Journal of Experimental Botany, Vol. 73, No. 2 pp. 555–570, 2022
https://doi.org/10.1093/jxb/erab289 Advance Access Publication 15 June 2021

RESEARCH PAPER

Transcriptional up-regulation of host-specific terpene metabolism in aphid-induced galls of *Pistacia palaestina*

Rachel Davidovich-Rikanati[1], Einat Bar[1], Gal Hivert[1,2], Xing-Qi Huang[3,4], Carolina Hoppen-Tonial[5,6], Vered Khankin[7], Karin Rand[1,8], Amal Abofreih[7], Joelle K. Muhlemann[3,*], José Abramo Marchese[5], Yoram Shotland[7], Natalia Dudareva[3,4,†], Moshe Inbar[8] and Efraim Lewinsohn[1,2,†]

[1] Institute of Plant Sciences, Newe Ya'ar Research Center, Agricultural Research Organization, The Volcani Center, Ramat Yishay, 30095, Israel
[2] Faculty of Agriculture, The Hebrew University of Jerusalem, Rehovot, Israel
[3] Department of Biochemistry, Purdue University, West Lafayette, IN 47907-1165, USA
[4] Purdue Center for Plant Biology, Purdue University, West Lafayette, IN 47907, USA
[5] Department of Agronomy, Federal University of Technology - Paraná, Pato Branco, 85503-390, Brazil
[6] Department of Agronomy, Federal Institute of Paraná, Palmas, 85555-000, Brazil
[7] Department of Chemical Engineering, Shamoon College of Engineering, Beer Sheva, 84100, Israel
[8] Department of Evolutionary & Environmental Biology, University of Haifa, Mount Carmel, Haifa, 3498838, Israel

* Present address: Department of Biology, Wake Forest University, Winston-Salem, NC 27109, USA.
† Correspondence: twefraim@agri.gov.il or dudareva@purdue.edu

Received 14 May 2021; Editorial decision 8 June 2021; Accepted 11 June 2021

Editor: Robert Hancock, The James Hutton Institute, UK

Abstract

Galling insects gain food and shelter by inducing specialized anatomical structures in their plant hosts. Such galls often accumulate plant defensive metabolites protecting the inhabiting insects from predation. We previously found that, despite a marked natural chemopolymorphism in natural populations of *Pistacia palaestina*, the monoterpene content in *Baizongia pistaciae*-induced galls is substantially higher than in leaves of their hosts. Here we show a general up-regulation of key structural genes in both the plastidial and cytosolic terpene biosynthetic pathways in galls as compared with non-colonized leaves. Novel prenyltransferases and terpene synthases were functionally expressed in *Escherichia coli* to reveal their biochemical function. Individual *Pistacia* trees exhibiting chemopolymorphism in terpene compositions displayed differential up-regulation of selected terpene synthase genes, and the metabolites generated by their gene products *in vitro* corresponded to the monoterpenes accumulated by each tree. Our results delineate molecular mechanisms responsible for the formation of enhanced monoterpene in galls and the observed intraspecific monoterpene chemodiversity displayed in *P. palaestina*. We demonstrate that gall-inhabiting aphids transcriptionally reprogram their host terpene pathways by up-regulating tree-specific genes, boosting the accumulation of plant defensive compounds for the protection of colonizing insects.

Keywords: *Baizongia pistaciae*, extended phenotype, gall-forming insects, monoterpene biosynthesis, *Pistacia palaestina*, plant defense compounds, prenyl transferases, terpene metabolism, terpene synthases.

Introduction

Interactions between plants and their environments often result in profound remodeling of plant anatomy and metabolism. A prime example of the way insects manipulate their plant hosts for their own benefit is the gall-forming habit. Galls are modified plant organs that provide the gall-inducing organisms with steady nutrient supplies and efficient protection from abiotic factors and natural enemies (Price *et al.*, 1987). Insect-induced galls often accumulate plant defensive chemicals such as phenolics and terpenes, with their concentrations being substantially higher in galls as compared with intact non-colonized plant tissues (Caputo *et al.*, 1979; Cornell, 1983; Hartley, 1998; Nyman and Julkunen-Tiitto, 2000; Rostás *et al.*, 2013; Rand *et al.*, 2014, 2017). Transcriptomic analysis revealed that gall-forming insects reprogram plant gene expression and metabolism, modifying plant host anatomy and physiology for their own benefit (Nabity *et al.*, 2013, 2016; Schultz *et al.*, 2019). Nevertheless, little is known about how gall-inducing insects manipulate plant defense metabolism.

Several aphid species (Hemiptera, Fordini) induce galls on plants of the genus *Pistacia* (Anacardiaceae). *Pistacia palaestina* Boiss., a common deciduous tree in the Eastern Mediterranean basin, serves as an obligate host for at least six gall-forming aphid species (Inbar *et al.*, 2004). Each aphid species forms a unique visually distinguishable gall type, which is often referred to as the extended phenotype of the insect (Inbar *et al.*, 2004). *Baizongia pistaciae* is an aphid that induces large (≥25 cm long), banana-like galls on the terminal buds of *P. palaestina* branches (Wool, 2012). Each gall may support thousands of phloem-feeding aphids during the dry summer season (Wool, 2012). Plant-derived chemical constituents of the galls protect colonizing aphids from predation (Rostás *et al.*, 2013). The monoterpene concentrations in galls are greatly enhanced as compared with leaves, and their compositions vary between individual trees and the gall tissues (Rand *et al.*, 2014, 2017).

Terpenes constitute the most diverse class of plant specialized metabolites (Gershenzon and Croteau, 1992; Croteau *et al.*, 2000; Keeling and Bohlmann, 2006). They play broad biological and ecological roles, mediating direct and indirect plant defenses, serving as growth regulators, stabilizers of membrane structures, and pollinator attractants (Kessler and Baldwin, 2001; Wittstock and Gershenzon, 2002; Erbilgin *et al.*, 2006; Heil, 2014). To date, many of the terpenoid pathway structural genes have been identified in the plant kingdom (Martin *et al.*, 2004; Keeling and Bohlmann, 2006; Degenhardt *et al.*, 2009; Chen *et al.*, 2011; Pazouki *et al.*, 2015). The key enzymes responsible for monoterpene formation in plants are monoterpene synthases, members of the terpene synthase (TPS) family (Bohlmann *et al.*, 1998). They often produce several monoterpenes from a single substrate, mostly geranyl diphosphate (GPP) (Davidovich-Rikanati *et al.*, 2007; Gutensohn *et al.*, 2013). GPP is formed by the condensation of isopentenyl diphosphate (IPP) with dimethylallyl diphosphate (DMAPP), both of which are predominantly derived from the plastidial methylerythritol phosphate (MEP) pathway (Tholl, 2015). Geranyl diphosphate synthases (GPPSs), the enzymes generating GPP, are members of the prenyltransferase (PRT) family, which have different architectures depending on the plant species (Nagegowda, 2010). Gymnosperm and angiosperm species possess homomeric GPPSs (Bouvier *et al.*, 2000; Schmidt and Gershenzon, 2007; Van Schie *et al.*, 2007; Hsiao *et al.*, 2008; Schmidt *et al.*, 2010; Rai *et al.*, 2013), while heteromeric GPPSs have been functionally characterized in some angiosperms including *Mentha*, *Solanum*, *Antirrhinum*, and *Humulus* (Burke *et al.*, 2004; Tholl *et al.*, 2004; Orlova *et al.*, 2009; Wang and Dixon, 2009; Hivert *et al.*, 2020). Heteromeric GPPSs consist of a catalytically inactive small subunit that upon interaction with a large subunit, favors GPP formation at the expense of larger prenylated phosphates such as geranylgeranyl diphosphate. To date, however, limited information is available about the molecular and biochemical platforms supporting terpenoid biosynthesis in *Pistacia* or any other Anacardiaceae species.

Here, we investigated the molecular and biochemical mechanisms that allow galling aphids to manipulate host chemical defenses. We found that gall-inhabiting aphids up-regulate key genes of both the plastidial and cytosolic terpenoid pathways, enabling accumulation of high monoterpene concentrations in galls. This is also accompanied by a prominent up-regulation of *PpPRT2*, a gene encoding a PRT promoting GPP formation in the presence of a large GPPS subunit. Moreover, tree-specific monoterpene polymorphism observed in *P. palaestina* is achieved by differential expression of different TPS genes in galls from distinct trees. Our results provide evidence that galling insects rely on specific plant metabolic capacities, which they transcriptionally manipulate for their own benefit.

Materials and methods

Plant material

Samples were collected from two naturally growing trees on Mount Carmel (further 'Carmel') and in the Misgav region at the lower Galilee (hear further 'Shechania') both in Northern Israel. Young galls in stage 1 (~4–6 weeks with size of 5–20 mm, inhabiting 1–5 aphids) were sampled in the spring of 2016 according to Rand *et al.* (2014, 2017). The galls were cut open and the aphids were carefully removed with a fine brush and gently washed with acetone. The samples were fast frozen with liquid nitrogen and stored at −80 °C until analysis. Three independent biological replicates from each tree and tissue were transcriptomically and metabolomically examined. Each biological replicate consisted of 1–3 galls or leaves that were pooled together for each individual replicate and ground before RNA extraction and volatile analyses.

Identification and quantification of volatile terpenes in galls and leaves using chiral GC-MS analyses

We ground 0.2 g of frozen leaf or gall tissues in the presence of liquid nitrogen, and extracted the powder by vigorous shaking overnight at room temperature with 2 ml of methyl *tert*-butyl ether (MTBE) containing 25 ppm of ethyl myristate as an internal standard (Rand *et al.*, 2014, 2017). The ethereal phase was separated, dried with anhydrous Na_2SO_4, and analyzed by GC-MS as follows. A 1 µl aliquot of the MTBE extract was injected into an Agilent GC-MS system (model 6890N/5973N, Agilent Technologies, Santa Clara, CA, USA) equipped with a Restek Rt™-bDEXsm Chiral column (30 m length 0.25 mm i.d., 0.25 µm film thickness, 2,3-di-*O*-methyl-6-*O*-*tert*-butyl dimethylsilyl β-cyclodextrin added into 14% cyanopropylphenyl/86% dimethyl polysiloxane). Helium (0.8 ml min^{-1}) was used as a carrier gas with splitless injection. The injector and detector temperatures were 230 °C. The following conditions were used: initial temperature 40 °C for 5 min, followed by a ramp of 2 °C min^{-1} to 110 °C, and 5 °C min^{-1} up to 220 °C (10 min). A quadrupole mass detector with electron ionization at 70 eV was used to acquire the MS data in the range of 41–350 m/z. The identification of the volatiles was assigned by comparison of spectral data with authentic standards and with the W10N11 and HPCH2205 GC-MS libraries. The amount of each component in the sample was calculated as (peak area×internal standard response factor) divided by (response factor×internal standard peak area) as described in Rand *et al.* (2014, 2017).

RNA extraction, sequencing, and bioinformatic analyses

RNA samples were isolated using a Spectrum™ Plant Total RNA Kit (Sigma Aldrich, St. Louis, MO, USA), followed by treatment with DNase I (Sigma Aldrich). cDNA libraries were constructed using the TruSeq RNA sample preparation kit (Illumina, San Diego, CA, USA) essentially as directed by the manufacturer, but with the following modifications. Fragmentation time in the eluted fragment prime step was reduced from 8 min to 4 min, resulting in cDNA ranging in length from 100 bp to 1000 bp. Subsequent AMPure purification was performed at a sample to AMPure ratio of 1.8:1 (v/v). The resulting libraries were composed of amplicons largely ranging in length from 200 bp to 1000 bp. The libraries were sequenced on an Illumina HiSeq2500 using SBS v3 reagents to produce 100 bp paired-end reads.

Assembly of reads was performed using Trinity software (Haas *et al.*, 2013). RNA sequence quality of one out of three samples of Shechania galls gave poor results and therefore they were not used for transcriptome analyses. Library construction generated a total of 428 824 968 reads (see Supplementary Table S1). These reads were subjected to quality and adaptor trimming, which removed 1 881 122 (0.4%) reads before final *de novo* assembly. The assembly resulted in a total of 42 150 and 43 852 contigs (Supplementary Table S1) with an N50 of 2400 bp and 1990 bp (Supplementary Table S2) for the Carmel and Shechania trees, respectively. These contigs were used for downstream bioinformatics analysis.

Additionally, a total of 401 million reads (94% of all clean reads) for Carmel trees, and 386 million reads (96% of all clean reads) for Shechania trees were perfectly mapped (mismatch=0) to the reference transcriptome (each read to the transcriptome of its own tree) by RSEM, which showed that the quality of these mapped genes was sufficient to conduct the subsequent analysis.

The RSEM program was used to determine expression levels for all transcripts. To identify transcripts involved in terpenoid biosynthesis, the transcriptomes were annotated using diverse databases (Supplementary Table S3) and queried by BLASTX using previously characterized genes implicated in the terpenoid biosynthetic pathway.

Genes of the MEP (2-C-methyl-D-erythritol 4-phosphate) plastidial pathway, the mevalonate (MVA) cytosolic pathway, as well as genes encoding isopentenyl diphosphate isomerase (IDI), PRTs, and mono-, sesqui-, and diterpene synthases from *Arabidopsis thaliana* and other plant systems (such as mango and citrus) were used as baits to identify the corresponding genes in the transcriptomes of both galls and leaves.

The transcriptomic data have been deposited in NCBI under submission number 2464236. The accession numbers of all the genes in this study are indicated in Supplementary Table S3.

Cloning of candidate TPSs and their heterologous expression in E. coli

The assembled sequences generated by the preliminary Illumina methodology were aligned using the Blast algorithm against the NCBI (http://www.ncbi.nlm.nih.gov) protein databases. This search allowed the identification of five full-length *P. palaestina* putative TPS sequences (PpTPS1, PpTPS4, PpTPS7, PpTPS11, and PpTPS12). TPS1 and PpTPS4 sequences were members of the TPS-b subfamily comprising angiosperm monoterpene synthase-like genes and were cloned from Carmel gall cDNA. An additional TPS, PpTPS3, found in the new Shechania assembly was cloned from Shechania gall cDNA.

Specific primers were designed to clone the full length of the corresponding ORFs into a pGEX4T1 vector (GE Healthcare, Chicago, IL, USA) using suitable restriction sites. cDNA derived from gall tissue RNA was synthesized with 200 ng of RNA by reverse transcription–PCR (RT–PCR) using the cDNA Verso Kit (Thermo Fisher Scientific, Waltham, MA, USA) and 25 pmol of oligo(dT) primer.

Specific primers corresponding to the *P. palaestina* TPSs and containing restriction enzyme sites were designed as follows: the PpTPS1 forward primer (5′-TCTGGATCCATGGCTTCTTGTATTGTATCA-3′) containing a *Bam*HI restriction site and the reverse primer (5′-GTTGTCGACTCATGGCTCTTTATTGTCAG-3′) containing a *Sal*I restriction site; PpTPS3 was cloned from Shechania gall cDNA into pET28a with the PpTPs3 forward primer (5′-GTTGTCGACTCATGGCTCTTTATTGTCAG-3′) containing a *Sal*I restriction site and reverse primer (5′-GTTGTCGACTCATGGCTCTTTATTGTCAG-3′) containing a *Not*I restriction site; the PpTPS4 forward primer (5′-TCTGGATCCATGGCGTTCTGCATTG-3′) containing a *Bam*HI restriction site and a reverse primer (5′-AGAGTCGACCTACAGTGGAATAGA-3′) containing a *Sal*I restriction site.

The primers were synthesized by Hylabs Company, Israel. After obtaining the cDNA, high-fidelity RT–PCR was performed with the specific primers for each TPS sequence in order to amplify them. For this, Phusion DNA Polymerase (Thermo Fisher Scientific) was used according to the manufacturer's instructions. For expression in *E. coli*, amplicons obtained by RT–PCR were inserted into the expression vectors as described above. PCR products and the plasmids were cleaved with FastDigest restriction enzymes (Thermo Fisher Scientific) according to the manufacturer's instructions. The amplification products were then ligated to the plasmid using the Rapid DNA Ligation Kit (Thermo Fisher Scientific). The recombinant plasmid was then sequenced for authentication and transformed into BL21 (DE3) pLys *E. coli*.

Heterologous expression of selected genes and functional analysis

For heterologous functional expression of TPSs, the plasmids containing the putative TPS sequences were transformed into *E. coli* BL21(DE3) pLysS competent cells. A starter culture was incubated overnight at 37 °C, diluted 1:50 (v/v) next day, and grown till an OD_{600} of 0.8, when 0.1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) was added for induction of expression, and continued to grow at 18 °C and 250 rpm overnight. The cells were collected by centrifugation and disrupted by 10×30 s treatments in an Ultrasonicator in the presence of suitable buffers according to protein Tag properties [PpTPS1, PpTPS4. PpTPS7, PpTPS11, and PpTPS12 harbored a glutathione *S*-transferase (GST) tag derived from the pGEX4t-1 vector, while PpTPS3 had a His tag from the

pET28a vector]. GST ligation buffer contained 140 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.3) with 0.25 mg of lysozyme. Cell debris were removed by centrifugation at 12 000 rpm and the supernatant was collected and purified on a GST column (GSTrap FastFlow, GE Healthcare) with GST elution buffer (50 mM Tris–HCl, 10 mM reduced glutathione, pH 8.0) following the manufacturer's protocol. His-tagged protein purification was as described by Gonda *et al.* (2018). Extracted proteins were desalted and concentrated using VivaSpin20 (Sartorius, Epsom, UK) then were measured by Bradford protein assay (Bradford, 1976).

To determine the potential TPS catalytic activities of the cloned genes, assays containing 200 µl of protein were incubated overnight at 30 °C with 10 µM GPP, farnesyl pyrophosphate (FPP), or geranylgeranyl pyrophosphate (GGPP) in a final volume of 400 µl using a buffer containing 50 mM Bis-Tris, pH 6.9, 1 mM DTT, 10% (v/v) glycerol, 0.1 mM $MnCl_2$, and 10 mM $MgCl_2$. As controls, we utilized reactions without the cofactors Mn_2^+ and Mg_2^+ or extracts using heat-inactivated enzyme (previously incubated at 100 °C for 5 min) or the short peptide product of an empty vector. The volatiles produced were extracted by solid-phase microextraction (SPME) according to the methodology described before using a chiral column for enantiomer separations. In breaf, pre-heated headspace (55 °C) samples were absorbed with an SPME fiber of 65 µm PDMS/DVB/CAR (polydimethylsiloxane/divinylbenzene, carboxene) Supelco (Bellefonte, PA, USA) for 15 min by an automatic HS-SPME MPS2 (Gerstel, Mülheim, Germany). Then, the SPME fiber was introduced into the injector port of the GC-MS apparatus for 5 min (splitless) for desorption of the volatiles. GC-MS conditions were as described above except that the initial temperature was set to 40 °C for 5 min, followed by a ramp of 5 °C min^{-1} to 230 °C (5 min).

For enzymatic di-terpene synthase assays, the samples were pre-heated to 60 °C, for 10 min and then adsorbed for 30 min by an automatic HS-SPME MPS2 system (Gerstel) by 65 µm PDMS/DVB/CAR fiber (Supelco) that was injected into the GC-MS (Agilent GC-MS 7890 system) for 5 min, for desorption of the volatiles in splitless mode. The GC was equipped with a Rxi-5sil MS column (30 m length, 0.25 mm i.d., 0.25 µm film thickness, stationary phase 95% dimethyl–5% diphenyl polysiloxane). Helium (11.68 psi) was used as a carrier gas, the injector temperature was 250 °C, and the detector temperature was 280 °C. The following conditions were used: initial temperature 50 °C for 1 min, followed by a ramp of 5 °C min^{-1} to 280 °C (10 min). A quadrupole mass detector with electron ionization at 70 eV was used to acquire the MS data in the range of 41–350 m/z. A mixture of straight-chain alkanes (C7–C23) was injected into the column under the above-mentioned conditions for determination of retention indices. The identification of the volatiles was assigned by comparison of their retention indices with those in the literature and by comparison of spectral data with a standard or with the W10N14 and QuadLib 2205 GC-MS libraries.

Functional expression of prenyltransferase genes

Two genes encoding putative PRTs, PpGGPPS24311 (resembling other plant GPPS.LSUs, later termed *PpPRT1*) and PpC_GPPS.SSU23269 (resembling the GPPS.SSU, later termed *PpPRT2*), were synthesized by HyLabs (Israel), with codon adaptation for bacterial expression and without the signal peptide targeting the encoded protein to the plastids. The genes were cloned into the pET30 (Novagen, Madison, WI, USA) expression vector fused to a His tag at the N-terminus and then transformed to *E. coli* BL-21 DE3 pLYsS cells (Novagen) for protein expression. The recombinant proteins were purified from the cell lysate by nickel affinity chromatography as described by Gonda *et al.* (2018). Due to the high tendency of PpPRT2 to form inclusion bodies, a non-nucleotide optimized version of PpPRT2 was also directly amplified from *Baizongia* gall cDNA derived from the Carmel tree omitting the signal peptide, and cloned into pMAL-c6T fusion protein vector (New England BioLabs, Ipswich, MA, USA) that enabled soluble recombinant protein fused with maltose-binding protein The latter fusion protein was isolated with maltose resin and later cleaved with TEV protease according to pMAL working protocols (New England BioLabs).

The enzymatic activity assays were performed in glass GC vials using 25 mM Bis-Tris buffer with 10 mM $MgCl_2$ and 2 mM DTT at pH 7.4, containing 40 µM IPP and 40 µM DMAPP as substrates. The assay volume was 300 µl. The reaction was incubated overnight at 30 °C and terminated by adding 40 µl of 1 N HCl for 30 min at 30 °C to attain full acid hydrolysis of the prenylphosphate products and enable their detection by GC-MS. The alcoholic products were determined by SPME analysis and separated on a Rtx-5 SIL MS (30 m×30.25 mm×30.25 µm) fused silica capillary column (Restek Co., Bellefonte, PA, USA) using a temperature gradient of 50 °C for 1 min, increasing to 180 °C at a rate of 5 °C min^{-1}, then to 320 °C at 20 °C min^{-1}. The detector temperature was set at 280 °C. Identification of the hydrolysis products was performed as indicated above for SPME analyses (Orlova *et al.*, 2009; Hivert *et al.*, 2020). The distribution of the prenylated products was based on the abundance of their acid hydrolysis products as calibrated with authentic standards as described before (Hivert *et al.*, 2020).

Quantitative real-time PCR

Quantitative PCR was performed using Absolute Blue qPCR SYBR Green ROX Mix (Thermo Fisher Scientific). The reactions were in a 10 µl final volume containing 1 µl of reverse-transcribed cDNA as template and specific primers as shown in Supplementary Table S4. The primers were synthesized by Hylabs Company, located in Israel.

Samples were run in triplicate in an optical 96-well plate. The following PCR conditions were applied for all reactions: denaturation at 95 °C for 15 min, followed by 40 cycles of 15 s at 95 °C and 1 min at 60 °C, and, finally, a dissociation curve with denaturation at 95 °C for 15 s, cooling at 60 °C for 1 min, and a gradual increase of 0.6 °C until 95 °C. The melting curves were obtained at the end of cycling from 55 °C to 95 °C.

The fold increase in expression levels between leaves and galls was calculated based on the 2−ΔΔCT method (Livak and Schmittgen, 2001) using ubiquitin as a reference gene. The differences in fold induction values were evaluated by statistical significance using Student's *t*-test with the software Assistat (Silva and Azevedo, 2016).

Results

Aphid-induced galls accumulate higher concentrations of monoterpenes as compared to non-colonized leaves

To better understand the mechanisms by which aphid-induced galls accumulate higher monoterpene concentrations than non-colonized leaves, we first analyzed the volatile terpene profiles of two individual trees and their corresponding aphid-induced galls. We have previously reported the mono- and sesquiterpene chemopolymorphism prevailing among *P. palaestina* individuals (Rand *et al.*, 2017). Since the enantiomeric distribution and chirality of the monoterpene components of *P. palaestina* tissues remained unknown, chiral chromatography was utilized to evaluate the extent of chemopolymorphism between the individual trees. The monoterpene concentrations in both trees were remarkably higher in galls relative to non-colonized leaves,

and their composition strongly reflected the monoterpene repertoire present in leaves (Table 1). Total monoterpene concentrations in galls of both trees increased 4- to 6-fold relative to leaves, while concentrations of individual monoterpenes increased from 1.5- to 76-fold. While some monoterpenes were detected in both trees, others were tree specific. Indeed, (+)-limonene was present in leaves and galls of the Carmel tree, but absent in Shechania tissues, which accumulated (–)-limonene (Table 1). Limonene, (+)-α-pinene, and myrcene were the main monoterpenes present in intact leaves of Carmel trees and their concentrations significantly increased in galls of the same tree. In addition, Carmel galls contained moderate (>2 μg g FW^{-1}) concentrations of (–)-α pinene, (+)-β-pinene, (–)-β-pinene, and (–)-limonene, the concentrations of which were significantly lower in leaves. In contrast, the monoterpene profile of Shechania leaves included substantial concentrations of β-ocimene, (–)-limonene, myrcene, (+)-α-pinene, and (–)-sabinene with other minor components (<1 μg g FW^{-1}).

In contrast to monoterpenes, sesquiterpenes were less abundant in both leaves and galls of both trees (Table 1). Sesquiterpenes were dominated by germacrene D and

Table 1. Distribution and content of volatile terpenes in leaves and *Baizongia pistaciae*-induced galls in two individual *P. palaestina*

Compound name	Carmel			Shechania			Product of gene
	Leaves	*Baizongia* galls	Fold increase in gall	Leaves	*Baizongia* galls	Fold increase in gall	
Monoterpenes (μg g^{-1} tissue)							
(+)-Limonene	187±2	669±98	3.6	ND	ND	ND	PpTPS1
(–)-Limonene	1.5±0.1	7±0.8	4.8	50.4±6	471±88	9.4	
(+)-α-Pinene	35±0.5	297±26	8.5	9.4±1.6	611±96	69	PpTPS1/ PpTPS4
(–)-α-Pinene	0.5±0.3	2.1±0.7	4	0.9±0.4	20.2±0.5	17	PpTPS3
Myrcene	20.6±1	56.4±3.8	2.7	10.3±6	786±356	76	
β-Ocimene	ND	ND	ND	320±53	480±50	1.5	
(–)-β-Pinene	1.8±0.5	9.7±2	6	0.4±0.2	33.4±7	76	PpTPS3
(+)-β-Pinene	1.7±0.3	5.4±0.2	1.5	0.2±0.1	7.7±0.3	45	
(–)-Sabinene	0.8±0.5	ND	*	8.5±2.5	99±34	12	PpTPS3
α-Thujene (±)	ND	ND	ND	0.9±0.7	12.6±2.1	12	PpTPS3
(–)-Camphene	ND	ND	ND	0.3±0.3	4.7±1.7	18	
(+)-δ-3-Carene	ND	0.8±0.8	*	ND	ND	ND	
α-Phellandrene (±)	ND	0.1±0.1	*	ND	6±3.5	*	
δ-2-Carene	ND	0.1±0.1	*	0.5±0.1	5±1.6	11	
γ-Terpinene	ND	ND	ND	0.4±0.1	3.85±1.2	10	PpTPS3
Total monoterpenes	**248.9**	**1077.6**	**4.3**	**402.2**	**2540.5**	**6.3**	
Sesquiterpenes (μg g^{-1} tissue)							
Germacrene D	103±1	13±5	0.12	16±11	9.3±3.4	0.6	PpTPS7
(–)-β-Caryophyllene	58±2	1.1±0.5	0.02	58±4	50±15	0.9	
(–)-α-Cubebene	0.5±0	0.03	0.06	2±2	10±4.6	5.0	
α-Humulene	3.5±0.1	0.08±0	0.02	7.3±2.7	7.4±1	1.0	
δ-Cadinene	2±0.09	0.3±0.09	0.15	2.7±0.9	5.4±1.3	2.0	PpTPS7
Uni sesq 119_49.1	0.3±0	ND	*	1.3±1	6.8±2.7	5.1	
Uni sesq 161_42.1	1.3±0	0.2±0.05	0.13	1.1±0.9	5±2.6	4.6	
β-Ylangene	4.3±0.1	0.5±0.2	0.12	0.6±0.4	0.2±0.1	0.3	
Uni sesq 161_45.1	3±0.05	0.26±0.1	0.09	0.5±0.4	0.24±0.13	0.5	
γ-Cadinene	2±0.07	0.2±0.06	0.10	2.0±0.9	2.1±0.6	1.0	PpTPS7
Uni sesq 161_44.2	1.0±0	0.14±0	0.14	1.0±1	3.2±1.7	3.0	
α-Copaene	1.4±0	0.11±0.1	0.08	0.8±0.5	2.6±1	3.4	
β-Copaene	3.0±0	0.3±0.1	0.10	0.4±0.3	0.15±0.1	0.4	
Germacrene B	1.6±0.1	ND	*	ND	ND	ND	PpTPS11
Aromadendrene	ND	ND	ND	ND	1.6±0.9	*	
Total sesquiterpenes	**182**	**16.2**	**0.09**	**93.7**	**104**	**1.1**	

Values represent means ±SE of three biological replicates, each consisting of 1–3 galls or leaves from the same tree. MTBE-extracted tissues were analyzed by GC-MS.
Uni sesq, unidentified sesquiterpene_main ion_RT.
* Fold increase not designated as concentrations in leaves were below detection limit. ND, not detected (limit of detection=0.01 μg g^{-1} tissue).
Compounds and genes present in both trees are shown in white; compounds and genes unique to Carmel are shown in blue; and compounds and genes unique to Shechania are shown in pink.

β-caryophyllene in both trees, but their concentrations were lower in galls of Carmel trees and did not significantly change in Shechania galls. In Shechania galls, (–)-cubebene, α-copaene, and δ-cadinene, as well as three out of four additional unidentified sesquiterpenes (Table 1) displayed higher concentrations relative to leaves. In conclusion, although individual trees displayed differences in the terpenoid compositions and concentrations, with some overlapping compounds, the monoterpene concentrations were significantly increased in galls relative to intact leaves in both trees while sesquiterpene concentrations displayed a more variable trend (Table 1).

Transcriptomic analysis identifies putative genes involved in terpenoid biosynthesis in P. palaestina *leaves and galls*

To examine the molecular and biochemical mechanisms responsible for the enhanced monoterpene accumulation in galls formed by *B. pistaciae* aphids, we generated transcriptomic datasets based on mRNA from leaves and galls of Carmel and Shechania trees and sequenced them on an Illumina Hiseq2000 platform using the paired-end approach with 100 bp read length. The throughput and quality of the RNA-seq data are shown in Supplementary Table S1. *De novo* transcriptome assembly resulted in 42 150 (Carmel) and 43 852 (Shechania) contigs representing genes expressed in gall and/or leaf tissues (Supplementary Table S2). Contig annotation was performed using public databases, including the NCBI nucleotide sequences (NT), NCBI non-redundant protein sequences (NR), Swiss-Prot, and the protein family annotation (PFAM). A high number of contigs (92.1% for Carmel and 90.8% for Shechania) were mapped to at least one of the databases. Supplementary Table S5 shows statistics of annotated contigs for each tree in different databases, while detailed gene annotations based on an NCBI Blast search are presented in Supplementary Table S3. Contigs encoding proteins with amino acid homology >98% were considered as the same gene, otherwise they were named differently (Fig. 1; Supplementary Table S3). This analysis revealed that ~10% of the contigs represent genes unique to the tree species in this study.

Structural genes of the cytosolic and plastidial terpenoid pathways in P. palaestina

Searches of annotated genes for core genes in the plastidial MEP and cytosolic MVA pathways, as well as for isoprenoid diphosphate PRTs and TPSs resulted in at least one full-length or nearly full-length (>95%) transcript for most genes involved in terpenoid biosynthesis (Fig. 1; Supplementary Table S3). In both trees, most proteins encoded by the same core genes in both pathways shared 98–100% amino acid identity. The genes encoding all six enzymes involved in the cytosolic MVA pathway were identified, namely acetoacetyl-CoA thiolase (AACT), 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA) synthase (HMGS), HMG-CoA reductase (HMGR), mevalonate kinase (MVK), phospho-mevalonate kinase (PMK), and diphosphomevalonate decarboxylase (MVD). Similar to Arabidopsis (Vranová *et al.*, 2013), transcriptomic analysis revealed two *AACT* genes encoding isoforms with 82% amino acid identity. One *HMGS* transcript was found in both trees, although the Shechania contig appeared to be truncated. Two *HMGR* gene homologs were detected in both trees with ~74% amino acid identity between the corresponding gene products. Single gene transcripts for *MVK*, *PMK*, and *MVD* were found in both trees, although the *MVD* contig in Carmel was truncated.

Similar to the MVA pathway, all seven genes of the plastidial MEP pathway were identified in the transcriptomic datasets (Fig. 1; Supplementary Table S3). They included 1-deoxy-D-xylulose 6-phosphate (DOXP) synthase (DXS), DOXP reductoisomerase (DXR), 2-C-methyl-D-erythritol cytidyltransferase (MCT), the latter found in Carmel only, 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase (CMK), 4-hydroxy-3-methyl-but-2-enyl diphosphate (HMB-PP) synthase (HDS), and HMB-PP reductase (HDR). Two distinct *DXS* genes were found (Supplementary Table S3; Supplementary Fig. S1) as well as two full-length IDI-like contigs (Supplementary Table S3; Supplementary Fig. S2).

Due to the unavailability of the *Pistacia* whole genome at the time, we performed *de novo* transcriptome assembly. Only recently, the genome of *Pistacia vera* L. was published (Zeng *et al.*, 2019). Still, most of the identified contigs in this study are in agreement and display high similarities to those present in the *P. vera* genome (Supplementary Table S3). Interestingly, other than *P. vera* homologs, most of the *P. palaestina* genes exhibited the highest similarity to citrus (Rutaceae) orthologs. Both Anacardiaceae and Rutaceae are in the order Sapindales.

Prenyltransferase genes

Our annotation analysis identified seven PRT-like genes in each individual tree, five of which (*PpPRT1–PpPRT5*) were similar to plant GPPSs/GGPPSs and two, *PpPRT6* and *PpPRT7*, shared similarity with FPPSs (Fig. 2A; Supplementary Table S3). Most PRT homologs display 98% amino acid identity between the two trees and thus were considered to be encoded by the same gene. To better predict the roles of the *Pistacia* PRTs in prenyldiphosphate biosynthesis, a phylogenetic analysis was performed using short-chain PRTs with known functions from other plants and the *Homo sapiens HsGGPPS1* as a distant ortholog (Fig. 2A; Supplementary Fig. S3). This analysis revealed that PpPRT1 and PpPRT3 cluster with bona fide GGPPS/LSU, some of which function as large GPPS subunits, as was shown in snapdragon (AmGGPPS.LSU) and tomato (LeGGPPS2) (Tholl *et al.*, 2004; Hivert *et al.*, 2020). This clade also includes catalytically inactive GPPS.LSU from mint (MpGPPS.LSU) that forms an active GPPS upon interaction with a small subunit (Burke *et al.*, 2004). PpPRT2 appeared

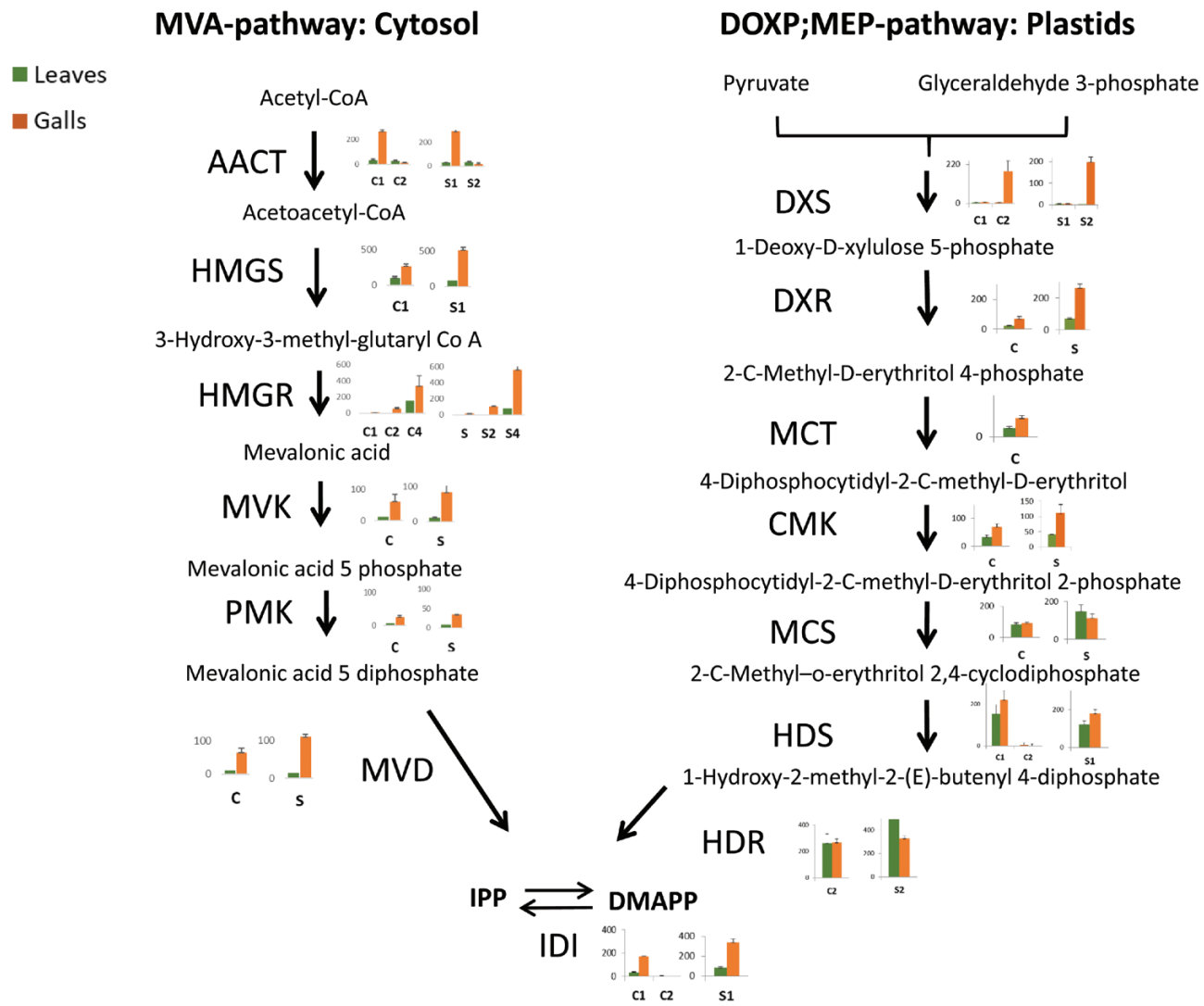

Fig. 1. Up-regulation of the terpene pathway structural genes leading to biosynthesis of volatile terpenoids in aphid-induced galls of *P. palaestina*. The cytosolic MVA pathway (left) is generally involved in providing IPP and DMAPP for the biosynthesis of sesqui- and triterpenes. The plastidial MEP (right, often called DOXP) pathway provides substrates for mono-, di-, and tetraterpene biosynthesis. Putative genes identified in *P. palaestina* transcriptomes of two individual wild trees (S, Shechania, C, Carmel) and their relative expression (FPKM values) in leaves (green bars) and aphid-induced galls (orange bars) are depicted next to the genes encoding the relevant enzymatic steps. Bars represent the mean of three replicates, except for Shechania galls with two valid replicates. Numbers represent the different isomorphs of the genes: (i.e. C1, Carmel isoform 1; S1, Shechania isoform 1). AACT, acetoacetyl-CoA thiolase; CDP-ME, 4-(cytidine 5′-diphospho)-2-C-methyl-D-erythritol; CDP-ME2P, 4-(cytidine 5′-diphospho)-2-C-methyl-D-erythritol phosphate; CMK, CDP-ME kinase; DMAPP, dimethylallyl diphosphate; DOXP, 1-deoxy-D-xylulose 5-phosphate; DXR, DOXP reductoisomerase; DXS, DOXP synthase; HDR, (E)-4-hydroxy-3-methylbut-2-enyl diphosphate reductase; HDS, (E)-4-hydroxy-3-methylbut-2-enyl diphosphate synthase; HMBPP, (E)-4-hydroxy-3-methylbut-2-enyl diphosphate; HMG-CoA, 3-hydroxy-3-methylglutaryl-CoA; HMGR, HMG-CoA reductase; HMGS, HMG-CoA synthase; IDI, isopentenyl diphosphate isomerase; IPP, isopentenyl diphosphate; MCT, 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase; MDS, 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase; ME-2,4cPP, 2-C-methyl-D-erythritol 2,4-cyclodiphosphate; MEP, 2-C-methyl-D-erythritol 4-phosphate; MVD, mevalonate diphosphate decarboxylase; MVK, mevalonate kinase; PMK, phosphomevalonate kinase. Details of the represented genes, annotation, and expression in individual samples, mean values ±SD, are listed in Supplementary Table S3.

to belong to the GPPS small subunit type I (SSUI) clade that contains catalytically inactive SSU of known heterodimeric GPPSs such as AmGPPS.SSU and MpGPPS.SSU, as well as *Phalaenopsis bellina* GPPS.SSU possessing autonomous GPPS activity (Burke *et al.*, 2004; Tholl *et al.*, 2004; Hsiao *et al.*, 2008; Orlova *et al.*, 2009). PpPRT4 clusters with small GPPS subunits from Arabidopsis and tomato, which together form a SSUII clade. Members of both clades were shown to function as 'modifiers' of the LSU products while the SSUII clade representatives in addition serve as 'accelerators' promoting the GPPS activity of the LSU (Wang and Dixon, 2009). GPPS. SSUs of the SSUI clade were predominantly identified in

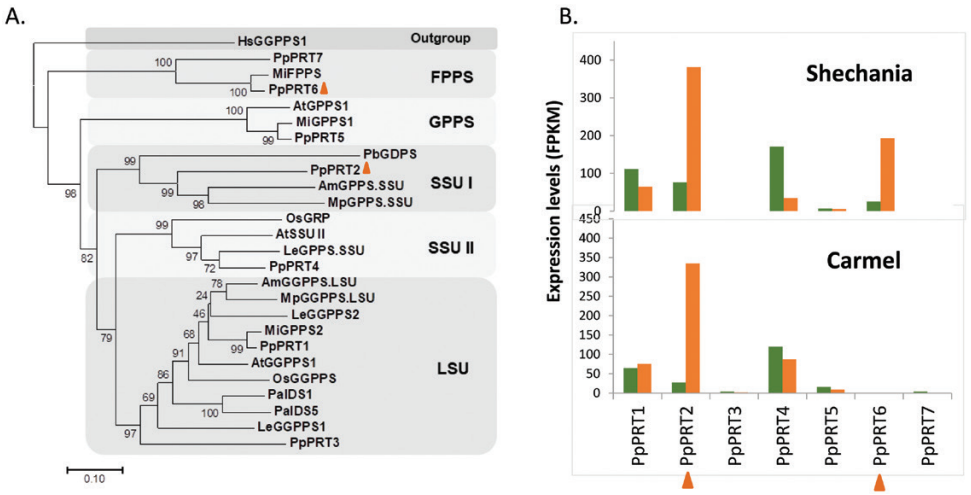

Fig. 2. Prenyltransferases in *P. palaestina*. (A) Phylogenetic tree of putative prenyltransferase genes identified in *P. palaestina* trees as compared with other known plant genes. The abbreviations of the protein sequences and their accession numbers are as follows: MpGPS.SSU (*M. piperita*, ABW86880), MpGPS.LSU (*M. piperita*, AF182828), AtGGPR (*A. thaliana*, AAG40013), AtGGPPS1 (*A. thaliana*, AEE86705), AtGPPS1 (*A. thaliana*, CAC16849), AmGPPS.SSU (*A. majus*, AAS82859), AmGPPS.LSU (*A. majus*, AY534687), LeGGPPS1 (*S. lycopersicum*, NM_001247158), LeGGPPS2 (*S. lycopersicum*, NM_001247373), MiFPPS (*M. indica*, AFJ53077), MiGPPS1 (*M. indica*, AFJ52721), MiGPPS2 (*M. indica*, AFJ52722), PbGDPS (*P. bellina*, ABV71395.1), PaIDS1 (*P. abies*, ACZ57571.1), PaIDS5 (*P. abies*, ACA21461.1), OsGGPPS (*O. sativa*, Os07g39270), OsGPR (*O. sativa*, Os02g44780), HsGGPPS1 (*H. sapiens*, AAH05252, outgroup). (B) Expression levels of putative prenyltransferase genes identified in *P. palaestina* leaves and aphid-induced galls. Green bars correspond to expression in leaves; orange bars indicate expression in galls. Orange triangles in both panels mark genes up-regulated in gall tissues. Bars represent the mean of three replicates ±SD except for two replicates for Schechania gall tissues.

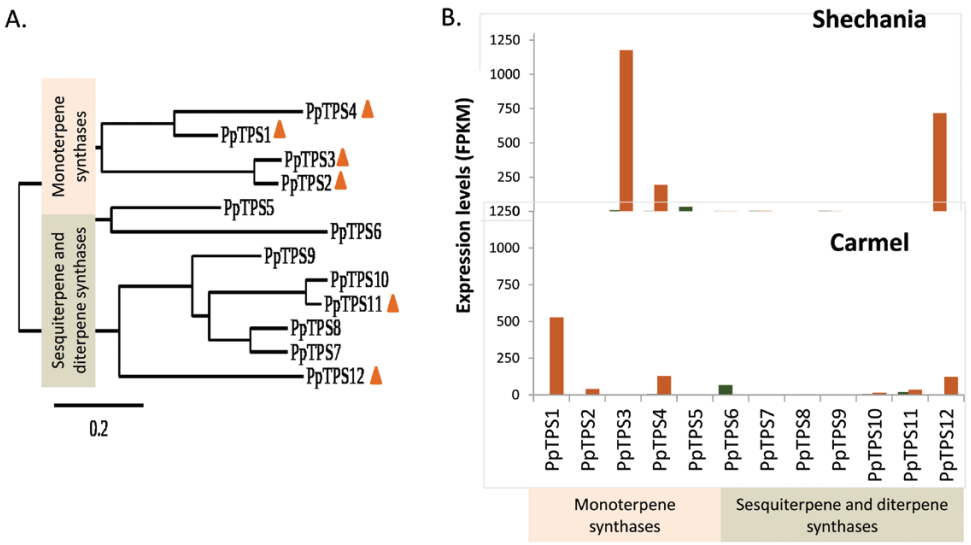

Fig. 3. Terpene synthases in *P. palaestina*. (A) Phylogenetic tree of putative terpene synthases identified by transcriptome analysis in two individual wild trees (Shechania and Carmel). Orange triangles mark genes up-regulated in gall tissues. (B) Relative expression values (FPKM) of PpTPSs in leaves (green bars) and galls (orange bars). Contigs sharing high similarity (98% amino acid identity) between the two trees were considered the same gene. Bars represent mean of three replicates ±SD except for two replicates for Schechania gall tissues.

monoterpene-producing plant species, in contrast to SSUII clade members which were less characterized but seem to have a wider expression (Wang and Dixon, 2009). PpPRT5 clustered with homomeric GPPSs including the Arabidopsis GPPS and mango GPPS1 (Bouvier *et al.*, 2000; Kulkarni *et al.*, 2013).

Interestingly, *PpPRT6* was only found in the transcriptome of the Shechania tree, while *PpPRT7* was detected only in the Carmel tree. Both genes belong to the FPPS clade and encode proteins with an amino acid similarity of 94% and 59%, respectively, to the mango MiFPPS (Kulkarni *et al.*, 2013). Clustal alignment of all PRTs found in the *Pistacia* transcriptome (Supplementary Fig. S3) revealed that, similarly to other SSUI clade members, PpPRT2 lacks two conserved Asp-rich motifs $DD(X)_{2-4}D$, important for prenyl-substrate binding (Koyama *et al.*, 1996; Wang and Ohnuma,

2000). In addition, two conserved CxxxC motifs, proposed to be involved in heteromeric interaction (Wang and Dixon, 2009), were identified in PpPRT1, PpPRT2, and PpPRT4, although in PpPRT2 the last cysteine was replaced by Ser, CxxxS, in the second Cys motif (Supplementary Fig. S3).

Terpene synthase (TPS) genes

A search for *P. palaestina* TPS genes resulted in a total of 12 contigs annotated as putative TPS genes. Similar to PRT homologs, *TPS* contigs encode proteins with 98% amino acid identity between the two trees. Phylogenetic analysis (Fig. 3A) and sequence alignment (Supplementary Fig. S4) reveal that all identified TPSs contain motifs typical for the TPS family members, including the conserved DDXXD motif involved in divalent cation binding (typically Mg^{2+} or Mn^{2+}) and, with the exception of PpTPS6 and PpTPS10, the RRX8W motif commonly found in cyclizing monoterpene synthases (Davis and Croteau, 2000). PpTPS1, PpTPS2, PpTPS3, PpTPS4, and PpTPS5 were annotated as monoterpene synthases (Supplementary Table S3) and predicted to contain a transit peptide (Supplementary Table S6). Like most angiosperm monoterpene synthases (Chen *et al.*, 2011) they belong to the TPS-b subfamily (Fig. 3A). PpTPS1 exhibited 64% amino acid identity with the *Zanthoxylum piperitum* β-phellandrene synthase (BBD88590.1). PpTPS2 and PpTPS3 shared, respectively, 58% and 59% amino acid identity with the *Camellia sinensis* (–)-alpha-terpineol synthase-like isoform X1 (XP_028062770.1). PpTPS4 showed 62% amino acid similarity with the *Citrus sinensis* limonene synthase, while PpTPS5 exhibited 63% amino acid identity with the *Quercus lobata* tricyclene synthase-like protein.

PpTPS6, PpTPS7, PpTPS8, PpTPS9, PpTPS10, and PpTPS11, exhibited 61–73% amino acid identity to the *Citrus sinensis* germacrene D synthase and α-farnesene synthase-like proteins, had no predicted transit peptide, and were therefore annotated as sesquiterpene synthases (Supplementary Tables S4, S5). PpTPS12 shared 71% amino acid identity with the *Ricinus communis* casbene synthase, a diterpene synthase induced upon fungal challenge, and contained an apparent plastidial transit peptide based on signal peptide predictor programs (Supplementary Table S6).

Three additional contigs corresponding to the oxidosqualene cyclase gene family (OSC1, OSC2, and OSC3), involved in the cyclization of triterpenes and sterol backbones such as cycloartenol and β-amyrin, were identified (Supplementary Table S3). These contigs were found in Carmel trees but were truncated in Shechania.

The galling habit leads to up-regulation of structural genes in both the plastidial and cytosolic terpenoid pathways in P. palaestina

The expression levels of key genes in both plastidial and cytosolic pathways were evaluated based on a number of counts corresponding to each contig (FPKM values) in galls and leaves of two trees (Fig. 1; Supplementary Table S3). Most of the MEP pathway genes leading to the formation of plastidial IPP and DMAPP for monoterpene biosynthesis were up-regulated in gall tissues relative to leaves (Fig. 1). The highest 240-fold increase in transcript levels was found for one of the two *DXS* genes (Fig. 1). DXS catalyzes the entry point and the probably important rate-limiting step of the MEP pathway (Phillips *et al.*, 2007; Vickers *et al.*, 2014; Hu *et al.*, 2017). In both trees, the expression of *DXR*, *MCT*, *CMK*, and *HDS* displayed a moderate 1.5- to 4-fold up-regulation in galls relative to leaves, while *MCS* and *HDR* expression remained unchanged (Fig. 1). In addition, *IDI* orthologs were 4- to 5-fold up-regulated in galls of both trees relative to leaves.

All genes of the MVA pathway, normally providing precursors for sesquiterpenes, triterpenes, and sterols, were also markedly up-regulated in galls, exhibiting a 2- to 44-fold increase. Transcripts of *HMGR*, coding for the putative rate-limiting enzyme of the MVA pathway (Chappell *et al.*, 1995; Vickers *et al.*, 2014), increased from 7- to 44-fold in galls. The expression of the gene encoding AACT, catalyzing the initial step of the MVA pathway, was also up-regulated 7-fold in galls relative to leaves.

Increases in expression of *PpPRT* genes were observed in galls of both trees, including *PpPRT2*, encoding a putative GPPS.SSU (see below) (Fig. 2B). In addition, expression of *PpPRT6* annotated as FPPS was up-regulated 7-fold in galls relative to leaves, but was found only in Shechania (Fig. 2B).

The transcript levels of putative monoterpene synthases showed 2- to 150-fold up-regulation in gall tissues (Fig. 3B). While expression of *PpTPS4* increased in galls of both trees, other TPSs displayed a tree-specific up-pregulation. *PpTPS1* and *PpTPS2* were uniquely induced in Carmel galls, while *PpTPS3* was only up-regulated in Shechania galls. Most putative sesquiterpene synthases displayed only minor transcript abundances and often their expression was down-regulated in galls relative to the corresponding leaves (Fig. 3B; Supplementary Table S3). Interestingly, transcript levels of *PpTPS12*, annotated as a putative diterpene synthase, were 123- to 650-fold up-regulated in galls. Significant increases in TPS transcripts were also verified by real-time PCR analysis performed on Carmel leaves and galls, which showed consistent results with RNA-seq (Supplementary Fig. S5).

Functional characterization of P. palaestina *prenyltransferases*

P. palaestina leaves and galls possess a GPPS.LSU-like enzyme not up-regulated in galls and generating GGPP *in vitro*

To better understand the biochemical platform acquired for enhanced monoterpene biosynthesis in galls (Table 1), we performed the functional characterization of PRTs that catalyze the formation of GPP, the immediate monoterpene precursor. Transcriptomic analyses revealed that out of seven putative PRTs, *PpPRT2* expression was increased by 5- to 12-fold

in galls relative to non-colonized leaves, as was expression of *PpPRT6*, which was found only in Shechania galls (Fig. 2B). PpPRT2 belongs to the GPPS.SSUI clade (Fig. 2A), thus we hypothesized that it may act as small subunit of GPPS, as it occurs in other plant heteromeric GPPSs. GPPS.SSUs are normally inactive by themselves and require interaction with a large subunit for their activity and GPP formation. From seven identified putative PRTs, two, PpPRT1 and PpPRT3, belong to the GGPPS/LSU clade (Fig. 2A) and only *PpPRT1* is expressed in gall and leaf tissues of both trees (Fig. 2B), suggesting that PpPRT1 could serve as a large subunit of *Pistacia* GPPS.

To test this hypothesis, PpPRT1 was produced in *E. coli* and the catalytic potential of the recombinant protein was analyzed. Recombinant PpPRT1 mainly produced GGPP from IPP and DMAPP *in vitro*, as was evidenced by the presence of the diterpenes β-springene, α-springene, and geranyllinalool (Fig. 4A), the products of GGPP acid hydrolysis (Hivert *et al.*, 2020). Additional hydrolysis products included low concentrations of limonene, (*E*)-β-ocimene, 3-carene, and terpineol, the acid hydrolysis products of GPP. Control assays with proteins from *E. coli* harboring an empty plasmid also revealed the presence of trace amounts of FPP and GPP probably as a result of minor endogenous PRT activities. Heat-inactivated protein and additional controls, including purified protein deprived of IPP and DMAPP, did not show any hydrolysis products.

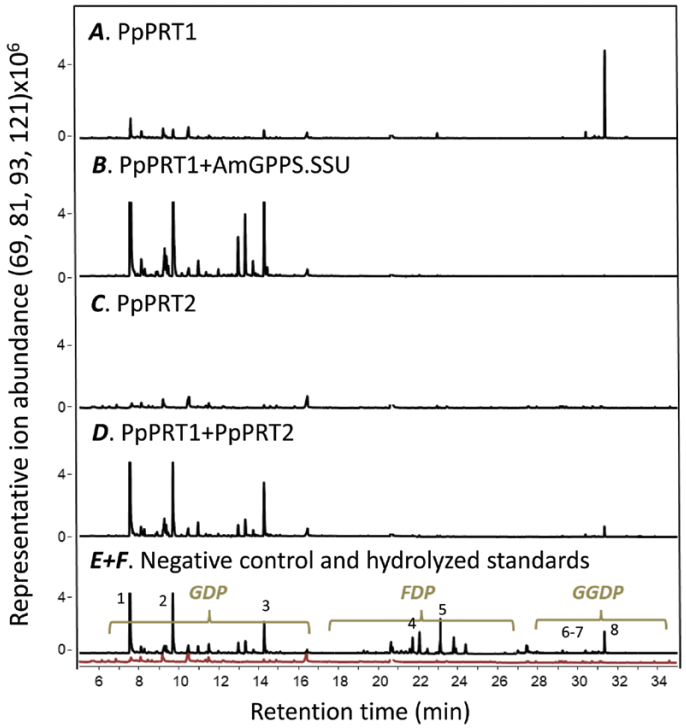

Fig. 4. Functional expression of *P. palaestina* prenyltransferases. (A) PpPTR1 (GGPPS.LSU homolog) alone. (B) Combined assay of PpPTR1 with *A. majus* GPPS.SSU (AmGPPS.SSU). (C) PpPTR2 (GGPPS.SSU homolog) alone. (D) Combined assay of PpPTR1 with PpPTR2. (E) Negative control reaction with protein derived from the expressed empty vector (red line). (F) GPP, FPP, and GGPP hydrolyzed standards. Prenyltransferases were expressed in *E. coli*, purified, and tested for their ability to form GGPP, FPP, and GPP from IPP and DMAPP as substrates. The products of the enzymatic assays (A–D) or standards (F) were acid hydrolyzed and analyzed by GC-MS. This procedure causes the isomerization of the GPP to several monoterpenes such as myrcene (1), linalool (2), and α-terpineol (3), FPP to β-farnesene (4) and bisabolene (5), and GGPP to β-springene (6), α-springene (7), and geranyl linalool (8), as indicated in the diagram. The results shown are typical of at least three separate replicates.

To determine whether PpPRT1 can interact with known GPPS.SSUs and change its product specificity from GGPP to GPP, PpPRT1 assays were performed in the presence of *Antirrhinum majus* AmGPPS.SSU (Fig. 4B) (Tholl *et al.*, 2004). Unlike PpPRT1 alone, a mixture of both proteins efficiently produced GPP as indicated by its hydrolysis products (Fig. 4B). Therefore, our functional analyses revealed that PpPRT1 is a bona fide GGPPS that provides GGPP for diterpene and carotenoid biosynthesis and it can modify its product specificity to favor GPP formation in the presence of a GPPS.SSU from a phylogenetically distant plant species. To test whether PpPRT2 could serve as a GPPS.SSU, the recombinant protein was produced in *E. coli* and incubated with IPP and DMAPP, which did not result in GPP formation (Fig 4C). However, inclusion of PpPRT1 in the assay mixtures containing PpPRT2 markedly favored GPP production (Fig. 4D) similar to assays harboring PpPRT1 and AmGPPS.SSU (Fig. 4B).

Tree-specific up-regulation of different TPS genes underlies the tree to tree differences in terpenoid volatile profiles

To investigate the molecular mechanisms leading to terpenoid polymorphism in trees (Table 1), we analyzed the biosynthetic capacity of the putative TPSs identified in the generated transcriptome datasets. *TPS* candidate genes were selected based on two criteria: (i) containing a complete ORF and (ii) expressed in both leaf and gall tissues, preferably with higher transcript abundance in galls. Therefore, we cloned the putative monoterpene synthase genes *PpTPS1*, *PpTPS2*, *PpTPS3*, and *PpTPS4*. We also aimed to characterize PpTPS12 that was highly up-regulated in galls as well as two sesquiterpene synthases (PpTPS7 and PpTPS11). After expression in *E. coli*, purified recombinant proteins were assayed using GPP, FPP, and GGPP as substrates to determine their potential mono-, sesqui-, and diterpene-forming activity (Fig. 5).

While we were unable to obtain active PpTPS2 protein, PpTPS1, PpTPS3, and PpTPS4 accepted GPP as substrate, generating specific monoterpenes (Fig. 5A–C). No discernible products were detected with either FPP or GGPP as substrates in lieu of GPP. PpTPS1, a protein encoded by a gene exclusively expressed in Carmel, catalyzed the formation of (+)-limonene as a major product, with low concentrations of (+)-α-pinene and myrcene (Fig. 5A). The observation that these compounds are the main monoterpenes present in intact leaves of Carmel trees, that drastically increased in galls,

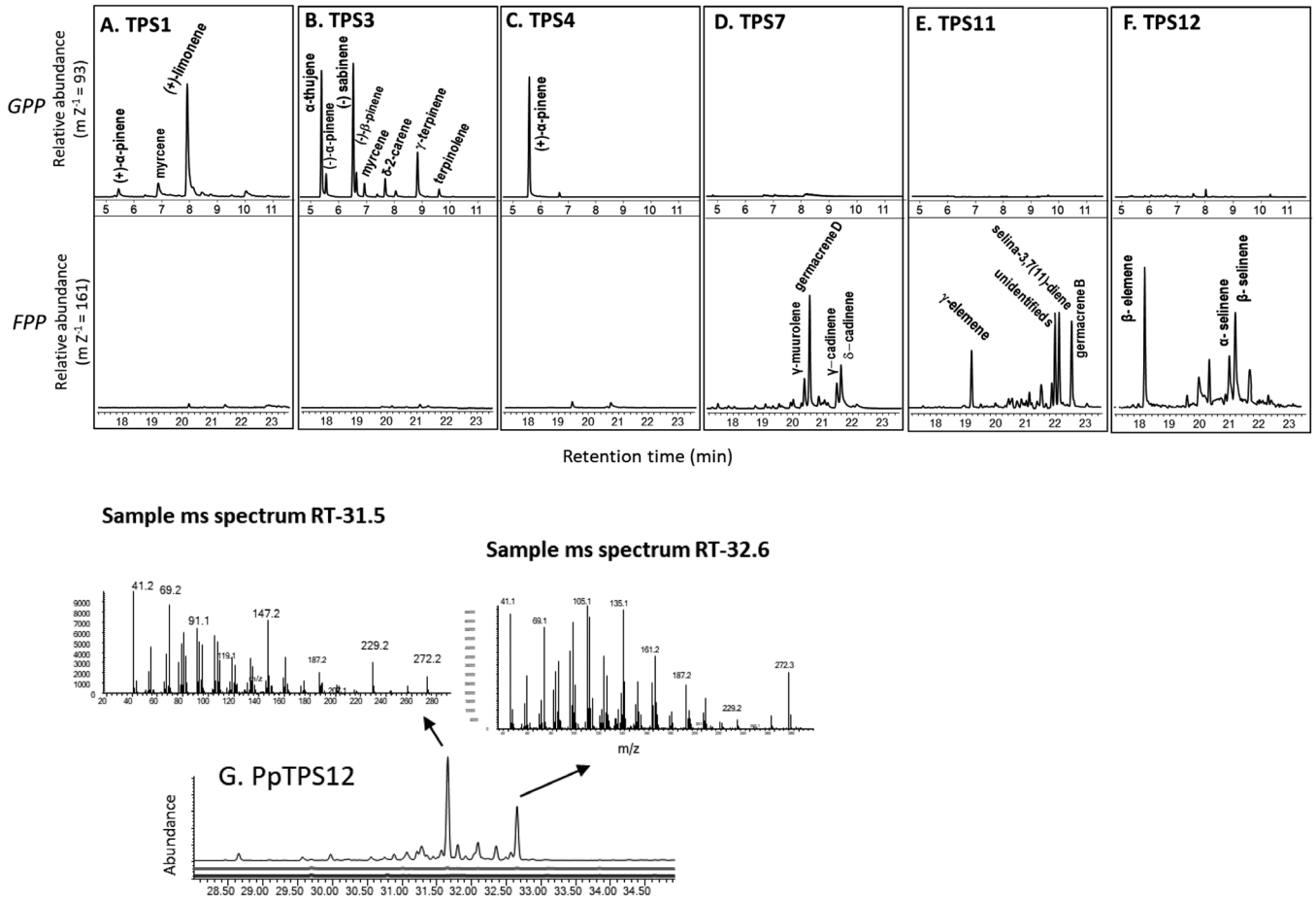

Fig. 5. Functional expression of *PpTPS* genes from *P. palaestina* leaves and galls. (A–F) SPME–GC separation of products synthesized from exogenously supplied geranyl diphosphate (top panels) or FPP (bottom panels) by purified proteins derived from *E. coli* cell lysates overexpressing the corresponding genes. (G) Mass spectra of two of the main diterpene products of PpTPS12 incubated with GGPP. Except for PpTPS12, no diterpene synthase activity was detected in any of the other assays. No enzymatic terpene products were detected in control assays either devoid of substrates, with empty vectors, or when heat-inactivated proteins were examined. Chromatogram abundance scales were adjusted to fit the figure.

combined with *PpTPS1* expression (Table 1; Fig. 3B), suggests that PpTPS1 is responsible for the monoterpene formation in this tree. PpTPS3, encoded by a gene found only in the Shechania transcriptome, formed an array of monoterpenes identified as α-thujene, (–)-α-pinene, (–)-sabinene, δ-2-carene, γ-terpinene, terpinolene, and myrcene (Fig. 5B). The same set of compounds was also identified in the metabolite profiles of Shechania leaves, with increased concentrations in galls (Table 1). PpTPS4, encoded by a gene expressed in both trees, exhibited monoterpene synthase activity generating (+)-α-pinene with traces of (+)-β-pinene (Fig. 5C). As both PpTPS1 and PpTPS4 form (+)-α-pinene, it is possible that PpTPS4 also contributes to (+)-α-pinene biosynthesis in Carmel tissues. If (+)-limonene, (+)-α-pinene, and myrcene are the products of only PpTPS1, then they would be increased in galls to the same degree. While (+)-limonene and myrcene amounts increased ~3-fold in Carmel galls versus leaves, the (+)-α-pinene concentrations were 8.5-fold higher in galls (Table 1), suggesting that both PpTPS1 and PpTPS4 contribute to monoterpene accumulation in Carmel trees.

Recombinant PpTPS7 and PpTPS11 produced multiple sesquiterpenes when incubated with FPP, but did not possess monoterpene synthase activity in the presence of GPP (Fig. 5D, E). PpTPS7 mainly generated germacrene D and lower concentrations of γ- and δ- cadinenes, that were also detected in leaves and in lower amounts in galls of both trees. PpTPS11 generated selina-3,7(11)-diene, germacrene B, and elemene; however, those compounds were not detected in the tissues (Table 1).

PpTPS12, a putative diterpene synthase, was cloned from Carmel gall cDNA, and functionally evaluated. Upon incubation with GGPP, PpTPS12 generated two unidentified diterpenes with retention times of 31.5 min and 32.6 min (Fig. 5G). PpTPS12 displayed no enzymatic activity with GPP as a substrate, but it catalyzed the formation of β-elemene, α-selinene, and β-selinene in the presence of FPP (Fig. 5F).

Discussion

Aphids enhance their defense ability by transcriptionally inducing monoterpene accumulation in galls

Many insect-induced galls accumulate plant specialized defense compounds that protect insects from biotic challenges without an apparent advantage to the host plant (Price *et al.*, 1987). Our results indicate that *B. pistaciae* aphids manipulate their hosts to boost the content of defensive monoterpene compounds in the galls they inhabit (Table 1). Galls display a marked up-regulation of key genes in both the plastidial and cytosolic terpenoid pathways (Figs 1–3). These observations complement previous studies showing enhanced monoterpene synthase enzymatic activity in galls (Rand *et al.*, 2017), as well as accumulation of higher monoterpene concentrations and changes in terpene compositions in galls relative to leaves (Rand *et al.*, 2014, 2017). Galls display an augmented monoterpene storage capacity in the form of specialized resin ducts that are absent in intact leaves (Martinez *et al.*, 2018). Aphids affect the host plant metabolism both locally within the gall tissues (Figs 1–3) and distally, by converting the galls into strong sinks for assimilates that are drawn to the galls from neighboring branches (Burstein *et al.*, 1994).

The molecular mechanisms enabling enhanced monoterpene accumulation in galls

Most genes of the MEP pathway, providing precursors for monoterpene biosynthesis, were up-regulated, albeit to a different degree, in galls of both trees (Fig. 1), indicating that the overall terpene biosynthetic capacity was enhanced in gall tissues, reinforcing elevated monoterpene production. *DXS* was one of the most up-regulated (200-fold) genes in galls. DXS catalyzes the first committed step and plays the major role in controlling carbon flux through the MEP pathway in plants (Wright *et al.*, 2014). In plants, DXSs comprise three discernible functional classes: class I and III DXSs participate in the biosynthesis of essential terpenoids including photosynthetic pigments and phytohormones, while type II DXSs are involved in the biosynthesis of specialized terpenoids, monoterpenes, and diterpenes (Phillips *et al.*, 2007; Zulak and Bohlmann, 2010; Tong *et al.*, 2015). *Pistacia DXS* genes up-regulated in gall tissues encode proteins that belong to class II DXSs (Supplementary Fig. S1), indicating an additional role for class II DXSs in monoterpene formation in the galls. These results further corroborate the pivotal role of DXS in regulating monoterpene biosynthesis in other systems such as in scented lily flowers (Hu *et al.*, 2017) and conifers (Zulak and Bohlmann, 2010). Other structural genes of the MEP pathway including *DXR*, *MCT*, *CMK*, and *HDR* were moderately up-regulated in galls displaying a 1.5- to 3-fold increase, while *MCS* and *HDR* expression remained unchanged as compared with leaves (Fig. 1).

IDI catalyzes the conversion of IPP to DMAPP and is one of the key enzymatic steps in the MEP pathway. Either one, two, or three molecules of IPP, the C_5 non-allylic substrate for terpene biosynthesis, are condensed with one molecule of the allylic DMAPP substrate to generate the C_{10}, C_{15}, and C_{20} prenylphosphate terpene precursors. While the molar ratio of DMAPP to IPP needed for monoterpene biosynthesis is 1:1, it decreases to 1:2 for sesqui- and triterpene biosynthesis, to 1:3 for di- and tetraterpene biosynthesis, and is even lower for the formation of long-chain prenols. Thus, the need for DMAPP is higher for monoterpene biosynthesis in comparison with larger terpenes. Therefore, IDIs have important roles in regulating the isomeric ratio of C_5 diphosphates used for biosynthesis of the major types of terpenes (Ramos-Valdivia *et al.*, 1997; Berthelot *et al.*, 2012; Zhou *et al.*, 2013) and the *IDI* up-regulation is often correlated with induced monoterpenoid metabolism. In both trees, IDI homologs possessed a predicted plastidial transit peptide (Supplementary Table S6) and their transcripts showed a 4-fold increase in gall tissues (Fig. 1), further supporting the role of IDI in enhanced terpenoid biosynthesis in *Pistacia* galls.

Monoterpene biosynthesis normally occurs in plastids (Rohmer, 1999; Lange *et al.*, 2000; Tholl, 2015) and relies on IPP and DMAPP derived from the plastidial MEP pathway. Therefore, the increase in expression of the structural genes in the MEP pathway in galls was consistent with elevated monoterpene concentrations in these structures (Table 1). However, a remarkable up-regulation of expression of genes of the cytosolic MVA pathway was also found in galls (Fig. 1). The MVA pathway generally provides precursors for sesqui- and triterpene biosynthesis in plants and other organisms. However, in contrast to monoterpenes, changes in sesquiterpene concentrations were found to be much less pronounced in aphid-induced galls as compared with leaves (Table 1; see also Rand *et al.*, 2014). In line with this, there was no observed up-regulation of sesquiterpene synthase genes in galls (Fig. 3B), suggesting that transcriptional up-regulation of the MVA pathway may provide additional precursors for monoterpene formation probably via IPP export from the cytosol to plastids. Although trafficking of IPP, and to a lesser extent of DMAPP, in and out of the plastids has been documented in many plant species (Bick and Lange, 2003; Dudareva *et al.*, 2005; Vranová *et al.*, 2013), it is still unknown if this process occurs in *Pistacia* or in other Anacardiaceae. Alternatively, the up-regulation of the cytosolic MVA pathway in galls could increase cytosolic IPP production for triterpene and sterol biosynthesis. Indeed, expression of *OSC1* and *OSC3* identified in the transcriptome of Carmel trees showed a 4- and 13-fold increase, respectively, in gall tissues (Supplementary Table S3), supporting this hypothesis. Although triterpene derivatives are prominent in *P. palaestina* (Caputo *et al.*, 1979) and other *Pistacia* spp. (Giner-Larza *et al.*, 2001), we do not have evidence indicating whether triterpene concentrations are elevated in gall tissues.

Prenyltransferase activity is transcriptionally altered in favor of GPP formation by specific induction of PpRT2

PRTs function at branch points of isoprenoid metabolism, determining the length of the formed prenyl diphosphate chains and controlling precursor flux into the different terpenoid classes (Wang and Dixon, 2009). Our generated transcriptome uncovered seven *PpPRT* genes that encode proteins belonging to each of the five distinct clades of *trans*-PRTs (Fig. 2A). Among them, only *PpPRT2* and *PpPRT6* showed significant up-regulation (5- to 12-fold) in galls relative to leaves, although the latter was found only in the Shechania tree (Fig. 2B; Supplementary Table S3).

The fact that up-regulation of *PpPRT2*, but not of *PpPRT1* encoding a bona fide GGPPS, occurs concomitantly with increased monoterpene content in galls *versus* leaves (Fig. 2B; Supplementary Table S3) suggests that PpPRT2 not only serves as a small subunit of heteromeric GPPS (Fig. 4D), but also plays a key regulatory role determining monoterpene concentrations by controlling GPP availability in galls. This is in accordance with other plant GPPS.SSUs and especially SSUI types (such as PpPRT2), which exhibit specific expression in monoterpene-rich tissues, such as the leaf glandular trichome of peppermint rich in menthol, and in snapdragon flowers rich in myrcene and ocimene (Burke *et al.*, 1999; Tholl *et al.*, 2004).

The molecular rationale for monoterpene chemodiversity in insect-induced galls

Our results show that gall-forming aphids induce the expression of many genes in the general upstream terpenoid biosynthetic pathways (Fig. 1) that provide precursors for monoterpene biosynthesis, and also induce PpPRT2 (Fig. 2B), a PRT directing precursor flux towards monoterpene formation. In addition, the galling habit results in the up-regulation of monoterpene synthases (Fig. 3B) responsible for the monoterpene diversity observed in *Pistacia* trees. Transcript levels of three TPS genes, *PpTPS1*, *PpTPS3*, and *PpTPS4*, encoding members of the TPS-b subfamily that encompasses most angiosperm monoterpene synthases (Chen *et al.*, 2011), were increased up to 130-fold in galls relative to non-colonized leaves, and their up-regulation, with the exception of *PpTPS4*, occurred in a tree-specific manner (Fig. 3B). PpTPS1 and PpTPS3 exhibit the formation of multiple products, while PpTPS4 produces only (+)-α-pinene (Fig. 5C). A positive correlation between expression of monoterpene synthases, the product specificities of the corresponding enzymes, and monoterpene accumulation, both in galls and in leaves of individual trees (Table 1), suggests that transcriptional regulation of a limited number of TPS genes encoding multiproduct enzymes is responsible for the chemopolymorphism displayed by individual *P. palaestina* trees. The observation that galls of individual trees contain basically the same monoterpenes as non-galled leaves, albeit at higher concentrations and often augmented by additional components (Table 1; Rand *et al.*, 2017) indicates that the monoterpene composition in the galls is determined by the metabolic capacity of the specific host. Insect manipulation through the galling habit causes enhanced formation of plant defensive compounds by capturing available mechanisms of the plant hosts normally utilized for plant self-defense, and recruiting them in turn for the insect's own benefit.

Gall terpene phenotypes are pre-determined by the existing genetic potential of individual plant hosts

Gall formation is an outcome of intimate and intricate interactions between a specific insect and its plant host. Gall morphological traits are clearly controlled by the inducing insects (e.g. Kurzfeld-Zexer *et al.*, 2015) and can be visibly recognized. The genetic variability within populations of *B. pistaciae* is limited (Ben-Shlomo and Inbar, 2012) and cannot be accounted for by the chemomorphological changes noted in galls. Moreover, the general gall morphology is clearly insect species specific, and it is often assumed that the host plant metabolic machinery, hijacked by the invading insects, provides some control of gall traits (e.g. Weis *et al.*, 1988; Giron *et al.*, 2016; Nabity, 2016; Schultz *et al.*, 2019). Our detailed comparison of gene expression patterns and monoterpene accumulation in two individual trees provides a rare opportunity to understand the influence and constraints of both the plant and the insects on gall traits. In both trees, the initial steps of both cytosolic and plastidial terpene biosynthesis are dramatically up-regulated in galls (Fig. 1). Additionally, *PpRT2*, encoding a PRT favoring the formation of GPP, and specific sets of TPS genes are up-regulated (Figs 2, 3). However, the two trees exhibit different expression of monoterpene synthase genes (Fig. 3B), which correlates with the distinct monoterpene profiles displayed by the hosts (Table 1). Therefore, the monoterpene profiles in galls are an augmented version of those found in leaves (Table 1), suggesting that the aphids colonizing these trees are able to utilize and up-regulate specific sets of the existing individual host-specific plant genes. We suggest that tissue reprograming leading to gall development and maintenance is modified by the ability of the aphids to up- or down-regulate the specifically available genetic resources of an individual host plant. Such host individuality may be associated with other gall traits and the ability of the plant to resist gall formation.

The ecological importance of chemodiversity in galls

Intraspecific chemodiversity is maintained by multiple environmental selective forces that probably confer on the plant marked ecological advantages (Moore *et al.*, 2014). The intraspecific chemodiversity in terpene content in *P. palaestina* is also manifested in the galls (Table 1; Fig. 3; Rand *et al.*, 2014, 2017), and it may also have an ecological benefit for the aphids. The divergent manipulation of host defense genes favors chemopolymorphism in the galls originating in different trees. This mechanism retains chemodiversity of the galls and may

further assist the aphids in coping with their own natural enemies. Our findings strongly support the 'gall defense hypothesis' by which the aphids manipulate their hosts to improve aphid ecological fitness without an apparent advantage to the hosts (Stone and Schönrogge, 2003; Inbar *et al.*, 2010; Rostás *et al.*, 2013).

Supplementary data

The following supplementary data are available at *JXB* online.

Fig. S1. Phylogenetic tree of putative 1-deoxy-D-xylulose 6-phosphate (DOXP) synthase (DXS) genes identified in two individual *P. palaestina* trees (S, Shechania; C, Carmel) as compared with other known genes from plants.

Fig. S2. Alignments of *P. palaestina* isopentenyl diphosphate isomerase (IDI) amino acid sequences.

Fig. S3. Alignments of *P. palaestina* isoprenyl diphosphate synthase (IDS) amino acid sequences with the homologous plant IDSs.

Fig. S4. Protein alignment of terpene synthase genes found in *P. palaestina* trees.

Fig. S5. Real-time PCR of selected terpene synthases from Carmel tree leaves and *Baizongia*-induced galls.

Table S1. Summary of RNA-seq data from six RNA libraries of galls and leaves of *Pistacia palaestina* trees from Carmel and Shechania.

Table S2. Length distribution of transcripts and unigenes.

Table S3. List of annotated genes in this study, their Blast annotation, and transcript abundance (FPKM) in all replicates.

Table S4. Primers used to amplify *Pistacia palaestina* PpTPS and the reference gene 'ubiquitin-conjugating enzyme' for RT–PCR analyses.

Table S5. Number of successful annotations for the unigenes from each tree against the different databases.

Table S6. Chloroplast signal peptide (Chlp) prediction according to amino acid sequences of characterized genes as predicted by three leading web tools.

Acknowledgements

This research was supported by The Israel Science Foundation (grant no. 940/08 to MI) and by research grant no. 2012241 from the BSF (United States–Israel Binational Science Foundation) to MI, ND, and EL. CHT received a 9 month SWE scholarship from the National Council for Scientific and Technological Development (CNPq, Brazil). We thank Rinat Guy for helpful technical assistance

Author contributions

RDR, ND, MI, and EL: conceptualization; EB, CHT, GH, KR, and JAM: methodology; RDR, EB, GH, XQH, JKM, VK, AA, JAM, and YS: investigation, data curation, formal analysis; MI, ND, and EL: funding acquisition; RDR, ND, MI, and EL: writing—review and editing; MI, ND and EL: project administration.

Data availability

The transcriptomic data have been deposited in NCBI under submission number 2464236. All data supporting the findings of this study are available within the paper and within its supplementary data published online.

References

Ben-Shlomo R, Inbar M. 2012. Patch size of gall-forming aphids: deme formation revisited. Population Ecology **54**, 135–144.

Berthelot K, Estevez Y, Deffieux A, Peruch F. 2012. Isopentenyl diphosphate isomerase: a checkpoint to isoprenoid biosynthesis. Biochimie **94**, 1621–1634.

Bick JA, Lange BM. 2003. Metabolic cross talk between cytosolic and plastidial pathways of isoprenoid biosynthesis: unidirectional transport of intermediates across the chloroplast envelope membrane. Archives of Biochemistry and Biophysics **415**, 146–154.

Bohlmann J, Meyer-Gauen G, Croteau R. 1998. Plant terpenoid synthases: molecular biology and phylogenetic analysis. Proceedings of the National Academy of Sciences, USA **95**, 4126–4133.

Bouvier F, Suire C, D'Harlingue A, Backhaus RA, Camara B. 2000. Molecular cloning of geranyl diphosphate synthase and compartmentation of monoterpene synthesis in plant cells. The Plant Journal **24**, 241–252.

Bradford MM. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein–dye binding. Analytical Biochemistry **72**, 248–254.

Burke C, Klettke K, Croteau R. 2004. Heteromeric geranyl diphosphate synthase from mint: construction of a functional fusion protein and inhibition by bisphosphonate substrate analogs. Archives of Biochemistry and Biophysics **422**, 52–60.

Burke CC, Wildung MR, Croteau R. 1999. Geranyl diphosphate synthase: cloning, expression, and characterization of this prenyltransferase as a heterodimer. Proceedings of the National Academy of Sciences, USA **96**, 13062–13067.

Burstein M, Wool D, Eshel A. 1994. Sink strength and clone size of sympatric, gall-forming aphids. European Journal of Entomology **91**, 57–61.

Caputo R, Mangoni L, Monaco P, Palumbo G. 1979. Triterpenes from the galls of *Pistacia palestina*. Phytochemistry **18**, 896–898.

Chappell J, Wolf F, Proulx J, Cuellar R, Saunders C. 1995. Is the reaction catalyzed by 3-hydroxy-3-methylglutaryl coenzyme A reductase a rate-limiting step for isoprenoid biosynthesis in plants? Plant Physiology **109**, 1337–1343.

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

Cornell HV. 1983. The secondary chemistry and complex morphology of galls formed by the Cynipinae (Hymenoptera): why and how? American Midland Naturalist **110**, 225–234.

Croteau R, Kutchan TM, Lewis NG, *et al.* 2000. Natural products (secondary metabolites). In: Buchanan B, Gruissem W, Jones R, eds. Biochemistry and molecular biology of plants. Rockville, MD: American Society of Plant Physiologists, 1250–1319.

Davidovich-Rikanati R, Sitrit Y, Tadmor Y, *et al.* 2007. Enrichment of tomato flavor by diversion of the early plastidial terpenoid pathway. Nature Biotechnology **25**, 899–901.

Davis EM, Croteau R. 2000. Cyclization enzymes in the biosynthesis of monoterpenes, sesquiterpenes, and diterpenes. In: Leeper FJ, Vederas JC, eds. Biosynthesis. Topics in Current Chemistry, vol 209. Berlin, Heidelberg: Springer, 53–95.

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

Dudareva N, Andersson S, Orlova I, Gatto N, Reichelt M, Rhodes D, Boland W, Gershenzon J. 2005. The nonmevalonate pathway supports

both monoterpene and sesquiterpene formation in snapdragon flowers. Proceedings of the National Academy of Sciences, USA **102**, 933–938.

Erbilgin N, Krokene P, Christiansen E, Zeneli G, Gershenzon J. 2006. Exogenous application of methyl jasmonate elicits defenses in Norway spruce (*Picea abies*) and reduces host colonization by the bark beetle *Ips typographus*. Oecologia **148**, 426–436.

Gershenzon J, Croteau R. 1992. Terpenoids. herbivores: their interactions with secondary plant metabolites, vol. 1: the chemical participants, 2nd edn. New York: Academic Press, 165–219.

Giner-Larza EM, Máñez S, Recio MC, Giner RM, Prieto JM, Cerdá-Nicolás M, Ríos JL. 2001. Oleanonic acid, a 3-oxotriterpene from *Pistacia*, inhibits leukotriene synthesis and has anti-inflammatory activity. European Journal of Pharmacology **428**, 137–143.

Giron D, Huguet E, Stone GN, Body M. 2016. Insect-induced effects on plants and possible effectors used by galling and leaf-mining insects to manipulate their host-plant. Journal of Insect Physiology **84**, 70–89.

Gonda I, Davidovich-Rikanati R, Bar E, *et al.* 2018. Differential metabolism of L-phenylalanine in the formation of aromatic volatiles in melon (*Cucumis melo* L.) fruit. Phytochemistry **148**, 122–131.

Gutensohn M, Orlova I, Nguyen TT, Davidovich-Rikanati R, Ferruzzi MG, Sitrit Y, Lewinsohn E, Pichersky E, Dudareva N. 2013. Cytosolic monoterpene biosynthesis is supported by plastid-generated geranyl diphosphate substrate in transgenic tomato fruits. The Plant Journal **75**, 351–363.

Haas BJ, Papanicolaou A, Yassour M, *et al.* 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nature Protocols **8**, 1494–1512.

Hartley SE. 1998. The chemical composition of plant galls: are levels of nutrients and secondary compounds controlled by the gall-former? Oecologia **113**, 492–501.

Heil M. 2014. Herbivore-induced plant volatiles: targets, perception and unanswered questions. New Phytologist **204**, 297–306.

Hivert G, Davidovich-Rikanati R, Bar E, Sitrit Y, Schaffer A, Dudareva N, Lewinsohn E. 2020. Prenyltransferases catalyzing geranyldiphosphate formation in tomato fruit. Plant Science **296**, 110504.

Hsiao YY, Jeng MF, Tsai WC, Chuang YC, Li CY, Wu TS, Kuoh CS, Chen WH, Chen HH. 2008. A novel homodimeric geranyl diphosphate synthase from the orchid *Phalaenopsis bellina* lacking a DD(X)2-4D motif. The Plant Journal **55**, 719–733.

Hu Z, Tang B, Wu Q, Zheng J, Leng P, Zhang K. 2017. Transcriptome sequencing analysis reveals a difference in monoterpene biosynthesis between scented *Lilium* 'Siberia' and unscented *Lilium* 'Novano'. Frontiers in Plant Science **8**, 1351.

Inbar M, Izhaki I, Koplovich A, Lupo I, Silanikove N, Glasser T, Gerchman Y, Perevolotsky A, Lev-Yadun S. 2010. Why do many galls have conspicuous colors? A new hypothesis. Arthropod-Plant Interactions **4**, 1–6.

Inbar M, Wink M, Wool D. 2004. The evolution of host plant manipulation by insects: molecular and ecological evidence from gall-forming aphids on *Pistacia*. Molecular Phylogenetics and Evolution **32**, 504–511.

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

Kessler A, Baldwin IT. 2001. Defensive function of herbivore-induced plant volatile emissions in nature. Science **291**, 2141–2144.

Koyama T, Tajima M, Sano H, Doi T, Koike-Takeshita A, Obata S, Nishino T, Ogura K. 1996. Identification of significant residues in the substrate binding site of *Bacillus stearothermophilus* farnesyl diphosphate synthase. Biochemistry **35**, 9533–9538.

Kulkarni R, Pandit S, Chidley H, Nagel R, Schmidt A, Gershenzon J, Pujari K, Giri A, Gupta V. 2013. Characterization of three novel isoprenyl diphosphate synthases from the terpenoid rich mango fruit. Plant Physiology and Biochemistry **71**, 121–131.

Kurzfeld-Zexer L, Lev-Yadun S, Inbar M. 2015. One aphid species induces three gall types on a single plant: comparative histology of one genotype and multiple extended phenotypes. Flora - Morphology, Distribution, Functional Ecology of Plants **210**, 19–30.

Lange BM, Rujan T, Martin W, Croteau R. 2000. Isoprenoid biosynthesis: the evolution of two ancient and distinct pathways across genomes. Proceedings of the National Academy of Sciences, USA **97**, 13172–13177.

Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. Methods **25**, 402–408.

Martin DM, Fäldt J, Bohlmann J. 2004. Functional characterization of nine Norway spruce TPS genes and evolution of gymnosperm terpene synthases of the TPS-d subfamily. Plant Physiology **135**, 1908–1927.

Martinez JI, Moreno-González V, Jonas-Levi A, Álvarez R. 2018. Quantitative differences detected in the histology of galls induced by the same aphid species in different varieties of the same host. Plant Biology **20**, 516–524.

Moore BD, Andrew RL, Külheim C, Foley WJ. 2014. Explaining intraspecific diversity in plant secondary metabolites in an ecological context. New Phytologist **201**, 733–750.

Nabity PD. 2016. Insect-induced plant phenotypes: revealing mechanisms through comparative genomics of galling insects and their hosts. American Journal of Botany **103**, 979–981.

Nabity PD, Haus MJ, Berenbaum MR, DeLucia EH. 2013. Leaf-galling phylloxera on grapes reprograms host metabolism and morphology. Proceedings of the National Academy of Sciences, USA **110**, 16663–16668.

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

Nyman T, Julkunen-Tiitto R. 2000. Manipulation of the phenolic chemistry of willows by gall-inducing sawflies. Proceedings of the National Academy of Sciences, USA **97**, 13184–13187.

Orlova I, Nagegowda DA, Kish CM, *et al.* 2009. The small subunit of snapdragon geranyl diphosphate synthase modifies the chain length specificity of tobacco geranylgeranyl diphosphate synthase in planta. The Plant Cell **21**, 4002–4017.

Pazouki L, Memari HR, Kännaste A, Bichele R, Niinemets Ü. 2015. Germacrene A synthase in yarrow (*Achillea millefolium*) is an enzyme with mixed substrate specificity: gene cloning, functional characterization and expression analysis. Frontiers in Plant Science **6**, 111.

Phillips MA, Walter MH, Ralph SG, *et al.* 2007. Functional identification and differential expression of 1-deoxy-D-xylulose 5-phosphate synthase in induced terpenoid resin formation of Norway spruce (*Picea abies*). Plant Molecular Biology **65**, 243–257.

Price PF, Fernandes GW, Waring GL. 1987. Adaptive nature of insect galls. Environmental Entomology **16**, 15–24.

Rai A, Smita SS, Singh AK, Shanker K, Nagegowda DA. 2013. Heteromeric and homomeric geranyl diphosphate synthases from *Catharanthus roseus* and their role in monoterpene indole alkaloid biosynthesis. Molecular Plant **6**, 1531–1549.

Ramos-Valdivia AC, van der Heijden R, Verpoorte R. 1997. Isopentenyl diphosphate isomerase: a core enzyme in isoprenoid biosynthesis. A review of its biochemistry and function. Natural Product Reports **14**, 591–603.

Rand K, Bar E, Ari MB, Davidovich-Rikanati R, Dudareva N, Inbar M, Lewinsohn E. 2017. Differences in monoterpene biosynthesis and accumulation in *Pistacia palaestina* leaves and aphid-induced galls. Journal of Chemical Ecology **43**, 143–152.

Rand K, Bar E, Ben-Ari M, Lewinsohn E, Inbar M. 2014. The mono- and sesquiterpene content of aphid-induced galls on *Pistacia palaestina* is not a simple reflection of their composition in intact leaves. Journal of Chemical Ecology **40**, 632–642.

Rohmer M. 1999. The discovery of a mevalonate-independent pathway for isoprenoid biosynthesis in bacteria, algae and higher plants. Natural Product Reports **16**, 565–574.

Rostás M, Maag D, Ikegami M, Inbar M. 2013. Gall volatiles defend aphids against a browsing mammal. BMC Evolutionary Biology **13**, 193.

Schmidt A, Gershenzon J. 2007. Cloning and characterization of isoprenyl diphosphate synthases with farnesyl diphosphate and geranylgeranyl diphosphate synthase activity from Norway spruce (*Picea abies*) and their relation to induced oleoresin formation. Phytochemistry **68**, 2649–2659.

Schmidt A, Wächtler B, Temp U, Krekling T, Séguin A, Gershenzon J. 2010. A bifunctional geranyl and geranylgeranyl diphosphate synthase is involved in terpene oleoresin formation in *Picea abies*. Plant Physiology **152**, 639–655.

Schultz JC, Edger PP, Body MJA, Appel HM. 2019. A galling insect activates plant reproductive programs during gall development. Scientific Reports **9**, 1833.

Silva F, Azevedo CA. 2016. The Assistat Software Version 7.7 and its use in the analysis of experimental data. African Journal of Agricultural Research **11**, 3733–3740.

Stone GN, Schönrogge K. 2003. The adaptive significance of insect gall morphology. Trends in Ecology & Evolution **18**, 512–522.

Tholl D. 2015. Biosynthesis and biological functions of terpenoids in plants. Advances in Biochemical Engineering/Biotechnology **148**, 63–106.

Tholl D, Kish CM, Orlova I, Sherman D, Gershenzon J, Pichersky E, Dudareva N. 2004. Formation of monoterpenes in *Antirrhinum majus* and *Clarkia breweri* flowers involves heterodimeric geranyl diphosphate synthases. The Plant Cell **16**, 977–992.

Tong Y, Su P, Zhao Y, Zhang M, Wang X, Liu Y, Zhang X, Gao W, Huang L. 2015. Molecular cloning and characterization of DXS and DXR genes in the terpenoid biosynthetic pathway of *Tripterygium wilfordii*. International Journal of Molecular Sciences **16**, 25516–25535.

van Schie CC, Ament K, Schmidt A, Lange T, Haring MA, Schuurink RC. 2007. Geranyl diphosphate synthase is required for biosynthesis of gibberellins. The Plant Journal **52**, 752–762.

Vickers CE, Bongers M, Liu Q, Delatte T, Bouwmeester H. 2014. Metabolic engineering of volatile isoprenoids in plants and microbes. Plant, Cell & Environment **37**, 1753–1775.

Vranová E, Coman D, Gruissem W. 2013. Network analysis of the MVA and MEP pathways for isoprenoid synthesis. Annual Review of Plant Biology **64**, 665–700.

Wang G, Dixon RA. 2009. Heterodimeric geranyl(geranyl)diphosphate synthase from hop (*Humulus lupulus*) and the evolution of monoterpene biosynthesis. Proceedings of the National Academy of Sciences, USA **106**, 9914–9919.

Wang KC, Ohnuma S. 2000. Isoprenyl diphosphate synthases. Biochimica et Biophysica Acta **1529**, 33–48.

Weis AE, Walton R, Crego CL. 1988. Reactive plant tissue sites and the population biology of gall makers. Annual Review of Entomology **33**, 467–486.

Wittstock U, Gershenzon J. 2002. Constitutive plant toxins and their role in defense against herbivores and pathogens. Current Opinion in Plant Biology **5**, 300–307.

Wool D. 2012. Autecology of *Baizongia pistaciae* (L.): a monographical study of a galling aphid. Israel Journal of Entomology **41–42**, 67–93.

Wright LP, Rohwer JM, Ghirardo A, Hammerbacher A, Ortiz-Alcaide M, Raguschke B, Schnitzler JP, Gershenzon J, Phillips MA. 2014. Deoxyxylulose 5-phosphate synthase controls flux through the methylerythritol 4-phosphate pathway in Arabidopsis. Plant Physiology **165**, 1488–1504.

Zeng L, Tu XL, Dai H, *et al.* 2019. Whole genomes and transcriptomes reveal adaptation and domestication of pistachio. Genome Biology **20**, 79.

Zhou Y, Nambou K, Wei L, Cao J, Imanaka T, Hua Q. 2013. Lycopene production in recombinant strains of *Escherichia coli* is improved by knockout of the central carbon metabolism gene coding for glucose-6-phosphate dehydrogenase. Biotechnology Letters **35**, 2137–2145.

Zulak KG, Bohlmann J. 2010. Terpenoid biosynthesis and specialized vascular cells of conifer defense. Journal of Integrative Plant Biology **52**, 86–97.