International Journal of
Molecular Sciences

MDPI

Review

Evolution of Terpene Synthases in Orchidaceae

Li-Min Huang [1], Hsin Huang [1], Yu-Chen Chuang [1], Wen-Huei Chen [1,2], Chun-Neng Wang [3] and Hong-Hwa Chen [1,2,*]

1 Department of Life Sciences, National Cheng Kung University, Tainan 701, Taiwan; limin925@gmail.com (L.-M.H.); n34545@gmail.com (H.H.); faseno@gmail.com (Y.-C.C.); wenhueic005@gmail.com (W.-H.C.)
2 Orchid Research and Development Center, National Cheng Kung University, Tainan 701, Taiwan
3 Department of Life Sciences, Institute of Ecology and Evolutionary Biology, National Taiwan University, Taipei 106, Taiwan; LEAFY@ntu.edu.tw
* Correspondence: hhchen@mail.ncku.edu.tw; Tel.: +886-6-275-7575 (ext. 58111); Fax: +886-6-235-6211

Abstract: Terpenoids are the largest class of plant secondary metabolites and are one of the major emitted volatile compounds released to the atmosphere. They have functions of attracting pollinators or defense function, insecticidal properties, and are even used as pharmaceutical agents. Because of the importance of terpenoids, an increasing number of plants are required to investigate the function and evolution of terpene synthases (*TPS*s) that are the key enzymes in terpenoids biosynthesis. Orchidacea, containing more than 800 genera and 28,000 species, is one of the largest and most diverse families of flowering plants, and is widely distributed. Here, the diversification of the *TPS*s evolution in Orchidaceae is revealed. A characterization and phylogeny of *TPS*s from four different species with whole genome sequences is available. Phylogenetic analysis of orchid *TPS*s indicates these genes are divided into *TPS-a*, *-b*, *-e/f*, and *g* subfamilies, and their duplicated copies are increased in derived orchid species compared to that in the early divergence orchid, *A. shenzhenica*. The large increase of both *TPS-a* and *TPS-b* copies can probably be attributed to the pro-duction of different volatile compounds for attracting pollinators or generating chemical defenses in derived orchid lineages; while the duplications of *TPS-g* and *TPS-e/f* copies occurred in a species-dependent manner.

Keywords: terpene synthase; Orchidaceae; evolution; phylogenetic tree

Citation: Huang, L.-M.; Huang, H.; Chuang, Y.-C.; Chen, W.-H.; Wang, C.-N.; Chen, H.-H. Evolution of Terpene Synthases in Orchidaceae. *Int. J. Mol. Sci.* **2021**, *22*, 6947. https://doi.org/10.3390/ijms22136947

Academic Editor: Jen-Tsung Chen

Received: 2 April 2021
Accepted: 23 June 2021
Published: 28 June 2021

1. Introduction

Terpenoids are the largest group of natural metabolites in the plant kingdom, including more than 40,000 different compounds, and have multiple physiological and ecological roles. Terpene metabolites are not only essential for plant growth and development (e.g., gibberellin phytohormones), but also important intermediaries in the various interactions of plants with the environment [1]. For example, chlorophylls and carotenoids are photosynthetic pigments, while brassinosteroids, gibberellic acid, and abscisic acid are plant hormones [2,3]. Terpenoids can be classified based on the number of isoprene units, such as hemiterpene (C5), monoterpene (C10), sesquiterpene (C15), diterpene (C20), sesterterpene (25), triterpene (C30), sesquarterpene (C35), and tetraterpene (C40) (Gershenzon and Dudareva, 2007). The increased number of cyclizations, possibly from a precursor with five additional carbon atoms, gives structural diversity. Terpenoid structures are extremely variable and most of them are low molecular weight like monoterpene (C_{10}), sesquiterpene (C_{15}), and diterpene (C_{20}) [4]. The approximate number of monoterpenes is 1000 and more than 7000 sesquiterpenes [5].

Terepene synthases (TPSs) are key enzymes in terpenoids biosynthesis. To date, TPSs have been studied in several typical plant genomes, such as *Arabidopsis thaliana* (Arabidopsis, 32 *TPS*s) [6], *Physcomitrella patens* (earthmoss, 1 *TPS*) [7], *Sorghum bicolor* (Sorghum, 24 TPSs) [8], *Vitis vinifera* (grape, 69 TPSs) [9], *Solanum lycopersicum* (tomato, 29 TPSs) [10],

Selaginella moellendorffii (spikemoss, 14 TPSs) [11], *Glycine max* (soybean, 23 TPSs) [12] *Populus trichocarpa* (poplar tree, 38 TPSs) [13], *Oryza sativa* (rice, 32 TPSs) [14], and *Dendrobium officinale* (Dendrobium orchid, 34 TPSs) [15]. According to the classification principle, TPSs can be generally classified into seven clades or subfamilies: *TPS-a*, *TPS-b*, *TPS-c*, *TPS-d*, *TPS-e/f*, *TPS-g*, and *TPS-h* [16]. *TPS-a*, *TPS-b*, and *TPS-g* are angiosperm-specific subfamilies, while the *TPS-e/f* subfamily is present in angiosperms and gymnosperms. *TPS-c* exists in land plants. *TPS-d* is a gymnosperm-specific subfamily, and the *TPS-h* subfamily only appears in *Selaginella moellendorffii* [16].

The full length of plant *TPS*s has three conserved motifs on C- and N-terminal regions. The conserved motif of N-terminal domain is R(R)X_8W (R, arginine, W, tryptophan and X, alternative amino acid) and the C-terminal domain contains two highly conserved aspartate-rich motifs. One of them is the DDXXD motif, which is involved in the coordination of divalent ion(s), water molecules, and the stabilization of the active site [17–19]. The second motif in the C-terminal domain is the NSE/DTE motif. These two motifs flank the entrance of the active site and function in binding a trinuclear magnesium cluster [20,21]. Most terpene synthases belong to monoterpene synthase (MTPSs) [22], sesquiterpene synthase (STPSs), and diterpene synthase (DTPSs) [23]. They all share three conserved domains in the active site, including 'DDXXD', 'DXDD', and 'EDXXD'. The 'R(R)X_8W' motif is also essential for monoterpene cyclization, while some MTPSs do not have it [16]. These circumstances can be seen in linalool synthase in rice (*Oryza sativa* L. cv. Nipponbare and Hinohikari) [24]; nerol synthase in soybean (*Glycine max* cv. 'Bagao'), which has a signal peptide and is believed to be functional in plastid [25]; and FaNES1, the cytosolic terpene synthase identified in strawberry, which is able to use cytosolic GDP and FDP to produce linalool and nerolidiol [26].

*TPS*s in the same subfamilies are similar in sequence and have similar functions. Based on the protein sequence, angiosperm STPSs and DTPSs belong to *TPS-a* subfamily and monoterpene synthases belong to *TPS-b* subfamily. Subfamilies in *TPS-c* and *e/f* have enzyme activities of DTPSs; Gymnosperm-specific *TPS-d* subfamily owns the enzyme activities for MTPSs, STPSs, and DTPSs. *TPS-g* encodes MTPSs, STPSs, and DTPSs that produce mainly acyclic terpenoids. *TPS-h* is *Selaginella moellendorffii*-specific subfamily and putative encodes DTPSs [16,27]. Recently, large amounts of TPSs have been identified by using BLAST and thus used for functional characterization assay to further confirm the activity of TPSs. The functions of TPSs can be mono- or multi-functional, and the enzymes can be highly identical to each other. For instance, the DTPs of levopimaradiene/abietadiene synthase and isopimaradiene synthase showed 91% identity in Norway spruce [28]. Moreover, the functional bifurcation of these two enzymes were proved to be caused by only four amino acid residues [28]. Some TPSs are responsible for producing compounds that are related to plant growth and development, such as gibberellin biosynthesis [29], others are responsible in secondary metabolism like monoterpenes and sesquiterpenes for pollination and defense [30,31]. Molecules catalyzed by *TPS* are usually further modified by cytochromes p450 (CYPs) to generate diverse structures [32].

Orchids show extraordinary morphological, structural, and physiological characteristics unique in the plant kingdom [33]. Containing more than 800 genera and 28,000 species, the Orchidaceae, classified in class Liliopsida, order Asparagales, is one of the largest and most diverse families of flowering plants [33]. They are widely distributed wherever sun shines except Antarctica, and with a variety of life forms from terrestrial to epiphytic [34]. According to molecular phylogenetic studies, Orchidaceae comprises five subfamilies, including Apostasioideae, Cypripedioideae, Vanilloideae, Orchidaideae, and Epidendroideae [35]. Orchids emit various volatile organic compounds (VOCs) to attract their pollinators, and/or the enemy of herbivores for olfactory capture. The emitted VOCs are plant secondary metabolites, and the major natural products include terpenoids, phenylpropenoids, benzeniods, and fatty acid derivatives. The floral scent composed of the VOCs plays an important role in plants, such as pollinator attraction, defense, and plant-to-plant communication, especially in insect-pollinated plants [30,36].

Floral VOCs are characterized into several orchids, including α- and β-pinene for *Cycnoches densiflorum* and *C. dianae* [37]; phenylpropanoids in *Bulbophyllum vinaceum* [38]; α-pinene and *e*-carvone oxide for *Catasetum* integerrimum [39]; *p*-dimethoxybenzene for *Cycnoches ventricosum* and *Mormodes lineata* [39]; β-bisabolene and 1,8-cineole for *Notylia barkeri* [39]; *e*-ocimene and linalool for *Gongora galeata* [39]; monoterpenes in *Orchis mascula* and *Orchis pauciflora* [40]; (Z)-11-eicosen-1-ol in *Dendrobium sinense* [41]; terpenoid of (E)-4,8-dimethylnona-1,3,7-triene (DMNT) in *Calanthe sylvatica* [42] and *Cyclopogon elatus* [43]; (*E*)-β-ocimene and (*E*)-epoxyocimene for *Catasetum cernuum* and *Gongora bufonia* [44]; and farnesol, methyl epi-jasmonate, nerolidol, and farnesene in *Cymbidium goeringii* [45].

Phalaenopsis spp. is very popular worldwide for its spectacular flower morphology and colors. Most *Phalaenopsis* orchids are scentless but some do emit scent VOCs [46]. The scented species have been extensively used as breeding parents for the production of scented cultivars, such as *P. amboinensis*, *P. bellina*, *P. javanica*, *P. lueddemanniana*, *P. schilleriana*, *P. stuartiana*, *P. venosa*, and *P. violace* [47]. *P. bellina* and *P. violacea* are two scented orchids that are very popular in breeding scented cultivars. *P. bellina* emits mainly monoterpenoids, including citronellol, geraniol, linalool, myrcene, nerol, and ocimene [47,48], while *P. violacea* emits monoterpenoids accompanied with a phenylpropanoid, cinnamyl alcohol [46]. The VOCs of *P. schilleriana* contain monoterpenoids as well, including citronellol, nerol, and neryl acetate [49]. Because of the importance of terpenoids in plants, an increasing number of plants are required to investigate the function and evolution of *TPS*s.

In the present review, we summarized the recent progress in the understanding of the biosynthesis and biological function of terpenoids, and the latest advances in research on the evolution and functional diversification of *TPS*s in Orchidaceae. TPSs from different orchid species are reported to explore the evolutionary history and the evolution diversification of Orchidaceae *TPS*s.

2. Terpenoids and Their Biosynthesis in Plants

There are two compartmentalized terpenoid biosynthesis pathways, the mevalonic acid (MVA) pathway that occurs in the cytosol, and the methylerythritol phosphate (MEP) pathway that occurs in plastids to produce isopentenyl diphosphate (IPP) and its allylic isomer-dimethylallyl diphosphate (DMAPP) converted by isopentenyl diphosphate isomerase (IDI) (Figure 1) [50–52]. There are four major steps involved in the biosynthesis of terpenoid, beginning with isoprene unit (IPP) formation, which has five carbons. Second, IPP combines to DMAPP by geranyl diphosphate synthase (GDPS), geranylgeranyl diphosphate synthases (GGDPS) or farnesyl diphosphate (FDPS), and generates geranyl diphosphate (GDP), farnesyl diphosphate (FDP) or geranylgeranyl diphosphate (GGDP), respectively [1,27,53,54]. Third, the C_{10}-C_{20} diphosphates go through cyclization and rearrangement to produce the basic carbon skeletons for terpenoids catalyzed by *TPS* [53]. The *TPS* family consists of enzymes that use GDP to form cyclic and acylic monoterpenes (C_{10}), FDP for sesquiterpene (C_{15}), and GGDP for diterpene (C_{20}) [16]. Moreover, FDP and GGDP can be dimerized to form the precursors of C_{30} and C_{40}. The final step converts terpenes into different skeletons by oxidation, reduction, isomerization, conjugation, and other transformation [53]. *TPS*s are the key enzymes in terpenoid biosynthesis.

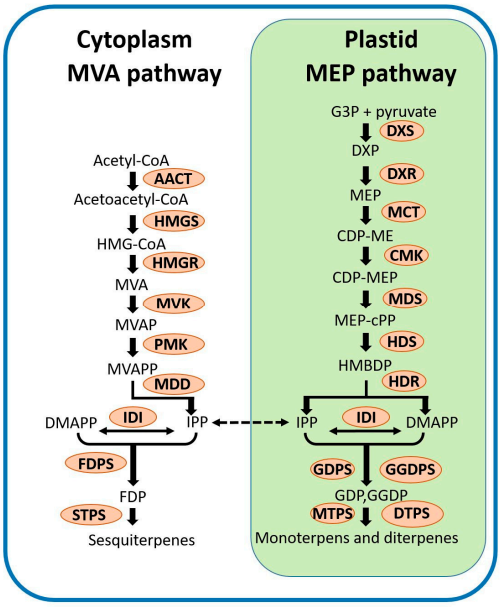


Figure 1. The MVA (left) and MEP (right) pathways responsible for IPP and DMAPP biosynthesis and monoterpene biosynthesis in plants. AACT, acetoacetyl-CoA thiolase; CMK, 4-(cytidine 5′ -diphospho)-2-*C*-methyl-d-erythritol kinase; DMAPP, dimethylallyl diphosphate; DXR, 1-deoxy-d-xylulose 5-phosphate reductoisomerase; DXS, 1-deoxyd- xylulose 5-phosphate synthase; FDP, farnesyl diphosphate; FPPS, farnesyl diphosphate synthase; G3P, d-glyceraldehyde 3-phosphate; GDPS, geranyl diphosphate synthase; GDP, geranyl diphosphate; HDR, (*E*)-4-hydroxy-3-methylbut-2-enyl diphosphate reductase; HDS, (*E*)-4-hydroxy-3-methylbut-2-enyl diphosphate synthase; HMGR, 3-hydroxy-3-methylglutaryl-CoA reductase; HMGS, 3-hydroxy-3-methylglutaryl- CoA synthase; IDI, isopentenyl diphosphate isomerase; IPP, isopentenyl diphosphate; MCT, 2-*C*-methyl-d-erythritol 4-phosphate cytidylyltransferase; MDD, mevalonate diphosphate decarboxylase; MDS, 2-*C*-methyld-erythritol 2,4-cyclodiphosphate synthase; MVK, mevalonate kinase; MVAP, mevalonate 5-phosphate; MVAPP, mevalonate diphosphate; PMK, phosphomevalonate kinase; TPS, terpene synthase.

3. The Evolution of *TPS* Genes in Orchidaceae Species

We chose the whole genome sequences of four orchids, including *A. shenzhenica* [54] in Apostasioideae subfamily; *Vanilla planifolia* [55] in Vanilloideae subfamily; and *D. catenatum* [56] and *P. equestris* [57] in Epidendroideae subfamily. There were two justifications for this selection. First, these four orchids are distributed into three different subfamilies, and their whole genome sequences are available in NCBI (https://www.ncbi.nlm.nih.gov/ (accessed on 6 January 2021).) and OrchidBase database [58] (http://orchidbase.itps.ncku.edu.tw/est/home2012.aspx (accessed on 9 August 2020).). Second, *A. ashenzhenica* is the most original orchid, and *P. equestris* is the first whole genome sequenced orchid. *V. planifolia* produces vanillin and is important in the food industry, and *D. catenatum* is a medicinal orchid and produces important secondary metabolites for pharmaceutical purpose. We isolated the *TPS* genes of Orchidaceae through KAAS (http://www.genome.jp/tools/kaas/ (accessed on 21 February 2017).) annotation and BLASTp from the whole genome sequences of four orchids. Each full-length *TPS* is characterized by two conserved domains with Pfam [59] ID PF01397 (N-terminal) and PF03936 (C-terminal) [17]. A total of 9, 27, 35, and 15 *TPS* genes were identified from the whole genome sequences of *A. shenzhenica*, *V. planifolia*, *D. catenatum*, and *P. equestris*, respectively. In addition, *P. aphrodite* with white, scentless flowers and *P. bellina* scented flowers are native species. Their floral transcriptomes are available in Orchidstra and OrchidBase transcriptome database, respectively. 17 *TPS* genes in *P. aphrodite* and 11 *TPS* genes in *P. bellina* were identified from the transcriptome database. The *TPS* genes were denoted with numbers *Ash-*, *KAG-*, *Dca-*, *Peq-*, *PATC-*, and *PbTPS-* identified from *A. shenzhenica*, *V. planifolia*, *D. catenatum*, *P. equestris*, *P. aphrodite*, and *P. bellina*, respectively.

*TPS*s in *P. equestris* and *D. officinale* have been reported [15,60]. These *TPS*s are divided into four subfamilies (*TPS-a*, *TPS-b*, *TPS-c*, and *TPS-e/f*). So, we further investigated *TPS* evolution in Orchidaceae and provided insight into *TPS*s at the genome level. In this review, the encoded amino acid sequences of identified orchid *TPS* genes were aligned with those from *Arabidopsis* and *Abies grandis*, and those from *Selaginella moellendorffii* were used as outgroups (Appendix A Table A1). The phylogenetic tree was constructed using Neighbor-Joining method with Jones–Taylor–Thornton model and pairwise deletion with 1000 bootstrap replicates by using MEGA7 software. The orchid *TPS*s are grouped into *TPS-a*, *-b*, *-e/f*, and *g* subfamilies (Figure 2). Most of the orchid *TPS*s belong to *TPS-a* and *TPS-b* subfamilies (89/115, Table 1). In the *TPS-a* subfamily, copies from dicot and monocot species formed distinct subgroups, which is in accordance to previous studies [15,16]. However, compared to angiosperm dicot species, which have more *TPS*s in *TPS-a* subfamily, orchid (monocot) *TPS*s have more members in *TPS-b* subfamily than in *TPS-a* subfamily. Within *TPS-b* subfamily, these orchid *TPS*s form distinct clades separated from those of *Arabidopsis* (dicot) *TPS*s (Figure 2). Taken together, the persistence of dicot and monocot distinct clades within *TPS-a* and *TPS-b* implies that these *TPS*s have diverged since the ancestor of angiosperm. On the other hand, most of the duplicated orchid *TPS-a* and *TPS-b* copies were species-dependent (i.e., paralogs duplicated within each species). In particular, the number of duplicated orchid *TPS-a* and *TPS-b* copies increased in *V. planifolia* and *D. catenatum* (Figure 2). These data suggest that *TPS-a* and *TPS-b* copies evolved in a species-dependent manner and may have been positively selected to generate exceptionally more multiple copies. *TPS-a* and *TPS-b* are angiosperm-specific subfamilies that are responsible for sesquiterpene or diterpene and monoterpene synthases. These orchid volatile terpenes have critical roles in producing floral scents in order to be attractive to pollinators and to respond to environmental stresses [15]. It is therefore not surprising that *TPS-a* and *TPS-b* subfamilies have diverged greatly in orchid species.

Our phylogenetic analysis also reveals that the orchid *TPS-e/f* subfamily has increased copy numbers compared to that from *A. thaliana* (Table 1; Figure 2). Orchid *TPS-g* subfamily can only be found in *A. shenzhenica* and *V. planifolia* (Table 1; Figure 2), whereas those Epidendroideae *TPS-g* members have perhaps been lost during evolution. There are no orchid *TPS*s in *TPS-c* group that host copalyl diphosphate synthases (CPS) of angiosperm [61]. *TPS-d* and *TPS-h* are gymnosperm and *Selaginella moellendorffii* specific, respectively [16]. Our analysis showed that no orchid *TPS*s were grouped in these subfamilies, in accordance with previous conclusions by Chen et.al, and Trapp et.al. [16,62].

Motifs of identified orchid TPS proteins were predicted using MEME software (https://meme-suite.org/meme/tools/meme (accessed on 19 March 2021).) (Figure 3A), and five major functional conserved motifs of *TPS*s (R(R)X_8W, EDXXD, RXR, DDXXD, and NSE/DTE) were elucidated (Figure 3B). The *TPS-a* subfamily that encodes STPSs is mainly found in both dicot and monocot plants [9,11,16,63]. In this subfamily, STPSs contain the non-conserved secondary "R" (arginine) of motif R(R)X_8W that functions in the initiation of the isomerization cyclization reaction [64], or in stabilizing the protein through electrostatic interactions [65]. Compared with *Arabidopsis*, most orchid *TPS*s contain motif R(R)X_8W, except PATC144727, Peq011664, Dca017107, and PATC155674 in *TPS-a* subfamily (Figure 4A). In contrast, the angiosperm-specific *TPS-b* subfamily that encodes MTPSs contains the highly conserved R(R)X_8W motif. All *TPS*s in *Arabidopsis* TPS-b subfamily contain conserved R(R)X_8W motif, except AtTPS02 (Figure 4B). However, several members of orchid TPS-b subfamily have lost the conserved R(R)X_8W motif (Figure 4B). Motifs EDXXD, RXR, DDXXD, and NSE/DTE are highly conserved in *TPS-a* and *-b* subfamilies, while the conserved R(R)X_8W motif of orchid *TPS*s is divergent in *TPS-b* subfamily.

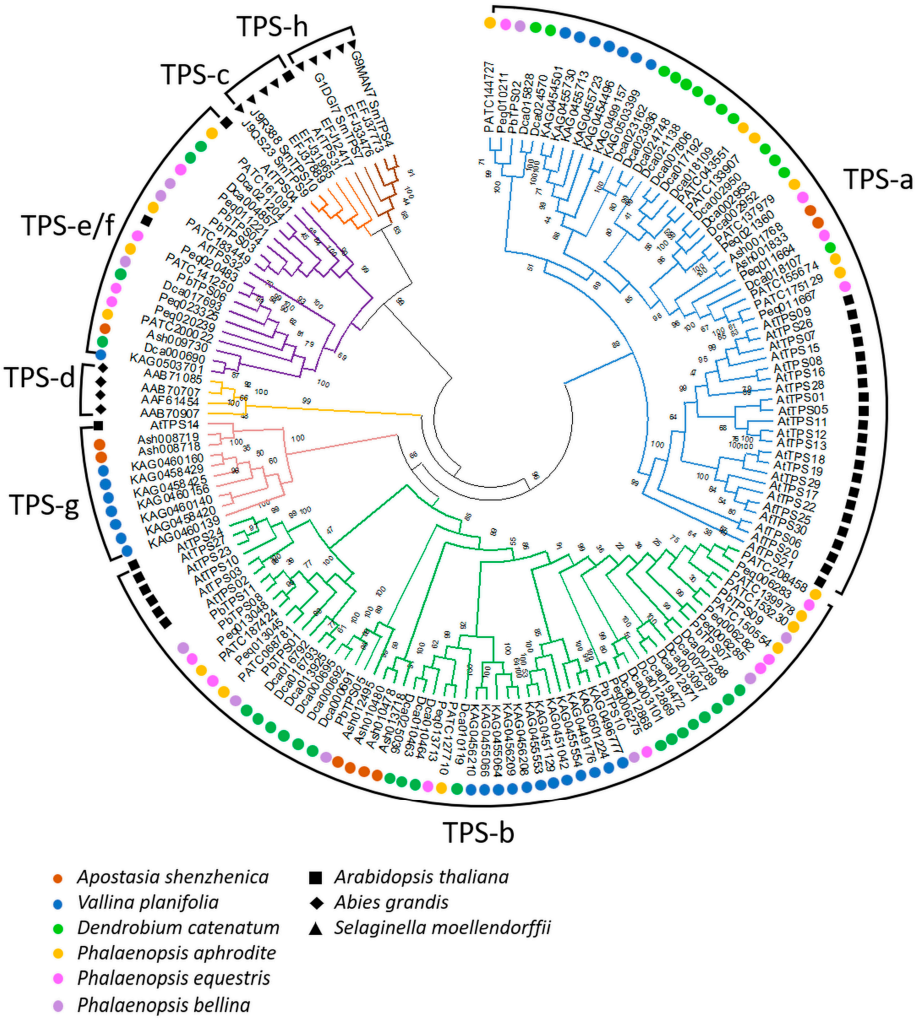
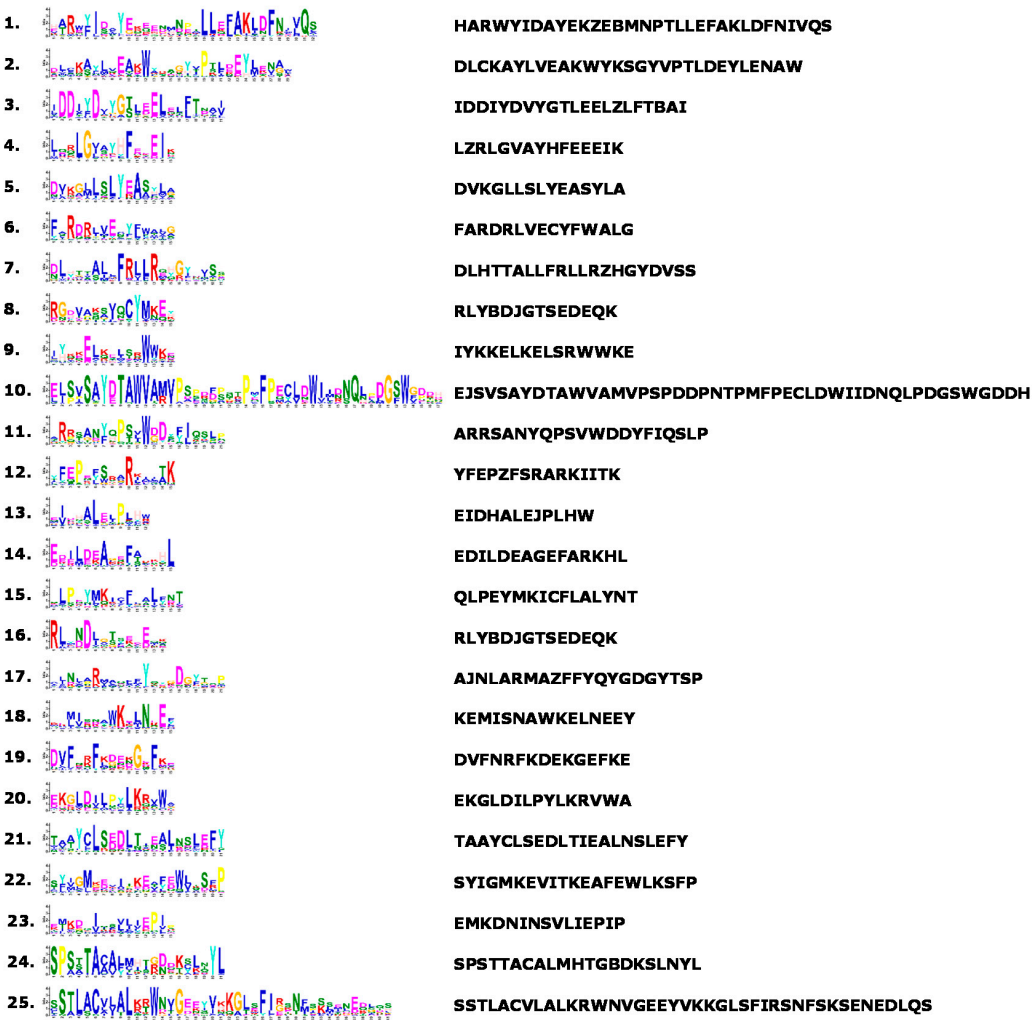


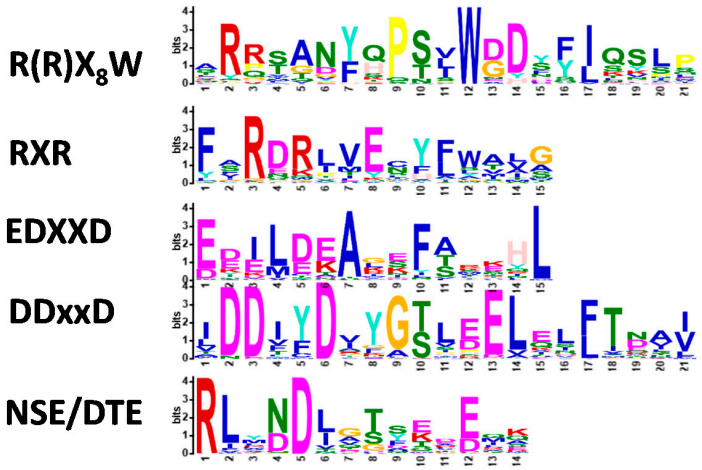
Figure 2. Phylogenetic analysis of terpene synthases. *TPS*s in Orchidaceae, including *A. shenzhenica*; *V. planifolia*; *D. catenatuml P. equestris Phalaenopsis aphrodite*; *P. bellina*, *Arabidopsis thaliana*, and *Abies grandis*; and *S. moellendorffii* were used. Sequence analysis was performed using MEGA 7.0 to create a tree using the nearest neighbor-joining method. The coding sequence was used for analysis. The numbers at each node represent the bootstrap values. Various colors mean distinct subfamilies and special symbols represent different plant species, with solid circles, tangle, diamond, and triangle illustrating Orchidaceae, *Arabidopsis thaliana*, *A. grandis*, and *S. moellendorffii*, respectively.

Table 1. The number of TPS*s* subfamilies in Orchidaceae and other plant species.

	***TPS* Subfamily**								
Species	***a***	***b***	***c***	***d***	***e/f***	***g***	***h***	**Total**	**Reference**
Apostasia shenzhenica	2	4	0	0	1	2	0	9	This research
Vallina planifolia	7	12	0	0	1	7	0	27	This research
Dendrobium catenatum	13	18	0	0	4	0	0	35	This research
Phalaenopsis equestris	4	7	0	0	4	0	0	15	This research
Phalaenopsis aphrodite	6	7	0	0	4	0	0	17	This research
Phalaenopsis bellina	1	7	0	0	3	0	0	11	This research
Arabidopsis thaliana	22	6	1	0	2	1	0	32	Aubourg et al. (2002) [6]
Solanum lycopersicum	12	8	2	0	5	2	0	29	Falara et al. (2011) [10]
Oryza sativa	18	0	3	0	9	2	0	32	Chen et al. (2014) [14]
Sorghum bicolor	15	2	1	0	3	3	0	24	Paterson et al. (2009) [8]
Vitis vinifera	30	19	2	0	1	17	0	69	Martin et al. (2010) [9]
Populus trichocarpa	16	14	2	0	3	3	0	38	Irmisch et al., (2014) [13]
Selaginella moellendorffii	0	0	3	0	3	0	8	14	Li et al., (2012) [11]

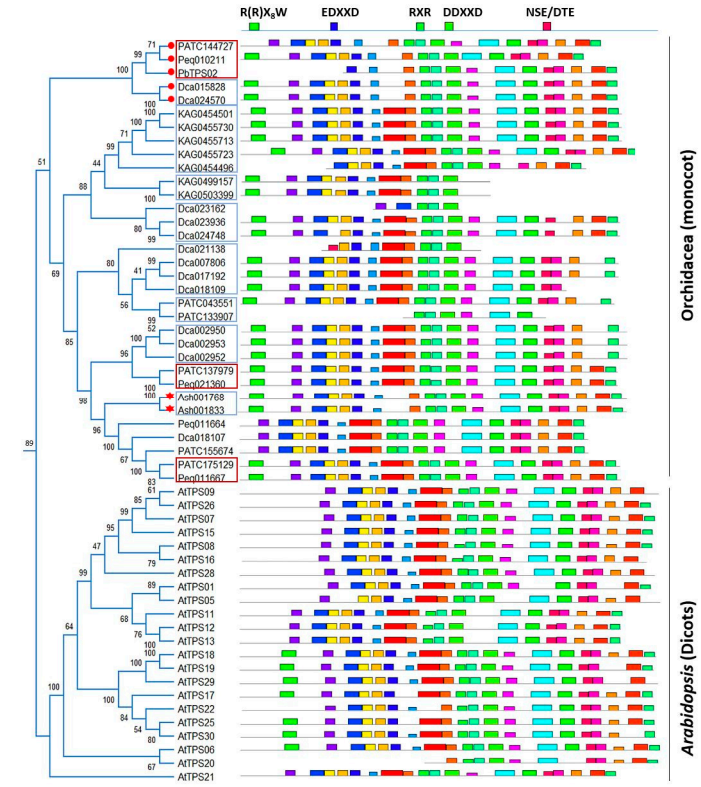


(**A**)

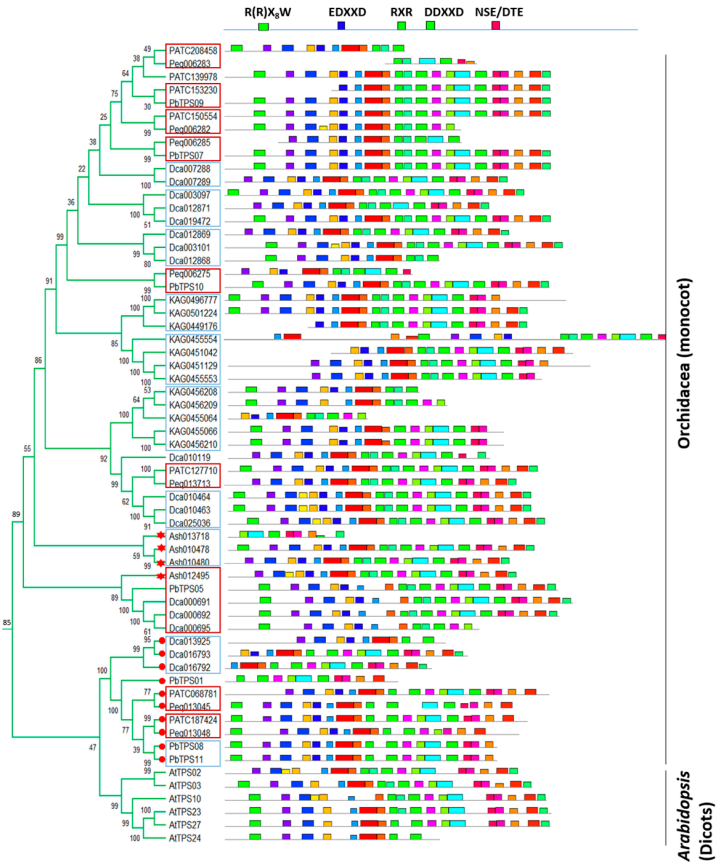


(**B**)

Figure 3. The amino acid sequences of the predicted motifs in *TPS* proteins. (**A**) Twenty-five classical motifs in *TPS* proteins were analyzed using the MEME tool. The width of each motif ranges from 6 to 50 amino acids. The font size represents the strength of conservation. (**B**) The amino acid sequences of five highly conserved motifs in *TPS* proteins.



(**A**) TPS-a



(**B**) TPS-b

Figure 4. *Cont.*

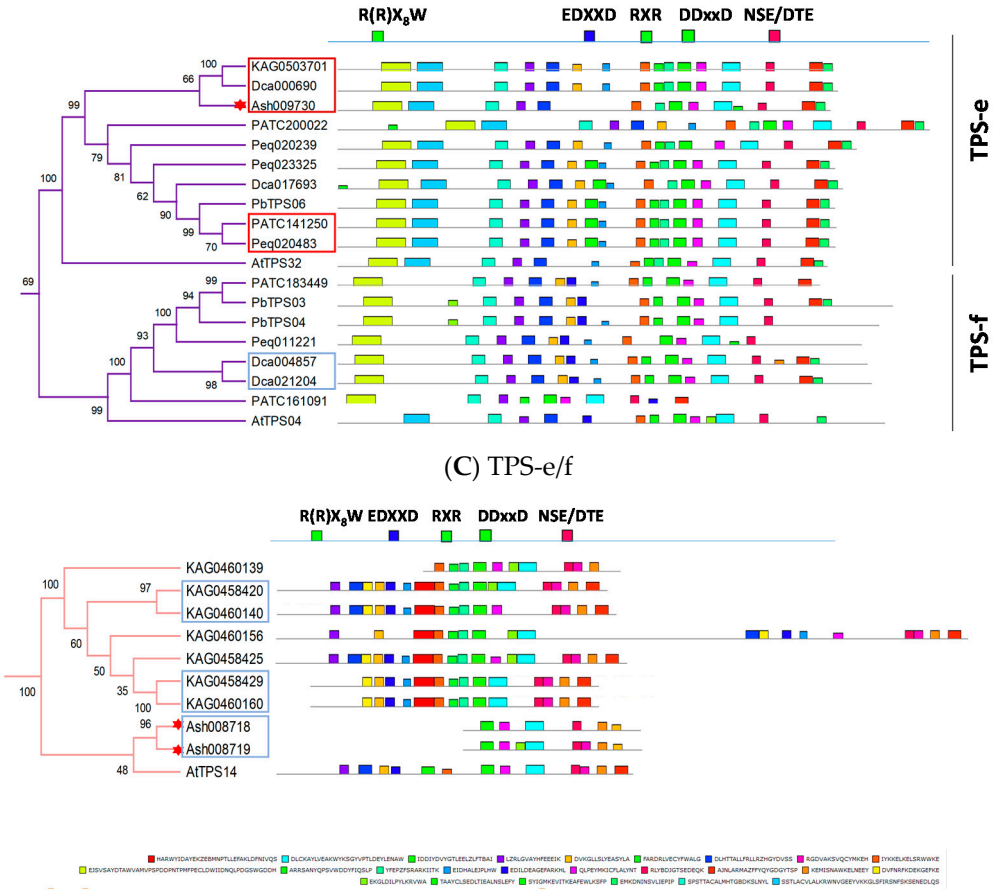


Figure 4. Motif structures of *TPS* proteins. (**A**–**D**) are *TPS-a*, *-b*, *-e/f*, and *-g* subfamilies, respectively. Twenty-five classical motifs in *TPS* proteins were analyzed by using the MEME tool. The width of each motif ranged from 6 to 50 amino acids. Different color blocks represent distinct motifs. Star indicates *TPS*s of *A. shenzhenica*, and the red solid circle indicates the out group of *Apostasia TPS*s. The red and blue rectangle squares reveal orthologous and paralogous gene pairs, respectively.

DTPSs are evolved from kaurene synthase (KS) and CPS. MTPSs and STPSs are evolved from ancestral DTPS through duplication and then sub- or neo-functionalization during evolution [66]. *A. shenzhenica* has clear evidence of whole-genome duplication that is shared by all orchids [54]. Yet, the copies of *TPS* in *A. shenzhenica* are among the fewest and are worthwhile for further investigation. For *Phalaenopsis* orchids, paralogs of *TPS* genes could be identified from each species, implying the duplications were attributed to their common ancestor, and some persisted or lost in current species (Figure 4). For example, *TPS-a* copies of *P. aphrodite*, *P. bellina*, and *P. equestris* species can be found (some lost) in three parallel clades of the phylogenetic tree (*PATC144727/Peq010211/PbTPS02*, *PATC137979/Peq021360*, and *PATC175129/Peq011667*) (red tangle, Figure 4A). Similarly, *TPS-b* copies of *P. aphrodite*, *P. bellina*, and *P. equestris* can be repeatedly identified (some lost) in eight parallel clades, indicating the *TPS-b* gene copy duplications could be traced back to the common ancestor of *Phalaenopsis* species (*PATC208458/Peq006283*, *PATC153230/PbTPS09*, *PATC150554/Peq006282*, *Peq006285/PbTPS07*, *Peq006275/PbTPS10*, *PATC127710/Peq013713*, *PATC068781/Peq013045* and *PATC187424/Peq013048*) (red tangle, Figure 4B).

Members of *TPS-e/f* subfamilies are mainly detected in angiosperm and conifers DTPSs of primary metabolism (i.e., gibberellin biosynthesis) [16,67]. Orchid *TPS-e/f* subfamilies comprise orthologous genes without R(R)X_8W (Figure 4C), which are consistent with *Arabidopsis*. The *Ash009730* in *TPS-e/f* subfamily, predicted to be KS, was grouped with *KAG0503701* and *Dca000690* (red retangle with red star, Figure 4C). No *TPS*s were found

in *A. shenzhenica* in *TPS-f* subclade. As copies of these orchid *TPS-e/f* subfamilies were duplicated within each species, the duplications seem to be species dependent.

TPS-g subfamily is closely related to the *TPS-b* but lacks the N-terminal "R(R)X_8W" motif and encodes MTPSs, STPSs, and DTPSs that produce mainly acyclic terpenoids [68,69]. A highly conserved arginine-rich RXR motif of sesquiterpene synthase reported that the motif is involved in producing a complex with the diphosphate group after the ionization of FPP in sesquiterpene biosynthesis [70]. *TPS-g* subfamily in *Arabidopsis* (AtTPS14) lacks both "R(R)X_8W" and "RXR" motifs. However, although *TPS*s of *V. planifolia* in *TPS-g* subfamily (those started with KAG in Figure 4D) lack the N-terminal "R(R)X_8W" motif, they still have the "RXR" motif (Figure 4D). This suggests that *TPS-g* subfamily of *V. planifolia* may have conserved enzyme activities that are capable of accepting a multi-substrate in terpene biosynthesis.

The pharmaceutical effective compounds in *D. catenatum*, a widely used Chinese herb, belong to terpenoid indole alkaloid (TIA) class [71], and many of them contain a terpene group. A sesquiterpene alkaloid-Dendrobine found in *Dendrobium* is believed to be responsible for its medical property [71]. Concomitantly, a significant increased number of *TPS-a TPS*s was detected in *D. catenatum*as as compared to that of other orchid species, which is responsible for sesquiterpene biosynthesis (Table 1). The increased number of *TPS-b* in *Dendrobium* may cause the floral fragrance in *D. catenatum* as well as the formation of TIA. *P. bellina* is a scented orchid with the main floral compounds of monoterpenes including linalool, geraniol, and their derivatives, which attract pollinators [48]. *PbTPS*s from the floral transcriptome database are majorly classified into the *TPS-b* subfamily (Table 1). Previously, the expression of both *PbTPS5* and *PbTPS10* were concomitant with the VOCs (monoterpene linalool and geraniol) emission in *P. bellina* [72]. This suggests that these genes may be involved in the biosynthesis of monoterpene in *P. bellina*. *TPS-e/f* enzymes have diverse functions, including linalool synthase, geranyllinalool synthase, and farnesene synthase in kiwifruit [73,74]. *TPS*s in the *TPS-e/f* subfamily are thought to be dicot-specific because so far no *TPS-e/f* activity has been reported in monocots. However, the number of TPS in *TPS-e/f* expands from 1 in *Apostasia* to 4 in *Phalaenopsis* (Table 1), suggesting that the duplication events of *TPS- b* and *TPS-e/f* have evolved in response to natural selection.

Together, our analyses suggest that orchid TPS*s* in each subfamily evolved from the early divergence orchid species, such as *A. shenzhenica* and/or *V. planifolia*. The large expansion of *TPS* copies in orchid groups such as *V. planifolia*, *D. catenatum*, and *Phalaenopsis* species might be due to high flexibility for adaptation and evolution through natural selection.

4. The Arrangement of *TPS*

The functional cluster phenomenon of *TPS* genes was detected in orchids. Orchid *TPS* gene clusters diverged with tandem or segmental duplications (Figure 5). Tandem duplication inferred that the duplication occurred in the same scaffold, such as *Ash012495* grouped with *Dca000691/Dca000692/Dca000697* cluster genes in *TPS-b* subfamily (Figures 4B and 5C). *TPS* genes duplicated on different scaffolds is thought to be segmental duplication, *e.x.*: *Ash008718/Ash008719* grouped with two cluster genes of *V. planifolia* (*KAG0458420/KAG0458425/KAG0458429 and KAG0460140/KAG0460156/KAG0460160*) in different scaffolds in the *TPS-g* subfamily (Figures 4D and 5A,B). We identified that 6, 24, 20, and 8 *TPS*s in *A. shenzhenica*, *V. planifolia*, *D. catenatum,* and *P. equestris*, respectively, form clusters in the same genome scaffold (Table 2, Figure 5A–D). In addition, these clusters were present with *TPS*s of the same subfamily and therefore the enhancement of functions was predicted. In *A. shenzhenica, V. planifolia*, *D. catenatum*, and *P. equestris*, *TPS* genes have three, nine, eight, and three clusters, respectively (Table 2, Figure 5). Each cluster contains two *TPS* genes in *A. shenzhenica*, while more genes are present in the clusters of *V. planifolia*, *D. catenatum,* and *P. equestris* (Figure 4). *TPS* genes in the same cluster usually belong to the same subfamily except that *V. planifolia* has one large scaffold containing *TPS* genes

of *TPS-a*, *TPS-b,* and *TPS-e/f* subfamilies, yet with huge distance between each subfamily cluster (44 Mb and 5 Mb, respectively). The percentages of clustered *TPS* genes were 66.7%, 81.5%, 57.1%, and 53.3% for *A. shenzhenica*, *V. planifolia, D. catenatum,* and *P. equestris*, respectively, while that was 40.6% in *Arabidopsis thaliana* (Table 2). The cluster density of orchid TPS*s* could infer the event of *TPS* gene duplication occurred during evolution. The genome sizes of *A. shenzhenica*, *V. planifolia*, *D. catenatum,* and *P. equestris* are 349 Mb, 7449 Mb, 1104 Mb, and 1064 Mb, respectively (Table 3). The cluster densities of TPS*s* in orchids were 47.3%, 78.6%, 50.5%, and 38.9% for *A. shenzhenica*, *V. planifolia*, *D. catenatum,* and *P. equestris*, respectively (Table 3). Interestingly, orchids have more clusters and higher *TPS* gene density as compared to that of *Arabidopsis*, with that of *V. planifolia* having the highest cluster gene density of *TPS* among the four orchids analyzed. Even though *TPS*s copies of derived orchids (*D. catenatum* and *Phalaenopsis* spp.) were increased compared with those in *A. shenzhenica*, the total number was not linked to the increased genome size.

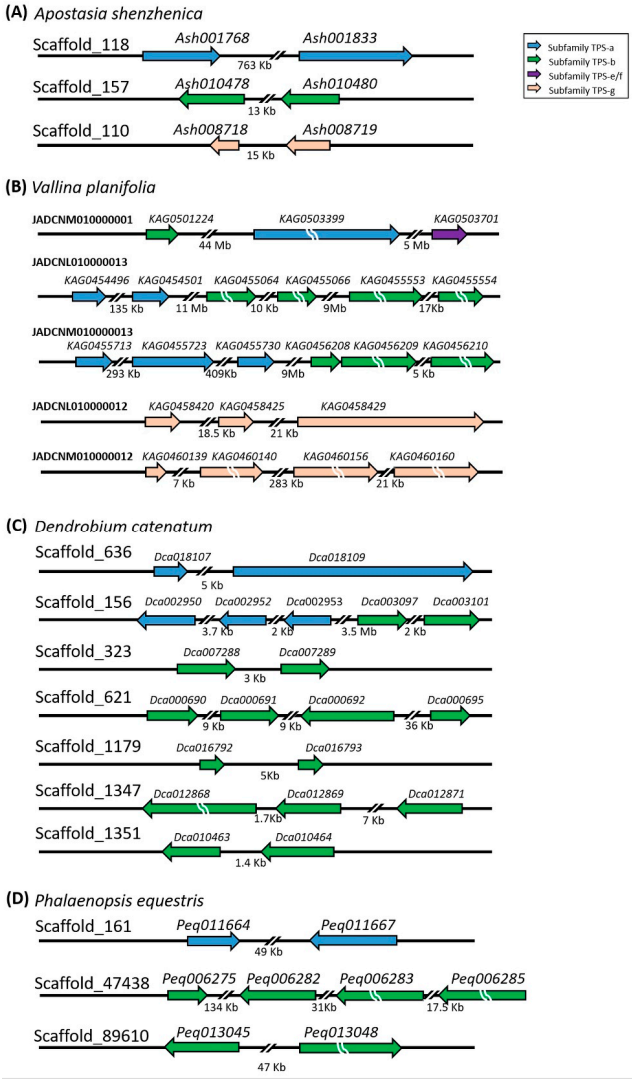


Figure 5. Gene clusters in Orchidaceae genome. Clustered genes in the genomic scaffolds of *A. shenzhenica* (**A**), *V. planifolia* (**B**), *D. catenatum* (**C**), and *P. equestris* (**D**), respectively. The *TPS* genes located on the scaffolds are identified from the assembled whole genome sequences of *A. shenzhenica*, *V. planifolia*, *D. catenatum,* and *P. equestris*. The direction of arrows illustrates the forward translation of genes in the scaffolds. Various colors indicate the distinct *TPS* subfamilies. Blue, green, purple, and bisque colors represent *TPS* genes in *TPS-a*, *-b*, *-e/f*, and *-g* subfamilies, respectively. Break lines indicate the shrink length of genes.

Table 2. The gene clusters of TPS*s* in the genome of Orchidaceae and *Arabidopsis thaliana*.

Species	Number of Clusters	Number of Scaffolds	Number of Clustered *TPS*s	Number of Total *TPS*s	Percentage of Clustered *TPS*s (%)
Apostasia shenzhenica	3	3	6	9	66.7
Vallina planifolia	7	5	22	27	81.5
Dendrobium catenatum	8	7	20	35	57.1
Phalaenopsis equestris	3	3	8	15	53.3
Arabidopsis thaliana [6]	5	5	13	32	40.6

Table 3. The gene density of TPS*s* in the genome of Orchidaceae and other plant species.

Species	Genome Size (Mb)	Cluster Length of *TPS*s (Kb)	Total Length of *TPS*s (Kb)	Cluster Density of *TPS*s (%)
Apostasia shenzhenica	349	26	56	47.3
Vallina planifolia	744	595	758	78.6
Dendrobium catenatum	1104	125	248	50.5
Phalaenopsis equestris	1064	62	158	38.9
Arabidopsis thaliana	120	43	109	39.9

In plants, gene clusters were often observed for metabolic pathways, such as gene clusters found in oat and *Arabidopsis* related to triterpene biosynthesis pathway [75]. Local duplication of *TPS* gene families in plants has been described and often results in tandem repeats, as an important driver for the expansion [16,76]. The genes related in terpene synthesis are usually lined together, forming functional clusters in plants [77]. The functional clusters of *TPS* genes have already been reported in several plant species, such as *Arabidopsis thaliana* [6], *Vitis vinifera* [9], *Solanum lycopersicum* [77], *Eucalpyus grandis* [78], and rice [79,80]. Genomic clusters of *TPS* genes in *E. grandis* are up to 20 genes [78]. In several *Solanum* species, the gene duplications and divergence give rise to *TPS* gene clusters for terpene biosynthesis [77]. A dense cluster of 45 *V. vinifera TPS*s are present on chromosome 18 [9]. *Arabidopsis TPS* genes are reported with the phenomenon of several gene clusters [6]. In addition, a gene cluster with three *TPS* members, including *Os08g07080*, *Os08g07100*, and *Os08g07120*, is observed in Asian rice *Oryza sativa* and also appears in various rice species including *O. glaberrima*, *O. rufipogon*, *O. nivara*, O. *barthii*, and *O. punctata*. [80]. Both conserved and species-specific expression patterns of the clustered rice *TPS*s indicate the functions in insect-damaged plants [80]. The expression of these rice *TPS* genes and their catalytic activities for emission patterns of volatile terpenes is induced by insect damage and is largely consistent [80]. Interestingly, the evolution of *TPS*s with other biosynthesis-related genes was also found to form unexpected connection with time passed. For instance, the evolution of *TPS/CYP* pairs is different in monocot and dicot [81]. *TPS/CYP* pairs duplicate with ancestral *TPS/CYP* pairs as templates to be evolved in dicots, but the evolutionary mechanism of monocot shows that the genome rearrangement of *TPS* and *CYP* occurred independently [81]. In *Solanum* spp., *TPS* forms functional clusters with *cis*-prenyl transferase [77]. Both tandem and segmental duplications significantly contribute toward family expansion and expression divergence and play important roles in the survival of these expanded genes. A functional gene cluster is a group of closely-related genes lined together in a genome, and the study of gene clusters is important for the understanding of evolution within species.

Together, the orchid *TPS* genes formed genomic clusters, and the clusters increased in *V. planifolia* and *D. catenatum*. Combining the results from phylogenetic analysis and functional gene clusters, orchid *TPS*s may be expanded by tandem or segmental duplications. Interestingly, the genome duplication events occurred all the way along the evolution from Apostasioideae to Vanilloideae and Epidendroideae; the *TPS* clusters and copy numbers increased in orchid lineages, such as the early divergence *A. shenzhenica*. The large expansion

of orchid *TPS* copies in *V. planifolia*, and *D. catenatum* species might have high flexibility in secondary biosynthesis through natural selection.

5. Conclusions

The basic evolution of *TPS* is from duplication and loss of *TPS* genes. In Orchidaceae, we discover that the duplication event of *TPS* occurred among all *TPS* subfamilies. *TPs-a*, *TPS-b*, and *TPS-e/f* subfamilies went through gene duplication, while *TPS-g* duplicated from Apostaceae to Vaniloideae, and then lost from Vaniloideae to Epidendroideae. The driving force of *TPS* evolution in each subfamily may be different. For example, in *TPS-a* and *TPS-b*, the necessity of generating volatile compounds for the interaction of orchids with their pollinators, producing chemical defenses and being responsive to environmental stress, may be the major reason for their rapid evolution. On the other hand, the duplications of *TPS-g* and *TPS-e/f* copies were mainly species dependent and the reason remains to be uncovered.

Author Contributions: L.-M.H. performed the phylogenetic analysis and motif prediction of *TPS*s; H.H. performed the gene arrangement analysis; Y.-C.C. performed the identification of orchid *TPS*s; W.-H.C. provided the suggestions for plant materials.; C.-N.W. provided discussion and composed the *TPS*s evolution; H.-H.C. conceived research plans and composed the article with assistances of all the authors, completed the writing, and served as the corresponding author for communication. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by a grant from Ministry of Science and Technology, Taiwan (MOST 107-2313-B-006-003-MY3) to H.-H.C.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: We thank the people that finished the whole genome sequence of the four orchid species, which allowed us to undertake this detail analysis.

Conflicts of Interest: No conflict of interest declared.

Appendix A

Table A1. *TPS* genes used in phylogenetic analysis.

Species	Gene ID	Accession Number of *TPS* Gene
Apostasia shenzhenica [1]	*Ash001768*	Ash001768
	Ash001833	Ash001833
	Ash008718	Ash008718
	Ash008719	Ash008719
	Ash009730	Ash009730
	Ash010478	Ash010478
	Ash010480	Ash010480
	Ash012495	Ash012495
	Ash013718	Ash013718
Vallina planifolia [2]	*KAG0449176*	KAG0449176
	KAG0451042	KAG0451042
	KAG0451129	KAG0451129
	KAG0454496	KAG0454496
	KAG0454501	KAG0454501
	KAG0455064	KAG0455064
	KAG0455066	KAG0455066
	KAG0455553	KAG0455553
	KAG0455554	KAG0455554
	KAG0455713	KAG0455713

Table A1. *Cont.*

Species	Gene ID	Accession Number of *TPS* Gene
	KAG0455723	KAG0455723
	KAG0455730	KAG0455730
	KAG0456208	KAG0456208
	KAG0456209	KAG0456209
	KAG0456210	KAG0456210
	KAG0458420	KAG0458420
	KAG0458425	KAG0458425
	KAG0458429	KAG0458429
	KAG0460139	KAG0460139
	KAG0460140	KAG0460140
	KAG0460156	KAG0460156
	KAG0460160	KAG0460160
	KAG0496777	KAG0496777
	KAG0499157	KAG0499157
	KAG0501224	KAG0501224
	KAG0503399	KAG0503399
	KAG0503701	KAG0503701
Dendrobium catenatum [1]	*Dca000690*	Dca000690
	Dca000691	Dca000691
	Dca000692	Dca000692
	Dca000695	Dca000695
	Dca002950	Dca002950
	Dca002952	Dca002952
	Dca002953	Dca002953
	Dca003097	Dca003097
	Dca003101	Dca003101
	Dca004857	Dca004857
	Dca007288	Dca007288
	Dca007289	Dca007289
	Dca007806	Dca007806
	Dca010119	Dca010119
	Dca010463	Dca010463
	Dca010464	Dca010464
	Dca012868	Dca012868
	Dca012869	Dca012869
	Dca012871	Dca012871
	Dca013925	Dca013925
	Dca015828	Dca015828
	Dca016792	Dca016792
	Dca016793	Dca016793
	Dca017192	Dca017192
	Dca017693	Dca017693
	Dca018107	Dca018107
	Dca018109	Dca018109
	Dca019472	Dca019472
	Dca021138	Dca021138
	Dca021204	Dca021204
	Dca023162	Dca023162
	Dca023936	Dca023936
	Dca024570	Dca024570
	Dca024748	Dca024748
	Dca025036	Dca025036
Phalaenopsis aphrodite [3]	*PATC043551*	PATC043551
	PATC068781	PATC068781
	PATC127710	PATC127710
	PATC133907	PATC133907
	PATC137979	PATC137979

Table A1. *Cont.*

Species	Gene ID	Accession Number of *TPS* Gene
	PATC139978	PATC139978
	PATC141250	PATC141250
	PATC144727	PATC144727
	PATC150554	PATC150554
	PATC153230	PATC153230
	PATC155674	PATC155674
	PATC161091	PATC161091
	PATC175129	PATC175129
	PATC183449	PATC183449
	PATC187424	PATC187424
	PATC200022	PATC200022
	PATC208458	PATC208458
Phalaenopsis equestris [1]	*Peq006275*	Peq006275
	Peq006282	Peq006282
	Peq006283	Peq006283
	Peq006285	Peq006285
	Peq010211	Peq010211
	Peq011221	Peq011221
	Peq011664	Peq011664
	Peq011667	Peq011667
	Peq013045	Peq013045
	Peq013048	Peq013048
	Peq013713	Peq013713
	Peq020239	Peq020239
	Peq020483	Peq020483
	Peq021360	Peq021360
	Peq023325	Peq023325
Phalaenopsis bellina [4]	*PbTPS01*	CL86.Contig1
	PbTPS02	CL214.Contig2
	PbTPS03	CL376.Contig6
	PbTPS04	CL376.Contig8
	PbTPS05	CL1323.Contig1
	PbTPS06	CL2295.Contig2
	PbTPS07	CL2800.Contig3
	PbTPS08	CL4514.Contig2
	PbTPS09	CL6288.Contig1
	PbTPS10	CL6288.Contig7
	PbTPS11	Unigene4722
Arabidopsis thaliana [2]	*AtTPS1*	At4g15870
	AtTPS2	At4g16730
	AtTPS3	At4g16740
	AtTPS4	At1g61120
	AtTPS5	At4g20230
	AtTPS6	At1g70080
	AtTPS7	At4g20200
	AtTPS8	At4g20210
	AtTPS9	At2g23230
	AtTPS10	At2g24210
	AtTPS11	At5g44630
	AtTPS12	At4g13280
	AtTPS13	At4g13300
	AtTPS14	At1g61680
	AtTPS15	At3g29190
	AtTPS16	At3g29110
	AtTPS17	At3g14490
	AtTPS18	At3g14520
	AtTPS19	At3g14540

Table A1. *Cont.*

Species	Gene ID	Accession Number of *TPS* Gene
	AtTPS20	At5g48110
	AtTPS21	At5g23960
	AtTPS22	At1g33750
	AtTPS23	At3g25830
	AtTPS24	At3g25810
	AtTPS25	At3g29410
	AtTPS26	At1g66020
	AtTPS27	At1g48820
	AtTPS28	At1g48800
	AtTPS29	At1g31950
	AtTPS30	At3g32030
	AtTPS31	At4g02780
	AtTPS32	At1g79460
Abies grandis [2]	*AAB70707*	AGU87910
	AAB70907	AF006193
	AAB71085	U87909
	AAF61454	AF139206
Selaginella moellendorffii [2]	*EFJ31965*	GL377573
	EFJ37889	GL377565
	J9QS23_SmTPS9	XM_002960304
	J9R388_SmTPS10	XM_024672072
	G9MAN7_SmTPS4	XM_024672355.
	G1DGI7_SmTPS7	XM_024689660
	EFJ12417	GL377639
	EFJ37773	GL377565
	EFJ33476	GL377571

[1] OrchidBase 4.0 (http://orchidbase.itps.ncku.edu.tw/est/home2012.aspx (accessed on 9 August 2020)). [2] NCBI database (https://www.ncbi.nlm.nih.gov/ (accessed on 9 August 2020). [3] Orchidstra 2.0 (http://orchidstra2.abrc.sinica.edu.tw/orchidstra2/index.php (accessed on 5 January 2021). [4] *P. bellina* trascriptome database (unpublished).

References

1. Tholl, D. Terpene synthases and the regulation, diversity and biological roles of terpene metabolism. *Curr. Opin. Plant Biol.* **2006**, *9*, 297–304. [CrossRef]
2. Pichersky, E.; Noel, J.P.; Dudareva, N. Biosynthesis of plant volatiles: Nature's diversity and ingenuity. *Science* **2006**, *311*, 808–811. [CrossRef]
3. Yamaguchi, S.; Sun, T.; Kawaide, H.; Kamiya, Y. The GA2 locus of Arabidopsis thaliana encodes ent-kaurene synthase of gibberellin biosynthesis. *Plant Physiol.* **1998**, *116*, 1271–1278. [CrossRef]
4. El Tamer, M.K.; Lücker, J.; Bosch, D.; Verhoeven, H.A.; Verstappen, F.W.; Schwab, W.; van Tunen, A.J.; Voragen, A.G.; De Maagd, R.A.; Bouwmeester, H.J. Domain swapping of Citrus limon monoterpene synthases: Impact on enzymatic activity and product specificity. *Arch. Biochem. Biophys.* **2003**, *411*, 196–203. [CrossRef]
5. Lange, B.M.; Severin, K.; Bechthold, A.; Heide, L. Regulatory role of microsomal 3-hydroxy-3-methylglutaryl-coenzyme A reductase for shikonin biosynthesis in Lithospermum erythrorhizon cell suspension cultures. *Planta* **1998**, *204*, 234–241. [CrossRef] [PubMed]
6. Aubourg, S.; Lecharny, A.; Bohlmann, J. Genomic analysis of the terpenoid synthase (AtTPS) gene family of Arabidopsis thaliana. *Mol. Genet. Genom.* **2002**, *267*, 730–745. [CrossRef] [PubMed]
7. Hayashi, K.; Kawaide, H.; Notomi, M.; Sakigi, Y.; Matsuo, A.; Nozaki, H. Identification and functional analysis of bifunctional ent-kaurene synthase from the moss Physcomitrella patens. *FEBS Lett.* **2006**, *580*, 6175–6181. [CrossRef]
8. Paterson, A.H.; Bowers, J.E.; Bruggmann, R.; Dubchak, I.; Grimwood, J.; Gundlach, H.; Haberer, G.; Hellsten, U.; Mitros, T.; Poliakov, A.; et al. The Sorghum bicolor genome and the diversification of grasses. *Nature* **2009**, *457*, 551–556. [CrossRef]
9. Martin, D.M.; Aubourg, S.; Schouwey, M.B.; Daviet, L.; Schalk, M.; Toub, O.; Lund, S.T.; Bohlmann, J. 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*, 226. [CrossRef] [PubMed]
10. Falara, V.; Akhtar, T.A.; Nguyen, T.T.; Spyropoulou, E.A.; Bleeker, P.M.; Schauvinhold, I.; Matsuba, Y.; Bonini, M.E.; Schilmiller, A.L.; Last, R.L.; et al. The tomato terpene synthase gene family. *Plant Physiol.* **2011**, *157*, 770–789. [CrossRef]

11. Li, G.; Kollner, T.G.; Yin, Y.; Jiang, Y.; Chen, H.; Xu, Y.; Gershenzon, J.; Pichersky, E.; Chen, F. Nonseed plant Selaginella moellendorffi has both seed plant and microbial types of terpene synthases. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 14711–14715. [CrossRef] [PubMed]
12. Liu, J.; Huang, F.; Wang, X.; Zhang, M.; Zheng, R.; Wang, J.; Yu, D. Genome-wide analysis of terpene synthases in soybean: Functional characterization of GmTPS3. *Gene* **2014**, *544*, 83–92. [CrossRef] [PubMed]
13. Irmisch, S.; Jiang, Y.; Chen, F.; Gershenzon, J.; Köllner, T.G. Terpene synthases and their contribution to herbivore-induced volatile emission in western balsam poplar (Populus trichocarpa). *BMC Plant Biol.* **2014**, *14*, 270. [CrossRef]
14. Chen, H.; Li, G.; Kollner, T.G.; Jia, Q.; Gershenzon, J.; Chen, F. Positive Darwinian selection is a driving force for the diversification of terpenoid biosynthesis in the genus Oryza. *BMC Plant Biol.* **2014**, *14*, 239. [CrossRef] [PubMed]
15. Yu, Z.; Zhao, C.; Zhang, G.; Teixeira da Silva, J.A.; Duan, J. Genome-Wide Identification and Expression Profile of TPS Gene Family in Dendrobium officinale and the Role of DoTPS10 in Linalool Biosynthesis. *Int. J. Mol. Sci.* **2020**, *21*, 5419. [CrossRef] [PubMed]
16. Chen, F.; Tholl, D.; Bohlmann, J.; Pichersky, E. 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*, 212–229. [CrossRef]
17. Starks, C.M.; Back, K.; Chappell, J.; Noel, J.P. Structural basis for cyclic terpene biosynthesis by tobacco 5-epi-aristolochene synthase. *Science* **1997**, *277*, 1815–1820. [CrossRef] [PubMed]
18. Rynkiewicz, M.J.; Cane, D.E.; Christianson, D.W. Structure of trichodiene synthase from Fusarium sporotrichioides provides mechanistic inferences on the terpene cyclization cascade. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 13543–13548. [CrossRef]
19. Whittington, D.A.; Wise, M.L.; Urbansky, M.; Coates, R.M.; Croteau, R.B.; Christianson, D.W. Bornyl diphosphate synthase: Structure and strategy for carbocation manipulation by a terpenoid cyclase. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 15375–15380. [CrossRef]
20. Christianson, D.W. Structural biology and chemistry of the terpenoid cyclases. *Chem. Rev.* **2006**, *106*, 3412–3442. [CrossRef]
21. Degenhardt, J.; Kollner, T.G.; Gershenzon, J. Monoterpene and sesquiterpene synthases and the origin of terpene skeletal diversity in plants. *Phytochemistry* **2009**, *70*, 1621–1637. [CrossRef]
22. Wise, M.L.; Croteau, R. Monoterpene biosynthesis. In *Comprehensive Natural Products Chemistry*; Elsevier: Amsterdam, The Netherlands, 1999; Volume 2, pp. 97–159.
23. MacMillan, J.; Beale, M.H. Diterpene biosynthesis. In *Comprehensive Natural Products Chemistry*; Elsevier: Amsterdam, The Netherlands, 1999; Volume 2, pp. 217–243.
24. Taniguchi, S.; Hosokawa-Shinonaga, Y.; Tamaoki, D.; Yamada, S.; Akimitsu, K.; Gomi, K. Jasmonate induction of the monoterpene linalool confers resistance to rice bacterial blight and its biosynthesis is regulated by JAZ protein in rice. *Plant Cell Environ.* **2014**, *37*, 451–461. [CrossRef]
25. Zhang, M.; Liu, J.; Li, K.; Yu, D. Identification and characterization of a novel monoterpene synthase from soybean restricted to neryl diphosphate precursor. *PLoS ONE* **2013**, *8*, e75972. [CrossRef] [PubMed]
26. Aharoni, A.; Giri, A.P.; Verstappen, F.W.; Bertea, C.M.; Sevenier, R.; Sun, Z.; Jongsma, M.A.; Schwab, W.; Bouwmeester, H.J. Gain and loss of fruit flavor compounds produced by wild and cultivated strawberry species. *Plant Cell* **2004**, *16*, 3110–3131. [CrossRef]
27. Bohlmann, J.; Crock, J.; Jetter, R.; Croteau, R. Terpenoid-based defenses in conifers: CDNA cloning, characterization, and functional expression of wound-inducible (E)-alpha-bisabolene synthase from grand fir (Abies grandis). *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 6756–6761. [CrossRef] [PubMed]
28. Keeling, C.I.; Weisshaar, S.; Lin, R.P.; Bohlmann, J. Functional plasticity of paralogous diterpene synthases involved in conifer defense. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 1085–1090. [CrossRef] [PubMed]
29. Yamaguchi, S. Gibberellin metabolism and its regulation. *Annu. Rev. Plant Biol.* **2008**, *59*, 225–251. [CrossRef]
30. Pichersky, E.; Gershenzon, J. The formation and function of plant volatiles: Perfumes for pollinator attraction and defense. *Curr. Opin. Plant Biol.* **2002**, *5*, 237–243. [CrossRef]
31. Unsicker, S.B.; Kunert, G.; Gershenzon, J. Protective perfumes: The role of vegetative volatiles in plant defense against herbivores. *Curr. Opin. Plant Biol.* **2009**, *12*, 479–485. [CrossRef]
32. Weitzel, C.; Simonsen, H.T. Cytochrome P450-enzymes involved in the biosynthesis of mono- and sesquiterpenes. *Phytochem. Rev.* **2015**, *14*, 7–24. [CrossRef]
33. Fay, M.F. Orchid conservation: How can we meet the challenges in the twenty-first century? *Bot. Stud.* **2018**, *59*, 16. [CrossRef] [PubMed]
34. Givnish, T.J.; Spalink, D.; Ames, M.; Lyon, S.P.; Hunter, S.J.; Zuluaga, A.; Doucette, A.; Caro, G.G.; McDaniel, J.; Clements, M.A.; et al. Orchid historical biogeography, diversification, Antarctica and the paradox of orchid dispersal. *J. Biogeogr.* **2016**, *43*, 1905–1916. [CrossRef]
35. Chase, M.W.; Cameron, K.M.; Barrett, R.L.; Freudenstein, J.V. DNA data and Orchidaceae systematics: A new phylogenetic classification. In *Orchid Conservation*; Natural History Publications: Kota Kinabalu, Malaysia, 2003; pp. 69–89.
36. Raguso, R.A.; Levin, R.A.; Foose, S.E.; Holmberg, M.W.; McDade, L.A. Fragrance chemistry, nocturnal rhythms and pollination "syndromes" in Nicotiana. *Phytochemistry* **2003**, *63*, 265–284. [CrossRef]
37. Gregg, K.B. Variation in floral fragrances and morphology: Incipient speciation in Cycnoches? *Bot. Gaz.* **1983**, *144*, 566–576. [CrossRef]

38. Tan, K.H.; Tan, L.T.; Nishida, R. Floral phenylpropanoid cocktail and architecture of Bulbophyllum vinaceum orchid in attracting fruit flies for pollination. *J. Chem. Ecol.* **2006**, *32*, 2429–2441. [CrossRef] [PubMed]
39. Cancino, A.D.M.; Damon, A. Fragrance analysis of euglossine bee pollinated orchids from Soconusco, south-east Mexico. *Plant Species Biol.* **2007**, *22*, 127–132.
40. Salzmann, C.C.; Schiestl, F.P. Odour and colour polymorphism in the food-deceptive orchid Dactylorhiza romana. *Plant Syst. Evol.* **2007**, *267*, 37–45. [CrossRef]
41. Brodmann, J.; Twele, R.; Francke, W.; Luo, Y.B.; Song, X.Q.; Ayasse, M. Orchid Mimics Honey Bee Alarm Pheromone in Order to Attract Hornets for Pollination. *Curr. Biol.* **2009**, *19*, 1368–1372. [CrossRef]
42. Delle-Vedove, R.; Juillet, N.; Bessiere, J.M.; Grison, C.; Barthes, N.; Pailler, T.; Dormont, L.; Schatz, B. Colour-scent associations in a tropical orchid: Three colours but two odours. *Phytochemistry* **2011**, *72*, 735–742. [CrossRef]
43. Wiemer, A.P.; More, M.; Benitez-Vieyra, S.; Cocucci, A.A.; Raguso, R.A.; Sersic, A.N. A simple floral fragrance and unusual osmophore structure in Cyclopogon elatus (Orchidaceae). *Plant Biol.* **2009**, *11*, 506–514. [CrossRef]
44. Nunes, C.E.; Gerlach, G.; Bandeira, K.D.; Gobbo-Neto, L.; Pansarin, E.R.; Sazima, M. Two orchids, one scent? Floral volatiles of Catasetum cernuum and Gongora bufonia suggest convergent evolution to a unique pollination niche. *Flora* **2017**, *232*, 207–216. [CrossRef]
45. Ramya, M.; Park, P.H.; Chuang, Y.-C.; Kwon, O.K.; An, H.R.; Park, P.M.; Baek, Y.S.; Kang, B.-C.; Tsai, W.-C.; Chen, H.-H. RNA sequencing analysis of Cymbidium goeringii identifies floral scent biosynthesis related genes. *BMC Plant Biol.* **2019**, *19*, 337. [CrossRef]
46. Kaiser, R.A. *On the Scent of Orchids*; ACS Publications: Washington, DC, USA, 1993; pp. 240–268.
47. Hsiao, Y.Y.; Pan, Z.J.; Hsu, C.C.; Yang, Y.P.; Hsu, Y.C.; Chuang, Y.C.; Shih, H.H.; Chen, W.H.; Tsai, W.C.; Chen, H.H. Research on orchid biology and biotechnology. *Plant Cell Physiol.* **2011**, *52*, 1467–1486. [CrossRef]
48. Hsiao, Y.Y.; Tsai, W.C.; Kuoh, C.S.; Huang, T.H.; Wang, H.C.; Wu, T.S.; Leu, Y.L.; Chen, W.H.; Chen, H.H. Comparison of transcripts in Phalaenopsis bellina and Phalaenopsis equestris (Orchidaceae) flowers to deduce monoterpene biosynthesis pathway. *BMC Plant Biol.* **2006**, *6*, 14. [CrossRef]
49. Awano, K.; Ichikawa, Y.; Tokuda, K.; Kuraoka, M. Volatile Components of the Flowers of Two Calanthe Species. *Flavour Fragr. J.* **1997**, *12*, 327–333. [CrossRef]
50. Lichtenthaler, H.K. The 1-deoxy-D-xylulose-5-phosphate pathway of isoprenoid biosynthesis in plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **1999**, *50*, 47–65. [CrossRef] [PubMed]
51. Newman, J.D.; Chappell, J. Isoprenoid biosynthesis in plants: Carbon partitioning within the cytoplasmic pathway. *Crit. Rev. Biochem. Mol. Biol.* **1999**, *34*, 95–106. [CrossRef]
52. Sapir-Mir, M.; Mett, A.; Belausov, E.; Tal-Meshulam, S.; Frydman, A.; Gidoni, D.; Eyal, Y. Peroxisomal Localization of Arabidopsis Isopentenyl Diphosphate Isomerases Suggests That Part of the Plant Isoprenoid Mevalonic Acid Pathway Is Compartmentalized to Peroxisomes. *Plant Physiol.* **2008**, *148*, 1219–1228. [CrossRef] [PubMed]
53. Ashour, M.; Wink, M.; Gershenzon, J. Biochemistry of terpenoids: Monoterpenes, sesquiterpenes and diterpenes. In *Annual Plant Reviews: Biochemistry of Plant Secondary Metabolism*; Wiley-Blackwell: Chichester, UK, 2010; pp. 258–303.
54. Zhang, G.Q.; Liu, K.W.; Li, Z.; Lohaus, R.; Hsiao, Y.Y.; Niu, S.C.; Wang, J.Y.; Lin, Y.C.; Xu, Q.; Chen, L.J.; et al. The Apostasia genome and the evolution of orchids. *Nature* **2017**, *549*, 379–383. [CrossRef] [PubMed]
55. Hasing, T.; Tang, H.; Brym, M.; Khazi, F.; Huang, T.; Chambers, A.H. A phased Vanilla planifolia genome enables genetic improvement of flavour and production. *Nat. Food* **2020**, *1*, 811–819. [CrossRef]
56. Zhang, G.-Q.; Xu, Q.; Bian, C.; Tsai, W.-C.; Yeh, C.-M.; Liu, K.-W.; Yoshida, K.; Zhang, L.-S.; Chang, S.-B.; Chen, F.; et al. The Dendrobium catenatum Lindl. genome sequence provides insights into polysaccharide synthase, floral development and adaptive evolution. *Sci. Rep.* **2016**, *6*, 19029. [CrossRef]
57. Cai, J.; Liu, X.; Vanneste, K.; Proost, S.; Tsai, W.C.; Liu, K.W.; Chen, L.J.; He, Y.; Xu, Q.; Bian, C.; et al. The genome sequence of the orchid Phalaenopsis equestris. *Nat. Genet.* **2015**, *47*, 65–72. [CrossRef]
58. Tsai, W.C.; Fu, C.H.; Hsiao, Y.Y.; Huang, Y.M.; Chen, L.J.; Wang, M.; Liu, Z.J.; Chen, H.H. OrchidBase 2.0: Comprehensive collection of Orchidaceae floral transcriptomes. *Plant Cell Physiol.* **2013**, *54*, e7. [CrossRef]
59. Sun, T.P.; Kamiya, Y. The Arabidopsis GA1 locus encodes the cyclase ent-kaurene synthetase A of gibberellin biosynthesis. *Plant Cell* **1994**, *6*, 1509–1518. [PubMed]
60. Tsai, W.C.; Dievart, A.; Hsu, C.C.; Hsiao, Y.Y.; Chiou, S.Y.; Huang, H.; Chen, H.H. Post genomics era for orchid research. *Bot. Stud.* **2017**, *58*, 61. [CrossRef] [PubMed]
61. Finn, R.D.; Coggill, P.; Eberhardt, R.Y.; Eddy, S.R.; Mistry, J.; Mitchell, A.L.; Potter, S.C.; Punta, M.; Qureshi, M.; Sangrador-Vegas, A.; et al. The Pfam protein families database: Towards a more sustainable future. *Nucleic Acids Res.* **2016**, *44*, D279–D285. [CrossRef] [PubMed]
62. Trapp, S.C.; Croteau, R.B. Genomic organization of plant terpene synthases and molecular evolutionary implications. *Genetics* **2001**, *158*, 811–832. [CrossRef] [PubMed]
63. Shalev, T.J.; Yuen, M.M.S.; Gesell, A.; Yuen, A.; Russell, J.H.; Bohlmann, J. An annotated transcriptome of highly inbred Thuja plicata (Cupressaceae) and its utility for gene discovery of terpenoid biosynthesis and conifer defense. *Tree Genet. Genomes* **2018**, *14*, 35. [CrossRef]

64. Williams, D.C.; McGarvey, D.J.; Katahira, E.J.; Croteau, R. Truncation of limonene synthase preprotein provides a fully active 'pseudomature' form of this monoterpene cyclase and reveals the function of the amino-terminal arginine pair. *Biochemistry* **1998**, *37*, 12213–12220. [CrossRef]
65. Hyatt, D.C.; Youn, B.; Zhao, Y.; Santhamma, B.; Coates, R.M.; Croteau, R.B.; Kang, C. Structure of limonene synthase, a simple model for terpenoid cyclase catalysis. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 5360–5365. [CrossRef]
66. Cao, R.; Zhang, Y.; Mann, F.M.; Huang, C.; Mukkamala, D.; Hudock, M.P.; Mead, M.E.; Prisic, S.; Wang, K.; Lin, F.Y.; et al. Diterpene cyclases and the nature of the isoprene fold. *Proteins Struct. Funct. Bioinform.* **2010**, *78*, 2417–2432. [CrossRef]
67. Keeling, C.I.; Dullat, H.K.; Yuen, M.; Ralph, S.G.; Jancsik, S.; Bohlmann, J. Identification and Functional Characterization of Monofunctional ent-Copalyl Diphosphate and ent-Kaurene Synthases in White Spruce Reveal Different Patterns for Diterpene Synthase Evolution for Primary and Secondary Metabolism in Gymnosperms. *Plant Physiol.* **2010**, *152*, 1197–1208. [CrossRef] [PubMed]
68. Alicandri, E.; Paolacci, A.R.; Osadolor, S.; Sorgona, A.; Badiani, M.; Ciaffi, M. On the Evolution and Functional Diversity of Terpene Synthases in the Pinus Species: A Review. *J. Mol. Evol.* **2020**, *88*, 253–283. [CrossRef]
69. Chen, X.; Kollner, T.G.; Shaulsky, G.; Jia, Q.; Dickschat, J.S.; Gershenzon, J.; Chen, F. Diversity and Functional Evolution of Terpene Synthases in Dictyostelid Social Amoebae. *Sci. Rep.* **2018**, *8*, 14361. [CrossRef]
70. Ker, D.S.; Pang, S.L.; Othman, N.F.; Kumaran, S.; Tan, E.F.; Krishnan, T.; Chan, K.G.; Othman, R.; Hassan, M.; Ng, C.L. Purification and biochemical characterization of recombinant Persicaria minor beta-sesquiphellandrene synthase. *Peer J.* **2017**, *5*, e2961. [CrossRef] [PubMed]
71. Tang, H.; Zhao, T.; Sheng, Y.; Zheng, T.; Fu, L.; Zhang, Y. Dendrobium officinale Kimura et Migo: A Review on Its Ethnopharmacology, Phytochemistry, Pharmacology, and Industrialization. *Evid. Based Complementary Altern. Med.* **2017**, *2017*, 19. [CrossRef] [PubMed]
72. Chuang, Y.C.; Hung, Y.C.; Tsai, W.C.; Chen, W.H.; Chen, H.H. PbbHLH4 regulates floral monoterpene biosynthesis in Phalaenopsis orchids. *J. Exp. Bot.* **2018**, *69*, 4363–4377. [CrossRef] [PubMed]
73. Nieuwenhuizen, N.J.; Wang, M.Y.; Matich, A.J.; Green, S.A.; Chen, X.; Yauk, Y.K.; Beuning, L.L.; Nagegowda, D.A.; Dudareva, N.; Atkinson, R.G. Two terpene synthases are responsible for the major sesquiterpenes emitted from the flowers of kiwifruit (Actinidia deliciosa). *J. Exp. Bot.* **2009**, *60*, 3203–3219. [CrossRef] [PubMed]
74. Chen, X.; Yauk, Y.-K.; Nieuwenhuizen, N.J.; Matich, A.J.; Wang, M.Y.; Perez, R.L.; Atkinson, R.G.; Beuning, L.L. Characterisation of an (S)-linalool synthase from kiwifruit (Actinidia arguta) that catalyses the first committed step in the production of floral lilac compounds. *Funct. Plant Biol.* **2010**, *37*, 232–243. [CrossRef]
75. Krokida, A.; Delis, C.; Geisler, K.; Garagounis, C.; Tsikou, D.; Peña-Rodríguez, L.M.; Katsarou, D.; Field, B.; Osbourn, A.E.; Papadopoulou, K.K. A metabolic gene cluster in Lotus japonicus discloses novel enzyme functions and products in triterpene biosynthesis. *New Phytol.* **2013**, *200*, 675–690. [CrossRef]
76. Jiang, S.Y.; Jin, J.; Sarojam, R.; Ramachandran, S. A Comprehensive Survey on the Terpene Synthase Gene Family Provides New Insight into Its Evolutionary Patterns. *Genome Biol. Evol.* **2019**, *11*, 2078–2098. [CrossRef] [PubMed]
77. Matsuba, Y.; Nguyen, T.T.; Wiegert, K.; Falara, V.; Gonzales-Vigil, E.; Leong, B.; Schafer, P.; Kudrna, D.; Wing, R.A.; Bolger, A.M.; et al. Evolution of a complex locus for terpene biosynthesis in solanum. *Plant Cell* **2013**, *25*, 2022–2036. [CrossRef] [PubMed]
78. Külheim, C.; Padovan, A.; Hefer, C.; Krause, S.T.; Köllner, T.G.; Myburg, A.A.; Degenhardt, J.; Foley, W.J. The Eucalyptus terpene synthase gene family. *BMC Genom.* **2015**, *16*, 450. [CrossRef] [PubMed]
79. Shimura, K.; Okada, A.; Okada, K.; Jikumaru, Y.; Ko, K.-W.; Toyomasu, T.; Sassa, T.; Hasegawa, M.; Kodama, O.; Shibuya, N.; et al. Identification of a Biosynthetic Gene Cluster in Rice for Momilactones. *J. Biol. Chem.* **2007**, *282*, 34013–34018. [CrossRef]
80. Chen, H.; Kollner, T.G.; Li, G.; Wei, G.; Chen, X.; Zeng, D.; Qian, Q.; Chen, F. Combinatorial Evolution of a Terpene Synthase Gene Cluster Explains Terpene Variations in Oryza. *Plant Physiol.* **2020**, *182*, 480–492. [CrossRef]
81. Boutanaev, A.M.; Moses, T.; Zi, J.; Nelson, D.R.; Mugford, S.T.; Peters, R.J.; Osbourn, A. Investigation of terpene diversification across multiple sequenced plant genomes. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, E81–E88. [CrossRef]