Ornamental
Plant Research

REVIEW

Open Access

https://doi.org/10.48130/opr-0024-0001
Ornamental Plant Research **2024**, 4: e003

Progress in phylogenetics, multi-omics and flower coloration studies in *Rhododendron*

Shuai Nie[1#*], Hai-Yao Ma[2#], Tian-Le Shi[2], Xue-Chan Tian[2], Yousry A. El-Kassaby[3], Ilga Porth[4], Fu-Sheng Yang[5,6] and Jian-Feng Mao[2,7]

[1] *Rice Research Institute, Guangdong Academy of Agricultural Sciences & Key Laboratory of Genetics and Breeding of High-Quality Rice in Southern China (Co-construction by Ministry and Province), Ministry of Agriculture and Rural Affairs & Guangdong Key Laboratory of New Technology in Rice Breeding, Guangzhou 510640, China*
[2] *National Engineering Research Center of Tree Breeding and Ecological Restoration, Beijing Advanced Innovation Center for Tree Breeding by Molecular Design, National Engineering Laboratory for Tree Breeding, Key Laboratory of Genetics and Breeding in Forest Trees and Ornamental Plants, Ministry of Education, College of Biological Sciences and Technology, Beijing Forestry University, Beijing 100083, China*
[3] *Department of Forest and Conservation Sciences, Faculty of Forestry, University of British Columbia, Vancouver, BC V6T 1Z4, Canada*
[4] *Département des Sciences du Bois et de la Forêt, Faculté de Foresterie, de Géographie et Géomatique, Université Laval, Québec, QC G1V 0A6, Canada*
[5] *State Key Laboratory of Plant Diversity and Specialty Crops, Institute of Botany, the Chinese Academy of Sciences, Beijing 100093, China*
[6] *University of Chinese Academy of Sciences, Beijing 100049, China*
[7] *Department of Plant Physiology, Umeå Plant Science Centre (UPSC), Umeå University, Umeå 90187, Sweden*
[#] These authors contributed equally: Shuai Nie, Hai-Yao Ma
[*] Corresponding author, E-mail: nieshuai@gdaas.cn

Abstract

The genus *Rhododendron* exhibits an immense diversity of flower colors and represents one of the largest groups of woody plants, which is of great importance for ornamental plant research. This review summarizes recent progress in deciphering the genetic basis for flower coloration in *Rhododendron*. We describe advances in phylogenetic reconstruction and genome sequencing of *Rhododendron* species. The metabolic pathways of flower color are outlined, focusing on key structural and regulatory genes involved in pigment synthesis. Gene duplications and losses associated with color diversification are discussed. In addition, the application of multi-omics approaches and analysis of gene co-expression networks to elucidate complex gene regulatory mechanisms is emphasized. This synthesis of current knowledge provides a foundation for future research on the evolution of flower color diversity within the *Rhododendron* lineage. Ultimately, these discoveries will support breeding endeavors aimed at harnessing the genetics of flower coloration and developing novel cultivars that exhibit desired floral traits.

Citation: Nie S, Ma HY, Shi TL, Tian XC, El-Kassaby YA, et al. 2024. Progress in phylogenetics, multi-omics and flower coloration studies in *Rhododendron*. *Ornamental Plant Research* 4: e003 https://doi.org/10.48130/opr-0024-0001

Species diversification in *Rhododendron*

Phylogeny

Accurate species delimitations and robust phylogenies are crucial for botanists, ecologists, and horticulturists alike to study and utilize plant resources. The genus *Rhododendron* (Ericaceae) is particularly appreciated in horticulture for its magnificent flowers[1–5]. At the same time, with about 1,000 species, it is the largest genus of woody plants in the northern hemisphere. *Rhododendron* (L.) was first described by the Swedish botanist Carl Linnaeus (Carl von Linné, 1707–1778) in 1753 with only five species. In addition, Linnaeus described the closely related genera *Azalea* and *Rhodora* with six and one species, respectively. Subsequently, a larger number of *Rhododendron*, *Azalea* and *Rhodora* species were described, including those from the temperate zones of the northern hemisphere. The first tropical species *R. malayanum* was described in 1822, then the German botanist Carl Ludwig Blume (1796–1862) established the genus *Vireya* with four species from Java and Celebes. However, the related genera *Azalea*, *Vireya*, and *Rhodora* were combined into the genus *Rhododendron*[5]. The re-defined genus *Rhododendron* was divided into eight sections by the Baltic-German botanist Karl Maximovich (1827–1891) in 1870, which had a significant impact on later classification systems, albeit with different names or categories[3].

In the late 19th and early 20th centuries, the extensive collections of *Rhododendron* in southwest China led to the discovery of numerous species and the establishment of many series. These achievements were compiled in a handbook entitled '*The species of Rhododendron*'[6], which included 43 series and about 850 species of temperate and subtropical zones. However, no formal classification system was established. It was not until 1949 that Sleumer proposed a comprehensive taxonomic system for *Rhododendron*, comprising eight subgenera and 13 sections, and presented a monograph covering all *Rhododendron* species[5], especially those from tropical Indochina, Malesia and Australia. Sleumer's system was later revised by Chamberlain et al., the main difference being the treatment of the subgenus *Therorhodion*[1], which had been removed from *Rhododendron* by Sleumer[5]. Chamberlain's treatment was largely confirmed by further phylogenetic analysis based on morphological characters and molecular data[2], and the molecular evidence supported the position of subgenus *Therorhodion* within the genus *Rhododendron*. Finally, morphologically

based classifications of *Rhododendron* have led to a consensus taxonomic framework that recognized the five major subgenera *Azaleastrum*, *Hymenanthes*, *Pentanthera*, *Rhododendron*, *Tsutsusi*, and the three minor ones *Candidastrum*, *Mumeazalea*, and *Therorhodion*[1].

Recently, the traditional classification systems have been challenged by new molecular findings. Based on molecular and morphological data, Goetsch et al. further revised Chamberlain's system, whereby the subgenera *Candidastrum*, *Mumeazalea*, and *Tsutsusi* were merged into subgenus *Azaleastrum* and the polyphyletic subgenus *Pentanthera* was dissolved by transferring the sections *Pentanthera* and *Rhodora* to the subgenus *Hymenanthes* and the sections *Sciadorhodion* and *Viscidula* to the subgenus *Azaleastrum*[1,2].

More recently, Xia et al. reconstructed a molecular phylogeny of *Rhododendron* comprising 200 species that represent all subgenera, sections, and nearly all multi-species subsections, using 3,437 orthologous nuclear genes generated from transcriptome data[4]. They provided the first robust and dated backbone phylogeny based on genome-level data across the genus. Xia et al. recognized five subgenera and 11 sections and resolved the relationships of most taxonomically problematic groups of *Rhododendron* and its close relatives[4] (Fig. 1). For example, subgenus *Choniastrum* was treated as a section of the subgenus *Tsutsusi*, and the genus *Menziesia* was merged into section *Sciadorhodion* under the subgenus *Tsutsusi*. They recognized the section *Schistanche*, which comprises 312 species, over 95% of which are endemic to the Malay Archipelago. A comparison with the recently published plastome phylogeny for *Rhododendron* revealed conflicts between nuclear and plastid phylogenetic topologies, suggesting that reticulation events may have occurred in the deep lineages of the genus[7].

Xia et al. have shown that phylo-transcriptomics is an extremely efficient approach for reconstructing phylogenies, even for plant groups with rapid radiations such as *Rhododendron*[4] (Fig. 1). In the updated phylogeny, subgenus *Therorhodion*, previously considered a distinct genus, is the earliest diverged branch and sister to all other *Rhododendron*. This group includes two species distributed in the boreal tundra of East Asia and North America. Then there is the subgenus *Tsutsusi*, a monophyletic clade consisting of the sections *Sciadorhodion*, *Choniastrum*, *Azaleastrum*, *Brachycalyx* and *Tsutsusi* and comprising around 130 species, which are mainly distributed in East Asia. The sections *Choniastrum* and *Azaleastrum* were previously recognized as separate subgenera based on morphological and/or molecular evidence[2]. In section *Sciadorhodion*, *R. albiflorum* from western North America is a sister species to species from northeast Asia[4]. Species of section *Tsutsusi* are important genetic resources for evergreen azaleas, which are popular as ornamental shrubs or pot azaleas. Some traditional Chinese and Japanese varieties have been cultivated for more than 400 years. Among the 50 *Tsutsusi* species native to Japan, some evergreen species such as *R. kaempferi*, *R. macrosepalum*, *R. indicum*, and *R. ripense* are highly decorative, and there are

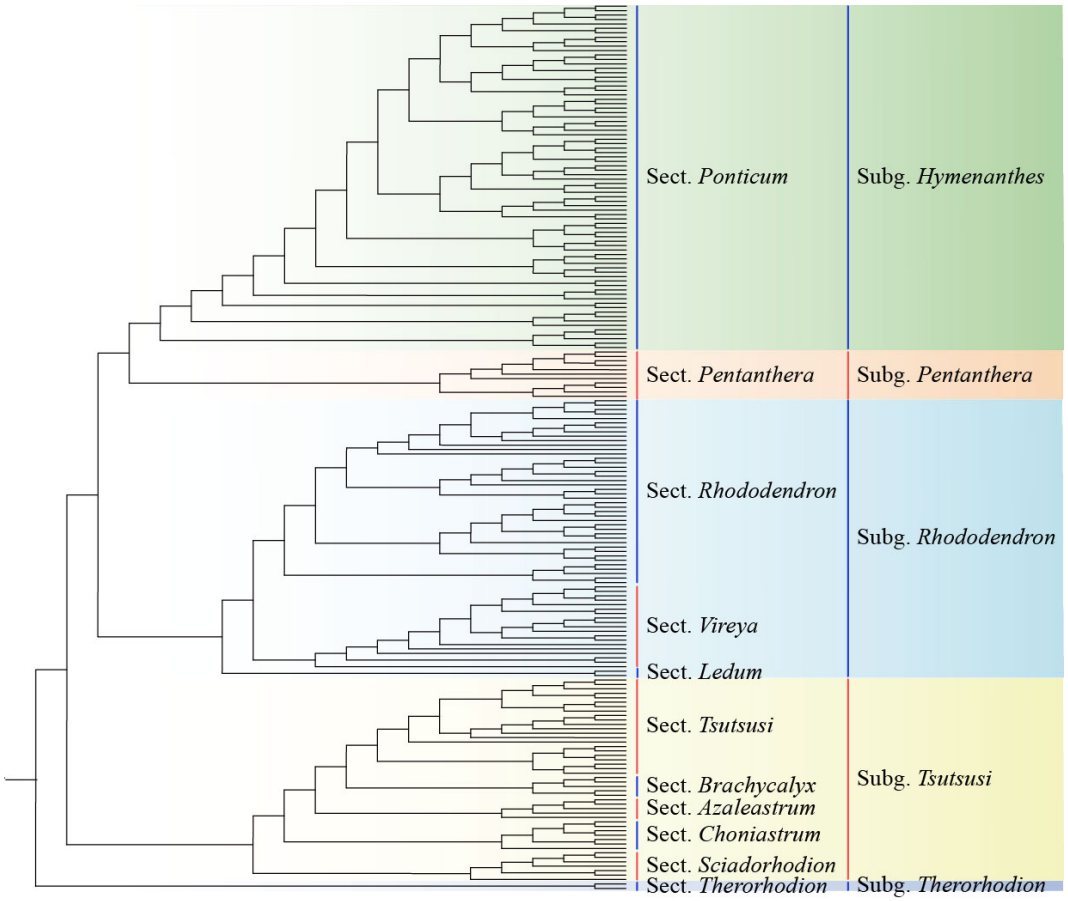

Fig. 1 Phylogenetic reconstruction in *Rhododendron* L. Sect., Section; Subg., Subgenus. (Modified from a previous publication[4]).

hundreds of azalea cultivars selected from natural populations and their hybrids[8,9]. The various molecular data is enabling the potential to unravel the complex phylogeny of the genus *Rhododendron* based on high-throughput sequencing. In a recent study, around 800,000 high-quality SNPs were obtained with restriction-site associated DNA sequencing (RAD-seq)[10]. Based on these large sets of SNPs, Shen et al. decoded the taxonomic relationships within the genus *Rhododendron*, reconstructed ancestral states of *Rhododendron*, and showcased the feasibility of applying RAD-seq to unravel intricate evolutionary relationships despite a significant amount of missing data[10]. With the further development of high-throughput sequencing, it is believed that deep whole-genome DNA sequencing will bring about many new discoveries in the future.

About 200 years ago, Belgian pot azaleas with a relatively narrow genetic base were developed from collector material from China and/or Japan, with several forms of *R. simsii* serving as potential parents, and *R. indicum*, *R. scabrum* and *R. mucronatum* may also have contributed to the modern *R. simsii* hybrids[8,11]. Azalea cultivars in China are categorized into four main types: (1) Eastern azaleas, which have small leaves and mostly originate from Japan; (2) Western azaleas (Belgian azaleas), including plants with large flowers and double and semi-double petals returning to China from Europe; (3) Hairy azaleas, which are characterized by large simple flowers and finely pubescent stems and leaves and are cultivated in China; and (4) Summer azaleas, which originate from Japan with cultivars blooming in early summer. However, the delimitation of these types is not supported by molecular evidence, and different cultivars may have the same name or a single plant may have multiple names[12].

There are over 500 species in the largest subgenus *Rhododendron*. These species are mainly distributed in the Himalayan-Hengduan Mountains (HHM) and the Malay Archipelago and can be identified by the scales on the back of their leaves. *R. hypoleucum* and *R. tomentosum* from circumboreal regions form a clade that is the first to diverge[4], and the sister clade comprises two reciprocal lineages: section *Schistanthe*, which occurs mainly in the Malay Archipelago, and section *Rhododendron*, which includes the Asian species (mainly in the HHM) and two European species (*R. ferrugineum* and *R. hirsutum*). Subgenus *Pentanthera* was previously considered as a section of subgenus *Hymenanthes*[2], but Xia et al. showed that it is related to subgenus *Hymenanthes*[4] and includes more than 20 deciduous species disjunctly distributed between North America, East Asia, and Western Europe. The subgenus *Hymenanthes*, on the other hand, comprises about 300 evergreen species, which are mainly distributed in the Hengduan Mountains.

Analyses of whole-genome resequencing data revealed that the relationships among species within the subgenus *Hymenanthes* were largely inconsistent with phylogenies inferred from different methods or datasets, especially for species from southwestern China, suggesting extensive hybridization and genomic admixture during the historical radiation of the subgenus, and repeated isolation and hybridization are thought to have promoted rapid diversification of species in mountainous regions[13]. In addition, most taxonomic treatments based on morphological characters have been rejected in this subgenus, and some subsections, such as the subsections *Fortunea* and *Argyrophylla*, have not been supported as monophyly[13].

Flower color evolution

Flower colors have been well studied by early evolutionary biologists in ecological and traditional genetic contexts, and it has been suggested that the attractive color display is mainly determined by three main groups of pigments, namely anthocyanins, carotenoids and betalains[14–17]. Rhododendrons are valued for their great ornamental beauty and a wide range of flower colors: red, purplish pink, purple, yellow, white, and various intermediate colors. Until now, the mechanism underlying flower coloration in *Rhododendron* has rarely been studied, with the main focus on cultivated azaleas. Co-pigmentation between anthocyanins and flavonols has been found to be responsible for petal coloration in species and cultivars of the subgenus *Tsutsutsi*, with color intensity being influenced by the amount of anthocyanins, and increasing amounts of flavonols leading to a color change from red to purple[18]. In many sections of *Rhododendron*, the carotenoids are not involved in the coloration of the petals but appear in the form of spots and blotches[18].

Flower color is often correlated with other floral traits, leading to the common use of the term 'pollination syndromes' to refer to adaption to pollinator regimes[19]. Recently, Du et al. studied the flower colors of 30 representative *Rhododendron* species from seven subgenera of the genus and identified seven anthocyanins (cyanidin 3-O-galactoside, cyanidin 3-O-arabinoside, cyanidin 3,5-di-O-glucoside, cyanidin 3-O-arabinoside-5-O-glucoside, delphinidin 3-O-glucoside, delphinidin 3-O-arabinoside-5-O-glucoside, malvidin 3-O-arabinoside-5-O-glucoside) were responsible for the different colors of *Rhododendron* flowers, leading to the conclusion that cyanidin-type anthocyanins were responsible for the red coloration and delphinidin- and malvidin-type anthocyanins for the purple coloration[20]. However, they also identified 21 flavonol derivatives that may not play a major role in flower coloration[20]. Their results improve our understanding of flower coloration, genetic variation and environmental adaptation in *Rhododendron*. It seems that *Rhododendron* has comparatively uniform pigment types but different amounts and components in different classification levels[20], which is supported by morphological and molecular evidence. Xia et al. found that red, purplish pink, purple and yellow flowers occurred repeatedly in the major clades and subclades of the *Rhododendron* phylogeny, indicating widespread parallel evolution of flower colors in this genus[4]. Variations in flower colors among closely related species are often associated with rapid shifts in specific pollinators[21–23], highlighting the rapid adaptation of plants to pollinators and species diversity in the species-rich genus *Rhododendron* driven by biotic factors.

The genomics of flower pigmentation regulation in *Rhododendron*

Progress in the sequencing of entire genomes for the genus

Over the past 20 years, the sequencing of plant genomes has increased rapidly, leading to an increase in the quantity and quality of publicly available genomic resources[24]. Genomic information on *Rhododendron*, one of the largest woody plant genera, has been acquired and disseminated by several research institutions. Since the publication of the draft genome

assembly of *R. delavayi* in 2017, a total of 14 genomes from 12 species have been sequenced[13,25–34]. Supplemental Table S1 shows that most genome assemblies (all but one) have reached the chromosome level, with three assemblies approaching a near-complete T2T level. Recently, the *R. vialii* genome assemblies have reached the haplotype level[35]. The improvement of assembly quality, especially in terms of completeness, is in line with the development of sequencing technologies. PacBio HiFi sequencing technology generates long-read sequencing datasets with average read lengths of 10–25 kb and an accuracy of over 99.5%[36,37]. Combined with Hi-C sequencing technology and newly developed assembly techniques, it will become easier to obtain more high-quality *Rhododendron* genome assemblies. This undoubtedly represents a great opportunity for *Rhododendron* species with high genome heterozygosity.

Among the five subgenera of *Rhododendron*, *Hymenanthes* has the most sequenced species (eight published genomes from seven species), followed by *Tsutsusi* (four genomes), and *Pentanthera* with one species (two assemblies of *R. molle*) (Supplemental Table S1). No genome assemblies were reported for the other two subgenera, *Therorhodion* and the larger subgenus *Rhododendron* (Supplemental Table S1). Moreover, only one of the sequenced species is deciduous (*R. molle*), while the remaining 11 species are evergreen Rhododendrons (Supplemental Table 1). Clearly, more importance should be given to larger and more evenly sampled genome projects. *Rhododendron* plants exhibit a wide range of flower colors, broadly classified into red, white, purple, and yellow color series[13,16,18,20,38,39]. The sequenced genomes covered four color series and provided the opportunity to investigate the genetic basis of flower color transitions at the genomic level (Supplemental Table S1). To fully utilize these resources, a website was developed to integrate and make these data accessible. Here we have compiled download links for different data types to facilitate access for researchers (Table 1). In addition, two *Rhododendron*-specific databases[40] are accessible: RPGD (http://bioinfor.kib.ac.cn/RPGD/) and http://rhododendron.plantgenie.org/.

Table 1. Data resources of 12 genomic projects in *Rhododendron*.

Species	Data resources
R. prattii	https://ngdc.cncb.ac.cn/gwh/Assembly/24363/show
R. molle v1	www.ncbi.nlm.nih.gov/datasets/genome/GCA_030770705.1
R. molle v2	https://ngdc.cncb.ac.cn/gwh/Assembly/29302/show
R. henanense	https://ngdc.cncb.ac.cn/gwh/Assembly/22219/show
R. ripense	https://plantgarden.jp/en/list/t224351/genome
R. ovatum	http://bioinfor.kib.ac.cn/RPGD/download_genome.html
R. griersonianum	www.ncbi.nlm.nih.gov/datasets/genome/GCA_018127125.1
R. simsii	www.ncbi.nlm.nih.gov/datasets/genome/GCA_014282245.1
R. williamsianum	http://bioinfor.kib.ac.cn/RPGD/download_genome.html
R. delavayi v1	http://bioinfor.kib.ac.cn/RPGD/download_genome.html
R. delavayi v2	http://bioinfor.kib.ac.cn/RPGD/download_genome.html
R. vialii	https://ngdc.cncb.ac.cn/gwh/Assembly/37538/show
R. irroratum	http://bioinfor.kib.ac.cn/RPGD/download_genome.html

Genome sizes and divergence

Genome size is commonly referred to as the amount of DNA present in the cell nucleus[41]. Genome size in land plants is influenced by several factors, such as whole genome duplication, polyploidization, and amplification and elimination of repetitive sequences[42–44]. All sequenced Rhododendrons are diploid with 13 chromosomes in the reported genome assemblies. The variation in genome size is not significant and ranges from ~500 Mb in *R. ripense* to ~700 Mb in *R. irroratum* (Supplemental Table S1). Comparative genomic analyses revealed that the increase in genome size of *R. molle* (653.46 Mb) compared to *R. simsii* (528.64 Mb) was due to the recent proliferation of long terminal repeat retrotransposons (LTR-RTs), particularly *Gypsy*, resulting in an increase of 125 Mb (19%)[26,31]. A positive correlation was observed between the length of repetitive sequences and genome size, suggesting that repetitive sequences may influence genome size variation in *Rhododendron* (Supplemental Table S1).

Rhododendron species show differentiation in inter-genomic sequences but are still conserved to a certain extent. With *R. simsii* genome assembly as a reference, a comparative analysis of the whole genome sequences revealed that 60% of the *R. molle* genome sequence exhibits inter-genomic collinearity, while 36% shows sequence differentiation[26,31]. LTR-RTs are considered to be the main driver of interspecific sequence variation. In particular, *Gypsy* elements could play an important role in the centromeric regions of chromosomes and contribute to the rapid evolution of interspecific centromeric sequences[26,31]. In the reported *Rhododendron* genome assemblies, LTR-RTs were identified as the predominant type of repetitive sequences, with *Gypsy* being the most abundant LTR-RTs element. It is worthwhile to investigate the evolutionary patterns of *Gypsy*-like LTR-RTs in *Rhododendron* and their effects on genome size and sequence variation.

Whole genome duplication (WGD) has been another focus in *Rhododendron*, thought to be another force contributing to changes in genome size and sequence. It is generally agreed that the genus *Rhododendron* has experienced two common WGD events: the first is the ancient whole genome triplication event (gamma) common to all core eudicots, while the second is the WGD event associated with the origin of the entire Ericales order. The occurrence of the most recent WGD is the subject of debate, with three main perspectives being advocated: (1) the Ericales[26,45–48]; (2) the core Ericales[49–51]; and (3) the Ericaceae and Actinidiaceae[31,51]. The complexity of gene loss and retention patterns after WGD makes it difficult to determine the origin and phylogenetic position of Ericales. Numerous techniques have been used to study and identify polyploidy events. These include non-phylogenetic methods such as karyotyping and synteny[52,53] as well as phylogenetic methods such as Ks-based analysis, least common ancestor (LCA) reconciliation, count-based, and gene networks approaches[54–57]. Despite the application of these detection methods in various studies[46–50,58,59], the exact timing of the ancient WGD event in the order Ericales remains unclear.

A fundamental problem is the difficulty of separating WGD from Ericales differentiation. A recent work using a gene-tree reconciliation algorithm has revealed the allopolyploid origin of Ericales[60]. However, no potential progenitors were identified in this study and the sample size was limited. While the proposed view is novel, ongoing genome sequencing projects in *Rhododendron* and its close relatives will eventually provide clarity on the evolutionary history of this duplication event.

Genes underpinning flower color

Genome sequencing and gene annotation techniques have opened up the possibility of identifying and recognizing genes

for flower colors throughout the genome. And multi-omics have opened new perspectives for understanding structural and regulatory genes and their interaction, i.e., the molecular mechanisms underlying flower pigmentation[61,62]. Anthocyanins have been identified as important pigments influencing color variations in *Rhododendron* based on transcriptomics and metabolome assessment of floral tissues at different developmental stages[39,63–66]. Other molecules such as flavonoids and carotenoids are also expected to play a role in pigment production[67,68]. Nie et al. showed that the yellow flower of *R. molle* can be attributed to (i) the synthesis of carotenoids and flavonols, (ii) the degradation of chlorophyll, and (iii) the complete absence of anthocyanin synthesis[26]. The metabolic pathways for carotenoids/chlorophyll/flavonol/anthocyanin were reconstructed in *R. simsii* and *R. molle* using whole-genome analysis (Figs 2 & 3)[26,31]. In these reconstructed metabolic pathways, a total of 15 enzymes related to anthocyanin synthesis were identified, including 11 crucial enzymes (*PAL, C4H, 4CL, CHS, CHI, F3H, F3'H, F3'5'H, DFR, ANS*, and *F3oGT*) involved in anthocyanin glycoside synthesis.

Transcription factors (TFs), which influence flower coloration by regulating the expression of pigment-encoding genes, are considered another category of flower color genes. The involvement of transcription factor families such as MYB, bHLH (basic helix-loop-helix), WD40 (WD40-repeat-containing protein), ERF and WRKY in *Rhododendron* pigment synthesis have been widely reported[69,70]. Wang et al. identified 68 RsWRKY gene members in *R. simsii*, which can be divided into 3 main groups and 7 subgroups[70]. Wu et al. revealed the conserved features during the evolution of R2R3-MYB in *Rhododendron*[69]. And Nie et al. identified the members and investigated the evolutionary patterns of five TF families in *R. simsii* and *R. molle*, respectively[26]. These studies have reconstructed the metabolic pathways and identified key regulatory genes, providing insights into the intricate processes that control the development of flower color in *Rhododendron*. These processes can potentially be used in various fields, including breeding programs to improve floral traits, developing novel cultivars with desired color patterns, and even exploring innovative biotechnological approaches to pigment manipulation in *Rhododendron* and other woody plants.

Duplication and loss of genes related to flower color

At the whole-genome level, numerous gene duplication events were detected in both *R. simsii* and *R. molle* (Fig. 4a), and DSD (dispersed gene duplication) being the most frequent and possibly linked to the expansion of LTR-RTs, leading to differences in gene number between these two *Rhododendron* species[26]. Among flower color-related genes, TD (tandem gene duplication) and PD (proximal gene duplication) were identified as the predominant duplication types (Table 2). The gene families associated with anthocyanin synthesis were larger than those involved in carotenoid and chlorophyll synthesis (Fig. 4b). Further analysis suggests that gene duplication, particularly TD and PD, may have facilitated the expansion of these pigment-related gene families. In *R. simsii* and *R. molle*, TD and PD accounted for 43.16% and 44.14% of anthocyanin-related genes, respectively. In contrast, these two duplication types accounted for only a small percentage of chlorophyll-related genes (22.22% and 26.56%), and an even smaller percentage of carotenoid-related genes (17.24% and 23.07%) (Table 2). It can be concluded that the expansion of anthocyanin-related gene families, especially TD and PD, in Rhododendrons is responsible for the production of the major pigments. Gene duplication contributes in different ways to the expansion of gene families for flower color. For the R2R3-MYB

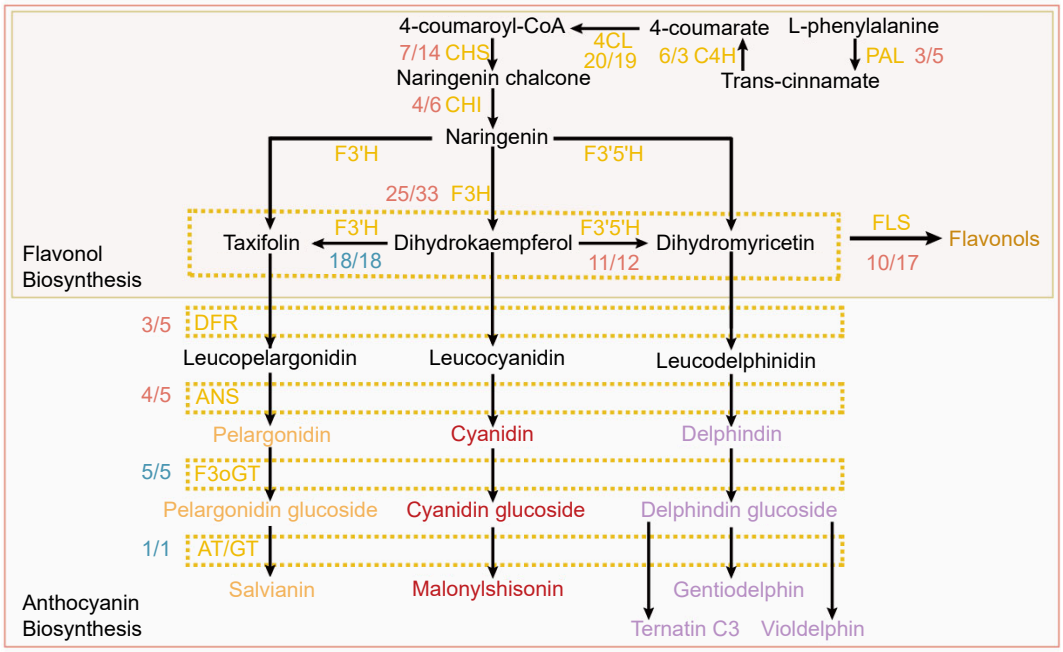

Fig. 2 Metabolic pathways of anthocyanins and flavonols in *R. simsii* and *R. molle*. Fractions next to gene abbreviations indicate the ratio of the number of expressed genes for *R. molle* vs *R. simsii*; the numerator is the gene number for *R. molle* and the denominator is the gene number for *R. simsii*. Equal gene numbers are colored blue, excess ratio in the gene number for *R. molle* is colored yellow, while excess ratio in *R. simsii* colored red.

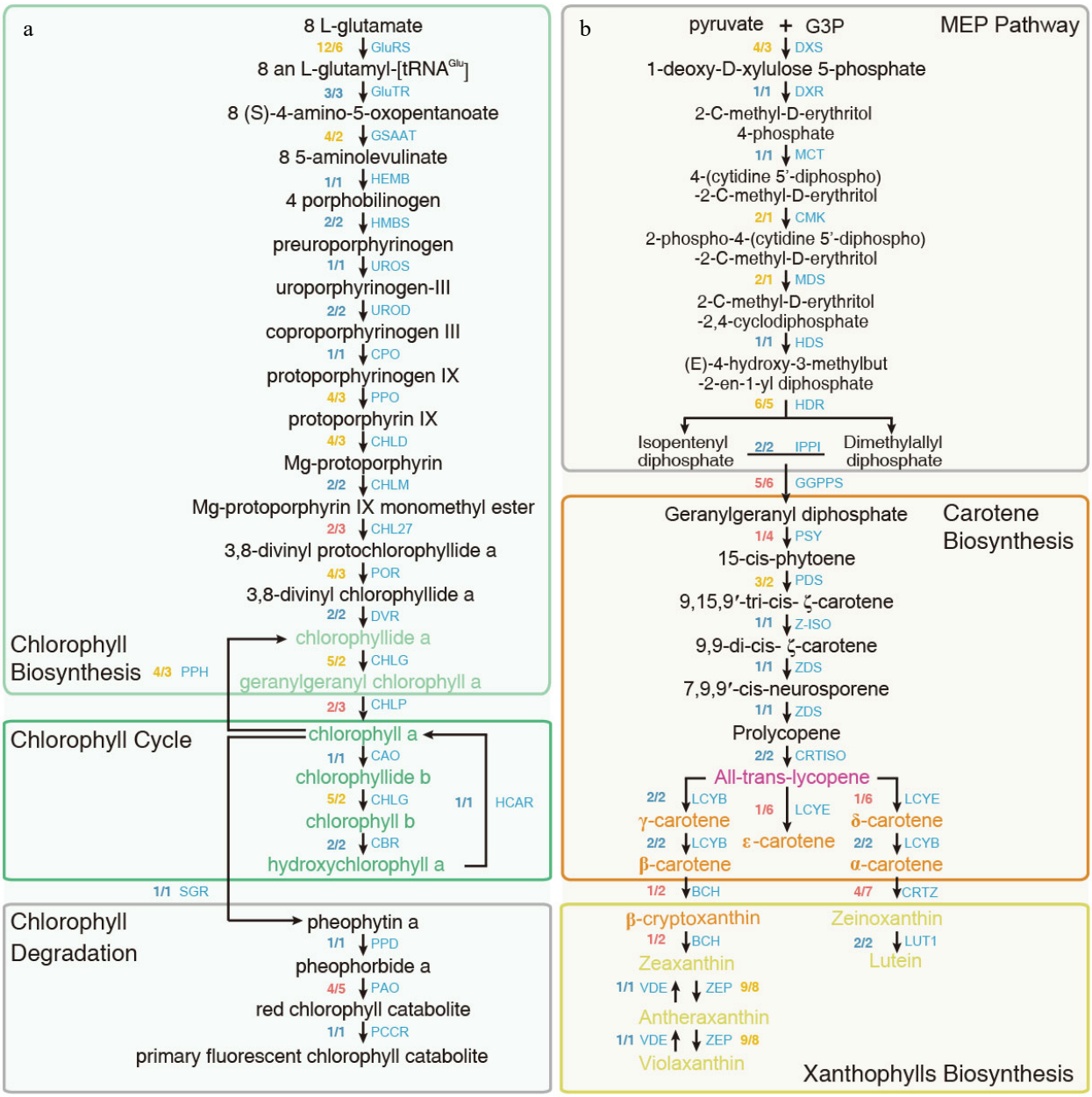

Fig. 3 Metabolic pathway of (a) chlorophylls and (b) carotenoids in *R. simsii* and *R. molle*. Fractions next to gene abbreviations indicate the ratio of the number of expressed genes for *R. molle* vs *R. simsii*; the numerator is the gene number for *R. molle* and the denominator is the gene number for *R. simsii*. Equal gene numbers are colored blue, excess ratio in the gene number for *R. molle* is colored yellow, while excess ratio in *R. simsii* colored red.

gene family, WGD contributed to more than 50% of the total number of genes in the three *Rhododendron* plants[69], and WGD also contributed to 47.37% (27 of the 57) for WRKY.

The contraction of gene families involved in anthocyanin synthesis is accompanied by pseudogenization in *R. simsii* and *R. molle*. Genes related to carotenoid and chlorophyll synthesis underwent relatively little pseudogenization, with only three to four pseudogenes identified in each *Rhododendron* species (Fig. 4c). In contrast, more pseudogenes were found in anthocyanin-related gene families, notably in *R. molle*, where 25 anthocyanin-related pseudogenes were discovered (Fig. 4c). Pseudogenization has led to a significant contraction of anthocyanin gene families in *R. molle*, with 13 contracted families accounting for 81.25% of all 16 families (Fig. 4d). In contrast, there were fewer contracted families and pseudogenes in the chlorophyll and carotenoid metabolic pathways (Fig. 4c & d).

Research to date on flower coloration genes is essentially limited to a handful of species, which is in stark contrast to the intricate interspecific color diversity observed in *Rhododendron*[66,71,72]. Interestingly, repetitive sequences could serve as key factors that facilitate the duplication and subsequent loss of color-determining genes. This unfolding perspective enriches our understanding of the evolution of both repetitive sequences and color diversity in *Rhododendron*. In future, deciphering the intricate evolutionary interplay between LTR-RTs, gene duplications and losses, and color

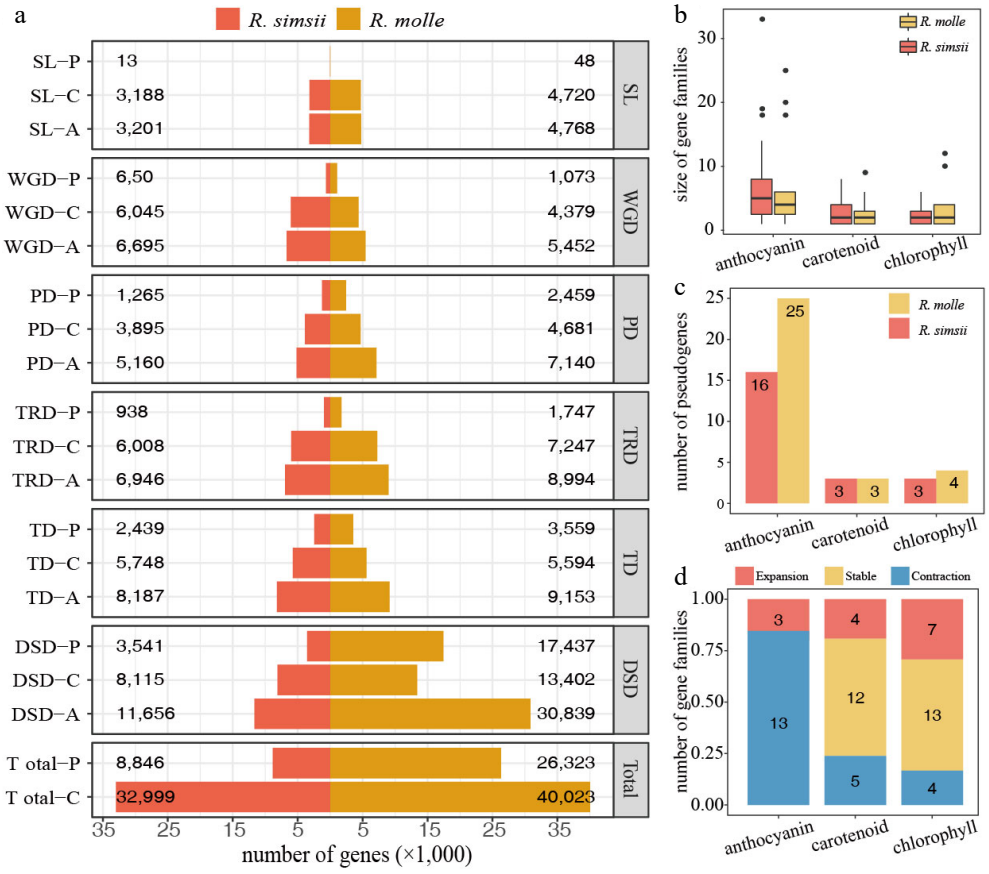

Fig. 4 Gene duplications and losses in *R. simsii* and *R. molle*. (a) Barplots show the counts between *R. simsii* and *R. molle* of coding genes and pseudogenes for the different duplication modes (WGD, TD, PD, TRD, DSD and SL: whole-genome, tandem, proximal, transposed and dispersed duplications, and singletons). Different suffixes on the left denote '-P' for Pseudogene, '-C' for Coding gene, and '-A' for all coding genes and pseudogenes. (b) Size distribution of enzyme coding gene families for the different pigment metabolic pathways. (c) Pseudogene numbers from the different metabolic pathways for anthocyanin, carotenoid and chlorophyll synthesis. We have linked a pseudogene to a specific enzyme, depending on which enzyme its parental coding gene encodes. (d) Barplots showing the counts of enzyme families from the different metabolic pathways for anthocyanin, carotenoid and chlorophyll synthesis. A fraction > 1 for enzyme encoding gene numbers (*R. molle*/*R. simsii*) defines as 'Expansion'; a fraction of 1 defines as 'Stable'; a fraction < 1 defines as 'Contraction'. (Modified from a previous publication[26]).

diversity in different *Rhododendron* species, aligned with phylogenetic frameworks, may become a focus of research.

Gene regulation of flower coloration

The metabolic pathways and regulatory networks that influence phenotypic traits such as color are central to our understanding of plant biology. The genus *Rhododendron*, one of the most prolific groups of woody plants, aroused great interest in the scientific community. This interest stems mainly from the great diversity of flower colors in wild species and cultivars[18,73]. So far, research on color differentiation in *Rhododendron* has primarily investigated the composition and concentration of pigments[18,20,63]. However, the underlying genetic and regulatory mechanisms that determine these vivid variations remain largely obscure.

In recent years, a number of structural genes responsible for the biosynthesis of flavonoids have been isolated from a variety of wild and cultivated *Rhododendron*. The spatiotemporal expression patterns of key regulatory genes were analyzed[13,39,68]. Pigment synthesis has been shown to be under the influence of specific transcription factors including MYB, bHLH, WD40, and their associated complexes, and members of these gene families have been identified and categorized at the whole-genome level in *Rhododendron* species such as *R. simsii*, *R. molle*, *R. delavayi*, *R. irroratum*, and *R. ovatum*[29,30,49,60,74]. In addition, the WRKY and ERF transcription factor families may also play an important role in establishing the temporal gene regulatory network that determines flower color.

The analysis of gene expression networks is an important tool that has been used extensively to study the nuances of flower coloration and differentiation. It is also important to keep in mind that flower color formation is a process that unfolds over time and therefore requires the study of gene expression and regulatory mechanisms at different stages of petal development. We have developed an analytical approach 'InterSpecificTOGCN' (https://github.com/JeffreyNIEgithub/InterSpecificTOGCN) to comprehensively decipher the patterns of the time-ordered gene co-expression networks (TO-GCNs). Here, the approach was applied to analyze the time-series transcriptomes of five timepoints (here T1 to T5) in the process of pigment formation under different Rhododendrons (red flowered *R. simsii* and yellow flowered *R. molle*) (Fig. 5a). We described the specific and consensus hierarchical TO-GCNs for two Rhododendrons and reconstructed seven levels (here

Table 2. Summary of genes in tandem (TD) or proximal (PD) gene clusters related to pigmentation metabolic pathway.

Pigment	Type	*R. simsii*	*R. molle* v2
Anthocyanin	ALL	125	104
Anthocyanin	PD	9 (7.20%)	15 (14.42%)
Anthocyanin	TD	46 (36.8%)	32 (30.76%)
Anthocyanin	TD/PD	55 (44.00%)	47 (45.19%)
Anthocyanin / flavonol	ALL	139	111
Anthocyanin / flavonol	PD	11 (7.91%)	17 (15.31%)
Anthocyanin / flavonol	TD	49 (35.25%)	32 (28.82%)
Anthocyanin / flavonol	TD/PD	60 (43.16%)	49 (44.14%)
Flavonol	ALL	112	90
Flavonol	PD	11 (9.82%)	14 (15.55%)
Flavonol	TD	42 (37.5%)	30 (33.33%)
Flavonol	TD/PD	53 (47.32%)	44 (48.88%)
Chlorophyll	ALL	54	64
Chlorophyll	PD	2 (3.70%)	7 (10.93%)
Chlorophyll	TD	10 (18.51%)	10 (15.62%)
Chlorophyll	TD/PD	12 (22.22%)	17 (26.56%)
Carotenoid	ALL	58	52
Carotenoid	PD	4 (6.89%)	4 (7.69%)
Carotenoid	TD	6 (10.34%)	8 (15.38%)
Carotenoid	TD/PD	10 (17.24%)	12 (23.07%)

ALL, all identified genes; TD, tandem duplicated genes; PD, proximal duplicated genes; TD/PD, tandem or proximal duplicated genes. (Modified from a previous publication[26]).

L1–L7) of TO-GCNs separately (Fig. 5b). Different levels were summarized into three major stages: initial (T1–T2; corresponding to L1–L3, flower appears white and green), transitional (T3; corresponding to L4, during green to yellow/red flower transition), and terminal (T4–T5; corresponding to L5–L7, for the determined yellow or red flower color) (Fig. 5c). The genes of carotenoid/flavonol biosynthesis related to yellow flower specificity formation were thought to be located in the terminal networks (L5 to L7) in *R. molle*-specific TO-GCNs, based on metabolome, gene function enrichment analysis and network complexity statistics for each stage[26,31]. We further identified hub genes of *SGR* (*Rhmol07G0115200*), *PSY* (*Rhmol10G0271900*), *PAL* (*Rhmol03G0128000*), and *F3'H*|*F3'5'H* (*Rhmol09G0029300*) families (Fig. 6a), which all represent top graph degrees of *R. molle*-specific TO-GCNs. By examining the network topology and TF binding site (TFBS) predictions, these four hub genes might be regulated with two TFs, *MYB_related* (*Rhmol 11G0261800*) and *B3* (*Rhmol04G0179100*), as common regulators in a hierarchical manner in *R. molle* (Fig. 6b & c)[26]. In *R. simsii*, two crucial structural genes, *F3oGT* (*Rhsim01G0008100*) and *F3'H*|*F3'5'H* (*Rhsim09G0023900*), and their potential regulators, *MYB_related* (*Rhsim03G0160200*) and *ERF* (*Rhsim03G 0176600*), were predicted within the specifically reconstructed hierarchical subnetworks related to anthocyanins with the same analysis principles (Fig. 6d)[26]. The genus *Rhododendron* displays a notable abundance of flower color intermediates, particularly in the case of red hues. It is of great importance to investigate the gene expression changes in the temporal development of these crucial flower color intermediates. Bioinformatics approaches such as TO-GCN have great potential to address these issues from the perspective of time-series gene co-expression network.

Furthermore, it is hypothesized that LTR-RTs may play an important role in the formation of interspecific gene regulatory patterns. Active LTR-RTs in *R. molle* could potentially induce flower color shifts by integrating in or near specific pigment biosynthesis genes and transcriptional regulatory factors, triggering changes in gene transcription and regulation[7]. The applicability and extent of this influence may be difficult to validate in the laboratory; however, their potential impact should not be ignored. As widespread and highly active components in the genome, LTR-RTs have great potential to promote the emergence of the enormous flower color diversity in *Rhododendron*.

In addition, an Agrobacterium-mediated genetic transformation system for *R. simsii* has been developed. However, the long growth cycle of woody plants has hindered the experimental validation of genes for flower color in *Rhododendron*. Therefore, the need to explore innovative methods to validate gene function is particularly evident in the study of *Rhododendron* color traits. This is not only essential for ongoing *Rhododendron*-specific research, but is also of great value for the study of reproductive traits in other woody species. The possible application of LTR-RTs as natural mutagens deserves special attention and research.

Flower color breeding in the age of multi-omics

As an ornamental horticultural plant, *Rhododendron* has a long history of artificial cultivation and has high visibility and a large industry chain worldwide[75]. There is a common saying in the Western gardening community: 'Without China's Rhododendrons, there would be no diversity in Western gardens'. *Rhododendron* breeding faces almost all the challenges of genetic breeding of woody plants, including long cycles, weak breeding foundations, and limited breeding methods[76]. The selection of flower color is of utmost importance in *Rhododendron* breeding. Wild Rhododendrons in nature exhibit a wide range of flower colors, while cultivated Rhododendrons are mostly red or white, with fewer yellow and blue varieties that are more valuable. Therefore, breeding yellow or blue Rhododendrons has become an important breeding goal[77].

From a global perspective, cultivation and breeding studies on Rhododendrons started earlier and more systematically in Japan, North America, and Europe[6,9]. China is considered the country with the richest resources of wild *Rhododendron* germplasm, with about 600 of the world's more than 1,200 species[78]. Collecting and studying germplasm resources is the first step towards the conservation and utilization *Rhododendron* resources, and previous research has made significant progress in this area[79–82]. A large number of new *Rhododendron* cultivars have been bred through traditional breeding methods such as selection and hybrid breeding, which will continue to be the main means of flower breeding in the future. However, advances in biotechnology, especially in molecular genetics and genetic engineering, offer new ideas and new opportunities for *Rhododendron* breeding and have great market potential.

The genus *Rhododendron* has undergone extensive interspecific hybridization within broad subgenera and a complex history of radiation evolution[4,13,25,83]. The improvement in sequence length and accuracy has made genome assembly cheaper, faster, and more accurate[84]. Functional genomics such as transcriptomics, metabolomics, proteomics, and phenomics have also become more accessible and comprehensive.

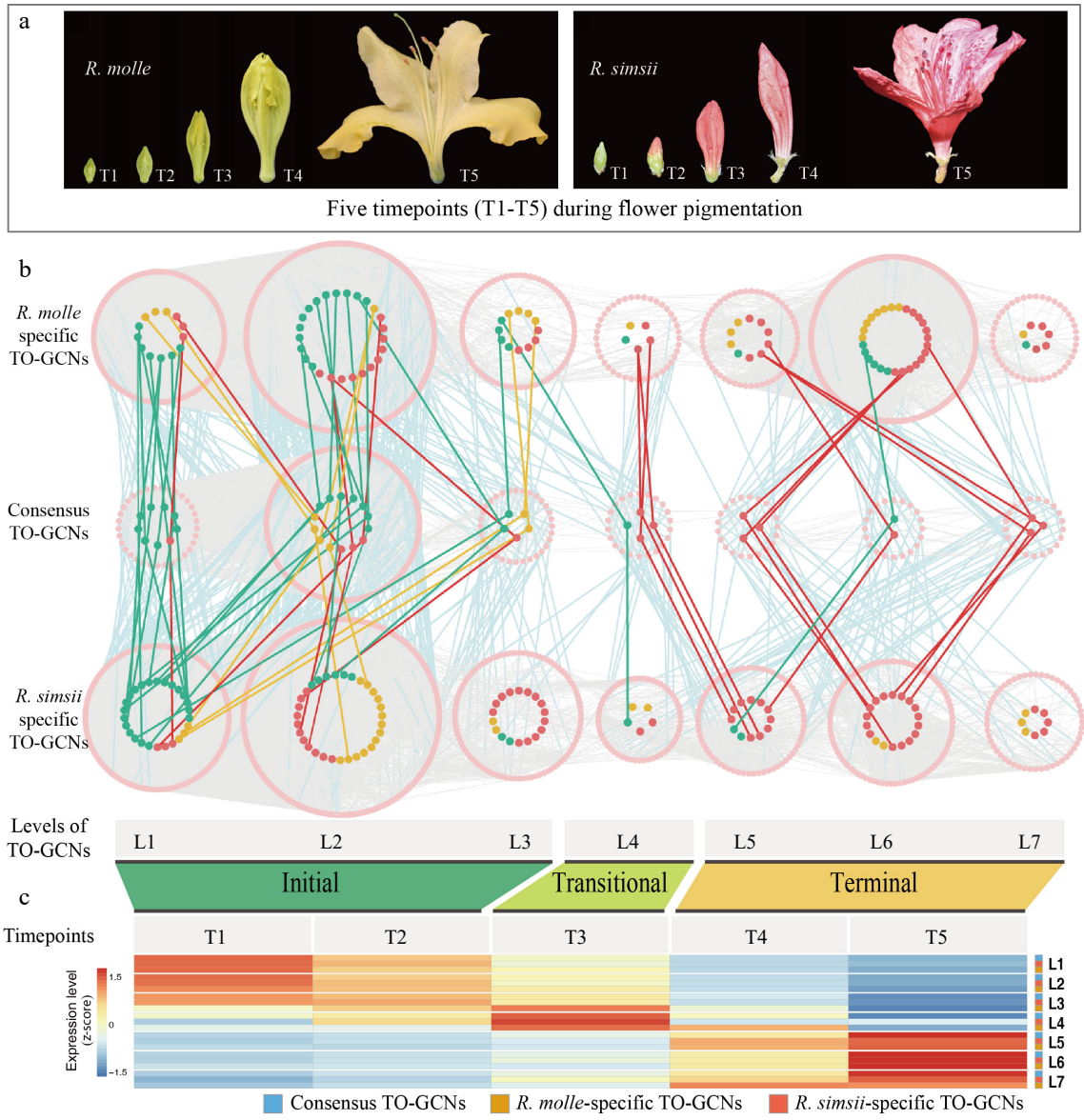

Fig. 5 Time-ordered gene co-expression networks during flower coloration of *R. molle* and *R. simsii*. (a) Five timepoints (T1–T5) during flower pigmentation. (b) Predicted regulatory network and the resolved gene regulation among transcription factors (TFs; pink points) and enzymatic genes involved in pathways of carotenoid (orange points), chlorophyll (green points) and anthocyanin/flavonol (red points) biosynthesis. L1 to L7 indicate seven levels identified in three hierarchical time-ordered gene co-expression networks (*R. molle*-specific, *R. simsii*-specific and the consensus TO-GCNs). Edges between enzymatic genes were not shown. (c) Heatmaps of average normalized TPMs (z-score) at each timepoint of flowering at each level identified in three TO-GCNs. Three stages of flower coloring were identified as the initial (T1–T2), transitional (T3) and terminal (T4–T5) stages, based on the expression profile. Low to high expression is indicated by a change in color from blue to red.

Advances in these techniques will help researchers to better understand the regulatory mechanisms of gene expression, signal transduction pathways, metabolite synthesis and pathways, protein composition and interactions, and phenotypic variations in *Rhododendron*[61,85]. Therefore, genomics-centered next-generation -omics will improve our understanding of Rhododendrons genetic diversity and evolution of in a more precise and efficient way. Due to the weak research base and the nature of Rhododendrons as woody plants, a complete breeding cycle for flower color can take up to 10 years. Flower color is a typical qualitative trait, and population genetics and linkage analysis are effective strategies to identify the genes responsible for flower color. Multi-omics-based breeding will therefore enable faster and more efficient methods for cultivating *Rhododendron* and other woody plants.

Author contributions

The authors confirm contribution to the paper as follows: study conception and design: Nie S, Yang FS, Mao JF; data collection: Nie S, Ma HY, Tian XC, Shi TL; analysis and interpretation of results: Nie S, Ma HY, Tian XC, Shi TL; draft manuscript preparation: Nie S, Yang FS, Ma HY; edited and improved the manuscript: Porth I, El-Kassaby YA and Mao JF. All authors reviewed the results and approved the final version of the manuscript.

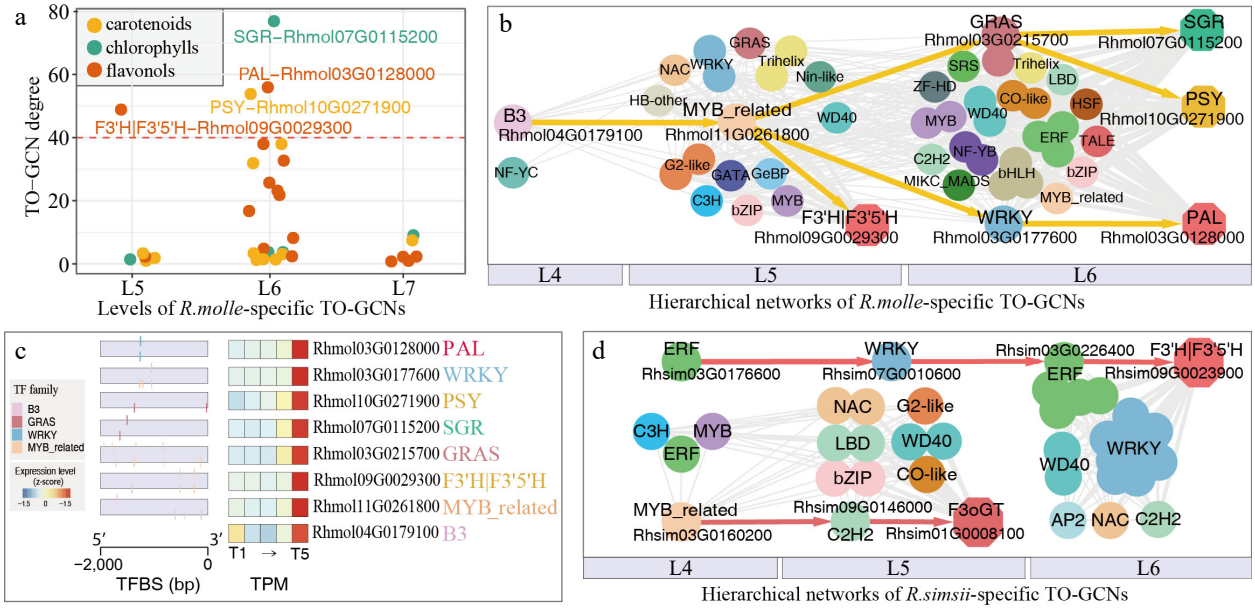

Fig. 6 Resolved hierarchical gene regulation during flower coloration of *R. molle* and *R. simsii*. (a) Degree of 34 enzymatic genes of the chlorophyll degradation and carotenoid/flavonol biosynthesis in two subnetworks. The degree cutoff for hub gene is 40. (b) Resolved hierarchical regulations for four hub genes *PSY*, *SGR*, *PAL* and *F3'H|F3'5'H* in *R. molle*. (c) Gene expressions (TPM) and TF binding site (TFBS) detected in the 2 kb upstream sequences of four hub genes and four potential regulators in *R. molle*. (d) Resolved hierarchical regulation for two hub genes *F3oGT* and *F3'H|F3'5'H* in *R. simsii*. (Modified from previous publications[26,31]).

Data availability

Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

Acknowledgments

This work was supported by the Science & Technology Fundamental Resources Investigation Program (Grant No. 2022FY101001).

Conflict of interest

The authors declare that they have no conflict of interest.

Supplementary Information accompanies this paper at (https://www.maxapress.com/article/doi/10.48130/opr-0024-0001)

Dates

Received 15 November 2023; Accepted 25 December 2023; Published online 2 February 2024

References

1. Chamberlain D, Hyam R, Argent G, Fairweather G, Walter KS. 1996. *The genus Rhododendron: its classification and synonymy*. Edinburgh, UK: Royal Botanic Garden Edinburgh. viii, 181 pp.
2. Goetsch L, Eckert AJ, Hall BD. 2005. The molecular systematics of *Rhododendron* (Ericaceae): a phylogeny based upon *RPB2* gene sequences. *Systematic Botany* 30:616–26
3. Geng Y. 2014. *The genus Rhododendron of China*. Shanghai: Shanghai Scientific and Technical Publishers. 312 pp.
4. Xia X, Yang M, Li C, Huang S, Jin W, et al. 2022. Spatiotemporal evolution of the global species diversity of *Rhododendron*. *Molecular Biology and Evolution* 39:msab314
5. Sleumer H. 1949. Ein system der gattung *Rhododendron* L. *Botanische Jahrbücherfür Systematik* 74:511–53
6. Stevenson JB. 1930. *The species of Rhododendron*. London, UK: The Rhododendron Society. 562 pp.
7. Mo Z, Fu C, Zhu M, Milne RI, Yang J, et al. 2022. Resolution, conflict and rate shifts: insights from a densely sampled plastome phylogeny for *Rhododendron* (Ericaceae). *Annals of Botany* 130:687–701
8. Kobayashi N. 2013. Evaluation and application of evergreen azalea resources of Japan. *Acta Horticulturae* 990:213–19
9. Kobayashi N, Handa T, Yoshimura K, Tsumura Y, Arisumi K, et al. 2000. Evidence for introgressive hybridization based on chloroplast dna polymorphisms and morphological variation in wild evergreen azalea populations of the kirishima mountains, Japan. *Edinburgh Journal of Botany* 57:209–19
10. Shen Y, Yao G, Li Y, Tian X, Li S, et al. 2023. RAD-seq data reveals robust phylogeny and morphological evolutionary history of the ornamentally important plant genus, *Rhododendron*. *Horticultural Plant Journal* In Press
11. De Riek J, De Keyser E, Calsyn E, Eeckhaut T, Van Huylenbroeck J, et al. 2018. Azalea. In *Ornamental Crops*, ed. Van Huylenbroeck J. Vol 11. Cham: Springer International Publishing. pp. 37–71. https://doi.org/10.1007/978-3-319-90698-0_11
12. Zhou H, Liao J, Xia Y, Teng Y. 2013. Determination of genetic relationships between evergreen azalea cultivars in China using AFLP markers. *Journal of Zhejiang University SCIENCE B* 14:299–308
13. Ma Y, Mao X, Wang J, Zhang L, Jiang Y, et al. 2022. Pervasive hybridization during evolutionary radiation of *Rhododendron* subgenus *Hymenanthes* in mountains of southwest China. *National Science Review* 9:nwac276
14. Sun T, Yuan H, Cao H, Yazdani M, Tadmor Y, et al. 2018. Carotenoid metabolism in plants: the role of plastids. *Molecular Plant* 11:58–74
15. Koes R, Verweij W, Quattrocchio F. 2005. Flavonoids: a colorful model for the regulation and evolution of biochemical pathways. *Trends in Plant Science* 10:236–42
16. Grotewold E. 2006. The genetics and biochemistry of floral pigment. *Annual Review of Plant Biology* 57:761–80

17. Tanaka Y, Sasaki N, Ohmiya A. 2008. Biosynthesis of plant pigments: anthocyanins, betalains and carotenoids. *The Plant Journal* 54:733–49
18. Spethmann W. 1980. Flavonoids and carotenoids of *Rhododendron* flowers and their significance for the classification of the genus *Rhododendron*. In *Contributions toward a Classification of Rhododendron*, eds Luteyn JL, O'Brien ME. New York: New York Botanical Garden. pp. 247–76
19. Rausher MD. 2008. Evolutionary transitions in floral color. *International Journal of Plant Sciences* 169:7–21
20. Du H, Lai L, Wang F, Sun W, Zhang L, et al. 2018. Characterisation of flower colouration in 30 *Rhododendron* species via anthocyanin and flavonol identification and quantitative traits. *Plant Biology* 20:121–29
21. de Camargo MGG, Lunau K, Batalha MA, Brings S, de Brito VLG, et al. 2019. How flower colour signals allure bees and hummingbirds: a community-level test of the bee avoidance hypothesis. *New Phytologist* 222:1112–22
22. Hopkins R, Rausher MD. 2012. Pollinator-mediated selection on flower color allele drives reinforcement. *Science* 335:1090–92
23. Sun S, Liao K, Xia J, Guo Y. 2005. Floral colour change in *Pedicularis monbeigiana* (Orobanchaceae). *Plant Systematics and Evolution* 255:77–85
24. Marks RA, Hotaling S, Frandsen PB, VanBuren R. 2021. Representation and participation across 20 years of plant genome sequencing. *Nature Plants* 7:1571–78
25. Ma H, Liu Y, Liu D, Sun W, Liu X, et al. 2021. Chromosome-level genome assembly and population genetic analysis of a critically endangered rhododendron provide insights into its conservation. *The Plant Journal* 107:1533–45
26. Nie S, Zhao S, Shi T, Zhao W, Zhang R, et al. 2023. Gapless genome assembly of azalea and multi-omics investigation into divergence between two species with distinct flower color. *Horticulture Research* 10:uhac241
27. Shirasawa K, Kobayashi N, Nakatsuka A, Ohta H, Isobe S. 2021. Whole-genome sequencing and analysis of two azaleas, *Rhododendron ripense* and *Rhododendron kiyosumense*. *DNA Research* 28:dsab010
28. Soza VL, Lindsley D, Waalkes A, Ramage E, Patwardhan RP, et al. 2019. The *Rhododendron* genome and chromosomal organization provide insight into shared whole-genome duplications across the heath family (Ericaceae). *Genome Biology and Evolution* 11:3353–71
29. Wang X, Gao Y, Wu X, Wen X, Li D, et al. 2021. High-quality evergreen azalea genome reveals tandem duplication-facilitated low-altitude adaptability and floral scent evolution. *Plant Biotechnology Journal* 19:2544–60
30. Wu X, Zhang L, Wang X, Zhang R, Jin G, et al. 2023. Evolutionary history of two evergreen *Rhododendron* species as revealed by chromosome-level genome assembly. *Frontiers in Plant Science* 14:1123707
31. Yang F, Nie S, Liu H, Shi T, Tian X, et al. 2020. Chromosome-level genome assembly of a parent species of widely cultivated azaleas. *Nature Communications* 11:5269
32. Zhang L, Xu P, Cai Y, Ma L, Li S, et al. 2017. The draft genome assembly of *Rhododendron delavayi* Franch var. *delavayi*. *GigaScience* 6:gix076
33. Zhou G, Li Y, Pei F, Gong T, Chen T, et al. 2022. Chromosome-scale genome assembly of *Rhododendron molle* provides insights into its evolution and terpenoid biosynthesis. *BMC Plant Biology* 22:342
34. Zhou X, Li J, Wang H, Han J, Zhang K, et al. 2022. The chromosome-scale genome assembly, annotation and evolution of *Rhododendron henanense* subsp. lingbaoense. *Molecular Ecology Resources* 22:988–1001
35. Chang Y, Zhang R, Ma Y, Sun W. 2023. A haplotype-resolved genome assembly of *Rhododendron vialii* based on PacBio HiFi reads and Hi-C data. *Scientific Data* 10:451
36. Cheng H, Concepcion GT, Feng X, Zhang H, Li H. 2021. Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nature Methods* 18:170–75
37. Tang L. 2019. Circular consensus sequencing with long reads. *Nature Methods* 16:958–58
38. De Loose R. 1969. The flower pigments of the belgian hybrids of *Rhododendron simsii* and other species and varieties from *Rhododendron* subseries *obtusum*. *Phytochemistry* 8:253–59
39. Liu X, Wang Y, Shen S. 2022. Transcriptomic and metabolomic analyses reveal the altitude adaptability and evolution of different-colored flowers in alpine *Rhododendron* species. *Tree Physiology* 42:1100–13
40. Liu N, Zhang L, Zhou Y, Tu M, Wu Z, et al. 2021. The *Rhododendron* Plant Genome Database (RPGD): a comprehensive online omics database for *Rhododendron*. *BMC Genomics* 22:376
41. Pellicer J, Hidalgo O, Dodsworth S, Leitch IJ. 2018. Genome size diversity and its impact on the evolution of land plants. *Genes* 9:88
42. Hidalgo O, Pellicer J, Christenhusz MJM, Schneider H, Leitch IJ. 2017. Genomic gigantism in the whisk-fern family (Psilotaceae): *Tmesipteris obliqua* challenges record holder *Paris japonica*. *Botanical Journal of the Linnean Society* 183:509–14
43. Kelly LJ, Renny-Byfield S, Pellicer J, Macas J, Novák P, et al. 2015. Analysis of the giant genomes of *Fritillaria* (Liliaceae) indicates that a lack of DNA removal characterizes extreme expansions in genome size. *New Phytologist* 208:596–607
44. Kovach A, Wegrzyn JL, Parra G, Holt C, Bruening GE, et al. 2010. The *Pinus taeda* genome is characterized by diverse and highly diverged repetitive sequences. *BMC Genomics* 11:420
45. Landis JB, Soltis DE, Li Z, Marx HE, Barker MS, et al. 2018. Impact of whole-genome duplication events on diversification rates in angiosperms. *American Journal of Botany* 105:348–63
46. Larson DA, Walker JF, Vargas OM, Smith SA. 2020. A consensus phylogenomic approach highlights paleopolyploid and rapid radiation in the history of Ericales. *American Journal of Botany* 107:773–89
47. Stull GW, Soltis PS, Soltis DE, Gitzendanner MA, Smith SA. 2020. Nuclear phylogenomic analyses of asterids conflict with plastome trees and support novel relationships among major lineages. *American Journal of Botany* 107:790–805
48. Wang Y, Chen F, Ma Y, Zhang T, Sun P, et al. 2021. An ancient whole-genome duplication event and its contribution to flavor compounds in the tea plant (*Camellia sinensis*). *Horticulture Research* 8:176
49. Chen H, Zeng Y, Yang Y, Huang L, Tang B, et al. 2020. Allele-aware chromosome-level genome assembly and efficient transgene-free genome editing for the autotetraploid cultivated alfalfa. *Nature Communications* 11:2494
50. Zhang C, Zhang T, Luebert F, Xiang Y, Huang C, et al. 2020. Asterid phylogenomics/phylotranscriptomics uncover morphological evolutionary histories and support phylogenetic placement for numerous whole-genome duplications. *Molecular Biology and Evolution* 37:3188–210
51. Chen J, Zheng C, Ma J, Jiang C, Ercisli S, et al. 2020. The chromosome-scale genome reveals the evolution and diversification after the recent tetraploidization event in tea plant. *Horticulture Research* 7:63
52. Barker MS, Arrigo N, Baniaga AE, Li Z, Levin DA. 2016. On the relative abundance of autopolyploids and allopolyploids. *New Phytologist* 210:391–98
53. Kellogg EA. 2016. Has the connection between polyploidy and diversification actually been tested? *Current Opinion in Plant Biology* 30:25–32
54. Linder CR, Rieseberg LH. 2004. Reconstructing patterns of reticulate evolution in plants. *American Journal of Botany* 91:1700–08
55. Lynch M, Conery JS. 2000. The evolutionary fate and consequences of duplicate genes. *Science* 290:1151–55

56. Page RDM. 1994. Maps between trees and cladistic analysis of historical associations among genes, organisms, and areas. *Systematic Biology* 43:58–77
57. Rabier CE, Ta T, Ané C. 2014. Detecting and locating whole genome duplications on a phylogeny: a probabilistic approach. *Molecular Biology and Evolution* 31:750–62
58. Shi T, Huang H, Barker MS. 2010. Ancient genome duplications during the evolution of kiwifruit (*Actinidia*) and related Ericales. *Annals of Botany* 106:497–504
59. Wei C, Yang H, Wang S, Zhao J, Liu C, et al. 2018. Draft genome sequence of *Camellia sinensis* var. *sinensis* provides insights into the evolution of the tea genome and tea quality. *Proceedings of the National Academy of Sciences of the United States of America* 115:E4151–E4158
60. Nie S, Tian X, Kong L, Zhao S, Chen Z, et al. 2022. Potential allopolyploid origin of Ericales revealed with gene-tree reconciliation. *Frontiers in Plant Science* 13:1006904
61. Depuydt T, De Rybel B, Vandepoele K. 2023. Charting plant gene functions in the multi-omics and single-cell era. *Trends in Plant Science* 28:283–96
62. Shen S, Zhan C, Yang C, Fernie AR, Luo J. 2023. Metabolomics-centered mining of plant metabolic diversity and function: past decade and future perspectives. *Molecular Plant* 16:43–63
63. Park CH, Yeo HJ, Kim NS, Park YE, Park SY, et al. 2018. Metabolomic profiling of the white, violet, and red flowers of *Rhododendron schlippenbachii* Maxim. *Molecules* 23:827
64. Wang S, Li Z, Jin W, Fang Y, Yang Q, Xiang J. 2018. Transcriptome analysis and identification of genes associated with flower development in *Rhododendron pulchrum* Sweet (Ericaceae). *Gene* 679:108–18
65. Xiao Z, Su J, Sun X, Li C, He L, et al. 2018. De novo transcriptome analysis of *Rhododendron molle* G. Don flowers by Illumina sequencing. *Genes & Genomics* 40:591–601
66. Ye L, Möller M, Luo Y, Zou J, Zheng W, et al. 2021. Differential expressions of anthocyanin synthesis genes underlie flower color divergence in a sympatric *Rhododendron sanguineum* complex. *BMC Plant Biology* 21:204
67. Thi Thanh Huyen D, Ureshino K, Thanh Van D, Miyajima I. 2016. Co-pigmentation of anthocyanin-flavonol in the blotch area of *Rhododendron simsii* planch. flowers. *The Horticulture Journal* 85:232–37
68. Li Z, Yang Q, Dong X, Zhu Y, Zhao S, et al. 2021. Transcriptome analysis of flower color variation in five *Rhododendron* species (Ericaceae). *Brazilian Journal of Botany* 44:685–95
69. Long F, Wu H, Li H, Zuo W, Ao Q. 2023. Genome-wide analysis of MYB transcription factors and screening of *MYBs* involved in the red color formation in *Rhododendron delavayi*. *International Journal of Molecular Sciences* 24:4641
70. Wang C, Ye D, Li Y, Hu P, Xu R, et al. 2023. Genome-wide identification and bioinformatics analysis of the WRKY transcription factors and screening of candidate genes for anthocyanin biosynthesis in azalea (*Rhododendron simsii*). *Frontiers in Genetics* 14:1172321
71. Chen Y, Ma T, Zhang L, Kang M, Zhang Z, et al. 2020. Genomic analyses of a "living fossil": the endangered dove-tree. *Molecular Ecology Resources* 20:756–69
72. Sun X, He L, Guo Z, Xiao Z, Su J, et al. 2022. Comparative transcriptome analyses reveal genes related to pigmentation in the petals of a flower color variation cultivar of *Rhododendron obtusum*. *Molecular Biology Reports* 49:2641–53
73. Wang Y, Zhang G, He J, Xu S, Liu X, et al. 2020. Research progress of *Rhododendron* flower color. *World Forestry Research* 33:19–24
74. Deng W, Zhang K, Busov V, Wei H. 2017. Recursive random forest algorithm for constructing multilayered hierarchical gene regulatory networks that govern biological pathways. *PLoS ONE* 12:e0171532
75. Leslie AC. 2008. *The international Rhododendron register and checklist (2004), fourth supplement*. UK: University Press, Cambridge. 27 pp. www.rhs.org.uk/plants/pdfs/plant-register-supplements/rhododendrons/4thrhodosupp.pdf
76. Zhuang P. 2012. Discuss on the *Rhododendron* geographical distribution types and their cause of formation in China. *Guihaia* 94:150–56
77. Lan X, Zhang L, Zhang J, Cui H, Jiang C, et al. 2012. Research progress of *Rhododendron* breeding. *Acta Horticulturae Sinica* 39:1829–38
78. Zhuang P. 2019. Progress on the fertility of *Rhododendron*. *Biodiversity Science* 27:327–38
79. Zhang C, Gao L, Xue R, Yang J. 2004. A general review of the research and conservation statue of Chinese *Rhododendron*. *Guangxi Sciences* 11:354–59, 362
80. Cheng J, Li M, Yuan T, Huang H, Yang G, et al. 2021. A dataset on wild *Rhododendron* and geographical distribution information in China. *Biodiversity Science* 29:1175–80
81. Fang L, Mao J, Xu D, Dong Y, Zhou Y, et al. 2021. Development of high quality EST-SSR markers in *Rhododendron obtusum* Hort. ex Wats. and their use in determining relationships among *Rhododendron* cultivars. *Genetic Resources and Crop Evolution* 68:3271–84
82. Zhang C, Huang C, Huang J, Wang L, Zhang J, et al. 2015. Investigation of germplasm resources of the genus *Rhododendron* in Baili nature reserve in Guizhou. *Plant Diversity* 37:357–64
83. Fu C, Mo Z, Yang J, Cai J, Ye L, et al. 2022. Testing genome skimming for species discrimination in the large and taxonomically difficult genus *Rhododendron*. *Molecular Ecology Resources* 22:404–14
84. Pucker B, Irisarri I, de Vries J, Xu B. 2022. Plant genome sequence assembly in the era of long reads: progress, challenges and future directions. *Quantitative Plant Biology* 3:e5
85. Yang Y, Saand MA, Huang L, Abdelaal WB, Zhang J, et al. 2021. Applications of multi-omics technologies for crop improvement. *Frontiers in Plant Science* 12:563953