



# Germacrene A synthase in yarrow (*Achillea millefolium*) is an enzyme with mixed substrate specificity: gene cloning, functional characterization and expression analysis

Leila Pazouki<sup>1\*</sup>, Hamid R. Memari<sup>2</sup>, Astrid Kännaste<sup>1</sup>, Rudolf Bichele<sup>3</sup> and Ülo Niinemets<sup>1,4</sup>

<sup>1</sup> Institute of Agricultural and Environmental Sciences, Estonian University of Life Sciences, Tartu, Estonia

<sup>2</sup> Biotechnology and Life Science Center and School of Agriculture, Shahid Chamran University, Ahvaz, Iran

<sup>3</sup> Molecular Pathology, Institute of Biomedicine and Translational Medicine, University of Tartu, Tartu, Estonia

<sup>4</sup> Estonian Academy of Sciences, Tallinn, Estonia

## Edited by:

Ikuro Abe, The University of Tokyo, Japan

## Reviewed by:

Dae-Kyun Ro, University of Calgary, Canada

Joong-Hoon Ahn, Konkuk University, South Korea

Tetsuo Kushi, Meiji University, Japan

Hirofumi Morita, The University of Tokyo, Japan

## \*Correspondence:

Leila Pazouki, Institute of Agricultural and Environmental Sciences, Estonian University of Life Sciences, Fr.R. Kreutzwaldi 5, EE-51014 Tartu, Estonia  
e-mail: leila.pazouki@emu.ee

Terpenoid synthases constitute a highly diverse gene family producing a wide range of cyclic and acyclic molecules consisting of isoprene (C<sub>5</sub>) residues. Often a single terpene synthase produces a spectrum of molecules of given chain length, but some terpene synthases can use multiple substrates, producing products of different chain length. Only a few such enzymes has been characterized, but the capacity for multiple-substrate use can be more widespread than previously thought. Here we focused on germacrene A synthase (GAS) that is a key cytosolic enzyme in the sesquiterpene lactone biosynthesis pathway in the important medicinal plant *Achillea millefolium* (*AmGAS*). The full length encoding gene was heterologously expressed in *Escherichia coli* BL21 (DE3), functionally characterized, and its *in vivo* expression was analyzed. The recombinant protein catalyzed formation of germacrene A with the C<sub>15</sub> substrate farnesyl diphosphate (FDP), while acyclic monoterpenes were formed with the C<sub>10</sub> substrate geranyl diphosphate (GDP) and cyclic monoterpenes with the C<sub>10</sub> substrate neryl diphosphate (NDP). Although monoterpene synthesis has been assumed to be confined exclusively to plastids, *AmGAS* can potentially synthesize monoterpenes in cytosol when GDP or NDP become available. *AmGAS* enzyme had high homology with GAS sequences from other Asteraceae species, suggesting that multi-substrate use can be more widespread among germacrene A synthases than previously thought. Expression studies indicated that *AmGAS* was expressed in both autotrophic and heterotrophic plant compartments with the highest expression levels in leaves and flowers. To our knowledge, this is the first report on the cloning and characterization of germacrene A synthase coding gene in *A. millefolium*, and multi-substrate use of GAS enzymes.

**Keywords:** *Achillea millefolium*, cytosolic terpene synthesis, enzyme assay, gene expression, germacrene A synthase gene, mixed substrate specificity, monoterpenes, sesquiterpenes

## INTRODUCTION

A large variety of volatile organic compounds (VOCs) are synthesized and released into the environment by plants (Pichersky and Gershenzon, 2002). Although VOCs include a wide range of hydrocarbons and oxygenated hydrocarbons, terpenoids consisting of isoprene, monoterpenes and sesquiterpenes constitute the largest class of VOCs in ambient atmosphere (Guenther et al., 1995, 2000; Fineschi et al., 2013). Overall, over 60,000 terpenes and derivatives are found in nature (Cheng et al., 2007; Bohlmann and Keeling, 2008). Terpenoids are synthesized by a variety of terpenoid synthases that are characterized by variation in substrate and product specificity and expression level in different tissues (Christianson, 2006, 2008; Cheng et al., 2007; Bohlmann and Keeling, 2008; Nagegowda, 2010; Rajabi et al., 2013). During recent decades, there has been major progress in identification and functional characterization of volatile terpenoid biosynthesis genes, enzymes and in metabolic engineering of terpenoid

synthesis, and this has contributed greatly to improved understanding of basic mechanisms and variability of terpenoid biosynthesis (Keeling and Bohlmann, 2006; Bohlmann and Keeling, 2008; Degenhardt et al., 2009; Nagegowda, 2010; Chen et al., 2011; Rajabi et al., 2013). However, we still lack information of gene structure, expression regulation and catalysis mechanisms for a large number of biologically and economically important terpenoid synthases.

Sesquiterpenes are synthesized by sesquiterpene synthases and play a variety of ecological roles in higher plants. Many sesquiterpenes are volatile compounds that are commonly emitted from flowers serving as attractants to pollinators (Morse et al., 2012), but also as repellents against nectar thieves (Junker and Bluethgen, 2008). In addition, sesquiterpene emissions from leaves of several plant species play important roles in direct and indirect chemical defense against pathogens and herbivores (Schnee et al., 2002; Cheng et al., 2007; Chappell and Coates,

2010). They can serve both as repellents (Huang et al., 2012; Scala et al., 2013) or as attractants of herbivore predators and parasitoids (Schnee et al., 2002). Sesquiterpenes are also synthesized and accumulated in underground organs like rhizomes and roots (De Kraker et al., 1998; Kovacevic et al., 2002; Rasmann et al., 2005) where they participate in attracting nematode predators (Rasmann et al., 2005).

Sesquiterpenes, including germacrenes, are particularly abundant in the Asteraceae family. In several species belonging to Asteraceae, germacrenes fulfill a central role in the formation of different sesquiterpene derivatives, in particular, sesquiterpene lactones (Adio, 2009a). Sesquiterpene lactones exhibit important pharmacological, physiological and ecological features. For example, artemisinin is an antimalarial sesquiterpene lactone produced by *Artemisia annua* (Ro et al., 2006; Keasling, 2012; Paddon et al., 2013). Production of this important pharmaceutical has recently been commercialized in heterologous systems (Ro et al., 2006; Keasling, 2012; Paddon et al., 2013). Sesquiterpene lactones also have antimigraine, antifungal and antibacterial properties and can protect against pests and herbivores (Picman, 1986). Recently, some important biofuels have been developed from sesquiterpene derivatives (McAndrew et al., 2011).

Among Asteraceae, the genus *Achillea* contains over 100 herbaceous species spread throughout the northern hemisphere. The aerial parts of species from this genus are widely used in herbal medicine for preparation of infusion with antiphlogistic and spasmolytic activity (Nemeth and Bernath, 2008). Different groups of sesquiterpene lactones have been reported from this genus, eudesmanolides, and guaianolides being the most common (Si et al., 2006). Aerial parts of *Achillea millefolium* L., one of the most wide-spread and important medicinal species, have long been used as a drug in traditional and modern medicine and in herbal teas, curing inflammation and gastrointestinal spasms (Chandler et al., 1982). Sesquiterpene lactones have been identified as major compounds in *A. millefolium* (Montsko et al., 2008) and a number of germacranolides and guaianolides has already been identified in this species (Glasl et al., 2002). Some other sesquiterpene lactones such as 8- $\alpha$ -angeloxy-artabsin, 8- $\alpha$ -tigloxy-artabsin, 8- $\alpha$ -angeloxy-3-oxa-artabsin, 8- $\alpha$ -tigloxy-3-oxa-artabsin, 8-desacetyl-matricarin and santolin have also been detected in *A. millefolium* by LC-MS (Montsko et al., 2008).

Sesquiterpene lactones are derived from sesquiterpene (+)-germacrene A in many plant species, including Asteraceae (De Kraker et al., 2001). Among germacrene A derived lactones, there are a number of pharmaceutically important compounds such as parthenolide in feverfew (*Tanacetum parthenium*) (Majdi et al., 2011), tenulin (yellow sneezeweed, *Helenium amarum*) and helentalin (sneezeweed, *Helenium autumnale*) (Bouwmeester et al., 2002). Furthermore, germacrene A itself and in particular its rearrangement product  $\beta$ -elemene have been shown to possess anticancer activity (Adio, 2009b).

Germacrene A is formed from farnesyl diphosphate (FDP) by germacrene A synthase (GAS) (De Kraker et al., 1998). The gene structure of GAS and the enzyme functional activity have been studied only in a few species (De Kraker et al., 1998; Bouwmeester et al., 2002; Majdi et al., 2011), and there is thus, limited information on biological variation in sequence structure, expression

and catalysis. Furthermore, there is overall limited information on key sesquiterpene synthases involved in physiological processes, in particular, on factors determining the substrate profiles of these enzymes. Recently, it has been demonstrated that some sesquiterpene synthases can catalyze both formation of sesquiterpenes with C15 substrate and monoterpenes with C10 substrate (Davidovich-Rikanati et al., 2008; Gutensohn et al., 2013; Rajabi et al., 2013), but it is unclear how general this finding is. The synthesis of hemiterpenes (C5), monoterpenes (C10) and diterpenes (C20) has been thought to occur in plastids, while that of sesquiterpenes (C15) and triterpenes (C30) to occur in cytosol (Cheng et al., 2007; Davidovich-Rikanati et al., 2008; Gutensohn et al., 2013; Rajabi et al., 2013). However, recent evidence suggests that multiple-substrate sesquiterpene synthases can catalyze monoterpene formation in cytosol (Davidovich-Rikanati et al., 2008; Gutensohn et al., 2013), providing a hugely exciting way of regulation of compound profiles, sesqui- vs. monoterpenes, by alterations in cytosolic pool sizes of different substrates. Alteration of product profiles as the result of substrate changes can have important consequences for terpenoid accumulation in aromatic species lacking specialized storage structures. Use of multiple substrates in functional characterization of terpenoid synthases is by far not a routine procedure (Davidovich-Rikanati et al., 2008; Gutensohn et al., 2013; Rajabi et al., 2013), and there is, as yet, no evidence of monoterpene synthase activity for GAS enzymes.

To gain insight into terpenoid synthesis in *A. millefolium* and its regulation, the objectives of this study were molecular identification and functional characterization of germacrene A synthase in *A. millefolium* and quantification of germacrene A synthase gene expression in different tissues. The results of this study demonstrate that *A. millefolium* GAS enzyme is a multi-substrate enzyme catalyzing formation of germacrene A, but also acyclic and cyclic monoterpenes depending on the substrate available.

## MATERIALS AND METHODS

### PLANT MATERIAL

Field-grown yarrow (*A. millefolium*) plants of local genotype (Tartu, Estonia, 58°23'N, 27°05'E) were transplanted in clay pots of 3 L and grown under controlled conditions in a growth chamber (16 h day length and day/night temperature of 25/18°C, incident quantum flux density of 400  $\mu\text{mol m}^{-2} \text{s}^{-1}$ ). Flowers, leaves, roots, rhizomes and stems were collected and immediately frozen in liquid nitrogen and stored at  $-80^\circ\text{C}$  for gene expression analysis (three biological replicates for each tissue were used). Fresh yarrow flowers (4 g dry weight) and leaves (7 g dry weight) were harvested for the analysis of volatiles from the field in August 2013.

Germacrene A has been previously found to accumulate in chicory (*Cichorium intybus* L.) roots (De Kraker et al., 1998). Due to lack of germacrene A as a reference standard we also analyzed chicory roots to get a baseline estimate of the sensitivity of germacrene A detection by our laboratory setup. Fresh roots of chicory were harvested in the field in October 2013. In the laboratory, chicory roots were cleaned and stored at  $-80^\circ\text{C}$  until chemical analyses.

### **IN VIVO SAMPLING OF VOLATILES FOR GAS-CHROMATOGRAPH MASS-SPECTROMETER (GC-MS) ANALYSES**

Fresh flowers and leaves of *A. millefolium* were enclosed in a 35 × 43 cm ovenproof polyethylene terephthalate bag (PEFT) (Stewart-Jones and Poppy, 2006; Niinemets et al., 2011), and conditioned at 30°C for 3–4 h under a light intensity of 1000  $\mu\text{mol m}^{-2} \text{s}^{-1}$ . A solid-phase microextraction (SPME) fiber of 65  $\mu\text{m}$  of polydimethylsiloxane/divinylbenzene (PDMS/DVB, Supelco, Bellefonte, PA, USA) was then inserted in the headspace for sampling of volatiles. Sampling with SPME has been previously demonstrated to provide excellent means to assess the composition of volatiles in *A. millefolium* (Cornu et al., 2001). After 20 min of sampling, the fiber was removed from the bag and immediately transferred to the injection port of the gas-chromatograph mass-spectrometer (GC-MS; GC 2010 and QP 2010 Plus, Shimadzu Corporation, Kyoto, Japan). Three biological replicates were used for collection of volatiles.

Roots of chicory (*C. intybus*) (1 g dry mass) were homogenized, and the homogenate was inserted in an ovenproof 10 × 15 cm PEFT bag for 1 h at 30°C. The SPME fiber was inserted into the headspace for 20 min and then immediately transferred into the injector of the GC-MS.

Separate samples were used to estimate dry (oven-drying at 70°C to a constant mass) to fresh mass ratio of each analyzed plant fraction.

### **GC-MS ANALYSIS**

Volatiles collected onto SPME-fiber were analyzed using the Shimadzu GC-MS system. A GC column ZB5-MS (0.25 mm i.d. × 30 m, 0.25  $\mu\text{m}$  film Zebron, Phenomenex, Torrance, CA, USA) was employed for separating the volatiles using the following temperature program: 40°C for 3 min, ramp of 7°C  $\text{min}^{-1}$  to 220°C followed by a 5 min hold. When developing the GC-MS protocol, various injector temperatures between 215°C and 120°C were tested. As demonstrated previously, high temperature caused the bulk of germacrene A to be converted into  $\beta$ -elemene through Cope rearrangement (De Kraker et al., 1998; Adio, 2009a). However, too low temperatures resulted in incomplete desorption. Thus, throughout the study we used an optimized GC-MS injector temperature of 150°C.

The mass spectrometer was operated in electron-impact mode at 70 eV and in the scan range  $m/z$  of 30–400 amu. The transfer line temperature was set at 240°C and ion-source temperature at 150°C. Terpenes were identified by comparing their mass spectra and retention indices (RIs) for ZB5-MS to the spectra available in the NIST library (National Institute of Standards and Technology) and using a catalog of essential oil components (Adams, 2001). Commercially available reference compounds were purchased from Sigma-Aldrich (St. Louis, MO, USA) at the highest purity available (>98%). Based on serial dilution of standards, we estimated that the analytical detection threshold for sesquiterpenes in headspace was better than 0.1 nmol  $\text{ml}^{-1}$ , and the minimum emission rate that could be detected was lower than 2 ng  $\text{g}^{-1} \text{DW h}^{-1}$  (ca. 50-fold lower than the typical detection threshold of ca. 0.1  $\mu\text{g g}^{-1} \text{DW h}^{-1}$ ). Thus, the analytical precision of our setup was suitable to detect emissions through the high to low emission range.

### **AMPLIFICATION OF GERMACRENE A SYNTHASE GENE AND RAPID AMPLIFICATION OF cDNA ENDS (RACE-PCR)**

Total RNA was extracted from different tissues using RNeasy Mini Kit (Qiagen, Venlo, The Netherlands). The RNA was checked by agarose gel electrophoresis (Sigma-Aldrich, St. Louis, MO, USA) and the quality was evaluated by Bioanalyzer 2100 (Agilent, Santa Clara, CA, USA). The reverse transcription reaction for cDNA synthesis was carried out using iScript cDNA Synthesis Kit (Bio-Rad, Hercules, CA, USA). Based on a comparison of sequences of germacrene A synthases, three degenerate primer pairs GAS1, GAS2, and GAS3 (Table 1), were designed for six conserved regions and the polymerase chain reaction (PCR) was performed. The amplicons were either run on agarose gel, or checked by agarose gel electrophoresis and they showed fragments of approximately 894 bp, 567 bp and 402 bp for GAS1, GAS2 and GAS3 fragments, respectively. The PCR product from GAS2 (green boxes in Figure 2) was purified and inserted into a pTZ57R/T vector and transformed to *E. coli* using InsTAclone PCR Cloning Kit (Thermo Scientific, Pittsburgh, PA, USA). Fourteen individual transformants were bidirectionally sequenced and finally assembled by MEGA 5 software (Tamura et al., 2011). Two rounds of 5' and 3' RACE were done using the 5'/3' RACE Kit (Roche Diagnostics, Indianapolis, IN, USA) according to the manufacturer's protocol. The single strand cDNA for 3' and 5' ends were synthesized from 1000 ng of total RNA extracted from yarrow flowers. Based on the partial coding sequence (CDS) of *AmGAS*, nested primers were designed (Table 1, RACE-GAS). PCR was conducted as specified in the previous section and the PCR products were sequenced, assembled by MEGA 5 software and full length cDNA of *A. millefolium* germacrene A synthase (*AmGAS*) was established (Tamura et al., 2011). The full length sequence of *A. millefolium* germacrene A synthase (*AmGAS*) was registered in GenBank, <http://www.ncbi.nlm.nih.gov/> with accession number KC145534 and integrated into UniProtKB/TrEMBL, <http://www.uniprot.org/> with accession number L7XCQ7.

### **PHYLOGENETIC TREE OF GERMACRENE A SYNTHASES AND MULTIPLE SEQUENCE ALIGNMENT**

Germacrene A cDNA from *A. millefolium* was translated to the corresponding amino acid sequence and aligned and compared with other related terpenoid synthase gene sequences for Asteraceae and GAS-like sequences in phylogenetically distant angiosperms in UniProtKB/TrEMBL, <http://www.uniprot.org/>. A phylogenetic tree (Figure 1) was generated by MEGA 5 software using the UPGMA method (Tamura et al., 2011). Multiple sequence alignment was done to visualize conserved sequences among germacrene A amino acid sequences in Asteraceae (Figure 2) with BioEdit software ver. 7 (<http://www.ebi.ac.uk/Tools/clustalw2/index.html>).

### **GENE CLONING AND CONSTRUCTION OF THE EXPRESSION VECTOR**

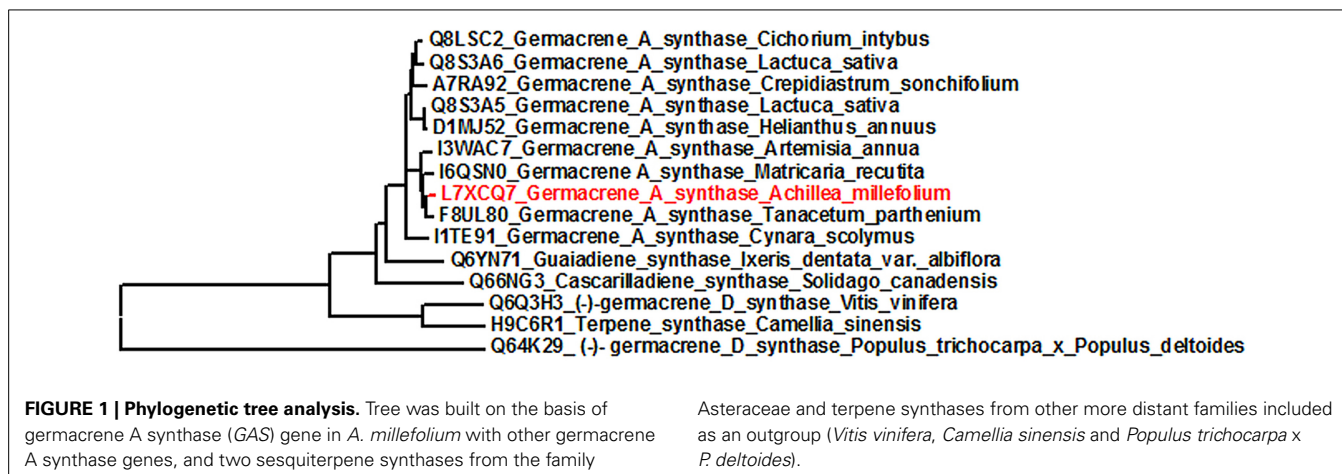
The *AmGAS* gene was amplified using specific primers containing restriction sites. Primer pairs were designed for amplification and cloning of full length *AmGAS* (~1700) gene (Table 1). Plasmid pET-26b (+) (Novagen, Madison, WI, USA) was used as the expression vector for *AmGAS*. The *AmGAS* PCR product and pET-26b (+) expression vector were digested with

**Table 1 | Sequence of primers used in this study.**

| Primer name <sup>a</sup> | Sequence (5'–3')   | Product size (bp) |
|--------------------------|--|-------------------|
| GAS1                     | GAS1-F: AGACCATTYCATCARGGGATGC<br>GAS1-R: CTCGTTDATATCYTTCCAYGCATTYTC  | 894               |
| GAS2                     | GAS2-F: TTYCCTCCTTCDGTATGGGGTGA<br>GAS2-R: TGG CAT CCC TTG ATG RAA TGG TC  | 567               |
| GAS3                     | GAS3-F: GCATCATTTCCDAGTAYATGAAG<br>GAS3-R: CCTGTAYACVACATCTATCATTC   | 402               |
| RACE-GAS                 | RACE-GAS-F: GATGAAGCCTCGGTTTTTCATCGAAGG<br>RACE-GAS-R: CACAAG GTAGTTTGTAAACCAAGGTGTCG  | 1150              |
| AmGAS                    | AmGAS-F: AATTCATGGCAGCGGTTCAAGCTACTACTGGTATC<br><i>Nco</i> <sup>b</sup><br>AmGAS-R: GATGCTCGAGTTAGTGGTGGTGGTGGTGGTGCACGGGTAGAGAATCCACAAAC<br><i>Xho</i> <sup>c</sup> | 1680              |
| GAPDH                    | GAPDH-F: ACTGGTGTCTTCACTGACAAGGA<br>GAPDH-R: GTA TCCCATTCGTTGTCGTACCA  | 135               |
| $\beta$ -actin           | $\beta$ -actin-F: ATGGAGAAGATCTGGCATCA<br>$\beta$ -actin-R: GGAAGCTGCTGGTATTTCATGAGAC  | 130               |
| RtGAS                    | RtGAS-F: CTCGGGTACTTTCAAGGAATCC<br>RtGAS-R: CTTTCGATGAAAA CCGAGGCTTC   | 123               |

<sup>a</sup>GAS1, GAS2, GAS3, degenerate primers for germacrene A synthase (GAS); RACE-GAS, GAS nested primer for rapid amplification of cDNA based on partial sequence of AmGAS; AmGAS, *Achillea millefolium* germacrene A synthase; GAPDH, glyceraldehyde 3-phosphate dehydrogenase (GenBank accession number KF286432);  $\beta$ -actin (GenBank accession number JX679606); RtGAS, real time PCR primer for AmGAS.

<sup>b,c</sup>Sequence of restriction enzymes.



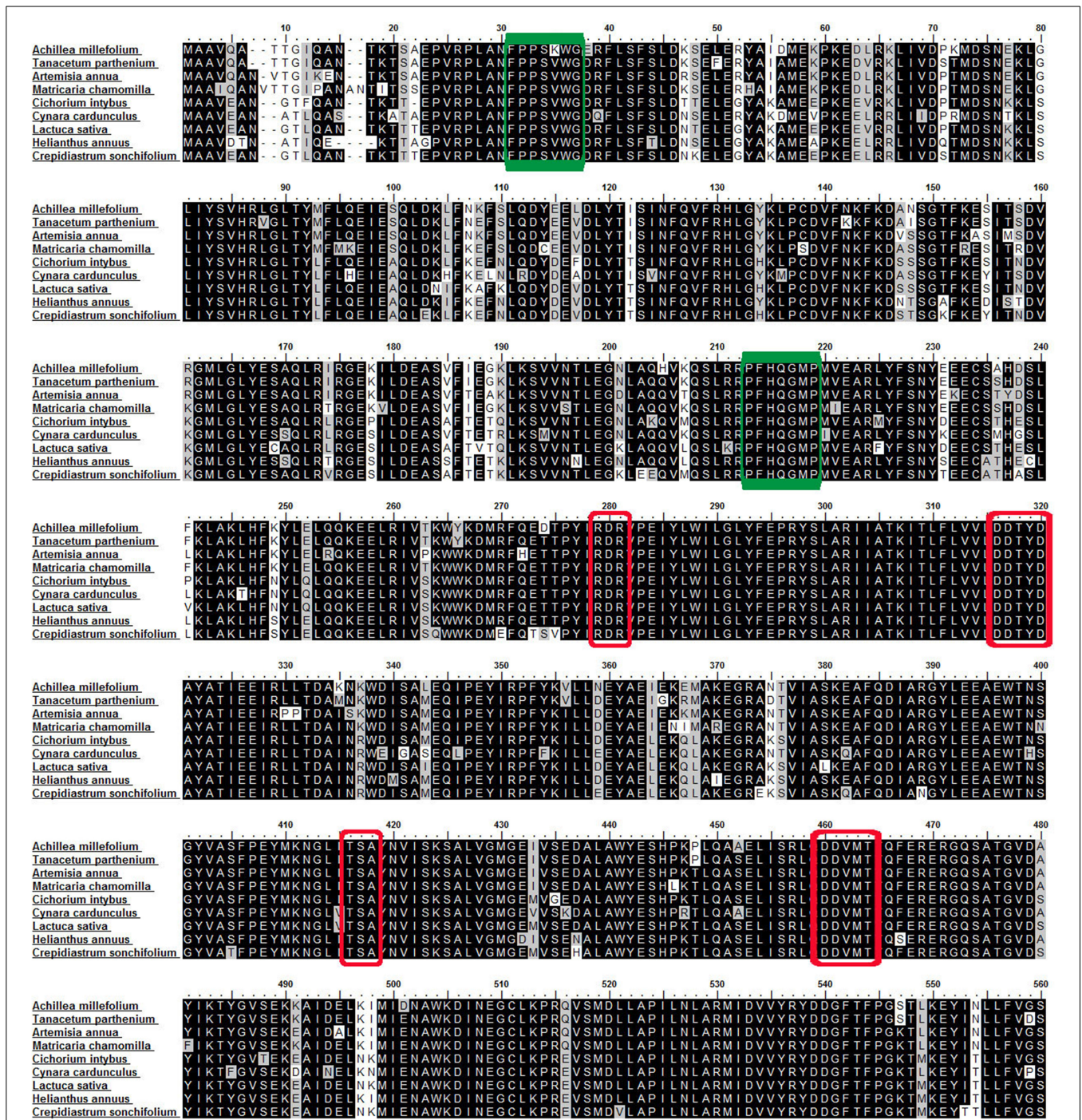
*Nco*I and *Xho*I restriction enzymes according to the manufacturer's protocol (New England Biolabs, Ipswich, MA, USA). The digested fragments were gel-purified and then *AmGAS* fragment was cloned into pET-26b (+) expression vector and transformed to *E. coli* BL21 (Novagen) using the calcium chloride transformation method (Sambrook and Russell, 2001). The lines were screened by culturing on a LB agar medium containing 50  $\mu$ g ml<sup>-1</sup> kanamycin. The obtained colonies were used in a colony PCR assay using *AmGAS*-specific primers (Table 1). The plasmids from positive colonies according to the PCR screening were digested with the same restriction enzymes used for cloning (*Nco*I and *Xho*I). The recombinant strains were selected and expression plasmid confirmed by sequencing analysis.

#### EXPRESSION OF RECOMBINANT GERMACRENE A SYNTHASE IN *ESCHERICHIA COLI*

A recombinant strain colony containing *AmGAS* gene was used in the protein expression experiment. A *E. coli* BL21 (DE3) strain containing pET-26b (+) vector was used as a control. To induce expression, isopropyl- $\beta$ -D-thiogalactoside (IPTG) was added to a final concentration of 1 mM to cultures with OD<sub>600</sub> (optical density at a wavelength of 600 nm) of 0.4. Cultures were incubated at 37°C for 2, 4, and 6 h.

#### ELECTROPHORETIC ANALYSIS OF RECOMBINANT GERMACRENE A SYNTHASE

Expression of *AmGAS* was confirmed by SDS-PAGE and western blotting. Bacterial samples were collected before



**FIGURE 2 | Multiple sequence alignment of germacrene A synthases.**

Amino acid sequence alignment of AmGAS with other germacrene A synthases from the Asteraceae family. The green boxes show position of

degenerate GAS primers used. The red boxes highlight conserved sequences (see Discussion). The alignment was conducted with BioEdit software ver. 7 (<http://www.ebi.ac.uk/Tools/clustalw2/index.html>).

(control) and after induction, and lysed in the sample buffer (100 mM Tris-HCl, pH = 8, 20% glycerol, 4% sodium dodecyl sulfate (SDS), 2%  $\beta$ -mercaptoethanol, 0.2% bromophenol blue). BlueStar prestained protein marker (Nippon Genetics Europe, Düren, Germany) was used as a size

standard. Samples were analyzed by 12% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) on a Bio-Rad Mini Protean electrophoresis unit. Protein bands were visualized by staining with Coomassie brilliant blue R-250.

Proteins were also transferred to a nitrocellulose membrane (Bio-Rad), for western blot analysis with 3, 3'-diaminobenzidine (DAB) liquid substrate system tetrahydrochloride (Sigma-Aldrich). The recombinant AmGAS has a His-tag in the C-terminus, and thus, the expression can be detected by anti His-tag peroxidase.

#### FUNCTIONAL CHARACTERIZATION OF GERMACRENE A SYNTHASE

For *in vitro* germacrene A production, cultures of pET-26b (+) plus AmGAS were grown to an OD<sub>600</sub> of 0.4, induced using IPTG (1 mM) and grown for 6 h. These cultures were pelleted by centrifugation for 5 min at 10,000 rpm and kept at -80°C. Frozen pellets were suspended in 1 mL of assay buffer selected for optimum pH 7 and ionic strength (25 mM Hepes pH 7.2, 100 mM KCl, 10 mM MnCl<sub>2</sub>, 10% glycerol, and 5 mM DTT) (Fischbach et al., 2001; Reiling et al., 2004; Rajabi et al., 2013) and lysed on ice by sonication for 1 min. Lysates were centrifuged at 17,530 RCF for 30 min at 4°C. A 200 μL of the supernatant was added into 800 μL of assay buffer in a 4 mL serum vial. We tested the use of C15 substrate farnesyl diphosphate (C15) and *cis*-configured C10 substrate neryl diphosphate (NDP) and *trans*-configured C10 substrate geranyl diphosphate (GDP). Four serum vials were considered for different substrates and then 2 μL of (1 mg mL<sup>-1</sup> aqueous solution) substrate (either FDP, GDP, mixture of FDP and GDP, mixture of FDP and NDP, mixture of GDP, NDP or NDP (Echelon Biosciences, Salt Lake City, UT, USA) was added to vials to start the reaction, and the vials were sealed. The vials were kept at 30°C for 50 min until collection of volatiles from the headspace.

To collect volatiles, a SPME fiber was inserted through the cap of the vial into the headspace for 1 min. After removal from the headspace, the fiber was transferred immediately into the injector of GC-MS and the analysis of volatiles was carried out as detailed in the section *GC-MS analyses*.

#### ISOLATION OF HOUSEKEEPING GENES AND PRIMER DESIGN FOR REAL-TIME PCR

Real-time PCR measurements for expression of genes of interest need to be normalized with respect to the housekeeping genes that are constitutively expressed in nearly all tissues and all physiological stages of an organism (Nicot et al., 2005; Maloukh et al., 2009). Two housekeeping genes,  $\beta$ -actin and glyceraldehyde 3-phosphate dehydrogenase (GAPDH), were used for this study. Based on comparison of sequences of  $\beta$ -actin and GAPDH of related species, two primer pairs were designed on the basis of conserved regions (Table 1). Two PCR products were purified and inserted in pTZ57R/T vector and transformed to *E. coli* using InsTAclone PCR Cloning Kit. Four individual transformants were sequenced and assembled by MEGA 5 software. The sequences of  $\beta$ -actin and GAPDH were registered in GenBank, <http://www.ncbi.nlm.nih.gov/>, with accession number JX679606.1 and KF286432, respectively.

The real-time PCR primers for AmGAS (Table 1, RtGAS) and housekeeping genes (Table 1) were designed on the basis of their sequences through GenScript real-time PCR primer design, <https://www.genscript.com/ssl-bin/app/primer>.

#### GENE EXPRESSION ANALYSIS OF GERMACRENE A SYNTHASE IN DIFFERENT TISSUES

RNA was extracted with three independent biological replicates from different tissues (flowers at different stages of development, leaf, stem, rhizome, and root) and quantified using a BioPhotometer plus (Eppendorf, Hamburg, Germany). First-strand cDNA was synthesized using iScript cDNA Synthesis Kit (Bio-Rad).

Quantitative PCR (qPCR) was performed with the Applied Biosystems Viia™ 7 real-time PCR system for different tissues using a qPCR iQ SYBR Green Supermix kit (Bio-Rad) according to manufacturer's instructions and using appropriate real-time PCR protocol for AmGAS (RtGAS) and housekeeping genes (Table 1). Every sample was run in three parallel reactions and the amplification specificity of primers was evaluated by melting curve analysis.

The relative gene expression levels were calculated using the comparative C<sub>t</sub> ( $\Delta\Delta C_t$ ) method (Schmittgen and Livak, 2008). According to this method, the relative gene expression is calculated as  $2^{-\Delta\Delta C_t}$ , where C<sub>t</sub> represents the threshold cycle.

## RESULTS

#### COMPOSITION OF VOLATILE BLEND OF *A. MILLEFOLIUM* FLOWERS AND LEAVES

The volatiles of *A. millefolium* detected in the emission blends were mostly monoterpenes (67% of total emissions for flowers and 59% of total emissions for leaves, and 21 compounds were above the detection threshold) and sesquiterpenes (17% for flowers and 19% for leaves, and 18 compounds were above the detection threshold, Table 2). In addition, lipoxygenase pathway volatiles (2% of total emissions for flowers and 11% of total emissions for leaves) and benzenoids, aliphatic compounds and their derivatives were found in minor proportions (Table 2). Among all the emitted compounds,  $\beta$ -pinene (36% of monoterpenes), (*E*)- $\beta$ -caryophyllene and germacrene D (Table 3) were the main floral and leaf volatiles.

#### DETERMINATION OF GERMACRENE A IN *A. MILLEFOLIUM* VOLATILE BLEND

Germacrene A is heat-labile and is converted to  $\beta$ -elemene upon heating (De Kraker et al., 1998). First, we used the roots of chicory (*C. intybus*) known to contain and emit germacrene A for optimization of sampling and GC analysis protocols. Based on this work, a GC injector temperature of 150°C was used in all GC-MS analyses, resulting in a significantly increased fraction of germacrene A detected compared to  $\beta$ -elemene, although a large part of germacrene A was still converted to  $\beta$ -elemene (Figure 3). Germacrene A mass-spectrum of *in vitro* analyses of AmGAS with FDP as substrate matched with the published spectrum, except for differences in the proportions of mass-fragments of 79 and 81 (Figures 3C,D). In earlier studies we have noticed similar differences with the identification of some sesquiterpenes using authentic standards, e.g., identification of germacrene D, suggesting that these minor differences were specific to the GC-MS device used.

Average germacrene A emission was between ca. 0.4 and 2.3% of total sesquiterpene emission in *A. millefolium* (Table 3).

**Table 2 | Relative proportions of volatiles (%) detected in the emissions of yarrow (*Achillea millefolium*) flowers and leaves.**

| Volatile                  | Retention index <sup>a</sup> |                    | Flower mean ± SE<br>n = 3 | Leaf mean ± SE<br>n = 3 |
|---------------------------|------------------------------|--------------------|---------------------------|-------------------------|
|                           | Lit.                         | Calc. <sup>b</sup> |                           |                         |
|                           |                              |                    |                           |                         |
| 1-Hexanol                 | 871                          | 858                | 0.047 ± 0.038             | 0.67 ± 0.38             |
| (3Z)-Hexenol              | 859                          | 863                | 0.39 ± 0.26               | 0.84 ± 0.63             |
| (3Z)-Hexenyl acetate      | 1005                         | 1008               | 1.32 ± 0.67               | 8.6 ± 4.1               |
| (3Z)-Hexenyl pentanoate   |                              | 1237               | –                         | 1.08 ± 0.88             |
| <b>MONOTERPENOIDS</b>     |                              |                    |                           |                         |
| Santolina triene          | 909                          | 902                | 1.02 ± 0.83               | 1.29 ± 1.06             |
| α-Thujene                 | 930                          | 926                | 0.309 ± 0.095             | 0.262 ± 0.098           |
| α-Pinene                  | 939                          | 932                | 4.93 ± 0.73               | 3.31 ± 0.67             |
| Camphene                  | 954                          | 949                | 0.28 ± 0.17               | 0.073 ± 0.030           |
| Sabinene                  | 975                          | 975                | 13.5 ± 3.9                | 15.4 ± 8.7              |
| β-Pinene                  | 979                          | 980                | 23.1 ± 1.9                | 11.9 ± 2.5              |
| Myrcene                   | 991                          | 988                | 1.22 ± 0.34               | 1.33 ± 0.32             |
| α-Phellandrene            | 1003                         | 1007               | 0.45 ± 0.37               | 0.91 ± 0.30             |
| Limonene                  | 1029                         | 1029               | 2.41 ± 0.19               | 2.46 ± 0.14             |
| β-Phellandrene            | 1030                         | 1033               | 1.19 ± 0.32               | 1.62 ± 0.42             |
| 1,8-Cineole               | 1031                         | 1036               | 7.8 ± 2.2                 | 0.79 ± 0.56             |
| (Z)-β-Ocimene             | 1037                         | 1035               | 2.92 ± 0.57               | 3.65 ± 1.61             |
| (E)-β-Ocimene             | 1050                         | 1046               | 5.4 ± 2.1                 | 12.5 ± 6.9              |
| γ-Terpinene               | 1060                         | 1058               | 1.39 ± 0.54               | 1.89 ± 0.70             |
| cis-Sabinene hydrate      | 1070                         | 1075               | 0.210 ± 0.088             | –                       |
| Terpinolene               | 1089                         | 1085               | 0.121 ± 0.015             | 0.157 ± 0.069           |
| trans-Pinocarveol         | 1139                         | 1147               | 0.121 ± 0.099             | 0.155 ± 0.080           |
| Camphor                   | 1146                         | 1153               | 0.064 ± 0.012             | 0.29 ± 0.14             |
| Pinocarvone               | 1165                         | 1168               | 0.164 ± 0.069             | 0.104 ± 0.051           |
| Borneol                   | 1169                         | 1179               | 0.191 ± 0.059             | 0.075 ± 0.062           |
| α-Terpineol               | 1189                         | 1197               | 0.415 ± 0.088             | 0.29 ± 0.13             |
| <b>SESQUITERPENES</b>     |                              |                    |                           |                         |
| δ-Elemene                 | 1338                         | 1340               | –                         | 0.29 ± 0.24             |
| α-Cubebene                | 1351                         | 1352               | 0.065 ± 0.053             | –                       |
| α-Copaene                 | 1377                         | 1382               | 0.138 ± 0.031             | 0.311 ± 0.059           |
| β-Bourbonene              | 1388                         | 1391               | 0.56 ± 0.16               | 0.29 ± 0.11             |
| β-Cubebene                | 1388                         | 1394               | 0.16 ± 0.13               | 0.044 ± 0.036           |
| β-Elemene                 | 1391                         | 1390               | 0.107 ± 0.087             | 0.44 ± 0.35             |
| α-Isocomene               | 1388                         | 1397               | –                         | 0.111 ± 0.057           |
| (E)-β-Caryophyllene       | 1419                         | 1428               | 6.7 ± 1.4                 | 7.78 ± 3.03             |
| trans-α-Bergamotene       | 1435                         | 1434               | 1.88 ± 0.99               | 0.226 ± 0.052           |
| (Z)-β-Farnesene           | 1443                         | 1455               | 0.53 ± 0.32               | 0.51 ± 0.13             |
| α-Humulene                | 1455                         | 1464               | 0.71 ± 0.13               | 0.77 ± 0.17             |
| cis-Muurolo-4(14),5-diene | 1467                         | 1470               | 0.111 ± 0.011             | 0.064 ± 0.028           |
| Germacrene D              | 1485                         | 1489               | 4.8 ± 1.4                 | 6.8 ± 3.4               |
| α-Zingiberene             | 1494                         | 1499               | 0.16 ± 0.13               | –                       |
| Bicyclgermacrene          | 1500                         | 1504               | 0.203 ± 0.055             | 0.26 ± 0.10             |
| (E,E)-α-Farnesene         | 1506                         | 1506               | 0.37 ± 0.22               | 0.65 ± 0.22             |

(Continued)

**Table 2 | Continued**

| Volatile                               | Retention index <sup>a</sup> |                    | Flower mean ± SE<br>n = 3 | Leaf mean ± SE<br>n = 3 |
|--|------------------------------|--------------------|---------------------------|-------------------------|
|  | Lit.                         | Calc. <sup>b</sup> |                           |                         |
|  |                              |                    |                           |                         |
| Germacrene A                           | 1509                         | 1512               | 0.074 ± 0.057             | 0.22 ± 0.15             |
| γ-Cadinene                             | 1514                         | 1521               | 0.057 ± 0.031             | 0.062 ± 0.051           |
| δ-Amorphene                            | 1512                         | 1525               | 0.39 ± 0.22               | 0.41 ± 0.23             |
| β-Sesquiphellandrene                   | 1523                         | 1529               | 0.042 ± 0.020             | 0.246 ± 0.085           |
| <b>OTHER VOLATILES</b>                 |                              |                    |                           |                         |
| 3-Methylbutanoic acid                  |                              | 853                | 1.47 ± 1.04               | –                       |
| 2-Methylbutanoic acid                  |                              | 862                | 0.59 ± 0.45               | –                       |
| 1-Nonene                               |                              | 893                | 7.7 ± 2.3                 | 5.4 ± 3.8               |
| Benzaldehyde                           | 960                          | 967                | 1.61 ± 0.73               | –                       |
| p-Cymene                               | 1025                         | 1028               | 2.40 ± 1.01               | 1.75 ± 0.66             |
| 4,8-Dimethyl-1,3-E,7-nonatriene (DMNT) |                              | 1115               | 0.110 ± 0.079             | 2.49 ± 1.32             |
| Methyl salicylate (MeSA)               | 1192                         | 1198               | 0.101 ± 0.015             | 1.23 ± 0.56             |

<sup>a</sup>Retention indices (Adams, 2001).<sup>b</sup>Retention indices calculated by injecting the hydrocarbon standard of C<sub>8</sub> to C<sub>20</sub> (Sigma-Aldrich, St. Louis, MO, USA) to GC-MS.

Assuming further that β-elemene detected in the emission blend is the conversion product of germacrene A, the emission estimates are ca. 1% for flowers and 7% for leaves (Table 3), suggesting that germacrene A is a minor component in the volatile blend of flowers, and a moderately high component in leaf sesquiterpene emissions (Table 3).

Next to the emissions we evaluated also the chemical composition and content of terpenoids in bud, flower and leaf extracts of *A. millefolium*. We observed statistically similar amounts (10.6 ± 2.4 μg g<sup>-1</sup> DW) of germacrene A in bud, flower and leaf extracts of *A. millefolium*.

#### CLONING OF GERMACRENE A SYNTHASE IN *A. MILLEFOLIUM*

*AmGAS* partial sequence was amplified by degenerate primers (green boxes in Figure 2 show position of the degenerate primers) and then the full length was obtained by 5' and 3' amplification of cDNA ends (RACE-PCR). The length of the coding sequence of *AmGAS* is 1680 bp, and it encodes a protein of 559 AA residues with the predicted molecular weight of 62 kD and isoelectric point (pI) of 5.24 ([http://web.expasy.org/compute\\_pi/](http://web.expasy.org/compute_pi/)). The overall length and lack of the characteristic chloroplast-targeting signal peptide suggests that *AmGAS* is functional in the cytosol.

The blast searches in NCBI and UniProtKB showed that *AmGAS* belongs to terpenoid synthase (TPS) gene subfamily TPS-a (Bohlmann et al., 1998), and has a high similarity with germacrene A synthases of two other members of Asteraceae, *T. parthenium* (F8UL80) and *A. annua* (I3WAC7) (Figure 1).

**Table 3 | Relative proportions (%) of sesquiterpenes detected in the emission of flowers and leaves of yarrow (*Achillea millefolium*).**

| Volatile                             | Retention index <sup>a</sup> |       | Flower mean $\pm$ SE<br><i>n</i> = 3 | Leaf mean $\pm$ SE<br><i>n</i> = 3 |
|--------------------------------------|------------------------------|-------|--------------------------------------|------------------------------------|
|                                      | Lit.                         | Calc. |                                      |                                    |
| $\delta$ -Elemene                    | 1338                         | 1340  |                                      | 0.98 $\pm$ 0.81                    |
| $\alpha$ -Cubebene                   | 1351                         | 1352  | 0.33 $\pm$ 0.27                      |                                    |
| $\alpha$ -Copaene                    | 1377                         | 1382  | 0.79 $\pm$ 0.27                      | 1.71 $\pm$ 0.20                    |
| $\beta$ -Bourbonene                  | 1388                         | 1391  | 3.18 $\pm$ 0.82                      | 1.50 $\pm$ 0.52                    |
| $\beta$ -Cubebene                    | 1388                         | 1394  | 0.82 $\pm$ 0.67                      | 0.15 $\pm$ 0.12                    |
| $\beta$ -Elemene                     | 1391                         | 1390  | 0.61 $\pm$ 0.49                      | 4.8 $\pm$ 3.8                      |
| $\alpha$ -Isocomene                  | 1388                         | 1397  |                                      | 0.51 $\pm$ 0.29                    |
| ( <i>E</i> )- $\beta$ -Caryophyllene | 1419                         | 1428  | 38.7 $\pm$ 5.8                       | 37.3 $\pm$ 14.2                    |
| <i>trans</i> - $\alpha$ -Bergamotene |                              |       | 13.0 $\pm$ 7.4                       | 1.22 $\pm$ 0.18                    |
| ( <i>Z</i> )- $\beta$ -Farnesene     |                              | 1455  | 3.8 $\pm$ 2.4                        | 2.63 $\pm$ 0.43                    |
| $\alpha$ -Humulene                   | 1455                         | 1464  | 3.87 $\pm$ 0.42                      | 4.07 $\pm$ 0.46                    |
| <i>cis</i> -Muurolo-4(14),5-diene    | 1467                         | 1470  | 0.69 $\pm$ 0.14                      | 0.26 $\pm$ 0.11                    |
| Germacrene D                         | 1485                         | 1489  | 27.0 $\pm$ 7.1                       | 32.4 $\pm$ 9.8                     |
| $\alpha$ -Zingiberene                | 1494                         | 1499  | 0.83 $\pm$ 0.68                      |                                    |
| Bicyclogermacrene                    | 1500                         | 1504  | 1.15 $\pm$ 0.28                      | 1.87 $\pm$ 1.1                     |
| ( <i>E,E</i> )- $\alpha$ -Farnesene  | 1506                         | 1506  | 2.2 $\pm$ 1.2                        | 5.1 $\pm$ 2.9                      |
| Germacrene A                         | 1509                         | 1512  | 0.42 $\pm$ 0.32                      | 2.3 $\pm$ 1.7                      |
| $\gamma$ -Cadinene                   | 1514                         | 1521  | 0.30 $\pm$ 0.16                      | 0.21 $\pm$ 0.17                    |
| $\delta$ -Amorphene                  | 1512                         | 1525  | 2.09 $\pm$ 0.99                      | 1.8 $\pm$ 0.6                      |
| $\beta$ -Sesquiphellandrene          | 1523                         | 1529  | 0.220 $\pm$ 0.099                    | 1.1 $\pm$ 0.4                      |
| Total emission of sesquiterpenes     |                              |       | 100                                  | 100                                |

<sup>a</sup>Retention indices as in **Table 2**.

Multiple sequence alignment of AmGAS amino acid sequences for several additional Asteraceae family members further showed conserved motifs of terpenoid synthases (**Figure 2**). Nevertheless, AmGAS also has a relatively high homology with other sesquiterpene synthases from Asteraceae, while the similarity is much less with other TPS-a gene subfamily members in other angiosperms (**Figure 1**).

#### GENE CLONING AND EXPRESSION OF GERMACRENE A SYNTHASE

Recombinant protein expression after induction was analyzed by SDS-PAGE and western blotting. Analysis of the recombinant protein expression by SDS-PAGE demonstrated a protein band at around 62 kD in induced recombinant strain samples containing pET-26b (+) plus AmGAS. This band corresponding to the calculated molecular mass of AmGAS protein was not observed in negative control and non-induced samples (**Figure 4A**). The western blotting also confirmed the expression (**Figure 4B**).

#### FUNCTIONAL CHARACTERIZATION OF RECOMBINANT GERMACRENE A SYNTHASE

Functional characterization of AmGAS *in vitro* was carried out by incubation with farnesyl diphosphate (FDP), the substrate for sesquiterpenes, and geranyl diphosphate (GDP) and neryl diphosphate (NDP), the substrates for synthesis of monoterpenes. Incubation with FDP yielded  $\beta$ -elemene and germacrene A as the main volatiles in the headspace with minor contributions

of  $\alpha$ - and  $\beta$ -selinene (**Table 3**, **Figure 3**). The percentage of germacrene A detected was greater for lower injector temperature (**Figure 3**), again suggesting that the bulk of  $\beta$ -elemene might reflect the heat conversion of germacrene A.

Incubation of AmGAS with GDP produced mainly aliphatic monoterpenes myrcene and *Z*- and *E*- $\beta$ -ocimene, but cyclic monoterpenes limonene and terpinolene were also produced at moderately high levels (**Figure 5A**, **Table 4**). AmGAS with NDP produced mainly limonene and terpinolene (**Figure 5B**, **Table 4**).

When equimolar concentrations of GDP and NDP were provided, AmGAS produced only monoterpenes (**Figure 5C**). With GDP and FDP or NDP and FDP, AmGAS produced both mono- and sesquiterpenes, whereas monoterpene production was favored over sesquiterpene production (**Figures 5D,E**).

#### RNA PROFILING OF GERMACRENE A SYNTHASE IN DIFFERENT TISSUES

Quantitative (real-time) PCR measurements of AmGAS were conducted with different tissues, including leaf, rhizome and root, and for flowers at different stages of development (bud, early flowering, full flowering and senescence). AmGAS was expressed in all organs, but the relative expression level was higher in flowers and leaves than in roots, stems and rhizomes (**Figure 6A**). However, flower developmental stage did not significantly alter the relative expression of AmGAS, although the variability was large (**Figure 6B**). The results were quantitatively identical by using the expression level of either  $\beta$ -actin or glyceraldehyde 3-phosphate dehydrogenase, the housekeeping genes selected, to normalize the AmGAS expression.

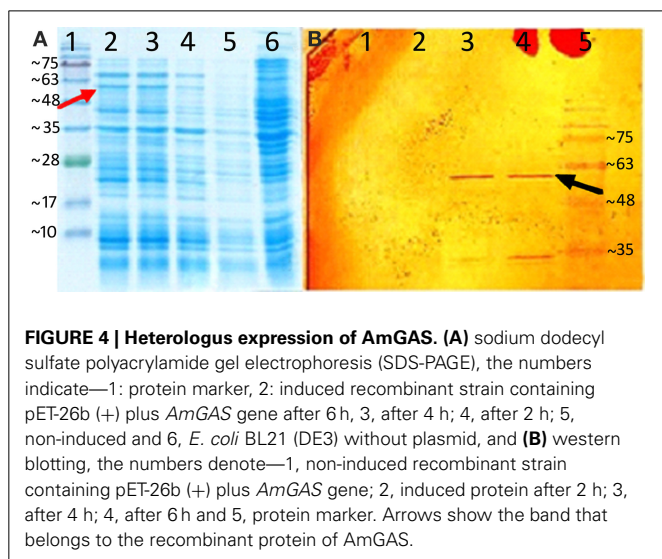
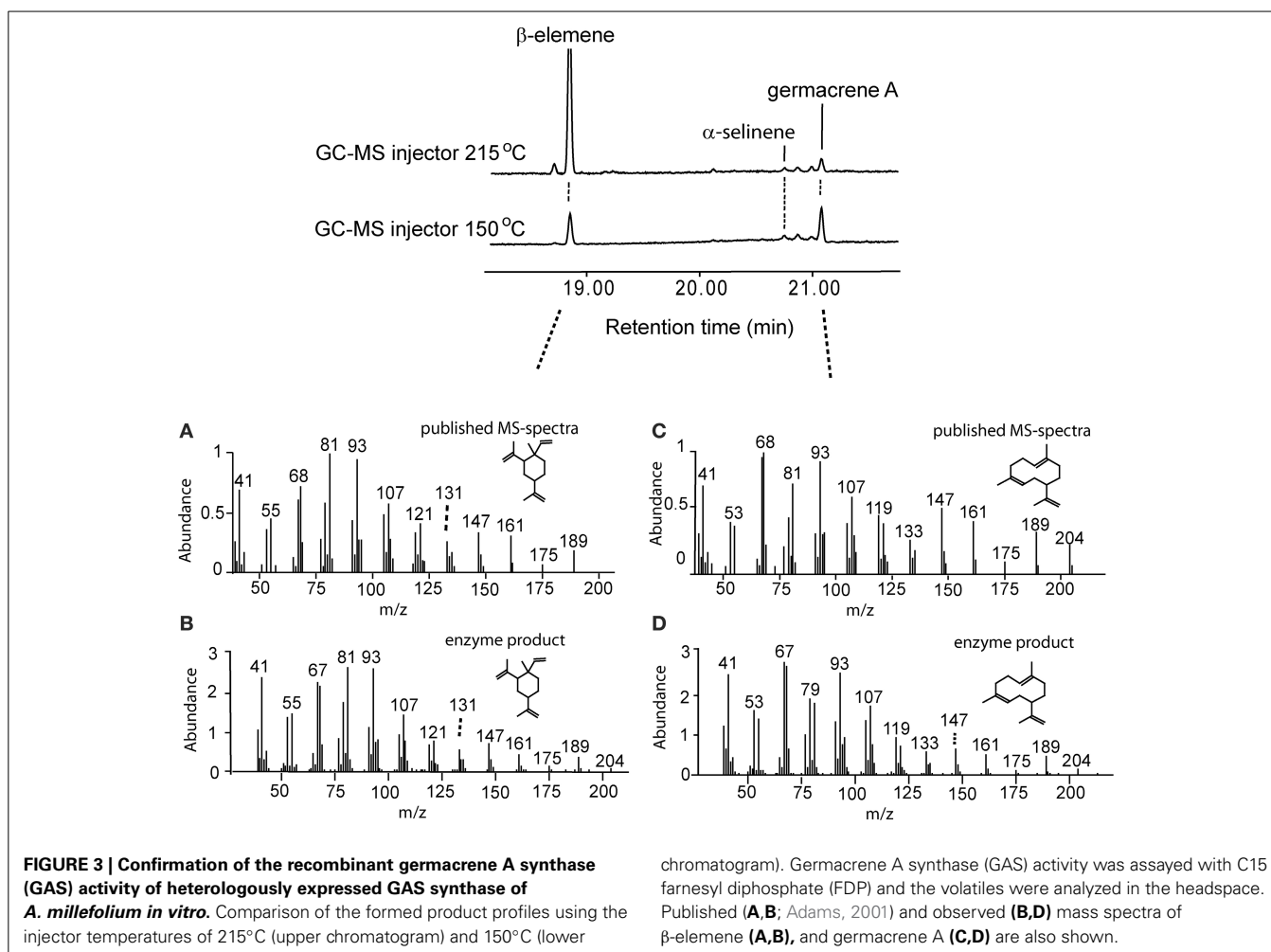
## DISCUSSION

### TERPENOID VOLATILES IN *A. MILLEFOLIUM* AND CONTRIBUTIONS OF GERMACRENE A AND $\beta$ -ELEMENE

Content and composition of secondary metabolites in *A. millefolium* tissues has been addressed in several publications (Orav et al., 2006; Gudaityte and Venskutonis, 2007; Raal et al., 2012; Dias et al., 2013). These studies have demonstrated occurrence of multiple chemotypes with varying chemical composition of the essential oil, although monoterpenes  $\beta$ -pinene,  $\alpha$ -pinene, sabinene, limonene, 1,8-cineole,  $\beta$ -phellandrene and ocimenes and sesquiterpenes (*E*)- $\beta$ -caryophyllene and germacrene D have often been observed as the main constituents of yarrow essential oil (Mockute and Judzentiene, 2003; Orav et al., 2006; Gudaityte and Venskutonis, 2007; Judzentiene and Mockute, 2010; Raal et al., 2012) agreeing with our observations (**Tables 2, 3**).

Germacrene A has been detected in past studies in trace level in only some chemotypes (Gudaityte and Venskutonis, 2007), and some studies have not detected germacrene A (Cornu et al., 2001; Lyakina, 2002; Orav et al., 2006; Raal et al., 2012). On the other hand,  $\beta$ -elemene was observed in all tested *A. millefolium* chemotypes at a significant level in the study of Gudaityte and Venskutonis (2007). However, all these past studies have used high GC injector temperatures of 230 to 250°C. Given that germacrene A is heat labile and is converted to  $\beta$ -elemene through the Cope rearrangement upon heating (De Kraker et al., 1998; Adio, 2009a), lack of germacrene A identification





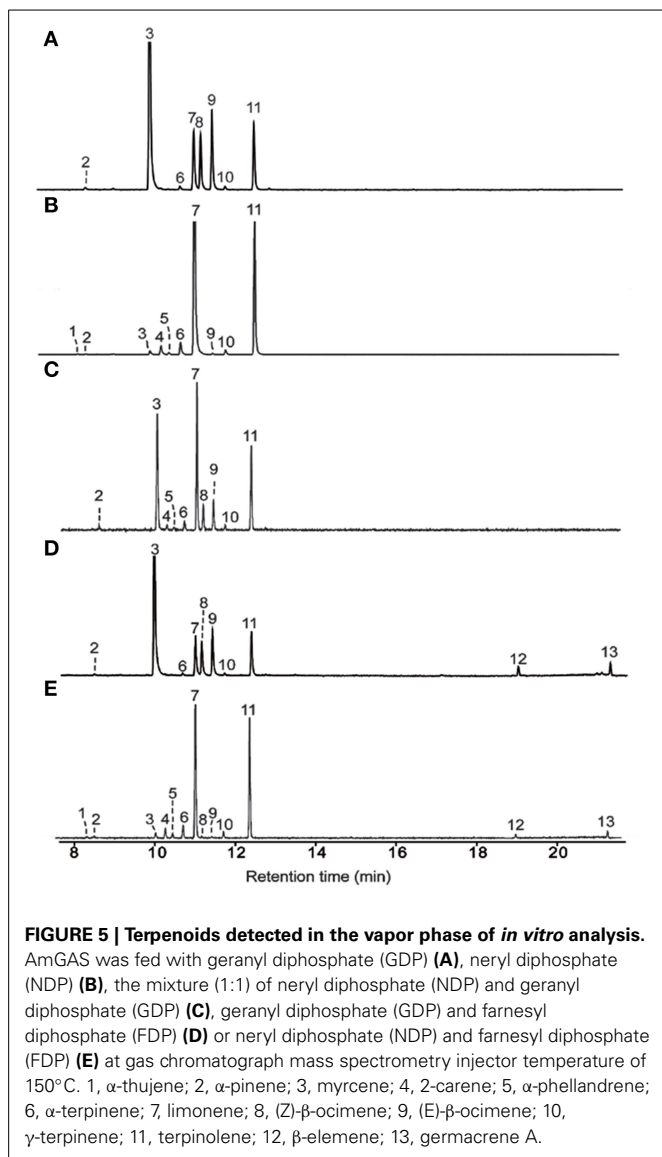
in several past studies can have resulted from excessive injector temperatures.

Here we analyzed *A. millefolium* volatiles in two different injector temperatures of 150°C and 215°C. Similarly to earlier findings

(De Kraker et al., 1998), moderately high injector temperature of 215°C caused the rearrangement of the bulk of germacrene A into  $\beta$ -elemene, while at the injector temperature of 150°C, much greater germacrene A detection yield was achieved (Figure 3). In fact, considering that  $\beta$ -elemene is the rearrangement product of germacrene A, the predicted contribution of germacrene A emission to total sesquiterpene emission was ca. 1% for flowers and 7% for leaves, indicating that germacrene A is a moderately important sesquiterpene in the emissions of *A. millefolium* (Table 3). However, the sum of germacrene A and  $\beta$ -elemene at the injector temperature of 150°C was less than that at the injector temperature of 215°C (Figure 3), suggesting imperfect desorption at this temperature. Thus, the release of germacrene A from *A. millefolium* can be even higher than detected by the modified procedure with mild injector temperature.

#### IDENTIFICATION OF GERMACRENE SYNTHASE GENE IN *A. MILLEFOLIUM* AND RNA PROFILING IN DIFFERENT TISSUES

Germacrene A synthase has been previously amplified in different Asteraceae family members (Figure 1) including *T. parthenium* (Majdi et al., 2011), *C. intybus* (Bouwmeester et al., 2002), *Solidago canadensis* (Prosser et al., 2002), *Helianthus annuus* (Göpfert et al., 2009, 2010), *Crepidiastrum sonchifolium* (Ren



et al., 2006), *Matricaria recutita* (Irmisch et al., 2012), *Lactuca sativa* (Bennett et al., 2002), *A. annua* (Bertea et al., 2006) and *Ixeris dentata* (Kim et al., 2005). Here we further amplified the germacrene A synthase gene (*AmGAS*) from important medicinal plant *A. millefolium*, cloned and expressed *AmGAS* in *E. coli* BL21 (DE3) and confirmed the bacterial expression by SDS-PAGE and western blotting (Figure 4). Thus, a new promising plant model system was developed to investigate the regulation and evolution of germacrenes' family of sesquiterpene synthesis.

RNA profiling of *AmGAS* in *A. millefolium* tissues showed different levels of germacrene A synthase in different tissues with the highest expression level observed in leaves and flowers and much lower expression level in rhizome, root and stem tissues (Figure 6). Organ-specific expression profile of *AmGAS* is in agreement with previous observations in other species having germacrene A synthases (De Kraker et al., 2001; Bouwmeester et al., 2002; Kim et al., 2005; Nguyen et al., 2010). Nevertheless, in some plant species such as *T. parthenium* even a more specialized GAS

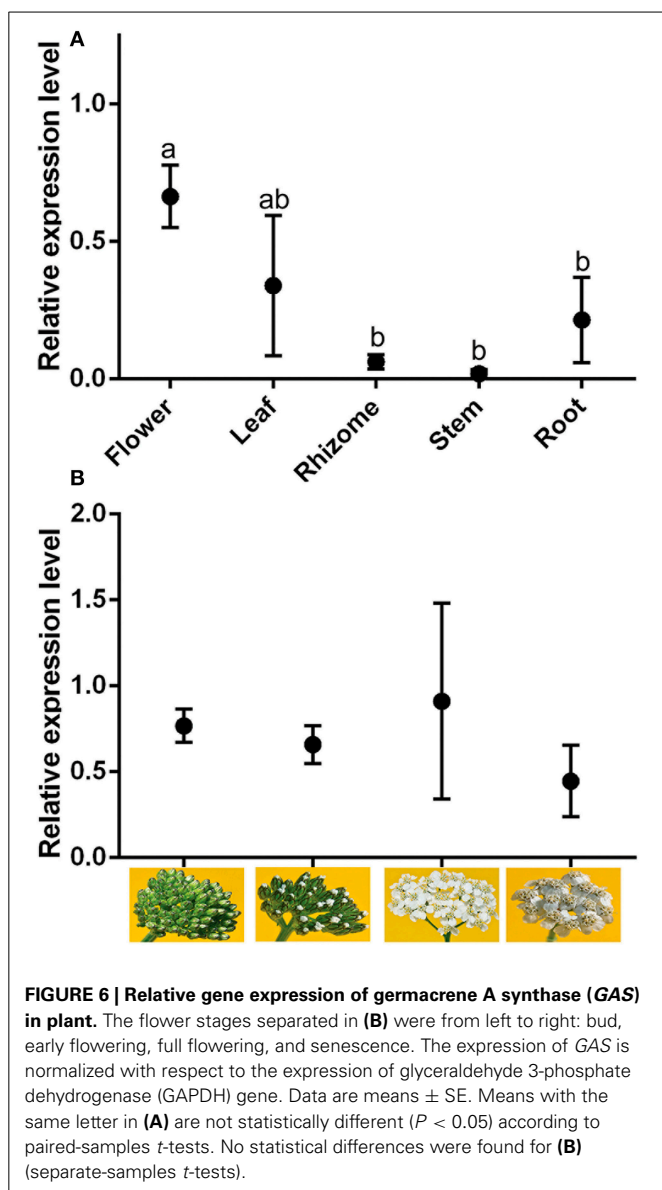
**Table 4 | Terpenoids detected in the headspace of *in vitro* analysis of recombinant *A. millefolium* germacrene A synthase (*AmGAS*) fed with geranyl diphosphate (GDP), neryl diphosphate (NDP) or farnesyl diphosphate (FDP) at gas chromatograph mass spectrometry (GC-MS) injector temperature of 150°C.**

| Substrate                      | Product                        | Retention index   |       | Relative proportion (%) |
|--------------------------------|--------------------------------|-------------------|-------|-------------------------|
|                                |                                | Lit.              | Calc. |                         |
| FDP                            | $\beta$ -elemene               | 1391              | 1390  | 47.5                    |
|                                | $\beta$ -selinene              | 1490              | 1493  | 4.9                     |
|                                | $\alpha$ -selinene             | 1498              | 1500  | 7.1                     |
|                                | germacrene A                   | 1509              | 1512  | 40.5                    |
| GDP                            | $\alpha$ -thujene              | 926               | 926   | <0.1                    |
|                                | $\alpha$ -pinene               | 939               | 932   | 1.2                     |
|                                | Camphene                       | 954               | 949   | 0.2                     |
|                                | myrcene                        | 991               | 988   | 51.8                    |
|                                | $\alpha$ -terpinene            | 1017              | 1016  | 0.6                     |
|                                | limonene                       | 1029              | 1029  | 11.5                    |
|                                | ( <i>Z</i> )- $\beta$ -ocimene | 1037              | 1035  | 8.8                     |
|                                | ( <i>E</i> )- $\beta$ -ocimene | 1050              | 1046  | 12.2                    |
|                                | $\gamma$ -terpinene            | 1060              | 1058  | 0.6                     |
|                                | terpinolene                    | 1089              | 1085  | 11.6                    |
|                                | linalool                       | 1097              | 1099  | 0.8                     |
|                                | $\alpha$ -terpineol            | 1189              | 1197  | 0.3                     |
|                                | NDP                            | $\alpha$ -thujene | 926   | 926                     |
| $\alpha$ -pinene               |                                | 939               | 932   | 0.2                     |
| $\alpha$ -fenchene             |                                | 953               | 941   | <0.1                    |
| camphene                       |                                | 954               | 949   | <0.1                    |
| myrcene                        |                                | 991               | 988   | 0.8                     |
| 2-carene                       |                                | 1002              | 997   | 2.1                     |
| $\alpha$ -phellandrene         |                                | 1003              | 1004  | 0.4                     |
| $\alpha$ -terpinene            |                                | 1017              | 1016  | 2.9                     |
| limonene                       |                                | 1029              | 1029  | 62.5                    |
| ( <i>Z</i> )- $\beta$ -ocimene |                                | 1037              | 1035  | 0.5                     |
| ( <i>E</i> )- $\beta$ -ocimene |                                | 1050              | 1046  | 0.2                     |
| $\gamma$ -terpinene            |                                | 1060              | 1058  | 1.0                     |
| terpinolene                    |                                | 1089              | 1085  | 29.1                    |
| $\alpha$ -terpineol            | 1189                           | 1197              | <0.1  |                         |

expression pattern has been found with the expression mainly confined to flowers and very low expression level or none in leaves and roots (Majdi et al., 2011). Future studies are needed to gain insight into regulatory elements responsible for organ-specific expression pattern and species differences in organ-specificity of expression.

#### ENZYME ASSAY OF GERMACRENE A SYNTHASE IN *A. MILLEFOLIUM*

AmGAS analysis with different substrates indicated that it is a multi-substrate enzyme that is capable of binding either C10 substrates GDP or NDP to form monoterpenes, or C15 substrate FDP to form sesquiterpenes (Table 4, Figures 3, 5). Although multiple substrates are not routinely used in functional characterization of terpenoid synthases, it has been demonstrated that several terpenoid synthases are capable of using multiple



substrates (Rajabi et al., 2013). For example, Steele et al. (1998a) showed that sesquiterpene  $\delta$ -selinene synthase and  $\gamma$ -humulene synthase from conifer *Abies grandis* could produce monoterpenes when incubated with GDP *in vitro*. Analogously, sweet basil (*Ocimum basilicum*)  $\alpha$ -zingiberene synthase can catalyze formation of several cyclic monoterpenes when GDP is provided as substrate (Davidovich-Rikanati et al., 2008). In apple (*Malus domestica*) sesquiterpene  $\alpha$ -farnesene synthase formed monoterpenes, in particular acyclic monoterpenes, *E*- $\beta$ -ocimene, myrcene and linalool when GDP was given as substrate (Green et al., 2007). This latter result is analogous to AmGAS reaction with GDP where mainly acyclic monoterpenes were produced in our study (Table 4).

It is interesting that AmGAS incubation with NDP resulted in production of cyclic monoterpenes, while incubation with GDP mainly resulted in production of acyclic monoterpene. This

indicates that substrate structure importantly drives the product profiles of AmGAS. It is plausible that the *trans*-substrate, GDP, ionizes mainly to linalyl cation, resulting in production of acyclic products, while the *cis*-substrate, NDP, ionizes to neryl cation and further to terpinyl cation leading to production of cyclic monoterpenes (Schillmiller et al., 2009). The linalyl cation can further isomerize to neryl cation, but the reverse, *cis-trans*-isomerization is likely sterically restricted as no acyclic monoterpenes were formed with NDP.

Despite AmGAS has the monoterpene synthase activity similarly to some other sesquiterpene synthases, the functional significance of this finding, especially the finding of the use of potential use of NDP, is not fully clear. Traditionally, monoterpene synthesis has been considered to occur in plastids, while sesquiterpene synthesis in cytosol (Dudareva et al., 2004, 2006; Pichersky et al., 2006; Tholl, 2006; Bohlmann and Keeling, 2008; Chen et al., 2011; Gutensohn et al., 2013; Rajabi et al., 2013). This understanding stems from the evidence of subcellular localization of pertinent terpenoid synthases and distribution of GDP (assumed to be mainly in chloroplasts) and FDP (assumed to be mainly in cytosol). Chloroplastic monoterpene synthases have a typical transit peptide at the N-terminal position which is responsible for chloroplast targeting. Thus they are 50–70 amino acids longer (600–650 amino acids) than sesquiterpene synthases which lack a transit peptide and contain 550–580 amino acids (Bohlmann et al., 1998; Rajabi et al., 2013). Lack of the transient peptide and overall length of AmGAS (559 amino acids), suggest that AmGAS is functionally active in the cytosol. As in our study, a greater affinity to GDP than to FDP has been observed for some other sesquiterpene synthases. For instance, a sesquiterpene synthase (LaBERS) from lavender, used GDP with a higher affinity than FDP and also produced monoterpenes, albeit with low rates (Landmann et al., 2007). It has been suggested that LaBERS has probably evolved from a monoterpene synthase by the loss of the plastidial signal peptide and by broadening its substrate spectrum.

On the other hand, there is recent evidence that multiple-substrate sesquiterpene synthases in cytosol can function as monoterpene synthases in cytosol when GDP becomes available (Davidovich-Rikanati et al., 2008; Gutensohn et al., 2013), presumably through the export of GDP from chloroplasts (Gutensohn et al., 2013). Previously, the cross-talk among chloroplastic and cytosolic isoprenoid synthesis pathways has been thought to occur at the level of C5 intermediate isopentenyl diphosphate (IDP) (Hemmerlin et al., 2003; Laule et al., 2003). However, the experimental evidence suggests that as yet unidentified IDP-transporter can also transport GDP (Bick and Lange, 2003). In fact,  $^{13}\text{C}$ -labeling suggests that chloroplast-derived GDP can be used in cytosolic sesquiterpene synthesis in chamomile (*M. recutita*), close relative of *A. millefolium*, (Adam and Zapp, 1998; Adam et al., 1999), suggesting that GDP can be available for cytosolic monoterpene synthesis in Asteraceae.

In transgenic tomato that expresses multiple-substrate sesquiterpene  $\alpha$ -zingiberene synthase in cytosol, monoterpene synthesis in cytosol was relatively small unless chloroplastic GDP pool was strongly enhanced by overexpressing plastidic GDP synthase (Gutensohn et al., 2013). This evidence opens up an exciting

opportunity that physiological conditions leading to buildup of chloroplastic GDP can enhance GDP transport to cytosol, leading to major enhancement of cytosolic monoterpene synthesis. In fact, our study indicated that AmGAS affinity to GDP is greater than to FDP as more monoterpenes were produced when both substrates were given in equimolar concentrations (**Figure 5D**). Clearly the substrate affinity, C10 vs. C15, can depend on multiple factors such as the concentration of metal cations and pH of the reaction medium (Green et al., 2007), but nevertheless this result suggests that the balance between sesqui- and monoterpenes can be importantly altered by GDP availability. Ocimene-type aliphatic sesquiterpenes synthesized by AmGAS when GDP is provided as substrate are classic stress-induced monoterpenes (Rodríguez-Saona et al., 2001; D' Alessandro and Turlings, 2005; Arimura et al., 2009; Copolovici et al., 2011, 2012) that in the case of some stresses are induced almost instantaneously in response to stress (Copolovici et al., 2012). Possible regulation of chemical profiles by enzyme substrate availability, FDP vs. GDP, provides a potential important control point for physiological regulation of cytosolic terpene synthesis.

There is also a long-standing enigma of how monoterpene synthesis proceeds in heterotrophic compartments of aromatic plants lacking specialized storage structures. Plastidic monoterpene synthesis, especially in leaves, is classically strongly linked to photosynthetic carbon metabolism (Niinemets et al., 2010; Li and Sharkey, 2013). In the case of aromatic plants such as *A. millefolium*, mono- and sesquiterpene contents of the essential oil are strongly correlated (Mockute and Judzentiene, 2003; Orav et al., 2006; Gudaityte and Venskutonis, 2007; Judzentiene and Mockute, 2010) and not necessarily correlated with the rate of carbon assimilation. Thus, the finding of mixed substrate specificity of AmGAS might indicate a more important role of cytosolic monoterpene synthesis in aromatic plants.

What could be the physiological significance of *cis*- vs. *trans*-isomers of substrates? Recently, a tomato monoterpene synthase has been sequenced that uses neryl diphosphate, the *cis*-isomer of GDP, as a substrate instead of GDP, to form several cyclic monoterpenes in trichomes (Schillmiller et al., 2009). Sallaud et al. (2009) further reported that a sesquiterpene synthase in tomato uses *Z,Z*-FDP instead of the usual *E,E*-FDP for the biosynthesis of type II sesquiterpenes in the trichome secretory cells. Clearly more work is needed to gain insight into the possible use of *cis*-substrates in species other than tomato.

#### PHYLOGENETIC ANALYSIS OF GERMACRENE A SYNTHASE

Multiple sequence alignment of AmGAS amino acid sequences with germacrene A synthases from other Asteraceae species showed high sequence similarity (**Figure 2**). The phylogenetic analysis showed a particularly close relationship between GAS from *A. millefolium* and *T. parthenium* and *A. annua* which is in accordance with high phylogenetic relatedness among these species (**Figure 1**). Germacrene A synthase from Asteraceae grouped in one single clad, which suggests a monophyletic origin of the gene. This is in agreement with the observation that occurrence of germacrene A is restricted to this family (Bouwmeester et al., 2002; Adio, 2009a; Majdi et al., 2011).

Germacrene A synthase is a two-domain,  $\alpha$ - $\beta$ -terpenoid synthase with the active center in  $\alpha$ -domain (C-terminus, 234–558

AA) exhibiting class I terpene synthase activity (Christianson, 2008; Rajabi et al., 2013).  $\beta$ -domain in N-terminus (32–245 AA) has lost the catalytic activity in mono- and sesquiterpene synthases, and seems to play a role in tertiary conformation of  $\alpha$ - $\beta$ -terpenoid synthases (Christianson, 2006; Aaron and Christianson, 2010). Thus, the 5' end of the GAS gene (N-terminus for the protein) shows considerable variation in gene structure and sequence which is in agreement with other two-domain,  $\alpha$ - $\beta$ -terpenoid synthase genes (Aubourg et al., 2002).

A number of conserved sequences of AmGAS with high homology to germacrene A synthase amino acids in other Asteraceae family members was detected. The second red box in **Figure 2** shows the conserved aspartate-rich motif of DDxxD (DDTYD Asteraceae family, position 316–320 AA) which is conserved in all plant terpenoid synthases (Steele et al., 1998a). The occurrence of this aspartate-rich motif (DDxxD) at the catalytic site is crucial in positioning the substrate for catalysis. Another metal binding motif is located on the opposite side of the active site (Christianson, 2006). This motif, designated as NSE/DTE motif, has apparently evolved from a second aspartate-rich motif conserved in prenyl transferases, although this NSE/DTE motif is less conserved in sesquiterpene synthases. In grand fir (*A. grandis*) sesquiterpene  $\delta$ -selinene and  $\gamma$ -humulene synthases, this motif is replaced by a second DDxxD motif (Steele et al., 1998a). Here we show that this motif is replaced by DDxxx (DDVMT) in germacrene A synthases of Asteraceae (the forth red box, position 460–464, **Figure 2**). This second DDxxD (or DDxxx) motif is also involved in catalysis (Steele et al., 1998b; Little and Croteau, 2002) and the formation of multiple products might be enhanced by this motif (Degenhardt et al., 2009).

In addition to these motifs, about 35 amino acids upstream of the first DDxxD motif there is a highly conserved arginine-rich, RxR (RDR in Asteraceae) motif (the first red box in **Figure 2**), that plays a role in the complexing of the diphosphate group after ionization of FDP (Starks et al., 1997). Also the third red box shows a conserved motif of TSA (position 416–418) that plays a substantial role in cyclization (Chang et al., 2005). High conservation of these motifs in germacrene A synthases from Asteraceae suggests that they have the same catalytic mechanism and are potentially mixed-substrate terpene synthases. Clearly further work with protein crystal structure is needed to gain insight into the catalysis of germacrene A synthases with different substrates and into the determinants of substrate specificity and product profiles with different substrates.

#### AUTHOR CONTRIBUTIONS

LP participated in designing and carrying out the experiments, analyzing the data and writing the manuscript; HM contributed to designing and describing the methods, interpreting the data and writing; AK, performed GC-MS analysis; RB, supported real time PCR experiment; ÜN contributed to designing and planning the experiment, interpreting the data and writing.

#### ACKNOWLEDGMENTS

The authors thank Dr. Mohammad Majdi for valuable suggestions during the experiments, Prof. Peter C. Harley for critical reading of the MS, Prof. Jarmo Holopainen (University of Eastern Finland, Kuopio, Finland) for the standard of

4,8-Dimethyl-1,3-*E*,7-nonatriene. This study was supported by the Estonian Ministry of Science and Education (institutional grant IUT-8-3), the European Commission through the European Regional Fund (the Center of Excellence in Environmental Adaptation), and the European Research Council (advanced grant 322603, SIP-VOL+).

## REFERENCES

- Aaron, J. A., and Christianson, D. W. (2010). Trinuclear metal clusters in catalysis by terpenoid synthases. *Pure Appl. Chem.* 82, 1585–1597. doi: 10.1351/PAC-CON-09-09-37
- Adam, K. P., Thiel, R., and Zapp, J. (1999). Incorporation of 1-[1-(13C)]deoxy-D-xylulose in chamomile sesquiterpenes. *Arch. Biochem. Biophys.* 369, 127–132. doi: 10.1006/abbi.1999.1346
- Adam, K. P., and Zapp, J. (1998). Biosynthesis of the isoprene units of chamomile sesquiterpenes. *Phytochemistry* 48, 953–959. doi: 10.1016/S0031-9422(97)00992-8
- Adams, R. P. (2001). *Identification of Essential Oil Components by Gas Chromatography/Quadrupole Mass Spectroscopy*. Carol Stream, IL: Allured Publishing.
- Adio, A. M. (2009a). Germacrene A–E and related compounds: thermal, photochemical and acid induced transannular cyclizations. *Tetrahedron* 65, 1533–1552. doi: 10.1016/j.tet.2008.11.050
- Adio, A. M. (2009b). (-)-Trans- $\beta$ -elemene and related compounds: occurrence, synthesis, and anticancer activity. *Tetrahedron* 65, 5145–5159. doi: 10.1016/j.tet.2009.04.062
- Arimura, G.-I., Matsui, K., and Takabayashi, J. (2009). Chemical and molecular ecology of herbivore-induced plant volatiles: proximate factors and their ultimate functions. *Plant Cell Physiol.* 50, 911–923. doi: 10.1093/pcp/pcp030
- Aubourg, S., Lechary, A., and Bohlmann, J. (2002). Genomic analysis of the terpenoid synthase (ATPS) gene family of *Arabidopsis thaliana*. *Mol. Genet. Genomics* 267, 730–745. doi: 10.1007/s00438-002-0709-y
- Bennett, M. H., Mansfield, J. W., Lewis, M. J., and Beale, M. H. (2002). Cloning and expression of sesquiterpene synthase genes from lettuce (*Lactuca sativa* L.). *Phytochemistry* 60, 255–261. doi: 10.1016/S0031-9422(02)00103-6
- Bertea, C. M., Voster, A., Verstappen, F. W. A., Maffei, M., Beekwilder, J., and Bouwmeester, H. J. (2006). Isoprenoid biosynthesis in *Artemisia annua*: cloning and heterologous expression of a germacrene A synthase from a glandular trichome cDNA library. *Arch. Biochem. Biophys.* 448, 3–12. doi: 10.1016/j.abb.2006.02.026
- Bick, J. A., and Lange, B. M. (2003). Metabolic cross talk between cytosolic and plastidial pathways of isoprenoid biosynthesis: unidirectional transport of intermediates across the chloroplast envelope membranes. *Arch. Biochem. Biophys.* 415, 146–154. doi: 10.1016/S0003-9861(03)00233-9
- Bohlmann, J., and Keeling, C. I. (2008). Terpenoid biomaterials. *Plant J.* 54, 656–669. doi: 10.1111/j.1365-313X.2008.03449.x
- Bohlmann, J., Meyer-Gauen, G., and Croteau, R. (1998). Plant terpenoid synthases: molecular biology and phylogenetic analysis. *Proc. Natl. Acad. Sci. U.S.A.* 95, 4126–4133. doi: 10.1073/pnas.95.8.4126
- Bouwmeester, H. J., Kodde, J., Verstappen, F. W. A., Altug, I. G., De Kraker, J.-W., and Wallaart, T. E. (2002). Isolation and characterization of two germacrene A synthase cDNA clones from chicory. *Plant Physiol.* 129, 134–144. doi: 10.1104/pp.001024
- Chandler, R. F., Hooper, S. N., and Harvey, M. J. (1982). Ethnobotany and phytochemistry of yarrow, *Achillea millefolium*, Compositae. *Econ. Bot.* 36, 203–223. doi: 10.1007/BF02858720
- Chang, Y.-J., Jin, J., Nam, H.-Y., and Kim, S.-U. (2005). Point mutation of (+)-germacrene A synthase from *Ixeris dentata*. *Biotechnol. Lett.* 27, 285–288. doi: 10.1007/s10529-005-0681-9
- Chappell, J., and Coates, R. M. (2010). “Sesquiterpenes,” in *Comprehensive Natural Products II*, eds M. Lew and L. Hung-Wen (Oxford: Elsevier), 609–641.
- Chen, F., Tholl, D., Bohlmann, J., and 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. doi: 10.1111/j.1365-313X.2011.04520.x
- Cheng, A.-X., Lou, Y.-G., Mao, Y.-B., Lu, S., Wang, L.-J., and Chen, X.-Y. (2007). Plant terpenoids: biosynthesis and ecological functions. *J. Integr. Plant Biol.* 49, 179–186. doi: 10.1111/j.1744-7909.2007.00395.x
- Christianson, D. W. (2006). Structural biology and chemistry of the terpenoid cyclases. *Chem. Rev.* 106, 3412–3442. doi: 10.1021/cr050286w
- Christianson, D. W. (2008). Unearthing the roots of the terpenome. *Curr. Opin. Chem. Biol.* 12, 141–150. doi: 10.1016/j.cbpa.2007.12.008
- Copolovici, L., Kännaste, A., Pazouki, L., and Niinemets, Ü. (2012). Emissions of green leaf volatiles and terpenoids from *Solanum lycopersicum* are quantitatively related to the severity of cold and heat shock treatments. *J. Plant. Physiol.* 169, 664–672. doi: 10.1016/j.jplph.2011.12.019
- Copolovici, L., Kännaste, A., Rimmel, T., Vislap, V., and 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. doi: 10.1007/s10886-010-9897-9
- Cornu, A., Carnat, A.-P., Martin, B., Coulon, J.-B., Lamaison, J.-L., and Berdague, J.-L. (2001). Solid-phase microextraction of volatile components from natural grassland plants. *J. Agric. Food Chem.* 49, 203–209. doi: 10.1021/jf0008341
- D’Alessandro, M., and Turlings, T. C. J. (2005). *In situ* modification of herbivore-induced plant odors: a novel approach to study the attractiveness of volatile organic compounds to parasitic wasps. *Chem. Senses* 30, 739–753. doi: 10.1093/chemse/bji066
- Davidovich-Rikanati, R., Lewinsohn, E., Bar, E., Iijima, Y., Pichersky, E., and Sitrin, Y. (2008). Overexpression of the lemon basil  $\alpha$ -zingiberene synthase gene increases both mono- and sesquiterpene contents in tomato fruit. *Plant J.* 56, 228–238. doi: 10.1111/j.1365-313X.2008.03599.x
- Degenhardt, J., Koellner, T. G., and Gershenzon, J. (2009). Monoterpene and sesquiterpene synthases and the origin of terpene skeletal diversity in plants. *Phytochemistry* 70, 1621–1637. doi: 10.1016/j.phytochem.2009.07.030
- De Kraker, J., Franssen, M., De Groot, A., König, W., and Bouwmeester, H. (1998). (+)-Germacrene A biosynthesis. The committed step in the biosynthesis of bitter sesquiterpene lactones in chicory. *Plant Physiol.* 117, 1381–1392. doi: 10.1104/pp.117.4.1381
- De Kraker, J.-W., Franssen, M. C. R., Dalm, M. C. F., De Groot, A., and Bouwmeester, H. J. (2001). Biosynthesis of germacrene A carboxylic acid in chicory roots. Demonstration of a cytochrome P450 (+)-germacrene A hydroxylase and NADP<sup>+</sup>-dependent sesquiterpenoid dehydrogenase(s) involved in sesquiterpene lactone biosynthesis. *Plant Physiol.* 125, 1930–1940. doi: 10.1104/pp.125.4.1930
- Dias, M. I., Barros, L., Duenas, M., Pereira, E., Carvalho, A. M., Alves, R. C., et al. (2013). Chemical composition of wild and commercial *Achillea millefolium* L. and bioactivity of the methanolic extract, infusion and decoction. *Food Chem.* 141, 4152–4160. doi: 10.1016/j.foodchem.2013.07.018
- Dudareva, N., Negre, F., Nagegowda, D. A., and Orlova, I. (2006). Plant volatiles: recent advances and future perspectives. *CRC Crit. Rev. Plant Sci.* 25, 417–440. doi: 10.1080/07352680600899973
- Dudareva, N., Pichersky, E., and Gershenzon, J. (2004). Biochemistry of plant volatiles. *Plant Physiol.* 135, 1893–1902. doi: 10.1104/pp.104.049981
- Fineschi, S., Loreto, F., Staudt, M., and Peñuelas, J. (2013). “Diversification of volatile isoprenoid emissions from trees: evolutionary and ecological perspectives,” in *Biology, Controls and Models of Tree Volatile Organic Compound Emissions*, eds Ü. Niinemets and R. K. Monson (Berlin: Springer), 1–20.
- Fischbach, R. J., Zimmer, W., and Schnitzler, J.-P. (2001). Isolation and functional analysis of a cDNA encoding a myrcene synthase from holm oak (*Quercus ilex* L.). *Euro. J. Biochem.* 268, 5633–5638. doi: 10.1046/j.1432-1033.2001.02519.x
- Glasl, S., Mucaji, P., Werner, I., Presser, A., and Jurenitsch, J. (2002). Sesquiterpenes and flavonoid aglycones from a Hungarian taxon of the *Achillea millefolium* group. *J. Biosci.* 57, 976–982. doi: 10.1515/znc-2002-11-1203
- Göpfert, J., Buelow, A.-K., and Spring, O. (2010). Identification and functional characterization of a new sunflower germacrene A synthase (HaGAS3). *Nat. Prod. Commun.* 5, 709–715.
- Göpfert, J. C., Macnevin, G., Ro, D.-K., and Spring, O. (2009). Identification, functional characterization and developmental regulation of sesquiterpene synthases from sunflower capitate glandular trichomes. *BMC Plant Biol.* 9:86. doi: 10.1186/1471-2229-9-86
- Green, S., Friel, E. N., Matich, A., Beuning, L. L., Cooney, J. M., Rowan, D. D., et al. (2007). Unusual features of a recombinant apple  $\alpha$ -farnesene synthase. *Phytochemistry* 68, 176–188. doi: 10.1016/j.phytochem.2006.10.017

- Gudaityte, O., and Venskutonis, P. R. (2007). Chemotypes of *Achillea millefolium* transferred from 14 different locations in Lithuania to the controlled environment. *Biochem. Syst. Ecol.* 35, 582–592. doi: 10.1016/j.bsc.2007.03.016
- Guenther, A., Geron, C., Pierce, T., Lamb, B., Harley, P., and Fall, R. (2000). Natural emissions of non-methane volatile organic compounds, carbon monoxide, and oxides of nitrogen from North America. *Atmos. Environ.* 34, 2205–2230. doi: 10.1016/S1352-2310(99)00465-3
- Guenther, A., Hewitt, C. N., Erickson, D., Fall, R., Geron, C., Graedel, T., et al. (1995). A global model of natural volatile organic compound emissions. *J. Geophys. Res. Atmos.* 100, 8873–8892. doi: 10.1029/94JD02950
- Gutensohn, M., Orlova, I., Nguyen, T. T. H., Davidovich-Rikanati, R., Ferruzzi, M. G., Sitrit, Y., et al. (2013). Cytosolic monoterpene biosynthesis is supported by plastid-generated geranyl diphosphate substrate in transgenic tomato fruits. *Plant J.* 75, 351–363. doi: 10.1111/tpj.12212
- Hemmerlin, A., Hoefler, J. F., Meyer, O., Tritsch, D., Kagan, I. A., Grosdemange-Billiard, C., et al. (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. doi: 10.1074/jbc.M302526200
- Huang, M., Sanchez-Moreiras, A. M., Abel, C., Sohrabi, R., Lee, S., Gershenzon, J., et al. (2012). The major volatile organic compound emitted from *Arabidopsis thaliana* flowers, the sesquiterpene (E)- $\beta$ -caryophyllene, is a defense against a bacterial pathogen. *New Phytol.* 193, 997–1008. doi: 10.1111/j.1469-8137.2011.04001.x
- Irmisch, S., Krause, S. T., Kunert, G., Gershenzon, J., Degenhardt, J., and Koellner, T. G. (2012). The organ-specific expression of terpene synthase genes contributes to the terpene hydrocarbon composition of chamomile essential oils. *BMC Plant Biol.* 12:84. doi: 10.1186/1471-2229-12-84
- Judzentiene, A., and Mockute, D. (2010). Essential oil composition of two yarrow taxonomic forms. *Cent. Eur. J. Biol.* 5, 346–352. doi: 10.2478/s11535-010-0011-7
- Junker, R. R., and Bluethgen, N. (2008). Floral scents repel potentially nectar-thieving ants. *Evol. Ecol. Res.* 10, 295–308.
- Keasling, J. D. (2012). Synthetic biology and the development of tools for metabolic engineering. *Metab. Eng.* 14, 189–195. doi: 10.1016/j.ymben.2012.01.004
- Keeling, C. I., and Bohlmann, J. (2006). Diterpene resin acids in conifers. *Phytochemistry* 67, 2415–2423. doi: 10.1016/j.phytochem.2006.08.019
- Kim, M. Y., Chang, Y. J., Bang, M. H., Baek, N. I., Jin, J., Lee, C. H., et al. (2005). cDNA isolation and characterization of (+)-germacrene A synthase from *Ixeris dentata* form *albiflora* Hara. *J. Plant Biol.* 48, 178–186. doi: 10.1007/BF03030406
- Kovacevic, N., Pavlovic, M., Menkovic, N., Tzakou, O., and Couladis, M. (2002). Composition of the essential oil from roots and rhizomes of *Valeriana panceicii* Halacsy & Bald. *Flavour Fragr. J.* 17, 355–357. doi: 10.1002/ffj.1100
- Landmann, C., Fink, B., Festner, M., Dregus, M., Engel, K.-H., and Schwab, W. (2007). Cloning and functional characterization of three terpene synthases from lavender (*Lavandula angustifolia*). *Arch. Biochem. Biophys.* 465, 417–429. doi: 10.1016/j.abb.2007.06.011
- Laule, O., Fürholz, A., Chang, H. S., Zhu, T., Wang, X., Heifetz, P. B., et al. (2003). Crosstalk between cytosolic and plastidial pathways of isoprenoid biosynthesis in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A.* 100, 6866–6871. doi: 10.1073/pnas.1031755100
- Li, Z., and Sharkey, T. D. (2013). “Molecular and pathway controls on biogenic volatile organic compound emissions,” in *Biology, Controls and Models of Tree Volatile Organic Compound Emissions*, eds Ü. Niinemets and R.K. Monson. (Berlin: Springer), 119–151.
- Little, D. B., and Croteau, R. B. (2002). Alteration of product formation by directed mutagenesis and truncation of the multiple-product sesquiterpene synthases  $\delta$ -selinene synthase and  $\gamma$ -humulene synthase. *Arch. Biochem. Biophys.* 402, 120–135. doi: 10.1016/S0003-9861(02)00068-1
- Lyakina, M. N. (2002). Study of the total terpenoid composition in *millefolium* homopathic tinctures prepared from fresh and dry material. *Pharm. Chem. J.* 36, 437–439. doi: 10.1023/A:1021266712077
- Majdi, M., Liu, Q., Karimzadeh, G., Malboobi, M. A., Beekwilder, J., Cankar, K., et al. (2011). Biosynthesis and localization of parthenolide in glandular trichomes of feverfew (*Tanacetum parthenium* (L.) Schulz Bip.). *Phytochemistry* 72, 1739–1750. doi: 10.1016/j.phytochem.2011.04.021
- Maloukh, L., Matousek, J., Van Bockstaele, E., and Roldán-Ruiz, I. (2009). Housekeeping gene selection for real time-PCR normalization in female hop (*Humulus lupulus* L.) tissues. *J. Plant Biochem. Biot.* 18, 53–58. doi: 10.1007/BF03263295
- McAndrew, R. P., Peralta-Yahya, P. P., Degiovanni, A., Pereira, J. H., Hadi, M. Z., Keasling, J. D., et al. (2011). Structure of a three-domain sesquiterpene synthase: a prospective target for advanced biofuels production. *Structure* 19, 1876–1884. doi: 10.1016/j.str.2011.09.013
- Mockute, D., and Judzentiene, A. (2003). Variability of the essential oils composition of *Achillea millefolium* ssp. *millefolium* growing wild in Lithuania. *Biochem. Syst. Ecol.* 31, 1033–1045. doi: 10.1016/S0305-1978(03)00066-8
- Montsko, G., Boros, B., Takatsy, A., Ohmacht, R., Glasl, S., Krenn, L., et al. (2008). Separation of sesquiterpenes from yarrow (*Achillea millefolium* L. s. l.) by LC-MS on non-porous columns. *Chromatographia* 67, 467–470. doi: 10.1365/s10337-008-0520-y
- Morse, A., Kevan, P., Shipp, L., Khosla, S., and Mcgarvey, B. (2012). The impact of greenhouse tomato (Solanales: Solanaceae) floral volatiles on bumble bee (Hymenoptera: Apidae) pollination. *Environ. Entomol.* 41, 855–864. doi: 10.1603/EN11234
- Nagegowda, D. A. (2010). Plant volatile terpenoid metabolism: biosynthetic genes, transcriptional regulation and subcellular compartmentation. *FEBS Lett.* 584, 2965–2973. doi: 10.1016/j.febslet.2010.05.045
- Nemeth, E., and Bernath, J. (2008). Biological activities of yarrow species (*Achillea* spp.). *Curr. Pharm. Des.* 14, 3151–3167. doi: 10.2174/138161208786404281
- Nguyen, D. T., Goepfert, J. C., Ikezawa, N., Macnevin, G., Kathiresan, M., Conrad, J., et al. (2010). Biochemical conservation and evolution of germacrene A oxidase in Asteraceae. *J. Biol. Chem.* 285, 16588–16598. doi: 10.1074/jbc.M110.111757
- Nicot, N., Hausman, J. F., Hoffmann, L., and Evers, D. (2005). Housekeeping gene selection for real-time RT-PCR normalization in potato during biotic and abiotic stress. *J. Exp. Bot.* 56, 2907–2914. doi: 10.1093/jxb/eri285
- Niinemetts, Ü., Kuhn, U., Harley, P. C., Staudt, M., Arneith, A., Cescatti, A., et al. (2011). Estimations of isoprenoid emission capacity from enclosure studies: measurements, data processing, quality and standardized measurement protocols. *Biogeosciences* 8, 2209–2246. doi: 10.5194/bg-8-2209-2011
- Niinemetts, Ü., Monson, R. K., Arneith, A., Ciccioli, P., Kesselmeier, J., Kuhn, U., et al. (2010). The leaf-level emission factor of volatile isoprenoids: caveats, model algorithms, response shapes and scaling. *Biogeosciences* 7, 1809–1832. doi: 10.5194/bg-7-1809-2010
- Orav, A., Arak, E., and Raal, A. (2006). Phytochemical analysis of the essential oil of *Achillea millefolium* L. from various European Countries. *Nat. Prod. Res.* 20, 1082–1088. doi: 10.1080/14786410500510849
- Paddon, C. J., Westfall, P. J., Pitera, D. J., Benjamin, K., Fisher, K., McPhee, D., et al. (2013). High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature* 496, 528–532. doi: 10.1038/nature12051
- Pichersky, E., and Gershenzon, J. (2002). The formation and function of plant volatiles: perfumes for pollinator attraction and defense. *Curr. Opin. Chem. Biol.* 5, 237–243. doi: 10.1016/S1369-5266(02)00251-0
- Pichersky, E., Noel, J. P., and Dudareva, N. (2006). Biosynthesis of plant volatiles: nature's diversity and ingenuity. *Science* 311, 808–811. doi: 10.1126/science.1118510
- Picman, A. K. (1986). Biological activities of sesquiterpene lactones. *Biochem. Syst. Biol.* 14, 255–281.
- Prosser, I., Phillips, A. L., Gittings, S., Lewis, M. J., Hooper, A. M., Pickett, J. A., et al. (2002). (+)-(-10R)-Germacrene A synthase from goldenrod, *Solidago canadensis*: cDNA isolation, bacterial expression and functional analysis. *Phytochemistry* 60, 691–702. doi: 10.1016/S0031-9422(02)00165-6
- Raal, A., Orav, A., and Arak, E. (2012). Essential oil content and composition in commercial *Achillea millefolium* L. herbs from different countries. *J. Essent. Oil Bear Pl.* 15, 22–31. doi: 10.1080/0972060X.2012.10644015
- Rajabi, M. H., Pazouki, L., and Niinemets, Ü. (2013). “The biochemistry and molecular biology of volatile messengers in trees,” in *Biology, Controls and Models of Tree Volatile Organic Compound Emissions*, eds Ü. Niinemets and R. K. Monson (Dordrecht: Springer Netherlands), 47–93.
- Rasmann, S., Köllner, T. G., Degenhardt, J., Hiltbold, I., Toepfer, S., Kuhlmann, U., et al. (2005). Recruitment of entomopathogenic nematodes by insect-damaged maize roots. *Nature* 434, 732–737. doi: 10.1038/nature03451
- Reiling, K. K., Yoshikuni, Y., Martin, V. J. J., Newman, J., Bohlmann, J., and Keasling, J. D. (2004). Mono and diterpene production in *Escherichia coli*. *Biotechnol. Bioeng.* 87, 200–212. doi: 10.1002/bit.20128
- Ren, J., Liu, Y.-Q., Yang, L., Ni, R., and You, S. (2006). Molecular cloning and bacterial expression of germacrene A synthase cDNA from *Crepidiastrum sonchifolium*. *Chem. Res. Chin. U* 22, 606–611. doi: 10.1016/S1005-9040(06)60172-4

- Ro, D. K., Ehrling, J., Keeling, C. I., Lin, R., Mattheus, N., and Bohlmann, J. (2006). Microarray expression profiling and functional characterization of AtTPS genes: duplicated *Arabidopsis thaliana* sesquiterpene synthase genes At4g13280 and At4g13300 encode root-specific and wound-inducible (Z)- $\gamma$ -bisabolene synthases. *Arch. Biochem. Biophys.* 448, 104–116. doi: 10.1016/j.abb.2005.09.019
- Rodriguez-Saona, C., Crafts-Brander, S. J., Pare, P. W., and Henneberry, T. J. (2001). Exogenous methyl jasmonate induces volatile emissions in cotton plants. *J. Chem. Ecol.* 27, 679–695. doi: 10.1023/A:1010393700918
- Sallaud, C., Rontein, D., Onillon, S., Jabes, F., Duffe, P., Giacalone, C., et al. (2009). A novel pathway for sesquiterpene biosynthesis from Z,Z-farnesyl pyrophosphate in the wild tomato *Solanum habrochaites*. *Plant Cell* 21, 301–317. doi: 10.1105/tpc.107.057885
- Sambrook, J., and Russell, D. W. (2001). *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory
- Scala, A., Allmann, S., Mirabella, R., Haring, M. A., and Schuurink, R. C. (2013). Green leaf volatiles: a plant's multifunctional weapon against herbivores and pathogens. *Int. J. Mol. Sci.* 14, 17781–17811. doi: 10.3390/ijms140917781
- Schilmiller, A. L., Schauvinhold, I., Larson, M., Xu, R., Charbonneau, A. L., Schmidt, A., et al. (2009). Monoterpenes in the glandular trichomes of tomato are synthesized from a neryl diphosphate precursor rather than geranyl diphosphate. *Proc. Natl. Acad. Sci. U.S.A.* 106, 10865–10870. doi: 10.1073/pnas.0904113106
- Schmittgen, T. D., and Livak, K. J. (2008). Analyzing real-time PCR data by the comparative C(T) method. *Nat. Protoc.* 3, 1101–1108. doi: 10.1038/nprot.2008.73
- Schnee, C., Köllner, T. G., Gershenzon, J., and Degenhardt, J. (2002). The maize gene terpene synthase 1 encodes a sesquiterpene synthase catalyzing the formation of (E)- $\beta$ -farnesene, (E)-nerolidol, and (E,E)-farnesol after herbivore damage. *Plant Physiol.* 130, 2049–2060. doi: 10.1104/pp.008326
- Si, X.-T., Zhang, M.-L., Shi, Q.-W., and Kiyota, H. (2006). Chemical constituents of the plants in the genus *Achillea*. *Chem. Biodivers.* 3, 1163–1180. doi: 10.1002/cbdv.200690119
- Starks, C. M., Back, K. W., Chappell, J., and Noel, J. P. (1997). Structural basis for cyclic terpene biosynthesis by tobacco 5-epi-aristolochene synthase. *Science* 277, 1815–1820. doi: 10.1126/science.277.5333.1815
- Steele, C. L., Crock, J., Bohlmann, J., and Croteau, R. (1998a). Sesquiterpene synthases from grand fir (*Abies grandis*) - Comparison of constitutive and wound-induced activities, and cDNA isolation, characterization and bacterial expression of  $\delta$ -selinene synthase and  $\gamma$ -humulene synthase. *J. Biol. Chem.* 273, 2078–2089. doi: 10.1074/jbc.273.4.2078
- Steele, C. L., Katoh, S., Bohlmann, J., and Croteau, R. (1998b). 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. doi: 10.1104/pp.116.4.1497
- Stewart-Jones, A., and Poppy, G. M. (2006). Comparison of glass vessels and plastic bags for enclosing living plant parts for headspace analysis. *J. Chem. Ecol.* 32, 845–864. doi: 10.1007/s10886-006-9039-6
- Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28, 2731–2739. doi: 10.1093/molbev/msr121
- Tholl, D. (2006). Terpene synthases and the regulation, diversity and biological roles of terpene metabolism. *Curr. Opin. Chem. Biol.* 9, 297–304. doi: 10.1016/j.pbi.2006.03.014
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Received: 03 December 2014; accepted: 11 February 2015; published online: 03 March 2015.
- Citation: Pazouki L, Memari HR, Kännaste A, Bichele R and Niinemets Ü (2015) Germacrene A synthase in yarrow (*Achillea millefolium*) is an enzyme with mixed substrate specificity: gene cloning, functional characterization and expression analysis. *Front. Plant Sci.* 6:111. doi: 10.3389/fpls.2015.00111
- This article was submitted to *Plant Metabolism and Chemodiversity*, a section of the journal *Frontiers in Plant Science*.
- Copyright © 2015 Pazouki, Memari, Kännaste, Bichele and Niinemets. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.