

RESEARCH

Open Access



Adaptation insights from comparative transcriptome analysis of two *Opisthopappus* species in the Taihang mountains

Ning Chen¹, Hao Zhang¹, En Zang¹, Zhi-Xia Liu¹, Ya-Fei Lan¹, Wei-Li Hao¹, Shan He¹, Xing Fan², Gen-Lou Sun^{3*} and Yi-Ling Wang^{1*}

Abstract

Opisthopappus is a major wild source of Asteraceae with resistance to cold and drought. Two species of this genus (*Opisthopappus taihangensis* and *O. longilobus*) have been employed as model systems to address the evolutionary history of perennial herb biomes in the Taihang Mountains of China. However, further studies on the adaptive divergence processes of these two species are currently impeded by the lack of genomic resources. To elucidate the molecular mechanisms involved, a comparative analysis of these two species was conducted. Among the identified transcription factors, the bHLH members were most prevalent, which exhibited significantly different expression levels in the terpenoid metabolic pathway. *O. longilobus* showed higher level of expression than did *O. taihangensis* in terms of terpenes biosynthesis and metabolism, particularly monoterpenoids and diterpenoids. Analyses of the positive selection genes (PSGs) identified from *O. taihangensis* and *O. longilobus* revealed that 1203 genes were related to adaptative divergence, which were under rapid evolution and/or have signs of positive selection. Differential expressions of PSG occurred primarily in the mitochondrial electron transport, starch degradation, secondary metabolism, as well as nucleotide synthesis and S-metabolism pathway processes. Several PSGs were obviously differentially expressed in terpenes biosynthesis that might result in the fragrances divergence between *O. longilobus* and *O. taihangensis*, which would provide insights into adaptation of the two species to different environments that characterized by sub-humid warm temperate and temperate continental monsoon climates. The comparative analysis for these two species in *Opisthopappus* not only revealed how the divergence occurred from molecular perspective, but also provided novel insights into how differential adaptations occurred in Taihang Mountains.

Keywords: *Opisthopappus*, Comparative transcriptome, Differentiation, Adaptation

Introduction

Adaptation is a topic of fundamental interest to researchers in ecology, evolution, and conservation [1–4]. When different populations experience heterogeneous environments, natural selection can drive phenotypic divergence while modulating underlying genomic architectures,

which induces populations to adapt to their habitats [3, 5]. Adaptative divergence has been well verified as a major mechanism that initiates evolutionary diversification and speciation [6, 7]. There are myriad variations in the combination of phenotypic traits and environmental factors [8]. Variation created by mutation, the raw material for evolutionary change, is translated into phenotypes by flux through metabolic pathways and by the topography and dynamics of molecular networks [9]. Thus, empirical adaptation tests are of principal importance as they explore the balance between the evolutionary

*Correspondence: genlou.sun@smu.ca; ylwangbj@hotmail.com

¹ College of Life Science, Shanxi Normal University, Taiyuan 030031, China

³ Department of Biology, Saint Mary's University, Halifax B3H3C3, Canada

Full list of author information is available at the end of the article



processes that shape populations (e.g., strength of selection in promoting local adaptation) relative to non-adaptive or even maladaptive factors such as gene flow, recombination, mutation, and genetic drift [10–12].

Transcriptome sequencing is a fast and cost-effective approach for characterizing entire RNA data of an organism [13]. It has been employed to identify the functional elements in the genomes, metabolic pathways, and differentially expressed genes (DEGs) of plants [14–21], particularly those that lack a sequenced reference genome, or non-model species [22]. Moreover, comparative transcriptome analysis has become a powerful tool for investigating plant growth, development, and physiology at the transcriptional and metabolic levels [23], which provides opportunities to estimate transcriptome-wide divergence and identify loci under selection [17]. Especially, comparative RNA-sequencing studies between intimately related species can not only provide additional genomic resources but also offer data related to the processes of speciation or adaptive evolution [16, 24].

The *Opisthopappus* Shih genus belongs to the family Asteraceae, which includes two closely related species (*O. taihangensis* Shih and *O. longilobus* (Ling) Shih) [25]. These two species generally grow within the cliff cracks and rock gaps of the Taihang Mountains in China. They have similar morphological characteristics, but can be distinguished based on the style of pinnatisect leaves and the presence or absence of bracteal leaves [26, 27]. For *O. longilobus*, most stems and leaves are pinnatifid with one pair of bracteal leaves; *O. taihangensis* has bipinnatifid stems and leaves with no bracteal leaves [27]. Based on our observations in a common-garden, the stems of *O. longilobus* are more brittle than those of *O. taihangensis*, and *O. longilobus* individuals are relatively smaller than those of *O. taihangensis*. Further, *O. taihangensis* more easily survives than *O. longilobus* when they were transplanted into a common garden. Meanwhile, both *O. taihangensis* and *O. longilobus* can exude unique fragrances. HPLC (High Performance Liquid Chromatography) and FTIR (Fourier Transform Infrared Spectrometry) analyses revealed that the chemical compositions and contents of chlorogenic acid, apigenin, and quercetin were different between the two species (data not published).

During the long-term evolution of *Opisthopappus*, the dramatic geological changes and significant topographical barriers of Taihang Mountains could limit the large-scale population expansion and gradually introduce heterogeneity within different populations of *O. taihangensis* and *O. longilobus* [28]. *O. taihangensis* and *O. longilobus* were likely diverged during the early Miocene under a strengthening East Asian monsoon [29]. Subsequently, the dramatic climatic transformation and

complex topography of the Taihang Mountains further promoted differentiations between the two species and their populations [29]. At the present, *O. taihangensis* and *O. longilobus* occupy different niches due to difference in growth habits [30]. According to the current geographical distribution of *O. taihangensis* and *O. longilobus*, we presumed that adaptive divergence processes might have played an additional role in maintaining the differentiation of the two *Opisthopappus* species that inhabit different regions of the Taihang Mountains [29]. Whether driven by morphological, genetic, or metabolic aspects, differentiation proceeded between *O. taihangensis* and *O. longilobus*. Previous genetic and phylogeographic studies on nuclear DNA, plastid DNA and transcriptome data have shown significant variation between the two species [28, 30–34]. However, the molecular mechanisms underlying these phenomena remains unclear, where a paucity of genetic resources has made further studies on the neutral and adaptive divergence processes of *Opisthopappus* is a challenging task.

In this study, transcriptome sequencing and comparative analysis of *O. taihangensis* and *O. longilobus* were conducted to reveal the expression of differentially genes under different selective pressures and the main metabolic pathways patterns in the two species. These analyses will provide useful data regarding their adaptation to specific habitats and enable the exploration of adaptive evolutionary signatures between the two species. It is anticipated that the data generated in this study will enhance our knowledge of the regulatory networks and pathways involved in assessing the roles of evolutionary factors on the adaptation, and speciation of *Opisthopappus*.

Materials and methods

Plant materials

O. taihangensis distributes mostly across Henan and Shanxi Provinces, which are located along the southern Taihang Mountains and belong to a sub-humid warm temperate climate; while *O. longilobus* generally grows in Hebei and Shanxi Provinces, which reside along the northern Taihang Mountains with a temperate continental monsoon climate.

To elucidate the molecular mechanisms of differentiation between *O. taihangensis* and *O. longilobus*, the perennial plants of the two species (Fig. 1A and B) were sampled from Shanxi Province where the two have adjacent distribution as closely as possible (Study protocol comply with relevant institutional, national, and international guidelines and legislation, and collection of *O. taihangensis* and *O. longilobus* have obtained the permission from the College of Life Sciences of Shanxi Normal University). During the same growth period of

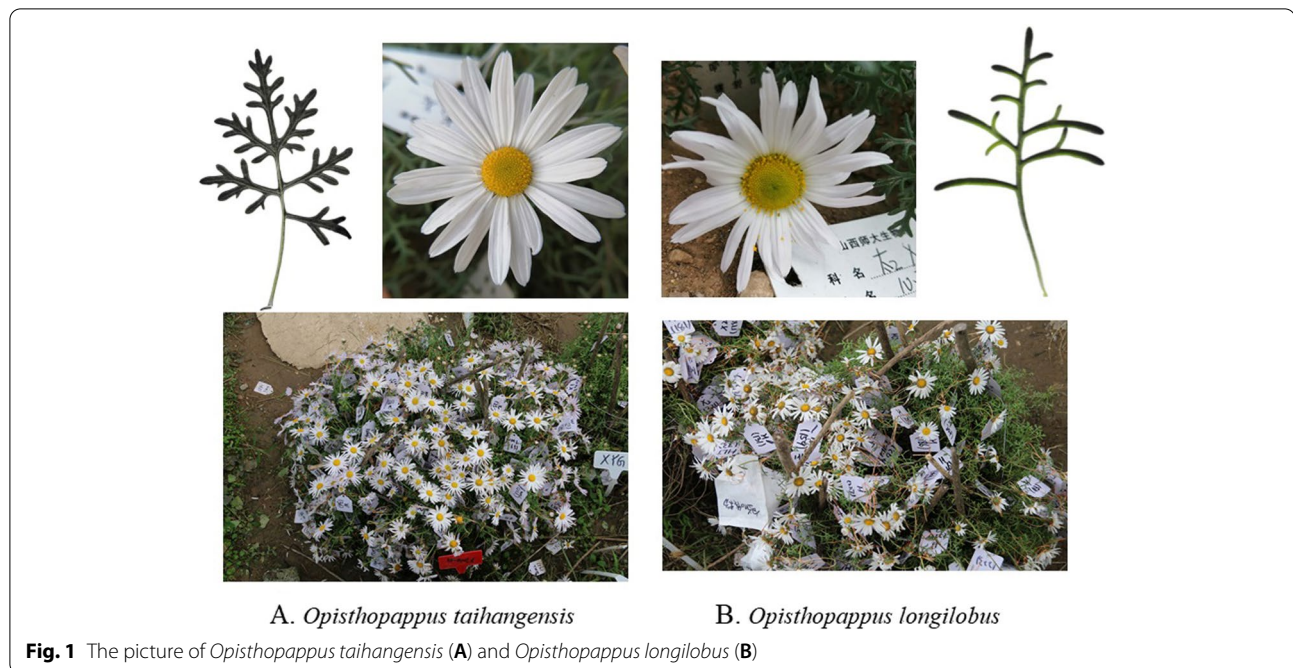


Fig. 1 The picture of *Opisthopappus taihangensis* (A) and *Opisthopappus longilobus* (B)

O. taihangensis and *O. longilobus* at July in 2015, fresh mature leaves from the same branch and position of each species were sampled. Each individual from the same population was collected from different locations at least 10 m apart. Then mixed leaves of five individuals were collected as a sample. Three samples (biological replicates) were tested for each population, including taihangensis1, taihangensis2 and taihangensis3 for *O. taihangensis* and longilobus1, longilobus2 and longilobus3 for *O. longilobus*. The samples were frozen in liquid nitrogen (immediately stored at -80°C) and then sent to Shanghai Personal Biotechnology Co., Ltd. The data were generated in accordance with a transcriptome analysis program in our previous studies [28].

TF identification and expression analysis

PlantTFDB (<http://planttfdb.cbi.pku.edu.cn/index.php>) is a plant transcription factor database that includes the sequences of 58 plant transcription factor families from 165 plant species [35]. Clean reads were assembled de novo using Trinity software, and the sequences of unigenes were Blastx aligned to the transcription factor database PlantTFDB, where the best of these with E values of less than $1e-5$, were screened as the annotation data of the unigenes. Candidates that contained DNA binding domains were recognized by GO annotation for final TF identification.

An in-depth transcriptome analysis would facilitate an improved perception of the species under study [15, 21, 36]. An advanced bioinformatics tool (MapMan) was

used to analyze and compare transcriptional responses between phylogenetically species, and especially useful at the level of functional categories in cross-species comparisons [37]. This analysis allows us to explore gene categories from large data sets to extract meaningful information. In this study, MapMan [38, 39] was employed to comprehensively interpret transcriptome data and visualize the functionalities of associated genes. Through the MapMan, it was evident that significant variations between two species were conspicuous with respect to differentially expressed genes [40]. At the same time, we employed metabolic overview installed in the MapMan tool to identify the primary and secondary metabolic pathways related to the genes associated with transcription factors (TFs) [41]. MapMan metabolism encompasses photosynthesis pathways, carbohydrate metabolism, N-dependent pathways (e.g., amino acid metabolism), as well as cell wall, lipid, and secondary metabolism. Meanwhile, the differentially expressed TFs and transporters were isolated by Fisher's Exact Test (P -value < 0.05) and enrichment fold ≥ 1.5 compared with the whole genome background.

PSG identification and expression analysis

Orthologous contigs and estimation of substitution rates

The assembled sequences were initially employed to predict the ORF (Open Read Frame) regions. Subsequently, the ORFs were blasted using OrthoMCL (<http://orthomcl.org/orthomcl/>) software. Orthologous pairs with identities of $< 60\%$ were excluded, and only 1:1

orthologous pair in both lineages were retained. Next, based on the BLAST results, Markov clustering was used with the Markov Cluster algorithm (MCL) method. Aligned sequences showing a 90% identity were defined as pairs of putative orthologues. The best-hit sequences of each cluster were then used in subsequent analyses.

The KAKS_CALCULATOR software [42, 43] was used to calculate nonsynonymous (dN) and synonymous (dS) substitution rates of the single-copy orthologous genes, and the dN/dS ratios of each putative orthologous pair using the YN [44] algorithm in PAML software. The dN/dS ratio is an estimate of natural selection acting on the genes. The estimated values and its proportion were assigned to one out of five selection scenarios: (i) dN/dS = 0–0.6 (strong negative selection), (ii) dN/dS = 0.61–0.9 (weak negative selection), (iii) dN/dS = 0.91–1.5 (neutral selection), (iv) dN/dS = 1.51–5 (weak positive selection), and (v) dN/dS > 5 (strong positive selection) [45]. Alignments with dS values of < 0.01 and dS or dN > 2 were discarded. Very high dN/dS (> 10) were also removed as they indicated bias.

According to the standard errors computed from YN00 for a Student's t-test, the statistical significance of the differences between dN and dS was analyzed [44]. The pair of orthologous with a *P*-value < 0.05 and dN/dS > 1.5 was designated as preliminary candidate under positive selection. A *P*-value < 0.01 and dN/dS > 5 were used as threshold for candidate genes with stronger positive selection.

GO classification analysis

The identified positive selection genes were processed to retrieve associated Gene Ontology (GO) terms describing biological processes, molecular functions, and cellular components using TBtools software. Expression levels of all the positive selection genes in replicates were assessed by mapping high quality (HQ) filtered reads using BOWTIE2 [46]. Mapped reads were further normalized using the Fragments Per Kilobase Per Millions (FPKM) method. The significance of the difference in GO term abundance between the two datasets was tested using the Fisher's exact test of the GOSSIP package [47] implemented in BLAST2GO [48].

Pathway enrichment analysis

To analyze the PSG pathways between the two species, we employed two annotation methods (KEGG and MapMan) to compare the expressional differentiation between *O. taihangensis* and *O. longilobus*. KEGG (<http://www.kegg.jp/>) [49] pathway annotation was performed using the BLASTX algorithm with *E*-values < 1.00E⁻⁵ and enabled 'sensitive mode' against the KEGG (Kyoto Encyclopedia of Genes and Genomes) gene peptide database. Subsequently, the corresponding KO identifiers,

EC numbers and pathway categories were parsed using customized Perl scripts. For the KEGG function classification, unigenes were annotated into five categories of KEGG metabolic pathways (cellular processes, environmental information processing, organismal systems, metabolic and genetic information processing).

For MapMan analysis [41], transcripts expressed in both species were annotated with the TAIR database (*Arabidopsis* homologs). A MapMan BIN file, with a hierarchical ontology system for basil genes, was prepared using Mercator [50, 51], by comparing transcripts against already-classified proteins. BIN level information and gene identifiers were also derived from the same database.

Results

Transcription factors (TFs)

Numerous studies have shown that transcription factors (TFs) are involved in the physiological processes and adaptation of plants, which can regulate gene expression during all developmental and growth stages. Of the 483 unigenes, 406 were transcription factors belonging to 41 transcription factors families, including bHLH, bZIP, etc. Among them, the bHLH family was the most abundant with 67 transcription factors, whereas the second most abundant family was FAR1 with 50 transcription factors, followed by the NAC (40 transcription factors), M-type (24 transcription factors), MYB and MYB related (24 transcription factors), and WRKY (22 transcription factors) families (Table 1).

While, families having 10–19 transcription factors included Nin-like (19 factors), B3 (17 factors), EIL (13 factors), C3H (11 factors), G2-like (10 factors), and Trihelix (10 factors). Other families had less than 10 transcription factors, such as BES1 (8 factors), SBP (7 factors), and TCP (4 factors). Seven families (ZF-HD, RAV, NF-YA, HRT-like, GATA, E2F/DP, and CAMTA) had only one factor. Among the 406 TFs, 309 were shared by *O. taihangensis* and *O. longilobus*, while 46 were exclusive to *O. longilobus* and 13 were exclusive to *O. taihangensis*.

To systematically assign functions to the 406 TF genes, we employed the diverse overview tools installed in MapMan. The metabolism overview assigned the 406 TF genes as follow: seven for amino acid metabolism, four for secondary metabolism, three for lipid metabolism, one for cell wall metabolism, one for major carbohydrate metabolism, one for photosystems, and one for fermentation (Fig. 2).

Significant expression differentiation between *O. taihangensis* and *O. longilobus* occurred in the terpenoid metabolic pathway, in which the transcription factors of *O. longilobus* were obviously more active than those of *O. taihangensis*. Furthermore, for the amino

Table 1 Transcription factors (TFs) identified in this study

TFs	Unigene counts	TFs	Unigene counts	TFs	Unigene counts	TFs	Unigene counts
bHLH	67	G2-like	10	C2H2	4	HD-ZIP	2
FAR1	50	Trihelix	10	CO-like	4	LBD	2
NAC	40	bZIP	10	Dof	4	STAT	2
M-type	24	MYB	9	S1Fa-like	4	CAMTA	1
WRKY	22	BES1	8	TCP	4	E2F/DP	1
Nin-like	19	GRAS	8	ERF	3	GATA	1
B3	17	GeBP	7	HB-other	3	HRT-like	1
MYB-related	15	SBP	7	NF-YB	3	NY-YA	1
EIL	13	YABBY	7	NF-YC	3	RAV	1
C3H	11	HSF	5	ARF	2	ZF-HD	1

acid metabolism pathways, the TFs were observed to be more numerously expressed in *O. taihangensis*, whereas in other pathways the TFs were almost equally expressed between the two species (Fig. 7A).

Positive selection genes(PSGs)

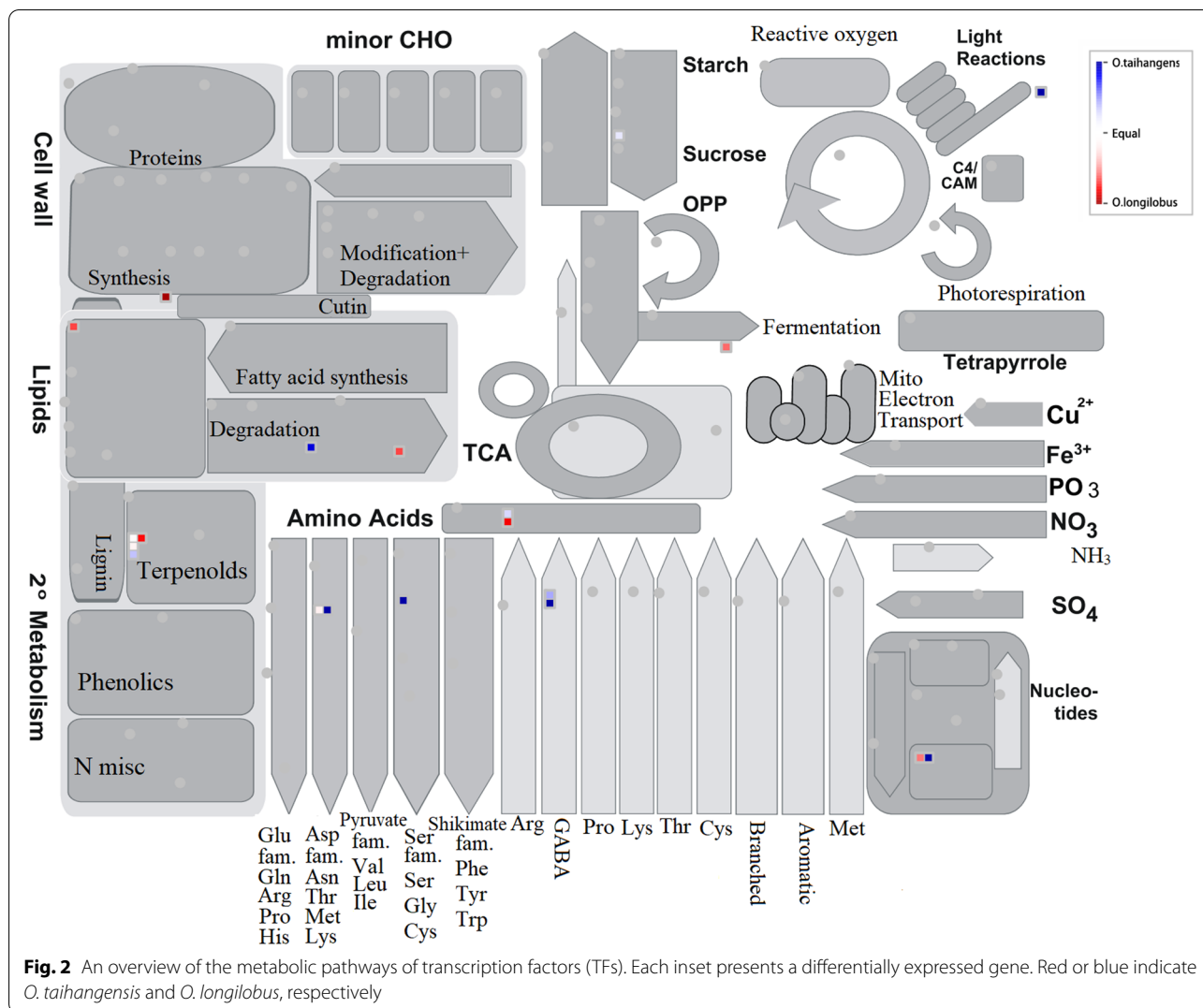
Positive selection genes (PSGs) are those under natural selection, which reflect the great adaptation of species to factors that are intimately linked to their survival under specific environmental conditions. Thus, PSGs were analyzed to explore the variations in the development and adaptation of the two *Opisthopappus* species.

Based on the predicted ORE, we identified 59,753 pairs of putative orthologous contigs between *O. taihangensis* and *O. longilobus* using ORTHOMCL. The one-to-one, reciprocal best method for elucidating orthologous proteins generated 50,587 putative orthologous pairs. After filtering the gene pairs annotated with different proteins in the Swiss-Prot database, 38,986 pairs of putative orthologs were finally identified and used in subsequent analyses. Of these, 11,598 pairs had only either synonymous or non-synonymous substitutions, and 38,986 pairs had both types of substitutions, for which the dN/dS ratios were calculated. Among the 38,986 pairs of putative orthologs, 1596 putative orthologous pairs had dN/dS ratios = 0.91–1.5, which indicated ongoing neutral evolution. And 36,187 with dN/dS ratios between 0.0001 and 0.9 suffered negative selection that might have experienced relaxed purifying selection and/or unfixed mutations without altering the encoded amino acid sequence during the speciation period.

There were five orthologous pairs with dN/dS ratios > 10 and 1203 with dN/dS ratios > 1.51 ($P < 0.05$). According to the frequency distribution plots of all dN/dS ratios of two species, we took a more appropriate threshold of 1.5 for the dN/dS ratio as an indicator of positive selection [17, 22, 45]

thus 1203 pairs with dN/dS values > 1.51 were identified as undergoing positive selection ($P < 0.05$) (Fig. 3A). Among the positive selection genes, 63 were shared by the two species, and 1125 were either exclusive to *O. taihangensis* or to *O. longilobus* (Fig. 3A). All positively selection genes were shown in a boxplot distribution with a median of 2.0308 (Fig. 3B). These results suggested divergent evolutionary processes in *O. taihangensis* and *O. longilobus*. Meanwhile, we clustered the above 1203 putative orthologous pairs into three main GO categories (biological process, cellular component, and molecular function). Within the biological process category, the term ‘menaquinone metabolic process’ and ‘menaquinone biosynthetic process’ were the most dominant. Within the cellular component category, ‘nuclear origin of replication recognition complex’ and ‘origin recognition complex’ represented the major subcategories. Within the molecular function category, the main functional subcategories were ‘2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1’ and ‘metal ion transmembrane transporter activity’ (Fig. 4A).

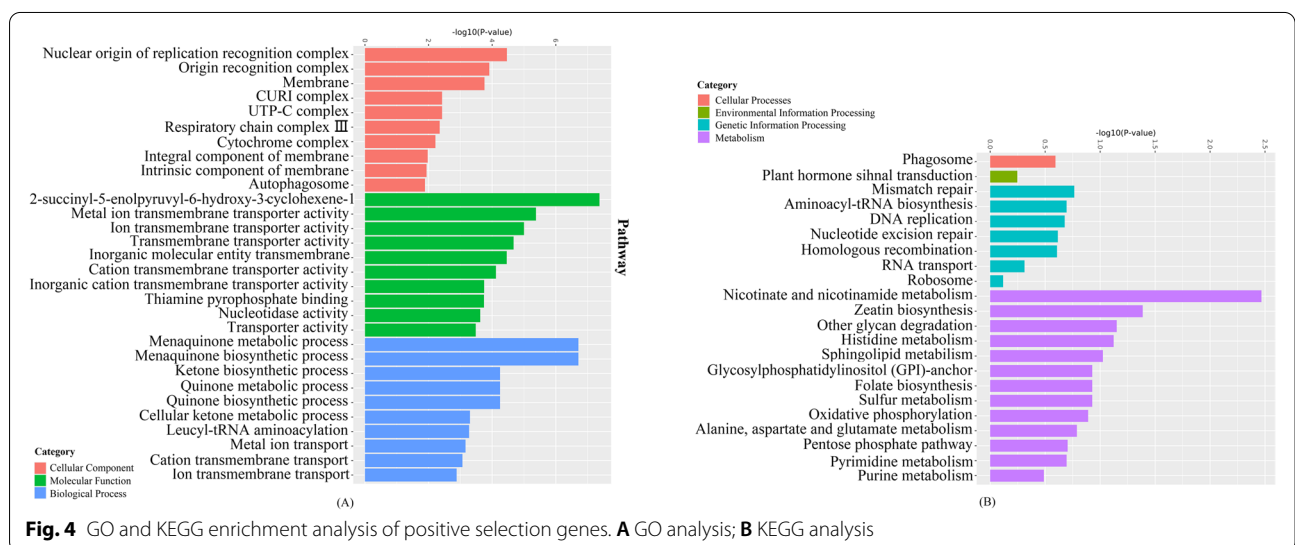
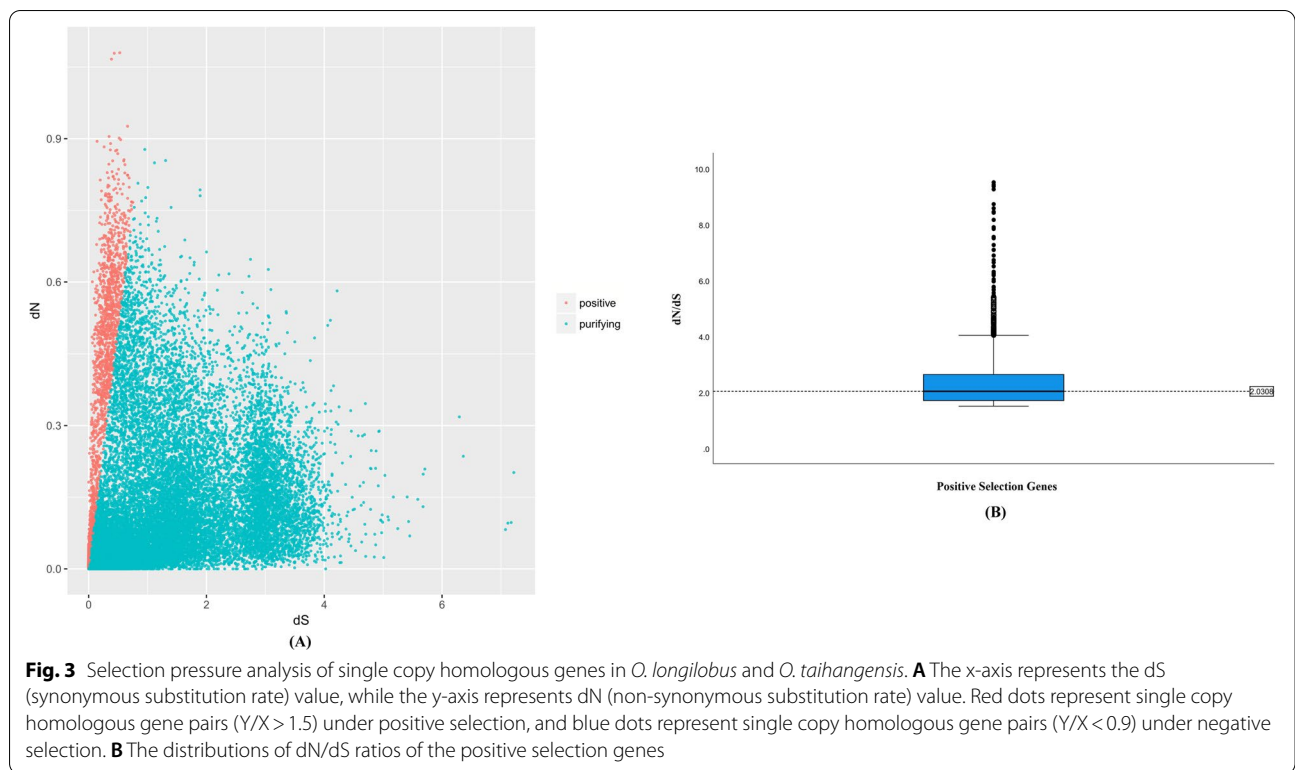
Furthermore, we used KEGG pathways to investigate the biological functions of positive selection genes. Overall, the PSGs were assigned to 22 KEGG pathways. The most represented pathways were “Nicotinate and nicotinamide metabolism”, “Zeatin biosynthesis (Metabolism of terpenoids and polyketides)”, “Other glycan degradation”, “Histidine metabolism”, and “Sphingolipid metabolism” (Fig. 4B). For a comprehensive assessment of the variations/similarities in the transcriptomes of both species, PSG transcripts were plotted using the MapMan tool, which were separated by bins based on their functional ontology (Fig. 5). The functional pathway classification of 1203 PSGs mapped to the “overview of metabolism”, with 4 to lipid metabolism, 12 to secondary metabolism, 7 to amino acid metabolism, 5 to mitochondrial electron transport, 4 to nucleotide metabolism, 3



to major carbohydrate metabolism, one to the tricarboxylic acid (TCA) cycle, 2 redox ascorbate and glutathione ascorbate, 1 to oxidative PP6-phosphogluconolactonase, one to C1-metabolism, one to cell wall metabolism, one to photosystems and one to tetrapyrrole synthesis.

Although major differences were not observed between *O. longilobus* and *O. taihangensis*, and most transcripts were shared by the two species, there were still a number of PSGs that differed. For example, the lipids pathway, major carbohydrate metabolism pathways, nucleotide metabolism pathways, tetrapyrrole metabolism pathways, mitochondrial electron transport pathways, presented higher expression levels in *O. longilobus* than in *O. taihangensis*. On the contrary, valine, leucine, and isoleucine biosynthesis in the amino acid pathway and C-1 metabolism pathways were obviously more active in *O. taihangensis* than *O. longilobus* (Fig. 5A).

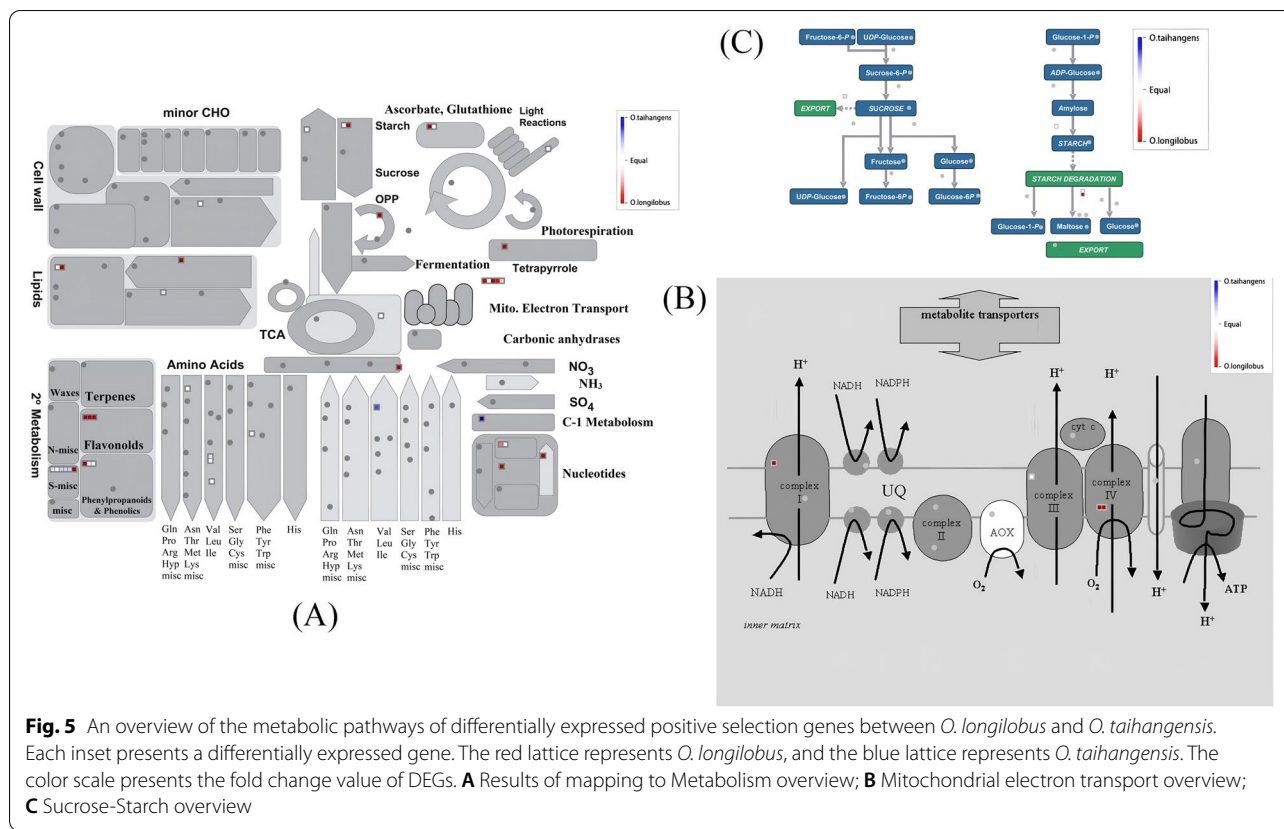
We simultaneously annotated the mitochondrial electron transport / ATP synthesis pathway, nucleotide metabolism pathway and sucrose-starch pathway. An overview of mitochondrial electron transport / ATP synthesis pathway showed that the various pathways presented mainly on NADH-DH complex I and cytochrome c oxidase of the mitochondrial electron transport chains complex I and IV involved the differentially expressed PSG trinity_dn142097_c0_g1 (prohibitin 3, PHB3) (Fig. 5B). In terms of nucleotide metabolism pathway, the differentially expressed PSG trinity_dn139928_c0_g1 (inositol monophosphatase family protein) for *O. longilobus* was significantly higher than those for *O. taihangensis* (Fig. 5A). An overview of major carbohydrate metabolism pathways suggested that expression level of the PSG trinity_dn108610_c0_g1 (encodes a beta-amylase targeted to the chloroplast) involved in beta-amylase



for *O. longilobus* was significantly higher than that for *O. taihangensis* (Fig. 5C).

Based on the above results, either TFs or PSGs, or both, can enrich the metabolic process pathway. We further annotated the secondary metabolite pathway of the metabolic process. The analysis of pathways enriched by MapMan revealed that terpenoid biosynthesis was simultaneously enriched at PSGs (Fig. 6A)

and TFs (Fig. 6B), and only the terpene biosynthesis pathway exhibited higher expression in *O. longilobus* than *O. taihangensis*. Except for the terpene biosynthesis pathway, the PSGs trinity_dn151066_c0_g1 (annotated as chalcone isomerase) and trinity_dn141682_c0_g1 (annotated as neoxanthin biosynthesis cofactor, NXD1) were primarily identified on the flavanone biosynthesis and xanthophyll biosynthesis pathways (Fig. 6A).



Furthermore, there were significantly different expression levels between the two species in these pathways. Terpenoids biosynthesis were enriched on mono-/sesquiterpene-/diterpene synthase where the expression levels of the PSG trinity_dn129788_c0_g1 (annotated as terpene synthase 14, TPS14) were higher in *O. longilobus* than *O. taihangensis* (Fig. 7A). Except for the monoterpene and diterpene biosynthesis pathways, the genes were mainly enriched on non-mevalonate pathway secondary metabolism pathways, especially ISPD (4-diphosphocytidyl-2-C-methyl-D-erythritol synthase) and ISPE (4-diphosphocytidyl-2-C-methyl-D-erythritol kinase), where the expression levels of the PSG trinity_dn155665_c0_g2 (encodes 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase, ISPE) were still higher in *O. longilobus* than *O. taihangensis* (Fig. 7A).

The terpenoid biosynthesis pathway was further analyzed by MapMan tools. The results indicated that the enriched pathway was carotenoid biosynthesis pathway with higher expression levels in *O. longilobus* than *O. taihangensis* for TFs. Moreover, this differentiation primarily occurred along the monoterpene and diterpene biosynthesis pathways. For these two pathways, the involved TFs all belonged to the bHLH family (Fig. 7B).

Discussion

Previous studies have confirmed significant differentiation and variation between *O. longilobus* and *O. taihangensis* based on geographic distribution, morphological characteristics and DNA information [26, 27, 29–33]. Subsequently, Chai [28, 34] used transcriptome data to reveal differentially expressed genes and screened DNA markers of the two species to further prove the differentiation and variation between the two species. To uncover the environment-related molecular mechanisms underlying the differentiation and variation of the two *Opisthospaprus* species, we performed a comparative analysis of genes for adaptive evolution occurring in transcriptome data, providing some clues to further elucidate the major differences under natural selection.

Adaptive characteristics of TFs

Transcription factors (TFs) are significant key elements of various physiological and biochemical pathways in higher plants. The activation and repression of TFs can primarily regulate the expression of genes related to growth, stress responses of plants, and other developmental processes [40, 52]. Thus, TF-based gene expression regulation allows plants to respond to changes in their environments [40, 53]. Based on transcriptome

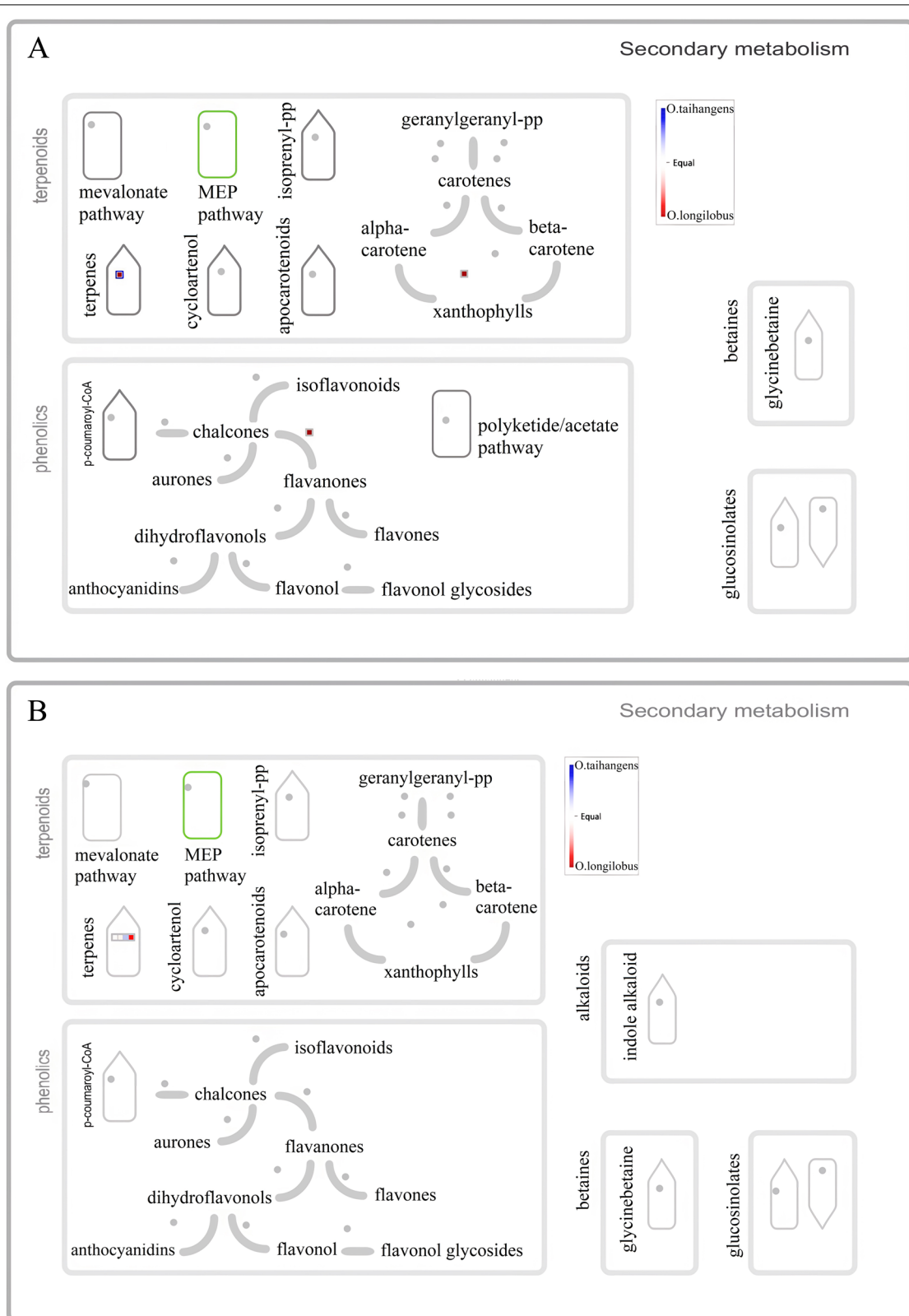


Fig. 6 Terpenoid biosynthesis overview in secondary metabolic pathways. **A** for positive selection genes (PSGs); **B** for transcription factors (TFs). Red and blue indicate *O. longilobus* and *O. taihangensis*, respectively

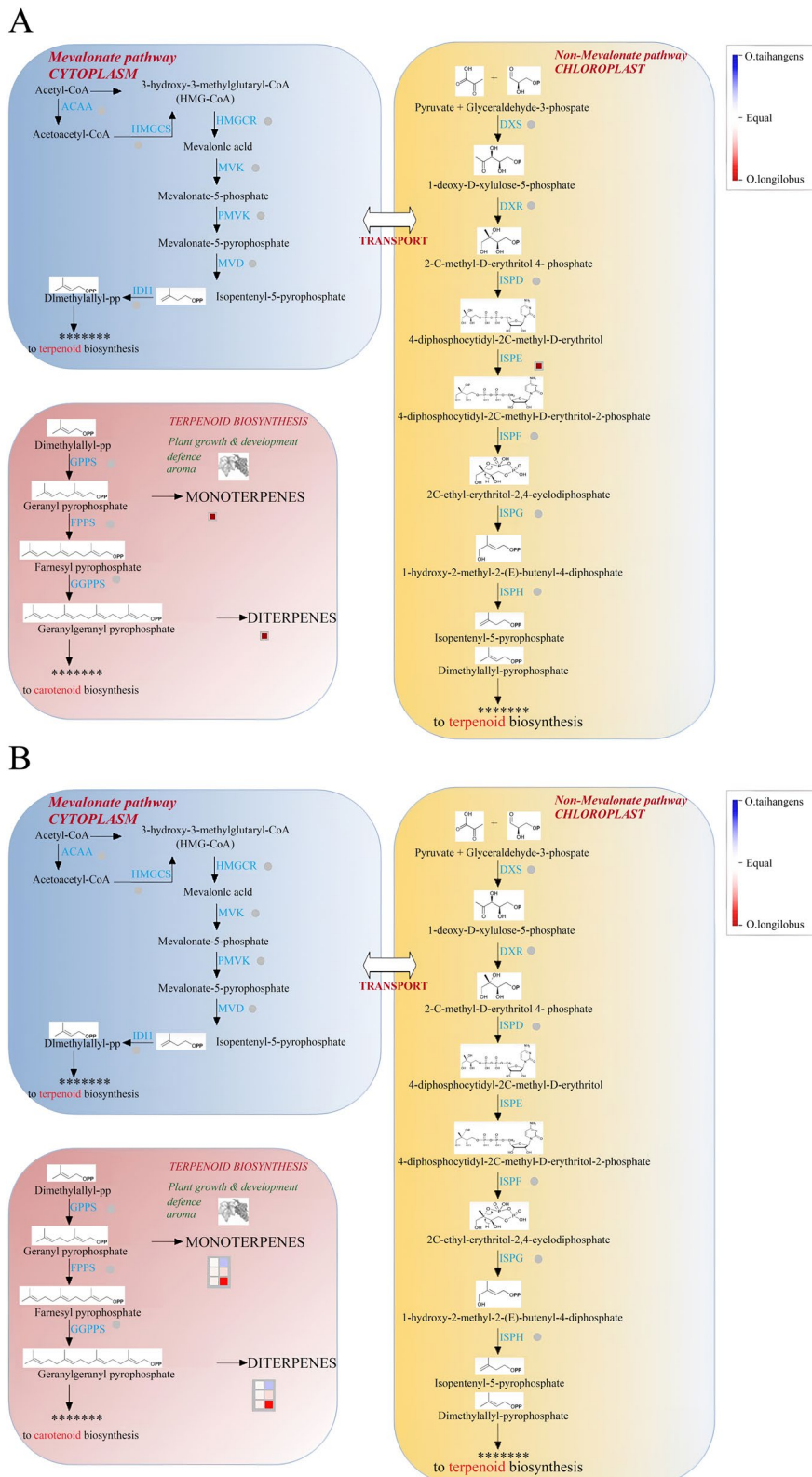


Fig. 7 Terpene biosynthesis pathways. **A:** for positive selection genes (PSGs); **B:** for transcription factors (TFs). Red and blue indicate *O. longilobus* and *O. taihangensis*, respectively

data, we initially obtained 406 TFs (Table 1). The bHLH, FAR1, NAC, MYB, and WRKY family members were highest among the obtained factors. These factors were found to be significantly expressed for lipid, secondary, amino acid, and nucleotide metabolism, which constituted the main biological processes. As many TFs were involved in the amino acid metabolism pathways (Fig. 2), the proactive responses to environmental/developmental cues were explicitly depicted in *O. taihangensis*.

In particular, *O. longilobus* had a higher expression level in the terpenoid metabolic pathway than did *O. taihangensis* (Fig. 6). Furthermore, the involved TFs were all bHLH family members (Fig. 6). Basic helix-loop-helix (bHLH) transcription factors, which are one of the largest families and a large superfamily in plants, play relevant roles in a variety of developmental and evolutionarily conserved processes, including cell-fate specification, tissue differentiation, and secondary metabolites [24, 54, 55]. The bHLH factor *VvMYC1* can regulate anthocyanin and/or proanthocyanidin (PA) synthesis of the flavonoid pathway and thus regulates the biosynthesis of terpenoids [56]. Several bHLH TFs within terpenoid biosynthesis pathway have been found in *Medicago truncatula* [57], *Chenopodium quinoa* [58], *Panax notoginseng* [59], and *Panax ginseng* [60]. The differential expression of bHLH family in the two *Opisthopappus* species indicated partly adaptive differentiation and variation.

Terpenoids are one of the most common compounds in the secondary metabolites of plants, which have numerous physiological and ecological functions, such as attracting pollinated insects, regulating growth and development, resisting environmental stress, and participating in the defense of pests [61]. Meanwhile terpenoids are important components of the fragrances of flowers. For the two species of *Opisthopappus* under study, these secondary metabolites are primarily emitted by their leaves and flowers. According to the Flora Reipublicae Popularis Sinicae, the surfaces of *O. longilobus* leaves are smooth and glabrous, while *O. taihangensis* is sparsely pubescent on both leaf surfaces, which include glandular trichomes and non-glandular trichomes [62]. The terpenoids produced in *O. taihangensis* can be released through glandular trichomes which give rise to the unique aroma for communication and defense [62]. For *O. longilobus*, owing to the lack of glandular trichomes on the surfaces of its leaves, terpenoid biosynthesis genes were more highly expressed to compensate for this deficit through the use of alternative emission strategies.

The other side, many bHLH genes have been shown to respond to various forms of stress such as drought, salt, and cold stresses [52, 57, 63]. Other TFs (e.g., NAC, MYB, and WRKY) have been investigated for their capacities to improve tolerance and resistance in many plants [64–68],

while participating in abiotic and biotic stress responses [18, 69, 70]. The expression of these TFs was significantly different between the two *Opisthopappus* species in overview of the metabolic pathways (Fig. 2) and secondary metabolic pathways of transcription factors (TFs) (Fig. 6B). This suggested that the differential expression of TFs (particularly bHLH family members) might play critical roles in adaptation of the two *Opisthopappus* species to heterogeneous environments.

Adaptability of PSGs

The average dN/dS ratios across the pairs of the *Opisthopappus* species were much lower than 0.9, suggesting that purifying selection had a general influence on the evolution of most protein-coding regions (ORF) of the two species, as has been observed in other plants [17, 71]. Actually, the genes that were most under the influence of purifying selection primarily contained structural or “housekeeping” genes, for example tyrosyl-tRNA synthetase. As these genes are involved in processes that are crucial for organisms, purifying selection can eliminate deleterious, nonsynonymous mutations [17, 72, 73]. Among the 38,986 orthologs shared between lineages, 1203 gene pairs significantly exhibited dN/dS ratios > 1.5 ($P < 0.05$) (Fig. 3A, 3B). These genes were involved in several biological functions (e.g., metabolic processes) and may constitute candidates that are under the effect of positive selection; thus, potentially associated with species divergence.

Adaptive divergence at molecular level may be reflected by an increased rate of non-synonymous changes within genes involved in adaptation [74]. Analyses of the PSGs identified from *O. taihangensis* and *O. longilobus* yielded significantly different landscapes of biological processes, cellular component and molecular function (Fig. 4). *O. taihangensis* and *O. longilobus* might experience divergent adaptation during their evolution, and genes related to adaptation was under rapid evolution and/or have signs of positive selection. These findings implicated that *Opisthopappus* were experienced the ongoing accelerated evolution under different environment. Thus, more than 1000 genes under positive selection between the two *Opisthopappus* species might have played a role in shaping the divergence of this genus, given that *O. longilobus* and *O. taihangensis* are exposed to sub-humid warm temperate, and temperate continental monsoon climates, respectively.

The comparative transcriptome analysis [28] indicated that 3,410 differentially expressed genes were mainly involved in lipid, carbohydrate and amino acid metabolism, xenobiotics biodegradation and metabolism. In this study, KEGG enrichment and MapMan analyses

suggested that significant GO terms of 1203 differentially expressed positive selection genes mainly related to metabolic regulation (Fig. 4B) were important among the *Opisthopappus* species (Fig. 5). Secondary metabolism was also significantly represented for *O. longilobus* and *O. taihangensis* in our PSG analysis (Fig. 6).

Beta-amylase, a member of family 14 of glycosyl hydrolases [75], can hydrolyze the α -1,4-glycosidic linkages in starch, removing successive maltose units from the non-reducing ends of the chains [76]. Starch is the main form of carbon storage in higher plants and it accumulates in different organs, which can be consumed by the cell in which they are produced, transported to nonphotosynthetic sink tissues, or stored for later use. According to studies using transgenic and mutant plants, beta-amylase seems to be important for normal degradation of the transient starch accumulated in chloroplasts [76]. Meanwhile, beta-amylases possibly facilitate the rapid starch degradation under heat and drought [77]. In this study, we found that the PSG expression levels involved in beta-amylase for *O. longilobus* were significantly higher than those for *O. taihangensis* (Fig. 5A). Starch degradation as a major response in conditions prohibiting photosynthesis or prolonged drought [77], therefore, *O. longilobus* seems better adapted to weak light or drought conditions than *O. taihangensis*.

Beta-amylase can convert starch into maltose and glucose as the predominant form in which sugars are translocated in plants presented in Fig. 5C. Consistent with this notion, *O. longilobus* exhibited higher levels of cellular glucose and fructose than did *O. taihangensis*. Collectively, these data signified that the higher expression levels of beta-amylase genes in *O. longilobus* might accelerate its cell growth processes (in contrast to *O. taihangensis*), which have important roles in the evolution of the two species. Ultimately, glycan degradation resulted in the accumulation of sucrose, which is critical for the protection of plants against xenobiotic and oxidative stresses [18], and which might facilitate the enhanced survival of *O. longilobus* for relatively more drought habitats over than *O. taihangensis*.

The electron transport chain (ETC, respiratory chain) is a series of protein complexes that transfer electrons from electron donors to electron acceptors via redox reactions (both reduction and oxidation occurring simultaneously) and couples this electron transfer with the transfer of protons (H^+ ions) across a membrane. The electron transport chain is built up of peptides, enzymes, and other molecules. In this study, the expression levels of relative PSGs involved in complex I (NADH ubiquinone oxidoreductase) and complex IV (cytochrome c oxidase, EC 1.9.3.1) for *O. longilobus* were higher than those for *O. taihangensis* (Fig. 5B). The

ETC sustains the major mitochondrial function of ATP generation, in relation to the metabolic dynamics of tricarboxylic acids (TCAs), acetyl-CoA, ADP, oxidized (NAD^+) or reduced (NADH) β -nicotinamide adenine dinucleotide, oxidized (FAD) or reduced (FADH₂) flavin adenine dinucleotide [78]. Thus, we concluded that the changes in ETC can affect nicotinate and nicotinamide metabolism, glycan degradation metabolism and transmembrane transport activity, which was supported by the KEGG and MapMan results (Figs. 4 and 5). The various expression levels of relative PSGs between these two species were also regarded as different adaptation signals to habitats.

Additionally, the responses to sulfur between *O. longilobus* and *O. taihangensis* were significantly different, which suggested that the two species possessed different stress tolerances for S in response to heterogeneous environments.

Adaptive function of Terpenoids

O. taihangensis and *O. longilobus* secrete chemicals and exude aromas to protect against biological and abiotic stresses [62]. These chemicals are primarily secondary metabolites, including polyphenols, alkaloids, and terpenoids. This scenario, particularly in terms of terpenoid biosynthesis, was unambiguously observed in the present study as a difference between the two species TFs and PSGs (Fig. 6). The terpenoid pathway is intricately regulated by endogenous and environmental factors that enable spatially and temporally controlled metabolite production [79–81]. Terpenes comprise the largest and the most diverse class of natural products, often play important roles in plant defenses by attracting the enemies or predators of herbivores and repelling herbivorous insect feeders [82].

Isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP) are the main intermediate compounds in the terpenoid pathway for terpene biosynthesis [20, 83]. Their precursor molecules (C₅) are generated via the process of isoprenogenesis [84, 85]. In plants, isoprenogenesis occurs through two discrete biosynthetic pathways: the mevalonic acid (MVA) pathway in the cytosol, and the 2-C-methyl-D-erythritol 4-phosphate/1-deoxy-D-xylulose 5-phosphate (MEP/DOXP) pathway within plastids. In the MEP pathways, the C₅ units are catalyzed step by step via a series of enzymes to synthesize IPP. The 4-diphosphocytidyl-2C-methyl-D-erythritol synthase (ISPD/CMS) and 4-diphosphocytidyl-2C-methyl-D-erythritol kinase (ISPE/CMK) catalyze the synthesis of 4-diphosphocytidyl-2C-methyl-D-erythritol (CDP-ME) and 4-diphosphocytidyl-2C-methyl-D-erythritol 2-phosphate (CDP-ME₂P), respectively. CDP-ME and

CDP-ME2P are the intermediate compounds of IPP biosynthesis.

IPP or DMAPP may be linked through head to tail condensation reactions to generate terpenes of different classes (e.g., mono, di, and triterpenes), which are catalyzed by geranyl diphosphate synthase (GPPS) to synthesize geranyl pyrophosphate (GPP). The GPP results in farnesyl pyrophosphate (FPP) and geranylgeranyl diphosphate (GGPP) via farnesyl pyrophosphate synthase (FPPS) and geranylgeranyl diphosphate synthase (GGPPS) enzymes, respectively. GPP and GGPP are substrates for monoterpene and diterpene biosynthesis in the terpene pathway [86–88]. Monoterpenes and diterpenes can impart the distinct flavors and aromas of plants [20, 89]. These terpenes are produced in plants as secondary bioactive metabolites, often for ecological adjustment and protection from microbial pathogens, fungi, pests, and predation [90]. An overview of terpenoid metabolism of the two species revealed that the putative PSG displayed different levels of expression for IPP biosynthesis, particularly in the synthesis of CDP-MK and CDP-ME2K (Fig. 7).

Meanwhile, both the PSGs and TFs involved in the downstream steps of the MEP pathway exhibited a relatively higher expression in contrast to the two species under study, which displayed different expression levels for monoterpene and diterpene biosynthesis (Fig. 7). These results indicated variable adaptability in the responses of the two species to different environments. During the evolutionary process, the genes (such as trinity_dn129788_c0_g1, annotated as terpene synthase 14, TPS14) involved in terpenoid biosynthesis pathway (in particular, monoterpene and diterpene biosynthesis), begun to diverge from their expression levels under different long-term environment stresses. Simultaneously, several bHLH TFs involved this metabolic process to further regulate the expression of these positive selection genes. Ultimately, the adaptive phenotypic characteristics occurred in *O. longilobus* and *O. taihangensis*, such as the presence of trichomes or not, and different fragrances.

Conclusion

This work presented a comparative genome-wide transcriptome analysis between two *Opisthopappus* species, toward the identification of biological processes and functional genes that facilitate differentiation and adaption of these two species during evolutionary processes. Under natural selection, *O. taihangensis* gradually modified its gene expression model (such as mitochondrial electron transport and starch degradation) to adapt to its surroundings, which differed

from *O. longilobus*. The screened several PSGs and TFs involved in the terpenoid biosynthesis of monoterpenoids and diterpenoids, and represented adaptive signal responses to heterogeneous environments. These results will not only shed light on how differentiations between the two *Opisthopappus* species occur, but also open the door to an increased understanding of how plants thrive in mountains environments. Additionally, our research lays a theoretical foundation for the further investigation of the molecular mechanisms of specific characteristics that have important academic and application value for other Asteraceae species.

Acknowledgements

This study was supported by the National Natural Science Foundation of China (31970358), the Research Project Supported by Shanxi Scholarship Council of China (2020-090) and the Graduate Education Innovation Project of Shanxi (2020SY323).

Author's contributions

Chen N, Genlou Sun and Yiling Wang designed experiments; Chen N and Yiling Wang carried out experiments, analyzed experimental results and data, developed analysis tools and wrote the main manuscript text. Hao Zhang, En Zang, Zhixia Liu, Yafei Lan, Weili Hao, Shan He, Xing Fan, Genlou Sun, Yiling Wang assisted with writing the manuscript. All authors reviewed the manuscript. The authors read and approved the final manuscript.

Funding

This study was supported by the National Natural Science Foundation of China (31970358), the Research Project Supported by Shanxi Scholarship Council of China (2020–090), the Natural Science Foundation of Shanxi Province, China (20210302124501) and the Graduate Education Innovation Project of Shanxi (2020SY323).

Availability of data and materials

Data deposited via the Genebank: accession SUB4002238: PRJNA471103.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

No conflict of interest exists in the submission of this manuscript, and the manuscript is approved by all authors for publication.

Author details

¹College of Life Science, Shanxi Normal University, Taiyuan 030031, China.

²Triticeae Research Institute, Sichuan Agricultural University, Chengdu 611130, China.

³Department of Biology, Saint Mary's University, Halifax B3H3C3, Canada.

Received: 27 October 2021 Accepted: 13 June 2022

Published online: 24 June 2022

References

1. Ferchaud AL, Hansen MM. The impact of selection, gene flow and demographic history on heterogeneous genomic divergence: threespine sticklebacks in divergent environments. *Mol Ecol*. 2016;25(1):238–59.
2. Svensson EI, Goedert D, Gómez-Llano MA, Spagopoulou F, Nava-Bolaños A, Booksmythe I. Sex differences in local adaptation: what

- can we learn from reciprocal transplant experiments? *Philos T R Soc B*. 2018;373:20170420.
3. Li YL, Xue DX, Zhang BD, Liu JX. Environmental selection in the catadromous roughskin sculpin *Trachidermus fasciatus*. *Genome Biol Evol*. 2019;11(7):1751–64.
 4. Enbody ED, Pettersson ME, Sprehn CG, Palmb S, Wickströmb H, Andersson L. Ecological adaptation in European eels is based on phenotypic plasticity[J]. *Proceedings of the National Academy of Sciences*. 2021;118(4):e2022620118.
 5. Zhang X, Rayner JG, Blaxter M, Bailey NW. Rapid parallel adaptation despite gene flow in silent crickets. *Nat Commun*. 2021;12:50.
 6. Delph LF. The study of local adaptation: a thriving field of research. *J Hered*. 2017;109:1–2.
 7. Ranade SS, García-Gil MR. Molecular signatures of local adaptation to light in Norway spruce. *Planta*. 2021;253(2):53.
 8. Honjo M.N, Kudoh H. *Arabidopsis halleri*: a perennial model system for studying population differentiation and local adaptation. *AoB Plants*. 2019;11:plz076.
 9. Olson-Manning CF, Wagner MR, Mitchell-Olds T. Adaptive evolution: evaluating empirical support for theoretical predictions[J]. *Nat Rev Genet*. 2012;13(12):867.
 10. Kawecki TJ, Ebert D. Conceptual issues in local adaptation. *Ecol Lett*. 2004;7:1225–41.
 11. Hereford J. A quantitative survey of local adaptation and fitness trade-offs. *Am Nat*. 2009;173:579–88.
 12. Whitlock MC. Modern approaches to local adaptation. *Am Nat*. 2015;186:51–4.
 13. Zhang HB, Xia EH, Huang H, Jiang JJ, Liu BY, Gao LZ. De novo transcriptome assembly of the wild relative of tea tree (*Camellia taliensis*) and comparative analysis with tea transcriptome identified putative genes associated with tea quality and stress response. *BMC Genomics*. 2015;16:298.
 14. Yan L, Liu Z, Xu H, Zhang X, Zhao A, Liang F, Ni Z. Transcriptome analysis reveals potential mechanisms for different grain size between natural and resynthesized allohexaploid wheats with near-identical AABB genomes. *BMC Plant Biol*. 2018;18:28.
 15. Kim J, Park SJ, Lee IH, Chu H, Penfold CA, Kim JH, Lim PO. Comparative transcriptome analysis in *Arabidopsis ein2/ore3* and *ahk3/ore12* mutants during dark-induced leaf senescence. *J Exp Bot*. 2018;69:3023–36.
 16. Li J, Ye C, Chang C. Comparative transcriptomics analysis revealing flower trichome development during flower development in two *Lonicera japonica*. *BMC Plant Biol*. 2020;20:341.
 17. Wang Y, Jiang W, Ye W, Fu C, Gitzendanner MA, Soltis PS, Qiu Y. Evolutionary insights from comparative transcriptome and transcriptome-wide coalescence analyses in *Tetragonia hemsleyana*. *BMC Plant Biol*. 2018;18:208.
 18. Ma Z, Lin S. Transcriptomic revelation of phenolic compounds involved in aluminum toxicity responses in roots of *Cunninghamia lanceolata* Hook. *Genes*. 2019;10:835.
 19. Onda Y, Mochida K, Yoshida T, Sakurai T, Seymour RS, Umekawa Y, Ito K. Transcriptome analysis of thermogenic *arum concinatum* reveals the molecular components of floral scent production. *Scie Rep*. 2015;5:8753.
 20. Torre S, Tattini M, Brunetti C, Guidi L, Gori A, Marzano C, Landi M. De novo assembly and comparative transcriptome analyses of red and green morphs of sweet basil grown in full sunlight. *PLoS ONE*. 2016;11(8):e0160370.
 21. Huang G, Zeng Y, Wei L, Yao Y, Dai J, Liu G, Gui Z. Comparative transcriptome analysis of mulberry reveals anthocyanin biosynthesis mechanisms in black (*Morus atropurpurea* Roxb.) and white (*Morus alba* L.) fruit genotypes. *BMC Plant Biol*. 2020;20:279.
 22. Zhang L, Yan HF, Wu W, Yu H, Ge XJ. Comparative transcriptome analysis and marker development of two closely related Primrose species (*Primula poissonii* and *Primula wilsonii*). *BMC Genomics*. 2013;14:39.
 23. Bonasio R, Zhang G, Ye C, Mutti NS, Fang X, Qin N, Liebig J. Genomic comparison of the ants *Camponotus floridanus* and *Harpegnathos saltator*. *Science*. 2010;329:1068–71.
 24. Bao L, Gao H, Zheng Z, Zhao X, Zhang M, Jiao F, Qian Y. Integrated transcriptomic and un-targeted metabolomics analysis reveals mulberry fruit (*Morus atropurpurea*) in response to sclerotinose pathogen *Ciboria shiraiana* infection. *Int J Mol Sci*. 2020;21:1789.
 25. Shih C. *Opisthopappus shih*-a new genus of compositae from China. *Acta Phytotaxon Sin*. 1979;32:110–2.
 26. Wang FZ, Tang J, Chen XQ, Liang SJ, Dai LK. *Liliaceae*. *Flora of China Editorial Committee of China Academy of Science*. Beijing: Science Press; 1978. p. 73–4.
 27. Ding BZ, Wang SY. *Flora of Henanica* [M]. Zhengzhou: Henan Science and Technology Press, 1998. p. 1789.
 28. Chai M, Wang S, He J, Chen W, Fan Z, Li J, Wang Y. De novo assembly and transcriptome characterization of *Opisthopappus* (*Asteraceae*) for population differentiation and adaptation. *Front Genet*. 2018;9:371.
 29. Ye H, Wang Z, Hou HM, Wu JH, Gao Y, Han W, Wang YL. Localized environmental heterogeneity drives the population differentiation of two endangered and endemic *Opisthopappus* Shih species. *BMC Ecol Evo*. 2021;21:56.
 30. Wang YL, Yan G.Q. Molecular phylogeography and population genetic structure of *O. longilobus* and *O. taihangensis* (*Opisthopappus*) on the Taihang Mountains. *PLoS One*. 2014;9:e104773.
 31. Wang YL. Chloroplast microsatellite diversity of *Opisthopappus shih* (*Asteraceae*) endemic to China. *Plant Syst Evol*. 2013;299:1849–58.
 32. Wang YL, Yan GQ. Genetic diversity and population structure of *Opisthopappus longilobus* and *Opisthopappus taihangensis* (*Asteraceae*) in China determined using sequence related amplified polymorphism markers. *Biochem Syst Ecol*. 2013;49:115–24.
 33. Wang YL, Zhang CQ, Lin LL, Yuan LH. Its sequence analysis of *Opisthopappus taihangensis* and *Opisthopappus longilobus*. *Acta Hortic Sin*. 2015;42:178–86.
 34. Chai M, Ye H, Wang Z, Zhou Y, Wu J, Gao Y, Wang Y. Genetic divergence and relationship among *Opisthopappus* species identified by development of EST-SSR markers. *Front Genet*. 2020;11:177.
 35. Jin J, Tian F, Yang D.C, Meng Y.Q, Kong L, Lou J, Gao G. PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Res*. 2017;45(D1):D1040–1045.
 36. Bathke J, Konzer A, Remes B, McIntosh M, Klug G. Comparative analyses of the variation of the transcriptome and proteome of *Rhodobacter sphaeroides* throughout growth. *BMC Genomics*. 2019;20:358.
 37. Usadel B, Poree F, Nagel A, et al. A guide to using MapMan to visualize and compare omics data in plants: a case study in the crop species, Maize[J]. *Plant Cell Environ*. 2010;32(9):1211–29.
 38. Jung KH, An G. Application of MapMan and rice net drives systematic analyses of the early heat stress transcriptome in rice seedlings. *J Plant Biol*. 2012;55(6):436–49.
 39. Chandran AK, Lee GS, Yoo YH, Yoon UH, Ahn BO, Yun DW, Jung KH. Functional classification of rice flanking sequence tagged genes using MapMan terms and global understanding on metabolic and regulatory pathways affected by *dxr* mutant having defects in light response. *Rice*. 2016;9:17.
 40. Rathinam M, Mishra P, Vasudevan M, Budhwar R, Mahato A, Prabha A.L, Sreevathsa R. Comparative transcriptome analysis of pigeonpea, *Cajanus cajan* (L.) and one of its wild relatives *Cajanus platycarpus* (Benth.). *Maesen PLoS One*. 2019;14(7):e0218731.
 41. Thimm O, Bläesing O, Gibon Y, Nagel A, Meyer S, Krüger P, Stitt M. MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. *Plant J*. 2004;37:914–39.
 42. Zhang Z, Li J, Zhao XQ, Wang J, Wong GK, Yu JM. Ka/Ks_Calculator: calculating Ka and Ks through model selection and model averaging. *Genom Proteom Bioinf*. 2006;4:259–63.
 43. Wang D, Zhang Y, Zhang Z, Zhu J, Yu J. Ka/Ks_Calculator 2.0: a toolkit incorporating gamma-series methods and sliding window strategies. *Genom Proteom Bioinform*. 2010;8:77–80.
 44. Yang Z, Nielsen R. Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models. *Mol Biol Evol*. 2000;17:32–43.
 45. Minaya M, Díaz-Pérez A, Mason-Gamer R, Pimentel M, Catalán P. Evolution of the beta-amylase gene in the temperate grasses: non-purifying selection, recombination, semiparalogy, homeology and phylogenetic signal [J]. *Mol Phylogenet Evol*. 2015;91:68–85.
 46. Langmead B, Salzberg S. Fast gapped-read alignment with bowtie 2. *Nat Methods*. 2012;9:357–9.

47. Blüthgen N, Brand K, Cajavec B, Swat M, Herzel H, Beule D. Biological profiling of gene groups utilizing gene ontology. *Genome Inform.* 2005;16:106–15.
48. Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics.* 2005;21:3674–6.
49. Kanehisa M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 2000;28:27–30.
50. Hanley ME, Lamont B, Fairbanks MM, Rafferty CM. Plant structural traits and their role in anti-herbivore defence. *Perspect Plant Ecol.* 2007;8:157–78.
51. Lohse M, Nagel A, Herter T, May P, Schroda M, Zrenner R, Usadel B. Mercator: a fast and simple web server for genome scale functional annotation of plant sequence data. *Plant Cell Environ.* 2014;37:1250–8.
52. Niu X, Guan Y, Chen S, Li H. Genome-wide analysis of basic helix-loop-helix transcription factors in *Brachypodium distachyon*. *BMC Genomics.* 2017;18:619.
53. Gu H, Yang Y, Xing M, Yue C, Wei F, Zhang Y, Huang J. Physiological and transcriptome analyses of *Opisthoppappus taihangensis* in response to drought stress. *Cell Biosci.* 2019;9:56.
54. Gyoja F. Basic helix-loop-helix transcription factors in evolution: Roles in development of mesoderm and neural tissues. *Genesis.* 2017;55(9):e23051.
55. Zheng Y, Wang P, Chen X, Sun Y, Yue C, Ye N. Transcriptome and metabolite profiling reveal novel insights into volatile heterosis in the tea plant (*Camellia sinensis*). *Molecules.* 2019;24:3380.
56. Hichri I, Heppel SC, Pillet J, Léon G, Czemplin S, Delrot VL, Bogs J. The basic helix-loop-helix transcription factor MYC1 is involved in the regulation of the flavonoid biosynthesis pathway in grapevine. *Mol Plant.* 2010;3:509–23.
57. Mertens J, Pollier J, Bossche RV, Lopez-Vidriero I, Franco-Zorrilla JM, Goossens A. The bHLH transcription factors TSAR1 and TSAR2 regulate triterpene saponin biosynthesis in *Medicago truncatula*. *Plant Physiol.* 2016;170:194–210.
58. Jarvis DE, Ho YS, Lightfoot DJ, Schmöckel SM, Li B, Borm TJA, Tester M. The genome of *Chenopodium quinoa*. *Nature.* 2017;542:307–12.
59. Zhang X, Ge F, Deng B, Shah T, Huang Z, Liu D, Chen C. Molecular cloning and characterization of PnbHLH1 transcription factor in *Panax notoginseng*. *Molecules.* 2017;22:1268.
60. Chu Y, Xiao S, Su H, Liao B, Zhang J, Xu J, Chen S. Genome-wide characterization and analysis of bHLH transcription factors in *Panax ginseng*. *Acta Pharm Sin B.* 2018;8(4):666–77.
61. Dudareva N, Klemplien A, Muhlemann JK, Kaplan I. Biosynthesis, function and metabolic engineering of plant volatile organic compounds. *New Phytol.* 2013;198:16–32.
62. Guo Y, Zhang T, Zhong J, Tingting Ba, Xu T, Zhang Q, Sun M. Identification of the volatile compounds and observation of the glandular trichomes in *Opisthoppappus taihangensis* and four species of *Chrysanthemum*. *Plants.* 2020;9:855.
63. Miao L, Gao Y, Zhao K, Kong L, Yu S, Li R, Yu X. Comparative analysis of basic helix-loop-helix gene family among *Brassica oleracea*, *Brassica rapa*, and *Brassica napus*. *BMC Genomics.* 2020;21:178.
64. Chen L, Song Y, Li S, Zhang L, Zou C, Yu D. The role of WRKY transcription factors in plant abiotic stresses. *BBA-Gene Regul Mech.* 2012;1819:120–8.
65. Liu C, Mao B, Ou S, Wang W, Liu L, Wu Y, Wang X. OsbZIP71, a bZIP transcription factor, confers salinity and drought tolerance in rice. *Plant Mol Biol.* 2014;84:19–36.
66. Liu J, Osbourn A, Ma P. MYB transcription factors as regulators of phenylpropanoid metabolism in plants. *Mol Plant.* 2015;8:689–708.
67. Zhang X, Zhang B, Li MJ, Yin XM, Huang LF, Cui YC, Xia X. OsMSR15 encoding a rice C2H2-type zinc finger protein confers enhanced drought tolerance in transgenic *Arabidopsis*. *J Plant Biol.* 2016;59:271–81.
68. Yoo YH, Chandran NAK, Park JC, Cho YS, Lee SW, An G, Jung KH. OsPhyB-mediated novel regulatory pathway for drought tolerance in rice root identified by a global RNA-Seq transcriptome analysis of rice genes in response to water deficiencies. *Front Plant Sci.* 2017;8:850.
69. Kumari M, Taylor GJ, Deyholos MK. Transcriptomic responses to aluminum stress in roots of *Arabidopsis thaliana*. *Mol Genet Genomics.* 2008;279:339–57.
70. Tang N, Zhang H, Li X, Xiao J, Xiong L. Constitutive activation of transcription factor OsbZIP46 improves drought tolerance in rice. *Plant Physiol.* 2012;158:1755–68.
71. Buschiazzo E, Ritland C, Bohlmann J, Ritland K. Slow but not low: genomic comparisons reveal slower evolutionary rate and higher dN/dS in conifers compared to angiosperms. *BMC Evol Biol.* 2012;12:8.
72. Hughes AL, Packer B, Welch R, Bergen AW, Chanock SJ, Yeager M. Widespread purifying selection at polymorphic sites in human protein-coding loci. *P Natl Acad Sci USA.* 2003;100:15754–7.
73. Renaut S, Nolte AW, Bernatchez L. Mining transcriptome sequences towards identifying adaptive single nucleotide polymorphisms in lake whitefish species pairs (*Coregonus* spp. Salmonidae). *Mol Ecol.* 2010;19:115–31.
74. Qiu Q, Zhang G, Ma T, Qian W, Wang J, Ye Z, Cao C, Hu Q, Kim J, Larkin DM. The yak genome and adaptation to life at high altitude. *Nat Genet.* 2012;44(8):946–9.
75. Henrissat B. A classification of glycosyl hydrolases based on amino-acid sequence similarities. *Biochem J.* 1991;280:309–16.
76. Lloyd JR, Kossmann J, Ritte G. Leaf starch degradation comes out of the shadows. *Trends Plant Sci.* 2005;10(3):130–7.
77. Tsamir-Rimon M, Ben-Dor S, Feldmesser E, Oppenheimer-Shaanan Y, David-Schwartz R, Samach A, Klein T. Rapid starch degradation in the wood of olive trees under heat and drought is permitted by three stress-specific beta amylases. *New Phytol.* 2021;229(3):1398–414.
78. Bohovych I, Khalimonchuk O. Sending out an SOS: mitochondria as a signaling hub. *Front Cell Dev Biol.* 2016;4:109.
79. Nieuwenhuizen NJ, Chen X, Wang MY, Matic AJ, Perez RL, Allan AC, Atkinson RG. Natural variation in monoterpene synthesis in kiwifruit: transcriptional regulation of terpene synthases by NAC and ETHYLENEINSENSITIVE3-like transcription factors. *Plant Physiol.* 2015;167:1243–58.
80. Wang Q, Reddy VA, Panicker D, Mao HZ, Kumar N, Rajan C, Sarojam R. Metabolic engineering of terpene biosynthesis in plants using a trichomespecific transcription factor MsYABBY5 from spearmint (*Mentha spicata*). *Plant Biotechnol J.* 2016;14:1619–32.
81. Costantini L, Kappel C.D, Trenti M, Battilana J, Emanuelli F, Sordo M, Grando M.S. Drawing links from transcriptome to metabolites: the evolution of aroma in the ripening berry of Moscato Bianco (*Vitis vinifera* L.). *Front Plant Sci.* 2017;8:780.
82. Niu L, Pan L, Zeng W, Lu Z, Cui G, Fan M, Li G. Dynamic transcriptomes of resistant and susceptible peach lines after infestation by green peach aphids (*Myzus persicae* Sulzer) reveal defence responses controlled by the Rm3 locus. *BMC Genomics.* 2018;19:846.
83. Chang W, Song H, Liu H, Liu P. Current development in isoprenoid precursor biosynthesis and regulation. *Curr Opin Chem Biol.* 2013;17:571–9.
84. Karami A, Khorshidi R, Jowkar A. Diurnal changes in the volatile oil compounds of rose-scented geranium (*Pelargonium graveolens*). *Anal Chem Lett.* 2015;5:103–8.
85. Ravindra NS, Kulkarni R. Essential oil yield and quality in rose-scented geranium: Variation among clones and plant parts. *Sci Hortic.* 2015;184:31–5.
86. Tholl D. Terpene synthases and the regulation, diversity and biological roles of terpene metabolism. *Curr Opin Plant Biol.* 2006;9:297–304.
87. Cheng AX, Lou YG, Mao YB, Lu S, Wang LJ, Chen XY. Plant terpenoids: biosynthesis and ecological functions. *J Integr Plant Biol.* 2007;49:179–86.
88. Narnoliya LK, Kaushal G, Singh SP, Sangwan RS. De novo transcriptome analysis of rosescented geranium provides insights into the metabolic specificity of terpene and tartaric acid biosynthesis. *BMC Genomics.* 2017;18:74.
89. Gang DR. An investigation of the storage and biosynthesis of phenylpropenes in sweet basil. *Plant Physiol.* 2001;125:539–55.
90. Nazzaro F, Fratianni F, De-Martino L, Coppola R, De-Feo V. Effect of essential oils on pathogenic bacteria. *Pharmaceuticals.* 2013;6:1451–74.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.