Zheng, 2016). Some evidence also hints at a role for the *arabinogalactan protein* (*AGP*) gene family. Glycosylated AGPs could be involved in egg cell–sperm recognition, and their expression within the nucellus cells of the mature egg cell may be involved in promoting the targeted pollen tube growth (AGPs' glycosylation-level dependent; Zhang & Zheng, 2016). *WUSCHEL-related homeobox* (*WOX*) transcription factor genes are implicated in conifer embryo formation and differentiation (patterning) which was tightly linked to polar auxin transport (*ibidem*).

As stated earlier, most genomic resources aimed at in-depth elucidation of the molecular basis of conifers' reproductive developmental stages were developed for *Pinus* (*Pinus tabuliformis*: Du et al., 2017; Niu, Yuan, Zhang, Chen, & Li, 2014; Niu et al., 2015; Niu et al., 2016; *Pinus sylvestris*: Avia, Kärkkäinen, Lagercrantz, & Savolainen, 2014); *Picea* (Carlsbecker et al., 2013; Vázquez-Lobo et al., 2007); and recently for *Cryptomeria japonica* (Katahata, Futamura, Igasaki, & Shinohara, 2014; Tsubomura, Kurita, & Watanabe, 2016). The flower development gene families (*MADS-BOX*, *LFY/NDLY*, phosphatidylethanolamine-binding protein (PEBP) *FLOWERING LOCUS T/TERMINAL FLOWER1*-like genes) are ancient, estimated to have been present before 300 Mya, predating the divergence of seed plants (Liu, Xiong, Li, & Fei, 2018; Moyroud et al., 2017). The first evidence of the importance of *MADS-box* genes for conifer reproductive organs dates back to 2002 (reviewed in Zhang & Zheng, 2016). The *bona fide FT* locus is absent from gymnosperms; therefore, it represents an innovation in the angiosperm lineage (Klintenäs, Pin, Benlloch, Ingvarsson, & Nilsson, 2012). However, cryptic *FT* gene homologs in gymnosperms were identified through phylogenetic reconstruction in three monophyletic clades (*FT-like*, *TFL1-like* and *MOTHER OF FT AND TFL1-like* PEBP genes). None of these cryptic *FT* genes induce flowering (Liu et al., 2018). In gymnosperms, the *TFL1-like* gene shows predominant expression in cone development of both sexes. For spruce, *FTL1* is expressed in male cone development and *FTL2* in female cone development. In addition, gymnosperm-derived *FTL2* genes are implicated in growth rhythm regulation (Liu et al.,

2018). Recent evidence shows an important link between *FT-like* and *MADS-box* gene families, and growth rhythm control, bud set and growth cessation (*FT-like*: e.g., *P. sylvestris*: Avia et al., 2014; *P. abies*: Karlgren, Gyllenstrand, Clapham, & Lagercrantz, 2013; *MADS-box*: Petterle, Karlberg, & Bhalerao, 2013). This suggests early molecular actors in reproductive development and seasonal growth cessation are similar.

Investigating aberrant reproductive structures for the analyses of gender expression in cones can provide important insights into the complete or partial transformation of male cones into female cones (homeosis). Of particular consideration is the molecular study on teratological reproductive structures. Niu et al. (2016) demonstrated that both male and female structures of bisexual cones were functional in *P. tabuliformis* (high pollen grain germination rate, seed producing conelets); however, the transcriptomes of the male structures from the bisexual cones were found to be distinct. The authors identified key regulators among *MADS-box* genes (i.e., *PtDAL1, PtDAL2, PtDAL3, PtDAL4, PtDAL10, PtDAL14, PtMADS1* and *PtMADS3*), as well as the two *LFY/NDLY* genes (the latter usually only expressed in female cones) that may have resulted in the presence of female structures in a male cone. Therefore, the observed sex transformation was the result of stark shifts in their gene expression patterns. Because male-to-female transformation in these bisexual cones was incomplete, the male-to-female homeosis did not involve a reduction in *PtDAL11* and *PtDAL13* expression, providing support that these two genes are essential to maintain the ancestral function in sex determination (male cone) and specify the pollen-bearing organs generated by this meristem (Niu et al., 2016). Another seed cone developmental mutant termed *P. abies acrocona*, characterized by early cone setting and homeotic transformations of vegetative shoots into female cones, revealed that gymnosperm *LFY/NDLY* genes are involved in reproductive initiation. This observation was based on the exceptionally high *PaNDLY* expression in the axis of the *acrocona* shoot at the transition from vegetative to female identity (Carlsbecker et al., 2013).

Finally, a single fertilization event within the ovule produces a diploid embryo that develops within a haploid female gametophyte. At early seed development, polyembryony is also an important reproductive feature in conifers, whereby multiple archegonia can be fertilized by different pollen grains. In all cases though, only the dominant embryo persists and matures while all others are aborted. The molecular basis of embryo persistence is unknown (Cairney & Pullman, 2007). The embryo suspensor stage is a critical stage in early embryonic development as it helps the embryo to grow within the female gametophyte, and to benefit from nutrient absorption while it enlarges. Gymnosperms contain genes of very similar sequence to angiosperm embryogenesis-regulating genes (Cairney & Pullman, 2007). The comparative synthesis by Cairney and Pullman revealed that in gymnosperm embryogenesis, subtle molecular interactions, spatially and temporally controlled gene expression, and few unique regulatory proteins can achieve differences in embryonic structure and development. One important example is the above-mentioned *WOX* transcription factor genes.

Recently, transcriptomic studies on embryogenesis in pines (*P. sylvestris*; *Pinus pinaster*; *P. lambertiana*; *Araucaria angustifolia*) and spruces (*P. abies*; *Picea balfouriana*) have been published (reviewed in Rodrigues et al., 2018). Nevertheless, in order to better understand gymnosperms' unique regulatory networks, any functional analysis of conifer developmental genes must be conducted by expressing these genes in a conifer. Therefore, the development of a robust, easy-to-use and broadly applicable transformation system for conifers constitutes a prerequisite to a better understanding of several aspects of this phylum's cell and molecular biology (Cairney & Pullman, 2007). Up until now, this has not been achieved. More recent studies revealed a crucial role for small noncoding RNAs, and some of their target genes were revealed for the regulation of seed development and in embryo development (Rodrigues & Miguel, 2017). Niu et al. (2015) identified such sRNAs specifically for male and female cones of *P. tabuliformis*, and with higher activities in the female than in the male reproductive structures. The miR156-SPLs, miR159-MYBs, miR172- AP2Ls, miR319-TCP, and miR396-GRFs interacting pairs found for this pine species coincided with those in angiosperms' reproductive development, suggesting ancient evolutionary histories of these sRNA regulatory pathways (Niu et al., 2015).

4 | WATER-CONDUCTING XYLEM TISSUES

4.1 | Cellulose/hemicelluloses synthases and their regulation

Cellulose, composed by a $(1 \rightarrow 4)$-β-d-glucan chain, is one of the most important components of the plant cell wall and is also a decisive factor controlling plant cell mechanical properties (Kumar, Atanassov, & Turner, 2017; Sarkar, Bosneaga, & Auer, 2009). The evolution of cellulose enhanced plant cell adaptation ability to respond to changing environments (Sarkar et al., 2009). Cellulose is biosynthesized by cellulose synthases (CesA) at the plasma membrane (Meents, Watanabe, & Samuels, 2018). The *CesA* family belongs to the *CesA* superfamily including *CesA-like* (*Csl*) gene families which harbor nine *Csl* subfamilies (*Csl*A/B/C/D/E/F/G/H/J) (Yin, Johns, Cao, & Rupani, 2014). Csl subfamily proteins catalyze the biosynthesis of hemicelluloses, another polysaccharide component in the plant cell wall with a backbone similar to cellulose (Cosgrove, 2005; Suzuki, Li, Sun, & Chiang, 2006). The *CesA* gene number in gymnosperms is similar to the one in angiosperms. For example, eight *CesA* genes were found in *Pinus radiata* (Krauskopf, Harris, & Putterill, 2005), 17 *CesA*s in *Cunninghamia lanceolata* (Huang et al., 2012), six in *P. taeda* (Neale et al., 2014), and nine *CesA*s in *G. montanum* (Wan et al., 2018), in comparison with 10 and 18 *CesA* genes in *Arabidopsis* and *Populus trichocarpa* genomes, respectively (Suzuki et al., 2006). Recently, in the *G. montanum* genome reference paper, it was suggested that large expansions in the *Csl*B/H subfamilies may explain the distinct growth characteristics in *Gnetum* when compared to other gymnosperms (Wan et al., 2018). It is interesting that *Csl*B/E/H/G that evolved from ancestral genes in ferns were lost in many

gymnosperms, such as *P. abies*, *P. taeda*, *G. biloba*, and other species (Yin et al., 2014).

Cellulose and hemicellulose biosyntheses are regulated at the transcriptional level (Li, Bashline, Lei, & Gu, 2014). In angiosperms, for example, at least 13 out of 126 MYB transcription factors were reported to be involved in cellulose formation by regulating *CesA/Csl* gene expression directly or indirectly in *Arabidopsis* (Zhang, Nieminen, Serra, & Helariutta, 2014). However, in gymnosperms, only 13 *Picea glauca* and five *P. taeda MYB* genes were identified, suggesting a much lower number of MYB genes than in *Arabidopsis* and *Populus* (Bedon, Grima-Pettenati, & Mackay, 2007). Some gymnosperm MYB genes, which have conserved functions (e.g., *PtMYB1* and *PtMYB4* in *P. taeda*), are expressed in the secondary xylem and involved in lignin biosynthesis as their homolog in *Arabidopsis* (Bedon et al., 2007; Patzlaff et al., 2003). Whether cellulose biosynthesis is regulated by MYB transcription factors is not clear in gymnosperms. However, the *CesA* genes' regulation network in gymnosperms might be less complex than in angiosperms. Cellulose biosynthesis is also affected by the content of lignin, another component of the plant cell wall (Endler & Persson, 2011). In *Populus*, artificial lignin biosynthesis inhibition is coupled with cellulose production and higher growth, suggesting cellulose synthase activity is restricted by substrate content (Hu et al., 1999). In *P. taeda*, spontaneous mutations in lignin biosynthesis (Songstad, Petolino, Voytas, & Reichert, 2017) also caused fast stem growth, suggesting cellulose synthase activity may be naturally regulated by lignin content in gymnosperms (Gill, Brown, & Neale, 2003).

4.2 | Vascular NAC domain

The difference between water-conducting xylem tissues (tracheids vs. vessels) is one of the main differences between gymnosperms and angiosperms (Wan et al., 2018). Tracheids, whose dual function is water transport and mechanical support, constitute the xylem tissue in gymnosperms. In angiosperms, xylem tissue is more complex and consists of vessels, fibers, and rays (Patten, Vassão, Wolcott, Davin, & Lewis, 2010, 2010). In angiosperms, vessel cells are differentiated from cambium cells that undergo secondary cell wall biosynthesis and programmed cell death (Zhang et al., 2014). VASCULAR-RELATED NAC-DOMAIN6 (VND6) and VND7 are the key transcription switches on vessel element formation in *Arabidopsis* via activation of the transcription cascade involved in secondary cell wall biosynthesis and programmed cell death (Kubo et al., 2005; Ohashi-Ito & Fukuda, 2014; Zhong, Lee, & Ye, 2010). There are seven *VND* genes encoded in the *Arabidopsis* genome, and *VND1-5* was also recently reported to be involved in vessel element formation (Tan et al., 2018; Zhou, Zhong, & Ye, 2014). All of the seven *VND* genes in *Arabidopsis* were specifically expressed in vessels and had conserved downstream targets controlling vessel formation (Zhong, Lee, Zhou, McCarthy, & Ye, 2008; Zhou et al., 2014). It is interesting that the *P. abies* and *G. montanum* genomes only encode two and one *VND* ortholog genes, respectively, and both of them are homologous to *VND4/5/6* (Nystedt et al., 2013; Wan et al., 2018). There are two

possible explanations for why vessels are absent in gymnosperms. In the first one, *VND1–3* and *VND7* may determine function in vessel formation (Wan et al., 2018), and at least, the dominant repression of VND7 showed a more severe phenotype than the dominant repression of VND6 (Kubo et al., 2005). In the second one, vessel formation requires VND gene expansion and their co-expression (Nystedt et al., 2013). Although the seven VNDs in *Arabidopsis* had conserved expression patterns and downstream genes, the expression level in vessels of different organs and activation strength were different (Zhou et al., 2014), suggesting the seven VNDs might coordinately work to regulate vessel formation.

5 | SECONDARY METABOLISM AND STRESS-RELATED GENE FAMILIES

5.1 | Abiotic stress—Dehydrins

Dehydrins are a group of proteins belonging to the late embryogenesis abundant (LEA) gene family that are highly hydrophilic and are commonly associated with acclimation to low temperature and other environmental stresses involving cellular dehydration in plants (Rorat, 2006). Dehydrins have been shown to be related to drought tolerance (Hu, Wang, Du, & Huang, 2010; Lopez, Banowetz, Peterson, & Kronstad, 2003; Suprunova et al., 2004) and low temperature acclimation (Danyluk et al., 1998; Gao & Lan, 2016; Strimbeck, Schaberg, Fossdal, Schröder, & Kjellsen, 2015) in several species. One study on *Picea obovata* showed a dehydrin accumulated to ~16× its initial level during acclimation from moderate to extreme low temperature tolerance (Kjellsen, Shiryaeva, Schröder, & Strimbeck, 2010). Multiple studies demonstrate a similar relationship in other taxonomic groups, suggesting a strong association between low temperature acclimation and accumulation of dehydrins across taxonomic groups (Arora & Wisniewski, 1994; Kontunen-Soppela & Laine, 2001; Liu et al., 2004; Renaut, Hoffmann, & Hausman, 2005; Rinne, Welling, & Kaikuranta, 1998). Furthermore, significant differentiation in allelic frequency has been observed at three dehydrin-associated loci between populations of *P. sylvestris* L. showing divergence for cold tolerance (Wachowiak, Balk, & Savolainen, 2009). However, their data suggest that nucleotide polymorphism in most *P. sylvestris* dehydrins cannot be directly related to adaptive variation in cold tolerance (Wachowiak et al., 2009).

The specific mode of action of dehydrins is unclear, but some studies suggest that dehydrins stabilize membranes and macromolecules in conditions of low water availability (Hanin et al., 2011). The size of the dehydrin gene family is highly variable ranging from two members in *Amborella* to more than 12 in *Malus domesticus* in angiosperms. Gymnosperms are less studied, but within Pinaceae, the dehydrin family appears to be much larger relative to angiosperms, with a total of 53 having been identified in *P. glauca* (Stival Sena, Giguère, Rigault, Bousquet, & Mackay, 2018). Subfunctionalization is thought to be the primary driver for the increased diversity of dehydrins in conifers over angiosperms (Stival Sena et al., 2018). In contrast, extant species of *Gnetum* have reduced numbers of *LEA*

genes (and dehydrins) when compared to other gymnosperms (Wan et al., 2018). *Gnetum* also differs from other gymnosperms in that it only exists in warm, mesic habitats (Wan et al., 2018), lending more evidence to the role dehydrins play in adaptation to water stress.

5.2 | Defense systems

Conifer defenses against pests and pathogens involve many different gene families, and many of them have been well-studied in terms of their occurrence within the genome (Warren et al., 2015; Zhou et al., 2019). Some of these include biosynthetic enzymes like the ones acting in oxygenation, phenoxy radical coupling or regio- or stereo-selective reactions resulting in an immense diversity of defense compounds (for induced defenses see: Keeling & Bohlmann, 2006; Kovalchuk et al., 2015; Oliva et al., 2015; Ralph, Yueh, et al., 2006; Ralph, Park, Bohlmann, & Mansfield, 2006; Ralph, Jancsik, & Bohlmann, 2007; Visser, Wegrzyn, Myburg, & Naidoo, 2018; for constitutive defenses see: Keeling & Bohlmann, 2006; Ralph et al., 2007; Porth et al., 2011; Porth et al., 2012; Porth et al., 2018). Regulatory genes include the Sg4C R2R3-MYB transcription factor that exhibits a significant gene family expansion in conifers (Bedon et al., 2010). Moreover, crucial biosynthetic genes for pest resistance (e.g., 3CAR; CYP720B4) feature high content of repetitive sequence regions and transposable elements, suggesting that diversification of the conifer TPS and P450 gene families may have been achieved by DNA transposon-mediated translocation mechanisms (Hamberger et al., 2009). Another important feature of conifer TPSs is their high potential for functional plasticity such that few changes in amino acids can create new potent defense molecules (Keeling, Weisshaar, Lin, & Bohlmann, 2008).

Because plants have a long evolutionary history of interaction with herbivores, hosts have acquired coevolved defenses (Futuyma & Agrawal, 2009). A special case is the gymnosperm *G. biloba*, which is largely herbivore-free. Ginkgo's foliage produces ginkgolides, a class of terpene trilactones known as a potent antifeeding defense (Mohanta et al., 2012; Pan, Ren, Chen, Feng, & Luo, 2016). In general, the most effective host tree defenses exist against local pests and pathogens, while host defenses weaken under relaxed or absent pathogen pressure. This is a recurrent problem with introduced foreign pest and pathogens, but also with native pests and pathogens expanding their natural ranges. As climate warms, these native species may expand their ranges northwards or to higher altitudes, where they may encounter "naïve" hosts. Moreover, native species may change their metabolism to support a more aggressive behavior, leading to unprecedented population growth and range expansions, and threatening local and new host trees in a pest's newly invaded habitat. A widely publicized example of current range expansion is the mountain pine beetle (*Dendroctonus ponderosae* Hopkins). This pest epidemic in western North America is now threatening the boreal forest (Cullingham et al., 2011).

Trees have developed different lines of defense that are more or less effective, and also alternative strategies such as tolerance. Anatomical and the associated chemical defenses in conifer bark

have been described (Franceschi, Krokene, Christiansen, & Krekling, 2005). Strength and rapidity of traumatic resinosis (direct defense) has often been associated with resistance. The physical structures studied in most detail are the parenchyma cells (locations of synthesis and storage of polyphenols), and the resin ducts (synthesis and storage of terpenes) that are located in the secondary phloem and the cambium. The traumatic resin canals are formed in the secondary xylem as a way of active defense. Upon attack, reallocation of resources from primary processes to active defense, or the mobilization of the resources for host tolerance, takes place. Indirect tree defense responses that involve the attraction of predators or herbivore parasitoids have also been documented. Moreover, trade-offs involving defense strategies involve display of chemical defenses, or rely on tolerance (Futuyma & Agrawal, 2009). In a recent study on the genomics of host defenses against the spruce shoot weevil (*Pissodes strobi* Peck), Porth et al. (2018) concluded that well-established terpenoid-related spruce defenses and tolerance to this herbivore might be mutually exclusive.

It has been postulated that drought-stressed conifers whose metabolism is diverted from growth to secondary compounds can rely more on constitutive, preformed defenses (Turtola, Manninen, Rikala, & Kainulainen, 2003). Also, it is well known that fast growing individuals are biased toward induced defenses (Steppuhn & Baldwin, 2008). Therefore, trade-offs between already established and induced defenses can be expected. These dynamics under different environmental conditions need to be better studied in the future, while current genomic studies usually represent a snap-shot situation aiming to identify few highly upregulated candidate genes from well-annotated conifer defense metabolic pathways such as the phenylpropanoid and methylerythritol phosphate/mevalonate (Hall, Yuen, et al., 2013a; Keeling et al., 2011; Porth et al., 2011; Shalev et al., 2018; Warren et al., 2015; Zhou et al., 2019). In addition, the genetic networks between defenses in conifers and their reproductive development seem to be intricate. With few exceptions, this important relationship has been largely ignored in conifer defense studies, mainly because the conifer reproductive genes (many are also gene family members) were under-studied; thus, their exact functioning remains elusive (see section on Reproductive Biology). In any case, it is known that certain signaling pathways (jasmonate, ethylene, auxin, gibberellin) required for developmental processes (such as those important in reproduction) can be co-opted for biotic stress responses (e.g., Du et al., 2017; Oliva et al., 2015; Thaler, Farag, Paré, & Dicke, 2002; Zi, Mafu, & Peters, 2014). Alternatively, these signaling pathways may evolve into new specialized pathways such as the conifer defensive resin production co-opted from gibberellin production (Zi et al., 2014).

Given the current knowledge about defensive gene family expansion in gymnosperms (Porth et al., 2011, 2012; Warren et al., 2015; Zhou et al., 2019), the challenge remains to identify the most potent defensive metabolites against herbivory or disease produced in these pathways. Here, we show examples of natural host defense compounds with proven dramatic negative impacts on pest development in *Picea*. Delvas, Bauce, Labbé, Ollevier, and Bélanger
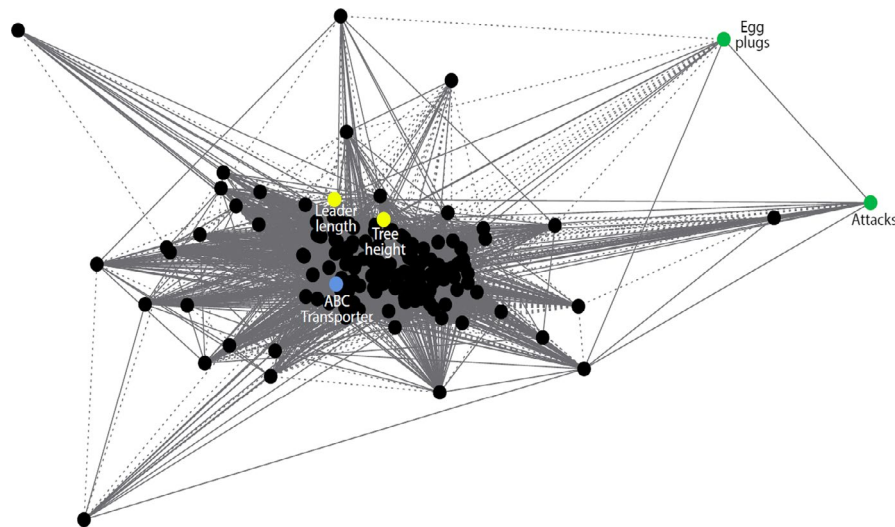
FIGURE 2 The white spruce PDR gene family member *WS0269_K02* identified as a core gene. Spruce PDR gene (ABC transporter, blue dot) identified as "core gene" (Porth et al., 2018) in the gene regulatory network with growth (yellow dots) or defense phenotypes (against the stem-boring spruce shoot weevil *Pissodes strobi*; green dots)

(2011) identified acetophenones that act against the budworm *Choristoneura fumiferana*. Robert et al., 2010 showed 3-carene and dehydroabietic acid acting against the weevil *P. strobi*. Moreover, (+) catechin was found effective against fungal pathogenicity in the form of *Heterobasidion parviporum* (Nemesio Gorriz et al., 2016). The work by Liu and Ekramoddoullah (2007, 2011) showed CC-NBS-LRR and TIR-NBS-LRR expression conveying resistance against white pine blister rust (caused by *Cronartium ribicola*) in western white pine.

For resistance breeding purposes, knowledge about defense metabolites' heritability (i.e., the extent of their genetic control) in the breeding population is required. The work by Méndez-Espinoza et al. (2018) on acetophenones' genetic parameters remains the only work on this aspect to date. In the case of genetic marker-assisted resistance breeding, it is important to ascertain the underlying genetic regulation for compounds of interest. This information is only available for a few metabolites. For example, Roach, Hall, Zerbe, and Bohlmann (2014) identified the related 3-carene synthase, 2-sabinene TPSs. Hamberger, Ohnishi, Hamberger, Seguin, and Bohlmann (2011) found the specific cytochrome P450 of the CYP720B family implicated in dehydroabietic acid synthesis. Mageroy et al. (2015) found β-glucosidase, the biosynthetic gene for acetophenones. Finally, Nemesio Gorriz et al. (2016) isolated a leucoanthocyanidin reductase for (+) catechin generation. Other studies have focused on identifying the biosynthesis locations for terpenoids (Abbott, Hall, Hamberger, & Bohlmann, 2010; Zulak & Bohlmann, 2010) and phenolic (Li et al., 2012) compounds to better target effective tree defenses in the future.

5.3 | A case study of functional pleiotropy with defense: the PDR ABC transporter family

Pleiotropic drug resistance (*PDR*) genes belong to a fungi and plant-specific gene family within the *ATP Binding Cassette (ABC)* gene superfamily (Crouzet, Trombik, Fraysse, & Boutry, 2006; Higgins, 1992; Lamping et al., 2010). The *PDR* gene family was named following the observation that members of its family confer resistance to various drugs; however, *PDR* genes are also involved in the transport of substrates not related to cell detoxification (Ito & Gray, 2006; Nuruzzaman, Zhang, Cao, & Luo, 2014; Pierman et al., 2017; Sasse et al., 2016). Three recent and completely independent studies on two spruces (*P. glauca*; *P. glauca × engelmannii*) and *P. taeda* are suggesting specific *PDR* genes as important key players in defense mechanisms against different herbivores (Mageroy et al., 2015; Porth et al., 2018) and pathogens (De la Torre et al., 2018). For example, research on spruce budworm (*C. fumiferana*) resistance identified gene *WS0269_K02* with high statistical support for its expression upregulation in budworm resistant versus nonresistant white spruces (Mageroy et al., 2015; information drawn from their Table S1). The same *WS0269_K02* gene was found in spruce shoot weevil (*P. strobi*) resistance (Porth et al., 2018, Figure 2). In pine, a closely related gene family member was identified for pitch canker disease (*Fusarium circinatum*) resistance (De la Torre et al., 2018). Because these genes' expressions were also correlated with drought resistance (De la Torre et al., 2018) and growth rate (Porth et al., 2018), genetic pleiotropic functioning of conifer *PDR* genes could be implied. Drought resistance and growth might share a genetic relationship to a certain extent, as trees impaired in drought tolerance and succumbing to drought stress are expected to show decreased growth (Salmon et al., 2019). It has further been postulated that drought-stressed conifers rely more on constitutive than on induced defenses (Turtola et al., 2003).

Our study found that the size of the PDR family in gymnosperms is smaller compared to angiosperms. This may indicate that gymnosperm species require less PDR transporters than angiosperms to cope with their environment. The identified conifer *PDR* gene sequences were further mapped to the *PDR* genes' phylogenetic tree for improved annotations (Figure 3). In the case of the white spruce

FIGURE 3 Phylogenetic analysis of vascular plant PDR protein sequences obtained by maximum likelihood. Protein sequences were arbitrarily grouped into nine clusters (I–IX) based on close sequence similarity. Red branches are leading to PDR sequences belonging to gymnosperm sequences. Sequence names have been pruned from the tree for the sake of clarity. Sequence names belonging to each cluster can be found in Table S1. Node support from 1,000 replicates is indicated for the basal nodes defining the nine putative PDR sequence clusters. For further details, see Appendix S2
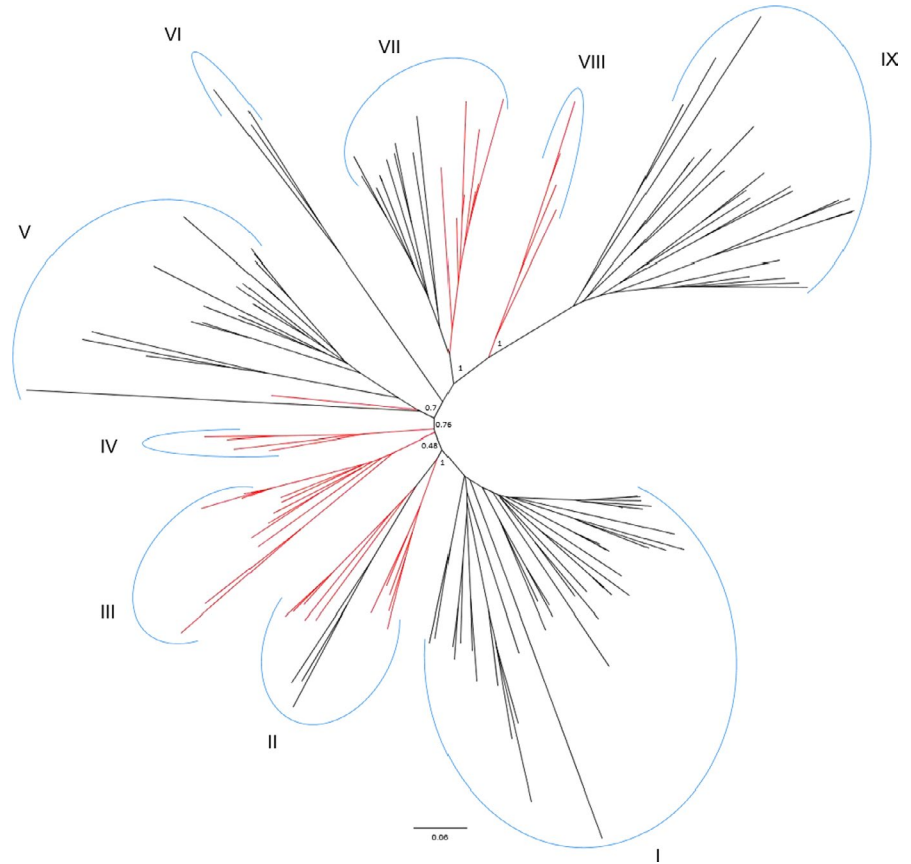
gene (identified by Mageroy et al., 2015 and Porth et al., 2018), *WS0269_K02* mapped to cluster IV, a gymnosperm-specific clade, and it was found to be putatively identical to the *P. abies* gene *Pab_MA_17319g0010*, Table S1. In the case of the *P. taeda* gene (c3387/f1p0/2274 identified in De la Torre et al., 2018), its transcript mapped to cluster II on the phylogenetic tree (sister to a group of the two conifer sequences Pab_MA_10427561g0010 and Pta_04241). Two hypotheses can be proposed to explain the diversification of the *PDR* gene family. First, *PDR* genes might have diversified by acquiring new physiological roles (neofunctionalization). Different plant species produce different metabolites, and diversification is expected in enzymes involved in the transport of these metabolites (Yazaki, 2006). Secondly, the differential expression of *PDR* genes in different tissues or during different developmental stages might have promoted their diversification (subfunctionalization). To fully grasp the evolution of the *PDR* gene family, more *PDR* gene sequences from additional species across the plant kingdom are needed to better resolve *PDR* gene evolution and relationships (this was beyond the scope of the present study).

6 | NONCODING AND SMALL RNAS

Noncoding RNAs are a class of RNAs not involved in protein coding, but with very important functions as regulators in plant life cycle, response to the environment, and phenotypic plasticity (Borges &

Martienssen, 2015; Shin & Shin, 2016). Noncoding RNAs can be divided into two categories, long noncoding RNAs (>200 nucleotides, nt) and small noncoding RNAs (sRNA) (20–24 nt) (Arikit, Zhai, & Meyers, 2013). Differences in sRNA size distribution can be observed between gymnosperms and angiosperms. The 21-nt sRNAs are dominant in gymnosperms such as *P. abies* (Nystedt et al., 2013), *Pinus contorta* (Morin et al., 2008), *Larix leptolepis* (Zhang et al., 2013), and *P. tabuliformis* (Niu et al., 2015), whereas 24-nt sRNAs represent the majority in angiosperms (Li et al., 2017; Morin et al., 2008). For a long time, 24-nt sRNAs were thought to be absent from gymnosperms, and now, we know they occur at low frequencies and are mainly restricted to reproductive tissues (Niu et al., 2015; Nystedt et al., 2013; Zhang et al., 2013). Therefore, the presence of 24-nt sRNAs may be important in the regulation of reproduction in gymnosperms (Niu et al., 2015). Because 21-nt sRNAs are involved in target gene silencing or protein translation inhibition and 24-nt sRNAs are functional on chromatin remodeling (Borges & Martienssen, 2015), it seems that sRNAs may play different regulatory roles in gymnosperm and angiosperm development, respectively.

In plants, the biogenesis of sRNA is dependent on dsRNA precursor cleavage mediated by DICER ribonucleases (DCL; Van Ex, Jacob, & Martienssen, 2011). DCL1 and DCL4 generate 21-nt sRNAs, while DCL3 generates 24-nt sRNAs. It is interesting that short variants of DCL1 were characterized in *P. lambertiana* (Gonzalez-Ibeas et al., 2016) and in the bryophyte *Physcomitrella patens* (Coruh, Shahid, & Axtell, 2014). This finding suggests that functional truncated

DCL1 might be the reason why 21-nt sRNAs are dominant in gymnosperms, although the relationship between conifer-specific 21-nt sRNA and short DCL1 is unclear (Gonzalez-Ibeas et al., 2016). A conifer-specific set of DCL1 proteins was found in *P. glauca, P. abies,* and *P. lambertiana* (Dolgosheina et al., 2008; Gonzalez-Ibeas et al., 2016). DCL3, which is involved in 24-nt sRNAs biogenesis, was characterized through *P. lambertiana* transcripts, primarily expressed in reproductive tissues (Gonzalez-Ibeas et al., 2016). Truncated DCL3 was also discovered in *P. glauca,* and its expression level upregulated in seed development indicated that the DCL3 variant and its expression level are responsible for 24-nt sRNA generation in *P. glauca* (Liu & El-Kassaby, 2017). The discovery of variant DCL partly explained the different sRNA size distribution between gymnosperms and angiosperm, although further confirmation is needed. The 24-nt sRNAs direct DNA methylation and affect histone modification which are related to chromatin condensation and silencing of transposable elements (Leitch & Leitch, 2012). The different silencing mechanisms were correlated with differences in genome sizes of angiosperms and gymnosperms (Dolgosheina et al., 2008; Leitch & Leitch, 2012).

7 | APPLICATIONS OF THE STUDY OF GYMNOSPERM GENE FAMILIES

Plant defense molecules are highly complex traits with nutritional value, flavor, and use in traditional medicine (Hamberger & Bak, 2013). Genes encoding natural product pathways often group together in biosynthetic gene clusters (Nützmann, Huang, & Osbourn, 2016). Some of the genes reviewed in this study are newly studied members of gene families that hold great potential for biotechnological applications related to commercial and pharmacological value. The plant-based dirigents (Pickel & Schaller, 2013), cytochrome P450s (Hamberger & Bak, 2013; Renault, Bassard, Hamberger, & Werck-Reichhart, 2014), and terpenoids (Bohlmann & Keeling, 2008; Singh & Sharma, 2015) have been recognized to hold potential for biotechnology. Because of the large metabolic diversity arising from the multitude of biochemical reactions for these gene families' members (P450s: Hamberger & Bak, 2013; TPSs: Boutanaev et al., 2015; Chen et al., 2011; Chen et al., 2018; conifer TPSs: Hall, Zerbe, et al., 2013b; Zerbe et al., 2012; dirigents: Pickel & Schaller, 2013), new plant natural product pathways are likely to be discovered and functionally described. Perhaps the most famous compound of terpenoid origin for human uses is taxol, a potent anticancer drug (Wani, Taylor, Wall, Coggon, & McPhail, 1971) and whose biosynthesis in *Taxus* spp. has been elucidated (Croteau, Ketchum, Long, Kaspera, & Wildung, 2006). Pseudolaric acid B, a diterpene acid (originating from *Pseudolarix kae*mpferi bark), is another well-known drug that reduces tumor growth, in particular for melanoma (Gong, Wang, Tashiro, Onodera, & Ikejima, 2005). With ~50,000 different molecules identified in extant plants, terpenoids are structurally and functionally the most diverse plant metabolic group. They are of substantial commercial and pharmacological value as essential oils, fragrances, colorants, drugs, coatings, and speciality plastics (Sainz

et al., 2016; Vranová, Coman, & Gruissem, 2012). For example, conifer terpenoid oleoresins are used by the naval stores industry to create rosin and turpentine used as adhesives, inks, solvents, and resins. Although their production was reduced due to the increase of less expensive petroleum-derived substitutes, terpenoid oleoresin might come back as an important source of "green" biofuels and bioproducts (Turner, Parrish, Zager, Fischedick, & Lange, 2018). Environmental and developmental factors affect the terpenoid pathway flux; understanding the complexity of the terpenoid pathway network in plants and its regulation remains a major challenge in terpenoid research but will facilitate future molecular breeding of agronomically useful traits (Vranová et al., 2012).

Some members of conifer gene families (such as the *PDR* gene family) can also be exploited for their potential to improve conifer tree growth on marginal or disturbed soils, thus providing an improved detoxification potential to employ conifers (i.e., spruces) in phytoremediation applications. In addition, functional characterization of *PDR* genes is required before biotechnology applications can be performed on the *PDR* gene family, particularly for long-lived trees (Lefevre, Baijot, & Boutry, 2015). Because PDRs have been shown to act in a variety of plant organs, above ground (foliage and reproductive structures) and below ground (in roots; Crouzet et al., 2006), one of the most intriguing applications besides phytoremediation is the PDR's potential in conferring improved resistance to biotic stressors (De la Torre et al., 2018; Mageroy et al., 2015; Porth et al., 2018). Also, a better knowledge of the genes and gene families conferring phenotypic variation is the first step to create plantations with improved varieties through marker-assisted breeding, genomic selection, or genetic modifications (CRISPR). For species with ecological importance, the identification of genes families involved in abiotic and biotic stress may contribute to identify species that are candidates to ecological restoration, or that may present increased potential to adapt to specific or changing climatic conditions.

8 | CONCLUSIONS

In this paper, we aim to understand how genes and gene families have contributed to the evolution of major functional differences in gymnosperms in comparison with its sister plant clade of flowering plants. Recently developed new reference genomes, transcriptomes, and genome-wide resources in gymnosperms have enabled large-scale comparisons of functional divergence within gymnosperms, and between angiosperms and gymnosperms. Information about the genomic architecture underlying phenotypic variation is key for any applied breeding and management of commercially important gymnosperm species. With the development of new genomic tools and analytical software, future approaches will include a higher contiguity of reference genomes; completion of structural and functional annotation of reference genomes; increase in the number and density of physical, linkage, or genetic maps; whole-genome re-sequencing of populations for GWAS studies; and genetic improvements through biotechnology.

ACKNOWLEDGEMENTS

DATA ARCHIVING STATEMENT

This manuscript has no associated data for data archiving.

ORCID

Amanda R. De La Torre https://orcid.org/0000-0001-6647-723X

REFERENCES

Abbott, E., Hall, D., Hamberger, B., & Bohlmann, J. (2010). Laser microdissection of conifer stem tissues: Isolation and analysis of high quality RNA, terpene synthase enzyme activity and terpenoid metabolites from resin ducts and cambial zone tissue of white spruce (*Picea glauca*). *BMC Plant Biology*, *10*(1), 106. https://doi.org/10.1186/1471-2229-10-106

Arikit, S., Zhai, J., & Meyers, B. C. (2013). Biogenesis and function of rice small RNAs from non-coding RNA precursors. *Current Opinion in Plant Biology*, *16*(2), 170–179. https://doi.org/10.1016/j.pbi.2013.01.006

Armenise, L., Simeone, M., Piredda, R., & Schirone, B. (2012). Validation of DNA barcoding as an efficient tool for taxon identification and detection of species diversity in Italian conifers. *European Journal of Forest Research*, *131*, 1337–1353. https://doi.org/10.1007/s10342-012-0602-0

Arora, R., & Wisniewski, M. E. (1994). Cold acclimation in genetically related (sibling) deciduous and evergreen peach (*Prunus persica* [L.] Batsch) (II. A 60-kilodalton bark protein in cold-acclimated tissues of peach is heat stable and related to the dehydrin family of proteins). *Plant Physiology*, *105*(1), 95–101. https://doi.org/10.1104/pp.105.1.95

Avia, K., Kärkkäinen, K., Lagercrantz, U., & Savolainen, O. (2014). Association of *FLOWERING LOCUS T/TERMINAL FLOWER 1-* like gene *FTL2* expression with growth rhythm in Scots pine (*Pinus sylvestris*). *New Phytologist*, *204*(1), 159–170. https://doi.org/10.1111/nph.12901

Bagal, U. R., Leebens-Mack, J. H., Lorenz, W. W., & Dean, J. F. (2012). The phenylalanine ammonia lyase (PAL) gene family shows a gymnosperm-specific lineage. *BMC Genomics*, *13*(3), S1. https://doi.org/10.1186/1471-2164-13-s3-s1

Bedon, F., Bomal, C., Caron, S., Levasseur, C., Boyle, B., Mansfield, S. D., … MacKay, J. (2010). Subgroup 4 R2R3-MYBs in conifer trees: Gene family expansion and contribution to the isoprenoid- and flavonoid-oriented responses. *Journal of Experimental Botany*, *61*(14), 3847–3864. https://doi.org/10.1093/jxb/erq196

Bedon, F., Grima-Pettenati, J., & Mackay, J. (2007). Conifer R2R3-MYB transcription factors: Sequence analyses and gene expression in wood-forming tissues of white spruce (*Picea glauca*). *BMC Plant Biology*, *7*(1), 17. https://doi.org/10.1186/1471-2229-7-17

Bohlmann, J., & Keeling, C. I. (2008). Terpenoid biomaterials. *The Plant Journal*, *54*(4), 656–669. https://doi.org/10.1111/j.1365-313X.2008.03449.x

Borges, F., & Martienssen, R. A. (2015). The expanding world of small RNAs in plants. *Nature Reviews Molecular Cell Biology*, *16*(12), 727–741. https://doi.org/10.1038/nrm4085

Boutanaev, A. M., Moses, T., Zi, J., Nelson, D. R., Mugford, S. T., Peters, R. J., & Osbourn, A. (2015). Investigation of terpene diversification across multiple sequenced plant genomes. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(1), E81–E88. https://doi.org/10.1073/pnas.1419547112

Cairney, J., & Pullman, G. S. (2007). The cellular and molecular biology of conifer embryogenesis. *New Phytologist*, *176*(3), 511–536. https://doi.org/10.1111/j.1469-8137.2007.02239.x

Carlsbecker, A., Sundström, J. F., Englund, M., Uddenberg, D., Izquierdo, L., Kvarnheden, A., … Engström, P. (2013). Molecular control of normal and *acrocona* mutant seed cone development in Norway spruce (*Picea abies*) and the evolution of conifer ovule-bearing organs. *New Phytologist*, *200*(1), 261–275. https://doi.org/10.1111/nph.12360

Casola, C., & Koralewski, T. E. (2018). Pinaceae show elevated rates of gene turnover that are robust to incomplete gene annotation. *The Plant Journal*, *95*(5), 862–876. https://doi.org/10.1111/tpj.13994

Chen, F., Ludwiczuk, A., Wei, G., Chen, X., Crandall-Stotler, B., & Bowman, J. L. (2018). Terpenoid secondary metabolites in bryophytes: Chemical diversity, biosynthesis and biological functions. *Critical Reviews in Plant Sciences*, *37*, 210–231. https://doi.org/10.1080/07352689.2018.1482397

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. *The Plant Journal*, *66*(1), 212–229. https://doi.org/10.1111/j.1365-313X.2011.04520.x

Coruh, C., Shahid, S., & Axtell, M. J. (2014). Seeing the forest for the trees: Annotating small RNA producing genes in plants. *Current Opinion in Plant Biology*, *18*, 87–95. https://doi.org/10.1016/j.pbi.2014.02.008

Cosgrove, D. J. (2005). Growth of the plant cell wall. *Nature Reviews Molecular Cell Biology*, *6*(11), 850–861. https://doi.org/10.1038/nrm1746

Crisp, M. D., & Cook, L. G. (2011). Cenozoic extinctions account for the low diversity of extant gymnosperms compared with angiosperms. *New Phytologist*, *192*(4), 997–1009. https://doi.org/10.1111/j.1469-8137.2011.03862.x

Croteau, R., Ketchum, R. E. B., Long, R. M., Kaspera, R., & Wildung, M. R. (2006). Taxol biosynthesis and molecular genetics. *Phytochemistry Reviews*, *5*, 75–97. https://doi.org/10.1007/s11101-005-3748-2

Crouzet, J., Trombik, T., Fraysse, A. S., & Boutry, M. (2006). Organization and function of the plant pleiotropic drug resistance ABC transporter family. *FEBS Letters*, *580*(4), 1123–1130. https://doi.org/10.1016/j.febslet.2005.12.043

Cullingham, C. I., Cooke, J. E. K., Dang, S., Davis, C. S., Cooke, B. J., & Coltman, D. W. (2011). Mountain pine beetle host-range expansion threatens the boreal forest. *Molecular Ecology*, *20*(10), 2157–2171. https://doi.org/10.1111/j.1365-294X.2011.05086.x

Danyluk, J., Perron, A., Houde, M., Limin, A., Fowler, B., Benhamou, N., & Sarhan, F. (1998). Accumulation of an acidic dehydrin in the vicinity of the plasma membrane during cold acclimation of wheat. *The Plant Cell*, *10*(4), 623–638. https://doi.org/10.1105/tpc.10.4.623

De La Torre, A. R., Birol, I., Bousquet, J., Ingvarsson, P. K., Jansson, S., Jones, S. J. M., … Bohlmann, J. (2014). Insights into conifer giga-genomes. *Plant Physiology*, *166*(4), 1724–1732. https://doi.org/10.1104/pp.114.248708

De La Torre, A. R., Li, Z., Van de Peer, Y., & Ingvarsson, P. K. (2017). Contrasting rates of molecular evolution and patterns of selection among gymnosperms and flowering plants. *Molecular Biology and Evolution*, *34*(6), 1363–1377. https://doi.org/10.1093/molbev/msx069

De La Torre, A. R., Lin, Y.-C., Van de Peer, Y., & Ingvarsson, P. K. (2015). Genome-wide analysis reveals diverged patterns of codon bias, gene expression, and rates of sequence evolution in Picea gene families. *Genome Biology and Evolution*, *7*(4), 1002–1015. https://doi.org/10.1093/gbe/evv044

De La Torre, A. R., Puiu, D., Crepeau, M. W., Stevens, K., Salzberg, S. L., Langley, C. H., & Neale, D. B. (2018). Genomic architecture of complex traits in loblolly pine. *New Phytologist*, 221(4), 1789–1801. https://doi.org/10.1111/nph.15535

Delvas, N., Bauce, É., Labbé, C., Ollevier, T., & Bélanger, R. (2011). Phenolic compounds that confer resistance to spruce budworm. *Entomologia Experimentalis Et Applicata*, 141, 35–44. https://doi.org/10.1111/j.1570-7458.2011.01161.x

Dolgosheina, E. V., Morin, R. D., Aksay, G., Sahinalp, S. C., Magrini, V., Mardis, E. R., ... Unrau, P. J. (2008). Conifers have a unique small RNA silencing signature. *RNA*, 14(8), 1508–1515. https://doi.org/10.1261/rna.1052008

Du, R., Niu, S., Liu, Y., Sun, X., Porth, I., El-Kassaby, Y. A., & Li, W. (2017). The gibberellin GID1-DELLA signalling module exists in evolutionarily ancient conifers. *Scientific Reports*, 7(1), 16637. https://doi.org/10.1038/s41598-017-11859-w

Endler, A., & Persson, S. (2011). Cellulose synthases and synthesis in Arabidopsis. *Molecular Plant*, 4(2), 199–211. https://doi.org/10.1093/mp/ssq079

Flores-Renteria, L., Vazquez-Lobo, A., Whipple, A. V., Pinero, D., Marquez-Guzman, J., & Dominguez, C. A. (2011). Functional bisporangiate cones in *Pinus johannis* (Pinaceae): Implications for the evolution of bisexuality in seed plants. *American Journal of Botany*, 98(1), 130–139. https://doi.org/10.3732/ajb.1000275

Fossdal, C. G., Yaqoob, N., Krokene, P., Kvaalen, H., Solheim, H., & Yakovlev, I. A. (2012). Local and systemic changes in expression of resistance genes, NB-LRR genes and their putative microRNAs in Norway spruce after wounding and inoculation with the pathogen *Ceratocystis polonica*. *BMC Plant Biology*, 12, 105. https://doi.org/10.1186/1471-2229-12-105

Franceschi, V. R., Krokene, P., Christiansen, E., & Krekling, T. (2005). Anatomical and chemical defenses of conifer bark against bark beetles and other pests. *New Phytologist*, 167(2), 353–376. https://doi.org/10.1111/j.1469-8137.2005.01436.x

Friesen, N., Brandes, A., & Heslop-Harrison, J. S. (2001). Diversity, origin, and distribution of retrotransposons (gypsy and copia) in conifers. *Molecular Biology and Evolution*, 18(7), 1176–1188. https://doi.org/10.1093/oxfordjournals.molbev.a003905

Futuyma, D. J., & Agrawal, A. A. (2009). Macroevolution and the biological diversity of plants and herbivores. *Proceedings of the National Academy of Sciences of the United States of America*, 106(43), 18054–18061. https://doi.org/10.1073/pnas.0904106106

Gao, J., & Lan, T. (2016). Functional characterization of the late embryogenesis abundant (LEA) protein gene family from *Pinus tabuliformis* (Pinaceae) in *Escherichia coli*. *Scientific Reports*, 6, 19467. https://doi.org/10.1038/srep19467

Geisler, K., Jensen, N. B., Yuen, M. M., Madilao, L., & Bohlmann, J. (2016). Modularity of conifer diterpene resin acid biosynthesis: P450 enzymes of different CYP720B clades use alternative substrates and converge on the same products. *Plant Physiology*, 171(1), 152–164. https://doi.org/10.1104/pp.16.00180

Gesell, A., Blaukopf, M., Madilao, L., Yuen, M. M. S., Withers, S. G., Mattsson, J., ... Bohlmann, J. (2015). The gymnosperm cytochrome P450 CYP750B1 catalyzes stereospecific monoterpene hydroxylation of (+)-sabinene in thujone biosynthesis in Western Redcedar. *Plant Physiology*, 168(1), 94–106. https://doi.org/10.1104/pp.15.00315

Gill, G. P., Brown, G. R., & Neale, D. B. (2003). A sequence mutation in the cinnamyl alcohol dehydrogenase gene associated with altered lignification in loblolly pine. *Plant Biotechnology Journal*, 1(4), 253–258. https://doi.org/10.1046/j.1467-7652.2003.00024.x

Gong, X. F., Wang, M. W., Tashiro, S. I., Onodera, S., & Ikejima, T. (2005). Pseudolaric acid B induces apoptosis through p53 and bax/Bcl-2 pathways in human melanoma a375–s2 cells. *Archives of Pharmacal Research*, 28, 68. https://doi.org/10.1007/BF02975138

Gonzalez-Ibeas, D., Martinez-Garcia, P. J., Famula, R. A., Delfino-Mix, A., Stevens, K. A., Loopstra, C. A.,... Wegrzyn, J. L. (2016). Assessing the gene content of the megagenome: Sugar pine (*Pinus lambertiana*). *G3*, 6(12), 3787–3802.

Gottsberger, G. (1988). The reproductive biology of primitive angiosperms. *Taxon*, 37(3), 630. https://doi.org/10.2307/1221105

Grover, C. E., & Wendel, J. F. (2010). Recent insights into mechanisms of genome size change in plants. *Journal of Botany*, 2010, 1–8. https://doi.org/10.1155/2010/382732

Guan, R., Zhao, Y., Zhang, H. E., Fan, G., Liu, X., Zhou, W., ... Chen, W. (2016). Draft genome of the living fossil *Ginkgo biloba*. *GigaScience*, 5(1), 49. https://doi.org/10.1186/s13742-016-0154-1

Hall, D. E., Yuen, M. M. S., Jancsik, S., Quesada, A., Dullat, H. K., Li, M., ... Bohlmann, J. (2013a). Transcriptome resources and functional characterization of monoterpene synthases for two host species of the mountain pine beetle, lodgepole pine (*Pinus contorta*) and jack pine (*Pinus banksiana*). *BMC Plant Biology*, 13, 80. https://doi.org/10.1186/1471-2229-13-80

Hall, D. E., Zerbe, P., Jancsik, S., Quesada, A. L., Dullat, H., Madilao, L. L., ... Bohlmann, J. (2013b). Evolution of conifer diterpene synthases: Diterpene resin acid biosynthesis in lodgepole pine and jack pine involves monofunctional and bifunctional diterpene synthases. *Plant Physiology*, 161(2), 600–616. https://doi.org/10.1104/pp.112.208546

Hamberger, B., & Bak, S. (2013). Plant P450s as versatile drivers for evolution of species-specific chemical diversity. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 368(1612), 20120426. https://doi.org/10.1098/rstb.2012.0426

Hamberger, B., Hall, D., Yuen, M., Oddy, C., Hamberger, B., Keeling, C. I., ... Bohlmann, J. (2009). Targeted isolation, sequence assembly and characterization of two white spruce (*Picea glauca*) BAC clones for terpenoid synthase and cytochrome P450 genes involved in conifer defence reveal insights into a conifer genome. *BMC Plant Biology*, 9, 106. https://doi.org/10.1186/1471-2229-9-106

Hamberger, B., Ohnishi, T., Hamberger, B., Seguin, A., & Bohlmann, J. (2011). Evolution of diterpene metabolism: Sitka spruce CYP720B4 catalyzes multiple oxidations in resin acid biosynthesis of conifer defense against insects. *Plant Physiology*, 157(4), 1677–1695. https://doi.org/10.1104/pp.111.185843

Hanin, M., Brini, F., Ebel, C., Toda, Y., Takeda, S., & Masmoudi, K. (2011). Plant dehydrins and stress tolerance. *Plant Signaling & Behavior*, 6(10), 1503–1509. https://doi.org/10.4161/psb.6.10.17088

Higgins, C. F. (1992). ABC transporters: From microorganisms to man. *Annual Review of Cell Biology*, 8(1), 67–113. https://doi.org/10.1146/annurev.cb.08.110192.000435

Hu, L., Wang, Z., Du, H., & Huang, B. (2010). Differential accumulation of dehydrins in response to water stress for hybrid and common bermudagrass genotypes differing in drought tolerance. *Journal of Plant Physiology*, 167(2), 103–109. https://doi.org/10.1016/j.jplph.2009.07.008

Hu, W. J., Harding, S. A., Lung, J., Popko, J. L., Ralph, J., Stokke, D. D., ... Chiang, V. L. (1999). Repression of lignin biosynthesis promotes cellulose accumulation and growth in transgenic trees. *Nature Biotechnology*, 17, 808–812. https://doi.org/10.1038/11758

Huang, H. H., Xu, L. L., Tong, Z. K., Lin, E. P., Liu, Q. P., Cheng, L. J., & Zhu, M. Y. (2012). De novo characterization of the Chinese fir (*Cunninghamia lanceolata*) transcriptome and analysis of candidate genes involved in cellulose and lignin biosynthesis. *BMC Genomics*, 13, 648. https://doi.org/10.1186/1471-2164-13-648

Ito, H., & Gray, W. M. (2006). A gain-of-function mutation in the arabidopsis pleiotropic drug resistance transporter PDR9 confers resistance to auxinic herbicides. *Plant Physiology*, 142(1), 63–74. https://doi.org/10.1104/pp.106.084533

Jaramillo-Correa, J. P., Verdu, M., & Gonzalez-Martinez, S. C. (2010). The contribution of recombination to heterozygosity differs among plant

evolutionary lineages and life-forms. *BMC Evolutionary Biology*, 10(1), 22. https://doi.org/10.1186/1471-2148-10-22

Karlgren, A., Gyllenstrand, N., Clapham, D., & Lagercrantz, U. (2013). FLOWERING LOCUS T/TERMINAL FLOWER1-like genes affect growth rhythm and bud set in Norway spruce. *Plant Physiology*, 163, 792–803. https://doi.org/10.1104/pp.113.224139

Katahata, S. I., Futamura, N., Igasaki, T., & Shinohara, K. (2014). Functional analysis of SOC1-like and AGL6-like MADS-box genes of the gymnosperm *Cryptomeria japonica*. *Tree Genetics and Genomes*, 10(2), 317–327. https://doi.org/10.1007/s11295-013-0686-9

Keeling, C. I., & Bohlmann, J. (2006). Diterpene resin acids in conifers. *Phytochemistry*, 67(22), 2415–2423. https://doi.org/10.1016/j.phytochem.2006.08.019

Keeling, C. I., Weisshaar, S., Lin, R. P., & Bohlmann, J. (2008). Functional plasticity of paralogous diterpene synthases involved in conifer defense. *Proceedings of the National Academy of Sciences of the United States of America*, 105, 1085–1090. https://doi.org/10.1073/pnas.0709466105

Keeling, C. I., Weisshaar, S., Ralph, S. G., Jancsik, S., Hamberger, B., Dullat, H. K., & Bohlmann, J. (2011). Transcriptome mining, functional characterization, and phylogeny of a large terpene synthase gene family in spruce (*Picea* spp.). *BMC Plant Biology*, 11(1), 43. https://doi.org/10.1186/1471-2229-11-43

Kjellsen, T. D., Shiryaeva, L., Schröder, W. P., & Strimbeck, G. R. (2010). Proteomics of extreme freezing tolerance in Siberian spruce (*Picea obovata*). *Journal of Proteomics*, 73(5), 965–975. https://doi.org/10.1016/j.jprot.2009.12.010

Klintenäs, M., Pin, P. A., Benlloch, R., Ingvarsson, P. K., & Nilsson, O. (2012). Analysis of conifer *FLOWERING LOCUS T / TERMINAL FLOWER1*-like genes provides evidence for dramatic biochemical evolution in the angiosperm *FT* lineage. *New Phytologist*, 196(4), 1260–1273. https://doi.org/10.1111/j.1469-8137.2012.04332.x

Kontunen-Soppela, S., & Laine, K. (2001). Seasonal fluctuation of dehydrins is related to osmotic status in Scots pine needles. *Trees*, 15(7), 425–430. https://doi.org/10.1007/s004680100124

Kovalchuk, A., Raffaello, T., Jaber, E., Keriö, S., Ghimire, R., Lorenz, W. W., … Asiegbu, F. O. (2015). Activation of defence pathways in Scots pine bark after feeding by pine weevil (*Hylobius abietis*). *BMC Genomics*, 16(1), 352. https://doi.org/10.1186/s12864-015-1546-9

Krauskopf, E., Harris, P. J., & Putterill, J. (2005). The cellulose synthase gene PrCESA10 is involved in cellulose biosynthesis in developing tracheids of the gymnosperm *Pinus radiata*. *Gene*, 350(2), 107–116. https://doi.org/10.1016/j.gene.2005.02.018

Kubo, M., Udagawa, M., Nishikubo, N., Horiguchi, G., Yamaguchi, M., Ito, J., … Demura, T. (2005). Transcription switches for protoxylem and metaxylem vessel formation. *Genes & Development*, 19(16), 1855–1860. https://doi.org/10.1101/gad.1331305

Kumar, M., Atanassov, I., & Turner, S. (2017). Functional analysis of cellulose synthase (CESA) protein class specificity. *Plant Physiology*, 173(2), 970–983. https://doi.org/10.1104/pp.16.01642

Lamping, E., Baret, P. V., Holmes, A. R., Monk, B. C., Goffeau, A., & Cannon, R. D. (2010). Fungal PDR transporters: Phylogeny, topology, motifs and function. *Fungal Genetics and Biology*, 47(2), 127–142. https://doi.org/10.1016/j.fgb.2009.10.007

Landis, J. B., Soltis, D. E., Li, Z., Marx, H. E., Barker, M. S., Tank, D. C., & Soltis, P. S. (2018). Impact of whole-genome duplication events on diversification rates in angiosperms. *American Journal of Botany*, 105(3), 348–363. https://doi.org/10.1002/ajb2.1060

Lefevre, F., Baijot, A., & Boutry, M. (2015). Plant ABC transporters: Time for biochemistry? *Biochemical Society Transactions*, 43(5), 931–936. https://doi.org/10.1042/BST20150108

Leitch, A. R., & Leitch, I. J. (2012). Ecological and genetic factors linked to contrasting genome dynamics in seed plants. *New Phytologist*, 194(3), 629–646. https://doi.org/10.1111/j.1469-8137.2012.04105.x

Leslie, A. B., Beaulieu, J. M., Rai, H. S., Crane, P. R., Donoghue, M. J., & Mathews, S. (2012). Hemisphere-scale differences in conifer evolutionary dynamics. *Proceedings of the National Academy of Sciences of the United States of America*, 109(40), 16217–16221. https://doi.org/10.1073/pnas.1213621109

Li, Q., Deng, C., Xia, Y., Kong, L., Zhang, H., Zhang, S., & Wang, J. (2017). Identification of novel miRNAs and miRNA expression profiling in embryogenic tissues of *Picea balfouriana* treated by 6-benzylaminopurine. *PLoS ONE*, 12, e0176112. https://doi.org/10.1371/journal.pone

Li, S., Bashline, L., Lei, L., & Gu, Y. (2014). Cellulose synthesis and its regulation. *The Arabidopsis Book*, 12, e0169. https://doi.org/10.1199/tab.0169

Li, S.-H., Nagy, N. E., Hammerbacher, A., Krokene, P., Niu, X.-M., Gershenzon, J., & Schneider, B. (2012). Localization of phenolics in phloem parenchyma cells of Norway spruce (*Picea abies*). *ChemBioChem*, 13(18), 2707–2713. https://doi.org/10.1002/cbic.201200547

Liu, J.-J., & Ekramoddoullah, A. K. M. (2007). The CC-NBS-LRR subfamily in *Pinus monticola*: Targeted identification, gene expression, and genetic linkage with resistance to *Cronartium ribicola*. *Phytopathology*, 97(6), 728–736. https://doi.org/10.1094/PHYTO-97-6-0728

Liu, J.-J., & Ekramoddoullah, A. K. M. (2011). Genomic organization, induced expression and promoter activity of a resistance gene analog (PmTNL1) in western white pine (*Pinus monticola*). *Planta*, 233, 1041–1053. https://doi.org/10.1007/s00425-011-1353-8

Liu, J.-J., Ekramoddoullah, A. K. M., Taylor, D., Piggott, N., Lane, S., & Hawkins, B. (2004). Characterization of Picg5 novel proteins associated with seasonal cold acclimation of white spruce (*Picea glauca*). *Trees*, 18(6), 649–657. https://doi.org/10.1007/s00468-004-0336-9

Liu, Y., & El-Kassaby, Y. A. (2017). Landscape of fluid sets of hairpin-derived 21-/24-nt-long small RNAs at seed set uncovers special epigenetic features in *Picea glauca*. *Genome Biology and Evolution*, 9(1), evw283. https://doi.org/10.1093/gbe/evw283

Liu, Z.-X., Xiong, H.-Y., Li, L.-Y., & Fei, Y.-J. (2018). Functional conservation of an AGAMOUS orthologous gene controlling reproductive organ development in the gymnosperm species *Taxus chinensis* var. mairei. *Journal of Plant Biology*, 61(1), 50–59. https://doi.org/10.1007/s12374-017-0154-4

Lopez, C. G., Banowetz, G. M., Peterson, C. J., & Kronstad, W. E. (2003). Dehydrin expression and drought tolerance in seven wheat cultivars. *Crop Science*, 43(2), 577–582. https://doi.org/10.2135/cropsci2003.5770

Lu, Y., Jin, B., Wang, L., Wang, Y., Wang, D., Jiang, X.-X., & Chen, P. (2011). Adaptation of male reproductive structures to wind pollination in gymnosperms: Cones and pollen grains. *Canadian Journal of Plant Science*, 91(5), 897–906. https://doi.org/10.4141/cjps2011-020

Mageroy, M. H., Parent, G., Germanos, G., Giguère, I., Delvas, N., Maaroufi, H., … Mackay, J. J. (2015). Expression of the β-glucosidase gene Pgβglu-1 underpins natural resistance of white spruce against spruce budworm. *The Plant Journal*, 81(1), 68–80. https://doi.org/10.1111/tpj.12699

Meents, M. J., Watanabe, Y., & Samuels, A. L. (2018). The cell biology of secondary cell wall biosynthesis. *Annals of Botany*, 121(6), 1107–1125. https://doi.org/10.1093/aob/mcy005

Méndez-Espinoza, C., Parent, G. J., Lenz, P., Rainville, A., Tremblay, L., Adams, G., … MacKay, J. (2018). Genetic control and evolutionary potential of a constitutive resistance mechanism against the spruce budworm (*Choristoneura fumiferana*) in white spruce (*Picea glauca*). *Heredity*, 121(2), 142–154. https://doi.org/10.1038/s41437-018-0061-6

Mohanta, T. K., Occhipinti, A., Atsbaha Zebelo, S., Foti, M., Fliegmann, J., Bossi, S., … Bertea, C. M. (2012). *Ginkgo biloba* responds to herbivory

by activating early signaling and direct defenses. *PLoS ONE*, *7*(3), e32822. https://doi.org/10.1371/journal.pone.0032822

Morin, R. D., Aksay, G., Dolgosheina, E., Ebhardt, H. A., Magrini, V., Mardis, E. R., ... Unrau, P. J. (2008). Comparative analysis of the small RNA transcriptomes of *Pinus contorta* and *Oryza sativa*. *Genome Research*, *18*(4), 571–584. https://doi.org/10.1101/gr.6897308

Moyroud, E., Monniaux, M., Thévenon, E., Dumas, R., Scutt, C. P., Frohlich, M. W., & Parcy, F. (2017). A link between LEAFY and B-gene homologues in *Welwitschia mirabilis* sheds light on ancestral mechanisms prefiguring floral development. *New Phytologist*, *216*(2), 469–481.

Neale, D. B., Martínez-García, P. J., De La Torre, A. R., Montanari, S., & Wei, X.-X. (2017). Novel insights into tree biology and genome evolution as revealed through genomics. *Annual Review of Plant Biology*, *68*(1), 457–483. https://doi.org/10.1146/annurev-arplant-042916-041049

Neale, D. B., Wegrzyn, J. L., Stevens, K. A., Zimin, A. V., Puiu, D., Crepeau, M. W., ... Langley, C. H. (2014). Decoding the massive genome of loblolly pine using haploid DNA and novel assembly strategies. *Genome Biology*, *15*(3), R59. https://doi.org/10.1186/gb-2014-15-3-r59

Nemesio Gorriz, M., Hammerbacher, A., Ihrmark, K., Källman, T., Olson, Å., Lascoux, M., ... Elfstrand, M. (2016). Different alleles of a gene encoding leucoanthocyanidin reductase (PaLAR3) influence resistance against the fungus *Heterobasidion parviporum* in *Picea abies*. *Plant Physiology*, *171*(4), 2671–2681. https://doi.org/10.1104/pp.16.00685

Nepi, M., Little, S., Guarnieri, M., Nocentini, D., Prior, N., Gill, J., ... von Aderkas, P. (2017). Phylogenetic and functional signals in gymnosperm ovular secretions. *Annals of Botany*, *120*(6), 923–936. https://doi.org/10.1093/aob/mcx103

Niu, S.-H., Liu, C., Yuan, H.-W., Li, P., Li, Y., & Li, W. (2015). Identification and expression profiles of sRNAs and their biogenesis and action-related genes in male and female cones of *Pinus tabuliformis*. *BMC Genomics*, *16*(1), 693. https://doi.org/10.1186/s12864-015-1885-6

Niu, S., Yuan, H., Sun, X., Porth, I., Li, Y., El-Kassaby, Y. A., & Li, W. (2016). A transcriptomics investigation into pine reproductive organ development. *New Phytologist*, *209*, 1278–1289. https://doi.org/10.1111/nph.13680

Niu, S., Yuan, L., Zhang, Y., Chen, X., & Li, W. (2014). Isolation and expression profiles of gibberellin metabolism genes in developing male and female cones of *Pinus tabuliformis*. *Functional & Integrative Genomics*, *14*(4), 697–705. https://doi.org/10.1007/s10142-014-0387-y

Nuruzzaman, M., Zhang, R., Cao, H.-Z., & Luo, Z.-Y. (2014). Plant pleiotropic drug resistance transporters: Transport mechanism, gene expression, and function. *Journal of Integrative Plant Biology*, *56*(8), 729–740. https://doi.org/10.1111/jipb.12196

Nützmann, H.-W., Huang, A., & Osbourn, A. (2016). Plant metabolic clusters – From genetics to genomics. *The New Phytologist*, *211*(3), 771–789. https://doi.org/10.1111/nph.13981

Nystedt, B., Street, N. R., Wetterbom, A., Zuccolo, A., Lin, Y.-C., Scofield, D. G., ... Jansson, S. (2013). The Norway spruce genome sequence and conifer genome evolution. *Nature*, *497*(7451), 579–584. https://doi.org/10.1038/nature12211

Ohashi-Ito, K., & Fukuda, H. (2014). Initiation of vascular development. *Physiologia Plantarum*, *151*(2), 142–146. https://doi.org/10.1111/ppl.12111

Oliva, J., Rommel, S., Fossdal, C. G., Hietala, A. M., Nemesio-Gorriz, M., Solheim, H., & Elfstrand, M. (2015). Transcriptional responses of Norway spruce (*Picea abies*) inner sapwood against *Heterobasidion parviporum*. *Tree Physiology*, *35*(9), 1007–1015. https://doi.org/10.1093/treephys/tpv063

Owens, J. N., Takaso, T., & Runions, C. J. (1998). Pollination in conifers. *Trends in Plant Science*, *3*, 479–485.

Pan, L., Ren, L., Chen, F., Feng, Y., & Luo, Y. (2016). Antifeedant activity of *Ginkgo biloba*. Secondary metabolites against *Hyphantria cunea* larvae: Mechanisms and applications. *PLoS ONE*, *11*(5), e0155682. https://doi.org/10.1371/journal.pone.0155682

Patten, A. M., Vassão, D. G., Wolcott, M. P., Davin, L. B., & Lewis, N. G. (2010). 3.*27 – Trees: A remarkable biochemical bounty. Comprehensive natural products II*. Oxford, UK: Elsevier; 2010: 1173–1296.

Patzlaff, A., McInnis, S., Courtenay, A., Surman, C., Newman, L. J., Smith, C., ... Campbell, M. M. (2003). Characterisation of a pine MYB that regulates lignification. *The Plant Journal*, *36*(6), 743–754. https://doi.org/10.1046/j.1365-313X.2003.01916.x

Pavy, N., Deschênes, A., Blais, S., Lavigne, P., Beaulieu, J., Isabel, N., ... Bousquet, J. (2013). The landscape of nucleotide polymorphism among 13,500 genes of the conifer *Picea glauca*, relationships with functions, and comparison with *Medicago truncatula*. *Genome Biology and Evolution*, *5*(10), 1910–1925. https://doi.org/10.1093/gbe/evt143

Pavy, N., Pelgas, B., Laroche, J., Rigault, P., Isabel, N., & Bousquet, J. (2012). A spruce gene map infers ancient plant genome re-shuffling and subsequent slow evolution in the gymnosperm lineage leading to extant conifers. *BMC Biology*, *10*, 84. https://doi.org/10.1186/1741-7007-10-84

Pennisi, E. (2009). On the origin of flowering plants. *Science*, *324*(5923), 28–31. https://doi.org/10.1126/science.324.5923.28

Peris, D., Pérez-de la Fuente, R., Peñalver, E., Delclòs, X., Barrón, E., & Labandeira, C. C. (2017). False blister beetles and the expansion of gymnosperm-insect pollination modes before angiosperm dominance. *Current Biology*, *27*(6), 897–904. https://doi.org/10.1016/j.cub.2017.02.009

Petterle, A., Karlberg, A., & Bhalerao, R. P. (2013). Daylength mediated control of seasonal growth patterns in perennial trees. *Current Opinion in Plant Biology*, *16*(3), 301–306. https://doi.org/10.1016/j.pbi.2013.02.006

Pickel, B., & Schaller, A. (2013). Dirigent proteins: Molecular characteristics and potential biotechnological applications. *Applied Microbiology and Biotechnology*, *97*(19), 8427–8438. https://doi.org/10.1007/s00253-013-5167-4

Pierman, B., Toussaint, F., Bertin, A., Lévy, D., Smargiasso, N., De Pauw, E., & Boutry, M. (2017). Activity of the purified plant ABC transporter NtPDR1 is stimulated by diterpenes and sesquiterpenes involved in constitutive and induced defenses. *The Journal of Biological Chemistry*, *292*(47), 19491–19502. https://doi.org/10.1074/jbc.M117.811935

Pires, N. D., & Dolan, L. (2012). Morphological evolution in land plants: New designs with old genes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1588), 508–518. https://doi.org/10.1098/rstb.2011.0252

Porth, I., & El-Kassaby, Y. (2014). Assessment of the genetic diversity in forest tree populations using molecular markers. *Diversity*, *6*, 283. https://doi.org/10.3390/d6020283

Porth, I., Hamberger, B., White, R., & Ritland, K. (2011). Defense mechanisms against herbivory in Picea: Sequence evolution and expression regulation of gene family members in the phenylpropanoid pathway. *BMC Genomics*, *12*(1), 608. https://doi.org/10.1186/1471-2164-12-608

Porth, I., White, R., Jaquish, B., Alfaro, R., Ritland, C., & Ritland, K. (2012). Genetical genomics identifies the genetic architecture for growth and weevil resistance in spruce. *PLoS ONE*, *7*, e44397. https://doi.org/10.1371/journal.pone.0044397

Porth, I., White, R., Jaquish, B., & Ritland, K. (2018). Partial correlation analysis of transcriptomes helps detangle the growth and defense network in spruce. *New Phytologist*, *218*(4), 1349–1359. https://doi.org/10.1111/nph.15075

Prunier, J., Caron, S., Lamothe, M., Blais, S., Bousquet, J., Isabel, N., & MacKay, J. J. (2017). Gene copy number variations in adaptive evolution: The genomic distribution of gene copy number variations revealed by genetic mapping and their adaptive role in an undomesticated species, white spruce (*Picea glauca*). *Molecular Ecology*, *26*(21), 5989–6001. https://doi.org/10.1111/mec.14337. PMID: 28833771

Prunier, J., Caron, S., & MacKay, J. J. (2017). CNVs into the wild: Screening the genomes of conifer trees (*Picea* spp.) reveals fewer gene copy number variations in hybrids and links to adaptation. *BMC Genomics*, *18*(1), 97. https://doi.org/10.1186/s12864-016-3458-8. PMID: 28100184

Prunier, J., Giguère, I., Ryan, N., Guy, R., Soolanayakanahally, R., Isabel, N., ... Porth, I. (2019). Gene copy number variations involved in balsam poplar (*Populus balsamifera* L.) adaptive variations. *Molecular Ecology*, *28*, 1476–1490. https://doi.org/10.1111/mec.14836. PMID: 30270494

Prunier, J., Verta, J.-P., & MacKay, J. J. (2016). Conifer genomics and adaptation: At the crossroads of genetic diversity and genome function. *New Phytologist*, *209*(1), 44–62. https://doi.org/10.1111/nph.13565

Ralph, S. G., Jancsik, S., & Bohlmann, J. (2007). Dirigent proteins in conifer defense II: Extended gene discovery, phylogeny, and constitutive and stress-induced gene expression in spruce (*Picea* spp.). *Phytochemistry*, *68*(14), 1975–1991. https://doi.org/10.1016/j.phytochem.2007.04.042

Ralph, S., Park, J.-Y., Bohlmann, J., & Mansfield, S. D. (2006). Dirigent proteins in conifer defense: Gene discovery, phylogeny, and differential wound- and insect-induced expression of a family of DIR and DIR-like genes in spruce (*Picea* spp.). *Plant Molecular Biology*, *60*(1), 21–40. https://doi.org/10.1007/s11103-005-2226-y

Ralph, S. G., Yueh, H., Friedmann, M., Aeschliman, D., Zeznik, J. A., Nelson, C. C., ... Bohlmann, J. (2006). Conifer defence against insects: Microarray gene expression profiling of Sitka spruce (*Picea sitchensis*) induced by mechanical wounding or feeding by spruce budworms (*Choristoneura occidentalis*) or white pine weevils (*Pissodes strobi*) reveals large-scale changes of the host transcriptome. *Plant, Cell & Environment*, *29*(8), 1545–1570. https://doi.org/10.1111/j.1365-3040.2006.01532.x

Renault, H., Bassard, J. E., Hamberger, B., & Werck-Reichhart, D. (2014). Cytochrome P450-mediated metabolic engineering: Current progress and future challenges. *Current Opinion in Plant Biology*, *19*, 27–34. https://doi.org/10.1016/j.pbi.2014.03.004

Renaut, J., Hoffmann, L., & Hausman, J.-F. (2005). Biochemical and physiological mechanisms related to cold acclimation and enhanced freezing tolerance in poplar plantlets. *Physiologia Plantarum*, *125*(1), 82–94. https://doi.org/10.1111/j.1399-3054.2005.00554.x

Rigault, P., Boyle, B., Lepage, P., Cooke, J. E. K., Bousquet, J., & MacKay, J. J. (2011). A white spruce gene catalog for conifer genome analyses. *Plant Physiology*, *157*(1), 14–28. https://doi.org/10.1104/pp.111.179663

Rinne, P., Welling, A., & Kaikuranta, P. (1998). Onset of freezing tolerance in birch (*Betula pubescens* Ehrh.) involves LEA proteins and osmoregulation and is impaired in an ABA-deficient genotype. *Plant, Cell & Environment*, *21*(6), 601–611. https://doi.org/10.1046/j.1365-3040.1998.00306.x

Roach, C. R., Hall, D. E., Zerbe, P., & Bohlmann, J. (2014). Plasticity and evolution of (+)-3-carene synthase and (–)-sabinene synthase functions of a Sitka spruce monoterpene synthase gene family associated with weevil resistance. *Journal of Biological Chemistry*, *289*(34), 23859–23869. https://doi.org/10.1074/jbc.M114.571703

Robert, J. A., Madilao, L. L., White, R., Yanchuk, A., King, J., & Bohlmann, J. (2010). Terpenoid metabolite profiling in Sitka spruce identifies association of dehydroabietic acid, (+)-3-carene, and terpinolene with resistance against white pine weevil. *Botany-Botanique*, *88*, 810–820. https://doi.org/10.1139/B10-049

Rodrigues, A. S., De Vega, J. J., & Miguel, C. M. (2018). Comprehensive assembly and analysis of the transcriptome of maritime pine developing embryos. *BMC Plant Biology*, *18*, 379. https://doi.org/10.1186/s12870-018-1564-2

Rodrigues, A. S., & Miguel, C. M. (2017). The pivotal role of small non-coding RNAs in the regulation of seed development. *Plant Cell Reports*, *36*(5), 653–667. https://doi.org/10.1007/s00299-017-2120-5

Rorat, T. (2006). Plant dehydrins — Tissue location, structure and function. *Cellular & Molecular Biology Letters*, *11*(4), 536. https://doi.org/10.2478/s11658-006-0044-0

Rudall, P. J., Hilton, J., Vergara-Silva, F., & Bateman, R. M. (2011). Recurrent abnormalities in conifer cones and the evolutionary origins of flower-like structures. *Trends in Plant Science*, *16*(3), 151–159. https://doi.org/10.1016/j.tplants.2010.11.002

Sainz, M. F., Souto, J. A., Regentova, D., Johansson, M. K. G., Timhagen, S. T., Irvine, D. J., ... Howdle, S. M. (2016). A facile and green route to terpene derived acrylate and methacrylate monomers and simple free radical polymerisation to yield new renewable polymers and coatings. *Polymer Chemistry*, *7*, 2882–2887. https://doi.org/10.1039/C6PY00357E

Salmon, Y., Dietrich, L., Sevanto, S., Hölttä, T., Dannoura, M., & Epron, D. (2019). Drought impacts on tree phloem: From cell-level responses to ecological significance. *Tree Physiology*, *39*(2), 173–191. https://doi.org/10.1093/treephys/tpy153

Sarkar, P., Bosneaga, E., & Auer, M. (2009). Plant cell walls throughout evolution: Towards a molecular understanding of their design principles. *Journal of Experimental Botany*, *60*(13), 3615–3635. https://doi.org/10.1093/jxb/erp245

Sasse, J., Schlegel, M., Borghi, L., Ullrich, F., Lee, M., Liu, G.-W., ... Kretzschmar, T. (2016). *Petunia hybrida* PDR2 is involved in herbivore defense by controlling steroidal contents in trichomes. *Plant, Cell & Environment*, *39*(12), 2725–2739. https://doi.org/10.1111/pce.12828

Shafer, A. B. A., Cullingham, C. I., Côté, S. D., & Coltman, D. W. (2010). Of glaciers and refugia: A decade of study sheds new light on the phylogeography of northwestern North America. *Molecular Ecology*, *19*(21), 4589–4621. https://doi.org/10.1111/j.1365-294X.2010.04828.x

Shalev, T. J., Yuen, M. M. S., Gesell, A., Yuen, A., Russell, J. H., & Bohlmann, J. (2018). An annotated transcriptome of highly inbred *Thuja plicata* (Cupressaceae) and its utility for gene discovery of terpenoid biosynthesis and conifer defense. *Tree Genetics & Genomes*, *14*(3), 35. https://doi.org/10.1007/s11295-018-1248-y

Shin, S.-Y., & Shin, C. (2016). Regulatory non-coding RNAs in plants: Potential gene resources for the improvement of agricultural traits. *Plant Biotechnology Reports*, *10*(2), 35–47. https://doi.org/10.1007/s11816-016-0389-4

Singh, B., & Sharma, R. A. (2015). Plant terpenes: Defense responses, phylogenetic analysis, regulation and clinical applications. *3 Biotech*, *5*(2), 129–151. https://doi.org/10.1007/s13205-014-0220-2

Soltis, P. S., & Soltis, D. E. (2016). Ancient WGD events as drivers of key innovations in angiosperms. *Current Opinion in Plant Biology*, *30*, 159–165. https://doi.org/10.1016/j.pbi.2016.03.015

Songstad, D. D., Petolino, J. F., Voytas, D. F., & Reichert, N. A. (2017). Genome editing of plants. *Critical Reviews in Plant Sciences*, *36*(1), 1–23. https://doi.org/10.1080/07352689.2017.1281663

Steppuhn, A., & Baldwin, I. T. (2008). Induced defenses and the cost-benefit paradigm. *Induced Plant Resistance to Herbivory*, 61–83. https://doi.org/10.1007/978-1-4020-8182-8_3

Stival Sena, J., Giguère, I., Rigault, P., Bousquet, J., & Mackay, J. (2018). Expansion of the dehydrin gene family in the Pinaceae is associated with considerable structural diversity and drought-responsive expression. *Tree Physiology*, *38*(3), 442–456. https://doi.org/10.1093/treephys/tpx125

Strimbeck, G. R., Schaberg, P. G., Fossdal, C. G., Schröder, W. P., & Kjellsen, T. D. (2015). Extreme low temperature tolerance in woody plants. *Frontiers in Plant Science*, *6*, https://doi.org/10.3389/fpls.2015.00884

Suprunova, T., Krugman, T., Fahima, T., Chen, G., Shams, I., Korol, A., & Nevo, E. (2004). Differential expression of dehydrin genes in wild barley, *Hordeum spontaneum*, associated with resistance to water deficit. *Plant, Cell & Environment*, *27*(10), 1297–1308. https://doi.org/10.1111/j.1365-3040.2004.01237.x

Suzuki, S., Li, L., Sun, Y.-H., & Chiang, V. L. (2006). The cellulose synthase gene superfamily and biochemical functions of xylem-specific cellulose synthase-like genes in *Populus trichocarpa*. *Plant Physiology*, *142*(3), 1233–1245. https://doi.org/10.1104/pp.106.086678

Tan, T. T., Endo, H., Sano, R., Kurata, T., Yamaguchi, M., Ohtani, M., & Demura, T. (2018). Transcription factors VND1-VND3 Contribute to cotyledon xylem vessel formation. *Plant Physiology*, *176*(1), 773–789. https://doi.org/10.1104/pp.17.00461

Thaler, J. S., Farag, M. A., Paré, P. W., & Dicke, M. (2002). Jasmonate-deficient plants have reduced direct and indirect defences against herbivores. *Ecology Letters*, *5*(6), 764–774. https://doi.org/10.1046/j.1461-0248.2002.00388.x

Tsubomura, M., Kurita, M., & Watanabe, A. (2016). Determination of male strobilus developmental stages by cytological and gene expression analyses in Japanese cedar (*Cryptomeria japonica*). *Tree Physiology*, *36*(5), 653–666. https://doi.org/10.1093/treephys/tpw001

Turner, G. W., Parrish, A. N., Zager, J. J., Fischedick, J. T., & Lange, B. M. (2018). Assessment of flux through oleoresin biosynthesis in epithelial cells of loblolly pine resin ducts. *Journal of Experimental Botany*, *70*(1), 217–230. https://doi.org/10.1093/jxb/ery338

Turtola, S., Manninen, A.-M., Rikala, R., & Kainulainen, P. (2003). Drought stress alters the concentration of wood terpenoids in Scots pine and Norway spruce seedlings. *Journal of Chemical Ecology*, *29*(9), 1981–1995. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14584671

Van Ex, F., Jacob, Y., & Martienssen, R. A. (2011). Multiple roles for small RNAs during plant reproduction. *Current Opinion in Plant Biology*, *14*(5), 588–593. https://doi.org/10.1016/j.pbi.2011.07.003

Vanneste, K., Maere, S., & Van de Peer, Y. (2014). Tangled up in two: A burst of genome duplications at the end of the Cretaceous and the consequences for plant evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*:20130353. https://doi.org/10.1098/rstb.2013.0353

Vázquez-Lobo, A., Carlsbecker, A., Vergara-Silva, F., Alvarez-Buylla, E. R., Piñero, D., & Engström, P. (2007). Characterization of the expression patterns of LEAFY/FLORICAULA and NEEDLY orthologs in female and male cones of the conifer genera Picea, Podocarpus, and Taxus: Implications for current evo-devo hypotheses for gymnosperms. *Evolution & Development*, *9*(5), 446–459. https://doi.org/10.1111/j.1525-142X.2007.00182.x

Visser, E. A., Wegrzyn, J. L., Myburg, A. A., & Naidoo, S. (2018). Defence transcriptome assembly and pathogenesis related gene family analysis in *Pinus tecunumanii* (low elevation). *BMC Genomics*, *19*(1), 632. https://doi.org/10.1186/s12864-018-5015-0

Von Aderkas, P., Prior, N. A., & Little, S. A. (2018). The evolution of sexual fluids in gymnosperms from pollination drops to nectar. *Frontiers in Plant Science*, *9*(1844), 1–21. https://doi.org/10.3389/fpls.2018.01844

Vranová, E., Coman, D., & Gruissem, W. (2012). Structure and dynamics of the isoprenoid pathway network. *Molecular Plant*, *5*(2), 318–333. https://doi.org/10.1093/mp/sss015

Wachowiak, W., Balk, P. A., & Savolainen, O. (2009). Search for nucleotide diversity patterns of local adaptation in dehydrins and other cold-related candidate genes in Scots pine (*Pinus sylvestris* L.). *Tree Genetics & Genomes*, *5*(1), 117–132. https://doi.org/10.1007/s11295-008-0188-3

Wan, T., Liu, Z.-M., Li, L.-F., Leitch, A. R., Leitch, I. J., Lohaus, R., … Wang, X.-M. (2018). A genome for gnetophytes and early evolution of seed plants. *Nature Plants*, *4*(2), 82–89. https://doi.org/10.1038/s41477-017-0097-2

Wang, X.-Q., & Ran, J.-H. (2014). Evolution and biogeography of gymnosperms. *Molecular Phylogenetics and Evolution*, *75*, 24–40. https://doi.org/10.1016/j.ympev.2014.02.005

Wani, M. C., Taylor, H. L., Wall, M. E., Coggon, P., & McPhail, A. T. (1971). Plant antitumor agents. VI. The isolation and structure of taxol, a novel antileukemic and antitumor agent from Taxus brevifolia. *Journal of the American Chemical Society*, *93*, 2325–2327.

Warren, R. L., Keeling, C. I., Yuen, M. M., Raymond, A., Taylor, G. A., Vandervalk, B. P., … Bohlmann, J. (2015). Improved white spruce (*Picea glauca*) genome assemblies and annotation of large gene families of conifer terpenoid and phenolic defense metabolism. *The Plant Journal*, *83*(2), 189–212. https://doi.org/10.1111/tpj.12886

Wegrzyn, J. L., Liechty, J. D., Stevens, K. A., Wu, L.-S., Loopstra, C. A., Vasquez-Gross, H. A., … Neale, D. B. (2014). Unique features of the loblolly pine (*Pinus taeda* L.) megagenome revealed through sequence annotation. *Genetics*, *196*(3), 891–909. https://doi.org/10.1534/genetics.113.159996

Xia, R., Xu, J., Arikit, S., & Meyers, B. C. (2015). Extensive families of miRNAs and *PHAS* loci in norway spruce demonstrate the origins of complex phasiRNA networks in seed plants. *Molecular Biology and Evolution*, *32*(11), 2905–2918. https://doi.org/10.1093/molbev/msv164

Yakovlev, I. A., Carneros, E., Lee, Y. K., Olsen, J. E., & Fossdal, C. G. (2016). Transcriptional profiling of epigenetic regulators in somatic embryos during temperature induced formation of an epigenetic memory in Norway spruce. *Planta*, *243*, 1237–1249. https://doi.org/10.1007/s00425-016-2484-8

Yazaki, K. (2006). ABC transporters involved in the transport of plant secondary metabolites. *FEBS Letters*, *580*(4), 1183–1191. https://doi.org/10.1016/j.febslet.2005.12.009

Yin, Y., Johns, M. A., Cao, H., & Rupani, M. (2014). A survey of plant and algal genomes and transcriptomes reveals new insights into the evolution and function of the cellulose synthase superfamily. *BMC Genomics*, *15*(1), 260. https://doi.org/10.1186/1471-2164-15-260

Zerbe, P., Chiang, A., Yuen, M., Hamberger, B., Hamberger, B., Draper, J. A., … Bohlmann, J. (2012). Bifunctional *cis*-abienol synthase from *Abies balsamea* discovered by transcriptome sequencing and its implications for diterpenoid fragrance production. *Journal of Biological Chemistry*, *287*(15), 12121–12131. https://doi.org/10.1074/jbc.M111.317669

Zhang, J., Nieminen, K., Serra, J. A. A., & Helariutta, Y. (2014). The formation of wood and its control. *Current Opinion in Plant Biology*, *17*, 56–63. https://doi.org/10.1016/j.pbi.2013.11.003

Zhang, J., Wu, T., Li, L., Han, S., Li, X., Zhang, S., & Qi, L. (2013). Dynamic expression of small RNA populations in larch (*Larix leptolepis*). *Planta*, *237*, 89–101. https://doi.org/10.1007/s00425-012-1753-4

Zhang, M., & Zheng, C. X. (2016). Archegonium and fertilization in Coniferopsida. *Trees*, *30*, 75–86. https://doi.org/10.1007/s00468-015-1349-2

Zhong, R., Lee, C., & Ye, Z. H. (2010). Global analysis of direct targets of secondary wall NAC master switches in arabidopsis. *Molecular Plant*, *3*, 1087–1103. https://doi.org/10.1093/mp/ssq062

Zhong, R., Lee, C., Zhou, J., McCarthy, R. L., & Ye, Z. H. (2008). A battery of transcription factors involved in the regulation of secondary cell wall biosynthesis in arabidopsis. *The Plant Cell Online*, *20*, 2763–2782.

Zhou, J., Zhong, R., & Ye, Z.-H. (2014). Arabidopsis NAC domain proteins, VND1 to VND5, are transcriptional regulators of secondary wall biosynthesis in vessels. *PLoS ONE*, *9*(8), e105726. https://doi.org/10.1371/journal.pone.0105726

Zhou, S. S., Xing, Z., Liu, H., Hu, X. G., Gao, Q., Xu, J., … Mao, J. F. (2019). In-depth transcriptome characterization uncovers distinct gene family expansions for *Cupressus gigantea* important to this long-lived species' adaptability to environmental cues. *BMC Genomics*, *20*(1), 213. https://doi.org/10.1186/s12864-019-5584-6. PubMed PMID: 30866823

Zi, J., Mafu, S., & Peters, R. J. (2014). To gibberellins and beyond! Surveying the evolution of (Di)terpenoid metabolism. *Annual Review*

of Plant Biology, *65*(1), 259–286. https://doi.org/10.1146/annurev-ar-plant-050213-035705

Zulak, K. G., & Bohlmann, J. (2010). Terpenoid biosynthesis and specialized vascular cells of conifer defense. *Journal of Integrative Plant Biology*, *52*(1), 86–97. https://doi.org/10.1111/j.1744-7909.2010.00910.x

SUPPORTING INFORMATION

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