Chapter 3
The Biochemistry and Molecular Biology of Volatile Messengers in Trees

Hamid Rajabi Memari, Leila Pazouki, and Ülo Niinemets

Abstract All tree species possess genes encoding terminal enzymes responsible for volatile isoprenoid synthesis. However, only in some species, these genes are expressed constitutively in leaves, while terpenoid emissions can be triggered by abiotic and biotic stress factor in essentially all species. This chapter analyses the biochemical diversity of volatile isoprenoid synthases and investigates the genomic modifications responsible for constitutive volatile production in trees. Plant terpenoids are up to three-domain proteins with either one active center in monofunctional synthases, or two active centers in bifunctional synthases. There is evidence of monophyletic origin of modern plant terpenoid synthases from a three-domain synthase in an ancient progenitor followed by extensive gene duplication and domain loss. The terpenoid synthase sequence similarity can be low among distant plant groups, but terpenoid tertiary structure is remarkably similar in different synthases, and this structural similarity is even conserved across domains of life. However, only minor changes in active center structure can lead to major changes in product profiles, indicating that presence of rich terpenoid genetic diversity constitutes an important means for rapid evolutionary adaptations to novel biotic interactions, and to new abiotic stresses in plant habitats.

H. Rajabi Memari
Genetic Engineering and Molecular Genetics, Biotechnology and Life Science Center and School of Agriculture, Shahid Chamran University of Ahvaz, Golestan BLV, Ahvaz, Iran

L. Pazouki (✉) • Ü. Niinemets
Institute of Agricultural and Environmental Sciences, Estonian University of Life Sciences, Kreutzwaldi 1, Tartu 51014, Estonia
e-mail: leila.pazouki@emu.ee; ylo.niinemets@emu.ee

3.1 Introduction

Humans have acknowledged the existence of terpenoids (isoprenoids) in plants from ancient times without any information about their chemistry and role in plants. The term terpene comes from the word turpentine, which is the resin that oozes out from the bark of pine trees after wounding. This sticky resin is rich in a variety of terpenoid compounds with widely differing chemical structure and physico-chemical characteristics (Crozier et al. 2006). All terpenoids consist of isoprene (C5) building blocks and are conditionally divided between primary metabolites such as carotenoids and hormones and secondary metabolites such as volatile terpenoids isoprene and monoterpenes (C10), and semivolatile terpenoids sesquiterpenes (C15) and diterpenes (C20) fulfilling a variety of functions, some of which are not yet fully understood (Modolo et al. 2009; Fineschi et al. 2013). Altogether, these volatiles and semivolatiles constitute the most import class of biogenic volatile organic compounds (BVOCs) (Kesselmeier and Staudt 1999; Dudareva et al. 2004; Holopainen 2004; Guenther et al. 2006; Laothawornkitkul et al. 2009).

Volatile terpenoids are released by different plant tissues including leaves, buds, flowers, fruits and roots (Dudareva and Pichersky 2008), but vegetative plant parts, in particular, leaves are believed to contribute the most to plant emissions due to high leaf mass fraction and generally the highest emission rates per foliage mass (Kesselmeier and Staudt 1999). Vegetative plant parts can release diverse mixtures of terpenoids, including isoprene, monoterpenes, sesquiterpenes and some diterpenes (Keeling and Bohlmann 2006a). These volatiles play different roles in plants including enhancement of abiotic stress tolerance by serving as antioxidants (Loreto et al. 2001b; Vickers et al. 2009a; Possell and Loreto 2013) and possibly also by modifying membrane fluidity (Sharkey and Singsaas 1995; Singsaas et al. 1997; Behnke et al. 2007), and serving as direct defences against omnivorous insects, e.g., oleoresin is a common direct defence in conifers against pathogens and herbivores (Bohlmann and Croteau 1999; Raffa et al. 2005; Heijari et al. 2008; Chen et al. 2011). In addition, these volatiles serve as indirect defences participating in within-plant and among-plant communication and stress priming and in communication with herbivore enemies (Arimura et al. 2000; Dicke and Bruin 2001; Dudareva et al. 2006; Pieterse and Dicke 2007; Choudhary et al. 2008; Dicke et al. 2009).

Due to the importance of volatile terpenoids in plant life, there is a continuing interest in plant terpenoid synthesis pathways and terpene synthase (Tps) genes as targets to increase plant stress tolerance by enhanced terpenoid content and emission of volatiles attracting herbivore parasites and predators (Pichersky and Gershenzon 2002; Degenhardt et al. 2003; Dudareva and Pichersky 2008). Understanding structural, functional and evolutionary features of Tps family in trees can help to discover defence systems in these plants and use this information for breeding and tree protection in forestry. Furthermore, some of these volatiles or their derivatives

have cosmetic or medicinal properties, which have made them as important targets for fragrance and pharmaceutical science and industry, for selection of promising genotypes and genetic engineering to overexpress the pathway in the given or a new host organism (Seigler 1998; Braun et al. 2001; Schepmann et al. 2001; Huang et al. 2004; Aharoni et al. 2005; Bohlmann and Keeling 2008; Saranitzky et al. 2009; Trusheva et al. 2010). Also, in last decades, the pathways of volatiles have been widely used and engineered to produce different types of biofuels (Peralta-Yahya et al. 2012).

From air chemistry and climate perspective, volatile isoprenoids significantly contribute to photochemical reactions in the atmosphere, participating in the formation of ozone, secondary organic aerosols and cloud condensation nuclei, thereby altering air quality, and solar radiation penetration (Huff Hartz et al. 2005; Guenther et al. 2006; Lee et al. 2006; Engelhart et al. 2008; Ashworth et al. 2013; Kulmala et al. 2013). For example, the hemiterpene (C5) isoprene (2-methyl-1,3-butadiene) is worldwide the most important volatile isoprenoid with global emissions about 440–660 Tg C year^{-1} (Guenther et al. 2006; Ashworth et al. 2013). Isoprene is emitted constitutively by leaves of several deciduous angiosperm tree species like *Salix* spp., *Populus* spp., and *Quercus* spp., but also from several North-American evergreen *Quercus* spp. (Kesselmeier and Staudt 1999). Furthermore, some gymnosperms and a number of herb, moss and fern species release isoprene as well (Sharkey and Yeh 2001; Sharkey et al. 2008). Carbon loss due to isoprene emission in constitutive emitters is typically between 1 and 2 % of photosynthetic carbon fixation, but this percentage may increase to more than 50 % under stress conditions (Sharkey and Yeh 2001).

During the past decades, major progress has been made in identification and functional characterization of volatile terpenoid biosynthesis genes, enzymes and in metabolic engineering, and this has greatly contributed to improved understanding of terpenoid biosynthesis (Crozier et al. 2006; Keeling and Bohlmann 2006b; Bohlmann and Keeling 2008; Degenhardt et al. 2009; Nagegowda 2010; Chen et al. 2011). Latest developments in molecular techniques, such as new technologies for identification of large genome parts up to full genomes and rapid assessment of plant transcriptome are strongly contributing to identifying and isolating new terpenoid genes, studying the synthase reaction mechanisms and understanding their function in plants. In this chapter, we briefly describe the key pathways for immediate substrates for plant terpenoid biosyntheses. Then we analyse isoprene, mono-sesqui- and diterpene biosynthesis, characteristics of involved terpenoid synthases, their regulation and corresponding gene families with special attention to trees species. So far, the research on plant terpenoids has mainly focused on herbaceous species. However, given that the bulk of volatiles released to the atmosphere is believed to come from woody species, clearly there is a pressing need to gain more insight into biochemical and genetic regulation of terpenoid biosynthesis in woody species.

3.2   Terpenoid Biosynthesis

The huge chemical diversity of plant isoprenoids is formed by an extensive array of enzymes that can be divided among three groups. The first group of enzymes serves as the interface between primary and secondary metabolism, being responsible for channeling of metabolites to terpenoid synthesis pathways (Modolo et al. 2009). The second group of enzymes forms the terpenoid molecule scaffolds in terpenoid pathways (Modolo et al. 2009). The third group of enzymes alters the terpenoid backbones, resulting in new molecules with different biological activities (Modolo et al. 2009), e.g., by hydroxylation, epoxidation, arylmigration, glycosylation, methylation, sulfation, acylation, prenylation, oxidation, and reduction (Gowan et al. 1995; Ro et al. 2005; Modolo et al. 2009). Here we briefly outline the basic isoprenoid synthesis pathways. More detailed summary of terpenoid synthesis pathways is provided by Rosenkranz and Schnitzler (2013) and Li and Sharkey (2013b).

3.2.1   Main Biosynthesis Pathways: MVA and MEP/DOXP

All terpenoid compounds are synthesized from the same precursors: isopentenyl diphosphate (IDP) and its isomer dimethylallyl diphosphate (DMADP) (Wanke et al. 2001; Crozier et al. 2006; Bohlmann and Keeling 2008; Modolo et al. 2009). These precursors are synthesized by two different pathways. The cytosolic mevalonic acid (MVA) pathway is present in most eukaryotes and is responsible for the synthesis of C15 (sesquiterpenoids) and C30 (triterpenoids such as sterols) terpenoid compounds in plants (Gershenzon and Croteau 1993). The second recently discovered 2-C-methyl-D-erythritol 4-phosphate/1-deoxy-D-xylulose 5-phosphate pathway (MEP/DOXP pathway) is present in several prokaryotes and in plastids of eukaryotic organisms (Rohmer et al. 1993; Eisenreich et al. 2004; Crozier et al. 2006). The MEP/DOXP pathway is responsible for the synthesis of isoprene (C5), mono- (C10), di- (C20) and tetraterpenoids (C40) in plants (Lichtenthaler et al. 1997a, b; Modolo et al. 2009). The two pathways operate almost independently, although there is a certain cross-talk among the two pathways at the level of IDP (e.g., Hemmerlin et al. 2003; Laule et al. 2003).

Isoprenoid synthesis through MVA pathway starts with condensation of three molecules of acetyl coenzyme A (acetyl-CoA) producing 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA). HMG-CoA is further reduced by HMG-CoA reductase to mevalonic acid (MVA), which is phosphorylated by two kinases forming mevalonate 5-diphosphate (MVADP). MVADP is converted into the terpenoid precursor, IDP, by mevalonate diphosphate carboxylase (Gershenzon and Croteau 1993; Eisenreich et al. 2004; Crozier et al. 2006). The cytosolic enzyme isopentenyl diphosphate isomerase (IDI) further catalyzes the reversible conversion between IDP and its isomer DMADP.

The plastidic MEP/DOXP pathway starts with the condensation of the substrates pyruvate and glyceraldehyde 3-phosphate (GAP) to DOXP. Carbon skeleton rearrangements and dehydration steps result in formation of 2-C-methyl-D-erythritol 4-phosphate (MEP). MEP is further converted to 2-C-methyl-D-erythritol 2,4-cyclodiphosphate and to 4-hydroxy-3-methylbut-2-enyl diphosphate (HMBDP). Finally, HMBDP is converted both to IDP and DMADP by HMBDP reductase, and the pool sizes of IDP and DMADP are further modulated by plastidic isopentenyl diphosphate isomerase (Eisenreich et al. 2004; Crozier et al. 2006; Li and Sharkey 2013b).

In the past years, major progress has been made in characterizing the enzymes of MEP/DOXP pathway in plants, but kinetic characteristics of all enzymes are still not available, although the available evidence suggests that they resemble those in bacterial counterparts (Harrison et al. 2013; Li and Sharkey 2013b). However, differently from bacteria, MEP/DOXP pathway in plants can directly accept electrons from photosynthetic electron transport chain. In particular, HMBDP synthase and HMBDP reductase can accept electrons from ferredoxin (Seemann et al. 2006; Seemann and Rohmer 2007), possibly explaining the tight coupling of MEP/DOXP pathway to light reactions of photosynthesis in the chloroplasts. Overall, synthesis of highly reduced terpenoids is energetically costly with synthesis of one isoprenoid C5 residue needing fixation of 6 molecules of CO_2, and consuming 20 ATP, and 14 NADPH molecules (Sharkey et al. 2008), underscoring the need for high plastidic ATP and NADPH status for DOXP/MEP pathway (Rasulov et al. 2011; Li and Sharkey 2013a, b).

3.2.2 From Precursors to Terpenoids

DMADP is the substrate for the synthesis of the smallest isoprenoids, the hemiterpenes isoprene and 2-methyl-3-buten-2-ol (MBO). IDP and DMADP are further substrates for short-chain prenyltransferases (Bohlmann and Keeling 2008; Modolo et al. 2009). The assembly of geranyl diphosphate, the backbone for monoterpenes, by head-to-head addition of IDP and DMADP is catalyzed by GDP synthase. Further farnesyl diphosphate (FDP), the backbone for sesquiterpenes, is formed by condensing GDP and IDP by FDP synthase. Geranylgeranyl diphosphate (GGDP) that is the substrate for diterpene synthesis is formed by condensing FDP and IDP by GGDP synthase. Further, tri- and tetraterpenoids are made by head-to-head condensation of two FDP and two GGDP molecules, respectively (Bohlmann and Keeling 2008; Modolo et al. 2009). The resulting terpenoid polymers are used as precursors by terpene synthases/cyclases and enter into synthesis of primary terpenoid compounds such as sterols, phytyl-chain of chlorophyll and carotenoids (Modolo et al. 2009). Terpenes can further be modified by hydroxylation and oxidization by cytochrome P450-dependent enzymes (Ro et al. 2005; Keeling and Bohlmann 2006a).

3.3   Terpenoid Synthases

Terpenoid synthases form a diverse class of enzymes catalyzing formation of molecules with different chain length, including hemiterpene synthases (C5), monoterpene synthases (C10), sesquiterpene synthases (C15), and diterpene synthases (C20). Being often the end-points of the pathway, they are the key terminal flux-controlling enzymes. So far, over 60,000 members of the terpenoid family are recognized, and a broad grouping of structures has been clarified (http://dnp.chemnetbase.com/) (Xie et al. 2012). There has been a major progress in understanding the function and structure of terpenoid synthases catalyzed by rapid developments in molecular biology techniques allowing for dissection of the structure of terpenoid synthase gene families and heterologous expression and study of recombinant terpenoid synthases (Degenhardt et al. 2009). A number of model terpene synthases has been characterized in detail (Hyatt et al. 2007; Köksal et al. 2010, 2011a, b; McAndrew et al. 2011; Zhou et al. 2012), but we are just starting to understand the size and structure of terpene synthase families in key organisms (Sect. 3.4). Furthermore, there is less information on tree terpenoid synthases than on synthases in herbaceous species, except perhaps for gymnosperms (Degenhardt et al. 2009). Nevertheless, the progress has been amazingly rapid as new biotechnology techniques are providing strong tools to understand the complex biochemical function and regulation of genes involved in the terpenoid pathways. The emergence of new high throughput techniques such as deep sequencing together with developments in computational bioinformatics is allowing shedding light on previously hidden aspects of genomes, transcriptomes, proteomes, metabolomes and finally terpenomes with unprecedented detail (Christianson 2008; Cane and Ishida 2012). Here we introduce the basic contemporary methods to study terpenoid synthases and analyse the basic functional structure of terpenoid synthases with emphasis on tree enzymes.

3.3.1   Identification and Analysis of the Functional Activity of Terpenoid Synthases

Due to simultaneous expression of multiple terpenoid synthases in plant tissues and relatively low product specificity of most synthases (Sect. 3.3.3.4), functional analysis of terpenoid synthases based on enzyme purification from crude leaf extracts is difficult, although a lot of pioneering work has been conducted using partially purified enzymes extracted from plants (e.g., Croteau and Karp 1977; Croteau et al. 1978; Dehal and Croteau 1988). Development of molecular biology techniques for identification and isolation of individual terpenoid synthases has opened completely new vistas for studying biochemistry, structure, genetics and evolution of terpenoid synthases. Development of RNA sequencing platforms in recent years using deep-sequencing technologies have changed the transcriptomics world and sometimes

predicate the death of micro-array and other transcriptome analysis technologies like serial analysis of gene expression (SAGE, Velculescu et al. 1995), cap analysis gene expression (CAGE, Shiraki et al. 2003), and massively parallel signature sequencing (MPSS, Brenner et al. 2000). However, the high throughput technologies have some disadvantages like high cost, inability to detect transcripts for isoforms and splice variation (Wang et al. 2009; Myllykangas et al. 2012).

The workflow for functional characterization of given terpene synthase typically consists of determination of the sequence of terpenoid synthase genes either on the basis of mRNA or genomic DNA, heterologous expression in a host system of the sequenced gene and functional characterization of the recombinant protein. Ultimately, the function can be further studied in a transgenic plant model system. Here these basic steps are briefly outlined.

3.3.1.1 Identification of Terpenoid Synthase Genes

In the infancy of terpenoid molecular and functional studies, identification of terpene synthase genes was a highly tedious task due to lack of information on homologous sequences for degenerate primer construction. Now, as more and more synthases have been sequenced, the rich existing genetic information allows for more rapid progress, albeit identification of terpenoid synthases with low level of homogeneity with those described so far, especially in organisms with little genome coverage is still difficult (Cane and Ishida 2012). By now, 51 genomes of higher plants (38 published) have been fully sequenced (as of Sept. 4, 2012, http://genomevolution.org/wiki/index.php/Sequenced_plant_genomes), making it possible to identifiy putative terpenoid synthases by genome mining. Genome sequence analysis of *Arabidopsis thaliana* has identified about 30 terpene synthase genes in this model organism (Aubourg et al. 2002), but in several vascular plant species much larger terpene synthase families have been detected (Sect. 3.4.2, Martin et al. 2010; Li et al. 2012), especially by using the widest range of terpenoid synthase sequences possible in homology searches (Li et al. 2012). In particular, information about the size and structure of tree terpene synthase gene families has been exponentially increasing since the first tree, *Populus trichocarpa* (Tuskan et al. 2006), genome completion, followed by other tree genome projects including *Carica papaya* (Ming et al. 2008), *Malus domestica* (Velasco et al. 2010), *Phoenix dactylifera* (Al-Dous et al. 2011), and *Pyrus bretschneideri* (Wu et al. 2013). New bioinformatics tools such as the use of profile-based hidden Markov models rather than pairwise searchers makes it possible to identify terpene synthase genes carrying remote sequence homology, thereby identifying putative terpene synthases with potential structural similarity of basic conserved functional domains (Gough et al. 2001; Wilson et al. 2009; Cane and Ishida 2012).

Although the number of fully sequenced genomes is rapidly increasing, the genome of only 11 tree species has been sequenced, most of them tree crops. Among sequenced trees, only *Populus trichocarpa* and *Eucalyptus grandis* can be conditionally considered as 'wild plants'. Thus, techniques alternative to genome

mining need to be applied to most tree species. New high throughput techniques (e.g., next generation sequencing, Liu et al. 2012) have opened up possibilities for fast characterization of transcriptome, and identification of expressed terpenoid synthases by transcriptome mining. Conifers are characterized by a particularly rich blend of terpene volatiles, suggesting a highly diverse terpenoid family, but conifer genome is especially complex, and full sequences of first conifer genomes are unlikely in the near future, although a number of genome projects has started (see e.g., http://pinegenome.org/). Thus, relatively few terpenoid synthases have been functionally characterized in conifer species so far, although more than in angiosperm trees (e.g., Wildung and Croteau 1996; Bohlmann et al. 1997, 1999; Hall et al. 2011). In these pioneering studies, different techniques were used to identify conifer terpene synthases, such as cDNA library screening and similarity-based PCR (Bohlmann et al. 1998a). Expressed sequence tag (EST) libraries have been available lately for loblolly pine (*Pinus taeda*) (Allona et al. 1998), Japanese cedar (*Cryptomeria japonica*) (Ujino-Ihara et al. 2000), white spruce (*Picea glauca*), interior spruce (*P. glauca × P. engelmannii*), and Sitka spruce (*Picea sitchensis*) (Holliday et al. 2008). Moreover, lots of attempts are currently in progress, trying to identify and characterize more terpene synthases in conifers using a combination of targeted cDNA cloning, large amounts of ESTs and full-length cDNA mining (Byun McKay et al. 2003; Miller et al. 2005; Keeling et al. 2011) as well as using other techniques such as bacterial artificial chromosome (BAC) technique (Hamberger et al. 2009). Use of these new methods has allowed identification of large terpenoid synthase families in several tree species, e.g., 69 actively expressed Tps genes (including monoterpene, sesquiterpene and diterpene synthases) have been identified in *Picea* species (Keeling et al. 2011).

3.3.1.2 Heterologous Expression of Tree Terpenoid Synthases in *E. coli* and in Plants

Functional characterization of terpene synthases is usually carried out by heterologous expression in *Escherichia coli* (Table 3.1 for a selected list of tree terpenoid synthases expressed in *E. coli*). However, expression of terpene synthase genes in *E. coli* carries potential problems. Some terpene synthase proteins are produced in cytosol, but are targeted to chloroplastic compartment, therefore having chloroplastic signal peptides. These transient sequences should be removed before cloning and expression in the host for sufficiently high expression of recombinant protein (Phillips et al. 2003). On the other hand, codon usage of eukaryotic genes is different from prokaryotic host, and the eucaryotic genome also contains rare codons (Fig. 3.1). The comparison between codon usages in the angiosperm *Salix discolor* and *Escherichia coli* shows the mean difference close to 30 % in codon usage, whereas the difference between the angiosperm *Populus trichocarpa* and the gymnosperm *Abies grandis* is about 10 %. Thus, for high level of expression, co-transformation of a plasmid encoding the rare tRNAs (e.g., for Arg) is needed (Hohn 1999; Martin et al. 2004). The codon optimization for gene expression of tree Tps

Table 3.1 List of selected tree terpenoid synthases expressed and characterized in *Escherichia coli* host system

Terpenoid synthase	Gene family[a]	UniProtKB/ Swiss-Prot entry	Species	Expression host	Substrate	Terpenoid products[b]	References
Hemiterpenes (C$_5$)							
Isoprene synthase	Tps-b	Q50L36	*Populus alba*	*E. coli* origami B (DE3)	DMADP	Isoprene (100 %)	Sasaki et al. (2005)
Isoprene synthase	Tps-b	Q9AR86	*Populus x canescens*	*E. coli* TG1	DMADP	Isoprene (100 %)	Miller et al. (2001)
2-Methyl-3-buten-2-ol (MBO) synthase	Tps-d1	F5CJS6	*Pinus sabiniana*	*E. coli* BL21 (DE3)-RIL	DMADP	Methylbutenol (99 %), isoprene (1 %)	Gray et al. (2011)
Monoterpenes (C$_{10}$)							
(−)-Camphene synthase	Tps-d1	Q948Z0	*Abies grandis*	*E. coli* BL21(DE3)	GDP	(−)-camphene (54 %), (−)-α-pinene (32 %), (−)-limonene	Bohlmann et al. (1999)
(+)-Carene synthase 1	Tps-d1	F1CKI6	*Picea sitchensis*	*E. coli*	GDP	(+)-3-carene (42 %), α-terpinolene (27 %), (+)-sabinene (11 %), (−)-β-phellandrene, myrcene, α-terpinene, γ-terpinene, (−)-α-pinene, α-thujene	Hall et al. (2011)
(+)-Carene synthase 2	Tps-d1	F1CKI8	*Picea sitchensis*	*E. coli*	GDP	(+)-3-carene (64 %), α-terpinolene (15 %), (+)-sabinene (7 %), (−)-β-phellandrene, α-terpinene, γ-terpinene, myrcene	Hall et al. (2011)
(−)-Limonene synthase	Tps-d1	O22340	*Abies grandis*	*E. coli* BL21(DE3)	GDP	(−)-limonene (>90 %), (−)-α-pinene, (−)-β-pinene	Bohlmann et al. (1997)

(continued)

Table 3.1 (continued)

Terpenoid synthase	Gene family[a]	UniProtKB/Swiss-Prot entry	Species	Expression host	Substrate	Terpenoid products[b]	References
(−)-Limonene/(−)-α-pinene synthase	Tps-d1	Q9M7C9	Abies grandis	E. coli BL21(DE3)	GDP	(−)-limonene (35 %), (−)-α-pinene (24 %), (−)-β-phellandrene (20 %), (−)-β-pinene (11 %), (−)-sabinene (10 %)	Bohlmann et al. (1999)
(+)-Limonene synthase	Tps-b	Q8L5K1	Citrus limon	E. coli BL21-CodonPlus-RIL	GDP	(+)-limonene (99 %)	Lücker et al. (2002)
(+)-Limonene synthase	Tps-b	Q8L5K3	Citrus limon	E. coli BL21-CodonPlus-RIL	GDP	(+)-limonene (99 %)	Lücker et al. (2002)
Limonene synthase	Tps-d1	Q675L1	Picea abies	E. coli BL21-CodonPlus	GDP	(−)-limonene (88 %), myrcene, (−)-α-pinene, (+)-limonene	Martin et al. (2004)
Linalool synthase	Tps-d1	Q675L2	Picea abies	E. coli BL21-CodonPlus	GDP	(−)-linalool (97 %), (+)-linalool, E-β-ocimene	Martin et al. (2004)
Myrcene synthase	Tps-d1	O24474	Abies grandis	E. coli XL1-Blue	GDP	Myrcene (100 %)	Bohlmann et al. (1997)
Myrcene synthase	Tps-b	Q93X23	Quercus ilex	E. coli TG1	GDP	Myrcene (97 %), limonene, β-pinene	Fischbach et al. (2001)
Myrcene synthase	Tps-d1	Q675K9	Picea abies	E. coli BL21-CodonPlus	GDP	Myrcene (100 %)	Martin et al. (2004)
(−)-β-Phellandrene synthase	Tps-d1	Q9M7D1	Abies grandis	E. coli BL21(DE3)	GDP	(−)-β-phellandrene (52 %), (−)-β-pinene (34 %), (−)-α-pinene (9 %), (−)-limonene	Bohlmann et al. (1999)
(+)-α-Pinene synthase	Tps-d1	Q84KL3	Pinus taeda	E. coli BL21(DE3)	GDP	(+)-α-pinene (97 %), (−)-α-pinene	Phillips et al. (2003)

(−)-α-Pinene synthase	Tps-d1	Q84KL6	*Pinus taeda*	*E. coli* BL21(DE3)	GDP	(−)-α-pinene (79 %), (−)-β-pinene, (−)-limonene, (+)-limonene, (+)-camphene, (−)-camphene, (+)-α-pinene, (+)-β-pinene	Phillips et al. (2003)
(−)-β-Pinene synthase	Tps-b	Q8L5K2	*Citrus limon*	*E. coli* BL21-CodonPlus-RIL	GDP	(−)-β-pinene (81 %), sabinene (11 %), (−)-α-pinene, (−)-limonene	Lücker et al. (2002)
(−)α/β-Pinene synthase	Tps-d1	Q675L3	*Picea abies*	*E. coli* BL21(DE3)	GDP	(−)-α-pinene (57 %), (−)-β-pinene (27 %), β-phellandrene (11 %), myrcene	Martin et al. (2004)
Pinene synthase	Tps-d1	Q6XDB5	*Picea sitchensis*	*E. coli* BL21-CodonPlus (DE3)	GDP	(−)-α-pinene (65 %), (−)-β-pinene (18 %),	Byun McKay et al. (2003)
(−)-Pinene synthase	Tps-d1	O24475	*Abies grandis*	*E. coli* XL1-Blue/*E. coli* XLOLR	GDP	(−)-α-pinene (42 %), (−)-β-pinene (58 %)	Bohlmann et al. (1997)
(+)-Sabinene synthase	Tps-d1	F1CKJ1	*Picea sitchensis*	*E. coli*	GDP	(+)-sabinene (54 %), α-terpinolene (31 %), (−)-α-pinene, α-terpinene, (−)-β-phellandrene, γ-terpinene, myrcene, (+)-3-carene	Hall et al. (2011)

Table 3.1 (continued)

Terpenoid synthase	Gene family[a]	UniProtKB/Swiss-Prot entry	Species	Expression host	Substrate	Terpenoid products[b]	References
γ-Terpinene synthase	Tps-b	Q8L5K4	Citrus limon	E. coli BL21-CodonPlus-RIL	GDP	γ-terpinene (71 %), (−)-limonene (9 %), (−)-α-pinene, (+)-β-pinene, terpinolene, α-thujene, (+)-limonene, α-terpinene	Lücker et al. (2002)
Terpinolene synthase	Tps-d1	Q9M7D0	Abies grandis	E. coli BL21(DE3)	GDP	Terpinolene (42 %), (−)-α-pinene (18 %), (−)-limonene (11 %), (−)-β-pinene (10 %)	Bohlmann et al. (1999)
(−)-α-Terpineol synthase	Tps-d1	Q84KL4	Pinus taeda	E. coli BL21(DE3)	GDP	α-terpineol (57 %), limonene (28 %), terpinolene (8 %), β-pinene, α-pinene, myrcene[c]	Phillips et al. (2003)
(+)-α-Terpineol synthase	Tps-b	B5A434	Santalum album	E. coli C41	GDP	(+)-α-terpineol (44 %), (−)-limonene (33.6 %), E-geraniol, linalool, myrcene, (−)-α-pinene, (+)-sabinene, α-terpinolene	Jones et al. (2008)
Sesquiterpenes (C$_{15}$)							
α-Bisabolene synthase	Tps-d3	O81086	Abies grandis	E. coli XL1-Blue	FDP	E-α-bisabolene (>99 %)	Bohlmann et al. (1998a)
					GDP	(+)-limonene (main product)	
E-α-Bisabolene synthase	Tps-d3	Q675L6	Picea abies	E. coli BL21-CodonPlus	FDP	E-α-bisabolene (100 %)	Martin et al. (2004)

	Tps type	Accession	Species	Host organism	Substrate	Products	Reference
α-Farnesene synthase	Tps-b	Q84LB2[d]	Malus domestica	E. coli DH5α	FDP	E,E-α-farnesene (>99 %), Z,E-α-farnesene (<1 %)	Pechous and Whitaker (2004)
					GDP	E-β-ocimene (68 %), (Z)-β-ocimene (32 %)	Pechous and Whitaker (2004)
α-Farnesene synthase	Tps-b	Q84LB2[e]	Malus domestica	E. coli BL21-CodonPlus-RIL	FDP	E,E-α-farnesene (96 %), Z,E-α-Farnesene (<1 %)	Green et al. (2007)
					GDP	E-β-ocimene (90 %), linalool, myrcene	Green et al. (2007)
E,E-α-Farnesene synthase	Tps-d1	Q675K8	Picea abies	E. coli BL21-CodonPlus	FDP	E,E-α-farnesene (100 %)	Martin et al. (2004)
(−)-Germacrene D synthase	Tps-a	Q64K29	Populus trichocarpa x P. deltoides	E. coli BL-21-CodonPlus	FDP	(−)-germacrene D (79 %), β-caryophyllene, α-humulene, alloaromadendrene and 4 non-identified terpenoids	Arimura et al. (2004)
γ-Humulene synthase	Tps-d2	O64405	Abies grandis	E. coli XL1-Blue	FDP	γ-humulene (dominant), sibirene, longifolene, β-himachalene, γ-himachalene, α-himachalene	Steele et al. (1998)
Longifolene synthase	Tps-d2	Q675L0	Picea abies	E. coli BL21-CodonPlus	FDP	Longifolene (61 %), α-longipinene (15 %), longicyclene, E-β-farnesene, longiborneol, cyclosativene, β-longipinene	Martin et al. (2004)

Table 3.1 (continued)

Terpenoid synthase (C$_{20}$)	Gene family[a]	UniProtKB/Swiss-Prot entry	Species	Expression host	Substrate	Terpenoid products[b]	References
Diterpenes (C$_{20}$)							
Abietadiene synthase	Tps-d3	Q38710	Abies grandis	E. coli BLR/E. coli XL1-Blue	GGDP	Abietadiene (mixture of double bond isomers)	Peters and Croteau (2002) and Stofer Vogel et al. (1996)
Isopimaradiene synthase	Tps-d3	Q675L5	Picea abies	E. coli BL21-CodonPlus	GGDP	Isopimara-7,15-diene (100 %)	Martin et al. (2004)
Levopimaradiene synthase	Tps-d3	Q947C4	Ginkgo biloba	E. coli	GGDP	Levopimaradiene (main product)	Schepmann et al. (2001)
Levopimaradiene synthase	Tps-d3	Q675L4	Picea abies	E. coli BL21-CodonPlus	GGDP	Levopimaradiene (37 %), abietadiene (32 %), neoabietadiene (23 %), palustradiene	Martin et al. (2004)
Taxadiene synthase		Q41594	Taxus brevifolia	E. coli JM109/ E. coli BL21(DE3)	GGDP	Taxadiene (mixture of double bond isomers)	Wildung and Croteau (1996), Williams et al. (2000) and Jin et al. (2005)

[a] According to Bohlmann et al. (1998b), Byun McKay et al. (2003), and Martin et al. 2004
[b] Only minor products ≥1 % are demonstrated
[c] Dominated by (−) isomer, except for α-pinene (100 % (+)-α-pinene)
[d] GenBank accession number AY182241.2, expressed with C-terminal Myc-tag (Pechous and Whitaker 2004)
[e] GenBank accession number AY787633, expressed with N-terminal His-tag (Green et al. 2007)
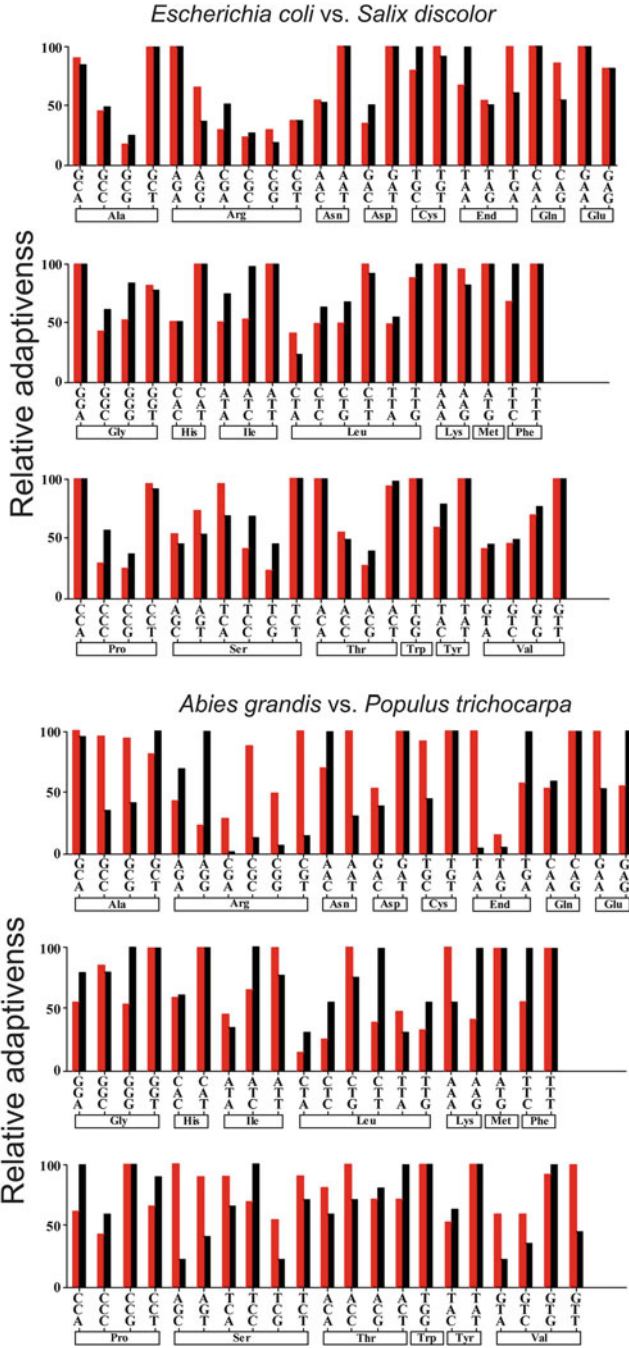
genes can also be achieved by using suitable host strains. Rosetta host strains are BL21 derivatives designed to enhance the expression of eukaryotic proteins that contain rare codons (like tree codons) which are seldom used in *E. coli*. This strategy can facilitate overexpression and characterization of different tree Tps genes in *E. coli* (Kane 1995).

The heterologous expression of Tps synthase in *E. coli* is followed by disruption of cellular contents of *E. coli* cultures carrying the transgenic construct, enzyme purification whenever needed, and assay of enzymatic activity (Martin et al. 2004; Sasaki et al. 2005; Majdi et al. 2011). Typically, the functional characterization of the protein involves incubation of the soluble recombinant enzyme with the substrate GDP, FDP, or GGDP in the presence of Mg^{2+} and/or Mn^{2+}, and analysis of the product profiles by gas chromatography – mass spectrometry (GC-MS) and identification of terpenoids by authentic standards (Bohlmann et al. 1997; Martin et al. 2004; Falara et al. 2011). Typically, an overlay of pentane or other hydrophobic solvent is used to trap the hydrophobic reaction products formed in the aqueous reaction mixture (e.g., Martin et al. 2004), or volatiles can also be sampled in the head-space using a solid-phase micro extraction (SPME) fiber (e.g., Falara et al. 2011) or using sample air from the headspace with a GC preconcentration trap (e.g., Fischbach et al. 2001).

Although heterologous expression offers an unique opportunity to work with the isolated protein without other potentially interfering proteins, the kinetic characteristics of the recombinant enzyme can differ from the native enzyme, but not necessarily substantially, e.g., as demonstrated by similar K_m values for GGDP for the diterpene taxadiene native and recombinant enzyme (Williams et al. 2000). Also, enzyme characteristics can depend on the transgenic construct, e.g., farnesene synthase expressed with C-terminal Myc-tag (Pechous and Whitaker 2004) and N-terminal His-tag (Green et al. 2007) had somewhat different product profiles (Table 3.1). Expression of poplar isoprene synthase either with N-terminal or C-terminal His-tag resulted in altered pH and temperature dependence and substrate specificity (Schnitzler et al. 2005). Thus, some caution is warranted when making inferences on the performance of native enzyme on the basis of measurements with recombinant protein.

Heterologous expression in *E. coli* can be followed by characterization of the functional role of the protein in plants as discussed in detail by Rosenkranz and Schnitzler (2013). For example, transgenic *Arabidopsis* model systems expressing isoprene synthase from white poplar (*Populus alba*) (Sasaki et al. 2007), grey poplar (*P. x canescens*) (Loivamäki et al. 2007a, 2008) and kudzu (*Pueraria lobata*) (Velikova et al. 2011), and transgenic tobacco (*Nicotiana tabacum*) expressing *P. alba* isoprene synthase gene (Vickers et al. 2009b, 2011) are available. Also, herbaceous model systems overexpressing certain mono-, sesqui- and diterpenes have been constructed (El Tamer et al. 2003; Besumbes et al. 2004; Wu et al. 2006). However, no such model systems were yet available for trees. In this book, Rosenkranz and Schnitzler (2013) first time describe successful introduction of poplar isoprene synthase gene into silver birch (*Betula pendula*), providing a new exciting model to test the role of isoprene synthase in plants.

Escherichia coli vs. *Salix discolor*

Abies grandis vs. *Populus trichocarpa*

Relative adaptivenss

3.3.2 Conserved Motifs and Functional Domains of Terpenoid Synthases
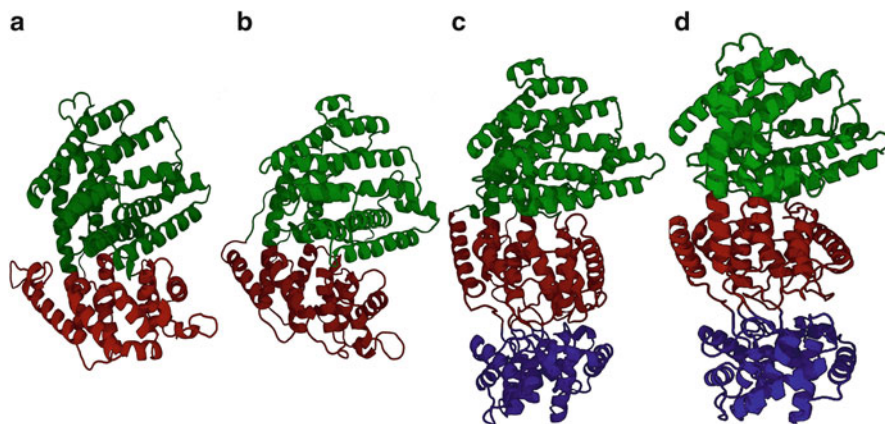
3.3.2.1 Conserved Motifs

According to the mechanism of catalysis, terpenoid synthases can be separated among two major classes. In the case of class I enzymes, the catalysis is initiated by metal-triggered ionization of the substrate diphosphate group, while for class II enzymes, the catalysis is initiated by protonation of an epoxide ring or carbon–carbon double bond (Christianson 2006, 2008; Aaron and Christianson 2010; Cao et al. 2010). In both cases, a highly reactive carbocation is formed that enters into isomerization and cyclization steps until the catalysis is terminated by either proton elimination or nucleophilic capture from the final carbocation (Aaron and Christianson 2010; Cao et al. 2010; Köksal et al. 2011a). Thus, the primary difference among the class I and class II terpenoid synthases is the initial step of the catalysis. These differences in catalytic mechanism also reflect different origin of terpenoid synthases with all type I synthases sharing the 'type I synthase fold', an α-helical structure, containing a core of bundled anti-parallel α-helices, while type II enzymes have a characteristic "α-barrel" structure (Aaron and Christianson 2010; Cao et al. 2010).

These differences in the catalytic mechanism are reflected in differences in conserved motifs and functional domains. Class I terpenoid synthases are characterized by the aspartate (D)-rich DDXXD and (N/D)DXX(S/T)XXXE (N is aspargine, S is serine and T is threonine, X can be any amino acid) motifs (Fig. 3.2) that are responsible for binding of divalent metal ions, in particular Mg^{2+} or Mn^{2+}; these metal ions are responsible for diphosphate elimination from the substrate, resulting in carbocation formation (Starks et al. 1997; Degenhardt et al. 2009; McAndrew et al. 2011). The DDXXD motif is typically located at the entrance position of the catalytic site and plays a prominent role in positioning the substrate for catalytic reaction, while (N/D)DXX(S/T)XXXE motif is located at the opposite site of the active site entry, with the underlined amino acids coordinating the metal cations (Little and Croteau 2002; Degenhardt et al. 2009; Cao et al. 2010; Köksal et al. 2011a). Mutational analysis of these motifs in diterpene abietadiene synthase

---

Fig. 3.1 Comparison of codon usage between (*upper panels*) the bacterium *E. coli* (*red bars*) and the angiosperm *Salix discolor* (*black bars*) and between (*lower panels*) the gymnosperm *Abies grandis* (*red bars*) and the angiosperm *Populus trichocarpa* (*black bars*). Relative adaptiveness is an index of usage of synonymous codons that scales the frequency of use of given codon relative to the optimal codon (most frequently used codon with the greatest translation efficiency) (Sharp and Li 1987). This index is calculated for a given codon j and for given amino acid i, w_{ij}, as x_{ij}/x_{imax}, where x_{ij} is the frequency of the use of the given codon and $x_{i,max}$ is that for the optimal codon (Sharp and Li 1987). w_{ij} facilitates comparison of the codon usage in different proteins and among different organisms. Codon frequency is based on NCBI GenBank (www.kazusa.or.jp/codon/). The data was extracted and analysed by graphical codon usage analyzer (http://gcua.schoedl.de/)

```
293  YRKRHVEYYSWVVMCIFEPEFSESRIAFAKTAILCTVLDDLYDTHATLHEIKIMTEGVRR  352
528  ARERPLEFYFLVAAGTYEPQYAKCRFLFTKVACLQTVLDDMYDTYGTLDELKLFTEAVRR  587
305  YRKCYVEFYFWMAAAISEPEFSGSRVAFTKIAILMTMLDDLYDTHGTLDQLKIFTEGVRR  364
      *:   :*:*   :.    **::: .*. *:* * * *:***:***:.**.::*::**.***

353  WDLSLTDDLPDYIKIAFQFFFNTVNELIVEIVKRQGRDMTTI-VKDCWKRYIESYLQEAE  411
588  WDLSFTENLPDYMKLCYQIYYDIVHEVAWEAEKEQGRELVSFF-RKGWEDYLLGYYEEAE  646
365  WDVSLVEGLPDFMKIAFEFWLKTSNELIAEAVKAQGQDMAAYIRKNAWERYLEAYLQDAE  424
      **:*:.: ***::*:.:::   .   .*:   *   * **::.:     :. *: *: .* ::**

412  WIATGHIPTFNEYIKNGMASSGMCILNLNPLLLLD-KLLPDNILEQIHSPSK--ILDLLE  468
647  WLAAEYVPTLDEYIKNGITSIGQRILLLSGVLIMDGQLLSQEALEKVDYPGRRVLTELNS  706
425  WIATGHVPTFDEYLNNGTPNTGMCVLNLIPLLLMG-EHLPIDILEQIFLPSR--FHHLIE  481
      *:*:  ::**::**::**   . *  :* *  :*::  : *   :  **::   *.:  :  .* .

469  LTGRIADDLKDFEDEKERGEMASSLQCYMKENPESTVENALNHIKGILNRSLEEFNWEFM  528
707  LISRLADDIKTYKAEKARGELASSIECYMKDHPECTEEEALDHIYSILEPAVKELTREFL  766
482  LASRLVDDARDFQAEKDHGDL-SCIECYLKDHPESTVEDALNHVNGLLGNCLLEMNWKFL  540
      * .*:.** :  :: ** :*:: *.::**:*:.**.* *:**:*: .:*   .: *:. :*:
```

Fig. 3.2 Two DDXXD metal binding motifs in sesquiterpene synthases in the gymnosperm grand fir (*Abies grandis*). The first sequence is for δ-selinene synthase (ag4, UniProtKB/Swiss-Prot entry O64404), the second for α-bisabolene synthase (ag1, O81086) and the third for γ-humulene synthase (ag5, O64405) (The sequences were aligned using UniProt protein sequence database, http://www.uniprot.org/align/)

in *Abies grandis* demonstrated that mutations in metal-coordinating amino acids strongly reduced the enzyme catalytic activity (Zhou and Peters 2009). However, for some terpene synthases, e.g., the sesquiterpene synthases γ-humulene synthase (ag5) and δ-selinene synthase (ag4) from *Abies grandis*, the sequence information indicates that (N/D)DXX(S/T)XXXE motif is replaced by another DDXXD motif (Back and Chappell 1996; Steele et al. 1998).

Class II terpene synthases contain a conserved DXDD motif which is responsible for protonation of the double bond in the initial reaction step (Cao et al. 2010; Köksal et al. 2011a; Zhou et al. 2012). Mutations in the central aspartic acid render the active site completely non-functional (Christianson 2006; McAndrew et al. 2011). As the diphosphate cleaving activity is missing in class II terpenoid cyclases, they commonly yield diphosphorylated products such as the diterpenoid copalyl diphosphate (CDP) formed by CDP synthase (Keeling et al. 2010; Köksal et al. 2011a) unless they use non-diphosphorylated substrates such as the bacterial squalene-hopene synthase (Wendt et al. 1997). Exceptionally, class I monoterpene bornyl diphosphate synthase from *Salvia officinalis* forms a diphosphorylated product, but in this unusual reaction mechanism, diphosphate residue is first cleaved and then again reincorporated (Whittington et al. 2002).

3.3.2.2 Functional Domains of Terpenoid Synthases

Terpenoid synthases consist of up to three functional domains, α-, β and γ-domain. α-domain bears class I terpene synthase activity, β-domain class II terpene synthase

Fig. 3.3 Structure of selected tree terpenoid synthases: (**a**) hemiterpene isoprene synthase from *Populus* x *canescens* (Protein Data Bank, http://www.rcsb.org/pdb, PDB ID: 3N0F) (Köksal et al. 2010), (**b**) sesquiterpene δ-cadinene synthase from *Gossypium arboreum* (PDB ID: 3G4F, Gennadios et al. 2009), (**c**) sesquiterpene α-bisabolene synthase from *Abies grandis* (PDB ID: 3SAE, McAndrew et al. 2011), and (**d**) diterpene abietadiene synthase from *Abies grandis* (PDB ID: 3S9V, Zhou et al. 2012). Different colours correspond to α- (green, C-terminus), β- (red, N-terminus) and γ-domain (blue). The active site for the terpenoid synthases in (**a–c**) is in the α-domain. The bifunctional abietadiene synthase has two active sites, one in the α-domain and the other in the β-domain. The illustrations were generated by Protein Homology/analogY Recognition Engine 2.0 (PHYRE 2.0, http://www.sbg.bio.ic.ac.uk/phyre2) (Kelley and Sternberg 2009) that uses the protein structure library stored in Protein Data Bank (http://www.rcsb.org)

activity, while γ-domain does not have any known catalytic site (Aaron and Christianson 2010; Cao et al. 2010; Köksal et al. 2011a). Although the domains form clearly separate folds in protein tertiary structure (Fig. 3.3), in the protein amino acid sequences, γ-domain sequence is typically embedded within the β-domain sequence (Wendt et al. 1997; Zhou et al. 2012). In nature, α-domain synthases and β-γ-domain synthases can be found in several organisms such as bacteria and fungi. For example, sesquiterpene pentalenene synthase from the actinobacterium *Streptomyces* (Cane et al. 1994; Lesburg et al. 1997; Caruthers et al. 2000), sesquiterpene aristolochene synthase from the fungus *Penicillium roqueforti* (Caruthers et al. 2000) and sesquiterpene trichodiene synthase from the fungus *Fusarium sporotrichioides* (Rynkiewicz et al. 2001) are α-domain only class I terpenoid synthases, whereas triterpenoid squalene-hopene cyclase from the firmicute *Alicyclobacillus acidocaldarius* (Wendt et al. 1997; Siedenburg and Jendrossek 2011) and diterpene tuberculosinol diphosphate synthase from the actinobacterium *Mycobacterium tuberculosis* (Nakano and Hoshino 2009) are β-γ-domain terpenoid synthases. To our knowledge, none of the plant species possesses β-γ-domain synthases, and most of the plant terpene synthases either contain α-β-domains or all the three domains, α-β-γ (Cao et al. 2010; Hillwig et al. 2011). These plant terpenoid synthases have been postulated to originate from a fusion of an α-domain

type and a β-γ-domain type terpenoid synthase in an ancient progenitor (Morrone et al. 2009; Cao et al. 2010, Sect. 3.4.1), resulting in formation of α-β-γ domain synthases followed by γ-domain loss, yielding α-β domain synthases (Hillwig et al. 2011).

Until recently, it was thought that plants do not possess single, α-domain, synthases. However, it was just discovered that phylogenetically old spikemoss *Selaginella muellendorffii* has both "plant-type" α-β-γ- and α-β-type terpenoid synthases, and "microbial-type" α-domain only terpenoid synthases, fundamentally altering our understanding of the structure of plant terpene synthase families (Li et al. 2012).

All proteins can be classified based on the functional domains. Structural Classification of Proteins (SCOP, http://scop.mrc-lmb.cam.ac.uk/scop/) database provides an hierarchical way to systematize proteins among classes, folds, superfamilies and families (Murzin et al. 2001; Andreeva et al. 2008). A terpenoid synthase superfamily embraces all terpenoid synthase protein domains sharing a common evolutionary origin (Wilson et al. 2009). As α-domain and β- or β-γ-domains of terpenoid synthases have different evolutionary origin (Morrone et al. 2009; Cao et al. 2010, Sect. 3.4.1), plant terpenoid synthase protein domains are divided between two superfamilies, the superfamily *Terpenoid synthases* that includes the α-domains of terpenoid synthases (for most plant terpenoid synthases the relevant domain family is 'Terpenoid cyclase C-terminal domain') and superfamily *Terpenoid cyclases/protein prenyltransferases* that includes β- or β-γ-domains of terpenoid synthases (family: 'Terpenoid cyclase N-terminal domain'). As plant terpenoids commonly consist of either α-β- or α-β-γ-domains, they typically belong simultaneously to both superfamilies. γ-domain, sequence of which is embedded within the β-domain sequence, is classified together with β-domain in SCOP (Gough et al. 2001), The superfamily can be highly diverse at the level of amino acid sequences, but the structures of terpenoid synthases within given superfamily are broadly similar (Fig. 3.4, Sect. 3.3.2.3).

3.3.2.3   Structural Alignment of Terpenoid Synthases in Trees

The three-dimensional protein structure reveals ultimate level of structural information directly related to its function. It is possible that any two proteins with low amino acid sequence similarity still have close structural homology, suggesting similar functional activity. In fact, for terpenoid synthases this appears to be the case. The terpenoid synthases are present in multiple domains of life, and the sequence homology can be quite low even within the same domain, and in particular, across the domains of life (Bohlmann et al. 1998b; Chen et al. 2011; Cane and Ishida 2012; Li et al. 2012). As noted in Sect. 3.3.1.1, genomes can be screened for proteins sharing even remote homology with powerful bioinformatics computational algorithms, identifying new terpenoid synthases (see also Sect. 3.3.2.2). This way, the spikemoss "microbial-type" terpenoid gene family has been recently discovered (Li et al. 2012).
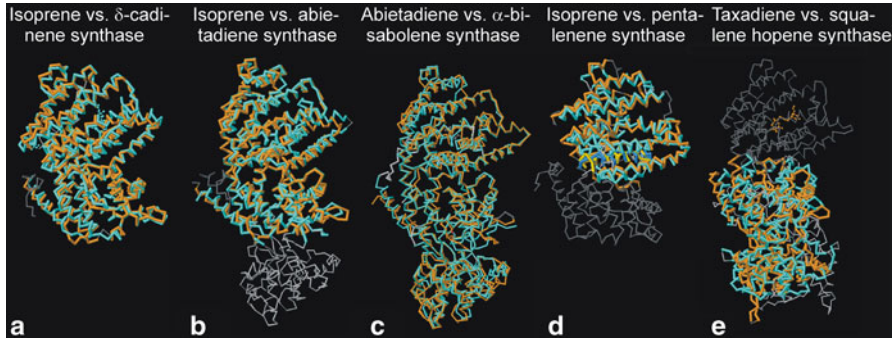
Fig. 3.4 Pairwise structural alignment of tree terpenoid synthases (**a**–**c**) and alignment of proteins structures across different domains of life (**d**, **e**). The proteins aligned in (**a**–**c**) are as described in Fig. 3.3. The sesquiterpene pentalenene synthase in (**d**) is from the actinobacterium *Streptomyces* sp. UC5319 (PDB ID: 1PS1, Lesburg et al. 1997), the diterpene taxadiene synthase in (**e**) is from the gymnosperm *Taxus brevifolia* (PDB ID: 3P5R, Köksal et al. 2011b), and the triterpenoid squalene-hopene cyclase is from the firmicute *Alicyclobacillus acidocaldarius* (PDB ID: 1SQC, Wendt et al. 1997). The alignment was conducted with the Protein Data Bank alignment tool (http://www.rcsb.org) (Berman et al. 2000; Prlic et al. 2010). Pentalenene synthase has only α-domain, isoprene and δ-cadinene synthases have α-β-domains, squalene-hopene cyclase has β-γ-domains and α-bisabolene, abietadiene and taxadiene synthases have α-β-γ-domains. The proteins are oriented as in Fig. 3.3 with the α-domain at the top and β- or γ-domain at the bottom. The olive colour is for the first aligned synthase and cyan for the second, and the grey parts stand for non-aligned sequence components. In (**c**) the sequence similarity is 52, and 97 % of δ-cadinene synthase (sequence length 515 amino acids, AA) is structurally aligned with isoprene synthase (531 AA). In (**b**), the sequence similarity is (36 %) and 95 % of isoprene synthase is aligned with abietadiene synthase (755 AA). In (**c**), the sequence similarity is 48 and 100 % of abietadiene synthase is aligned with α-bisabolene synthase (780 AA). In (**d**), the sequence similarity is 22, and 92 % of pentalenene synthase (304 AA) is aligned with isoprene synthase. In (**e**), the sequence similarity is 17 and 67 % of squalene-hopene synthase is aligned with taxadiene synthase (750 AA)

At present, only a few X-ray crystal structures of plant Tps proteins are available (Starks et al. 1997; Kampranis et al. 2007; Degenhardt et al. 2009; Gennadios et al. 2009; Köksal et al. 2011a, b). In the case of trees, crystal structures are available for isoprene synthase from grey poplar (*P. x canescens*) (Köksal et al. 2010), α-β-domain sesquiterpene synthase δ-cadinene synthase from tree cotton (*Gossypium arboreum*) (Gennadios et al. 2009) and α-β-γ-domain sesquiterpene α-bisabolene synthase from *Abies grandis* (McAndrew et al. 2011), and two α-β-γ-domain diterpene synthases, bifunctional (class I and class II) abietadiene synthase from *A. grandis* (Zhou et al. 2012) and monofunctional (class I) taxadiene synthase from *Taxus brevifolia* (Köksal et al. 2011b) (Fig. 3.3). The three gymnosperm synthases are all from terpene synthase family Tps-d1, while poplar isoprene synthase belongs to Tps-b family and tree cotton δ-cadinene synthase to Tps-a family (Li and Sharkey 2013b for discussion of classification of terpenes into gene families). For α-β-domain monoterpene synthases, the crystal structures are available only for herbs, including Tps-b family monoterpene limonene (Hyatt et al. 2007) and bornyl diphosphate (Whittington et al. 2002) synthases.

Despite limited coverage, available crystal structures have provided major insight into the structure of catalytically active sites, metal binding motifs, substrate recognition and subsequently the function of proteins. Data on protein three-dimensional structure can be employed for protein structural alignment. Structural alignment is a robust tool to compare proteins' tertiary structure and reveal hidden evolutionary relationships, especially for proteins with high evolutionary distance, and consequently with little similarity in their nucleic acid or amino acid sequences. In such cases, the evolutionary relationships between proteins cannot be easily detected by standard sequence alignment techniques, but structural alignment can be used to gain insight into functional relationships among proteins with low sequence homology. For example, despite low level of sequence similarity, poplar isoprene synthase (Tps-b family) and tree cotton δ-cadinene synthase (Tps-a family) exhibit high structural similarity (Fig. 3.4a). High structural similarity is also evident among poplar isoprene synthase and grand fir abietadiene synthase (Tps-d family) α-β-domains (Fig. 3.4b). Furthermore, significant structural similarity is evident between plant and microbial terpenoid synthases. Albeit the sequence similarity is only about 20 % (Fig. 3.4d, e), α-domain alignment of poplar isoprene synthase and bacterial pentalenene synthase, and β-γ-domain alignment of Pacific yew (*Taxus brevifolia*) taxadiene synthase and bacterial squalene-hopene cyclase are remarkably good. Overall, this evidence again emphasizes the strong structural similarity within given domains of terpene synthases across plants and even across the kingdoms of life (Aaron and Christianson 2010; Cao et al. 2010; Cane and Ishida 2012).

3.3.3 Characteristics of Key Plant Terpenoid Synthases

3.3.3.1 Isoprene Synthase

Isoprene synthase (IspS) is the terminal enzyme completing chloroplastic isoprene synthesis through MEP/DOXP pathway (Sharkey and Yeh 2001; Sharkey et al. 2008). IspS crystal structure for recombinant protein from grey poplar (*P. x canescens*) was recently characterized, and it was demonstrated that it is a classic two domain, α-β, terpenoid synthase. The C-terminal class I terpenoid synthase fold (α-domain) possesses the catalytic activity, while the N-terminal class II terpenoid synthase domain (β-domain) possesses no known catalytic activity (Köksal et al. 2010). Formation of isoprene from DMADP occurs through a syn-periplanar elimination mechanism via an allylic carbocation intermediate as in other class I terpenoid synthases (Köksal et al. 2010). The enzyme requires Mg^{2+} for catalytic activity and has a relatively broad alkaline pH optimum between 7 and 8.5, and a temperature optimum between 40 and 45 °C (Monson et al. 1992; Sasaki et al. 2005; Schnitzler et al. 2005). IspS has a high K_m value for DMADP in vivo of ca. 0.3 mM (Rasulov et al. 2009), and there is evidence of allosteric regulation (Schnitzler et al. 2005) and competitive inhibition by GDP, the substrate for monoterpenoid synthases

(Köksal et al. 2010). Multiple isoprene synthase genes have been demonstrated in some poplar species, but the role of these paralogous genes is not yet clear (Vickers et al. 2010). Further details of isoprene synthase are provided in Rosenkranz and Schnitzler (2013) and Li and Sharkey (2013b).

All of the sequenced IspS synthase genes so far have suggested to posses specific chloroplastic signal sequences (Miller et al. 2001; Sasaki et al. 2005; Sharkey et al. 2005; Fortunati et al. 2008; Vickers et al. 2009a, 2010). Localization of IspS to chloroplasts in constitutive emitters has also been confirmed by chloroplast extractions (Wildermuth and Fall 1996; Wildermuth and Fall 1998), immunogold-labelling (Schnitzler et al. 2005), and chloroplast-allocation of green fluorescent protein fused with isoprene synthase in transformed constructs (Sasaki et al. 2005). In fact, transgenic tobacco (*Nicotiana tabacum*) expressing isoprene synthase in cytosol appeared to be essentially void of isoprene emission (Vickers et al. 2011).

3.3.3.2 Terpene Synthases

Terpenes are synthesized by terpene synthases from one of three common prenyl diphosphate precursors formed by the fusion of DMADP with one or more isopentenyl diphosphate (IDP) molecules, catalyzed by prenyltransferases (Chappell 1995; Koyama and Ogura 1999; Dewick 2002).

Typically, tree mono- and sesquiterpene synthases are α-β-domain proteins with only the α-domain (class I) terpene synthase active site functional (Aaron and Christianson 2010; Cao et al. 2010). These terpene synthases are ca. 550–650 amino acids long with sesquiterpene synthases functionally active in the cytosol being characteristically 50–70 amino acids shorter than hemi- and monoterpene synthases that possess a N-terminal plastid-targeting sequence (Bohlmann et al. 1998a; Degenhardt et al. 2009).

Most diterpene synthases and some sesquiterpene synthases are three-domain proteins, 800–870 amino acids long. Again, diterpene synthases are longer than three-domain sesquiterpene synthases due to N-terminal plastid-targeting sequence (Bohlmann et al. 1998a), but some of these three-domain sesquiterpene synthases might contain N-terminal signal sequence (Bohlmann et al. 1998b; Martin et al. 2004). Despite having three domains, only class I terpene synthase active site in the α-domain is functional in sesquiterpenes due to lack of DXDD motif as in α-bisabolene synthase in *Abies grandis* (McAndrew et al. 2011). In the case of diterpenes, commonly either only class I or class II terpene synthase activity is present and the other active site is rendered non-functional (Keeling et al. 2010). Among the diterpenes, taxadiene synthase in Pacific yew (*Taxus brevifolia*) has class I terpene synthase activity (Köksal et al. 2011b), while *ent*-copalyl diphosphate (CDP) synthase in *Arabidopsis* has class II terpene synthase activity (Köksal et al. 2011a).

A few bifunctional diterpene synthases having both class I and class II terpene synthase activities have been reported to date in trees, including abietadiene

synthase from *Abies grandis* (Vogel et al. 1996; Peters et al. 2001; Zhou et al. 2012), *cis*-abienol synthase from *Abies balsamea* (Zerbe et al. 2012b), levopimaradiene synthases from *Gingko biloba* (Schepmann et al. 2001), and *Picea abies* (Martin et al. 2004) and abietadiene/levopimaradiene synthase from *Pinus taeda* (Ro and Bohlmann 2006). In bifunctional diterpene synthases, β-domain with class II activity typically forms a diphosphorylated diterpenoid intermediate (e.g., CDP) that freely diffuses to the second, class I active site in the α-domain, where the diphosphate group is cleaved and given diterpenoid (or typically, a spectrum of terpenoids) is formed (Zhou et al. 2012).

While it is generally thought that mono- and diterpene synthases are functional in plastids, and sesquiterpene synthases in cytosol, there is surprisingly little information on subcellular location of terpene synthases other than the presence or absence of plastid targeting sequences in the expressed proteins (Nagegowda 2010). Available evidence does suggest that monoterpenes are functionally active in plastids (Nagegowda 2010), but α-terpineol synthase in *Magnolia* appeared to be targeted both to chloroplasts and mitochondria (Lee and Chappell 2008), and there are terpene synthases, capable of formation of both mono- and sesquiterpenes depending on substrate (Sect. 3.3.3.3), that can be localized in the cytosol or in the plastids (Aharoni et al. 2004; Nagegowda et al. 2008). On the other hand, while sesquiterpene synthases are generally located in the cytosol, presence of N-terminal residue in some three-domain sesquiterpene synthases in conifers *Abies grandis* and *Picea abies* (Bohlmann et al. 1998b; Martin et al. 2004) suggests that they might be targeted to chloroplasts. Recently, three-domain sesquiterpene santalene/bergamotene synthase from wild tomato (*Solanum habrochaite*s) was shown to be targeted to chloroplasts, suggesting that sesquiterpene synthesis can occur via MEP/DOXP pathway (Sallaud et al. 2009). On the other hand, further biochemical modifications in plastid-synthesized terpenoids, e.g., by cytochrome P450-dependent oxidases typically take place in cytosol (Haudenschild et al. 2000; Ro and Bohlmann 2006; Hamberger et al. 2009). Cleary, more work is needed to gain insight into subcellular location of various terpene synthases, but the evidence suggests that three-domain sesquiterpene synthases with strong homology to diterpenoid synthases could be located in plastids.

In general, terpene synthases have much higher substrate affinity than isoprene synthase with K_m values as low as 0.9 μM (linalool synthase, Pichersky et al. 1995), 7.6 μM (linalool/nerolidol synthases, Nagegowda et al. 2008) to 84 μM (myrcene synthase, Fischbach et al. 2001) for GDP and 1.4 μM (santalene synthase, Jones et al. 2011) – 23 μM (linalool/nerolidol synthases, Nagegowda et al. 2008) for FDP, and 3–10 μM for GGDP (taxadiene synthase, Hezari et al. 1995; Williams et al. 2000). Typically, the pH optimum of terpene synthases is neutral to somewhat alkaline between 6 and 7.5, but the optimum tends to be sharper than that for isoprene synthase (Cori and Rojas 1985; Alonso and Croteau 1993; Bohlmann et al. 1998b). The temperature optimum of terpene synthases seems to be lower than that for isoprene synthase with reported values around 40 °C for conifer *Picea abies* and broad-leaved evergreen angiosperm *Quercus ilex* (Fischbach et al. 2000, 2001).

3.3.3.3 Substrate Specificity of Terpenoid Synthases

Typically, monoterpene synthases use geranyl diphosphate (GDP), sesquiterpene synthases farnesyl diphosphate (FDP) and diterpene synthases geranylgeranyl diphosphate (GGDP) as the only substrate (e.g., Fischbach et al. 2001; Martin et al. 2004 for the tests of multiple substrates). However, a few plant terpene synthases have reported to form terpenoids of different chain length depending on substrate. Among these multifunctional enzymes, α-bisabolene synthase from the gymnosperm tree *Abies grandis* forms sesquiterpene *E*-α-bisabolene with FDP as substrate and monoterpene (+)-limonene with GDP (Bohlmann et al. 1998a, Table 3.1). α-Farnesene synthase from the angiosperm tree *Malus domestica* forms sesquiterpene *E,E*-α-farnesene with FDP and *E*-β-ocimene with GDP (Pechous and Whitaker 2004; Green et al. 2007, Table 3.1). For both sesquiterpene synthases, the enzymes preferably form sesquiterpenes when FDP and GDP are supplied simultaneously (Pechous and Whitaker 2004; Green et al. 2007). Sesquiterpene santalene synthase from *Santalum* species can form a spectrum of monoterpenes when incubated with GDP, but the affinity to GDP is much less than to FDP (Jones et al. 2011). In snapdragon (*Antirrhinum majus*) (Nagegowda et al. 2008) and strawberry (*Fragaria ananassa*) (Aharoni et al. 2004) nerolidol/linalool synthases were shown to form monoterpenes linalool and other acyclic monoterpenes with GDP and nerolidol with FDP. Recently a myrcene/isoprene synthase was characterized in *Humulus lupulus* that formed myrcene with GDP, and isoprene with DMADP (Sharkey et al. 2013).

Capacity to form different products depending on substrate has also been reported for bacterial terpenoid synthases and seems to be widespread in nature (Hamano et al. 2002; Siedenburg and Jendrossek 2011). So far, the available evidence suggests that this trait is rare among plants, but clearly it is recommended to routinely use multiple substrates to test for broad-spectrum substrate use capacity in functional characterization of terpene synthases.

3.3.3.4 Product Specificity of Terpenoid Synthases

Plants produce a huge variety of terpenoids with diverse composition among and within species. This high variety reflects expression of several terpene synthases (Sects. 3.3.3.2 and 3.4.2) as well as multiple products of single terpene synthases. For example, γ-humulene synthase from *Abies grandis* has been shown to produce 52 sesquiterpene olefins (Steele et al. 1998). However, all synthases have a certain specificity to form some main products with greater probability. Typically, the synthases also are stereospecific, forming preferably certain stereoisomers (Table 3.1, Croteau 1987; Prosser et al. 2004; Christianson 2006, 2008; Degenhardt et al. 2009). In fact, the capacity to produce multiple products is not universal, and some terpenoid synthases are almost completely specific, forming only one product or nearly so (Table 3.1). For example, isoprene synthases tend to form only isoprene, with the only known exception being a myrcene/isoprene synthase that can form

acyclic monoterpenes when GDP is the substrate (Sect. 3.3.3.3, Sharkey et al. 2013), but there are also several highly specific mono-, sesqui- and diterpene synthases (Table 3.1).

Predicting product profiles based on the terpenoid synthase full amino acid sequences is currently not possible as the correlative patterns are weak (Degenhardt et al. 2009). Even with the genes having high sequence homology, some enzymes might produce multiple products while others form one single product; highly homologous enzymes forming multiple products might produce strongly different product spectra (Degenhardt et al. 2009). Depending on the reaction mechanism (Markovnikov vs. anti-Markovnikov addition to a double bond), more or less stable carbocation can be formed in the initial reaction steps, thereby affecting the potential range of products formed (Christianson 2006). Also, active site structure, presence of certain fixed and protected dipoles that can stabilise the carbocation, and steric limitations play an important role in the product formation (Christianson 2006, 2008). Clearly, more X-ray crystal structures of tree terpene synthases are needed to gain further insight into the reaction mechanisms in the highly diverse terpenoid synthase family in trees.

3.4 Origin and Size of Terpenoid Synthase Gene Families

The vast array of different terpenoid compounds in nature results both from the low product specificity of many terpenoid synthases (Sect. 3.3.3.4) and from large number of terpenoid synthases present in most plant species studied. Presence of a vast number of terpenoid synthase genes can reflect divergent evolutionary paths for catalytic activities of terpenoid synthase genes leading to variations in terpenoid compounds formed (Bohlmann et al. 1998a; Aubourg et al. 2002; Dornelas and Mazzafera 2007), and resulting in novel functions that increase species resistance to insects, pathogens, and herbivores (Trapp and Croteau 2001) as well as to abiotic stresses (Owen and Peñuelas 2005; Fineschi et al. 2013; Possell and Loreto 2013), overall increasing species fitness. On the other hand, synthesis of similar compounds by phylogenetically widely distant plant species, e.g., among angiosperms and gymnosperms possessing terpenoid synthases with low homogeneity, also indicates convergent evolution of terpenoid synthases, and in some cases, multiple events of evolution and loss of capacity to form the given compound such as isoprene (Bohlmann et al. 1998a; Aubourg et al. 2002; Dornelas and Mazzafera 2007; Monson et al. 2013). Here we analyse the evidence of the origin of terpenoid synthases in plants and the size of terpenoid synthase gene families with emphasis on tree terpenoid synthases. Evolution of terpenoid synthases, including classification of plant terpenoid synthase gene families, is addressed in other chapters of the book (Fineschi et al. 2013; Li and Sharkey 2013b).

3.4.1 Origin of Plant Terpenoid Synthase Gene Families

It has been suggested that all modern plant terpenoid synthases originate from a common three-domain, α-β-γ-diterpene synthase in an ancient progenitor, followed by independent gene duplications and evolution of new genes in different organisms. As such an ancestor gene in plants, a bifunctional *ent*-copalyl diphosphate synthase/*ent*-kaurene synthase (PpCPS/KS) of the moss *Physcomitrella patens* has been suggested (Hayashi et al. 2006). This enzyme produces *ent*-kaurene which is a common precursor of gibberellins and endogenous diterpenes derived from this compound (Hayashi et al. 2010). The active site in the β-domain has class II terpene synthase activity and is responsible for *ent*-copalyl diphosphate synthesis from GGDP, while the active site in the α-domain forms *ent*-kaurene from *ent*-copalyl diphosphate. Additional bifunctional kaurene synthases have been discovered in the phylogenetically old species liverwort *Jungermannia subulata* (Kawaide et al. 2011) and spikemoss *Selaginella muellendorffii* (Li et al. 2012).

In phylogenetically younger gymnosperms and angiosperms, *ent*-kaurene is produced by two distinct mono-functional enzymes, *ent*-copalyl diphosphate synthase (CPS) and *ent*-kaurene synthase (KS) (Keeling et al. 2010; Zerbe et al. 2012a). Monofunctional diterpene synthases still have α-β-γ three-domain structure, but the active site in either α- or β-domain has lost the functional activity. As noted in Sect. 3.3.3.2, such three-domain structure is characteristic also to some sesquiterpene synthases, but most sesquiterpene, and all mono- and hemiterpene synthases have lost γ-domain sequence and have α-β domain configuration. Phylogenetic analyses suggest that the formation of α-β-domain terpenoid synthases via loss of the γ-domain has occurred several times during evolution (Hillwig et al. 2011).

An interesting question is what are the immediate ancestors for the α-domain and β-γ-domain in the common higher plant terpenoid α-β-γ ancestor? Analysis of the phylogenetic signals separately for α- and β-domains across a broad-spectrum of plant terpenoid synthases actually does not confirm the postulated origin of all terpenoid synthases from the moss *Physcomitrella patens* diterpene synthase gene (Fig. 3.5). If this were the ancestor of all multidomain plant terpenoid synthases, phylogenetic signals in α- and β-domains should reflect this, but the sequence homologies rather suggest that there must have been another common ancestor for moss, liverwort, spikemoss, gymnosperm and angiosperm terpene synthases (Fig. 3.5). Interestingly, *ent*-copalyl diphosphate and *ent*-kaurene synthases have been discovered in the bacterium *Bradyrhizobium japonicum* (Morrone et al. 2009). These terpenoid synthases share some homology with plant terpenoid synthases, and it has been suggested that there might be even an ancient progenitor for modern terpenoid synthases across the domains of life (Morrone et al. 2009).

Particularly interesting is the position of recently discovered spikemoss *Selaginella muellendorffii* α-domain only 'microbial'-type terpene synthases (Li et al. 2012). These sequences carry homology with both 'modern plant'
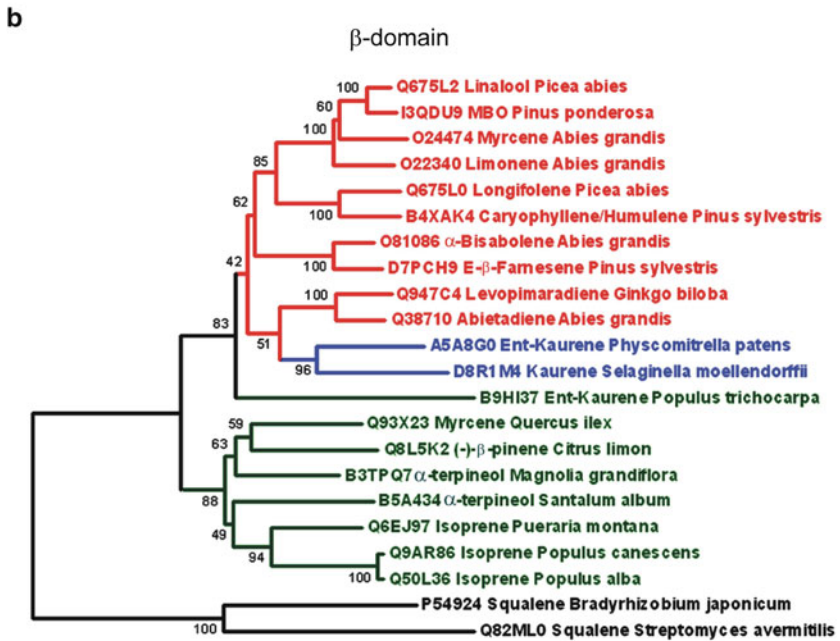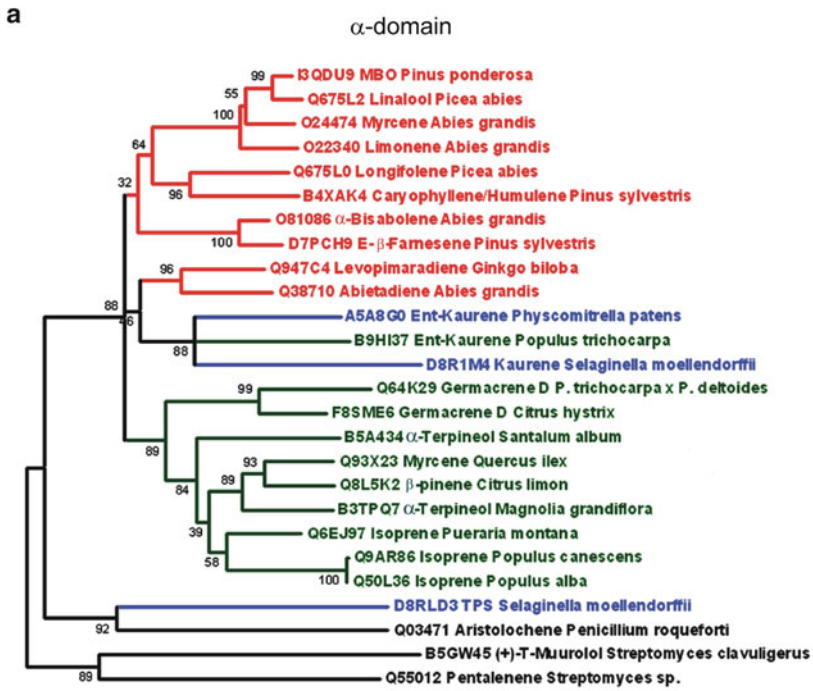
multi-domain terpenoid synthases and microbial α-domain only synthases, possibly providing the missing link to modern plant terpenoid origin. As more information of genomes becomes available, we will be possibly able to more clearly identify the origin of terpenoid synthases in plants.

3.4.2 Size of Terpenoid Synthase Gene Families

Given the possible monophyletic origin of 'higher plant type terpene synthases', continuous and extensive gene duplication followed by functional and structural specialization is responsible for the high diversity of terpenoid synthase gene families in plants (Fryxell 1996; Bohlmann and Keeling 2008; Chang and Duda 2012). The Viridiplantae have the largest terpenoid synthase families among living organisms (Fig. 3.6), but the size of terpenoid gene families varies strongly among plant species. According to genome mining by profile-based hidden Markov models, the size of terpenoid gene family varies from one in the moss *Physcomitrella patens* to 86 in the angiosperm tree *Eucalyptus grandis* (Fig. 3.6). These estimates based on similarity screens of full genomes broadly agree with independent estimates in other studies (Bohlmann and Keeling 2008; Chen et al. 2011). However, it is difficult to detect genes with low level of homogeneity. For instance, in *Selaginella muellendorffii*, 16–18 terpene synthases have been suggested based on genome screens for higher plant, α-β- or α-β-γ-domain, terpene synthases (Chen et al. 2011),

Fig. 3.5 Phylogenetic trees of selected tree terpene synthases (*red font* corresponding to gymnosperms and *green font* to angiosperms) and representative plant (*blue font*, moss *Physcomitrella patens* and spikemoss *Selaginella muellendorffii*) and bacterial and fungal outgroups based on α-domain (C-terminal part of the sequence, (**a**) and β-domain (N-terminal part of the sequence, (**b**). It has been suggested that all modern plant terpenoids originate from a fusion of α-domain and β-γ-domain terpenoid synthases in an ancient progenitor (Morrone et al. 2009; Cao et al. 2010). Thus, α-domain and β-domain might carry somewhat different phylogenetic signals (Sharkey et al. 2013). UniProtKB/Swiss-Prot entry codes are also given with species name (http://www.uniprot.org/). Isoprene and MBO (2-methyl-3-buten-2-ol) are hemiterpenes; linalool, limonene, myrcene, β-pinene and α-terpineol are monoterpenes; aristolochene, α-bisabolene, caryophyllene, β-farnesene, germacrene D, humulene, longifolene, T-muurolol, and pentalenene are sesquiterpenes; abietadiene, levopimaradiene, and *ent*-kaurene are diterpenes; and hopene and squalene are triterpenes. The microbial synthases in (**a**) only have α-domain as the *S. muellendorffii* 'microbial type' terpene synthase D8LD3 (Li et al. 2012), while the microbial and fungal squalene/hopene synthases in (**b**) have β-γ-domains only. The domains were separated using *Abies grandis* three-domain abietadiene synthase domain structure (Q38710, Zhou et al. 2012) as a seed. After alignment, signal peptides were deleted, and whenever pertinent, the sequence part corresponding to γ-domain was deleted. α-domain and β-domain parts of the sequence were separately aligned, truncated after alignment to a common length, and the trees were constructed by MEGA5 software using the maximum likelihood method (Tamura et al. 2011). The numbers next to the branches refer to the actual bootstrap values of branches and characterize the reliability of the branching (bootstrap consensus trees are demonstrated). The higher the score, the more reliable is the branching at that point (Tamura et al. 2011)
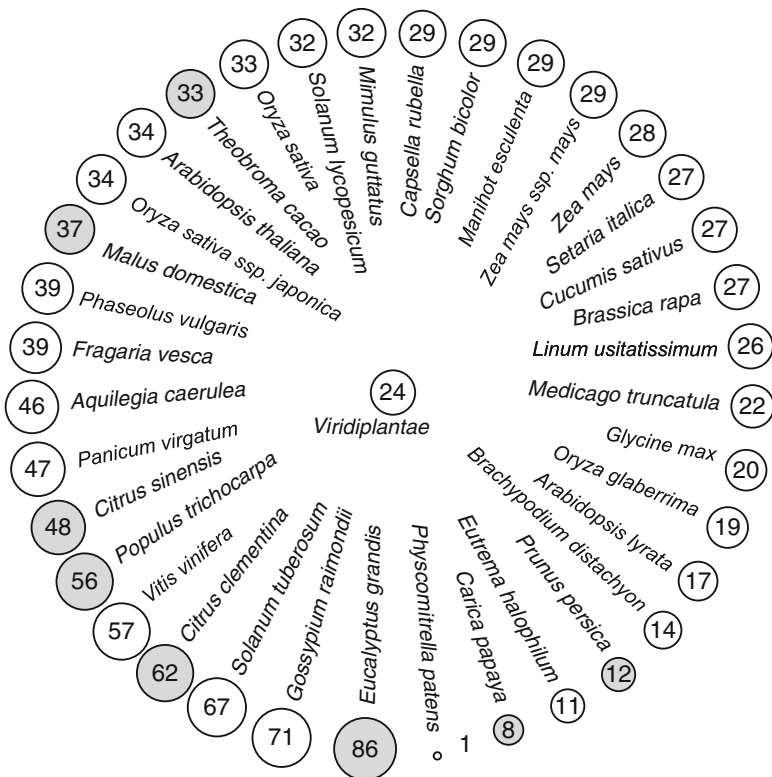
Fig. 3.6 Protein superfamily 'Terpenoid synthases' domains in Viridiplantae clade analysed using SUPERFAMILY 1.75 (http://supfam.cs.bris.ac.uk/SUPERFAMILY/) that includes essentially all protein domains from all available sequenced organisms, including 36 sequenced Viridiplantae genomes (Gough et al. 2001; Wilson et al. 2009). Superfamily is defined as a collection of domains (functional units of proteins) having structural and functional evidence of common evolutionary ancestor. Plant terpenoid synthases characteristically have either two (α and β) or three (α, β and γ) domains (Figs. 3.3 and 3.4, Hillwig et al. 2011; Köksal et al. 2011a, b). The homologous domains in SUPERFAMILY are identified by hidden Markov models, which is a profile-based method with high selectivity (Gough et al. 2001). Here the protein domains from the domain family of 'Terpenoid cyclase C-terminal domain' that contains the α-domain of the plant α-β- and α-β-γ-domain terpenoid synthases are demonstrated. In the figure, the radius of the circle reflects the size of the terpenoid synthase superfamily in logarithmic scale (numbers inside denote the number of terpenoid synthase domains in the given organism), the inner circle shows the average number of terpenoid synthase domains in Viridiplantae. Trees are denoted by *grey filling*

but recently 48 putative "bacterial" type, α-domain only synthases were identified in *Selaginella* genome, and functional activity of several of them was characterized (Li et al. 2012). Thus, we might have vastly underestimated the size of terpene gene families in plants.

The largest terpenoid synthase gene families among sequenced Viridiplantae are observed in trees (Fig. 3.6). Among top seven sequenced plants with the

greatest terpenoid synthase families, four, *Eucalyptus grandis*, *Citrus clementina*, *C. sinensis* and *Populus trichocarpa* are trees, and *Vitis vinifera* is a woody vine (Fig. 3.6, Martin et al. 2010). Furthermore, in addition to raw data generated by genome projects, functional characterization and analysis of terpenoid genes by transcriptome mining has highlighted large terpenoid gene families also in several gymnosperms (Martin and Bohlmann 2005; Chen et al. 2011; Keeling et al. 2011). However, some tree species such as *Carica papaya* have small terpenoid synthase gene families (Fig. 3.6). Currently, the functional implications of variations in the size of terpenoid synthase gene families have not yet been fully elucidated, especially in trees. However, it is reasonable to expect that presence of multiple terpenoid synthase genes forming different compounds and compound spectra and presence of paralogs with similar compound spectra, but potentially with differing regulatory promoter elements is associated with greater diversity of volatile product profiles and more diverse response patterns to biotic and abiotic stresses (Keeling et al. 2008; Bohlmann et al. 2011). Given that only minor modifications, at the level of single amino acid or a few amino acids, may be needed to change the terpenoid synthase function (Kampranis et al. 2007; Keeling et al. 2008; Hall et al. 2011), presence of duplicated genes can constitute an important gene pool for rapid adaptation to new biotic interactions, 'new' abiotic stresses in given plant habitat, and novel stress combinations or to changes in stress severities, thereby helping plants to adapt to their environment (Hall et al. 2011). In addition to variations among species, there is evidence of important within-species variation in the terpenoid gene family, in agreement with adaptive genomic modifications (Hall et al. 2011; Gonzales-Vigil et al. 2012).

3.5 Regulation of Terpenoid Synthesis

The quantities of metabolites ultimately produced are influenced by genetic features and environmental signals. So, the metabolite contents in plants are dynamic and controlled by internal and external stimuli. These changes are regulated by gene expression networks involved in targeting pathways to specific cells and organs and determining the overall expression level of the pathway (Grotewold 2008). Studying one or some genes is not enough for having a deep insight into such networks across a variety of tissues and treatments. Transcriptome and metabolome data provided by next generation technologies have opened new opportunities for understanding the biosynthesis and regulation of plant metabolites (Crispin and Wurtele 2013). Perhaps with this comprehensive data, we will be able to start unresolving the complex regulatory networks, not only focusing on the synthesis of terpenoid end-products, but addressing the entire cascade of events altering profoundly plant 'normal' metabolism, from stress signals to elicitation of pathways.

Substrate-level regulation of terpenoid synthesis has been addressed in the chapters of Li and Sharkey (2013b) and Monson (2013). Here we analyse the regulation of terpenoid synthesis driven by variations in gene expression. Terpenoid

synthase genes show a high variety of temporal and spatial expression patterns; some of these compounds are constitutively expressed (Steele et al. 1998), some of them induced by both biotic and abiotic stresses (Yin et al. 1997; Byun-McKay et al. 2006; Tholl 2006; Fares et al. 2008; Chen et al. 2011), some are expressed in leaf mesophyll and leaf surface trichomes (Köllner et al. 2004; van Schie et al. 2007; Falara et al. 2011), and others in flowers (Chen et al. 2003; Dudareva et al. 2003; Falara et al. 2011). This level of control has been termed as 'genetic control' (Monson 2013).

3.5.1 'Constitutively' Expressed Synthases

Regulation of terpenoid synthase gene expression is controlled by the regulatory elements in the gene promoter region that interact with a variety of transcription factors. Rosenkranz and Schnitzler (2013) provide an in-depth overview of the poplar isoprene synthase promoter region (*PcIspS*). Isoprene synthase in emitting species is a constitutively expressed gene. However, *PcIspS* contains circadian, heat and stress-dependent regulatory elements (Loivamäki et al. 2007b; Cinege et al. 2009) and thus, expression of isoprene synthase varies during the day, during the season and is responsive to changes in temperature (Mayrhofer et al. 2005; Sasaki et al. 2005; Wiberley et al. 2008, 2009). So far, it is however, unclear how isoprene synthase activity is regulated in plants grown under different atmospheric CO_2 concentrations (Wilkinson et al. 2009; Sun et al. 2012).

There is much less information available on regulation of expression of terpene synthases. In particular, there are several constitutively expressed terpene synthase genes analogous to isoprene synthase, such as myrcene synthase in broad-leaved evergreen Mediterranean oaks (Fischbach et al. 2001), which seem to be regulated similarly to isoprene synthase, i.e., responding to ambient temperature and light conditions (Fischbach et al. 2002; Staudt et al. 2003), and might be responsive to CO_2 as well (Loreto et al. 2001a), but so far the promoter regions of these enzymes have not been characterized.

In most conifers, there are also several constitutively expressed enzymes for oleoresin formation. Oleoresin in these species is accumulated in specific resin ducts or galls in the bark, sapwood, and needles, and serves as a major physico-chemical defence barrier against insects and pathogens (Bohlmann and Croteau 1999; Trapp and Croteau 2001; Martin and Bohlmann 2005; Raffa et al. 2005). Release of oleoresin after insect attack repels and deters insects; emissions of mono- and sesquiterpenes provide indirect defence against herbivores, but diterpenes and terpene oxidation products provide direct defence by forming a physical barrier at the place of insect attack (Martin et al. 2003; Miller et al. 2005; Raffa et al. 2005). The promoter regions of constitutively expressed terpenoid synthases in conifers have not yet been characterized. In herbs, there has been some progress in identifying promoters targeting the gene expression to glandular trichomes (Tissier 2012), but again, terpenoid-specific promoters have not yet been identified.

Information about terpenoid-specific transcription factors is currently also very limited (Tholl 2006; Chen et al. 2011), and obviously understanding terpenoid synthase regulation should constitute a priority for future studies.

3.5.2 Stress-Induced Terpene Synthases

Emissions of a large number of mono- and sesquiterpenes are elicited in response to numerous biotic stresses such as herbivory, oviposition and fungal inoculation, but also to abiotic stresses through activation of signalling responses triggered by reactive oxygen species. Stress-induced production of traumatic resin plays a prominent role as a physical and toxic direct defence to insects and pathogens (Martin et al. 2002; Miller et al. 2005; Raffa et al. 2005), while volatile terpenoid emissions can serve as repellents to herbivores and attractants to enemies of herbivores (Thaler 2002; Hilker et al. 2005; Miller et al. 2005; Raffa et al. 2005; Dicke et al. 2009). There is a plethora of examples of stress-elicited modifications in terpenoid synthesis. For detailed consideration of these responses, we refer to the chapters of Trowbridge and Stoy (2013) and Holopainen et al. (2013) and the review of Keeling and Bohlmann (2006b) and only briefly mention a few of these elicited responses. Feeding by the white pine weevil (*Pissodes strobi*) on Sitka spruce (*P. sitchensis*) induced traumatic resin accumulation in stems and also induced expression of several terpene synthase transcripts, such as the (−)-pinene synthase (Byun McKay et al. 2003; Miller et al. 2005). In addition, (−)-linalool synthase activity increased after weevil feeding, resulting in linalool emissions (Miller et al. 2005). Analogously, monoterpene synthase activities in needles of ponderosa pine (*Pinus ponderosa*), lodgepole pine (*Pinus contorta*), and white fir (*Abies concolor*) were enhanced after feeding by tiger moth (*Halisdota ingens*) larvae (Litvak and Monson 1998). Treatments with methyl jasmonate (MeJA), elicitor causing responses similar to herbivore attacks, have been widely used to study influence of biotic stress effects on terpenoid synthesis (Keeling and Bohlmann 2006b; Bohlmann 2008). In several conifers, MeJA treatment has shown to result in formation of traumatic resin ducts, terpenoid accumulation, and induction of prenyltransferase and terpene synthase activities, and altered volatile terpenoid emission profiles (e.g., Martin et al. 2002, 2003; Miller et al. 2005).

There are different patterns of timing of elicitation of various terpenoid synthases. In response to wounding in the gymnosperm tree *Abies grandis*, monoterpene synthase genes were elicited first, followed by sesqui- and diterpene synthases (Steele et al. 1998). In the angiosperm tree *Alnus glutinosa*, feeding by common white wave (*Cabera pusaria*) larvae resulted in simultaneous elicitation of emissions of mono-, sesqui and homoterpenes, but sesquiterpene emissions were characterized by biphasic emission kinetics and the level of elicitation differed for different volatile terpenes (Copolovici et al. 2011). Such variations in temporal patterns can be simulated based on the theory of recursive action of regulators on the target gene(s) over time (Vu and Vohradsky 2007) as demonstrated in the

chapter of Grote et al. (2013). In addition, there are also genotypic differences in the degree of elicitation of given synthases (Byun-McKay et al. 2006), suggesting that modifications in gene regulation patterns can constitute a further important adaptive mechanism.

Overall, the phenomenological evidence consistently indicates the biotic stress-dependent enhancement of terpenoid synthesis pathway and modifications of terpenoid profiles. Although there has been a significant progress in biochemical and molecular characterization of the induced terpenoid responses (Huber et al. 2004; Martin and Bohlmann 2005; Keeling and Bohlmann 2006b; Tholl 2006; Chen et al. 2011), regulation of stress-induced terpene synthesis is still poorly known. In white spruce (*Picea glauca*), putative *cis*-acting elements such as MeJA- and wound-response elements, and promoter-enhancing sequences have been identified for monoterpene 3-carene synthase by BAC cloning (Hamberger et al. 2009), but the inducible terpenoid promoters have not yet been functionally tested. Potato (*Solanum tuberosum*) wound- and insect-inducible promoter of proteinase inhibitor protein was used to demonstrate local elicitation of terpenoid synthase activity in transgenic *Picea glauca* after mechanical wounding (Godard et al. 2007). This transgenic system provides encouraging platform for proof-of-concept studies, and possibly will also help to elucidate the regulation of inducible terpenoid synthases in native systems.

3.6   Conclusions

More than 60,000 terpenoid compounds have been described in living organisms. This huge chemical diversity results from a large number of terpenoid synthases and relatively low product specificity of many terpenoid synthases. High genetic richness of terpenoid synthases present in many plant species constitutes an important gene pool for adaptation. Especially, given that the product profiles of several terpenoid synthases can be altered with only minor modifications in active center structure, gene duplication of ancient progenitor terpenoid synthases has catalyzed the vast evolutionary adjustment in terpenoid profiles (Sect. 3.4.2). On the other hand, millions years of evolution have led to low sequence similarity among terpenoid synthases in distant organisms. However, sequence similarity is not necessarily associated with protein function, as the classic class I and class II terpene synthase tertiary structures are remarkably similar even across the domains of life. In fact, widely divergent terpene synthases at sequence level can make the same products and have similar product profiles. Thus, there is evidence of convergent evolution in terpenoid synthases among many plant groups.

So far, terpenoid synthase gene family structure is available only for a few tree species, with particularly limited information for key forest species, albeit highly detailed information has been gained by next generation transcriptome sequencing techniques and several full genome sequences for important tree species have become available (Sect. 3.4.2). As more genomic data become available, we will be

able to gain more conclusive insight into the variations in terpenoid synthase gene family size and diversity. Even among the sequenced organisms, recent evidence of polyphyletic origin of terpenoid synthases has been found, suggesting that genetic diversity of plant terpenoid synthases may be even larger than we have thought before (Li et al. 2012).

There is rich phenomenological evidence of changes in constitutively expressed synthase activities and induction of non-constitutively expressed terpenoid synthases by different abiotic and biotic stresses for many tree species (Sect. 3.5). However, the information on regulatory elements of terpenoid synthases is currently very limited, and clearly more work is needed to gain insight into regulation of the expression of terpenoid synthases as driven by environmental variability and stress.

References

Aaron JA, Christianson DW (2010) Trinuclear metal clusters in catalysis by terpenoid synthases. Pure Appl Chem 82:1585–1597

Aharoni A, Giri AP, Verstappen FWA, Bertea CM, Sevenier R, Sun Z, Jongsma MA, Schwab W, Bouwmeester HJ (2004) Gain and loss of fruit flavor compounds produced by wild and cultivated strawberry species. Plant Cell 16:3110–3131

Aharoni A, Jongsma MA, Bouwmeester HJ (2005) Volatile science? Metabolic engineering of terpenoids in plants. Trends Plant Sci 10:594–602

Al-Dous EK, George B, Al-Mahmoud ME, Al-Jaber MY, Wang H, Salameh YM, Al-Azwani EK, Chaluvadi S, Pontaroli AC, DeBarry J, Arondel V, Ohlrogge J, Saie IJ, Suliman-Elmeer KM, Bennetzen JL, Kruegger RR, Malek JA (2011) De novo genome sequencing and comparative genomics of date palm (*Phoenix dactylifera*). Nat Biotechnol 29:521–527

Allona I, Quinn M, Shoop E, Swope K, St Cyr S, Carlis J, Riedl J, Retzel E, Campbell MM, Sederoff R, Whetten RW (1998) Analysis of xylem formation in pine by cDNA sequencing. Proc Natl Acad Sci USA 95:9693–9698

Alonso WR, Croteau R (1993) Prenyltransferases and cyclases. In: Lea PJ (ed) Enzymes of secondary metabolism, vol 9, Methods in plant biochemistry. Academic, London/San Diego/New York/Boston/Sydney/Tokyo/Toronto, pp 239–260

Andreeva A, Howorth D, Chandonia J-M, Brenner SE, Hubbard TJP, Chothia C, Murzin AG (2008) Data growth and its impact on the SCOP database: new developments. Nucleic Acids Res 36:D419–D425

Arimura G-i, Ozawa R, Shimoda T, Nishioka T, Boland W, Takabayashi J (2000) Herbivory-induced volatiles elicit defence genes in lima bean leaves. Nature 406:512–515

Arimura G-i, Huber DPW, Bohlmann J (2004) Forest tent caterpillars (*Malacosoma disstria*) induce local and systemic diurnal emissions of terpenoid volatiles in hybrid poplar (*Populus trichocarpa* x *deltoides*): cDNA cloning, functional characterization, and patterns of gene expression of (−)-germacrene D synthase, PtdTPS1. Plant J 37:603–616

Ashworth K, Boissard C, Folberth G, Lathière J, Schurgers G (2013) Global modeling of volatile organic compound emissions. In: Niinemets Ü, Monson RK (eds) Biology, controls and models of tree volatile organic compound emissions, vol 5, Tree physiology. Springer, Berlin, pp 451–487

Aubourg S, Lecharny A, Bohlmann J (2002) Genomic analysis of the terpenoid synthase (AtTPS) gene family of *Arabidopsis thaliana*. Mol Genet Genomics 267:730–745

Back KW, Chappell J (1996) Identifying functional domains within terpene cyclases using a domain-swapping strategy. Proc Natl Acad Sci USA 93:6841–6845

Behnke K, Ehlting B, Teuber M, Bauerfeind M, Louis S, Hänsch R, Polle A, Bohlmann J, Schnitzler J-P (2007) Transgenic, non-isoprene emitting poplars don't like it hot. Plant J 51:485–499

Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The protein data bank. Nucleic Acids Res 28:235–242

Besumbes Ó, Sauret-Güeto S, Phillips MA, Imperial S, Rodríguez-Concepción M, Boronat A (2004) Metabolic engineering of isoprenoid biosynthesis in *Arabidopsis* for the production of taxadiene, the first committed precursor of taxol. Biotechnol Bioeng 88:168–175

Bohlmann J (2008) Insect-induced terpenoid defenses in spruce. In: Schaller A (ed) Induced plant resistance to herbivory. Springer, Berlin, pp 173–187

Bohlmann J, Croteau R (1999) Diversity and variability of terpenoid defences in conifers: molecular genetics, biochemistry and evolution of the terpene synthase gene family in grand fir (*Abies grandis*). Novartis Found Symp 223:132–149

Bohlmann J, Keeling CI (2008) Terpenoid biomaterials. Plant J 54:656–669

Bohlmann J, Steele CL, Croteau R (1997) Monoterpene synthases from grand fir (*Abies grandis*). cDNA isolation, characterization and functional expression of myrcene synthase, (−)-4S-limonene synthase and (−)-(1S,5S)-pinene synthase. J Biol Chem 272:21784–21792

Bohlmann J, Crock J, Jetter R, Croteau R (1998a) Terpenoid-based defenses in conifers: cDNA cloning, characterization, and functional expression of wound-inducible (E)-α-bisabolene synthase from grand fir (*Abies grandis*). Proc Natl Acad Sci USA 95:6756–6761

Bohlmann J, Meyer-Gauen G, Croteau R (1998b) Plant terpenoid synthases: molecular biology and phylogenetic analysis. Proc Natl Acad Sci USA 95:4126–4133

Bohlmann J, Phillips M, Ramachandiran V, Katoh S, Croteau R (1999) cDNA cloning, characterization, and functional expression of four new monoterpene synthase members of the *Tpsd* gene family from grand fir (*Abies grandis*). Arch Biochem Biophys 368:232–243

Bohlmann J, Hall D, Robert J, Keeling C, Domanski D, Quesada AL, Jancsik S, Kuzyk M, Hamberger B, Borchers C (2011) Biomarkers and gene copy number variation for terpenoid traits associated with insect resistance in Sitka spruce: an integrated genomic, proteomic, and biochemical analysis of (+)-3-carene biosynthesis. BMC Proc 5:027

Braun EL, Dias AP, Matulnik TJ, Grotewold E (2001) Transcription factors and metabolic engineering: novel applications for ancient tools. In: Romeo JT, Saunders JA, Matthews BF (eds) Regulation of phytochemicals by molecular techniques. Elsevier Science, Beltsville, pp 79–109

Brenner S, Johnson M, Bridgham J, Golda G, Lloyd DH, Johnson D, Luo S, McCurdy S, Foy M, Ewan M, Roth R, George D, Eletr S, Albrecht G, Vermaas E, Williams SR, Moon K, Burcham T, Pallas M, DuBridge RB, Kirchner J, Fearon K, J-i M, Corcoran K (2000) Gene expression analysis by massively parrellel signature sequencing (MPSS) on microbead arrays. Nat Biotechnol 18:630–634

Byun McKay A, Hunter WL, Godard KA, Wang SX, Martin DM, Bohlmann J, Plant AL (2003) Insect attack and wounding induce traumatic resin duct development and gene expression of (−)-pinene synthase in Sitka spruce. Plant Physiol 133:368–378

Byun-McKay A, Godard KA, Toudefallah M, Martin DM, Alfaro R, King J, Bohlmann J, Plant AL (2006) Wound-induced terpene synthase gene expression in Sitka spruce that exhibit resistance or susceptibility to attack by the white pine weevil. Plant Physiol 140:1009–1021

Cane DE, Ishida H (2012) Exploration and mining of the bacterial terpenome. Acc Chem Res 45:463–472

Cane DE, Sohng J-K, Lamberson CR, Rudnicki SM, Wu Z, Lloyd MD, Oliver JS, Hubbard BR (1994) Pentalenene synthase. Purification, molecular cloning, sequencing, and high-level expression in *Escherichia coli* of a terpenoid cyclase from *Streptomyces* UC531g. Biochemistry 33:5846–5857

Cao R, Zhang Y, Mann FM, Huang C, Mukkamala D, Hudock MP, Mead ME, Prisic S, Wang K, Lin F-Y, Chang T-K, Peters RJ, Odfield E (2010) Diterpene cyclases and the nature of the isoprene fold. Proteins Struct Funct Bioinform 78:2417–2432

Caruthers JM, Kang I, Rynkiewicz MJ, Cane DE, Christianson DW (2000) Crystal structure determination of aristolochene synthase from the blue cheese mold, *Penicillium roqueforti*. J Biol Chem 275:25533–25539

Chang D, Duda TF Jr (2012) Extensive and continuous duplication facilitates rapid evolution and diversification of gene families. Mol Biol Evol 29:2019–2029

Chappell J (1995) Biochemistry and molecular biology of the isoprenoid biosynthetic pathways in plants. Annu Rev Plant Physiol Plant Mol Biol 46:521–547

Chen F, Tholl D, D'Auria JC, Farooq A, Pichersky E, Gershenzon J (2003) Biosynthesis and emission of terpenoid volatiles from *Arabidopsis* flowers. Plant Cell 15:1–14

Chen F, Tholl D, Bohlmann J, Pichersky E (2011) The family of terpene synthases in plants: a mid-size family of genes for specialized metabolism that is highly diversified throughout the kingdom. Plant J 66:212–229

Choudhary DK, Johri BN, Prakash A (2008) Volatiles as priming agents that initiate plant growth and defence responses. Curr Sci 94:595–604

Christianson DW (2006) Structural biology and chemistry of the terpenoid cyclases. Chem Rev 106:3412–3442

Christianson DW (2008) Unearthing the roots of the terpenome. Curr Opin Chem Biol 12:141–150

Cinege G, Louis S, Hänsch R, Schnitzler J-P (2009) Regulation of isoprene synthase promoter by environmental and internal factors. Plant Mol Biol 69:593–604

Copolovici L, Kännaste A, Remmel T, Vislap V, Niinemets Ü (2011) Volatile emissions from *Alnus glutinosa* induced by herbivory are quantitatively related to the extent of damage. J Chem Ecol 37:18–28

Cori O, Rojas MC (1985) Carbocyclases from *Citrus limonum*. In: Law JH, Rilling HC, Kaplan NO (eds) Steroids and isoprenoids part A, vol 110, Methods in enzymology. Academic, Orlando/San Diego/New York/London/Toronto/Montreal/Sydney/Tokyo, pp 406–417

Crispin MC, Wurtele ES (2013) Use of metabolomics and transcriptomics to gain insights into the regulation and biosynthesis of medicinal compounds: *Hypericum* as a model. In: Chandra S, Lata H, Varma A (eds) Biotechnology for medicinal plants. Springer, Berlin, pp 395–411

Croteau R (1987) Biosynthesis and catabolism of monoterpenoids. Chem Rev 87:929–954

Croteau R, Karp F (1977) Biosynthesis of monoterpenes: partial purification and characterization of 1,8-cineole synthetase from *Salvia officinalis*. Arch Biochem Biophys 179:257–265

Croteau R, Hooper CL, Felton M (1978) Biosynthesis of monoterpenes. Partial purification and characterization of a bicyclic monoterpenol dehydrogenase from sage (*Salvia officinalis*). Arch Biochem Biophys 188:182–193

Crozier A, Clifford MN, Ashihara H (eds) (2006) Plant secondary metabolites, occurrence, structure and role in the human diet. Blackwell Publishing Ltd, Oxford

Degenhardt J, Gershenzon J, Naldwin IT, Kessler A (2003) Attracting friends and feast on foes: engineering terpene emission to make crop plants more attractive to herbivore enemies. Curr Opin Biotechnol 14:169–176

Degenhardt J, Koellner TG, Gershenzon J (2009) Monoterpene and sesquiterpene synthases and the origin of terpene skeletal diversity in plants. Phytochemistry 70:1621–1637

Dehal SS, Croteau R (1988) Partial purification and characterization of two sesquiterpene cyclases from sage (*Salvia officinalis*) which catalyze the respective conversion of farnesyl pyrophosphate to humulene and caryophyllene. Arch Biochem Biophys 261:346–356

Dewick PM (2002) The biosynthesis of C_5-C_{25} terpenoid compounds. Nat Prod Rep 19:181–222

Dicke M, Bruin J (2001) Chemical information transfer between plants: back to the future. Biochem Syst Ecol 29:981–994

Dicke M, van Loon JJA, Soler R (2009) Chemical complexity of volatiles from plants induced by multiple attack. Nat Chem Biol 5:317–324

Dornelas MC, Mazzafera P (2007) A genomic approach to characterization of the *Citrus* terpene synthase gene family. Genet Mol Biol 30:832–840

Dudareva N, Pichersky E (2008) Metabolic engineering of plant volatiles. Curr Opin Biotechnol 19:181–189

Dudareva N, Martin D, Kish CM, Kolosova N, Gorenstein N, Faldt J, Miller B, Bohlmann J (2003) (E)-β-ocimene and myrcene synthase genes of floral scent biosynthesis in snapdragon: function and expression of three terpene synthase genes of a new terpene synthase subfamily. Plant Cell 15:1227–1241

Dudareva N, Pichersky E, Gershenzon J (2004) Biochemistry of plant volatiles. Plant Physiol 135:1893–1902

Dudareva N, Negre F, Nagegowda DA, Orlova I (2006) Plant volatiles: recent advances and future perspectives. Crit Rev Plant Sci 25:417–440

Eisenreich W, Bacher A, Arigoni D, Rohdich F (2004) Biosynthesis of isoprenoids via the non-mevalonate pathway. Cell Mol Life Sci 61:1401–1426

El Tamer MK, Smeets M, Holthuysen N, Lücker J, Tang A, Roozen J, Bouwmeester HJ, Voragen AGJ (2003) The influence of monoterpene synthase transformation on the odour of tobacco. J Biotechnol 106:15–21

Engelhart GJ, Asa-Awuku A, Nenes A, Pandis SN (2008) CCN activity and droplet growth kinetics of fresh and aged monoterpene secondary organic aerosol. Atmos Chem Phys 8:3937–3949

Falara V, Akhtar TA, Nguyen TTH, Spyropoulou EA, Bleeker PM, Schauvinhold I, Matsuba Y, Bonini ME, Schilmiller AL, Last RL, Schuurink RC, Pichersky E (2011) The tomato terpene synthase gene family. Plant Physiol 157:770–789

Fares S, Loreto F, Kleist E, Wildt J (2008) Stomatal uptake and stomatal deposition of ozone in isoprene and monoterpene emitting plants. Plant Biol 10:44–54

Fineschi S, Loreto F, Staudt M, Peñuelas J (2013) Diversification of volatile isoprenoid emissions from trees: evolutionary and ecological perspectives. In: Niinemets Ü, Monson RK (eds) Biology, controls and models of tree volatile organic compound emissions, vol 5, Tree physiology. Springer, Berlin, pp 1–20

Fischbach RJ, Zimmer I, Steinbrecher R, Pfichner A, Schnitzler J-P (2000) Monoterpene synthase activities in leaves of *Picea abies* (L.) Karst. and *Quercus ilex* L. Phytochemistry 54:257–265

Fischbach RJ, Zimmer W, Schnitzler J-P (2001) Isolation and functional analysis of a cDNA encoding a myrcene synthase from holm oak (*Quercus ilex* L.). Eur J Biochem 268:5633–5638

Fischbach RJ, Staudt M, Zimmer I, Rambal S, Schnitzler J-P (2002) Seasonal pattern of monoterpene synthase activities in leaves of the evergreen tree *Quercus ilex*. Physiol Plant 114:354–360

Fortunati A, Barta C, Brilli F, Centritto M, Zimmer I, Schnitzler J-P, Loreto F (2008) Isoprene emission is not temperature-dependent during and after severe drought-stress: a physiological and biochemical analysis. Plant J 55:687–697

Fryxell KJ (1996) The coevolution of gene family trees. Trends Genet 12:364–369

Gennadios HA, Gonzalez V, Di Costanzo L, Li A, Yu F, Miller DJ, Allemann RK, Christianson DW (2009) Crystal structure of (+)-δ-cadinene synthase from *Gossypium arboreum* and evolutionary divergence of metal binding motifs for catalysis. Biochemistry 48:6175–6183

Gershenzon J, Croteau RB (1993) Terpenoid biosynthesis: the basic pathway and formation of monoterpenes, sesquiterpenes, and diterpenes. In: Moore TS (ed) Lipid metabolism in plants. CRC Press, Boca Raton/Ann Arbor/London/Tokyo, pp 339–388

Godard K-A, Byun-McKay A, Levasseur C, Plant A, Séguin A, Bohlmann J (2007) Testing of a heterologous, wound- and insect-inducible promoter for functional genomics studies in conifer defense. Plant Cell Rep 26:2083–2090

Gonzales-Vigil E, Hufnagel DE, Kim J, Last RL, Barry CS (2012) Evolution of TPS20-related terpene synthases influences chemical diversity in the glandular trichomes of the wild tomato relative *Solanum habrochaites*. Plant J 71:921–935

Gough J, Karplus K, Hughey R, Chothia C (2001) Assignment of homology to genome sequences using a library of hidden Markov models that represent all proteins of known structure. J Mol Biol 313:903–919

Gowan E, Lewis BA, Turgeon R (1995) Phloem transport of antirrhinoside, an iridoid glycoside, in *Asarina scandens* (Scrophulariaceae). J Chem Ecol 21:1781–1788

Gray DW, Breneman SR, Topper LA, Sharkey TD (2011) Biochemical characterization and homology modeling of methylbutenol synthase and implications for understanding hemiterpene synthase evolution in plants. J Biol Chem 286:20582–20590

Green S, Friel EN, Matich A, Beuning LL, Cooney JM, Rowan DD, MacRae E (2007) Unusual features of a recombinant apple α-farnesene synthase. Phytochemistry 68:176–188

Grote R, Monson RK, Niinemets Ü (2013) Leaf-level models of constitutive and stress-driven volatile organic compound emissions. In: Niinemets Ü, Monson RK (eds) Biology, controls and models of tree volatile organic compound emissions, vol 5, Tree physiology. Springer, Berlin, pp 315–355

Grotewold E (2008) Transcription factors for predictive plant metabolic engineering: are we there yet? Curr Opin Biotechnol 19:138–144

Guenther A, Karl T, Harley P, Wiedinmyer C, Palmer PI, Geron C (2006) Estimates of global terrestrial isoprene emissions using MEGAN (model of emissions of gases and aerosols from nature). Atmos Chem Phys 6:3181–3210

Hall DE, Robert JA, Keeling CI, Domanski D, Quesada AL, Jancsik S, Kuzyk MA, Hamberger B, Borchers CH, Bohlmann J (2011) An integrated genomic, proteomic and biochemical analysis of (+)-3-carene biosynthesis in Sitka spruce (*Picea sitchensis*) genotypes that are resistant or susceptible to white pine weevil. Plant J 65:936–948

Hamano Y, Kuzuyama T, Itoh N, Furihata K, Seto H, Dairi T (2002) Functional analysis of eubacterial diterpene cyclases responsible for biosynthesis of a diterpene antibiotic, terpentecin. J Biol Chem 277:37098–37104

Hamberger B, Hall D, Yuen M, Oddy C, Hamberger B, Keeling CI, Ritland C, Ritland K, 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 Biol 9:106

Harrison SP, Morfopoulos C, Dani KGS, Prentice IC, Arneth A, Atwell BJ, Barkley MP, Leishman MR, Loreto F, Medlyn BE, Niinemets Ü, Possell M, Peñuelas J, Wright IJ (2013) Volatile isoprenoid emissions from plastid to planet. New Phytol 197:49–57

Haudenschild C, Schalk M, Karp F, Croteau R (2000) Functional expression of regiospecific cytochrome P450 limonene hydroxylases from mint (*Mentha* spp.) in *Escherichia coli* and *Saccharomyces cerevisiae*. Arch Biochem Biophys 379:127–136

Hayashi K-i, Kawaide H, Notomi M, Sakigi Y, Matsuo A, Nozaki H (2006) Identification and functional analysis of bifunctional ent-kaurene synthase from the moss *Physcomitrella patens*. FEBS Lett 580:6175–6181

Hayashi K-i, Horie K, Hiwatashi Y, Kawaide H, Yamaguchi S, Hanada A, Nakashima T, Nakajima M, Mander LN, Yamane H, Hasebe M, Nozaki H (2010) Endogenous diterpenes derived from ent-kaurene, a common gibberellin precursor, regulate protonema differentiation of the moss *Physcomitrella patens*. Plant Physiol 153:1085–1097

Heijari J, Nerg A-M, Kainulainen P, Vuorinen M, Holopainen JK (2008) Long-term effects of exogenous methyl jasmonate application on Scots pine (*Pinus sylvestris*) needle chemical defence and diprionid sawfly performance. Entomol Exp Appl 128:162–171

Hemmerlin A, Hoeffler JF, Meyer O, Tritsch D, Kagan IA, Grosdemange-Billiard C, Rohmer M, Bach TJ (2003) Cross-talk between the cytosolic mevalonate and the plastidial methylerythritol phosphate pathways in tobacco bright yellow-2 cells. J Biol Chem 278:26666–26676

Hezari M, Lewis NG, Croteau R (1995) Purification and characterization of taxa-4(5),11(12)-diene synthase from Pacific yew (*Taxus brevifolia*) that catalyzes the first committed step of taxol biosynthesis. Arch Biochem Biophys 322:437–444

Hilker M, Stein C, Schroder R, Varama M, Mumm R (2005) Insect egg deposition induces defence responses in *Pinus sylvestris*: characterisation of the elicitor. J Exp Biol 208:1849–1854

Hillwig ML, Xu M, Toyomasu T, Tiernan MS, Wei G, Cui G, Huang L, Peters RJ (2011) Domain loss has independently occurred multiple times in plant terpene synthase evolution. Plant J 68:1051–1060

Hohn T (1999) Cloning and expression of terpene synthase genes. In: Cane DE (ed) Comprehensive natural products chemistry: isoprenoids including steroids and carotenoids, vol 2. Pergamon Press, Oxford, pp 155–200

Holliday JA, Ralph SG, White R, Bohlmann J, Aitken SN (2008) Global monitoring of autumn gene expression within and among phenotypically divergent populations of Sitka spruce (*Picea sitchensis*). New Phytol 178:103–122

Holopainen JK (2004) Multiple functions of inducible plant volatiles. Trends Plant Sci 9:529–533

Holopainen JK, Nerg A-M, Blande JD (2013) Multitrophic signalling in polluted atmospheres. In: Niinemets Ü, Monson RK (eds) Biology, controls and models of tree volatile organic compound emissions, vol 5, Tree physiology. Springer, Berlin, pp 285–314

Huang SH, Duke RK, Chebib M, Sasaki K, Wada K, Johnston GAR (2004) Ginkgolides, diterpene trilactones of *Ginkgo biloba*, as antagonists at recombinant $\alpha 1\gamma 2\gamma 2L$ GABA$_A$ receptors. Eur J Pharmacol 494:131–138

Huber DPW, Ralph S, Bohlmann J (2004) Genomic hardwiring and phenotypic plasticity of terpenoid-based defenses in conifers. J Chem Ecol 30:2399–2418

Huff Hartz KE, Rosenørn T, Ferchak SR, Raymond TM, Bilde M, Donahue NM, Pandis SN (2005) Cloud condensation nuclei activation of monoterpene and sesquiterpene secondary organic aerosol. J Geophys Res Atmos 110:D14208. doi:14210.11029/12004JD005754

Hyatt DC, Youn B, Zhao Y, Santhamma B, Coates RM, Croteau RB, Kang C (2007) Structure of limonene synthase, a simple model for terpenoid cyclase catalysis. Proc Natl Acad Sci USA 104:5360–5365

Jin Q, Williams DC, Hezari M, Croteau R, Coates RM (2005) Stereochemistry of the macrocyclization and elimination steps in taxadiene biosynthesis through deuterium labeling. J Org Chem 70:4667–4675

Jones CG, Keeling CI, Ghisalberti EL, Barbour EL, Plummer JA, Bohlmann J (2008) Isolation of cDNAs and functional characterisation of two multi-product terpene synthase enzymes from sandalwood, *Santalum album* L. Arch Biochem Biophys 477:121–130

Jones CG, Moniodis J, Zulak KG, Scaffidi A, Plummer JA, Ghisalberti EL, Barbour EL, Bohlmann J (2011) Sandalwood fragrance biosynthesis involves sesquiterpene synthases of both the terpene synthase (Tps)-a and Tps-b subfamilies, including santalene synthases. J Biol Chem 286:17445–17454

Kampranis SC, Ioannidis D, Purvis A, Mahrez W, Ninga E, Katerelos NA, Anssour S, Dunwell JM, Degenhardt J, Makris AM, Goodenough PW, Johnson CB (2007) Rational conversion of substrate and product specificity in a *Salvia* monoterpene synthase: structural insights into the evolution of terpene synthase function. Plant Cell 19:1994–2005

Kane JF (1995) Effects of rare codon clusters on high-level expression of heterologous proteins in *Escherichia coli*. Curr Opin Biotechnol 6:494–500

Kawaide H, Hayashi K-i, Kawanabe R, Sakigi Y, Matsuo A, Natsume M, Nozaki H (2011) Identification of the single amino acid involved in quenching the ent-kauranyl cation by a water molecule in *ent*-kaurene synthase of *Physcomitrella patens*. FEBS J 278:123–133

Keeling CI, Bohlmann J (2006a) Diterpene resin acids in conifers. Phytochemistry 67:2415–2423

Keeling CI, Bohlmann J (2006b) Genes, enzymes and chemicals of terpenoid diversity in the constitutive and induced defence of conifers against insects and pathogens. New Phytol 170:657–675

Keeling CI, Weisshaar S, Lin RPC, Bohlmann J (2008) Functional plasticity of paralogous diterpene synthases involved in conifer defense. Proc Natl Acad Sci USA 105:1085–1090

Keeling CI, Dullat HK, Yuen M, Ralph SG, Jancsik S, Bohlmann J (2010) Identification and functional characterization of monofunctional *ent*-copalyl diphosphate and *ent*-kaurene synthases in white spruce reveal different patterns for diterpene synthase evolution for primary and secondary metabolism in gymnosperms. Plant Physiol 152:1197–1208

Keeling CI, Weisshaar S, Ralph SG, Jancsik S, Hamberger B, Dullat HK, Bohlmann J (2011) Transcriptome mining, functional characterization, and phylogeny of a large terpene synthase gene family in spruce (*Picea* spp.). BMC Plant Biol 11:43

Kelley LA, Sternberg MJE (2009) Protein structure prediction on the Web: a case study using the Phyre server. Nat Protoc 4:363–371

Kesselmeier J, Staudt M (1999) Biogenic volatile organic compounds (VOC): an overview on emission, physiology and ecology. J Atmos Chem 33:23–88

Köksal M, Zimmer I, Schnitzler J-P, Christianson DW (2010) Structure of isoprene synthase illuminates the chemical mechanism of teragram atmospheric carbon emission. J Mol Biol 402:363–373

Köksal M, Hu H, Coates RM, Peters RJ, Christianson DW (2011a) Structure and mechanism of the diterpene cyclase *ent*-copalyl diphosphate synthase. Nat Chem Biol 7:431–433

Köksal M, Jin Y, Coates RM, Croteau R, Christianson DW (2011b) Taxadiene synthase structure and evolution of modular architecture in terpene biosynthesis. Nature 469:116–120

Köllner TG, Schnee C, Gershenzon J, Degenhardt J (2004) The sesquiterpene hydrocarbons of maize (*Zea mays*) form five groups with distinct developmental and organ-specific distributions. Phytochemistry 65:1895–1902

Koyama T, Ogura K (1999) Isopentenyl diphosphate isomerase and prenyltransferases. In: Cane DE (ed) Comprehensive natural product chemistry: isoprenoids including carotenoids and steroids, vol 2. Pergamon Press, Oxford, pp 69–96

Kulmala M, Nieminen T, Chellapermal R, Makkonen R, Bäck J, Kerminen V-M (2013) Climate feedbacks linking the increasing atmospheric CO_2 concentration, BVOC emissions, aerosols and clouds in forest ecosystems. In: Niinemets Ü, Monson RK (eds) Biology, controls and models of tree volatile organic compound emissions, vol 5, Tree physiology. Springer, Berlin, pp 489–508

Laothawornkitkul J, Taylor JE, Paul ND, Hewitt CN (2009) Biogenic volatile organic compounds in the Earth system. New Phytol 183:27–51

Laule O, Fürholz A, Chang HS, Zhu T, Wang X, Heifetz PB, Gruissem W, Lange M (2003) Crosstalk between cytosolic and plastidial pathways of isoprenoid biosynthesis in *Arabidopsis thaliana*. Proc Natl Acad Sci USA 100:6866–6871

Lee S, Chappell J (2008) Biochemical and genomic characterization of terpene synthases in *Magnolia grandiflora*. Plant Physiol 147:1017–1033

Lee A, Goldstein AH, Kroll JH, Ng NL, Varutbangkul V, Flagan RC, Seinfeld JH (2006) Gas-phase products and secondary aerosol yields from the photooxidation of 16 different terpenes. J Geophys Res Atmos 111:D17305

Lesburg CA, Zhai G, Cane DE, Christianson DW (1997) Crystal structure of pentalenene synthase: mechanistic insights on terpenoid cyclization reactions in biology. Science 277:1820–1824

Li Z, Sharkey T (2013a) Metabolic profiling of the methylerythritol phosphate pathway reveals the source of post-illumination isoprene burst from leaves. Plant Cell Environ 36:429–437

Li Z, Sharkey TD (2013b) Molecular and pathway controls on biogenic volatile organic compound emissions. In: Niinemets Ü, Monson RK (eds) Biology, controls and models of tree volatile organic compound emissions, vol 5, Tree physiology. Springer, Berlin, pp 119–151

Li G, Köllner TG, Yin Y, Jiang Y, Chen H, Xu Y, Gershenzon J, Pichersky E, Chen F (2012) Nonseed plant *Selaginella moellendorffii* has both seed plant and microbial types of terpene synthases. Proc Natl Acad Sci USA 109:14711–14715

Lichtenthaler HK, Rohmer M, Schwender J (1997a) Two independent biochemical pathways for isopentenyl diphosphate and isoprenoid biosynthesis in higher plants. Physiol Plant 101:643–652

Lichtenthaler HK, Schwender J, Disch A, Rohmer M (1997b) Biosynthesis of isoprenoids in higher plant chloroplasts proceeds via a mevalonate-independent pathway. FEBS Lett 400:271–274

Little DB, Croteau RB (2002) Alteration of product formation by directed mutagenesis and truncation of the multiple-product sesquiterpene synthases δ-selinene synthase and γ-humulene synthase. Arch Biochem Biophys 402:120–135

Litvak ME, Monson RK (1998) Patterns of induced and constitutive monoterpene production in conifer needles in relation to insect herbivory. Oecologia 114:531–540

Liu L, Li Y, Li S, Hu N, He Y, Pong R, Lin D, Lu L, Law M (2012) Comparison of next-generation sequencing systems. J Biomed Biotechnol 2012:251364

Loivamäki M, Gilmer F, Fischbach RJ, Sörgel C, Bachl A, Walter A, Schnitzler J-P (2007a) *Arabidopsis*, a model to study biological functions of isoprene emission? Plant Physiol 144:1066–1078

Loivamäki M, Louis S, Cinege G, Zimmer I, Fischbach RJ, Schnitzler J-P (2007b) Circadian rhythms of isoprene biosynthesis in grey poplar leaves. Plant Physiol 143:540–551

Loivamäki M, Mumm R, Dicke M, Schnitzler J-P (2008) Isoprene interferes with the attraction of bodyguards by herbaceous plants. Proc Natl Acad Sci USA 105:17430–17435

Loreto F, Fischbach RJ, Schnitzler J-P, Ciccioli P, Brancaleoni E, Calfapietra C, Seufert G (2001a) Monoterpene emission and monoterpene synthase activities in the Mediterranean evergreen oak *Quercus ilex* L. grown at elevated CO_2. Glob Change Biol 7:709–717

Loreto F, Mannozzi M, Maris C, Nascetti P, Ferranti F, Pasqualini S (2001b) Ozone quenching properties of isoprene and its antioxidant role in leaves. Plant Physiol 126:993–1000

Lücker J, El Tamer MK, Schwab W, Verstappen FWA, van der Plas LHW, Bouwmeester HJ, Verhoeven HA (2002) Monoterpene biosynthesis in lemon (*Citrus limon*). cDNA isolation and functional analysis of four monoterpene synthases. Eur J Biochem 269:3160–3171

Majdi M, Liu Q, Karimzadeh G, Malboobi MA, Beekwilder J, Cankar K, de Vos R, Todorovic S, Simonovic A, Bouwmeester H (2011) Biosynthesis and localization of parthenolide in glandular trichomes of feverfew (*Tanacetum parthenium* L. Schulz Bip.). Phytochemistry 72:1739–1750

Martin D, Bohlmann J (2005) Molecular biochemistry and genomics of terpenoid defenses in conifers. In: Romeo JT (ed) Recent advances in phytochemistry, vol 39. Elsevier, Amsterdam, pp 29–56

Martin D, Tholl D, Gershenzon J, Bohlmann J (2002) Methyl jasmonate induces traumatic resin ducts, terpenoid resin biosynthesis, and terpenoid accumulation in developing xylem of Norway spruce stems. Plant Physiol 129:1003–1018

Martin D, Bohlmann J, Gershenzon J, Francke W, Seybold SJ (2003) A novel sex-specific and inducible monoterpene synthase activity associated with a pine bark beetle, the pine engraver, *Ips pini*. Naturwissenschaften 90:173–179

Martin DM, Faldt J, Bohlmann J (2004) Functional characterization of nine Norway spruce TPS genes and evolution of gymnosperm terpene synthases of the TPS-d subfamily. Plant Physiol 135:1908–1927

Martin DM, Aubourg S, Schouwey MB, Daviet L, Schalk M, Toub O, Lund ST, Bohlmann J (2010) Functional annotation, genome organization and phylogeny of the grapevine (*Vitis vinifera*) terpene synthase gene family based on genome assembly, FLcDNA cloning, and enzyme assays. BMC Plant Biol 10:226

Mayrhofer S, Teuber M, Zimmer I, Louis S, Fischbach RJ, Schnitzler J-P (2005) Diurnal and seasonal variation of isoprene biosynthesis-related genes in grey poplar leaves. Plant Physiol 139:474–484

McAndrew RP, Peralta-Yahya PP, DeGiovanni A, Pereira JH, Hadi MZ, Keasling JD, Adams PD (2011) Structure of a three-domain sesquiterpene synthase: a prospective target for advanced biofuels production. Structure 19:1876–1884

Miller B, Oschinski C, Zimmer W (2001) First isolation of an isoprene synthase gene and successful expression of the gene from poplar in *E. coli*. Planta 213:483–487

Miller B, Madilao LL, Ralph S, Bohlmann J (2005) Insect-induced conifer defense. White pine weevil and methyl jasmonate induce traumatic resinosis, de novo formed volatile emissions, and accumulation of terpenoid synthase and putative octadecanoid pathway transcripts in Sitka spruce. Plant Physiol 137:369–382

Ming R, Yu Q, Moore PH, Paull RE, Chen NJ, Wang M-L, Zhu YJ, Schuler MA, Jiang J, Paterson AH (2008) Genome of papaya, a fast growing tropical fruit tree. Tree Genet Genomes 8:445–462

Modolo LV, Reichert AI, Dixon RA, Osbourn AE, Lanzotti V (2009) Introduction to the different classes of biosynthetic enzymes. In: Osbourn AE, Lanzotti V (eds) Plant-derived natural products synthesis, function, and application. Springer, New York, pp 143–163

Monson RK (2013) Metabolic and gene expression controls on the production of biogenic volatile organic compounds. In: Niinemets Ü, Monson RK (eds) Biology, controls and models of tree volatile organic compound emissions, vol 5, Tree physiology. Springer, Berlin, pp 153–179

Monson RK, Jaeger CH, Adams WW III, Driggers EM, Silver GM, Fall R (1992) Relationships among isoprene emission rate, photosynthesis, and isoprene synthase activity as influenced by temperature. Plant Physiol 98:1175–1180

Monson RK, Jones RT, Rosenstiel TN, Schnitzler J-P (2013) Why only some plants emit isoprene. Plant Cell Environ 33:503–516. doi:10.1111/pce.12015

Morrone D, Chambers J, Lowry L, Kim G, Anterola A, Bender K, Peters RJ (2009) Gibberellin biosynthesis in bacteria: separate ent-copalyl diphosphate and ent-kaurene synthases in *Bradyrhizobium japonicum*. FEBS Lett 583:475–480

Murzin AG, Brenner SE, Hubbard T, Chothia C (2001) SCOP: a structural classification of proteins database for the investigation of sequences and structures. J Mol Biol 247:536–540

Myllykangas S, Buenrostro J, Ji HP (2012) Overview of sequencing technology platforms. In: Rodríguez-Ezpeleta N, Hackenberg M, Aransay AM (eds) Bioinformatics for high throughput sequencing. Springer, Berlin, pp 11–25

Nagegowda DA (2010) Plant volatile terpenoid metabolism: biosynthetic genes, transcriptional regulation and subcellular compartmentation. FEBS Lett 584:2965–2973

Nagegowda DA, Gutensohn M, Wilkerson CG, Dudareva N (2008) Two nearly identical terpene synthases catalyze the formation of nerolidol and linalool in snapdragon flowers. Plant J 55:224–239

Nakano C, Hoshino T (2009) Characterization of the Rv3377c gene product, a type-B diterpene cyclase, from the *Mycobacterium tuberculosis* H37 genome. Chembiochem 10:2060–2071

Owen SM, Peñuelas J (2005) Opportunistic emissions of volatile isoprenoids. Trends Plant Sci 10:420–426

Pechous SW, Whitaker BD (2004) Cloning and functional expression of an (E,E)-α-farnesene synthase cDNA from peel tissue of apple fruit. Planta 219:84–94

Peralta-Yahya PP, Zhang F, del Cardayre SB, Keasling JD (2012) Microbial engineering for the production of advanced biofuels. Nature 488:320–328

Peters RJ, Croteau RB (2002) Abietadiene synthase catalysis: conserved residues involved in protonation-initiated cyclization of geranylgeranyl diphosphate to (+)-copalyl diphosphate. Biochemistry 41:1836–1842

Peters RJ, Ravn MM, Coates RM, Croteau RB (2001) Bifunctional abietadiene synthase: free diffusive transfer of the (+)-copalyl diphosphate intermediate between two distinct active sites. J Am Chem Soc 123:8974–8978

Phillips MA, Wildung MR, Williams DC, Hyatt DC, Croteau R (2003) cDNA isolation, functional expression, and characterization of (+)-α-pinene synthase and (−)-α-pinene synthase from loblolly pine (*Pinus taeda*): stereocontrol in pinene biosynthesis. Arch Biochem Biophys 411:267–276

Pichersky E, Gershenzon J (2002) The formation and function of plant volatiles: perfumes for pollinator attraction and defense. Curr Opin Plant Biol 5:237–243

Pichersky E, Lewinsohn E, Croteau R (1995) Purification and characterization of S-linalool synthase, an enzyme involved in the production of floral scent in *Clarkia breweri*. Arch Biochem Biophys 316:803–807

Pieterse CMJ, Dicke M (2007) Plant interactions with microbes and insects: from molecular mechanisms to ecology. Trends Plant Sci 12:564–569

Possell M, Loreto F (2013) The role of volatile organic compounds in plant resistance to abiotic stresses: responses and mechanisms. In: Niinemets Ü, Monson RK (eds) Biology, controls and models of tree volatile organic compound emissions, vol 5, Tree physiology. Springer, Berlin, pp 209–235

Prlic A, Bliven S, Rose PW, Bluhm WF, Bizon C, Godzik A, Bourne PE (2010) Pre-calculated protein structure alignments at the RCSB PDB website. Bioinformatics 26:2983–2985

Prosser I, Altug IG, Phillips AL, König WA, Bouwmeester HJ, Beale MH (2004) Enantiospecific (+)- and (−)-germacrene D synthases, cloned from goldenrod, reveal a functionally active variant of the universal isoprenoid-biosynthesis aspartate-rich motif. Arch Biochem Biophys 432:136–144

Raffa KF, Aukema BH, Erbilgin N, Klepzig KD, Wallin KF (2005) Interactions among conifer terpenoids and bark beetles across multiple levels of scale: an attempt to understand links between population patterns and physiological processes. In: Romeo JT (ed) Chemical ecology and phytochemistry of forest ecosystems. Recent advances in phytochemistry, vol 39. Elsevier, Amsterdam, pp 79–118

Rasulov B, Copolovici L, Laisk A, Niinemets Ü (2009) Postillumination isoprene emission: in vivo measurements of dimethylallyldiphosphate pool size and isoprene synthase kinetics in aspen leaves. Plant Physiol 149:1609–1618

Rasulov B, Hüve K, Laisk A, Niinemets Ü (2011) Induction of a longer-term component of isoprene release in darkened aspen leaves: origin and regulation under different environmental conditions. Plant Physiol 156:816–831

Ro D-K, Bohlmann J (2006) Diterpene resin acid biosynthesis in loblolly pine (*Pinus taeda*): functional characterization of abietadiene/levopimaradiene synthase (PtTPS-LAS) cDNA and subcellular targeting of PtTPS-LAS and abietadienol/abietadienal oxidase (PtAO, CYP720B1). Phytochemistry 67:1572–1578

Ro DK, Arimura GL, Lau SYW, Piers E, Bohlmann J (2005) Loblolly pine abietadienol/abietadienal oxidase PtAO (CYP720B1) is a multifunctional, multisubstrate cytochrome P450 monooxygenase. Proc Natl Acad Sci USA 102:8060–8065

Rohmer M, Knani M, Simonin P, Sutter B, Sahm H (1993) Isoprenoid biosynthesis in bacteria: a novel pathway for the early steps leading to isopentenyl diphosphate. Biochem J 295:517–524

Rosenkranz M, Schnitzler J-P (2013) Genetic engineering of BVOC emissions from trees. In: Niinemets Ü, Monson RK (eds) Biology, controls and models of tree volatile organic compound emissions, vol 5, Tree physiology. Springer, Berlin, pp 95–118

Rynkiewicz MJ, Cane DE, Christianson DW (2001) Structure of trichodiene synthase from *Fusarium sporotrichioides* provides mechanistic inferences on the terpene cyclization cascade. Proc Natl Acad Sci USA 98:13543–13548

Sallaud C, Rontein D, Onillon S, Jabès F, Duffé P, Giacalone C, Thoraval S, Escoffier C, Herbette G, Leonhardt N, Causse M, Tissier A (2009) A novel pathway for sesquiterpene biosynthesis from Z,Z-farnesyl pyrophosphate in the wild tomato *Solanum habrochaites*. Plant Cell 21:301–317

Saranitzky E, White CM, Baker EL, Baker WL, Coleman CI (2009) Feverfew for migraine prophylaxis: a systematic review. J Dietary Suppl 6:91–103

Sasaki K, Ohara K, Yazaki K (2005) Gene expression and characterization of isoprene synthase from *Populus alba*. FEBS Lett 579:2514–2518

Sasaki K, Saito T, Lämsä M, Oksman-Caldentey K-M, Suzuki M, Ohyama K, Muranaka T, Ohara K, Yazaki K (2007) Plants utilize isoprene emission as a thermotolerance mechanism. Plant Cell Physiol 48:1254–1262

Schepmann HG, Pang J, Matsuda SPT (2001) Cloning and characterization of *Ginkgo biloba* levopimaradiene synthase, which catalyzes the first committed step in ginkgolide biosynthesis. Arch Biochem Biophys 392:263–269

Schnitzler J-P, Zimmer I, Bachl A, Arend M, Fromm J, Fischbach R (2005) Biochemical properties of isoprene synthase in poplar (*Populus x canescens*). Planta 222:777–786

Seemann M, Rohmer M (2007) Isoprenoid biosynthesis via the methylerythritol phosphate pathway: GcpE and LytB, two novel iron-sulphur proteins. Comptes Rendus Chimie 10:748–755

Seemann M, Bui BTS, Wolff M, Miginiac-Maslow M, Rohmer M (2006) Isoprenoid biosynthesis in plant chloroplasts via the MEP pathway: direct thylakoid/ferredoxin-dependent photoreduction of GcpE/IspG. FEBS Lett 580:1547–1552

Seigler DS (1998) Plant secondary metabolism. Springer, New York

Sharkey TD, Singsaas EL (1995) Why plants emit isoprene. Nature 374:769

Sharkey TD, Yeh SS (2001) Isoprene emission from plants. Annu Rev Plant Physiol Plant Mol Biol 52:407–436

Sharkey TD, Yeh S, Wiberley AE, Falbel TG, Gong D, Fernandez DE (2005) Evolution of the isoprene biosynthetic pathway in kudzu. Plant Physiol 137:700–712

Sharkey TD, Wiberley AE, Donohue AR (2008) Isoprene emission from plants: why and how. Ann Bot 101:5–18

Sharkey TD, Gray DW, Pell HK, Breneman SR, Topper L (2013) Isoprene synthase genes form a monophyletic clade of acyclic terpene synthases in the Tps-b terpene synthase family. Evolution 67:1026–1040. doi:10.1111/evo.12013

Sharp PM, Li W-H (1987) The codon adaptation index – a measure of directional synonymous codon usage bias, and its potential applications. Nucleic Acids Res 15:1281–1295

Shiraki T, Kondo S, Katayama S, Waki K, Kasukawa T, Kawaji H, Kodzius R, Watahiki A, Nakamura M, Arakawa T, Fukuda S, Sasaki D, Podhajska A, Harbers M, Kawai J, Carninci P, Hayashizaki Y (2003) Cap analysis gene expression for high-throughput analysis of transcriptional starting point and identification of promoter usage. Proc Natl Acad Sci USA 100:15776–15781

Siedenburg G, Jendrossek D (2011) Squalene-hopene cyclases. Appl Environ Microbiol 77:3905–3915

Singsaas EL, Lerdau M, Winter K, Sharkey TD (1997) Isoprene increases thermotolerance of isoprene-emitting species. Plant Physiol 115:1413–1420

Starks CM, Back KW, Chappell J, Noel JP (1997) Structural basis for cyclic terpene biosynthesis by tobacco 5-epi-aristolochene synthase. Science 277:1815–1820

Staudt M, Joffre R, Rambal S (2003) How growth conditions affect the capacity of *Quercus ilex* leaves to emit monoterpenes. New Phytol 158:61–73

Steele CL, Katoh S, Bohlmann J, Croteau R (1998) Regulation of oleoresins in grand fir (*Abies grandis*). Differential transcriptional control of monoterpene, sesquiterpene, and diterpene synthase genes in response to wounding. Plant Physiol 116:1497–1504

Stofer Vogel B, Wildung MR, Vogel G, Croteau R (1996) Abietadiene synthase from grand fir (*Abies grandis*). cDNA isolation, characterization, and bacterial expression of a bifunctional diterpene cyclase involved in resin acid biosynthesis. J Biol Chem 271:23262–23268

Sun Z, Niinemets Ü, Hüve K, Noe SM, Rasulov B, Copolovici L, Vislap V (2012) Enhanced isoprene emission capacity and altered light responsiveness in aspen grown under elevated atmospheric CO_2 concentration. Glob Change Biol 18:3423–3440

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28:2731–2739

Thaler JS (2002) Effect of jasmonate-induced plant responses on the natural enemies of herbivores. J Anim Ecol 71:141–150

Tholl D (2006) Terpene synthases and the regulation, diversity and biological roles of terpene metabolism. Curr Opin Plant Biol 9:297–304

Tissier A (2012) Interactions among conifer terpenoids and bark beetles across multiple levels of scale: an attempt to understand links between population patterns and physiological processes. In: Çiftçi YO (ed) Transgenic plants – advances and limitations. InTech, Rijeka, pp 353–378

Trapp SC, Croteau RB (2001) Genomic organization of plant terpene synthases and molecular evolutionary implications. Genetics 158:811–832

Trowbridge AM, Stoy PC (2013) BVOC-mediated plant-herbivore interactions. In: Niinemets Ü, Monson RK (eds) Biology, controls and models of tree volatile organic compound emissions, vol 5, Tree physiology. Springer, Berlin, pp 21–46

Trusheva B, Todorov I, Ninova M, Najdenski H, Daneshmand A, Bankova V (2010) Antibacterial mono- and sesquiterpene esters of benzoic acids from Iranian propolis. Chem Cent J 4:8

Tuskan GA, DiFazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S, Rombauts S, Salamov A, Schein J, Sterck L, Aerts A, Bhalerao RR, Bhalerao RP, Blaudez D, Boerjan W, Brun A, Brunner A, Busov V, Campbell M, Carlson J, Chalot M, Chapman J,

Chen G-L, Cooper D, Coutinho PM, Couturier J, Covert S, Cronk Q, Cunningham R, Davis J, Degroeve S, Déjardin A, dePamphilis C, Detter J, Dirks B, Dubchak I, Duplessis S, Ehlting J, Ellis B, Gendler K, Goodstein D, Gribskov M, Grimwood J, Groover A, Gunter L, Hamberger B, Heinze B, Helariutta Y, Henrissat B, Holligan D, Holt R, Huang W, Islam-Faridi N, Jones S, Jones-Rhoades M, Jorgensen R, Joshi C, Kangasjärvi J, Karlsson J, Kelleher C, Kirkpatrick R, Kirst M, Kohler A, Kalluri U, Larimer F, Leebens-Mack J, Leplé J-C, Locascio P, Lou Y, Lucas S, Martin F, Montanini B, Napoli C, Nelson DR, Nelson C, Nieminen K, Nilsson O, Pereda V, Peter G, Philippe R, Pilate G, Poliakov A, Razumovskaya J, Richardson P, Rinaldi C, Ritland K, Rouzé P, Ryaboy D, Schmutz J, Schrader J, Segerman B, Shin H, Siddiqui A, Sterky F, Terry A, Tsai C-J, Uberbacher E, Unneberg P, Vahala J, Wall K, Wessler S, Yang G, Yin T, Douglas C, Marra M, Sandberg G, Van de Peer Y, Rokhsar D (2006) The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). Science 313:1596–1604

Ujino-Ihara T, Yoshimura K, Ugawa Y, Yoshimaru H, Nagasaka K, Tsumura Y (2000) Expression analysis of ESTs derived from the inner bark of *Cryptomeria japonica*. Plant Mol Biol 43:451–457

van Schie CCN, Haring MA, Schuurink RC (2007) Tomato linalool synthase is induced in trichomes by jasmonic acid. Plant Mol Biol 64:251–263

Velasco R, Zharkikh A, Affourtit J, Dhingra A, Cestaro A, Kalyanaraman A, Fontana P, Bhatnagar SK, Troggio M, Pruss D, Salvi S, Pindo M, Baldi P, Castelletti S, Cavaiuolo M, Coppola G, Costa F, Cova V, Dal Ri A, Goremykin V, Komjanc M, Longhi S, Magnago P, Malacarne G, Malnoy M, Micheletti D, Moretto M, Perazzolli M, Si-Ammour A, Vezzulli S, Zini E, Eldredge G, Fitzgerald LM, Gutin N, Lanchbury J, Macalma T, Mitchell JT, Reid J, Wardell B, Kodira C, Chen Z, Desany B, Niazi F, Palmer M, Koepke T, Jiwan D, Schaeffer S, Krishnan V, Wu C, Chu VT, King ST, Vick J, Tao Q, Mraz A, Stormo A, Stormo K, Bogden R, Ederle D, Stella A, Vecchietti A, Kater MM, Masiero S, Lasserre P, Lespinasse Y, Allan AC, Bus V, Chagne D, Crowhurst RN, Gleave AP, Lavezzo E, Fawcett JA, Proost S, Rouze P, Sterck L, Toppo S, Lazzari B, Hellens RP, Durel C-E, Gutin A, Bumgarner RE, Gardiner SE, Skolnick M, Egholm M, Van de Peer Y, Salamini F, Viola R (2010) The genome of the domesticated apple (*Malus* x *domestica* Borkh.). Nat Genet 42:833–839

Velculescu VE, Zhang L, Vogelstein B, Kinzler KW (1995) Serial analysis of gene expression. Science 270:484–487

Velikova V, Várkonyi Z, Szabó M, Maslenkova L, Nogues I, Kovács L, Peeva V, Busheva M, Garab G, Sharkey TD, Loreto F (2011) Increased thermostability of thylakoid membranes in isoprene-emitting leaves probed with three biophysical techniques. Plant Physiol 157:905–916

Vickers CE, Gershenzon J, Lerdau MT, Loreto F (2009a) A unified mechanism of action for volatile isoprenoids in plant abiotic stress. Nat Chem Biol 5:283–291

Vickers CE, Possell M, Cojocariu CI, Velikova VB, Laothawornkitkul J, Ryan A, Mullineaux PM, Hewitt CN (2009b) Isoprene synthesis protects transgenic tobacco plants from oxidative stress. Plant Cell Environ 32:520–531

Vickers CE, Possell M, Hewitt CN, Mullineaux PM (2010) Genetic structure and regulation of isoprene synthase in poplar (*Populus* spp.). Plant Mol Biol 73:547–558

Vickers CE, Possell M, Laothawornkitkul J, Ryan AC, Hewitt CN, Mullineaux PM (2011) Isoprene synthesis in plants: lessons from a transgenic tobacco model. Plant Cell Environ 34:1043–1053

Vogel BS, Wildung MR, Vogel G, Croteau R (1996) Abietadiene synthase from grand fir (*Abies grandis*). cDNA isolation, characterization, and bacterial expression of a bifunctional diterpene cyclase involved in resin acid biosynthesis. J Biol Chem 271:23262–23268

Vu TT, Vohradsky J (2007) Nonlinear differential equation model for quantification of transcriptional regulation applied to microarray data of *Saccharomyces cerevisiae*. Nucleic Acids Res 35:279–287

Wang Z, Gerstein M, Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet 10:57–63

Wanke M, Skorupinska-Tudek K, Swiezewska E (2001) Isoprenoid biosynthesis via 1-deoxy-D-xylulose 5-phosphate/2-C-methyl-D-erythritol 4-phosphate (DOXP/MEP) pathway. Acta Biochim Pol 48:663–672

Wendt KU, Poralla K, Schulz GE (1997) Structure and function of a squalene cyclase. Science 277:1811–1815

Whittington DA, Wise ML, Urbansky M, Coates RM, Croteau RB, Christianson DW (2002) Bornyl diphosphate synthase: structure and strategy for carbocation manipulation by a terpenoid cyclase. Proc Natl Acad Sci USA 99:15375–15380

Wiberley AE, Donohue AR, Meier ME, Westphal MM, Sharkey TD (2008) Regulation of isoprene emission in *Populus trichocarpa* leaves subjected to changing growth temperature. Plant Cell Environ 31:258–267

Wiberley AE, Donohue AR, Westphal MM, Sharkey TD (2009) Regulation of isoprene emission from poplar leaves throughout a day. Plant Cell Environ 32:939–947

Wildermuth MC, Fall R (1996) Light-dependent isoprene emission. Characterization of a thylakoid-bound isoprene synthase in *Salix discolor* chloroplasts. Plant Physiol 112:171–182

Wildermuth MC, Fall R (1998) Biochemical characterization of stromal and thylakoid-bound isoforms of isoprene synthase in willow leaves. Plant Physiol 116:1111–1123

Wildung MR, Croteau R (1996) A cDNA clone for taxadiene synthase, the diterpene cyclase that catalyzes the committed step of taxol biosynthesis. J Biol Chem 271:9201–9204

Wilkinson MJ, Monson RK, Trahan N, Lee S, Brown E, Jackson RB, Polley HW, Fay PA, Fall R (2009) Leaf isoprene emission rate as a function of atmospheric CO_2 concentration. Glob Change Biol 15:1189–1200

Williams DC, Wildung MR, Jin AQ, Dalal D, Oliver JS, Coates RM, Croteau R (2000) Heterologous expression and characterization of a "pseudomature" form of taxadiene synthase involved in paclitaxel (taxol) biosynthesis and evaluation of a potential intermediate and inhibitors of the multistep diterpene cyclization reaction. Arch Biochem Biophys 379:137–146

Wilson D, Pethica R, Zhou Y, Talbot C, Vogel C, Madera M, Chothia C, Gough J (2009) SUPERFAMILY – sophisticated comparative genomics, data mining, visualization and phylogeny. Nucleic Acids Res 37:D380–D386

Wu S, Schalk M, Clark A, Miles RB, Coates R, Chappell J (2006) Redirection of cytosolic or plastidic isoprenoid precursors elevates terpene production in plants. Nat Biotechnol 24:1441–1447

Wu J, Wang Z, Shi Z, Zhang S, Ming R, Zhu S, Khan MA, Tao S, Korban SS, Wang H, Chen NJ, Nishio T, Xu X, Cong L, Qi K, Huang X, Wang Y, Zhao X, Wu J, Deng C, Gou C, Zhou W, Yin H, Qin G, Sha Y, Tao Y, Chen H, Yang Y, Song Y, Zhan D, Wang J, Li L, Dai M, Gu C, Wang Y, Shi D, Wang X, Zhang H, Zeng L, Zheng D, Wang C, Chen M, Wang G, Xie L, Sovero V, Sha S, Huang W, Zhang S, Zhang M, Sun J, Xu L, Li Y, Liu X, Li Q, Shen J, Wang J, Paull RE, Bennetzen JL, Wang J, Zhang S (2013) The genome of pear (*Pyrus bretschneideri* Rehd.). Genome Res 23:396–408. doi:10.1101/gr.144311.144112

Xie X, Kirby J, Keasling JD (2012) Functional characterization of four sesquiterpene synthases from *Ricinus communis* (castor bean). Phytochemistry 78:20–28

Yin SH, Mei L, Newman J, Back K, Chappell J (1997) Regulation of sesquiterpene cyclase gene expression – characterization of an elicitor- and pathogen-inducible promoter. Plant Physiol 115:437–451

Zerbe P, Chiang A, Bohlmann J (2012a) Mutational analysis of white spruce (*Picea glauca*) *ent*-kaurene synthase (PgKS) reveals common and distinct mechanisms of conifer diterpene synthases of general and specialized metabolism. Phytochemistry 74:30–39

Zerbe P, Chiang A, Yuen M, Hamberger B, Hamberger B, Draper JA, Britton R, Bohlmann J (2012b) Bifunctional *cis*-abienol synthase from *Abies balsamea* discovered by transcriptome sequencing and its implications for diterpenoid fragrance production. J Biol Chem 287:12121–12131

Zhou K, Peters RJ (2009) Investigating the conservation pattern of a putative second terpene synthase divalent metal binding motif in plants. Phytochemistry 70:366–369

Zhou K, Gao Y, Hoy JA, Mann FM, Honzatko RB, Peters RJ (2012) Insights into diterpene cyclization from structure of bifunctional abietadiene synthase from *Abies grandis*. J Biol Chem 287:6840–6850