Horticulture Research

Horticulture Research, 2022, **9**: uhab020
https://doi.org/10.1093/hr/uhab020

Article

Characterization of Terpene synthase variation in flowers of wild *Aquilegia* species from Northeastern Asia

Song Yang[1], Ning Wang[1], Shadrack Kimani[1,4], Yueqing Li[1], Tingting Bao[1], Guogui Ning[2], Linfeng Li[3], Bao Liu[1], Li Wang[1,*] and Xiang Gao[1,*]

[1]Key Laboratory of Molecular Epigenetics of MOE and Institute of Genetics & Cytology, Northeast Normal University, Changchun 130024, China
[2]Key Laboratory of Horticultural Plant Biology, Ministry of Education, College of Horticulture and Forestry Sciences, Huazhong Agricultural University, Wuhan 430070, China
[3]Ministry of Education Key Laboratory for Biodiversity Science and Ecological Engineering, Institute of Biodiversity Science, School of Life Sciences, Fudan University, Shanghai 200438, China
[4]School of Pure and Applied Sciences, Karatina University, Karatina, Kenya
*Corresponding authors. E-mail: gaoxiang424@163.com; wanglee57@163.com

Abstract

There are several causes for the great diversity in floral terpenes. The terpene products are determined by the catalytic fidelity, efficiency and plasticity of the active sites of terpene synthases (TPSs). However, the molecular mechanism of TPS in catalyzing terpene biosynthesis and its evolutionary fate in wild plant species remain largely unknown. In this study, the functionality of terpene synthases and their natural variants were assessed in two Northeastern Asia endemic columbine species and their natural hybrid. Synoptically, *TPS7*, *TPS8*, and *TPS9* were highly expressed in these *Aquilegia* species from the Zuojia population. The *in vitro* and *in vivo* enzymatic assays revealed that *TPS7* and *TPS8* mainly produced (+)-limonene and β-sesquiphellandrene, respectively, whereas *TPS9* produced pinene, similar to the major components released from *Aquilegia* flowers. Multiple sequence alignment of *Aquilegia* TPS7 and TPS8 in the Zuojia population revealed amino acid polymorphisms. Domain swapping and amino acid substitution assays demonstrated that 413A, 503I and 529D had impacts on TPS7 catalytic activity, whereas 420G, 538F and 545 L affected the ratio of β-sesquiphellandrene to β-bisabolene in TPS8. Moreover, these key polymorphic amino acid residues were found in *Aquilegia* species from the Changbai Mountain population. Interestingly, amino acid polymorphisms in TPSs were present in individuals with low expression levels, and nonsynonymous mutations could impact the catalytic activity or product specificity of these genes. The results of this study will shed new light on the function and evolution of *TPS* genes in wild plant species and are beneficial to the modification of plant fragrances.

Introduction

Volatile terpenes, mainly consisting of monoterpenes and sesquiterpenes, are among the most abundant components of floral scent and play crucial roles in plant development and chemical ecology, especially in the interactions between plants and stresses [1–6]. The structurally diverse terpenes found in nature result from the mid-sized terpene synthase (TPS) family [7, 8], which directly converts geranyl diphosphate (GPP), neryl diphosphate (NPP) or farnesyl diphosphate (FPP) into versatile mono−/sesqui-terpenes [1–6]. In evolutionary history, *TPS* genes have undergone species-specific duplication and divergence, which contributes to the plasticity (the ability of TPS to use a broad range of substrates to generate multiple products) and adaptation of terpene metabolism [7, 9, 10]. Generally, a landmark feature of *TPS* genes is that they often lie in clusters of tandemly duplicated genes, including both functional genes and pseudogenes, evidently demonstrates that duplicated *TPS* genes have possibly led to phenotypic variations, although instances of convergent evolution have been observed in maize TPS23 [7, 11–14]. Chen et al. 2011 concluded that the TPSs from angiosperms and gymnosperms could be cataloged into seven clades, effectively supporting the evident evolutionary plasticity of the TPS family [7].

Recent studies on the functional characterization of several plant *TPS* genes have revealed the existence of remarkable plasticity in terpenoid biosynthesis in higher plants [5, 7, 12, 13, 15–22]; thus, the prediction of enzymatic products solely based on the protein similarity of terpene synthases is often difficult or impossible. Broadly, typical TPSs possess conserved structural features, such as $RR(X)_8W$ in the N-terminus and a metal cofactor binding domain, such as DDXXD or NSE/DTE in the C-terminus [7, 12, 15, 23]. However, these conserved motifs are necessary only for their catalytic features and cannot fully explain the diversity or infidelity (the ability of a TPS to generate multiple

Received: March 21, 2021; **Accepted:** October 2, 2021; **Published:** 20 January 2022

products) of the enzymatic products of TPS proteins. In contrast, other amino acids within or adjacent to the active site might play even more important roles in determining product diversity or infidelity, as even a single amino acid variation in or near the active site was found to have dramatic effects on catalytic activity and product specificity [24]. Previously, domain exchange and point mutation assays have been utilized to identify the roles of residues that affect product profiles in specific regions [25–29], suggesting that TPS evolution is likely centered on the recombination of functional domains or substitution of key amino acids. Despite those attempts to understand the product specificity or infidelity driven by key amino acids in TPSs, little is known about how TPSs catalytically function in closely related wild plant species [14, 24].

Aquilegia (Columbine) is a perennial *Ranunculus* genus comprising over 70 closely related species that occupy a wide range of ecosystems across Asia, Europe and North America. *Aquilegia* is best known for its floral spurred petals [30], and the genus is widely cultivated as an ornamental plant. The genus consists of very closely related taxa with small and almost identical genomes [31, 32]; thus, the expansion of genomic resources for a single genus lays a strong foundation for understanding the biochemistry and genetics from numerous aspects. *Aquilegia* has been employed in countless ecological and evolutionary studies, including those on herkogamy, radiation speciation due to pollinator shifts, mating system evolution, floral development, and floral color evolution [32–38]. In addition to floral color and floral shape, floral scents acting as olfactory signals also play important roles in attracting pollinators and are beneficial for long-distance attraction [39, 40]. However, no effort has been devoted to understanding the molecular mechanisms that contribute to floral volatile diversification within *Aquilegia* species.

Aquilegia oxysepala and *Aquilegia japonica* have been identified as geographically representative *Aquilegia* species in Northeastern Asia [41]. In this study, two *Aquilegia* species and their natural hybrid from Changbai Mountain (42.4°N, 128.15°E) and Zuojia District (43.87°N, 126.57°E) were employed to functionally characterize *TPS* genes and evaluate the basic amino acids driving the biosynthesis of diverse floral volatile terpenes. The expression levels and enzymatic products of three TPS proteins perfectly mirrored the major volatile terpenes released from *Aquilegia* flowers. Conceivably, nonsynonymous amino acid mutations were observed when the expression levels of the *TPS* genes changed, and several key amino acids were identified to account for the catalytic activity and product specificity of *Aquilegia* TPS proteins. Our findings provide a foundation for further exploration of the functional characterization and evolution of *TPS* genes in wild plant species and pave the way to deciphering the roles of floral volatile terpenes in the fitness and radiation speciation of *Aquilegia* species.

Results

Isolation of *TPS*s involved in floral volatile terpene synthesis of *Aquilegia* species endemic to Northeastern Asia

According to Chen et al. 2011, TPSs responsible for volatile terpenes usually cluster into the TPS-a subclade composed of sesquiterpene synthases, the TPS-b subclade mainly responsible for monoterpene synthesis and the TPS-g family, including TPSs that function in both monoterpene and sesquiterpene synthesis in angiosperms [7]. A total of forty-two (42) *TPS*s were identified from the genome of *Aquilegia coerulea*. In detail, ten (10), twenty-seven (27) and five (5) fell into the TPS-a, TPS-b and TPS-g subclades, respectively (Fig. 1a). To identify the floral *TPS*s expressed in *Aquilegia* species endemic to Northeastern Asia, flowers of *A. oxysepala*, *A. japonica* and their hybrid (Fig. S1) were collected at different developmental stages (S1-S4) from Zuojia (43.87°N, 126.57°E) and subjected to RNA-Seq. As a result, most *TPSs* had relatively similar but low transcript levels, suggesting that they hold minor roles in floral volatile terpene biosynthesis (Fig. S2). In addition, quantitative real-time PCR (qPCR) was carried out to confirm the *TPS* expression patterns in the fully opened flowers (S4). Consistent with the RNA-Seq results, *TPS7*, *TPS8*, *TPS9* and *TPS27* showed remarkably high expression levels, indicating that their encoded proteins might have pivotal roles in determining the specific volatile terpenes detected in *Aquilegia* flowers (Fig. 1b, S3).

To further characterize the TPSs, the full-length cDNA sequences of the most highly expressed *TPS7* and *TPS8* were taken as representatives and isolated from seven (7) individuals of *A. oxysepala* or *A. japonica* and one (1) natural hybrid. Amino acid sequence alignment of TPS7 and TPS8 revealed conserved DDXX(D/E) and (N,D)DXX(S,T,G) XXXE (NSE/DTE), which functioned in binding Mg^{2+} or Mn^{2+} cofactors, and terpene cyclization related $RR(X)_8W$ (Figs. S4, S5). Interestingly, residue polymorphisms were observed in the TPSs, indicating that the polymorphisms might result in functional differences in these *Aquilegia* proteins. The biosynthesis of monoterpenes and sesquiterpenes is thought to be compartmentalized, with monoterpenes being produced in the plastids where GPP or NPP is synthesized and sesquiterpenes being formed in the cytosol where FPP is generated [7, 42]. Consistent with this regularity, AoTPS7 and AoTPS9 were confined to plastids, similar to most known monoterpene synthases, while AjTPS8 and AoTPS27 were localized in the cytoplasm, suggesting their different roles in plant terpene synthesis (Fig. 1c).

The expression levels of *TPS*s are consistent with the volatile terpene emissions of *Aquilegia* flowers

Intricate floral-scent bouquets occur in many plant species. In the *A. oxysepala* population from the Zuojia location, the leading floral volatile terpenes identified

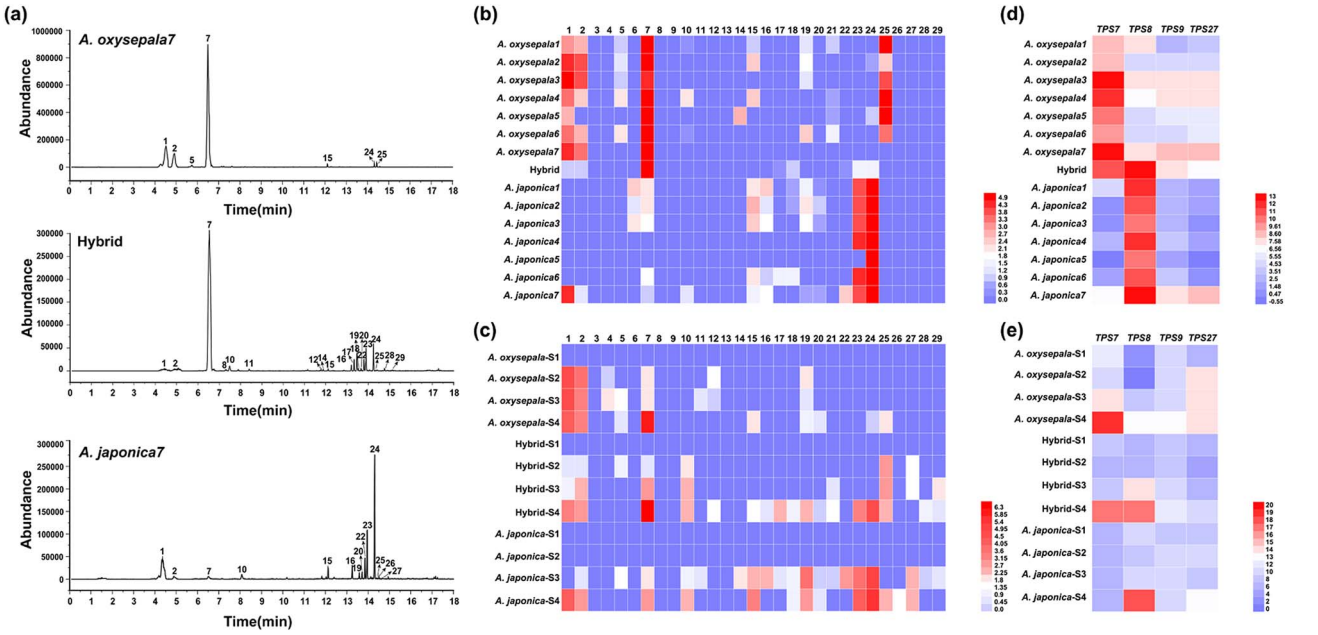


Figure 2. Volatile terpene emissions and TPS expression of Aquilegia individuals from the Zuojia population. a Floral volatile terpene profiles of representative Aquilegia individuals. b Volatile terpenes emitted from fully blooming flowers of Aquilegia individuals. Data represent the relative peak area in specific individuals. c Volatile terpenes released from flowers at different developmental stages of representative A. oxysepala 7#, A. japonica 7#, and their hybrid. Data represent the relative content calculated by the standard curve of (+)-limonene. d Relative expression levels of TPS7, TPS8, TPS9 and TPS27 in flowers of Aquilegia individuals. e Relative expression levels of TPS7, TPS8, TPS9 and TPS27 in flowers at different developmental stages of representative A. oxysepala 7#, A. japonica 7#, and their hybrid. The transcripts were normalized by AqIPP2 (GenBank KC854337) and compared with the lowest expression level of TPS in specific species. All the data were calculated as log2. Red and blue boxes indicate high and low expression levels, respectively. The detailed terpene contents are listed in Table S1 and Table S2. Here, 1 to 29 represent different volatiles: 1, (1R)-(+)-α-pinene; 2, (1S)-(−)-β-pinene; 3, α-phellandrene; 4, β-phellandrene; 5, β-pinene; 6, myrcene; 7, (+)-limonene; 8, 3-carene; 9, (E)-β-ocimene; 10, linalool; 11, α-terpineol; 12, an unidentified sesquiterpene; 13, (−)-α-copaene; 14, α-bergamotene; 15, an unidentified sesquiterpene; 16, (Z)-β-farnesene; 17, an unidentified sesquiterpene; 18, α-curcumene; 19, (E)-β-farnesene; 20, himachalene; 21, β-caryophyllene; 22, zingiberene; 23, β-bisabolene; 24, β-sesquiphellandrene; 25, α-caryophyllene; 26, an unidentified sesquiterpene; 27, cubebene; 28, α-muurolene; and 29, an unidentified sesquiterpene.

Comparatively, the relatively lower transcript levels of *TPS9* and *TPS27* might correlate with the relatively lower pinene or caryophyllene levels (Fig. 2d). Moreover, the transcripts of the four *Aquilegia TPS* genes increased with flower development (Fig. 2e). To validate the results, transcripts of the highly expressed *TPS7, TPS8* and *TPS9* were also normalized by *β-actin,* and similar expression patterns were observed (Fig. S8). Together, the results indicated that the four *Aquilegia* TPSs were mainly responsible for the main volatile terpenes released from *Aquilegia* flowers.

The *in vitro* enzymatic products of AoTPS7, AjTPS8 and AoTPS9 mirror the major volatile terpenes released from *Aquilegia* flowers

It is usually challenging to define the direct involvement of each *TPS* gene in the detected floral volatile blend unless *TPS*-encoded enzymes are functionally evaluated in detail. To elucidate the functionality of TPS proteins and their roles in the biosynthesis of the emitted floral volatile terpenes in *Aquilegia* species, the highly expressed *AoTPS7* (*A. oxysepala TPS7*), *AjTPS8* (*A. japonica TPS8*), *AoTPS9* (*A. oxysepala TPS9*) and *AoTPS27* (*A. oxysepala TPS27*) from *A. oxysepala* or *A. japonica* were cloned and subjected to further biochemical analysis. To prepare recombinant proteins, the four *TPS* genes were expressed

in the *E. coli* BL21(DE3) strain. Recombinant AoTPS7, AjTPS8, and AoTPS9 were successfully purified, whereas AoTPS27 failed to be expressed as a soluble protein (Fig. 3a). Enzymatic assays revealed that the recombinant proteins could convert GPP, NPP, (*E,E*)-FPP or (*Z,Z*)-FPP into a variety of products, suggesting that these terpene synthases were generally versatile enzymes (Fig. 3b–3d; Table S4). In detail, AoTPS7 could utilize GPP and NPP as substrates to generate (+)-limonene, accounting for 97% (15238.96 ng) and 99% (8759.98 ng) of the total products, respectively. In addition, AoTPS7 could transform (*E,E*)-FPP rather than (*Z,Z*)-FPP into (+)-α-longipinene (4.62 ng, 20%), α-bergamotene (3.10 ng, 14%) and other sesquiterpenes (Fig. 3b). AjTPS8 could catalyze GPP into (+)-limonene (3767.41 ng, 40%), β-pinene (1310.7 ng, 14%), (1*S*)-(−)-β-pinene (940.74 ng, 10%) and (1*R*)-(+)-α-pinene (818.95 ng, 9%) in order of amount. When NPP was used as the substrate, (+)-limonene (2272.94 ng, 62%), (1*S*)-(−)-β-pinene (391.75 ng, 11%) and (1*R*)-(+)-α-pinene (303.35 ng, 8%) were among the main products. Using (*E,E*)-FPP as a substrate, β-sesquiphelldrene (3192.4 ng, 52%) was identified as the most abundant sesquiterpene, followed by β-bisabolene (1231.19 ng, 20%), an unidentified sesquiterpene (792.38 ng, 13%) and zingiberene (359.64 ng, 6%). When (*Z,Z*)-FPP was used as the substrate, β-bisabolene (341.78 ng, 27%), an unidentified sesquiterpene (241.33 ng, 19%), α-himachalene (177.72 ng, 14%), β-sesquiphellandrene (199.15 ng, 16%) and α-caryophyllene (122.6 ng, 10%) were produced (Fig. 3c). AoTPS9 could convert GPP to (1*S*)-(−)-β-pinene derivatives (4923.49 ng, 55%), (1*R*)-(+)-α-pinene (1356.23 ng, 15%) and other monoterpenes. Relatively higher quantities of (1*S*)-(−)-β-pinene (2237.77 ng, 60%) and an unidentified sesquiterpene (36.43 ng, 72%) were produced by AoTPS9 when using NPP and (*E,E*)-FPP as substrates, respectively (Fig. 3d). Similar to AoTPS7, AoTPS9 could not catalyze (*Z,Z*)-FPP into any detected sesquiterpenes (Fig. 3d). In contrast, proteins extracted from E. coli BL21(DE3) expressing the empty pET32a vector could not catalyze GPP, NPP, (*E,E*)-FPP or (*Z,Z*)-FPP (Fig. S9).

In summary, the results indicated that the recombinant *Aquilegia* TPS proteins displayed different catalytic features, and the enzymatic products were basically consistent with major volatile compounds emitted by *Aquilegia* flowers. Notably, TPS7, TPS8 and TPS9 might mainly control the synthesis of (+)-limonene, β-sesquiphellandrene and pinene isomers, respectively, considering their subcellular localizations and available substrates in *Aquilegia* species.

The catalytic activity of the *Aquilegia* TPS7 protein is largely affected by polymorphic amino acids

As mentioned above, amino acid polymorphisms were found among TPS proteins from different *Aquilegia* individuals, which might affect the catalytic properties of TPSs. For example, sequences of the highly expressed AoTPS7 and HTPS7 (hybrid TPS7) were almost identical (only one residue difference), whereas the less expressed AjTPS7 had 7 amino acid substitutions when compared to AoTPS7 (Fig. 4a). To verify the function of the TPS7 proteins from different *Aquilegia* species, *AoTPS7*, *HTPS7* and *AjTPS7* were transformed into Nicotiana tabacum *(cv. K326)* leaves and several transgenic lines were obtained (Fig. S10a). Headspace solid-phase microextraction gas chromatography–mass spectrometry (HS-SPME-GC–MS) analysis showed that transgenic tobacco overexpressing *AoTPS7* and *HTPS7* released more (+)-limonene than those overexpressing *AjTPS7* (Figs. 4b, S10b). Furthermore, *in vitro* enzymatic assays were carried out to explore whether the polymorphisms of TPS7 affected the production of (+)-limonene. Consistent with the overexpression assays in tobacco, recombinant AoTPS7 ((+)-limonene, 3060.60 ± 948.44 ng) catalyzed GPP to produce increased amounts of (+)-limonene compared to AjTPS7 ((+)-limonene, 4.97 ± 3.63 ng). To minimize the key amino acids, AoTPS7 and AjTPS7 were split at residue 308 into N (containing 4 mutant residues) and C (containing 3 mutant residues) termini, respectively. Two kinds of chimeric proteins, N(Ao)C(Aj) and N(Aj)C(Ao), were purified from E. coli BL21(DE3) and employed in enzymatic assays against GPP. The chimeric protein

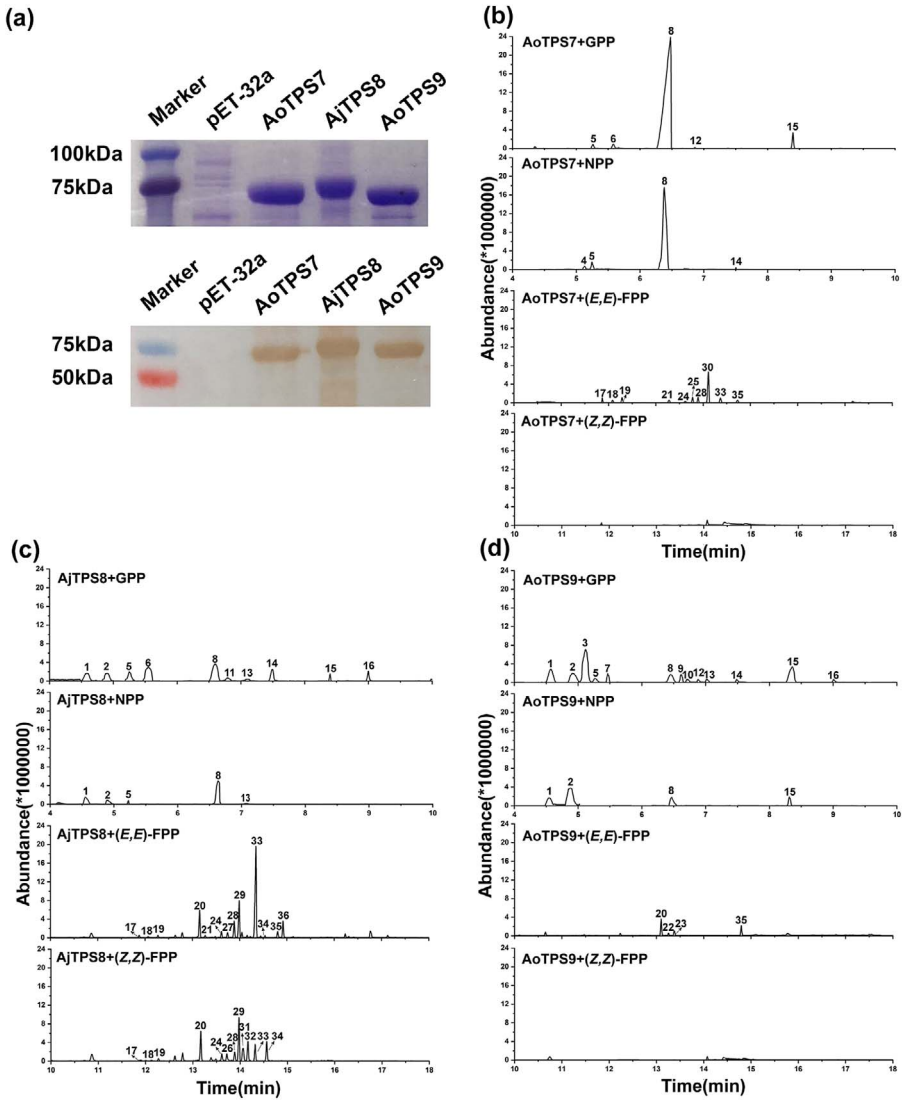


Figure 3. In vitro enzymatic analysis of AoTPS7, AjTPS8 and AoTPS9 using four acyclic prenyl diphosphate substrates. a SDS–PAGE and western blotting analysis of the purified Aquilegia TPS proteins. The recombinant proteins were indicated by anti-His-Tag antibody. b Enzymatic products of AoTPS7. c Enzymatic products of AjTPS8. d Enzymatic products of AoTPS9. The x-axis represents the retention time, and the y-axis represents the abundance of each compound. The enzymatic products are detailed in Table S4. Numbers above the peaks represent the following: 1, (1R)-(+)-α-pinene; 2, (1S)-(−)-β-pinene; 3, (1S)-(−)-β-pinene derivative; 4, α-phellandrene; 5, β-phellandrene; 6, β-pinene; 7, myrcene; 8, (+)-limonene; 9, (3R)-(+)-isosylvestrene; 10, 3-carene; 11, (Z)-β-ocimene; 12, γ-terpinene; 13, terpinolene; 14, linalool; 15, α-terpineol; 16, geraniol; 17, α-bergamotene; 18, an unidentified sesquiterpene; 19, an unidentified sesquiterpene; 20, an unidentified sesquiterpene; 21, an unidentified sesquiterpene; 22, β-chamigrene; 23, α-farnesene; 24, (E)-β-farnesene; 25, β-himachalene; 26, α-cedrene; 27, (Z,E)-α-farnesene; 28, zingiberene; 29, β-bisabolene; 30, (+)-α-longipinene; 31, α-patchoulene; 32, α-himachalene; 33, β-sesquiphellandrene; 34, α-caryophyllene; 35, (E)-nerolidol; and 36, farnesol methyl ether.

N(Aj)C(Ao) had higher efficiency in producing (+)-limonene than AjTPS7, suggesting that the three (3) amino acid residues in the C-terminus of AoTPS7 contributed significantly to its increased catalytic activity (Fig. 4c, 4d). Additionally, the 3 residues in the C-terminus of AjTPS7 were back-mutated to V413A, L503I and G529D by different combinations. Enzymatic assays showed that the V413A and L503I mutants scarcely affected the catalytic activity of TPS7, while the G529D mutant marginally increased the activity compared to the parental counterpart (Fig. 4c, 4d). Subsequently, a bioassay of the variant AjTPS7 protein harboring double mutations (V413AL503I) produced a very low amount of (+)-limonene, while the AjTPS7 protein carrying the V413AG529D and L503IG529D mutations yielded higher amounts of (+)-limonene, at 1944.73 ± 241.80 ng and 1763.71 ± 6.12 ng, respectively, than the parent AjTPS7 (4.97 ± 3.63 ng) (Fig. 4c, 4d), indicating that 413A, 503I and 529D were vital for the catalytic activity of the TPS7 protein. However, the triple mutant V413AL503IG529D, corresponding to N(Aj)C(Ao), resulted in less (+)-limonene than the double mutants V413AG529D and L503IG529D (Fig. 4c, 4d), indicating that the 4 mutant residue in the N-terminus should not be overlooked in determining the catalytic activity of TPS7. Moreover, we tried to obtain the kinetic parameters of

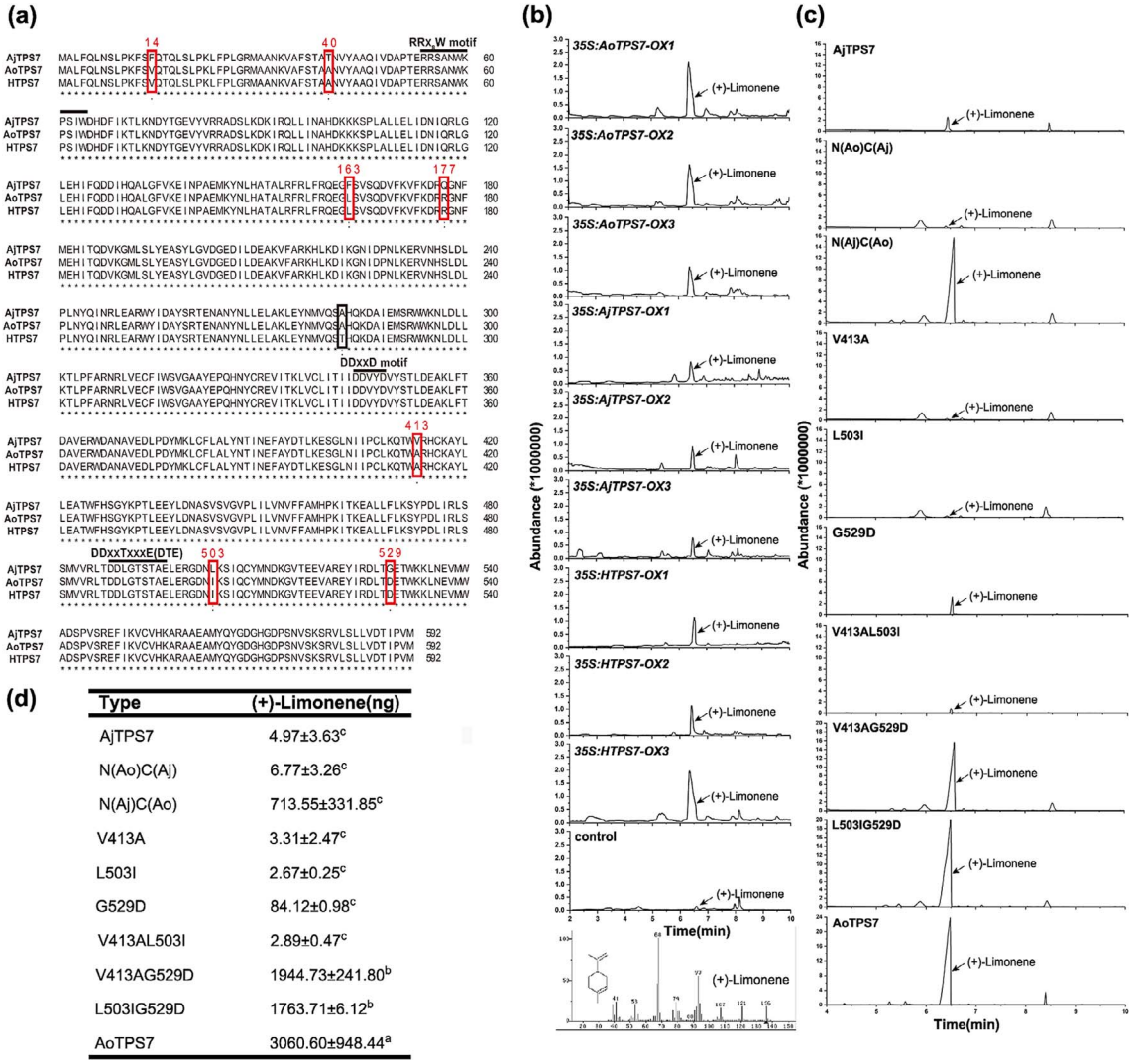
Aquilegia TPS7 proteins. For example, the k_m, V_{max}, and K_{cat} of AoTPS7 were 4.77 μM, 0.007 mM s^{-1} and 0.02 s^{-1}, respectively (Fig. S11). However, AjTPS7 produced too little (+)-limonene to calculate the kinetic parameters. Considering that the results presented may be enough to compare the differences in these enzymes, we tentatively abandoned the pursuit of accurate kinetic parameters for all the proteins. Overall, the results suggested that the catalytic activity of AjTPS7 could be significantly affected by several residue variations.

Three amino acids of TPS8 likely determine the ratio of β-sesquiphellandrene to β-bisabolene

AoTPS8 and HTPS8 shared identical sequences (Fig. S5), whereas six variant residues were observed between AoTPS8 and AjTPS8 (Figs. S5, 5a). To verify the function of the TPS8 proteins, *AoTPS8* and *AjTPS8* were first transformed into *N. tabacum (cv. K326)*, but no obvious terpene changes were detected. Alternatively, *AoTPS8* and *AjTPS8* were transiently expressed in *Nicotiana benthamiana* leaves, which resulted in the generation of diverse sesquiterpenes mainly composed of β-sesquiphellandrene and β-bisabolene (Figs. 5b, S12). Interestingly, the ratio of β-sesquiphellandrene to β-bisabolene produced by AoTPS8 and AjTPS8 varied obviously, as AoTPS8 generated more β-bisabolene than β-sesquiphellandrene, and vice versa for AjTPS8 (Figs. 5b, S12). Consistently, the *in vitro* enzymatic assays also confirmed the catalytic properties of these two *Aquilegia* TPS8 proteins (Fig. 5c; Table S5). For instance, the ratio of β-sesquiphellandrene to β-bisabolene produced by AoTPS8 was ~7% to ~34%, whereas the ratio became ~52% to ~18% for AjTPS8 (Fig. 5d). Moreover, the six residue variations made TPS8 produce mostly β-sesquiphellandrene and β-bisabolene. β-sesquiphellandrene and β-bisabolene constituted ~41% of the total products generated by AoTPS8, while these main products of AjTPS8 increased to ~69% (Fig. 5d; Table S5). To further investigate the key amino acids responsible for the change in this ratio, site-directed mutation was conducted on AoTPS8 sequentially. Using (*E, E*)-FPP as a substrate, the A420G, V538F and F545L mutants of AoTPS8 were found to significantly influence the yield of β-sesquiphellandrene and β-bisabolene, whereas Q14R, L444F and V556I variations were not (Fig. 5d; Table S5). Furthermore, AoTPS8 mutant proteins carrying double mutations and triple mutations resulted in significant yield alterations of β-sesquiphellandrene to β-bisabolene, further validating that these residues were crucial for β-bisabolene and β-sesquiphellandrene biosynthesis (Fig. 5d; Table S5). Notably, the triple mutation (A420GV538FF545L) was enough to produce a similar amount or ratio of β-bisabolene and β-sesquiphellandrene as AjTPS8, suggesting that the residues at positions 420, 538 and 545 played vital roles in the biosynthesis of β-sesquiphellandrene and β-bisabolene in AjTPS8.

Correlation analysis between volatiles of *A. oxysepala* and *A. japonica* from the Changbai Mountain population and *TPS* gene expression

As *Aquilegia* is widely distributed, we are quite interested in the correlation between volatile release and the underlying *TPSs* among different populations. Consequently, the volatile terpenes of *A. oxysepala* and *A. japonica* from Changbai Mountain (42.4°N, 128.15°E) were detected. *A. oxysepala* from Changbai Mountain released relatively more (+)-limonene (~65% versus ~42%) than that from Zuojia. Comparatively, *A. oxysepala* from both areas had similar terpene release patterns (Figs. 2, 6a; Tables S1, S6). In contrast, *A. japonica* from Changbai Mountain released ~96% (+)-limonene, while the sesquiterpenes β-sesquiphellandrene and β-bisabolene constituted ~79% of the total volatile terpenes of *A. japonica* from the Zuojia population (Figs. 2, 6a; Tables S1, S6). Therefore, it was reasonable to conclude that the volatile terpenes exhibited geography-specific variations among different *Aquilegia* populations. To verify the molecular basis of the geography-specific volatile terpene emissions, the transcripts of *TPS7* and *TPS8* were measured in different *Aquilegia* individuals from the Changbai Mountain population. *TPS7* was found to be highly expressed, which corresponded well with the (+)-limonene release from both *A. oxysepala* and A. japonica in this geographic area (Fig. 6b). As the key amino acid residues found in *Aquilegia* TPS7 and TPS8 were determinants of their catalytic properties, variations in these sites were detected in different *Aquilegia* individuals from both Zuojia and the Changbai Mountain



(d)

Type	(+)-Limonene(ng)
AjTPS7	4.97±3.63^{c}
N(Ao)C(Aj)	6.77±3.26^{c}
N(Aj)C(Ao)	713.55±331.85^{c}
V413A	3.31±2.47^{c}
L503I	2.67±0.25^{c}
G529D	84.12±0.98^{c}
V413AL503I	2.89±0.47^{c}
V413AG529D	1944.73±241.80^{b}
L503IG529D	1763.71±6.12^{b}
AoTPS7	3060.60±948.44^{a}

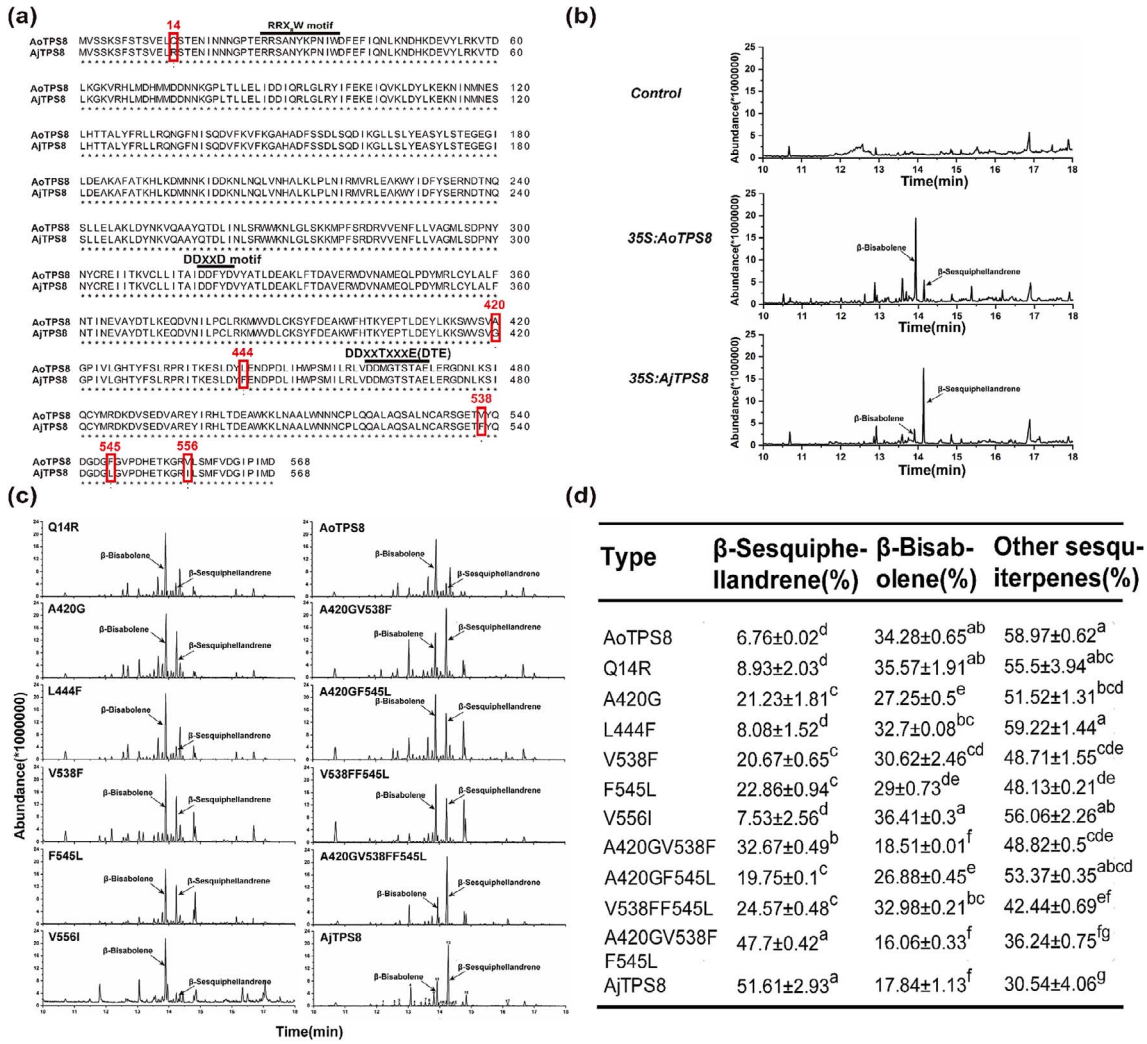
Figure 4. Possible pivotal amino acid residues of TPS7 for producing (+)-limonene. a Amino acid sequence alignment of AjTPS7 (from A. japonica 7#), AoTPS7 (from A. oxysepala 7#) and HTPS7. The conserved motifs are indicated by lines. The polymorphic residues between AjTPS7 and AoTPS7 are marked with red boxes, while polymorphic residues between HTPS7 and AoTPS7 are indicated by black boxes. The numbers in each line indicate the positions of the last residue. b Volatile terpenes detected in control and transgenic tobacco overexpressing TPS7. TPS7 was stably transformed into Nicotiana tabacum (cv. K326). The seedlings in tissue culture bottles were directly used for volatile terpene detection. The lower panel indicates the typical mass spectrum of (+)-limonene. c Enzymatic assays of AoTPS7, AjTPS7, and a series of mutant TPS7 proteins based on the AjTPS7 backbone. GPP was employed as the substrate in these assays. d Tabulated contents of (+)-limonene catalyzed by AoTPS7, AjTPS7 and mutant TPS proteins. Data represent the mean ± SD of three replicates. The (+)-limonene content was calculated by a standard curve. One-way ANOVA was carried out to compare significant differences (Duncan, $P < 0.05$).

population. As shown in Fig. 6c, the mutations in TPS7 were species dependent instead of geography related, as all the mutations only appeared in *A. japonica* (Fig. 6c). For TPS8, only *A. oxysepala* from Zuojia changed from AVF to GFL (Fig. 6c). Considering the expression levels of *TPS7* and *TPS8* in different populations, mutations were often present in individuals with low expression levels in specific species, suggesting that there might be some correlations between amino acid mutations and gene expression.

Together, these results suggest that TPS orthologs have divergent roles within *Aquilegia* populations endemic to Northeast Asia, and they are likely to be subfunctionalized due to nonsynonymous amino acid mutations when their expression levels become low.

Discussion

The quality and quantity of volatile terpenes have been extensively studied in both angiosperms and gymnosperms. Although previous studies have attempted to examine genus-specific floral volatile constituents in anthers, pollen and stamens of the Ranunculaceae family [43], little is known about *Aquilegia*-specific floral chemistry. Previous studies have demonstrated that floral volatile chemistry is species-specific with the possibility for great variation among related species or different selection events of similar species/varieties [44]. Abundant terpenes are synthesized solely in specific taxa and probably evolved to cope with different ecological habitats, thus forming crucial cues



Type	β-Sesquiphellandrene(%)	β-Bisabolene(%)	Other sesquiterpenes(%)
AoTPS8	6.76±0.02^{d}	34.28±0.65^{ab}	58.97±0.62^{a}
Q14R	8.93±2.03^{d}	35.57±1.91^{ab}	55.5±3.94^{abc}
A420G	21.23±1.81^{c}	27.25±0.5^{e}	51.52±1.31^{bcd}
L444F	8.08±1.52^{d}	32.7±0.08^{bc}	59.22±1.44^{a}
V538F	20.67±0.65^{c}	30.62±2.46^{cd}	48.71±1.55^{cde}
F545L	22.86±0.94^{c}	29±0.73^{de}	48.13±0.21^{de}
V556I	7.53±2.56^{d}	36.41±0.3^{a}	56.06±2.26^{ab}
A420GV538F	32.67±0.49^{b}	18.51±0.01^{f}	48.82±0.5^{cde}
A420GF545L	19.75±0.1^{c}	26.88±0.45^{e}	53.37±0.35^{abcd}
V538FF545L	24.57±0.48^{c}	32.98±0.21^{bc}	42.44±0.69^{ef}
A420GV538F F545L	47.7±0.42^{a}	16.06±0.33^{f}	36.24±0.75^{fg}
AjTPS8	51.61±2.93^{a}	17.84±1.13^{f}	30.54±4.06^{g}

Figure 5. Possible amino acid residues responsible for the product plasticity in Aquilegia TPS8. a Amino acid sequence alignment between AoTPS8 (from A. oxysepala 7#) and AjTPS8 (from A. japonica 7#). The conserved TPS motifs are indicated by lines, and polymorphic amino acids are marked with red boxes. The numbers in each line indicate the positions of the last residue. b Volatile compounds detected in the control and transgenic Nicotiana benthamiana transiently overexpressing TPS8. AtFPS2, which encodes Arabidopsis thaliana farnesyl diphosphate synthase in FPP synthesis, was coinfiltrated into tobacco leaves. The tobacco leaves were sampled and analyzed by GC–MS analysis. Tobacco leaves infiltrated with AtFPS2 alone were used as controls. c Enzymatic assays of AoTPS8, AjTPS8, and a series of mutant TPS8 proteins based on the AoTPS8 backbone. (E, E)-FPP was employed as the substrate in these assays. The enzymatic products are detailed in Table S5. d Relative contents of β-sesquiphellandrene and β-bisabolene catalyzed by AoTPS8, AjTPS8 and mutant TPS proteins. Data represent the mean ± SD of two replicates calculated from relative peak areas. One-way ANOVA was carried out to compare significant differences (Duncan, $P < 0.05$).

for attracting pollinators, repelling enemies such as insects, mollusks and mammals or becoming toxic to insects, bacteria and fungi [45, 46]. The *Aquilegia* species show obvious differences in vegetative and floral traits with specializations for different ecological niches and pollinators [37, 47–49]. Therefore, evaluation of divergent floral volatiles in *A. oxysepala* and *A. japonica* and their hybrid may provide insight into pollinator shift and speciation. For example, studies found that *Aquilegia pubescens* and *Aquilegia formosa* have adapted to different pollinators. Bees and hummingbirds visited *A. formosa*, whereas hawkmoths exclusively visited *A. pubescens* [33]. Presently, *A. oxysepala* and *A. japonica* from Northeastern Asia were also found to have quite different terpene release patterns (Figs. 2, 6). However, whether the different floral volatiles released by *A. oxysepala* and *A. japonica* affect their pollinators or benefit them in specific ecological niches remains to be further investigated.

The enormous architectural diversity of floral volatile terpenes is driven by *TPS* genes ubiquitously present in terrestrial plants [7]. The enzymes encoded by these genes catalyze the synthesis of monoterpenes (C-10), sesquiterpenes (C-15), and diterpenes (C-20) using acyclic substrates geranyl diphosphate (GPP; C-10), farnesyl diphosphate (FPP; C-15), and geranylgeranyl diphosphate (GGPP; C-20), in that order. Varying numbers of *TPS* genes were observed in plants, ranging from a single gene in Physcomitrella patens to over a hundred genes in both Eucalyptus grandis and Eucalyptus grandis, which resulted in the vast volatile terpene product spectrum [7, 50–52]. In the representative *A. coerulea*, a total of 42 *TPS*s were identified. Gene expression analysis indicated

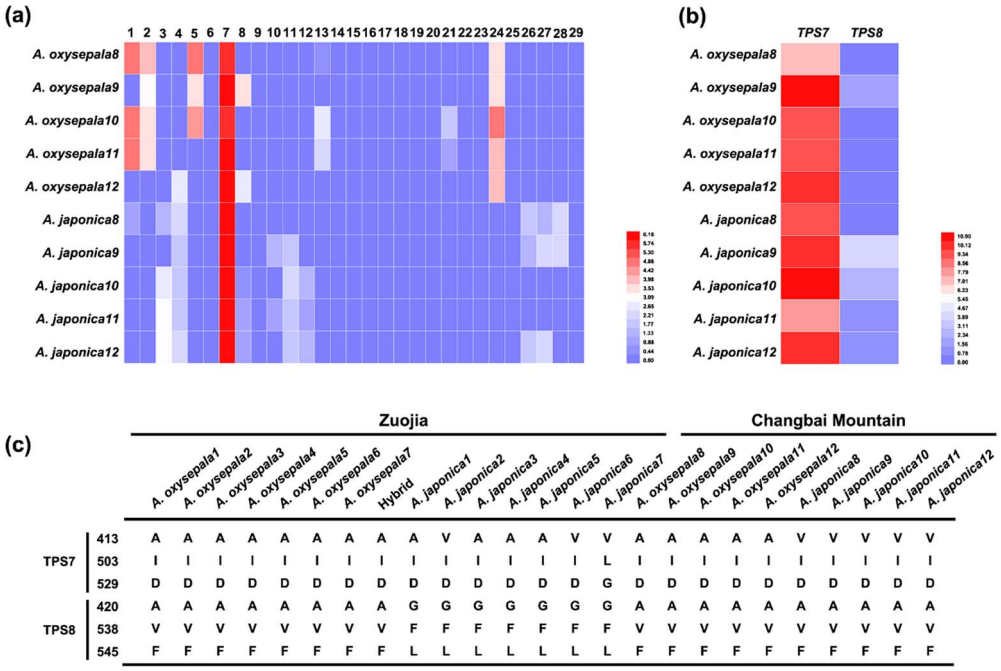


Figure 6. Volatile terpene release, main TPS expression and key residue polymorphisms of A. oxysepala and A. japonica from Changbai Mountain. a Volatile terpenes emitted from flowers of A. oxysepala and A. japonica individuals from the Changbai Mountain population. Data represent the relative content calculated from peak areas and then log2 transformed. The detailed terpene contents are listed in Table S6. Here, 1 to 29 represent different volatiles: 1, (1R)-(+)-α-pinene; 2, (1S)-(−)-β-pinene; 3, α-phellandrene; 4, β-phellandrene; 5, β-pinene; 6, myrcene; 7, (+)-limonene; 8, 3-carene; 9, (E)-β-ocimene; 10, linalool; 11, α-terpineol; 12, an unidentified sesquiterpene; 13, (−)-α-copaene; 14, α-bergamotene; 15, an unidentified sesquiterpene; 16, (Z)-β-farnesene; 17, an unidentified sesquiterpene; 18, α-curcumene; 19, (E)-β-farnesene; 20, himachalene; 21, β-caryophyllene; 22, zingiberene; 23, β-bisabolene; 24, β-sesquiphellandrene; 25, α-caryophyllene; 26, an unidentified sesquiterpene; 27, cubebene; 28, α-muurolene; and 29, an unidentified sesquiterpene. b Relative expression levels of TPS7 and TPS8 in Aquilegia individuals from the Changbai Mountain population. Data represent changes relative to the lowest expression level of TPS in specific species. Data are the mean of three replicates and calculated as log2. c Key amino acid residue polymorphisms of TPSs among Aquilegia individuals from both Zuojia and Changbai Mountain. A. oxysepala 1#-7#, A. japonica 1#-7# and the hybrid are from Zuojia. A. oxysepala 8#-12# and A. japonica 8#-12# are from Changbai Mountain. The number in front of each line indicates the key residue position.

that only some *TPSs* were highly expressed in flowers (Fig. 1), which seemed paradoxical to the diverse terpenes detected. This paradox could be partly interpreted by the versatile TPS proteins, as a single TPS could transform a specific kind of substrate into multiple products, and most TPS proteins had broad substrate spectra. For example, the highly expressed *TPS7*, *TPS8* and *TPS9* characterized in *Aquilegia* could catalyze either GPP, FPP or NPP into a series of volatile terpenes (Fig. 3). Moreover, amino acid alterations in the suggested TPS active sites could have significant effects on catalytic features, such as functional plasticity and specificities, which was validated by the site-directed mutations of TPS7 and TPS8 (Figs. 4, 5). Evolutionarily, the occurrence of single/multiple mutations within a gene might enable organisms to adapt to new or changing environments. For instance, when a gene is duplicated and then one copy of the duplicated gene can be functionally modified by mutation or even become a novel gene with a completely different function from the original gene. Theoretically, the large diversity of *TPS* genes might have evolved from a parental gene via three routes: gene silencing, neofunctionalization and subfunctionalization [53]. Consequently, sustained *TPS* gene duplication accompanied by functional divergence has been considered as a crucial factor responsible for the chemodiversity of floral volatile terpenes [7, 10]. However, the low expression of *TPS* genes detected in *Aquilegia* could not be arbitrarily considered to be silenced, as some of these genes might be induced by specific stimuli, such as biotic or abiotic stresses, or expressed in other tissues or developmental stages in addition to flowers [54, 55]. Furthermore, TPSs play important roles in the determination of immense variations in the terpenome among and within terrestrial plant taxa [1, 23, 56]. Notwithstanding these general conceptions, concrete illustrations of how duplicated *TPS* genes have evolved across multiple plant species populations are still scarce.

Generally, the functional characterization of *Aquilegia* TPSs and their key amino acids identified from this taxon exhibit broad plasticity, which indicates a crucial cornerstone toward rapid functional divergence within and among *Aquilegia* species. The TPS product specificity is prominently variable and elementarily depends on how easily the substrate carbocation can be stabilized in the active pocket or center of the enzymes [57, 58]. Most terpene synthases have active pockets with a propensity toward catalytically accepting more than one acyclic substrate ranging from IPP (C-5) to FPP (C-15) to synthesize multiterpene profiles [59]. The TPSs with verified potential for multisubstrate usage are extensively divergent, spreading across the seven terpene

synthase subfamilies (TPS-a, −b, −c -d, −e/f, −g and TPS-h), further accentuating the capability that the multisubstrate usage of TPSs could be widespread more than previously believed. For example, detailed analysis of the substrate specificities of TPS isolated from Vitis vinifera has shown the existence of at least nine multisubstrate TPSs: two are mono−/sesqui−/diterpenes, and the others are mono−/sesquiterpene synthases [12, 38]. Similarly, *Santalum* species have seven functionally confirmed C-10/C-15 multisubstrate terpene synthases. The versatile *Aquilegia* TPSs further indicated that the capacity to use multiple substrates may be widespread across TPS families in multiple species [59, 60]. However, it seems that the subcellular compartments confine the TPS in utilizing specific substrates due to the subcellular localizations of monoterpene-related TPSs in plastids (GPPs or NPPs) versus sesquiterpene-related TPSs in the cytosol (FPP) [7, 42]. For instance, although plastid-localized AoTPS7 and cytosol-localized AjTPS8 can catalyze GPP, NPP or FPP *in vitro*, not all the products can be detected *in planta* (Figs. 1-5).

In general, the observed expression levels of *TPS* genes, subcellular localizations of TPS proteins, and their catalytic properties accounted for the emission patterns of *Aquilegia* floral volatile terpenes. Simply, an illustrative hypothetical model was constructed to interpret the molecular mechanism underlying the floral volatile terpene emissions in *A. oxysepala* and *A. japonica* from Changbai Mountain and Zuojia District (Fig. 7). *TPS* genes are responsible for the floral volatile terpenes within *Aquilegia* populations endemic to Northeast Asia, and they will be subfunctionalized when their expression levels are changed. Generally, the lowly expressed orthologs appear to accumulate mutations easily, among which the nonsynonymous amino acid mutations will contribute to the functional divergence of the *TPS* isogenes.

Nonsynonymous amino acid mutations responsible for the catalytic activity or product specificity of TPS proteins have also been observed elsewhere. For instance, some amino acid motif variations in the closely related TPS4 and TPS5 genes from two maize cultivars drove the stereospecificity of multiproduct sesquiterpene synthases [61]. N338A in 1,8-cineole synthase of *Salvia fructicosa*, Y402L in β-farnesene synthase of Artemisia annua and TPSs 4/10 in maize demonstrated that the byproducts of TPSs were determined not only by residues forming the center active pocket but also by adjacent residues impacting the conformation [29, 62–64]. In a similar manner, our study revealed that 413A, 503I and 529D perhaps influenced the catalytic activity of AjTPS7, while 420G, 538F and 545 L affected the ratio of β-sesquiphellandrene to β-bisabolene. Consistent with the opinion that organisms are always evolving, floral volatiles may be employed by plants as a strategy to cope with perpetually changing environmental settings. Our results provide evidence for a possible mechanism whereby *Aquilegia* or other angiosperms, instead of modifying the committed genes in the earlier terpene biosynthesis pathway, may employ a simple but effective scaffold that can, with only very few substitutions, provide diverse terpene release. Furthermore, our results showed some correlations between amino acid mutations and gene expression. As a result, we suggest that the change in the gene expression level occurs prior to the amino acid mutation, as not all the less expressed *AjTPS7* from Zuojia populations had VID or VIG mutations when compared with AID in Zuojia *AoTPS7* (Fig. 6). However, more investigations are still necessary to fully investigate this issue. The combination of product specificity and plasticity observed in *Aquilegia* terpene synthases may be a snapshot of an evolutionary mechanism that maintains a dynamic state between desirable specific activities and a capacity for rapid change.

Materials and methods

Plant materials and growth conditions

Plant materials of the two columbine species, *A. oxysepala* and *A. japonica*, and their natural hybrid were collected from Zuojia (43.87°N, 126.57°E) and Changbai Mountain (42.4°N, 128.15°E) in Northeast China (Figs. S1, S13). After volatile compound sampling, the floral samples obtained were promptly frozen in liquid nitrogen and stored at −80°C until further use.

To determine the correlation between the *TPS* expression levels and floral volatile terpene emissions, *Aquilegia* plants were cultivated in a greenhouse under a photoperiod of 16 h of light followed by 8 h in the dark, and flowers were separated into four developmental stages (Fig. S1). Wild-type plants of Arabidopsis thaliana ecotype Columbia (Col-0) used to generate protoplasts for subcellular localization were grown in a greenhouse at 22°C with an alternating photoperiod of 16/8 h light/dark. Rosette leaves of *Arabidopsis* plants grown for 3 to 4 weeks were used for protoplast isolation. For the *in vivo* activity assay of TPS proteins, tobacco plants were grown in a greenhouse under natural light settings. The expanded leaves of 4-week-old tobacco plants were used for *Agrobacterium GV3101* infiltration experiments.

Analysis of floral volatile compounds in *A. oxysepala*, *A. japonica* and their hybrid

The volatile terpenes released from *Aquilegia* flowers were analyzed following a previously published method [18, 20]. Briefly, fresh flowers from *A. oxysepala*, *A. japonica* and their hybrid were sealed in a transparent and odor-free plastic device. A 100-mm fused silica fiber coated with DVB/CAR/PDMS (Sigma-Aldrich) was used to trap floral volatile compounds for 2 h at room temperature [18, 65] and then directly used for gas chromatography–mass spectrometry (GC/MS). The fibers were thermally desorbed at 260°C for 20 seconds in an Agilent 5975-6890 N GC/MS apparatus (Agilent Technologies) equipped with an HP-1 MS Fused-Silica Capillary Column (0.25 mm diameter, 30 m length, and

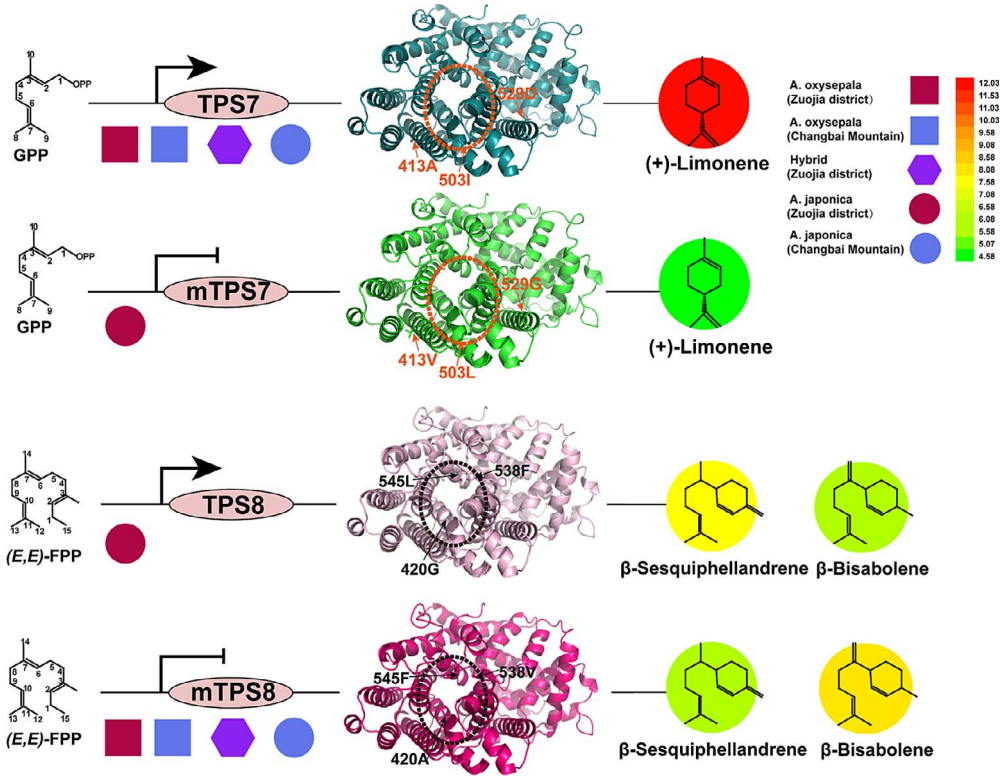


Figure 7. Proposed model for volatile terpene biosynthesis in flowers of Aquilegia populations endemic to Northeastern Asia. The parent TPS7 was highly expressed in A. oxysepala (Zuojia and the Changbai mountain population), the hybrid and A. japonica (the Changbai mountain population). The active parent TPS7 generated a high amount of (+)-limonene, whereas the mTPS7 mutated protein was hardly expressed in A. japonica (the Zuojia population), although it produced a low amount of (+)-limonene. The parent TPS8 was highly expressed in A. japonica (the Zuojia population) when compared with the less expressed mTPS8 in the other populations. Moreover, the parent TPS8 yielded a high amount of β-sesquiphellandrenes, while the mutant mTPS8 produced more β-bisabolenes. The arrows represent high expression. The squares and circles under TPSs indicate A. oxysepala and A. japonica, respectively. The different background colors of the same shape indicate the same species from different populations. The hexagon indicates the natural hybrid from the Zuojia population. The three-dimensional models of the TPSs were created according to the crystal structure of 5-epi-aristolochene synthase (PDB ID: 5IL3) from tobacco61. The active cavity is illustrated in dotted circles. Polymorphic amino acids in or adjacent to the active cavity are marked. The different background colors of (+)-limonene, β-sesquiphellandrenes and β-bisabolenes represent different contents produced by the enzymes.

0.25 μm film thickness). The detection parameter was held at 60°C for 3 min, subsequently raised at a rate of 5°C min^{-1} to 100°C for 1 min, and then increased at a rate of 10°C min^{-1} to 250°C for 10 min. Prospective floral volatile constituents were identified by equating mass spectra from GC/MS analysis with the mass spectra archived in NIST 2008 and a reference standard sample when applicable. The absolute content of volatile terpenoids was quantitatively analyzed according to the standard curve of (+)-limonene or β-caryophyllene.

TPS gene identification, sequencing analysis and cloning

Using FhTPS1 (GenBank NO. AFP23421.1) as a bait sequence, the predicted *TPS* genes were first mined from the assembled genome database of *A. coerulea* (Phytozome v12.1, https://phytozome-next.jgi.doe.gov/info/Acoerulea_v3_1) [32] by using the TBLASTN algorithm and subsequently subjected to a manual NCBI-BLASTX search. The best hits were taken as *TPS* candidate genes. In addition, RNA-seq analysis of flowers at different developmental stages was also carried out to screen the potential *TPS* genes in the selected *Aquilegia* species. Briefly, total RNA was first isolated from samples by employing the OminiPlant RNA Kit (DNase I) (CWBIO, Beijing, PRC) following the manufacturer's specified procedure. The RNA samples were either sent to BioMarker, Beijing, PRC for RNA-seq following the company's standard process or subjected to cDNA synthesis. The cDNA sample was made in a final 25 μl reaction system including total RNA (1 μg), Oligo d(T)15 primers, and M-MLV Reverse Transcriptase (Promega, Madison, US) according to the manufacturer's manual. Specific primers (Table S7) were designed to amplify the full-length cDNA sequences of *TPS*s. The sequences were ligated to the pESI-Blunt vector (Yeasen, Shanghai, PRC) and then subjected to sequencing confirmation.

Multiple sequence alignment of TPSs was performed by Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/) with default parameters. For phylogenetic evaluation, the aligned sequences were subjected to MEGA-X to generate a neighbor-joining tree with bootstrap analysis (1000 replicates) and gap handling by pairwise deletion [66]. The following GenBank accession numbers were used: AtTPS10 (ACF41947), AtTPS14 (NP001185286), AtTPS21 (NP001190374), AtGA1 (Q38802.1), and AtGA2

(Q9SAK2.1). Transcript names of the 42 A. coerulea TPSs used in the phylogenetic analysis and deposited in Phytozome v12.1 are listed in Table S7.

Quantitative real-time PCR analysis

To evaluate the expression profiles of the *TPS* genes in *A. oxysepala*, *A. japonica* and their hybrid, a qPCR assay was conducted in a 10 μl reaction volume containing 5 μl of TB Green® Premix Ex Taq (TaKaRa, Kusatsu, Japan), 0.5 μM of each oligonucleotide and 1 μl of cDNA template. The specific qPCR primers used in this study are listed in Table S7. The *β-actin* and *IPP2* genes were utilized as endogenous controls [67, 68]. Relative gene expression was calculated using the $2^{-\Delta\Delta C_T}$ formula [69]. All measurements presented in this study were prepared in triplicate.

Subcellular localization of *Aquilegia* TPS proteins

The previously used *35S:FhPAP1-GFP* in the backbone of pUC19 was digested by *Nde*I and *Cla*I [70]. The whole open reading frames (ORFs) of *AoTPS7, AjTPS8, AoTPS9*, and *AoTPS27* without termination codons were subcloned into the aforementioned digested pUC19 by a Minerva Super Fusion Cloning Kit (US Everbright® Inc., Suzhou, PRC) to generate *35S:AoTPS7-GFP*, *35S:AjTPS8-GFP, 35S:AoTPS9-GFP*, and *35S:AoTPS27-GFP*, respectively. The plasmids were extracted using the GoldHi EndoFreePlasmid Maxi Kit (CWBIO, Beijing, PRC) according to the instructions provided and then transfected into *Arabidopsis* protoplasts as described previously [71]. The GFP signals were visualized by fluorescence microscopy.

Heterologous expression of TPS proteins in E. coli and *in vitro* enzymatic assay

To express recombinant proteins in E. coli , the ORFs of *Aquilegia TPS*s were amplified with gene-specific primers (Table S7) and then subcloned into the pET32a vector. Afterward, an empty vector and vectors appropriately ligated with different *TPS* genes were used for transformation of E. coli strain BL21 (DE3). Recombinant TPS proteins were induced by 0.25 mM isopropyl-*β*-D-thiogalactopyranoside (IPTG) at 16°C for 16 h in a constant shaking incubator. Subsequently, the cells were harvested by centrifugation, resuspended in phosphate-buffered saline (PBS) and broken down by sonication. The crude proteins were then purified on a Ni-TED Sefinose™ column (Sangon Biotech, Shanghai, PRC) and analyzed by SDS-PAGE. Moreover, the recombinant proteins were confirmed by western blotting analysis using Anti His-Tag Mouse Monoclonal Antibody (CWBIO, Beijing, PRC) following the standard protocol. The following enzymatic assays were performed as previously described with slight revisions [18, 72]. Briefly, the 300 μl reaction mixture consisted of 50 mM HEPES buffer (pH 7.4), 2 mM GPP (Sigma Aldrich), *(E,E)*-FPP (Sigma Aldrich), NPP (Echelon Biosciences) or *(Z,Z)*-FPP (Echelon Biosciences), 75 mM $MgCl_2$, 5 mM dithiothreitol, 5% glycerol and 60 μg protein. The mixture was incubated at 30°C for 2 h. The volatile products were captured by silica fiber and then analyzed by GC–MS as previously mentioned. Based on an *in vitro* enzymatic assay system, the kinetic parameters of AoTPS7 were calculated.

Chimera construction and site-directed mutagenesis

To create both chimeric cDNAs and site-directed mutants, PCR was used together with a revised version of the overlap extension strategy detailed in other studies [27, 28]. The PCR-based mutagenesis protocol (*Fast* Mutagenesis System, TransGen Biotech, PRC) was executed on *pET32a-AjTPS7* and *pET32a-AoTPS8* by employing primers containing the desired mutations (Table S7). The *TPS7* Chimeras N(Ao)C(Aj) and N(Aj)C(Ao) were created by overlapping PCR fragments of the targeted region (split at residue 308), which were employed to evaluate the role of C-terminal regions in the determination of enzyme activity and product specificity in *Aquilegia* TPS[27]. Briefly, complementary primers Chimera-F (Table S7) were designed to anneal to specific regions of the parent *AoTPS7* and *AjTPS7* to construct overlapping fragments. Subsequently, they were fused and amplified to generate the final chimeric construct. All chimeric constructs were subcloned back into the pET32a vector, and their sequences were verified as described elsewhere [27, 28].

In vivo characterization of terpene synthase

The obtained *TPS7* genes were subcloned into the pBI121 binary vector and then transformed into *Agrobacterium* competent cells (strain GV3101). *Agrobacterium* harboring *TPS7* or *TPS8* was infiltrated into the juvenile leaves of 4-week-old *N. tabacum* (*cv. K326*) [20]. Then, the leaves infiltrated by *Agrobacterium* were cut into small pieces and cultured on MS medium to obtain transgenic seedlings [70]. The volatile terpenes of transgenic seedlings in the tissue culture vessels were captured by silica fiber and then analyzed by GC–MS as previously mentioned. Moreover, *AoTPS8*, *AjTPS8* and *AtFPS2* (*A. thaliana* farnesyl diphosphate synthase 2, At4g17190) were subcloned into the pEAQ-HT vector, which was reported to dramatically enhance target protein expression [73]. *Agrobacterium* harboring *AoTPS8* or *AjTPS8* was mixed with that carrying *AtFPS2* (1:1) and infiltrated into the juvenile leaves of *Nicotiana benthamiana*. Four days later, the transiently transformed leaves were cut into small pieces and collected into odor-free bottles to detect volatile compounds.

Acknowledgments

We thank Professor Hongzhi Kong (Institute of Botany Chinese Academy of Sciences: Beijing, China) and Feng Chen (Department of Plant Sciences, University of Tennessee) for their discussion and comments on the manuscript. The pEAQ-HT vector was kindly shared by Professor Wei Li (Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences,

Shenzhen, China). This work was supported by the National Key R & D Program of China (2019YFD1000400; 2016YFD0101900), the National Natural Science Foundation of China [31972445, 31670382], and the Department of Science and Technology of Jilin Province [20190201299JC]. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author Contributions

SY, NW, SK, YL and TB performed the experiments and analyzed the data with the help of GN, and LL helped to confirm and collect the *Aquilegia* species. The manuscript was drafted by SK, SY, and XG and revised by XG and BL. XG designed the experiments and supervised the research with LW. All authors have participated in this research and approved the final manuscript.

Data availability

RNA-seq data have been deposited in the NCBI repository (SRR15204285- SRR15204292). The sequences were uploaded to GenBank with the following numbers: AoTPS7, MZ666959; AjTPS7, MZ666960; HTPS7, MZ666963; AoTPS8, MZ666961; AjTPS8, MZ670776; and AoTPS9, MZ666964. Other data supporting the findings of this study are available within the paper and within its supplementary materials published online.

Conflicts 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.

Supplementary data

Supplementary data is available at *Horticulture Research Journal* online.

References

1. Chen F, Tholl D, D'Auria JC. *et al.* Biosynthesis and emission of terpenoid volatiles from Arabidopsis flowers. *Plant Cell*. 2003;**15**: 481–94.
2. Besser K, Harper A, Welsby N. *et al.* Divergent regulation of terpenoid metabolism in the trichomes of wild and cultivated tomato species. *Plant Physiol*. 2009;**149**:499–514.
3. Degenhardt J, Köllner TG, Gershenzon J. Monoterpene and Sesquiterpene synthases and the origin of Terpene skeletal diversity in plants. *Phytochemistry*. 2009;**70**:1621–37.
4. Nagegowda DA. Plant volatile terpenoid metabolism: biosynthetic genes, transcriptional rRegulation and subcellular compartmentation. *FEBS Lett*. 2010;**584**:2965–73.
5. Tholl D, Lee S. Terpene specialized metabolism in *Arabidopsis thaliana*. *The Arabidopsis Book*. 2011;**9**:e0143
6. Pichersky E, Raguso RA. Why do plants produce so many terpenoid compounds? *New Phytol*. 2018;**220**:692–702.
7. Chen F, Tholl D, Bohlmann J. *et al.* 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*. 2011;**66**:212–29.
8. Zhou F, Pichersky E. More is better: the diversity of terpene metabolism in plants. *Curr Opin Plant Biol*. 2020;**55**:1–10.
9. Pichersky E, Gang DR. Genetics and biochemistry of secondary metabolites in plants: an evolutionary perspective. *Trends Plant Sci*. 2000;**5**:439–45.
10. Jiang S, Jin J, Sarojam R. *et al.* Comprehensive survey on the Terpene synthase gene family provides new insight into its evolutionary patterns. *Genome Biol Evol*. 2019;**11**:2078–98.
11. Köllner TG, Matthias H, Lenk C. *et al.* A maize (E)-β-caryophyllene synthase implicated in indirect defense responses against herbivores is not expressed in most American maize varieties. *Plant Cell*. 2008;**20**:482–94.
12. Martin DM, Aubourg S, Schouwey MB. *et al.* Functional annotation, genome organization and phylogeny of the grapevine (Vitis vinifera) terpene synthase gene family based on genome assembly, FLcDNA cloning, and enzyme assays. *BMC Plant Biol*. 2010;**10**:1–22.
13. Falara V, Akhtar TA, Nguyen TTH. *et al.* The tomato terpene synthase gene family. *Plant Physiol*. 2011;**157**:770–89.
14. Chen H, Köllner TG, Li G *et al.* Combinatorial evolution of a terpene synthase gene cluster explains terpene variations in Oryza. *Plant Physiol*. 2020;**182**:480–92.
15. Nieuwenhuizen NJ, Green SA, Chen X. *et al.* Functional genomics reveals that a compact terpene synthase gene family can account for terpene volatile production in apple. *Plant Physiol*. 2013;**161**:787–804.
16. Irmisch S, Jiang Y, Chen F. *et al.* Terpene synthases and their contribution to herbivore-induced volatile emission in western balsam poplar (*Populus trichocarpa*). *BMC Plant Biol*. 2014;**14**:1–16.
17. Külheim C, Padovan A, Hefer C. *et al.* The *eucalyptus* terpene synthase gene family. *BMC Genomics*. 2015;**16**:1–18.
18. Gao F, Liu B, Li M. *et al.* Identification and characterization of terpene synthase genes accounting for volatile terpene emissions in flowers of *freesia* x *hybrida*. *J Exp Bot*. 2018;**69**:4249–65.
19. Nawade B, Yahyaa M, Reuveny H. *et al.* Profiling of volatile terpenes from almond (*Prunus dulcis*) young fruits and characterization of seven terpene synthase genes. *Plant Sci*. 2019;**287**: 110187–11.
20. Bao T, Shadrack K, Yang S. *et al.* Functional characterization of terpene synthases accounting for the volatilized-terpene heterogeneity in *Lathyrus odoratus* cultivar flowers. *Plant Cell Physiol*. 2020;**4**:1–35.
21. Luck K, Chen X, Norris AM. *et al.* The reconstruction and biochemical characterization of ancestral genes furnish insights into the evolution of terpene synthase function in the poaceae. *Plant Mol Biol*. 2020;**104**:203–15.
22. Zhang L, Chen F, Zhang X. *et al.* The water lily genome and the early evolution of flowering plants. *Nature*. 2020;**577**:79–84.
23. Nagegowda DA, Gupta P. Advances in biosynthesis, regulation, and metabolic engineering of plant specialized terpenoids. *Plant Sci*. 2020;**294**:110457.
24. Chen H, Li G, Köllner TG. *et al.* Positive Darwinian selection is a driving force for the diversification of terpenoid biosynthesis in the genus *Oryza*. *BMC Plant Biol*. 2014;**14**:1–12
25. Back K, Chappell J. Identifying functional domains within terpene cyclases using a domain-swapping strategy. *Proc Natl Acad Sci U S A*. 1996;**93**:6841–5

26. El Tamer MK, Lü J, Bosch D. *et al.* Domain swapping of citrus Limon monoterpene synthases: impact on enzymatic activity and product specificity. *Arch Biochem Biophys*. 2003;**411**: 196–203
27. Peters RJ, Croteau RB. Alternative termination chemistries utilized by monoterpene cyclases: chimeric analysis of bornyl diphosphate, 1,8-cineole, and sabinene synthases. *Arch Biochem Biophys*. 2003;**417**:203–11.
28. Katoh S, Hyatt D, Croteau R. Altering product outcome in *Abies grandis* (–)-limonene synthase and (–)-limonene/(–)-α-pinene synthase by domain swapping and directed mutagenesis. *Arch Biochem Biophys*. 2004;**425**:65–76.
29. Köllner TG, Degenhardt J, Gershenzon J. The product specificities of maize terpene synthases TPS4 and TPS10 are determined both by active site amino acids and residues adjacent to the active site. *Plants-Basel*. 2020;**9**:1–12.
30. Puzey JR, Gerbode SJ, Hodges SA. *et al.* Evolution of spur-length diversity in *aquilegia* petals is achieved solely through cell-shape anisotropy. *P Roy Soc B-Biol Sci*. 2012;**279**:1640–5
31. Xie J, Zhao H, Li K. *et al.* A chromosome-scale reference genome of Aquilegia oxysepala var kansuensis. *Hortic Res*. 2020;**7**:1–13
32. Filiault D, Ballerini ES, Mandáková T. *et al.* The *aquilegia* genome provides insight into adaptive radiation and reveals an extraordinarily polymorphic chromosome with a unique history. *elife*. 2018;**7**:1–31.
33. Hodges SA, Whittall JB, Fulton M. *et al.* Genetics of floral traits influencing reproductive isolation between *aquilegia Formosa* and *Aquilegia pubescens*. *Am Nat*. 2002;**159**:S51–60.
34. Herlihy CR, Eckert CG. Evolutionary analysis of a key floral trait in *Aquilegia canadensis* (Ranunculaceae): genetic variation inherkogamy and its effect on the mating system. *Evolution*. 2007;**61**:1661–74.
35. Whittall JB, Hodges S. Pollinator shifts drive increasingly long nectar spurs in columbine flowers. *Nature*. 2007;**447**: 706–9.
36. Hodges SA, Derieg NJ. Adaptive radiations: from field to genomic studies. *Proc Natl Acad Sci U S A*. 2009;**106**:9947–54.
37. Kramer EM. *Aquilegia*: a new model for plant development, ecology, and evolution. *Annu Rev Plant Biol*. 2009;**60**: 261–77.
38. Zhu R, Gao Y, Xu L. *et al.* Genetic diversity of *aquilegia* (Ranunculaceae) species and cultivars assessed by AFLPs. *Genet Mol Res*. 2011;**10**:817–27.
39. Adler LS, Irwin RE. Nectar alkaloids decrease pollination and female reproduction in a native plant. *Oecologia*. 2012;**168**: 1033–41.
40. Parachnowitsch AL, Raguso RA, Kessler A. Phenotypic selection to increase floral scent emission, but not flower size or colour in bee-pollinated penstemon digitalis. *New Phytol*. 2012;**195**: 667–75.
41. Li L-F, Wang H-Y, Pang D. *et al.* Phenotypic and genetic evidence for ecological speciation of *Aquilegia japonica* and *A. oxysepala*. *New Phytol*. 2014;**204**:1028–40.
42. Dudareva N, Klempien A, Muhlemann JK. *et al.* Biosynthesis, function and metabolicengineering of plant volatile organic-compounds. *New Phytol*. 2013;**198**:16–32.
43. Jürgens A, Dötterl S. Chemical composition of anther volatiles in Ranunculaceae: genera-specific profiles in *anemone*, *aquilegia*, *Caltha*, *Pulsatilla*, *Ranunculus*, and *Trollius* species. *Am J Bot*. 2004;**91**:1969–80.
44. Klahre U, Gurba A, Hermann K. *et al.* Pollinator choice in *petunia* depends on two major genetic loci for floral scent production. *Curr Biol*. 2011;**21**:730–9.
45. Chorianopoulos N, Kalpoutzakis E, Aligiannis N. *et al.* Essential oils of *Satureja*, *Origanum*, and *thymus* species: chemical composition and antibacterial activities against foodborne pathogens. *J Agric Food Chem*. 2004;**52**:8261–7.
46. Terzi V, Morcia C, Faccioli P. *et al.* *In vitro* antifungal activity of the tea tree (*Melaleuca alternifolia*) essential oil and its major components against plant pathogens. *Lett Appl Microbiol*. 2007;**44**: 613–8.
47. Kramer EM, Hodges SA. *Aquilegia* as a model system for the evolution and ecology of petals. *Philos Trans R Soc Lond*. 2010;**365**: 477–90.
48. Fior S, Li M, Oxelman B. *et al.* Spatiotemporal reconstruction of the *aquilegia* rapid radiation through next-generation sequencing of rapidly evolving cpDNA regions. *New Phytol*. 2013;**198**: 579–92.
49. Sharma B, Yant L, Hodges SA. *et al.* Understanding the development and evolution of novel floral form in *aquilegia*. *Curr Opin Plant Biol*. 2014;**17**:22–7.
50. Gao Y, Honzatko RB, Peters RJ. Terpenoid synthase structures: a so far incomplete view of complex catalysis. *Nat Prod Rep*. 2012;**29**:1153–75.
51. Garms S, Chen F, Boland W. *et al.* A single amino acid determines the site of deprotonation in the active center of sesquiterpene synthases SbTPS1 and SbTPS2 from *sorghum bicolor*. *Phytochemistry*. 2012;**75**:6–13.
52. Tholl D. Biosynthesis and biological functions of terpenoids in plants. *Biotechnology of Isoprenoids*. 2015;**148**:63–106.
53. Wagner A. The fate of duplicated genes: loss or new function? *BioEssays*. 1998;**20**:785–8.
54. Boncan DAT, Tsang SSK, Li C. *et al.* Terpenes and Terpenoids in plants: interactions with environment and insects. *Int J Mol Sci*. 2020;**21**:7382.

55. Rosenkranz M, Chen Y, Zhu P. *et al*. Volatile terpenes – mediators of plant-to-plant communication. *Plant J*. 2021;**108**:617–31.
56. Tholl D, Chen F, Petri J. *et al*. Two sesquiterpene synthases are responsible for the complex mixture of sesquiterpenes emitted from Arabidopsis flowers. *Plant J*. 2005;**42**:757–71.
57. Christianson DW. Structural biology and chemistry of the terpenoid cyclases. *Chem Rev*. 2006;**106**:3412–42.
58. Bohlmann J, Keeling CI. Terpenoid biomaterials. *Plant J*. 2008;**54**: 656–69.
59. Pazouki L, Niinemets Ü. Multi-substrate terpene synthases: their occurrence and physiological significance. *Front Plant Sci*. 2016;**7**: 1–16.
60. Jones CG, Moniodis J, Zulak KG. *et al*. Sandalwood fragrance biosynthesis involves sesquiterpene synthases of both the terpene synthase (TPS)-a and TPS-b subfamilies, including santalene synthases. *J Biol Chem*. 2011;**286**:17445–54.
61. Köllner TG, Schnee C, Gershenzon J. *et al*. The variability of sesquiterpenes emitted from two *Zea mays* cultivars is controlled by allelic variation of two terpene synthase genes encoding stereoselective multiple product enzymes. *Plant Cell*. 2004;**16**: 1115–31.
62. Kampranis SC, Loannidis D, Purvis A. *et al*. Rational conversion of substrate and product specificity in a salvia monoterpene synthase: structural insights into the evolution of terpene synthase function. *Plant Cell*. 2007;**19**:1994–2005.
63. Salmon M, Laurendon C, Vardakou M. *et al*. Emergence of terpene cyclization in *Artemisia annua*. *Nat Commun*. 2015;**6**:4–13.
64. Karunanithi PS, Zerbe P. Terpene synthases as metabolic gatekeepers in the evolution of plant terpenoid chemical diversity. *Front Plant Sci*. 2019;**10**:1–23.
65. Dudareva N, Negre F, Nagegowda DA. *et al*. Plant volatiles : recent advances and future perspectives. *Crit Rev Plant Sci*. 2006;**25**: 417–40.
66. Kumar S, Stecher G, Li M. *et al*. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol*. 2018;**35**:1547–9.
67. Sharma B, Kramer EM. Aquilegia B gene homologs promote petaloidy of the sepals and maintenance of the C domain boundary. *EvoDevo*. 2017;**8**:22.
68. Sharma B, Kramer EM. Virus-Induced Gene Silencing in the Rapid Cycling Columbine Aquilegia coerulea "Origami". In: Becker A, (ed.), *Virus-Induced Gene Silencing*. Clifton, N.J.: Humana Press; New York: Springer, 2013.
69. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C}$T method. *Methods*. 2001;**25**:402–8.
70. Li Y, Shan X, Tong L. *et al*. The conserved and particular roles of the R2R3-MYB regulator FhPAP1 from *Freesia hybrida* in flower anthocyanin biosynthesis. *Plant Cell Physiol*. 2020;**61**: 1365–80.
71. Yang Z, Li Y, Gao F. *et al*. MYB21 interacts with MYC2 to control the expression of terpene synthase genes in flowers of *F. hybrida* and *A. thaliana*. *J Exp Bot*. 2020;**71**:4140–58.
72. Yahyaa M, Matsuba Y, Brandt W. *et al*. Identification, functional characterization, and evolution of terpene synthases from a basal dicot. *Plant Physiol*. 2015;**169**:1683–97.
73. Sainsbury F, Thuenemann EC, Lomonossoff GP. pEAQ: versatile expression vectors for easy and quick transient expression of heterologous proteins in plants. *Plant Biotechnol J*. 2009;**7**: 682–93.