Genome-wide analysis of terpene synthase gene family in Rosa chinensis

Ying Kong
Beijing Academy of Science and Technology
Huan Wang
Beijing Academy of Science and Technology
Lixin Lang
Beijing Academy of Science and Technology
Xiaoying Dou
Beijing Academy of Science and Technology
Jinrong Bai (bjr301@126.com)
Beijing Academy of Science and Technology

Research Article

Keywords: butterfly rose, monoterpene, sesquiterpene, NUDIX, senescence

Posted Date: April 15th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1505144/v1

Abstract

Background

Terpenoids are economically and ecologically important compounds and vital constituents of rose flower fragrance and rose essential oil. Terpene synthase (TPS) is a primary enzyme for terpenoid biosynthesis. However, only a few TPS genes have been characterized in rose, owing to the lack of a systematic TPS gene analysis. Additionally, the key gene *RcNUDX1* was analyzed because of the presence of unique geraniol pathways in rose.

Results

The present study identified 74 putative *RcTPS* genes through a comprehensive bioinformatics analysis; of these genes, 49 genes were predicted to encode functional TPS enzymes. The TPS-a subfamily had the largest members, and tandem duplication was the main mechanism for driving TPS gene family expansion. The collinearity analysis with the genome of other Rosaceae plants and grape indicated that four *RcTPS* genes may be derived from the dicotyledonous plant ancestors. During the flower opening process of butterfly rose (*Rosa chinensis* 'Mutabilis', MU), *RcTPS* genes exhibited two main expression trends, namely high and low, in old and fresh petals. Similarly, dark treatment decreased the expression of four *RcTPS*s and increased the expression of three *RcTPS*s in the MU petals. Most of the monoterpenes were released more from old MU petals than from fresh MU petals, which may be related to *RcLIN-NERS3* and *RcNUDX1-1a4* genes. Analysis of volatile emission, gene expression, sequence structure, and collinearity indicated that *RcTPS18* may encode (*E*,*E*)-α-farnesene synthase. The three genes whose expression levels increased at the late stage and after dark treatment exhibited similar expression patterns with some senescence marker genes, which indicated that these genes might be involved in petal senescence or cell death.

Conclusions

The present study revealed the differential *RcTPS* gene expression patterns in rose petals. These findings may be useful for elucidating the molecular mechanism of terpenoid metabolism in rose and are vital for future studies on terpene regulation.

Background

Terpenoids constitute the largest and structurally most diverse class of plant metabolites that defend against predators and attract beneficial organisms [1–3]. All plant organs such as leaves, branches, flowers, and roots can emit terpene volatiles to ensure healthy plant growth [4]. In most plants, the volatile terpenoids are constructed from two C_5 precursors, namely isopentenyl diphosphate (IPP) and its isomer dimethylallyl diphosphate (DMAPP), which are produced through either the methylerithritol phosphate

pathway (MEP) in the chloroplast or the mevalonate pathway (MVA) in the cytosol [2]. Then, the IPP and DMAPP units are condensed by prenyltransferases (PTs) to form direct terpene precursors, such as geranyl diphosphate (C_{10}, GPP), farnesyl diphosphate (C_{15}, FPP), or geranylgeranyl diphosphate (C_{20}, GGPP) [5, 6]. Subsequently, terpene synthase (TPS, EC 4.2.3), which is the primary enzyme in the terpenoid biosynthetic pathways, converts the precursors into various terpene products such as monoterpene (C_{10}), sesquiterpene (C_{15}), and diterpene (C_{20}) [7]. These products can undergo further modifications under the action of various enzymes such as dehydrogenases, methyltransferases, acyltransferases, and glycosyltransferases to form highly diverse metabolites [8, 9].

The TPS gene family is a mid-sized family, and different sized TPS families, with gene members ranging from approximately 20 to 150, have evolved in different plant species [10, 11]. Both tandem and segmental duplication are considered as the main mechanisms driving TPS expansion [3]. The TPS gene family members have been characterized in many eudicot plants such as Arabidopsis (*Arabidopsis thaliana*) [12], grapevine (*Vitis vinifera*) [13], tomato (*Solanum lycopersicum*) [8, 14], and apple (*Malus domestica*) [11]. The canonical TPSs, typified by highly conserved Asp-rich motifs, are traditionally divided into two categories: type I TPSs (having DDxxD and NSE/DTE motifs) and type II TPSs (commonly possessing a DxDD motif) [15]. Seven subfamilies, namely TPS-a (sesquiterpenes), TPS-b (cyclic monoterpenes and hemiterpenes), TPS-c (copalyl diphospate and *ent*-kaurene), TPS-d (gymnosperm specific), TPS-e/f (*ent*-kaurene and other diterpenes, as well as some mono- and sesquiterpenes), TPS-g (acyclic monoterpenes), and TPS-h (*Selaginella* specific), have been clustered based on the phylogenetic analysis using full-length amino acid sequences [3, 16].

Rose is widely cultivated as a garden plant for the cut-flower industry, and it is a crucial floriculture crop that provides essential oil to the fragrance, food, and cosmetic industries [9, 17]. Floral fragrance is a vital characteristic of ornamental roses [18]. Hundreds of compounds have been found to contribute to rose scent comprised terpenoids, phenylpropanoids, and fatty acid derivatives etc. [19]. The terpenoid volatiles in the floral rose scent are mainly monoterpenes and sesquiterpenes, as well as their derivatives. Some fragrant rose cultivars produce a large amount of geraniol, citronellol, and nerol, all of which are the major components of rose scent and rose oil; the aldehyde and acetate ester derivatives of these compounds have also been identified [20, 21]. Other monoterpenoids including α-pinene, β-pinene, limonene, and linalool are emitted at low levels in rose floral scent [21]. Sesquiterpenes including germacrene D, δ-cadinene, α-copaene, α-cubenene, β-cubenene, β-elemene, and β-caryophyllene are emitted from rose flowers, whereas germacrene D was highly released from some rose cultivars [19, 21–23].

Several TPS genes involved in the biosynthesis of rose volatile terpenes have been identified. The first TPS gene cloned in rose was germacrene D synthase (*RhGDS*), which catalyzes the substrate FPP to germacrene D as a unique product [19]. Three TPS genes, namely *RcLINS*, *RcLIN-NERS1*, and *RcLIN-NERS2*, have been characterized, and the expression levels of these genes are low in rose petals [24]. *RcLINS* belonging to the TPS-b subgroup is responsible for the presence of a small amount of (*3R*)-(-)-linalool in rose scent. The bifunctional *RcLIN-NERS1* and *RcLIN-NERS2*, belonging to the TPS-g subgroup, produce (*3S*)-(+)-linalool and nerolidol when incubated with GPP and FPP, respectively, whereas *RcLIN-*

NERS3 has been identified as a pseudogene. Moreover, a novel TPS-independent pathway for monoterpene biosynthesis was described in rose. An enzyme of the Nudix hydrolase family (RhNUDX1) localized in the cytoplasm was reported to be involved in geraniol biosynthesis [25]. Another study showed that *RwNUDX1-2* was involved in the biosynthesis of a group of sesquiterpenoids [26].

Rosa chinensis 'Mutabilis' (butterfly rose, MU), with single petals, is an ancient Chinese rose cultivar. Studies have indicated that monoterpene alcohols released from MU flowers, including geraniol, nerol, and linalool, and the monoterpene contents in MU exhibit a significant increase from unopened buds to floral maturity and further accumulation during senescence [22, 27]. However, the mechanism of volatile sesquiterpene release from MU flowers is unclear. The present study conducted a genome-wide identification of the *R. chinensis* TPS genes. Phylogenetic analysis, conserved motif analysis, chromosomal mapping, and collinearity analysis of *RcTPS* genes were performed. Furthermore, RNA sequencing (RNA-seq) was performed to investigate the expression patterns of *RcTPS*s and related genes during flower development. The effect of dark treatment was also determined. Furthermore, the expression of differential volatile terpenes and *RcTPS*s was analyzed. These results may provide insights for further characterization of the putatively functional TPS genes in *Rosa* species.

Results

RcTPS gene identification and sequence features

The present study originally obtained 81 nonredundant candidate gene models corresponding to PF01397 or PF03936. After aligning the candidate sequences by using MAFFT software, the sequences that did not contain the typical TPS domains were manually removed. A total of 74 putative TPS genes were identified, 49 of which were predicted to encode functional TPS enzymes. These genes encoding proteins longer than 500 amino acids (aa) contained typical TPS domains comprising either the Mg^{2+} binding (DDxxD/E) and NSE/DTE regions or the DxDD motif (Figure S1). The remaining 25 TPS genes, most of which encoded polypeptides shorter than 400 amino acids, were categorized as partial/pseudo genes as they contained partial or mutant TPS domains. The genes were then renamed based on their order on the linkage groups (Table S1) [28]. Among the 49 complete RcTPS proteins, the molecular weight (MW) of the proteins ranged from 57.2 to 98.5 kDa, whereas the isoelectric point (pI) ranged from 5.06 to 6.17. The candidate TPS genes identified were *RcGDS* (*RcTPS30*, homologous gene of *RhGDS*), *RcLINS* (*RcTPS-p9* + *RcTPS-p10*), *RcLIN-NERS1* (*RcTPS20*), *RcLIN-NERS2* (*RcTPS22*), and *RcLIN-NERS3* (*RcTPS23*).

To explore the evolutionary relationship of the *Rosa* TPS family, a maximum likelihood phylogenetic tree was constructed using all putative and characterized RcTPSs (Fig. 1a). *ent*-kaurene synthase from *Physcomitrella patens* (PpCPS/KS), an ancient bifunctional TPS, was considered as the root of the phylogenetic tree [29]. The phylogenetic analysis separated RcTPSs into five groups, namely TPS-c, -e/f, -g, -b, and -a subfamilies. Among these 49 putative functional RcTPSs, 32 belonged to the TPS-a subfamily, which was found to be the largest subfamily. The number of complete RcTPSs in the

subfamilies TPS-b, -g, -e/g, and -c were fewer, with 8, 4, 3, and 2 genes respectively. A comparison of the subfamily gene numbers with those in other eudicot plants exhibited that the gene number of the TPS-a subfamily in the *R. chinensis* genome was much higher than that in the other plants, which suggested a substantial species-specific expansion for the TPS-a subfamily in *R. chinensis* (Table 1).

Subsequently, 10 conserved motifs were identified in the rose TPS genes through the MEME program (Fig. 1b). Compared with *RcTPS* genes within the same subfamily, the *RcTPS-p* genes comprised incomplete motifs. The DDxxD and NSE/DTE motifs were represented by Motifs 1 and 7, respectively (Figure S2), and were found to be present in all putative functional RcTPSs, except those belonging to the TPS-c subfamily. Although the figure illustrates that the TPS-c genes also contained Motif 1, they did not contain the DDxxD motif. The characteristic motif DxDD was located in Motif 5; however, it was not intuitive (Figure S2). The submitted sequence failed to reflect its characteristic domains due to the presence of only a few TPS-c genes.

The gene structure analysis confirmed that every complete *RcTPS* gene had a TPS open reading frame of expected size and organization (Fig. 1c). The members of TPS-c and -e/f subfamilies belonged to class I TPSs, which have a large number of exon regions ranging from 11 to 15, encoding larger proteins with the number of aa residues ranging from 724 to 857. The TPS-a/b/g subfamily members belonged to class III TPSs, which normally comprise 6–7 exon regions, encoding smaller proteins with the number of aa residues ranging from 501 to 601.

Table 1 Statistics of the TPS subfamily gene numbers in the genome of *Rosa chinensis* and other plants

Species	(putative) Functional TPS						Partial/pseudo TPS
	TPS-a	TPS-b	TPS-c	TPS-e/f	TPS-g	All	
Rosa chinensis	32	8	2	3	4	**49**	25
Malus 'Golden Delicious' [11]	4	1	1	1	3	**10**	45
Populus trichocarpa [30]	16	14	2	3	3	**38**	19
Glycine max [31]	6	5	3	2	7	**23**	7
Vitis vinifera [13]	30	19	2	1	17	**69**	83
Arabidopsis thaliana [10]	22	6	1	2	1	**32**	8
Solanum lycopersicum [8]	15	9	2	6	2	**34**	18

Chromosomal localization and gene duplication

The 74 rose TPS and related genes (10 complete *PT* and 10 *RcNUDX1* genes) in the *R. chinensis* genome were mapped to the seven chromosomes of the *R. chinensis* genome sequence assembly based on their localization information (Fig. 2). The chromosomal localization analysis exhibited the uneven distribution of the 49 *RcTPS* and 25 *RcTPS-p* genes across all seven chromosomes. RcChr 5 contained the largest number of TPS genes, including 18 *RcTPS* and 8 *RcTPS-p* genes, which suggest the multiple duplication and recombination events on this chromosome [30]. Most pseudogenes were distributed near the putative full-length TPS genes. Six tandemly duplicated genes were present in the *R. chinensis* genome, which occurred in the TPS-a, -b, and -g subfamilies, forming five gene clusters. No segmentally duplicated TPS genes were detected in the *R. chinensis* genome, indicating that tandem duplication is the predominant form of expansion for the TPS gene family in *R. chinensis*. Additionally, four of the seven full-length TPT genes, namely *RcFPPS1*, *RcGPPS*, *RcSPPS*, and *RcFPPS2*, were mapped on RcChr 5. Some *RcTPS* genes were localized in the vicinity of putative rose PT genes, which indicated that some *RcTPS* and PT genes probably evolved through genomic duplication [31].

RcTPS gene collinearity analysis

The comparative collinearity maps of *R. chinensis* associated with other representative species were constructed to further infer the phylogenetic mechanisms of the TPS gene family (Fig. 3). Four *RcTPS* genes, namely *RcTPS18*, *RcLIN-NERS1*, *RcTPS27*, and *RcTPS-p8*, exhibited genomic shuffling across the Rosaceae species and grape, indicating that these genes were derived probably from the ancestors of dicotyledonous plants. Unlike the other three genes with only one collinear gene pair in each plant genome, *RcLIN-NERS1* had three collinear gene pairs in the *R. rugosa* genome and three collinear gene pairs in the strawberry genome. *RcLIN-NERS1* and its two collinear genes, namely *FaNES1* and *FaNES2*, in strawberry are bifunctional terpene synthase that can efficiently convert GPP and FPP into linalool and nerolidol, respectively [24, 32]. This indicates that collinear TPS genes may exhibit similar catalytic function in relative plants. *RcTPS27* (TPS-a) and *RcLIN-NERS1* (TPS-b) in *R. chinensis* shared the same collinear TPS genes in *R. rugosa* and strawberry, indicating that TPS-a (*RcTPS27*) is probably derived from TPS-g. *RcTPS-p8*, a putative pseudogene that encodes a short protein (455 aa), had a conserved collinear region in all representative plants, which might be due to the breakage or loss of genes during evolution.

Effect of developmental stages and dark treatment on *RcTPS* expression

Nine *RcTPS* genes belonging to the TPS-a, -b, and -g subfamilies were expressed in MU opening flower samples, whereas the other two *RcTPS* genes were only expressed in the buds. These genes exhibited great differences in the expression levels and patterns. *RcGDS*, which was responsible for catalyzing the synthesis of germacrene D as the only product, exhibited the highest average expression level [19]. *RcTPS32* and *RcTPS33* exhibited the second and third highest expression levels, respectively. Although the *RcTPS32* and *RcTPS33* coding sequences were identical, their expression levels varied. In all samples

of opened MU flowers, the *RcTPS32* expression level was approximately thrice that of *RcTPS33*. The expression of other *RcTPS*s in petals and buds was relatively low (Fig. 4a).

The oscillations in *RcTPS* expression in different developmental stages of MU flowers were analyzed. These genes exhibited several expression patterns. The first group comprised *RcLINS*, *RcTPS18*, *RcLIN-NERS1*, and *RcGDS*, whose expression peaked in the buds about to open (S3) or in the early opening flowers (D1) and declined in old flowers. The second group comprised *RcTPS39*, whose peak expression was on the second day after flowering (D2). The third group comprised four *RcTPS* genes, whose expression levels increased when the flowers opened and peaked in old flowers. The last group comprising *RcTPS8* and *RcTPS9* was expressed only in the buds about to open (S3) and hardly expressed in open flowers (Fig. 4b). The effect of dark treatment on *RcTPS* expression in flowering samples was also studied. Compared with control samples (LD), dark treatment resulted in the decreased expression of four genes, namely *RcLINS*, *RcLIN-NERS1*, *RcLIN-NERS3*, and *RcGDS*, and increased expression of three genes, namely *RcTPS32*, *RcTPS33*, and *RcTPS46.* Additionally, it exhibited no significant effect on *RcTPS18* and *RcTPS39* expression (Fig. 4c).

Nine *RcTPS*s expressed in the opening MU petals were selected to further validate the RNA-seq results in three samples using real-time quantitative polymerase chain reaction (qRT-PCR). The results confirmed the differential TPS gene expression patterns in different samples (Figure S3). The concordance between qRT-PCR and RNA-seq results demonstrated the reliability of RNA-seq data in the present study.

Expression profiles of terpene-related genes

The expression profiles of related genes involved in the synthesis of terpenoid volatiles were analyzed, and the *RhUBI2* expression level was used as an internal control. Most genes involved in the MEP pathway exhibited decreased expression in old petals, whereas most genes involved in the MVA pathway exhibited peak expression levels in old flowers and buds or fresh flowers (Fig. 4a). *RcIDI* and *RcIDK* exhibited different expression patterns, reaching peak expression in buds and senescent flowers, respectively. Among the five TPT genes involved in the biosynthesis of floral volatile terpenes, *RcGGPPS1* and *RcSSUII* exhibited highest expression in buds or fresh flowers. *RcGPPS* expression peaked in senescent flowers, whereas the expression of two *RcFPPS* genes peaked in buds and old flowers. Dark treatment resulted in increased *RcGGPPS1* expression and decreased *RcFPPS1* expression, whereas no significant effect was observed on the expression of the other three TPT genes (Fig. 4c).

The six homologous genes of *RcNUDX1* involved in geraniol synthesis were highly expressed in MU petals (Fig. 4a). All six *RcNUDIX1* genes exhibited significantly increased expression after flowering. Four of the *RcNUDIX1* genes exhibited constant expression throughout the flowering stage, whereas the other two genes exhibited significantly higher expression in old flowers than in fresh flowers (Fig. 4c). Furthermore, the expression levels of all six *RcNUDX1* genes significantly increased after dark treatment (Fig. 4b).

Phylogenetic analysis of *RcTPS*s expressed in butterfly rose petals

Phylogenetic analysis was performed using the maximum likelihood method and included 11 *RcTPS* genes expressed in the butterfly rose petals or buds and other representative TPS genes (Fig. 5). Of these, four *RcTPS-b* genes were grouped into three clusters. *RcTPS8* and *RcTPS9* were clustered with ocimene synthases, whereas *RcLINS* was clustered with other linalool synthases. *RcTPS18* was clustered together with other angiosperm α-farnesene synthases (AFSs) from apple (*MdAFS*) [33], *Prunus campanulata* (*PcTPS7*) [34], peach (*PpTPS2*) [35], *Glycine max* (*GmAFS*) [36], *Populus trichocarpa* (*PtTPS2*) [37], tomato (*SlTPS27*) [8], grape (*VvGwbOciF*) [13], and *Ricinus communis* (*RcSeTPS7*) [38], forming an AFS cluster. All genes in this cluster were predicted or identified as being localized in the cytoplasm. Sequence analysis showed that *RcTPS18* exhibited 59% to 66% identity with other AFSs of Rosaceae plants. RcTPS18 and other AFSs exhibited conserved structural features including the DDxxD motif, NSD/DTE motif, and H-α1 loop. The last motif demonstrated function in the binding of the metal ion K^+ in MdAFS (Figure S4) [39]. Moreover, *RcTPS18* and *PpTPS2* were collinear gene pairs located in a large collinear region with other TPS, TPT, and P450 genes (Figure S5). Thus, *RcTPS18* might encode (*E*,*E*)-α-farnesene synthase, which is consistent with its annotation in the genome.

The five *RcTPS-a* genes expressed in rose petals were grouped into two clusters. *RcTPS39* was clustered with some genes that can catalyze monoterpenoid products such as *FvPINS* (strawberry) [32], *PcTPS2* and *PcTPS5* (*P. campanulata*) [34], *PdTPS1* (*Prunus dulcis*) [40], and *MdPIN/CAM* (apple) [11]. An evolutionary analysis on the TPS-a genes of Poaceae exhibited that some TPS-a members can convert GPP into monoterpenes derived from an initial C6-C1 closure [41]. Further studies must be conducted to investigate the ability of RcTPS39 to catalyze monoterpene products. The other four *RcTPS-a* genes were mainly clustered with TPSs that catalyzed (*E*,*E*)-FPP to C_{15} products by an initial C10-C1 or C11-C1 closure. RcTPS32, RcTPS32, and RcTPS46 might have the ability to convert (*E*,*E*)-FPP to produce sesquiterpenes.

Differential *RcTPS* expression and volatile terpenes in two flower developmental stages

Based on the differential expression patterns of *RcTPS* and related genes, the butterfly rose petals on the anthesis day (D1, fresh flower) and day 3 post-anthesis (D3, old flower) were selected for differential gene expression analysis ($Log_2FC \geq 0.8$ and $P < 0.05$) within nine *RcTPS* and six *RcNUDX1* genes (Fig. 6a). Compared with D1 samples, four *RcTPS*s and one *RcNUDX1* were upregulated and three *RcTPS* genes were downregulated in D3 samples.

Then, the volatile terpenes emitted from these two samples were analyzed. In total, 24 monoterpenoids and 30 sesquiterpenoids were identified (Fig.6b), whose content was higher than the content of those detected from headspace volatiles due to the high extraction temperature (Supplemental Table S4). The levels of almost all monoterpenes emitted from D3 samples were higher than the levels of those from D1 samples, whereas different sesquiterpenes exhibited peak release from D1 or D3 samples (Fig. 6c). As the variable importance in projection (VIP) value of most volatile terpenes was very low, differential metabolites were screened based on $log_2FC \geq 0.8$ and $VIP \geq 0.6$. Compared with D1 samples, 14

compounds in the D3 sample including two sesquiterpenes and 12 monoterpenoids were upregulated, whereas five sesquiterpenoids were downregulated.

The relationship between differentially expressed genes (DEGs) and differential volatiles was analyzed. *RcGDS* expression was higher in D1 samples than in D3 samples, which was consistent with the decreased germacrene D emission in D3 samples (40% of D1 samples). However, germacrene D was not classified as a differential volatile because of its low VIP value. Increased *RcLINS* and decreased *RcLINNERS3* were expressed in D1 samples, whereas the linalool emission level was higher in D3 samples (130% of D1 samples). The chirality of linalool could not be detected due to the detection method limitations. Thus, it is difficult to analyze the correlation between linalool and these two genes. The emission abundances of geraniol, nerol, and some monoterpenes [myrcene, (*Z*)-ocimene, and (*E*)-β-ocimene] from the D3 samples were 1.3–2.1 times those from the D1 samples, which may be related to the upregulated expression of *RcNUDX1-1a4* in D3 samples.

The two sesquiterpenes upregulated in D3 samples were not identified, and the RcTPS32/33 and RcTPS46 functions could be speculated. *RcTPS18* expression was downregulated in D3 samples, and the (*E*,*E*)-α-farnesene release in D3 samples also decreased (67% of D1 samples). However, (*E*,*E*)-α-farnesene was classified as a nonsignificant volatile because of its low VIP value. *Rosa* 'Qinglian Xueshi' that released high levels of (*E*,*E*)-α-farnesene was selected as control for further analysis. The emission amounts of (*E*,*E*)-α-farnesene in the three rose samples had the same trend as the *RcTPS18* expression levels (Fig. 6d). Although the emission of five sesquiterpenes was upregulated in D1 samples, the genes responsible for their synthesis remain unknown.

Similar expression patterns of *RcTPS32, RcTPS33,* and *RcTPS46* genes

To analyze the similar expression patterns of *RcTPS32*, *RcTPS33*, and *RcTPS46* genes, the following three criteria were applied to identify these genes: (1) the average FPKM value in the petal samples at five developmental stages was ≥ 1, (2) the expression of genes had to be upregulated in D3 samples and downregulated in D1 and S3 samples, and (3) the expression of genes was significantly different between D2 and bagging treatment samples (LD and DD). The DEGs between different samples were screened out (Fig. 7a). Venn analysis exhibited similarity in the expression patterns of 91 genes with those of *RcTPS32*, *RcTPS33*, and *RcTPS46* (Fig. 7b). These genes included some cell wall metabolism-related genes (e.g., β-xylosidase, β-galactosidase, cellulose synthase-like, pectinesterase/pectinesterase inhibitor), sugar metabolism-related genes (e.g., trehalose-phosphate synthase and UDP-glycosyltransferase), reactive oxygen species scavenging enzymes and antioxidants (e.g., glutathione S-transferase and peroxidases), purple acid phosphatase (PAP) genes, and four transcription factors (namely *RcMYB108* [42], *RcWRKY55* [43], *RcNAC062* [44], and *RhPMP1* [45]) (Fig. 7c). These four transcription factors were selected and further validated through qRT-PCR, confirming their expression patterns in different samples (Figure S6).

Discussion

Expansion and evolution of the TPS gene family in the *R. chinensis* genome

The present study documented that the *R. chinensis* genome comprised 49 putative functional TPSs, indicating that the number of TPSs in *R. chinensis* is higher than those reported in apple (10), poplar (38), and soybean (23) and lower than that in grapevine (69) (Table 1). Unlike the TPS gene family in *Mentha longifolia* (Lamiaceae) that has segmentally duplicated TPS genes [46], no such TPS genes were observed in the rose genome. Similarly, segmentally duplicated TPS genes were detected in neither grapevine nor peach [3]. The organization and diversity of the TPS gene family were affected significantly by the genome duplication events [11]. Grapevine, peach, and *R. chinensis* exhibited only the core eudicot-specific gamma whole-genome triplication with no recent polyploidization [47]. Thus, the TPS gene family expansion in *R. chinensis* is driven primarily by tandem duplication.

The TPS-a subfamily is substantially expanded in *R. chinensis* with 32 *RcTPS*s, which are more than the TPS-a genes in grapevine and other rosid species [3]. Among the five TPS gene clusters formed in the rose genome, three were found to be composed of all TPS-a genes. Only a part of the genes in one gene cluster, which was located on chromosome 5 and comprised *RcTPS32*, *RcTPS33*, and *RcTPS34*, was expressed in MU petals. The RcTPS32 and RcTPS33 protein sequences were identical, which indicated recent gene duplication [48]. Two identical TPS genes distributed adjacently on the same gene cluster were also observed in arabidopsis [49]. Despite sharing a sequence similarity of 71%, the *RcTPS34* expression was undetectable in rose petals. The differential expression of these genes indicated the functional divergence of the duplicated members.

Four complete TPS-g genes that were distributed in a small segment (65 kb) of the chromosome were present in *R. chinensis*. *RcLIN-NERS1* exhibited collinearity with other plants, which indicates its ancient origin (Fig. 3). However, *RcLIN-NERS1* and the other three TPS-g genes differed in gene structure (Fig. 1) and were located on different phylogenetic clusters (Fig. 5). Moreover, *RcLIN-NERS1* and *RcLIN-NERS3* exhibited different expression patterns during the flowering stage (Fig. 4b), indicating the divergence of RcTPS-g gene functions.

In the RcTPS-b subfamily, both *RcTPS-p8* and *RcTPS18* had collinear TPS gene pairs in other plants (Fig. 3). As a putative pseudogene, the collinear genes of *RcTPS-p8* in *R. rugosa* (Chr6.5828, 587 aa) and peach (ONI18546, 629 aa) may be functional, although those in grapevine (VIT_12s0059g02710, 519 aa) and strawberry (FvH4_6g43710, 144 aa) were also putative pseudogenes. Thus, *RcTPS-p8* might have lost fragments from its ancestral gene during evolution. *RcTPS18*, another collinearly conserved TPS-b gene, was located on a special phylogenetic cluster, which differed from other RcTPS-b genes (Fig. 5). Some TPSs in this cluster can catalyze GPP and FPP to acyclic monoterpenes or sesquiterpenes, respectively. Although several TPSs have broad substrate specificity and catalyze several substrates *in vitro*, their function *in vivo* may be limited due to their subcellular localization [50]. The RcTPS18 protein sequence was predicted to be located in the cytoplasm. Thus, it may use FPP as a substrate to catalyze

the formation of acyclic terpenoids, demonstrating that the TPS-b gene subfamily has undergone complex gene loss and duplication events [51].

Different TPS expression profiles during flower developmental stages

Some *RcTPS* genes exhibited peak expression in fresh flowers or buds that are about to open. In most plants producing floral scents, volatile emission peaks when the flowers are ready for pollination and decreases afterwards [52]. Corresponding to this phenomenon, the TPS genes encoding scent biosynthetic enzymes typically peak 1–2 days ahead of emission of the corresponding compound, and the related TPS expression decreases during petal senescence stages when scent emission declines with the decrease in *LoTPS1* and *LoTPS2* expression in *Lilium* 'Siberia' [53-55]. Previous studies have reported that germacrene D emitted from the petals of *Rosa* 'Fragrant Cloud' reached a maximum value in mature petals and then decreased [19, 21], which is similar to the *RcGDS* expression pattern observed in the present study.

Notably, some *RcTPS*s exhibited increased expression in late flowering petals. In *Osmanthus fragrans* flowers, the expression of *OfTPS2* that exclusively produced linalool increased from the full flowering stage to the late full flowering stage. Only a small amount of linalool and its oxides were released at the late full flowering stage, whereas more glycosylated linalool and its oxides were accumulated in the flower [56]. Some other plants also released or accumulated higher terpenoids in old flowers. For example, higher levels of 1,8-cineole and β-ocimene were emitted from senescent ginger (*Hedychium coronarium*) flowers, and maximum monoterpenes accumulated in old flowers of wild *Rosa* species [27, 57]. Additionally, higher levels of sesquiterpenes were released in senescent butterfly and 'Honesty' rose flowers [23, 27]. Further studies are required to investigate the correlation between increased TPS gene expression in senescent petals and increased terpene emission because glycosylation might be involved in regulating the release of floral volatiles.

Effect of light on *RcTPS* expression

According to our observations, many bees visited newly opened MU flowers (D1) and remained there through the morning. In the evening, the petals of fresh MU flowers closed and opened again the next day, proving that MU mainly attracts daytime pollinators. Similarly, the *R. damascena* and *R. hybrida* flowers exhibited maximum emission during the daytime [23, 58]. *RcTPS* expression in MU petals was suppressed under dark treatment. This may be because some TPS genes are directly controlled by light. For example, the *RcGDS* expression level decreased significantly in the dark environment, which is consistent with the previous results indicating that *RcGDS* expression was directly regulated by light [59]. Light-induced HY5, a basic leucine zipper transcription factor, is crucial for light-mediated transcriptional regulation of TPS in *AtTPS03* (ocimene/farnenese synthase) [60] and β-pinene synthase *QH6* [61]. Additionally, *RcDXS2-1* (homologous gene of *RrDXS*) and *RcDXR* involved in the MEP pathway were downregulated by dark treatment (Fig. 4c), which may be related to the downregulated expression of light signal components in the dark environment. Two of the light signaling components, HY5 and PIF, have been demonstrated to differentially multitarget the genes in initial steps of the MEP pathway, which

would limit the flux through the MEP pathway [62, 63]. Significant differences were observed in the expression of *RcHY5* and two *RcPIFs* genes under different light conditions (Fig. S7) [64]. Further studies are warranted to investigate their involvement in the MEP pathway regulation.

Potential functions of some *RcTPS*s in rose petals

The expression levels of three genes, namely *RcTPS32*, *RcTPS33*, and *RcTPS46*, peaked in old petals, all of which exhibited increased expression after bagging treatment (light or dark conditions). Their expression patterns were significantly different from those of other *RcTPS* genes, indicating that they may have other functions. Some genes exhibited similar expression patterns related to the senescence or death of plant cells [4]. *RhMYB108* is a characterized senescence marker gene that exhibits higher expression in senescent petals than in young petals [42]. *RhPMP1* was upregulated by ethylene, which induced *RhAPC3b* expression, resulting in asymmetric growth of the petal base [45]. β-Galactosidase expression increased in senescent rose and *Petunia* petals, whereas β-xylosidase expression increased in senescent carnations (*Dianthus caryophyllus*), both of which may be involved in cell wall degradation [65-67]. The *PAP* genes are broadly involved in liberating Pi from phosphomonoesters in senescing tissues such as leaves, thus helping in Pi uptake and distribution in the plant [68].

After bagging treatment, MU flowers exhibited some characteristics similar to those of old flowers, such as shrinkage and easy-to-fall petals. This may be because the aluminum foil bag or semi-transparent bag used in the treatment prevented the emission of flower volatiles into the surrounding environment, resulting in an enrichment of ethylene and other volatiles under flower headspace circumstances. Ethylene is a major endogenous signal that can accelerate petal senescence. Additionally, some studies have indicated that monoterpenes play a key role in senescence by suppressing cell division, membrane disruption, and oxidative stress, eventually leading to abscission. Geraniol is a potent inducer of apoptosis-like cell death in plant cells, and glutathione S-transferase is a type of geraniol responsive factor in the cultured soybean cells. Although the catalytic terpenoid products of these three *RcTPS*s are still unclear, they might be involved in petal senescence and cell death.

Conclusions

The present study identified 74 putative *RcTPS* genes, including 49 putative functional and 25 partial/pseudo TPSs in the *R. chinensis* genome. These genes were distributed on all seven chromosomes and divided into five subfamilies (TPS-a, -b, -g, -c, and -e/f). The TPS-a subfamily had 32 members, all of which exhibited a species-specific expansion compared with other plants. Tandem duplication contributed greatly to the increase in the number of TPS genes in *R. chinensis*. Three complete *RcTPS* genes and one *RcTPS-p* gene have probably originated from ancient dicotyledonous plants. Only nine *RcTPS* genes were expressed in the opening petals of butterfly rose, and they exhibited different expression patterns at different developmental stages or under dark treatment. The *Rosa* ‘Qinglian Xueshi’ petals exhibited greater emission of (*E*,*E*)-a-farnenese and higher expression of *RcTPS18* than butterfly rose petals. A combined analysis of volatile emissions, sequence homology, and conserved

structural features indicated that *RcTPS18*, a member of the TPS-b subfamily, encodes (*E*,*E*)-α-farnesene synthase. Three *RcTPS*s whose functions are not characterized exhibited increased expression in senescent flowers and dark treated samples, and these genes might be involved in petal senescence and cell death. The present study provided valuable insights into the terpenoid biosynthesis mechanism in rose flowers.

Materials And Methods

Identification and sequence analysis of the *RcTPS* gene family

The complete rose (*R. chinensis* 'Old Blush') genome sequence was obtained from the official website (https://lipm-browsers.toulouse.inra.fr/pub/RchiOBHm-V2/). The consensus protein sequence of the TPS hidden Markov model (HMM) was downloaded from Pfam (PF03936 and PF01397) to identify TPS genes in the rose genome [31]. This HMM profile was then used as a query to search the rose genome, resulting in the identification of all candidate sequences containing either PF01397 or PF03936, with an E-value < $1e^{-3}$. All candidate rose TPS genes were aligned using MAFFT software before manually figuring out the conserved regions. The genes with all conserved regions and expected length were classified as complete RcTPS genes, whereas those with incomplete or mutated conserved regions were classified as partial/pseudo (TPS-p) genes. Signal peptides were predicted using the programs TargetP, ChloroP, and LOCALIZER [8, 73].

Phylogenetic tree construction and structural analysis

The protein sequences of identified TPSs were aligned using MUSCLE software. Then, the maximum likelihood phylogenetic tree was constructed using the IQ-TREE program on an online platform (www.omicshare.com/tools). The exon–intron organization of *R. chinensis* TPS genes was illustrated using the online program Gene Structure Display Server (GSDS: http://gsds.cbi.pku.edu.cn) [74]. Furthermore, another online program MEME was used for identification of conserved motifs in the TPS protein sequences. The optimized MEME parameters were set as default. However, the optimum width of each motif was 10–60 aa residues. The MEME results were visualized using TBtools software [75, 76].

Identification of genes involved in the terpene biosynthesis pathways

The terpene-related genes involved in the MVA and MEP pathways were identified. The single or two copied genes were searched from annotated documents of the rose genome and double checked by comparison with the homologous genes in *Arabidopsis* and other characterized genes (Supplementary Table S5) [77, 78], namely 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase (MCT), 4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol kinase (CMK), 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase (MDS), 4-hydroxy-3-methylbut-2-enyl-diphosphate synthase (HDS), 4-hydroxy-3-methylbut-2-enyl diphosphate reductase (HDR), isopentenyl pyrophosphate isomerase (IPI), acetyl-CoA C-acetyltransferase (AACT), hydroxymethylglutaryl-CoA synthase (HMGA), mevalonate kinase (MK), phosphomevalonate kinase (PMK), and diphosphomevalonate decarboxylase (MDC). The HMM profile (PF00368) was used

as a query to search the rose genome for the genes belonging to the 3-hydroxy-3-methylglutaryl coenzyme A reductase (HMGR) gene family [79]. The amino acid sequences of candidate genes were aligned using MAFFT software and then manually confirmed based on the presence of the conserved domain. Six homologous *RcNUDX1* genes were identified [26], and genes belonging to the PT gene family (including *cis*- and *trans*-PT) were identified in our previous study (unpublished).

Chromosomal localization and collinearity analyses

The rose TPS genes were mapped on the chromosomes according to their positions in the annotated genome documents by using TBtools software [76]. The tandemly duplicated genes were confirmed based on three criteria: (a) length of alignable sequence covered > 70% of the longer gene; (b) similarity of aligned regions > 70%; (3) close chromosome location (< 100 kb) and few separated genes (≤ 5) [80]. Collinearity analysis within *R. chinensis* was conducted, and segmentally duplicated genes were identified in the collinear segments. The whole-genome sequences and annotation documents of grapevine, peach (*Prunus persica*), strawberry (*Fragaria vesca*), and *R. rugosa* were downloaded, and the interspecific collinearity analysis between rose and these plants was performed using TBtools software to determine the interspecies collinear relationships among orthologous TPS genes [76].

Plant materials

Two old Chinese rose cultivars, butterfly rose (*R. chinensis* 'Mutabilis,' MU) and *Rosa* 'Qinglian Xue shi' (QL), were collected from Kunming Yang Chinese Rose Gardening Co., Ltd. and planted in the germplasm garden of our institute under open field conditions (116°43'N, 40°16'E) for 2–3 years. The plant materials were identified by Ying Kong and Jinrong Bai. Although the breeding time is unknown, it is recognized that these two old rose cultivars have been bred for hundreds of years. They are not endangered or protected plants, and no other specific permissions were required for collection or research.

According to our observation, the MU flowering process lasts approximately 4 days. Different floral developmental stages of MU petals, namely bud about to open (S3), first day of anthesis (D1), second day of anthesis (D2), third day of anthesis (D3), and fourth day of anthesis (D4), were analyzed (Fig. 8). The upper half of rose petals was sampled at 8:00 am–9:00 am on sunny days and frozen in liquid nitrogen, and the samples were stored at −80°C. Three biological replicates collected on different days were used as samples.

To study the effect of light on MU flowers, the buds about to open were wrapped in aluminum foil in the afternoon for continuous dark treatment (DD) [64]. Simultaneously, a translucent paper bag (sketch tracing paper, Rotring, 78% light transmittance) was used as a control treatment for the flower buds (light-dark conditions, LD). The petals were sampled in the morning of the second day when the buds were supposed to bloom (Fig. 8).

RNA-seq analysis

Seven butterfly rose samples at different developmental stages and under different treatment were used for RNA-seq. The samples stored at −80°C were sent to Guangzhou Gene Denovo Biological Technology Co., Ltd. (Guangzhou, China) to perform RNA isolation, RNA-seq library preparation, and sequencing [82]. The libraries of three biological replicates were prepared independently. After removing low-quality reads, the clean reads were mapped to the *R. chinensis* reference genome (https://lipm-browsers.toulouse.inra.fr/pub/RchiOBHm-V2/), and the FPKM (fragments per kilobase million) value was used to determine the gene expression levels. The raw sequence data reported in this paper have been deposited in the Genome Sequence Archive [83] in National Genomics Data Center [84], China National Center for Bioinformation / Beijing Institute of Genomics, Chinese Academy of Sciences (GSA: CRA006521) that are publicly accessible at https://ngdc.cncb.ac.cn/gsa.

Volatile sampling and gas chromatography–mass spectrometry (GC-MS) analysis

The fresh petals (D1) and old petals (D3) of butterfly rose were selected for floral volatile analysis, whereas the fully bloomed QL flowers collected at the same time were used as control. After the samples were ground into powder in liquid nitrogen, 1 g (1 mL) of the powder was transferred immediately to a 20-mL headspace vial (Agilent, Palo Alto, CA, USA) containing NaCl-saturated solution to inhibit any enzymatic reaction [85, 86]. The vials were sealed using crimp-top caps with TFE-silicone headspace septa (Agilent) and warmed at 100°C for 5 min. Then, 120-μm divinylbenzene, carboxen, or polydimethylsilioxan fiber (Agilent) was exposed to the sample headspace for 15 min at 100°C.

VOC identification and quantification were conducted using an Agilent Model 8890 GC and a 5977B mass spectrometer (Agilent) equipped with a DB-5MS (30 m × 0.25 mm × 0.25 μm) capillary column. After sampling, desorption was performed at 250°C for 5 min in the split-less mode of the GC apparatus. Helium was used as the carrier gas at a linear velocity of 1.2 mL/min. The oven temperature was programmed from 40°C (3.5 min), increasing at 10°C/min to 100°C, at 7°C/min to 180°C, at 25°C/min to 280°C, and hold for 5 min. Other GC-MS analytical conditions used were as described by Gong et al. [86]. Volatile compounds were identified by comparing the mass spectra with the data system library (MWGC or NIST) and retention index.

qRT-PCR analysis

Nine *RcTPS*s and four DEGs were used for qRT-PCR validation. *RhUBI2* was used as an internal control [45]. Total RNA was extracted using the Spin Column Plant Total RNA Purification Kit (Sangon Biotech Co., Ltd., Shanghai, China). The first strand of cDNA was synthesized from 0.6 μg of total RNA by using the Maxima First Strand cDNA Synthesis Kit (Thermo Scientific, USA). Primers were designed using Primer Premier 5.0 software and synthesized by Sangon Biotech (Shanghai, China) (Table S6). qRT-PCR was performed on a LightCycler480 II instrument (Roche, New York, USA). Each reaction was conducted in a 20-μL mixture containing 10 μL of 2 × SybrGreen qPCR Master Mix, 7.2 μL of RNase-free H_2O, 2 μL of cDNA, 0.4 μL of forward primer, and 0.4 μL of reverse primer. The PCR machine was programmed as

follows: 3 min at 95°C, 10 min at 95°C; 45 cycles of 5 s at 95°C, 30 s at 60°C. The relative gene expression was calculated using the $2^{-\Delta\Delta CT}$ method [87].

Data analyses

Differential expression analysis was performed using edgeR in OmicShare tools (www.omicshare.com/tools), and the FDR was used to determine the threshold of the P-value. Heat map analysis and orthogonal partial least square discriminant analysis (OPLS-DA) were performed, and the variable importance in projection (VIP) of volatiles was calculated for the screening of differential terpenes [88, 89]. One-way analysis of variance (ANOVA) and *t*-test were performed using SPSS 23.0 software (IBM, Chicago, IL, USA).

Declarations

Acknowledgements

We thank Yuyong Yang, the chairman of Kunming Yang Chinese Rose Gardening Co., Ltd., for his assistance in the experiments.

Funding

This research has been supported by the National Natural Science Foundation of China (31401901), China National Key R&D Program (2019YFD1000400), and the Program for Innovative Research Team in BJAST (IG201404N).

Authors' contributions

YK and JRB designed the experiment. YK, HW, LXL and XYD performed the experiments. YK and HW analyzed the data and wrote the paper. JRB revised this paper. All authors have read and approved the manuscript.

Availability of data and materials

The datasets generated and/or analysed during the current study are available in the NGDC repository (https://ngdc.cncb.ac.cn/gsa/browse/CRA006521). All databases in this study are available to the public.

Ethics approval and consent to participate

All methods were carried out in accordance with relevant guidelines and regulations.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Rudolf JD, Chang C-Y. Terpene synthases in disguise: enzymology, structure, and opportunities of non-canonical terpene synthases. Nat Prod Rep. 2020, 37(3):425–463.
2. Chen X, Köllner TG, Jia Q, Norris A, Santhanam B, Rabe P, et al. Terpene synthase genes in eukaryotes beyond plants and fungi: Occurrence in social amoebae. Proceedings of the National Academy of Sciences. 2016, 113(43):12132–12137.
3. 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(8):2078–2098.
4. Fineschi S, Loreto F, Staudt M, Peñuelas J: Diversification of Volatile Isoprenoid Emissions from Trees: Evolutionary and Ecological Perspectives. In Niinemets Ü, Monson RK, editors. Biology, Controls and Models of Tree Volatile Organic Compound Emissions. Dordrecht: Springer Netherlands; 2013: p. 1–20.
5. Wang C, Chen Q, Fan D, Li J, Wang G, Zhang P. Structural analyses of short-chain prenyltransferases identify an evolutionarily conserved GFPPS clade in Brassicaceae plants. Mol Plant. 2016, 9(2):195–204.
6. Chang H-Y, Cheng T-H, Wang AHJ. Structure, catalysis, and inhibition mechanism of prenyltransferase. IUBMB Life. 2021, 73(1):40–63.
7. Yue Y, Yu R, Fan Y. Characterization of two monoterpene synthases involved in floral scent formation in *Hedychium coronarium*. Planta. 2014, 240(4):745–762.
8. Zhou F, Pichersky E. The complete functional characterization of the terpene synthase family in tomato. New Phytol. 2020, 226(5):1341–1360.
9. Shalit M, Guterman I, Volpin H, Bar E, Tamari T, Menda N, et al. Volatile ester formation in roses. Identification of an acetyl-coenzyme A. geraniol/citronellol acetyltransferase in developing rose petals. Plant Physiol. 2003, 131(4):1868–1876.
10. 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(1):212–229.
11. Nieuwenhuizen NJ, Green SA, Chen X, Bailleul EJD, Matich AJ, Wang MY, et al. Functional genomics reveals that a compact terpene synthase gene family can account for terpene volatile production in apple. Plant Physiol. 2013, 161(2):787–804.
12. Aubourg S, Lecharny A, Bohlmann J. Genomic analysis of the terpenoid synthase (*AtTPS*) gene family of *Arabidopsis thaliana*. Mol Genet Genomics. 2002, 267(6):730–745.

13. Martin DM, Aubourg S, Schouwey MB, Daviet L, Schalk M, Toub O, 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):226.
14. Falara V, Akhtar TA, Nguyen TTH, Spyropoulou EA, Bleeker PM, Schauvinhold I, et al. The tomato terpene synthase gene family. Plant Physiol. 2011, 157(2):770–789.
15. Zhou F, Pichersky E. More is better: the diversity of terpene metabolism in plants. Curr Opin Plant Biol. 2020, 55:1–10.
16. Külheim C, Padovan A, Hefer C, Krause ST, Köllner TG, Myburg AA, et al. The *Eucalyptus* terpene synthase gene family. BMC Genomics. 2015, 16(1):450.
17. Baudino S, Hugueney P, Caissard J-C. Evolution of scent genes. In Pichersky E, Dudareva N, editors. Biology of Plant Volatiles. Boca Raton: CRC Press; 2020: p. 217–234.
18. Tholl D, Gershenzon J. The flowering of a new scent pathway in rose. Science. 2015, 349(6243):28–29.
19. Guterman I, Shalit M, Menda N, Piestun D, Dafny-Yelin M, Shalev G, et al. Rose scent: genomics approach to dscovering novel floral fragrance-related genes. Plant Cell. 2002, 14(10):2325–2338.
20. Lewinsohn E, Vainstein A, Weiss D. An integrated genomics approach to identifying floral scent genes in rose. In Pichersky E, Dudareva N, editors. Biology of floral scent. Boca Raton: CRC Press; 2006: p. 91–102.
21. Shalit M, Shafir S, Larkov O, Bar E, Kaslassi D, Adam Z, et al. Volatile compounds emitted by rose cultivars: Fragrance perception by man and honeybees. Isr J Plant Sci. 2004, 52(3):245–255.
22. Joichi A, Yomogida K, Awano K-i, Ueda Y. Volatile components of tea-scented modern roses and ancient Chinese roses. Flavour Frag J. 2005, 20(2):152–157.
23. Helsper JPFG, Davies JA, Bouwmeester HJ, Krol AF, van Kampen MH. Circadian rhythmicity in emission of volatile compounds by flowers of *Rosa hybrida* L. cv. Honesty. Planta. 1998, 207(1):88–95.
24. Magnard J-L, Bony AR, Bettini F, Campanaro A, Blerot B, Baudino S, et al. Linalool and linalool nerolidol synthases in roses, several genes for little scent. Plant Physiol Bioch. 2018, 127:74–87.
25. Magnard J-L, Roccia A, Caissard J-C, Vergne P, Sun P, Hecquet R, et al. Biosynthesis of monoterpene scent compounds in roses. Science. 2015, 349(6243):81–83.
26. Sun P, Dégut C, Réty S, Caissard J-C, Hibrand-Saint Oyant L, Bony A, et al. Functional diversification in the *Nudix hydrolase* gene family drives sesquiterpene biosynthesis in *Rosa* × *wichurana*. Plant J. 2020, 104(1):185–199.
27. Dani KGS, Fineschi S, Michelozzi M, Trivellini A, Pollastri S, Loreto F. Diversification of petal monoterpene profiles during floral development and senescence in wild roses: relationships among geraniol content, petal colour, and floral lifespan. Oecologia. 2021, 197, 957–969.
28. Xie T, Chen C, Li C, Liu J, Liu C, He Y. Genome-wide investigation of WRKY gene family in pineapple: evolution and expression profiles during development and stress. BMC Genomics. 2018, 19(1):490.

29. Ma L-T, Lee Y-R, Liu P-L, Cheng Y-T, Shiu T-F, Tsao N-W, et al. Phylogenetically distant group of terpene synthases participates in cadinene and cedrane-type sesquiterpenes accumulation in *Taiwania cryptomerioides*. Plant Sci. 2019, 289:110277.
30. Irmisch S, Jiang Y, Chen F, Gershenzon J, Köllner TG. Terpene synthases and their contribution to herbivore-induced volatile emission in western balsam poplar (*Populus trichocarpa*). BMC Plant Biol. 2014, 14(1):270.
31. Liu J, Huang F, Wang X, Zhang M, Zheng R, Wang J, et al. Genome-wide analysis of terpene synthases in soybean: functional characterization of GmTPS3. Gene. 2014, 544(1):83–92.
32. Aharoni A, Giri AP, Verstappen FWA, Bertea CM, Sevenier R, Sun Z, et al. Gain and loss of fruit flavor compounds produced by wild and cultivated strawberry species. Plant Cell. 2004, 16(11):3110.
33. Pechous SW, Whitaker BD. Cloning and functional expression of an (*E*,*E*)-α-farnesene synthase cDNA from peel tissue of apple fruit. Planta. 2004, 219(1):84–94.
34. Huang K-F, Wen C-H, Lee Y-R, Chu F-H. Cloning and characterization of terpene synthase genes from Taiwan cherry. Tree Genet Genomes. 2019, 15(4):51.
35. Liu H, Cao X, Liu X, Xin R, Wang J, Gao J, et al. UV-B irradiation differentially regulates terpene synthases and terpene content of peach. Plant Cell Environ. 2017, 40(10):2261–2275.
36. Lin J, Wang D, Chen X, Köllner TG, Mazarei M, Guo H, et al. An (*E*,*E*)-α-farnesene synthase gene of soybean has a role in defence against nematodes and is involved in synthesizing insect-induced volatiles. Plant Biotechnol J. 2017, 15(4):510–519.
37. Danner H, Boeckler GA, Irmisch S, Yuan JS, Chen F, Gershenzon J, et al. Four terpene synthases produce major compounds of the gypsy moth feeding-induced volatile blend of *Populus trichocarpa*. Phytochemistry. 2011, 72(9):897–908.
38. Xie X, Kirby J, Keasling JD. Functional characterization of four sesquiterpene synthases from *Ricinus communis* (Castor bean). Phytochemistry. 2012, 78:20–28.
39. Green S, Squire CJ, Nieuwenhuizen NJ, Baker EN, Laing W. Defining the potassium binding region in an apple terpene synthase. J Biol Chem. 2009, 284(13):8661–8669.
40. Nawade B, Yahyaa M, Reuveny H, Shaltiel-Harpaz L, Eisenbach O, Faigenboim A, et al. Profiling of volatile terpenes from almond (*Prunus dulcis*) young fruits and characterization of seven terpene synthase genes. Plant Sci. 2019, 287:110187.
41. Luck K, Chen X, Norris AM, Chen F, Gershenzon J, Köllner TG. 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(1):203–215.
42. Zhang S, Zhao Q, Zeng D, Xu J, Zhou H, Wang F, et al. RhMYB108, an R2R3-MYB transcription factor, is involved in ethylene- and JA-induced petal senescence in rose plants. Hortic Res. 2019, 6(1):131.
43. Liu X, Li D, Zhang S, Xu Y, Zhang Z. Genome-wide characterization of the rose (*Rosa chinensis*) WRKY family and role of RcWRKY41 in gray mold resistance. BMC Plant Biol. 2019, 19(1):522.

44. Geng L, Su L, Fu L, Lin S, Zhang J, Liu Q, Jiang X. Genome-wide analysis of the rose (Rosa chinensis) NAC family and characterization of RcNAC091. Plant Mol Biol. 2022, 108(6): 605–619.
45. Cheng C, Yu Q, Wang Y, Wang H, Dong Y, Ji Y, et al. Ethylene-regulated asymmetric growth of the petal base promotes flower opening in rose (*Rosa hybrida*). Plant Cell. 2021, 33(4):1229–1251.
46. Chen Z, Qi X, Yu X, Zheng Y, Liu Z, Fang H, et al. Genome-wide analysis of terpene synthase gene family in *Mentha longifolia* and catalytic activity analysis of a single terpene synthase. Genes. 2021, 12(4):518.
47. Chen F, Su L, Hu S, Xue J-Y, Liu H, Liu G, et al. A chromosome-level genome assembly of rugged rose (*Rosa rugosa*) provides insights into its evolution, ecology, and floral characteristics. Hortic Res. 2021, 8(1):141.
48. Chen H, Köllner TG, Li G, Wei G, Chen X, Zeng D, et al. Combinatorial evolution of a terpene synthase gene cluster explains terpene variations in *Oryza*. Plant Physiol. 2020, 182(1):480.
49. Chen F, Ro D-K, Petri J, Gershenzon J, Bohlmann J, Pichersky E, et al. Characterization of a root-Specific arabidopsis terpene synthase responsible for the formation of the volatile monoterpene 1,8-cineole. Plant Physiol. 2004, 135(4):1956.
50. Dhandapani S, Tjhang JG, Jang I-C. Production of multiple terpenes of different chain lengths by subcellular targeting of multi-substrate terpene synthase in plants. Metab Eng. 2020, 61:397–405.
51. Ding G, Zhang S, Ma B, Liang J, Li H, Luo Y, et al. Origin and functional differentiation of (*E*)-β-ocimene synthases reflect the expansion of monoterpenes in angiosperms. J Exp Bot. 2020, 71(20):6571–6586.
52. Lynch JH, Pichersky E, Dudareva N. Floral scent metabolic pathways and their regulation. In Pichersky E, Dudareva N, editors. Biology of plant volatiles. Boca Raton: CRC Press; 2020. p. 147–164.
53. Shi S, Duan G, Li D, Wu J, Liu X, Hong B, et al. Two-dimensional analysis provides molecular insight into flower scent of *Lilium* 'Siberia'. Sci Rep. 2018, 8(1):5352.
54. Abbas F, Ke Y, Yu R, Fan Y. Functional characterization and expression analysis of two terpene synthases involved in floral scent formation in *Lilium* 'Siberia'. Planta. 2019, 249(1):71–93.
55. Abbas F, Ke Y, Zhou Y, Ashraf U, Li X, Yu Y, et al. Molecular cloning, characterization and expression analysis of *LoTPS2* and *LoTPS4* involved in floral scent formation in oriental hybrid *Lilium* variety 'Siberia'. Phytochemistry. 2020, 173:112294.
56. Zeng X, Liu C, Zheng R, Cai X, Luo J, Zou J, et al. Emission and accumulation of monoterpene and the key terpene synthase (TPS) associated with monoterpene biosynthesis in *Osmanthus fragrans* Lour. Front Plant Sci. 2016, 6:1232.
57. Yue Y, Yu R, Fan Y. Transcriptome profiling provides new insights into the formation of floral scent in *Hedychium coronarium*. BMC Genomics. 2015, 16(1):470.
58. Picone JM, Clery RA, Watanabe N, MacTavish HS, Turnbull CGN. Rhythmic emission of floral volatiles from *Rosa damascena semperflorens* cv. 'Quatre Saisons'. Planta. 2004, 219(3):468–478.

59. Hendel-Rahmanim K, Masci T, Vainstein A, Weiss D. Diurnal regulation of scent emission in rose flowers. Planta. 2007, 226(6):1491–1499.
60. Michael R, Ranjan A, Kumar RS, Pathak PK, Trivedi PK. Light-regulated expression of terpene synthase gene, *AtTPS03*, is controlled by the bZIP transcription factor, HY5, in *Arabidopsis thaliana*. Biochem Bioph Res Co. 2020, 529(2):437–443.
61. Zhou F, Sun T-H, Zhao L, Pan X-W, Lu S. The bZIP transcription factor HY5 interacts with the promoter of the monoterpene synthase gene *QH6* in modulating its rhythmic expression. Front Plant Sci. 2015, 6:304.
62. Chenge-Espinosa M, Cordoba E, Romero-Guido C, Toledo-Ortiz G, León P. Shedding light on the methylerythritol phosphate (MEP)-pathway: long hypocotyl 5 (HY5)/phytochrome-interacting factors (PIFs) transcription factors modulating key limiting steps. Plant J. 2018, 96(4):828–841.
63. Cordoba E, Salmi M, León P. Unravelling the regulatory mechanisms that modulate the MEP pathway in higher plants. J Exp Bot. 2009, 60(10):2933–2943.
64. Zhang Y, Wu Z, Feng M, Chen J, Qin M, Wang W, et al. The circadian-controlled PIF8–BBX28 module regulates petal senescence in rose flowers by governing mitochondrial ROS homeostasis at night. Plant Cell. 2021, 33(8):2716–2735.
65. Hajizadeh H, Razavi K, Mostofi Y, Mousavi A, Cacco G, Zamani Z, et al. Identification and characterization of genes differentially displayed in *Rosa hybrida* petals during flower senescence. Sci Hortic. 2011, 128(3):320–324.
66. O'Donoghue EM, Somerfield SD, Watson LM, Brummell DA, Hunter DA. Galactose metabolism in cell walls of opening and senescing petunia petals. Planta. 2008, 229(3):709.
67. Hoeberichts FA, van Doorn WG, Vorst O, Hall RD, van Wordragen MF. Sucrose prevents up-regulation of senescence-associated genes in carnation petals. J Exp Bot. 2007, 58(11):2873–2885.
68. Bhadouria J, Giri J. Purple acid phosphatases: roles in phosphate utilization and new emerging functions. Plant Cell Rep. 2021, 41(1):33–51.
69. Khaskheli AJ, Ahmed W, Ma C, Zhang S, Liu Y, Li Y, et al. RhERF113 functions in ethylene-induced petal senescence by modulating cytokinin content in rose. Plant Cell Physiol. 2018, 59(12):2442–2451.
70. Korankye EA, Lada R, Asiedu S, Caldwell C. Plant senescence: the role of volatile terpene compounds (VTCs). Am J Plant Sci. 2017, 8(12):3120–3139.
71. Izumi S, Takashima O, Hirata T. Geraniol is a potent inducer of apoptosis-like cell death in the cultured shoot primordia of *Matricaria chamomilla*. Biochem Bioph Res Co. 1999, 259(3):519–522.
72. Ashida Y, Matsushima A, Hirota T, Watanabe J. Geraniol-inducible glutathione S-transferase in cultured soybean cells. Biosci Biotech Bioch. 2002, 66(1):168–170.
73. Sperschneider J, Catanzariti A-M, DeBoer K, Petre B, Gardiner DM, Singh KB, et al. LOCALIZER: subcellular localization prediction of both plant and effector proteins in the plant cell. Sci Rep. 2017, 7(1):44598.

74. Hu B, Jin J, Guo A-Y, Zhang H, Luo J, Gao G. GSDS 2.0: an upgraded gene feature visualization server. Bioinformatics. 2015, 31(8):1296–1297.
75. Cao Y, Jia H, Xing M, Jin R, Grierson D, Gao Z, et al. Genome-wide analysis of MYB gene family in Chinese bayberry (*Morella rubra*) and identification of members regulating flavonoid biosynthesis. Front Plant Sci. 2021, 12:1244.
76. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, et al. TBtools: An integrative toolkit developed for interactive analyses of big biological data. Mol Plant. 2020, 13(8):1194–1202.
77. Feng L, Chen C, Li T, Wang M, Tao J, Zhao D, et al. Flowery odor formation revealed by differential expression of monoterpene biosynthetic genes and monoterpene accumulation in rose (*Rosa rugosa* Thunb.). Plant Physiol Bioch. 2014, 75:80–88.
78. Rodríguez-Concepción M, Campos N, Ferrer A, Boronat A. Biosynthesis of isoprenoid precursors in Arabidopsis. In Bach TJ, Rohmer M, editors. Isoprenoid Synthesis in Plants and Microorganisms: New Concepts and Experimental *Approaches*. New York, NY: Springer New York; 2013: p. 439–456.
79. Li W, Liu W, Wei H, He Q, Chen J, Zhang B, et al. Species-specific expansion and molecular evolution of the 3-hydroxy-3-methylglutaryl coenzyme A reductase (HMGR) gene family in plants. PLoS One. 2014, 9(4):e94172.
80. Zhao P, Wang D, Wang R, Kong N, Zhang C, Yang C, et al. Genome-wide analysis of the potato *Hsp20* gene family: identification, genomic organization and expression profiles in response to heat stress. BMC Genomics. 2018, 19(1):61.
81. Ling H-Q, Ma B, Shi X, Liu H, Dong L, Sun H, et al. Genome sequence of the progenitor of wheat A subgenome *Triticum urartu*. Nature. 2018, 557(7705):424–428.
82. Song J, Zhang Y, Song S, Su W, Chen R, Sun G, et al. Comparative RNA-Seq analysis on the regulation of cucumber sex differentiation under different ratios of blue and red light. Bot Stud. 2018, 59(1):21.
83. Chen T, Chen X, Zhang S, Zhu J, Tang B, Wang A, et al. The genome sequence archive family: toward explosive data growth and diverse data types. Genomics, Proteomics Bioinf. 2021, 19(4):578–583.
84. CNCB-NGDC members and partners. Database resources of the national genomics data center, China national center for bioinformation in 2022. Nucleic Acids Res. 2022, 50(D1):D27-D38.
85. Wei G, Tian P, Zhang F, Qin H, Miao H, Chen Q, et al. Integrative analyses of nontargeted volatile profiling and transcriptome data provide molecular Insight into VOC diversity in cucumber plants (*Cucumis sativus*). Plant Physiol. 2016, 172(1):603–618.
86. Gong C, Diao W, Zhu H, Umer MJ, Zhao S, He N, et al. Metabolome and transcriptome integration reveals insights into flavor formation of 'Crimson' watermelon flesh during fruit development. Front Plant Sci. 2021, 12:629361.
87. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. Methods. 2001, 25(4):402–408.
88. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics. 2016, 32(18):2847–2849.

89. Kong Y, Wang H, Lang L, Dou X, Bai J. Metabolome-based discrimination analysis of five *Lilium* bulbs associated with differences in secondary metabolites. Molecules. 2021, 26(5):1340.

Figures

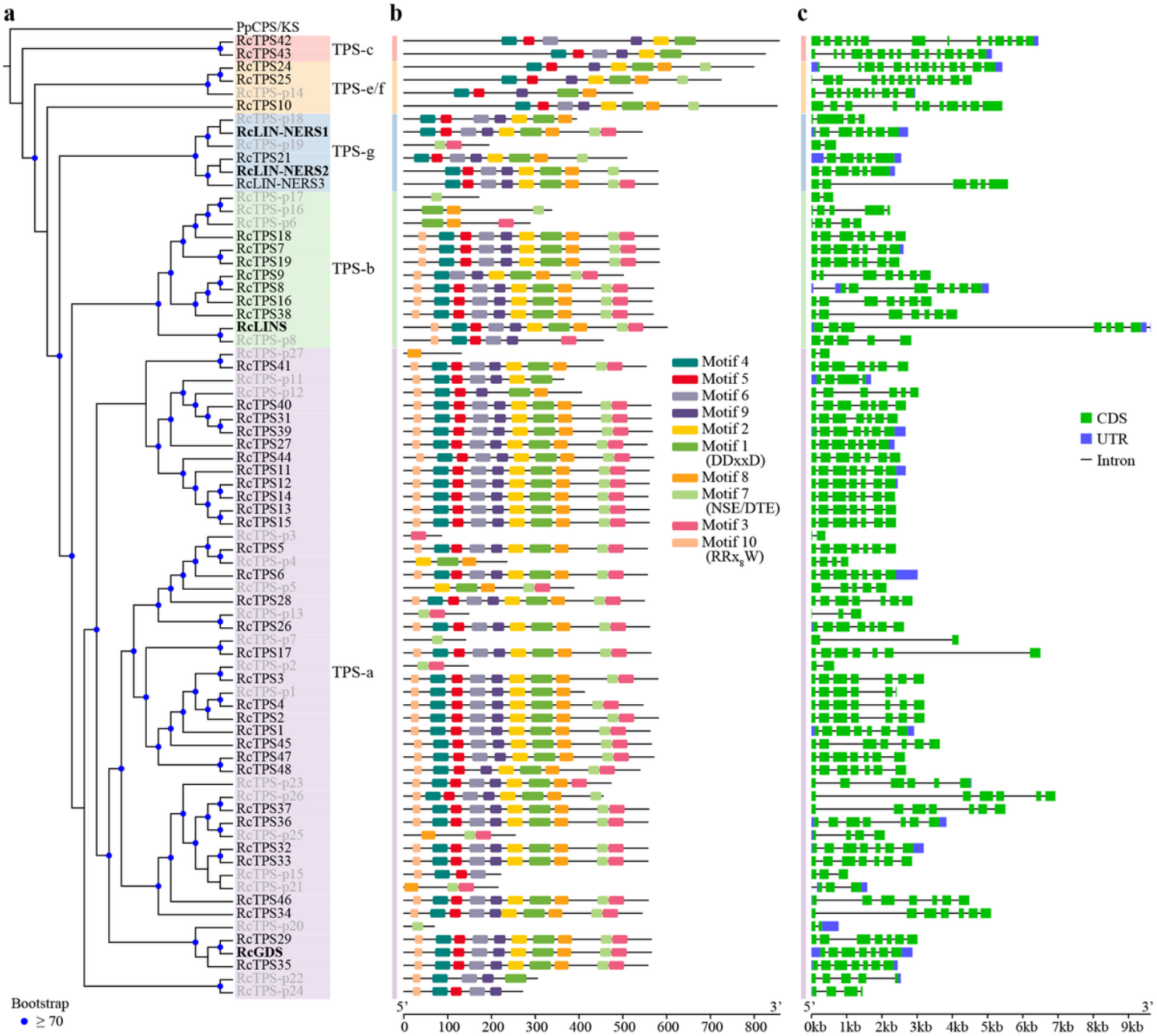

Figure 1

Phylogenetic relationships (a), conserved motifs (b), and gene structure analysis (c) of the *R. chinensis* TPS gene family. A maximum likelihood phylogenetic tree was constructed by aligning the amino acid sequences of 74 putative TPS proteins in the *R. chinensis* genome. Black letters represent putative complete TPS genes, and gray letters represent putative partial/pseudo TPS (TPS-p) genes. The

characterized genes are in bold, except *RcLIN-NERS3* because its cloned sequence is incomplete. Five subfamilies are illustrated with different colors. Ten conserved motifs are illustrated in different colors, and their specific sequence information is provided in Figure S2.

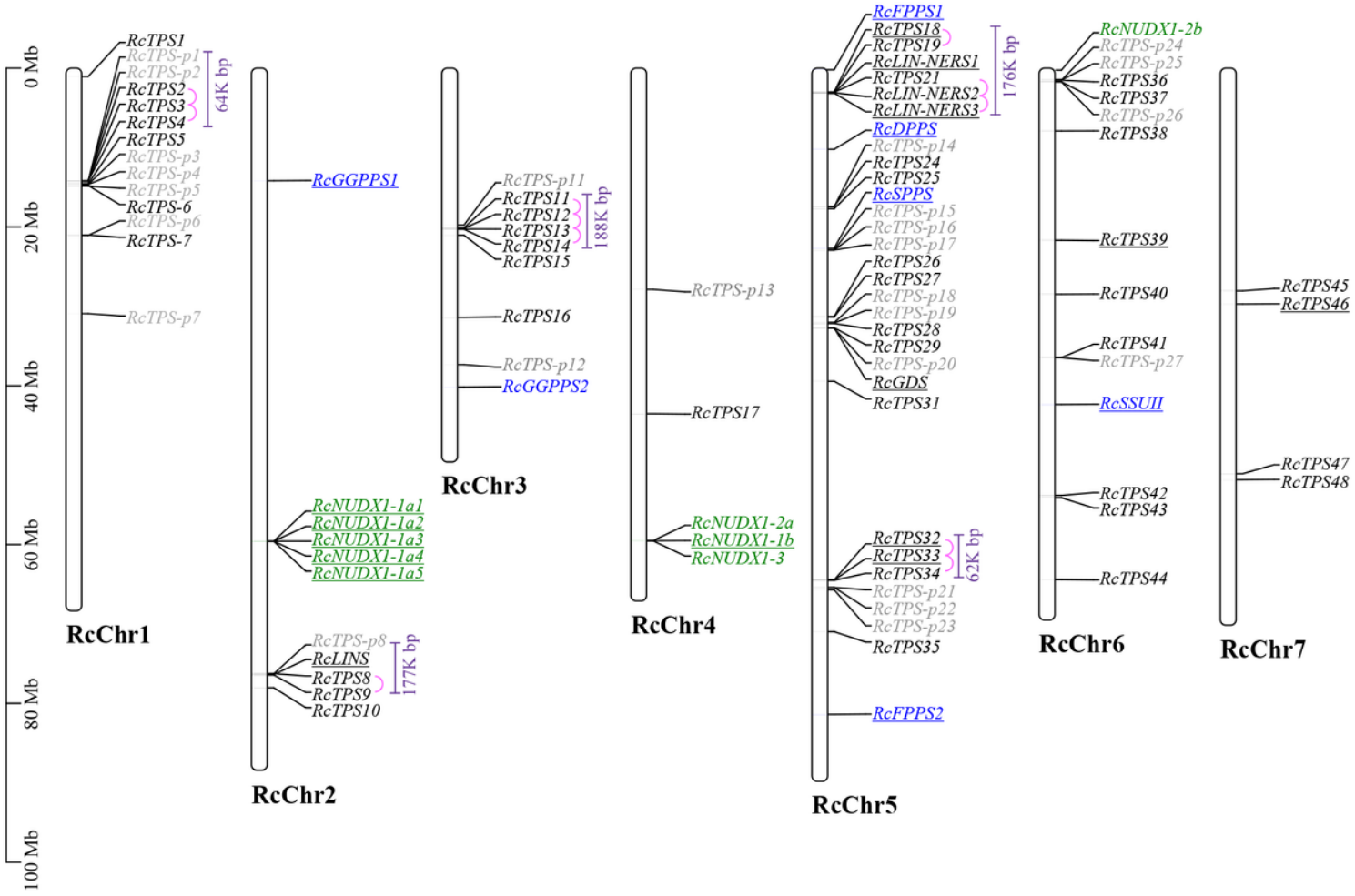

Figure 2

Chromosomal distribution of rose TPS and other genes in the *R. chinensis* genome. Black letters represent putative complete TPS genes, gray letters represent putative partial or pseudo TPS (*TPS-p*) genes, blue letters represent *cis*-prenyltransferase (CPT) and *trans*-prenyltransferase (TPT) genes, and green letters represent *RcNUDX1* genes. The yellow dots indicate genes expressed in open flowers of butterfly rose. The tandemly duplicated TPS genes are indicated in pink lines, and gene clusters are indicated in purple lines.

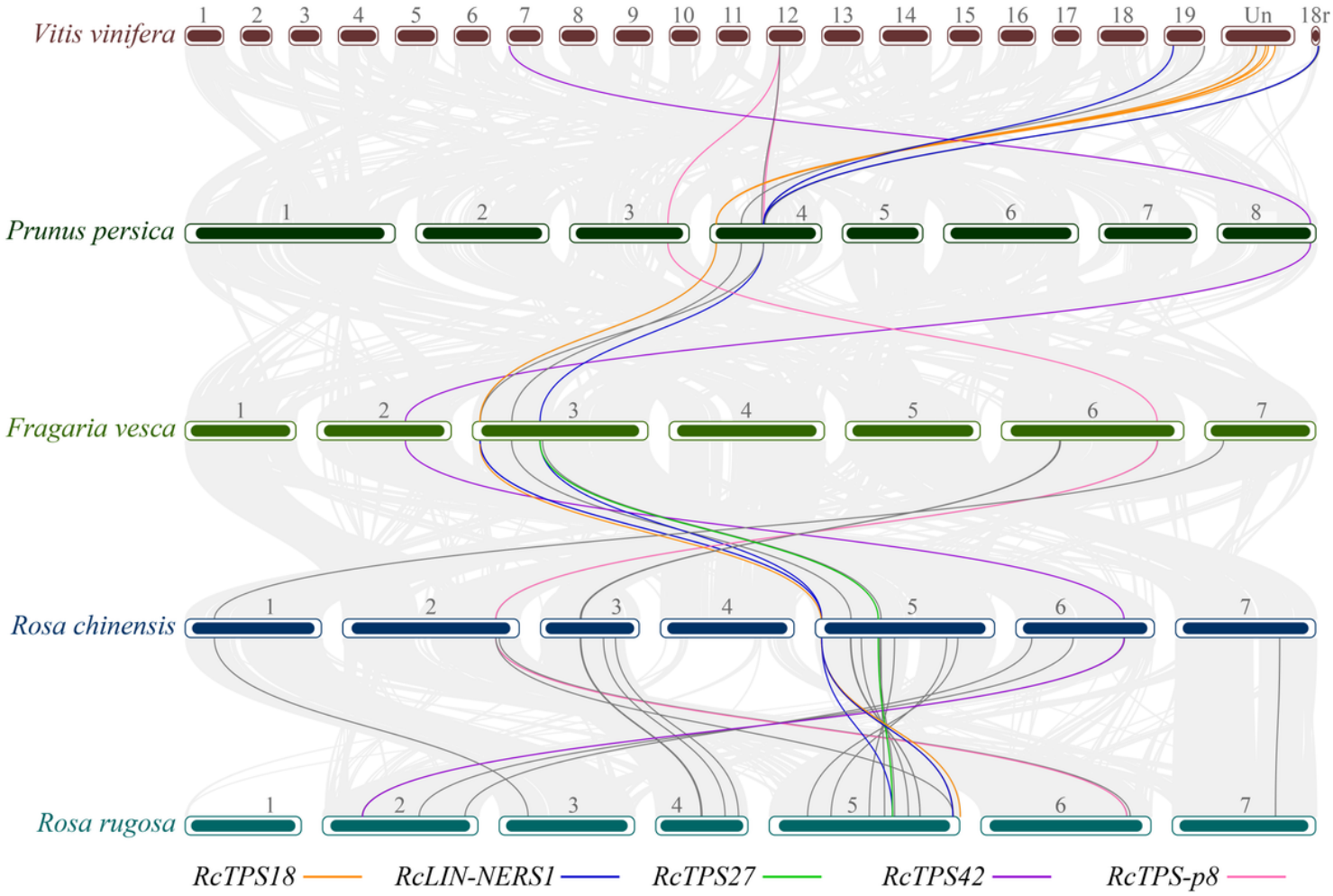

Figure 3

Collinearity analysis of TPS genes among *R. chinensis* and other representative plant species. Gray lines in the background indicate the collinear blocks between rose and other plant genomes, whereas the color and black lines highlight the syntenic TPS gene pairs.

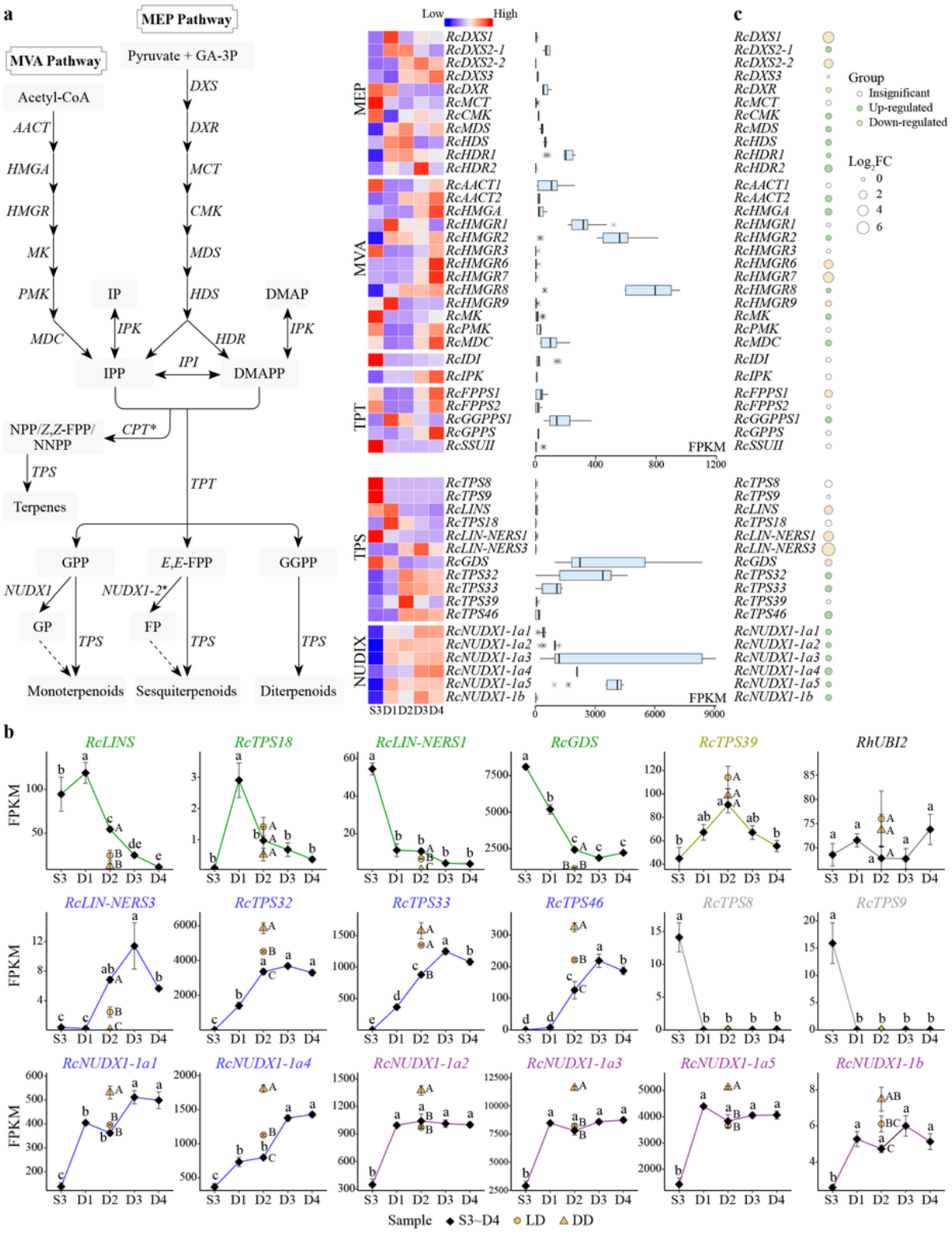

Figure 4

Expression profiles of *RcTPS* and terpene-related genes at different flower developmental stages and under dark treatment. (a) Terpene biosynthesis pathways in rose [15, 26] and expression levels of *RcTPS* and terpene-related genes. The genes marked with * indicate that the average FPKM of all members of this family was less than 1 in MU samples. The corresponding IDs of these genes are listed in Table S2. (b) *RcTPS* and *RcNUDX1* gene expression levels in different samples of *R. chinensis* 'Mutablis' petals.

Bars represent the standard error (n = 3). The genes with different expression patterns are illustrated in different colors. Different lowercase letters indicate statistically significant differences among samples from different developmental stages (P < 0.05). Different uppercase letters indicate statistically significant differences among D2, LD, and DD samples (P < 0.05). (c) Dot plot of the differential expression genes between LD (CK) and DD samples.

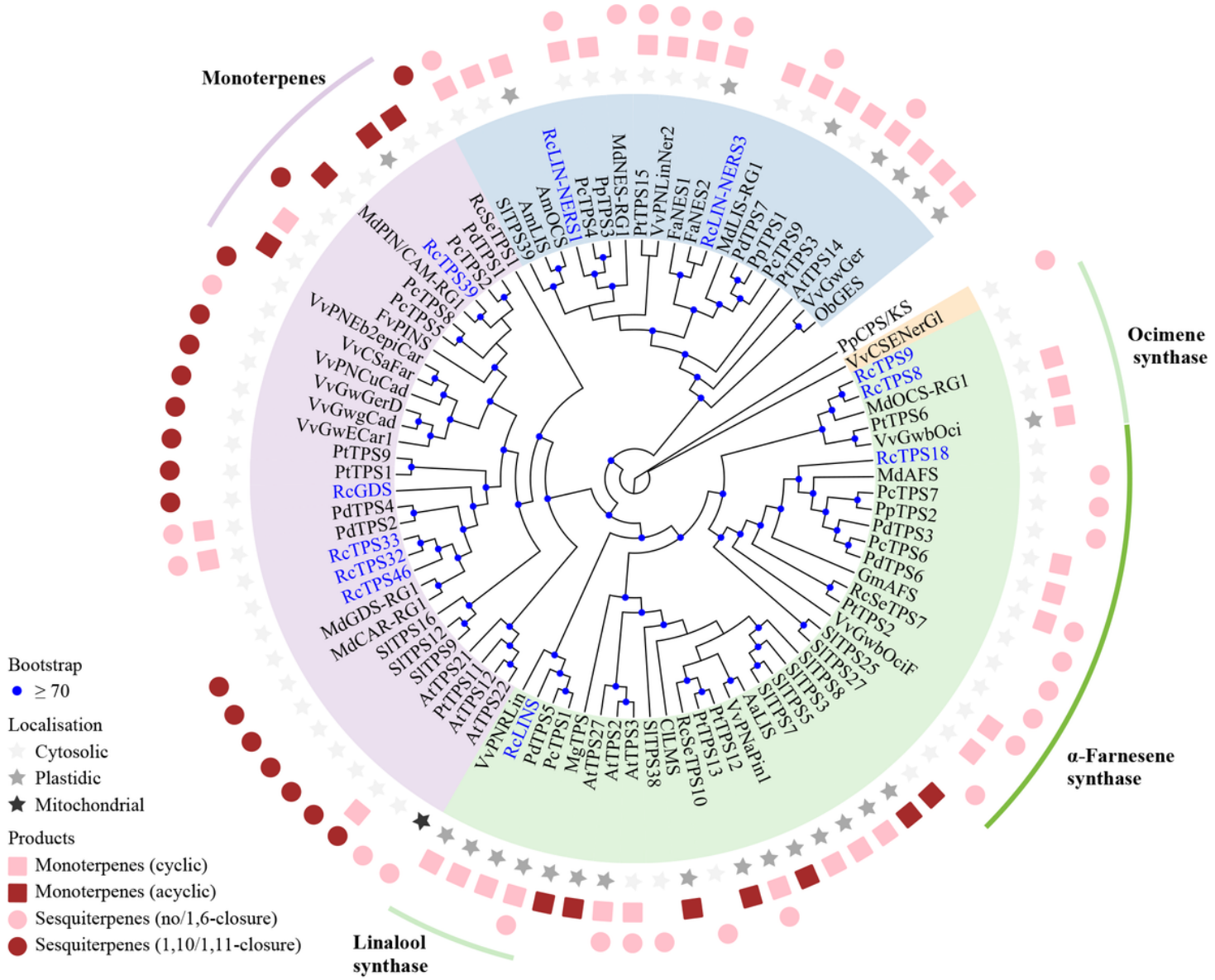

Figure 5

Maximum likelihood phylogeny of 11 RcTPSs and other characterized TPSs. RcTPSs are highlighted in blue. *ent*-kaurene synthase from *Physcomitrella patens* (PpCPS/KS) was used as the root of the tree. The sequences used in this analysis are listed in Table S3. The subfamilies are illustrated with different colors: TPS-a (purple), TPS-b (green), TPS-e/f (orange), and TPS-g (blue).

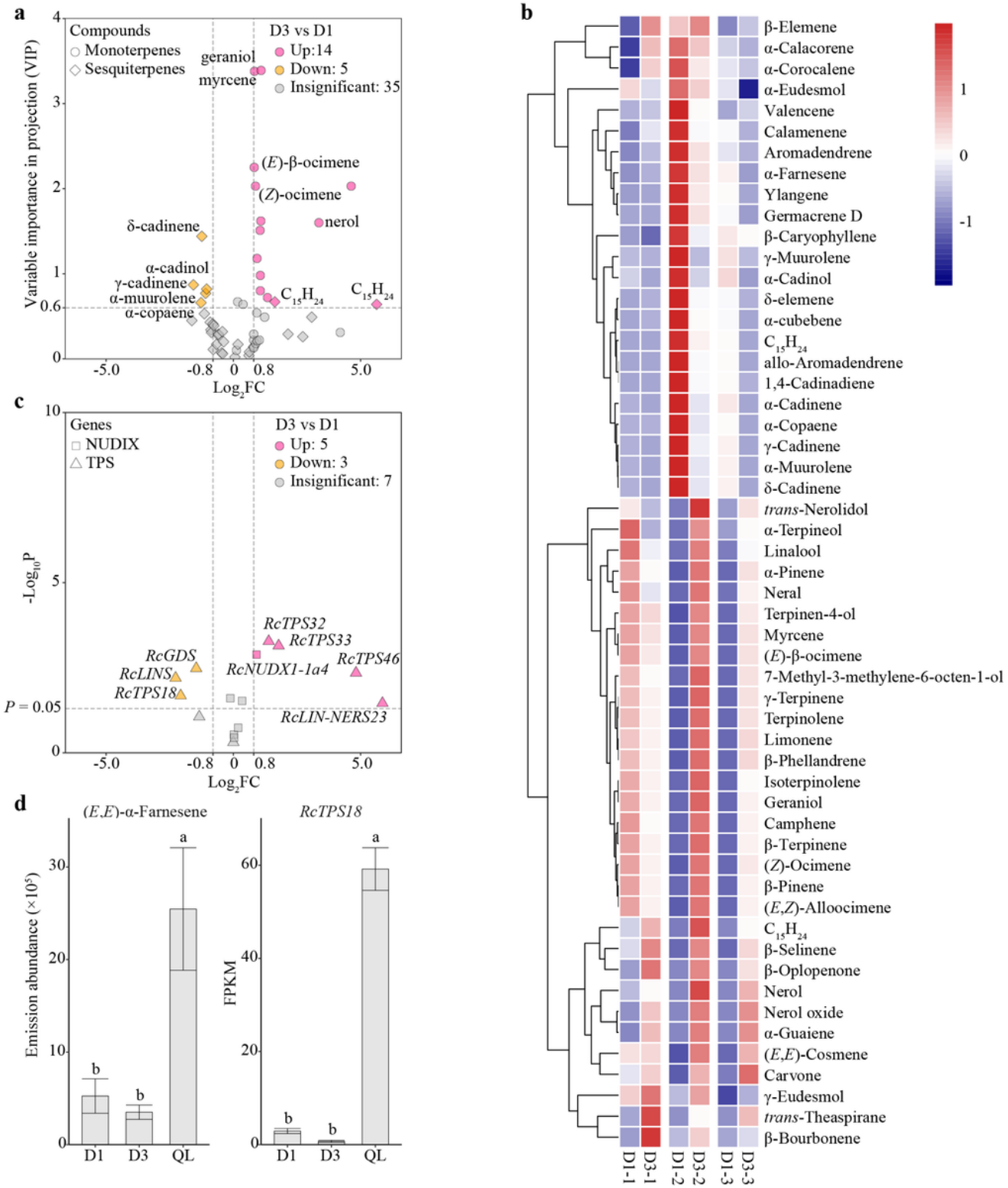

Figure 6

Differential metabolites and terpene-related genes in rose petals. (a) Volcano plot of differential metabolites between D3 and D1 samples. (b) Heatmap of terpene volatiles in D3 and D1 samples. (c) Volcano plot of differential terpene-related genes between D3 and D1 samples. (d) Emission abundance of (*E*,*E*)-α-farnesene and *RcTPS18* expression levels in three rose samples. QL: *Rosa* 'Qinglian Xueshi'.

Data are presented as the mean ± standard error (n = 3). Different lowercase letters indicate statistically significant differences (P < 0.05).

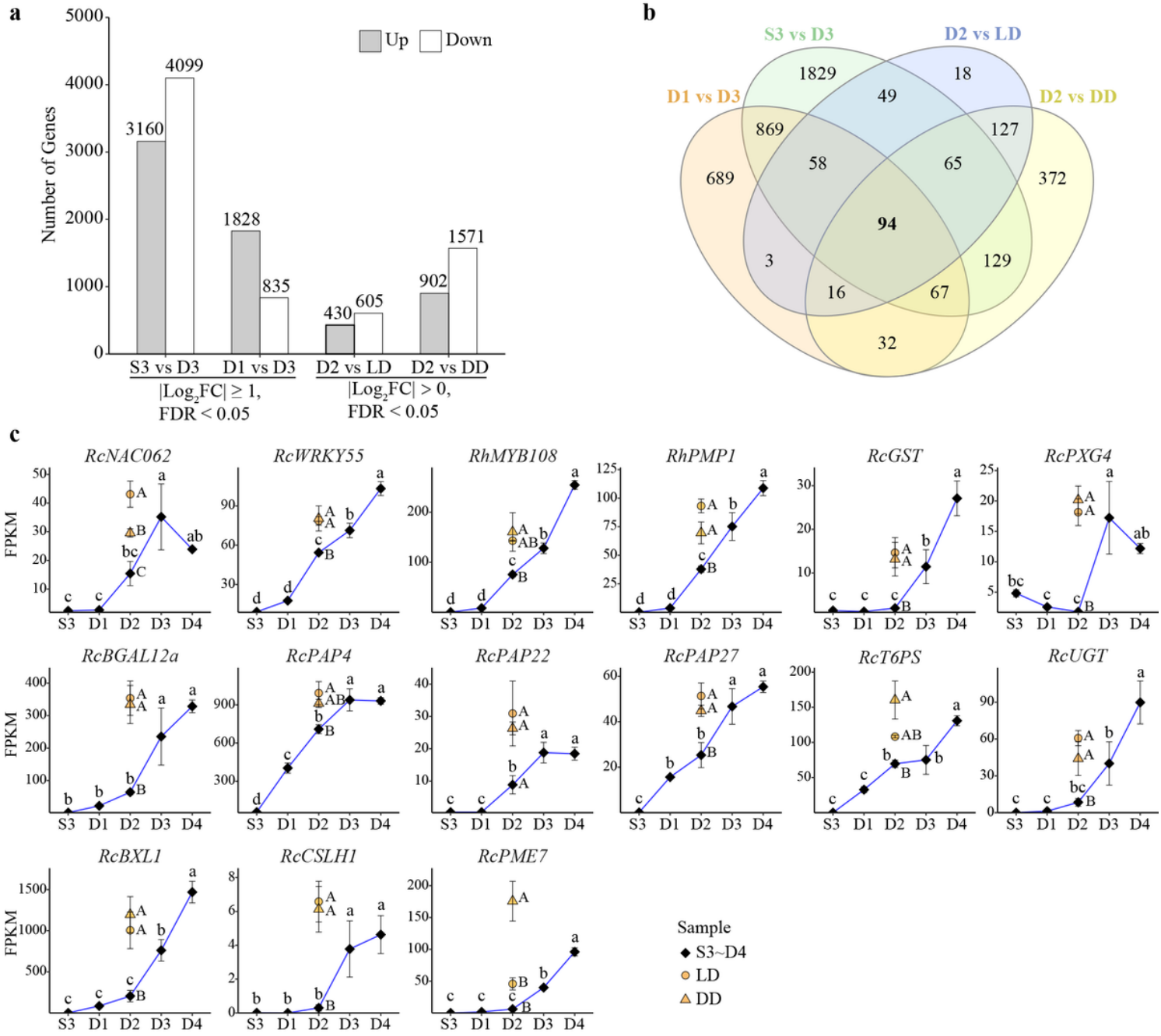

Figure 7

Gene expression comparisons among different samples. (a) Gene expression profile of DEGs. (b) Venn diagram of DEG numbers. (c) Expression analysis of some genes exhibiting similar expression patterns with those of *RcTPS32*, *RcTPS33*, and *RcTPS46*. Bars represent the standard error (n = 3). Different lowercase letters indicate statistically significant differences between samples from different developmental stages (P < 0.05). Different uppercase letters indicate statistically significant differences among D2, LD, and DD samples (P < 0.05). The genome ids of genes mentioned in this figure are as follows: *RcNAC062* (Chr5g0009331), *RcWRKY55* (Chr7g0241021), *RhMYB108* (Chr4g0426181), *RhPMP1*

(Chr1g0382431), *RcGST* (Chr7g0215041), *RcPXG4* (Chr4g0443291), *RcBGAL12a* (Chr2g0093301), *RcBGAL12b* (Chr2g0093311), *RcPAP4* (Chr4g0442051), *RcPAP22* (Chr3g0472921), *RcPAP27* (Chr1g0319551), *RcUGT* (Chr7g0212611), *RhAPC3b* (Chr2g0091631), *RhLAC* (Chr1g0384031), *RhSAG12* (Chr5g0010401), *RhXTH6* (Chr2g0109241).

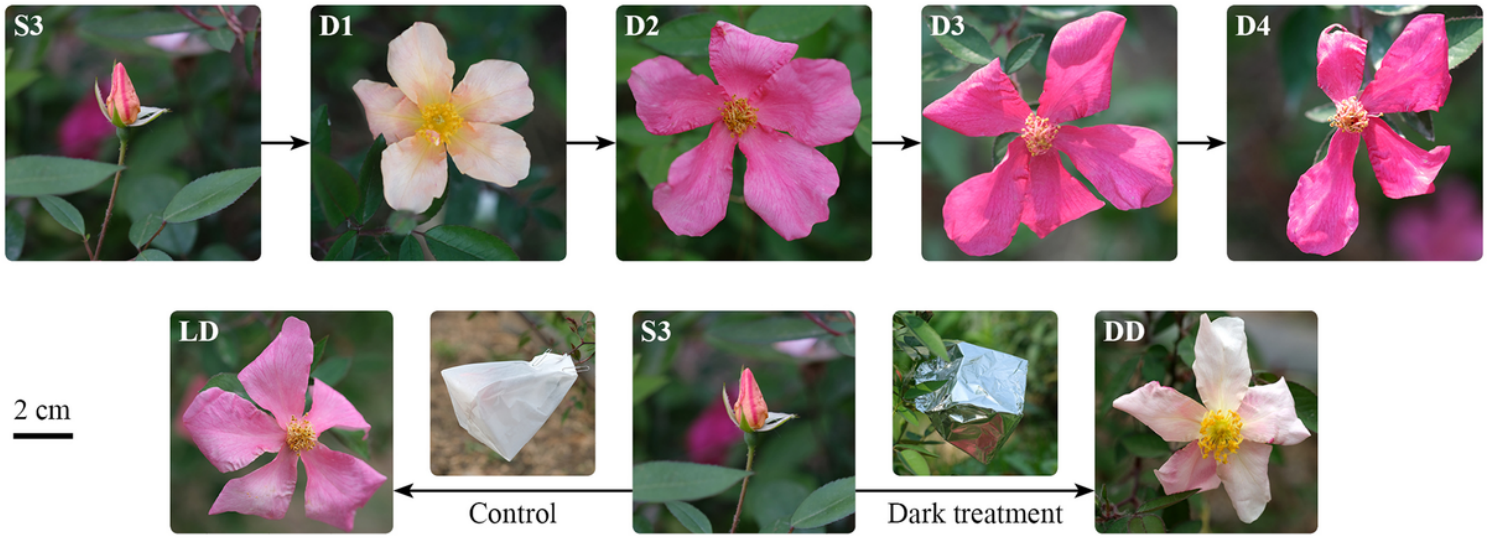

Figure 8

Petal samples of butterfly rose (*R. chinensis* 'Mutabilis') used in the present study. S3: bud about to open; D1: first day of anthesis; D2: second day of anthesis; D3: third day of anthesis; D4: fourth day of anthesis; LD: bagging with semi-transparent paper under light-dark conditions; DD: bagging with aluminum foil under continuous dark treatment.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile1.docx
- Additionalfile2.xlsx