



Comparative Transcriptomics and Metabolites Analysis of Two Closely Related *Euphorbia* Species Reveal Environmental Adaptation Mechanism and Active Ingredients Difference

OPEN ACCESS

Edited by:

Xueqing Fu,
Shanghai Jiao Tong University, China

Reviewed by:

Ling Li,
Shanghai Jiao Tong University, China
Xiaojian Yin,
China Pharmaceutical
University, China
Xu Lu,
China Pharmaceutical
University, China

*Correspondence:

Lu-Qi Huang
huangluqi01@126.com
Xiu-Teng Zhou
zxt_0508@163.com
Ye Shen
shenye@nrc.ac.cn

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Plant Metabolism and Chemodiversity,
a section of the journal
Frontiers in Plant Science

Received: 26 March 2022

Accepted: 02 May 2022

Published: 31 May 2022

Citation:

Zheng H, Yu M-Y, Han Y, Tai B, Ni S-F,
Ji R-F, Pu C-J, Chen K, Li F-Q, Xiao H,
Shen Y, Zhou X-T and Huang L-Q
(2022) Comparative Transcriptomics
and Metabolites Analysis of Two
Closely Related *Euphorbia* Species
Reveal Environmental Adaptation
Mechanism and Active Ingredients
Difference.
Front. Plant Sci. 13:905275.
doi: 10.3389/fpls.2022.905275

Han Zheng^{1†}, Mu-Yao Yu^{1†}, Yang Han^{1†}, Badalahu Tai^{1,2}, Sheng-Fa Ni³, Rui-Feng Ji¹,
Chun-Juan Pu¹, Kang Chen¹, Fu-Quan Li⁴, Hua Xiao⁴, Ye Shen^{1*}, Xiu-Teng Zhou^{1*} and
Lu-Qi Huang^{1*}

¹ State Key Laboratory of Dao-di Herbs, National Resource Center for Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing, China, ² Mongolian Medicine College, Inner Mongolia Minzu University, Tongliao, China, ³ Anhui University of Science and Technology, Huainan Xinhua Hospital, Huainan, China, ⁴ Hulunbeier Mongolian Medical Hospital, Hulunbeier, China

Roots of *Euphorbia fischeriana* and *Euphorbia ebracteolata* are recorded as the source plant of traditional Chinese medicine “Langdu,” containing active ingredients with anticancer and anti-AIDS activity. However, the two species have specific patterns in the geographic distribution. Compared with *E. ebracteolata*, *E. fischeriana* distributes in higher latitude and lower temperature areas and might have experienced cold stress adaptation. To reveal the molecular mechanism of environmental adaptation, RNA-seq was performed toward the roots, stems, and leaves of *E. fischeriana* and *E. ebracteolata*. A total of 6,830 pairs of putative orthologs between the two species were identified. Estimations of non-synonymous or synonymous substitution rate ratios for these orthologs indicated that 533 of the pairs may be under positive selection ($Ka/Ks > 0.5$). Functional enrichment analysis revealed that significant proportions of the orthologs were in the TCA cycle, fructose and mannose metabolism, starch and sucrose metabolism, fatty acid biosynthesis, and terpenoid biosynthesis providing insights into how the two closely related *Euphorbia* species adapted differentially to extreme environments. Consistent with the transcriptome, a higher content of soluble sugars and proline was obtained in *E. fischeriana*, reflecting the adaptation of plants to different environments. Additionally, 5 primary or secondary metabolites were screened as the biomarkers to distinguish the two species. Determination of 4 diterpenoids was established and performed, showing jolkinolide B as a representative component in *E. fischeriana*, whereas ingenol endemic to *E. ebracteolata*. To better study population genetics, EST-SSR markers were generated and tested in 9 species of *Euphorbia*. A total of 33 of the 68 pairs were screened out for producing clear fragments in at least four species, which will furthermore facilitate the studies on the genetic improvement and phylogenetics of this rapidly adapting taxon. In this study, transcriptome and metabolome analyses revealed the evolution of genes related to cold stress tolerance,

biosynthesis of TCA cycle, soluble sugars, fatty acids, and amino acids, consistent with the molecular strategy that genotypes adapting to environment. The key active ingredients of the two species were quantitatively analyzed to reveal the difference in pharmacodynamic substance basis and molecular mechanism, providing insights into rational crude drug use.

Keywords: *Euphorbia*, comparative transcriptomics, metabolome, environmental adaptation, diterpenoid, EST-SSR

INTRODUCTION

Euphorbia is the largest genus in the Euphorbiaceae family, consisting of more than 2,000 species, with 77 species distributed in China (Li et al., 2008). Roots of *Euphorbia fischeriana* Steud (Li et al., 2021b) and *Euphorbia ebracteolata* Hayata (Yang et al., 2021) were used as the same kind of traditional Chinese medicine “Lvru” (now known as Langdu) for more than 2,000 years of treating swelling and ulcer of scabies (Zhao et al., 1996). As main active ingredients of Langdu, jolkinolide B (abietane type, multicyclic diterpenoid) and its derivatives have been proven to display potent anticancer activity (Wang et al., 2009, 2017a); ingenol’s esters (ingenane type, bicyclic diterpenoid) have a great potential in treating human immunodeficiency virus (HIV) (Miana et al., 2015; Huang et al., 2019) and actinic keratosis (Parker et al., 2017), such as prostratin, Picato.

Although *E. fischeriana* and *E. ebracteolata* have been used interchangeably, the graphic distribution pattern of these two Langdu species is different (Li et al., 2008). Compared with *E. ebracteolata* growing in eastern and central China, *E. fischeriana* growing in northern and northeast China (higher latitude) has the habitat characteristics of lower annual average temperature (data from www.nmc.cn). In particular, *E. fischeriana* will be subjected to a strong cold stress in winter. Under this long-term environmental factor, the two species exhibit typical patterns of adaptive evolution and explosive speciation. From a genetic perspective, during speciation among closely related species, genes evolving rapidly are more differentiated than the rest, which is thought to be responsible for habitat differentiation and adaptation (Zhang et al., 2013; Mao et al., 2016; Zhao et al., 2016). However, few genomic resources are available for *E. fischeriana* and *E. ebracteolata*, leading to the unavailability of positive selection gene detection and the study of important loci. A few studies have been reported toward mechanism of ontogenesis (Prenner and Rudall, 2007), evolution of major structural characters (Horn et al., 2012), and diversity of species (Frajman et al., 2019) in *Euphorbia* (Euphorbiaceae), but neither of the two Langdu species was involved. Although the transcriptome of *E. fischeriana* has been sequenced (Barrero et al., 2011), it is not enough to explain the adaptive growth. With paucity of genetic data such as genome sequences and associated molecular markers, stress resistance or evolutionary analysis toward Langdu remains a challenge. In addition, studies have reported that the accumulation of compounds including soluble sugar, amino acids, and fatty acids contributes to adaptation toward abiotic stress (Wanner and Junttila, 1999; Duan et al., 2012; Wei et al., 2018). So far, chemical studies on Langdu have

only been conducted on the active components of the drug, and few studies have been conducted on the compounds coping with stress resistance.

Transcriptome, characterized by its low cost and high efficiency, can not only provide additional genome resources and information about the process of speciation or adaptive evolution, such as time estimation of divergence, or detection of adaptive genes (Koenig et al., 2013; Zhao et al., 2018), but also provide expressed sequence tag—simple sequence repeat (EST-SSR) markers for species identification and germplasm evaluation (Varshney et al., 2005; Zhao et al., 2019; Li et al., 2021a). Furthermore, using metabolomics to study the metabolic changes under stress can reveal the response mechanism of plants to the changes in external environment or genes (Fiehn, 2002; De Vos et al., 2007). Therefore, we should not only reveal the molecular mechanism of adaptation to abiotic stress, but also study the effect of metabolite accumulation on adaptation to abiotic stress and the difference in active ingredients between the two plants. As the first step toward answering these questions, we obtained transcripts and metabolites for *E. fischeriana* and *E. ebracteolata*, furthermore carrying out a comprehensive analysis. 1) Transcripts of two Langdu species were identified, and their genetic differences were compared, providing additional genetic resources for Langdu breeding or evolutionary analysis; 2) evolutionary dynamics of two species were determined, obtaining an estimated time of differentiation, as well as the characteristics of adaptive evolution between the two species; 3) the differential metabolites of the two species were identified through metabolomics, and the chemical mechanism of adaptation toward stress was analyzed; 4) A UPLC-MS/MS method determining 4 diterpenoids simultaneously was established to detect diterpenoid contents in root and leaf of the two Langdu species; 5) and genus-specific EST-SSR markers based on the two species of Langdu were developed in preliminarily.

MATERIALS AND METHODS

Plant Materials

We collected two Langdu species at fructescence. *E. fischeriana* is from hulun buir (48°34′57.18″N, 119°54′07.56″E, alt.746 m, Inner Mongolia); *E. ebracteolata* is from Jiyuan (35°12′46.56″N, 112°25′55.58″E, alt.708 m, Henan). Fresh leaves, stems, and roots from twelve individual plants for both species were stored in liquid nitrogen until total RNA extraction and metabolites analysis.

Microscope Observation

Roots of *E. fischeriana* and *E. ehracteolata* were first embedded with paraffin and then cut into slices of 5 μm . Slices were dewaxed using xylene for 20 min two times, ethanol for 5 min two times, and 75% ethanol for 5 min and then washed with water. Dewaxed slices were soaked in saffron dye for 2 h and decolorized with 50, 70, and 80% ethanol solution for 5 s. Fast green and ethanol were then used to dye and decolor, respectively. Slices were permeabilized using xylene and mounted with neutral gum. Microscopic features were captured using an Olympus BX51 microscope.

RNA-Seq and Data Analysis

To provide a preliminary indication of genetic variation within species, the high-quality mRNA from leaves, stems, and roots from three individual plants for each species was isolated using Plant RNA Purification Reagent (Invitrogen, Life Technologies, USA) following the manufacturer's protocol. Both RNA-seq library preparation and paired-end sequencing were performed using an Illumina HiSeq 6000 platform (Zheng et al., 2021). All unigenes were first subjected to BLASTX (Altschul et al., 1997) against the National Center for Biotechnology Information (NCBI), including the non-redundant protein (Nr) database and non-redundant nucleotide sequence (Nt) databases with an *E*-value threshold of 10^{-5} . The predicted gene name for each contig was assigned according to the best BLASTX hit. Protein structure and function were annotated with Swiss-Prot and Protein family (Pfam) database (Finn et al., 2010). Based on the annotations in NR, BLAST2GO v2.5 (Conesa et al., 2005) was used to obtain GO annotations for the aligned unigene sequences with an *E*-value threshold of 10^{-6} , and the Web Gene Ontology Annotation Plot (WEGO) software (Ye et al., 2006) was used to establish GO functional classifications for all unigenes. The unigenes were aligned to the euKaryotic Ortholog Groups/Cluster of Orthologous Groups (KOG/COG) database to predict and classify possible functions, and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways database (Kanehisa et al., 2008) was used to obtain pathway annotations (*E*-value threshold 10^{-10}).

Identification of Gene Orthologous Groups and Calculation of Ka/Ks

The Coding sequence (CDS) of each putative unigene were extracted according to the BLASTX results, and the orientation of the unaligned sequences was determined using ESTScan software. The CDSs extracted from the respected unigene were translated into amino acid sequences using the standard codon table. Self-to-self BLASTP was conducted for all amino acid sequences with a cutoff *E*-value of 10^{-5} . Based on the predicted CDS regions of both *E. ehracteolata* and *E. fischeriana* transcriptomes, ORTHOMCL version 2.0.9 (Li et al., 2003) with default settings was used to reconstruct the clusters of orthologous groups (COGs). Pairs of putative orthologous genes were identified based on the reciprocal best matches with an *E*-value threshold of 10^{-20} . Only pairs of sequences that mapped unambiguously to the same protein in Swiss-Prot database were selected as orthologous genes. The protein-coding sequences

with unexpected stop codons in the BLAST hit region and/or shorter than 150 bp in length were removed.

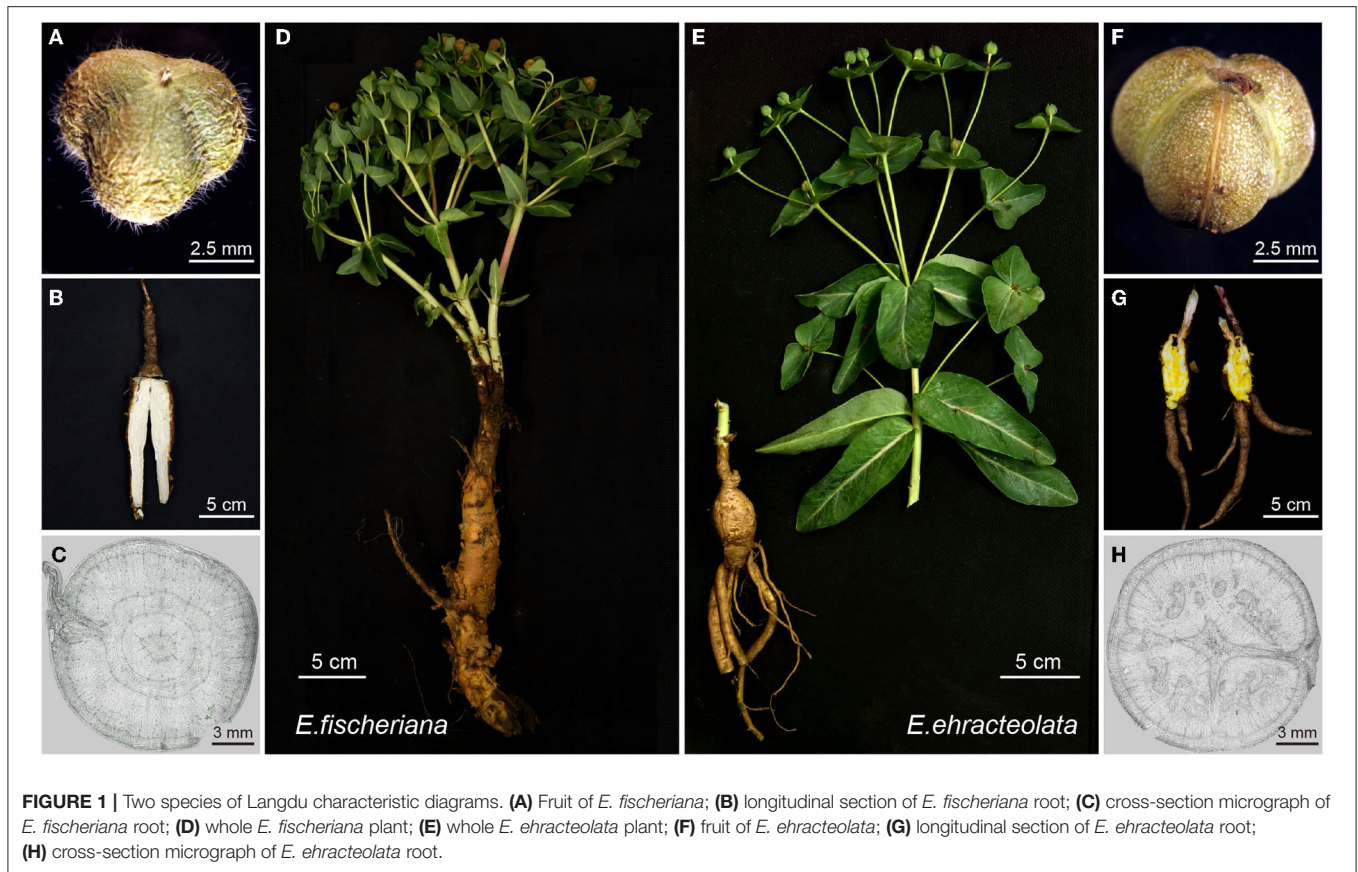
Then, the *Ks* value of the obtained orthologous genes and the formula $T = K/2r$ were used to estimate divergence time (*T*) between *E. fischeriana* and *E. ehracteolata* (Graur et al., 2000). "*K*" is a genetic divergence expressed in terms of mean number of synonymous substitutions between orthologs; "*r*" is the mean rate of synonymous substitution and is considered to be 1.5×10^{-8} substitutions/synonymous site/year for all dicots (Koch et al., 2000). *Ka/Ks* calculation was performed with PAML package (Yang, 2007) using default settings. Based on the *Ka/Ks* value and a threshold at 0.5, the orthologs were sub-categorized into two datasets: a test set with *Ka/Ks* above 0.5, and a reference dataset with *Ka/Ks* value < 0.5. The significance of the difference in GO term abundance between the two datasets was tested using the Fisher's exact test with the GOSSIP package (Bluthgen et al., 2005) implemented in BLAST2GO V.2.6 (Conesa et al., 2005).

Metabolome Analysis

About 50 mg of fresh pulverized roots and leaves of *E. fischeriana* and *E. ehracteolata* was accurately weighed and extracted with 1.0 ml 80% methanol for LC-MS. Extraction was done at room temperature (RT) for 3 h in an orbital shaker set at 220 rpm. Resulting extracts were spun down at 3,000 g for 15 min to sediment tissue material. About 200 μl of supernatant was transferred to vial for LC-MS analysis using an established LC-MS-based approach (Su et al., 2016; Kong et al., 2017). The LC-MS detection data were extracted and preprocessed using SIEVE software (Thermo), and the data were normalized and later edited in Excel 2013. Finally, it is organized into a two-dimensional data matrix form, which contains information such as retention time (RT), compound molecular weight (CompMW), observations (sample name), amount of material extracted (ID), and peak intensity. A total of 1,220 features at (ESI+) ion mode and 1,506 features at (ESI-) ion mode after editing were performed multivariate analysis (MVA) using SIMCA-P software (version 13.0, Umetrics AB, Umea, Sweden). Principal component analysis (PCA) and partial least squares-discriminate analysis (PLS-DA) were employed to identify biochemical patterns. The metabolites that differed between the two classes were quantified using a combination of VIP statistics (threshold > 1) of the OPLS-DA model and *t*-tests ($p < 0.05$). Compounds were identified by a comparison of *m/z* or precise molecular mass at <http://metlin.scripps.edu>.

Determination of Secondary Metabolites

Roots and leaves of *E. fischeriana* and *E. ehracteolata* were first freeze-dried then pulverized into powder. About 0.15 g of the fine powder was accurately weighed and extracted with 15 ml ethanol in an ultrasonic water bath at 30°C for 30 min. After cooling to room temperature, the extraction solution was filtered through a 0.22- μm millipore filter and injected directly into the UPLC system for analysis. Jolkinolide A (ChemFaces, CAS:37905-07-0, lot no: CFS201902), jolkinolide B (Nature Standard, CAS:37905-08-1, lot no:4069), jolkinolide E (ChemFaces, CAS:54494-34-7, lot no: CFS201901), and ingenol (Nature Standard, CAS:30220-46-3, lot no:3925) were used as the standard substances and were



dissolved in ethanol to obtain stock solutions of 1 mg/ml^{-1} . A number of 4 reference solutions were mixed and diluted with ethanol to obtain a series of mixture solutions. All the solutions were stored at -20°C before use.

Quantitation of active components was performed on an ultra-high-performance liquid chromatography coupled with tandem mass spectrometry (UPLC-MS/MS, LCMS 8045, Shimadzu, Japan). An ACQUITY BEH C_{18} column ($2.1 \text{ mm} \times 100 \text{ mm}$, $1.7 \mu\text{m}$) was used to separate diterpenoids. The mobile phase was composed of formic acid in water (0.05%, V/V) as solvent A and formic acid in ACN (0.05%, V/V) as solvent B, with a flow rate of 0.4 ml/min^{-1} at 40°C . The optimized gradient elution was as follows: 0–6 min, 30–93%B; 6–9 min, 93%B; 9–9.2 min, 93–30%B. About $1 \mu\text{l}$ of the solution was injected. The mass system was equipped with an electrospray ionization (ESI) source operating in both positive and negative ion modes, using multiple reaction monitoring (MRM) mode. Optimized parameters are listed in **Supplementary Table S1**. The nebulizer and drying gas were 99.95% nitrogen, and their flow rates were 3.0 and 10.0 L/min^{-1} , respectively. The heating gas was 99.95% air with a flow rate of 10.0 L/min^{-1} . The collision gas was 99.99% argon with a pressure of 270 kPa . Other parameters were as follows: interface voltage 4.0 kV , interface temperature 300°C , desolvation line temperature 250°C , and heat block temperature 400°C . The precision [relative standard deviation (RSD) $< 5.77\%$] and accuracy (recovery was from 95.9 to 105.1%) of this

method met the requirements for quantitative determination. Linearity was verified using coefficients of determination (R^2), which were all > 0.999 within the adopted linear range (**Supplementary Material**).

Development and Detection of EST-SSR Markers

The B MISA (<http://pgrc.ipk-gatersleben.de/misa/>) (Dieringer and Schlötterer, 2003) was used to identify and localize microsatellite motifs in the two Langdu species and SSRs were considered to contain motifs of two to six nucleotides and a minimum of five contiguous repeat units. The alignments of 6,153 pairs of orthologs were extracted as the input file for the MISA program. Using the detailed information on SSR loci obtained from the output of the MISA program, primers for each SSR-containing sequence with a repetitive at least 15 bp in length were designed with Program Primer Premier 5 (PREMIER Biosoft Int., Palo Alto, CA). To validate the SSRs identified *in silico* identified SSRs, 68 primer pairs shared between the two Langdu species were synthesized [Sangon Biotech (Shanghai) Co., Ltd., Shanghai, China] and validated by polymerase chain reaction (PCR) in 9 species, including *Euphorbia fischeriana* (Hailar, Inner Mongolia), *Euphorbia ebracteolata* (Jiyuan, Henan province), *Euphorbia lathyris* (Lijiang, Yunnan province), *Euphorbia pekinensis* (Chuzhou, Anhui province), *Euphorbia sieboldiana* (Wang qing, Jilin province), *Euphorbia kansui*

TABLE 1 | Summary of assembly results for *E. fischeriana* and *E. ehracteolata* using trinity.

Sequence	<i>E. fischeriana</i>			<i>E. ehracteolata</i>		
	Root	Stem	Leaf	Root	Stem	Leaf
Raw reads	50845208	46010830	49087484	61659938	56456010	51479006
Clean reads	49966652	45210234	48124262	60048398	54973026	50043212
Q20 (%)	97.09	97.07	96.92	97.17	97.2	97.13
GC (%)	43.54	42.53	42.83	42.62	42.62	42.51
Total number of contigs/unigenes		149,207/112,487			143,857/101,211	
Length range of contigs/unigenes		200–15,623			200–16,340	
N50 value of contigs /unigenes		1,559/1,051			1,635/1,155	
Mean length of contigs/unigenes		813/631			885/686	
Median length of contigs/unigenes		398/325			459/366	

(Yongji, Shanxi Province), *Euphorbia humifusa* (Yongji, Shanxi Province), *Euphorbia helioscopia* (Jiyuan, Henan province), and *Euphorbia esula* (Ewenke, Inner Mongolia autonomous region). According to the Flora of China (Li et al., 2008), samples are all commonly used Chinese medicine. Fresh leaves were dried in silica gel at the time of collection. Voucher specimens were deposited in National Resource Center for Chinese Materia Medica (Beijing, China).

Genomic DNA was extracted from the dry leaves using Plant Genomic DNA Kit (TIANGEN). PCRs were performed in a 25 μ l volume containing 50 ng of template genomic DNA. The PCRs were carried out under the following conditions: initial denaturation at 94°C for 2 min, 35 cycles at 94°C for 30 s, 55°C for 30 s, 72°C for 40 s, and a final extension at 72°C for 10 min. The separation of alleles was performed on a 3% polyacrylamide gel with a 500-bp DNA marker (TaKaRa) to calculate the length of the EST-SSR amplicons.

RESULTS

Differences in Habitats and Phenotypes Between the Two Langdu Species

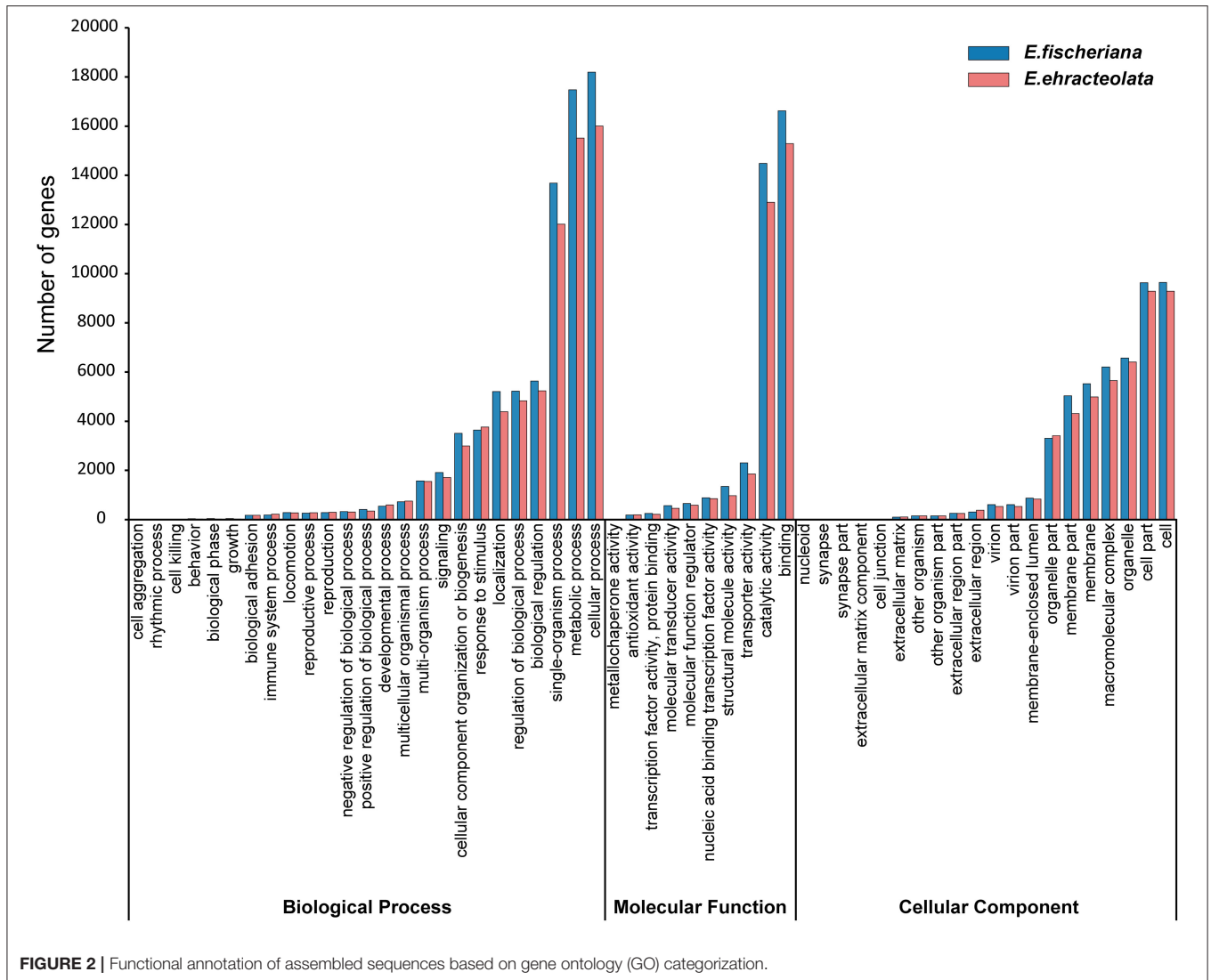
To study these two closely related species, we collected wild *E. fischeriana* in Hailar, Inner Mongolia and wild *E. ehracteolata* in Jiyuan, Henan province, which are the main producing areas of two species, respectively. The average annual temperature in Hailar is -2.0°C , and that in Jiyuan is 14.6°C (data from www.nmc.cn). Therefore, in terms of habitat distribution, *E. fischeriana* is more tolerant of coldness than *E. ehracteolata*. This adaptive distribution might also result in phenotypic differences between the two Langdu species. Classification of the two species majorly depends on whether the ovary and capsule have trichome (Figures 1A,F) (Li et al., 2008), but the identification and comparison of the experimental materials showed that their roots (medicinal part) are also significantly different. The root of *E. fischeriana* is elongated and enlarged with multiple layers of semi-exfoliated epidermis, whereas *E. ehracteolata* is fusiform and the outer epidermis is tight (Figures 1D,E). In root of *E. fischeriana*, latex is white (Figure 1B) with concentric ring vascular bundle (Figure 1C),

whereas yellow latex (Figure 1G) and abnormal vascular bundle (Figure 1H) are obtained in *E. ehracteolata*. The medicinal materials from the two species are similar in appearance, mostly in transverse slices or oblique slices, collectively referred to as “white Langdu,” which is easy to be confused. Therefore, we further developed biomarkers to assist the identification of medicinal materials.

De novo Assembly and Functional Annotation

To reveal the molecular mechanism of adaptive distribution, cDNA libraries of the root, stem, and leaf of *E. fischeriana* and *E. ehracteolata* were sequenced, with 169594954 and 145943522 raw reads obtained (Table 1). The average Q20 percentage (percentage of sequences with sequencing error rate lower than 1%) was 97.17 and 97.03%. The average GC percentage was 42.58 and 42.97%. Using the trinity *de novo* assembly program, we assembled the short read sequences from the three tissues into 149027 transcripts, with a mean length of 813 bp and an N50 length (the contig size such that 50% of the entire assembly is contained in contigs equal to or longer than this value) of 1,559 bp for the *E. fischeriana*. For *E. ehracteolata*, 143,857 contigs longer than 200 bp, with a mean length of 885 bp and an N50 length of 1,635 bp, were generated. In total, all contigs were connected into 112,487 unigenes with a mean length of 631 bp and an N50 of 1,051 bp for *E. fischeriana* and 101,211 unigenes with a mean length of 686 and an N50 value of 1,115 for *E. ehracteolata*.

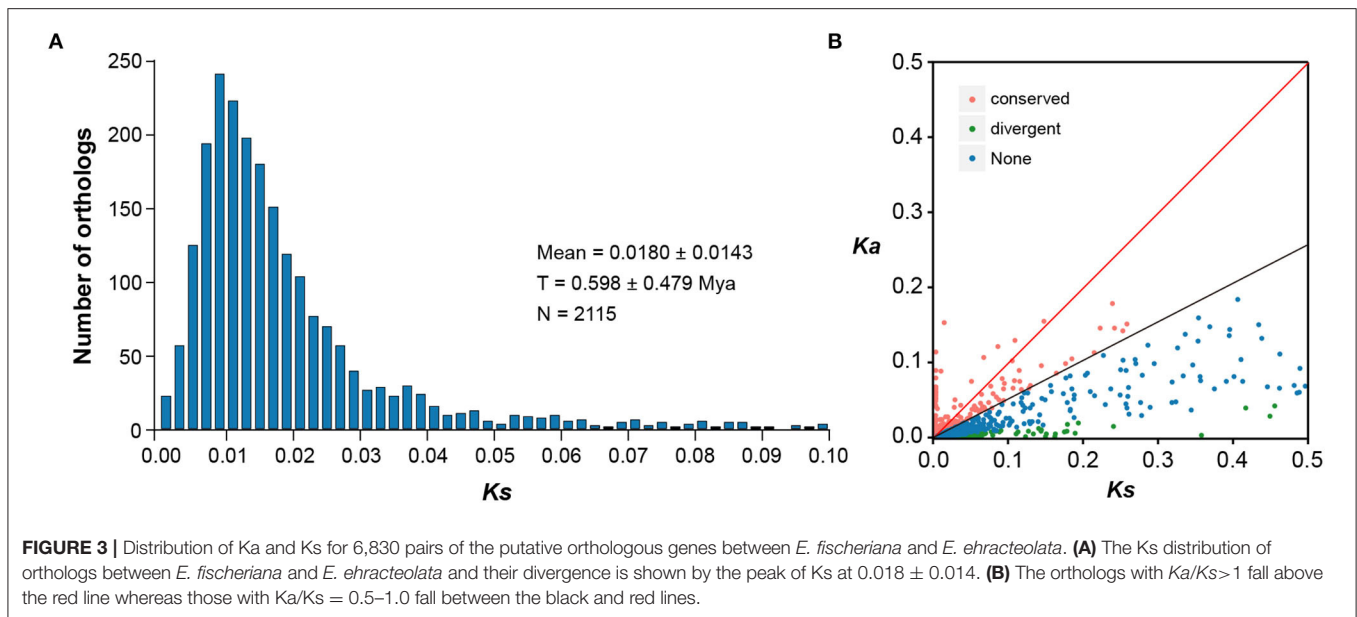
To obtain comprehensive gene function information, we carried out gene function annotation of seven databases, including Nr, Nt, Pfam, KOG/COG, Swiss-Prot, KEGG, and GO (Supplementary Table S2). The all unigenes were assigned putative gene descriptions based on the BLAST (E -value $\leq 1 \times 10^{-5}$) search against the NCBI non-redundant (Nr) protein database. A total of 36,204 (32.18%) unigenes for *E. fischeriana* and 38,018 (37.56%) unigenes for *E. ehracteolata* were shown significant similarity with proteins in the Nr database. There were 48,033 (42.70%) unigenes for *E. fischeriana* and 44,919 (44.38%) for *E. ehracteolata*



with at least one significant match to the above databases. For both species, a BLASTX top-hit species distribution of gene annotations showed highest homology to *Jatropha curcas* [*E. fischeriana* (38.49%) vs. *E. ehracteolata* (39.02%)]. Based on Nr annotations, we used the GO classification system to functionally categorize unigenes. A total of 28,472 (28.13%) unigenes for *E. ehracteolata* and 32,598 (28.97%) unigenes for *E. fischeriana* were assigned to at least one GO term annotation. The unigenes were assigned to three main GO categories (**Figure 2**): biological process (*E. ehracteolata*: 71,495, 47.30%; *E. fischeriana*: 79,428, 47.91%), molecular function (33,335, 22.06 vs. 37,307, 22.51%), and cellular component (46,314, 30.64 vs. 49,036, 29.58%). These categories were similarly distributed in both species.

For the biological process category, the two mostly highly represented terms among the 25 level-2 categories were cellular process, metabolic process, and single-organism process; for the molecular function category, among the 10 level-2 categories,

binding and catalytic activity were overrepresented; for the 20 level-2 categories in the cellular component category, cell part, and organelle were the most abundant terms (**Figure 2**). These categories were similarly distributed in both species. Specifically, compared with *E. ehracteolata*, more genes in *E. fischeriana* were involved in biological process and molecular function, including metabolic process (15,517, 10.27 vs. 17,479, 10.54%), localization (4,390, 2.91 vs. 5,209, 3.14%), cellular component organization or biogenesis (2,994, 1.98 vs. 3,508, 2.12%), positive regulation of biological process (351, 0.23 vs. 416, 0.25%), molecular transducer activity (451, 0.30 vs. 571, 0.34%), structural molecule activity (975, 0.65 vs. 1,347, 0.81%), transporter activity (1,853, 1.23 vs. 2,305, 1.39%), and so on. However, genes involved in biological phase (34, 0.02 vs. 42, 0.03%), growth (36, 0.02 vs. 46, 0.03%), immune system process (183, 0.11 vs. 226, 0.15%), and developmental process (544, 0.33 vs. 598, 0.40%) showed a conversed trend, comparing *E. fischeriana* with *E. ehracteolata*.



Orthologous Genes and Substitution Rates Between Two Langdu Species

Based on the predicted CDS regions of both *E. fischeriana* and *E. ehracteolata* transcriptomes, we identified initial putative orthologous pairs. After removing the pairs with unexpected stop codons in the BLAST hit region and/or shorter than 150 bp in length, 6,830 ortholog pairs were retained. Out of the 6,830 ortholog pairs, 3,193 pairs had both non-synonymous (Ka) and synonymous (Ks) substitutions and thus were allowed for the calculation of Ka/Ks ratios. A peak of Ks value distribution between *E. fischeriana* and *E. ehracteolata* was observed at 0.0180 ± 0.0143 (Figure 3A), and the low level of Ks value indicates close relationship. It was estimated by the formula $T = K/2r$ that the age of the speciation event between *E. fischeriana* and *E. ehracteolata* is ~ 0.598 Mya, which falls in the Middle Pleistocene. Among the 3,193 ortholog pairs, 533 pairs showed high Ka/Ks values (Figure 3B), in which 191 pairs have a $Ka/Ks > 1$ ($p < 0.05$), indicating positive selection, whereas 342 pairs have a Ka/Ks between 0.5 and 1, indicating weak positive selection. By contrast, most of the orthologous pairs (2,658) showed a Ka/Ks ratio < 0.5 , indicating that these gene pairs are likely under purifying selection.

Gene Functions Under Positive Selection and Implications for Adaptive Evolution Between Two Langdu Species

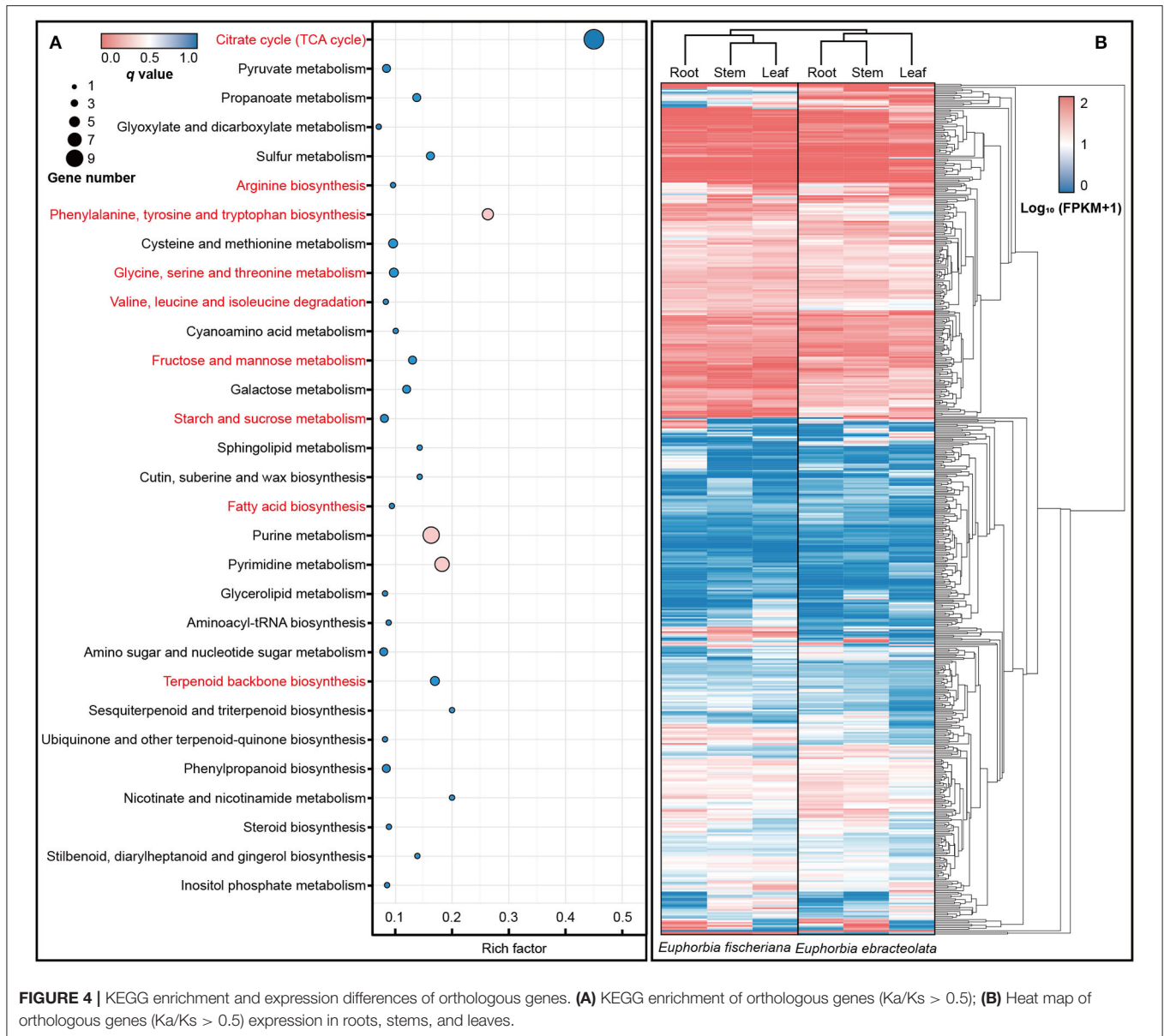
In enrichment analysis, we used the orthologs into one dataset with $Ka/Ks > 0.5$ (Swanson et al., 2004). In KEGG database, the positive selection orthologous genes were enriched to 57 pathways, including (1) citrate cycle (TCA cycle); (2) amino acid biosynthesis and metabolism including arginine biosynthesis, phenylalanine, tyrosine and tryptophan biosynthesis, cysteine and methionine metabolism, glycine, serine and threonine metabolism, as well as valine, leucine, and isoleucine degradation;

(3) fructose and mannose metabolism; (4) starch and sucrose metabolism; (5) fatty acid biosynthesis; and (6) terpenoid backbone biosynthesis (Figure 4A). The expression levels of orthologous genes in pathway are also different in two Langdu species (Figure 4B).

In an analysis of GO terms with at least five hits, 31 GO terms annotated to 76 pairs of orthologs were found to be over-represented (p -value < 0.05) in the test dataset. For the selected genes, we used BLASTX search to find their orthologous genes, and genes with function in stress tolerance, development, TCA-related, and so on were filtered (Supplementary Table S3). In particular, half of terms involved genes were related to stress tolerance, including cold stress, heat stress, and drought stress. Terms containing cold stress tolerance gene accounted for 30%, including *AtGCN1* (Wang et al., 2017b), *AtYLS9* (Griebel et al., 2022), *BrRZFP* (Jung et al., 2013), *AtPTP1* (Liu et al., 2016), and *AtPFD* (Perea-Resa et al., 2017). In terms of growth and development, we enriched genes related to trichome and root hair development, such as *AtBLT* (Kasili et al., 2011; Mazie and Baum, 2016), *AtVTI13* (Larson et al., 2014), *AtNRP2* (Zhu et al., 2017), and *AtWER1* (Wang et al., 2019). TCA-related genes were also enriched, including *AtSAM1* (Sekula et al., 2020), *AtLIP1* (Wang et al., 2022), *AtPES1* (Lippold et al., 2012), *AtCCDC* (Lohmeier-Vogel et al., 2008), and *AtBASS2* (Furumoto et al., 2011; Mueller et al., 2014; Lee et al., 2017), involving amino acid biosynthesis, catabolism of triacylglycerols, fatty acid, starch metabolism, pyruvate transport, etc.

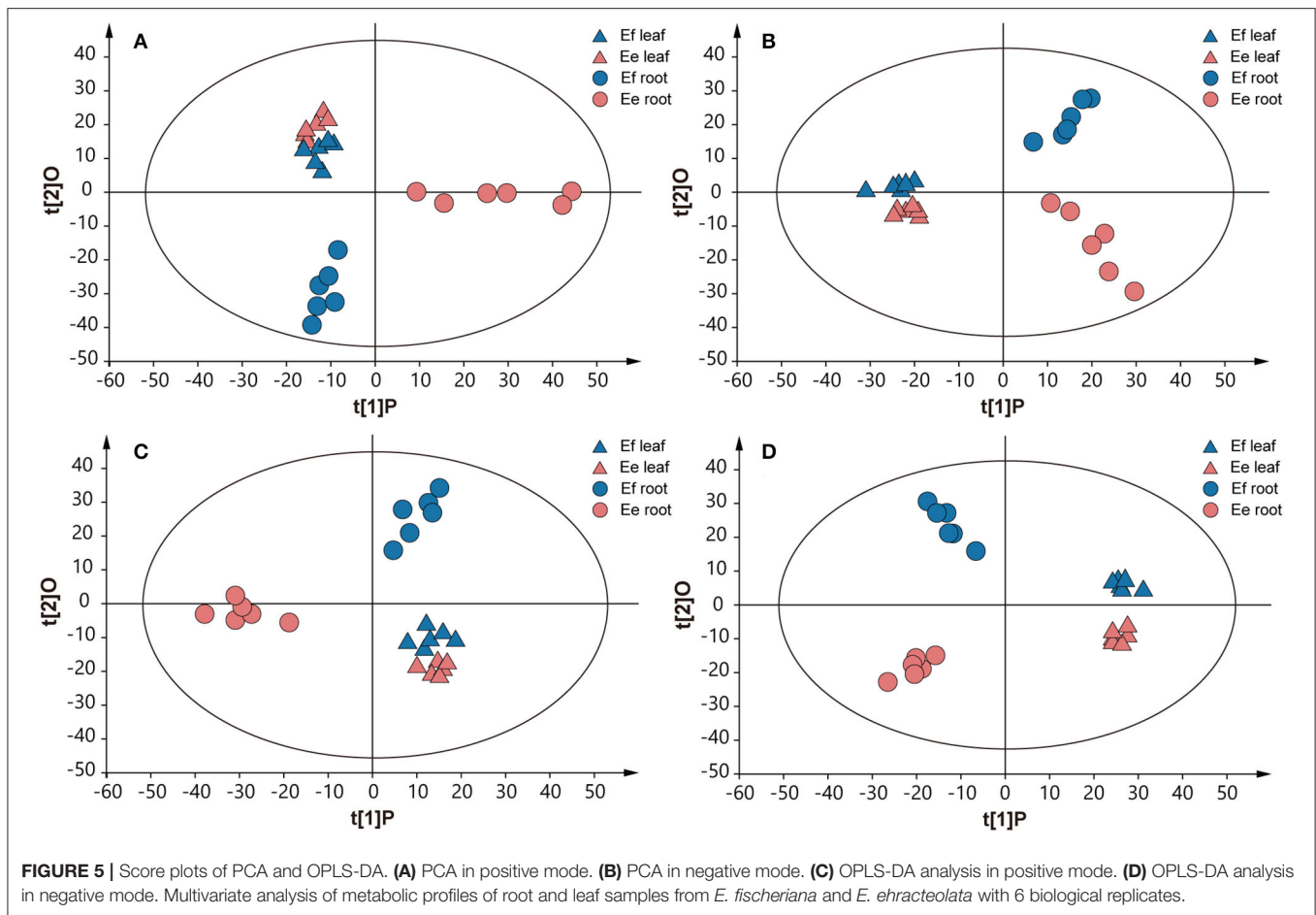
Identification of Metabolites Differentially Accumulated in Two Langdu Species

Due to the above genetic differences, we further studied the metabolic differences between the two species, to reveal the effects of genotypes on chemical phenotypes. The chemical



components in roots and leaves were detected through UPLC-MS analysis, obtaining final data containing 1,220 features in ESI+ mode and 1,506 features in ESI- mode. An unsupervised principal component analysis (PCA) model was used, since it can represent the intrinsic characteristics of the data. The result revealed a clear separation of metabolite samples between *E. fischeriana* and *E. ebracteolata* ($R^2X = 0.724$, $Q^2 = 0.561$ in ESI+ mode; $R^2X = 0.598$, $Q^2 = 0.387$ in ESI- mode) (Figures 5A,B). The PCA score plot revealed that the leaves from *E. fischeriana* (blue triangle) and *E. ebracteolata* (red triangle) are not perfectly separated by the 12 samples. Similar to the appearance of the two Langdu leaves, there is less difference in the composition and content of their leaves. But the roots of *E. fischeriana* (blue circle) and *E. ebracteolata* (red circle) were

clearly separated. The results indicated that there were significant differences in the compounds of Langdu roots. The PCA model was also applied to obtain a preliminary overview of general similarities and differences between collections. To identify discriminating metabolites and differentiate the four groups, we used the corresponding OPLS-DA analysis (Figures 5C,D). The two Langdu species can be separated clearly by only one predictive component deriving from a more sophisticated orthogonal partial least squares discriminate analysis (OPLS-DA) model, since noisy information irrelevant to species was removed prior to model building (Bylesjo et al., 2006). OPLS-DA analysis showed potential markers in different Langdu plants. A total of 98 discriminating metabolites ($\text{VIP} > 1.0$, $p < 0.05$), including 46 in positive mode ($R^2X = 0.758$, $R^2Y = 0.965$, $Q^2 =$



0.917) and 52 in negative mode ($R^2X = 0.593$, $R^2Y = 0.981$, $Q^2 = 0.926$), were identified in Langdu roots. The permutation result validated the stability and reliability of this OPLS-DA model. Subsequently, relying on the three criteria—variable importance in the projection (VIP) value of OPLS-DA model ≥ 1.5 , p -value of t -test ≤ 0.01 —5 metabolites could be presumably considered as candidate biomarkers. In addition, jolkinolide B and ingenol showed specific accumulation in the two species and could also serve as the biomarkers (**Supplementary Table S4**).

To reveal the effect of environments on the metabolites, the tentatively identified compounds were assigned in the common metabolic pathways according to the literature works (Schauer et al., 2006; Duan et al., 2012). Combined with previous KEGG and GO enrichment results of adaptive evolution genes, we found that the contents of malic acid, fumaric acid, GABA, amino acids, and other compounds in TCA in *E. fischeriana* were significantly lower than in *E. ehracteolata*. But some soluble sugars, such as glucose, sucrose, raffinose, xylulose, and mannitol, were increased in *E. fischeriana* in comparison with the *E. ehracteolata*. Specifically, proline in *E. fischeriana* was significantly higher than that in *E. Ehracteolata*, but unsaturated fatty acids including linoleic acid and α -linolenic acid were lower (**Figure 6**).

Differences in Expression Levels of Orthologous Genes and Active Ingredients

The differences in active ingredients between two Langdu species may result from the differences in genotype and gene expression. *DXR* (OG06739) and *MK* (OG07880) were found in the positive selection library ($Ka/Ks > 0.5$) mentioned in the previous study, indicating that these two genes had adaptive mutations. We also analyzed the expression levels of MVA and MEP pathways and terpene synthase homologous genes globally and found that the gene expressive abundance of *E. fischeriana* was higher than that of *E. ehracteolata*, especially *DXS*, *DXR*, *CMK*, *HDS*, *HDR*, *HMGS*, *HMGR*, *MK*, and diterpene synthase genes *CPS* and *KSL* in the roots. However, the expression level of *CBS* in *E. ehracteolata* root was significantly higher than that of *E. fischeriana*, and even *CBS* (OG17313) was not expressing in *E. fischeriana* root.

To analyze the differences in active ingredients, a UPLC-MS/MS method simultaneously determining 4 diterpenoids was established, and the root and leaf tissues of the two Langdu species were detected (**Figure 7**). The content of jolkinolide A was about 0.31 mg/g^{-1} in the roots of *E. fischeriana* and 0.10 mg/g^{-1} in the roots of *E. ehracteolata*. Jolkinolide B was significantly accumulated in the roots of *E. fischeriana* (1.13

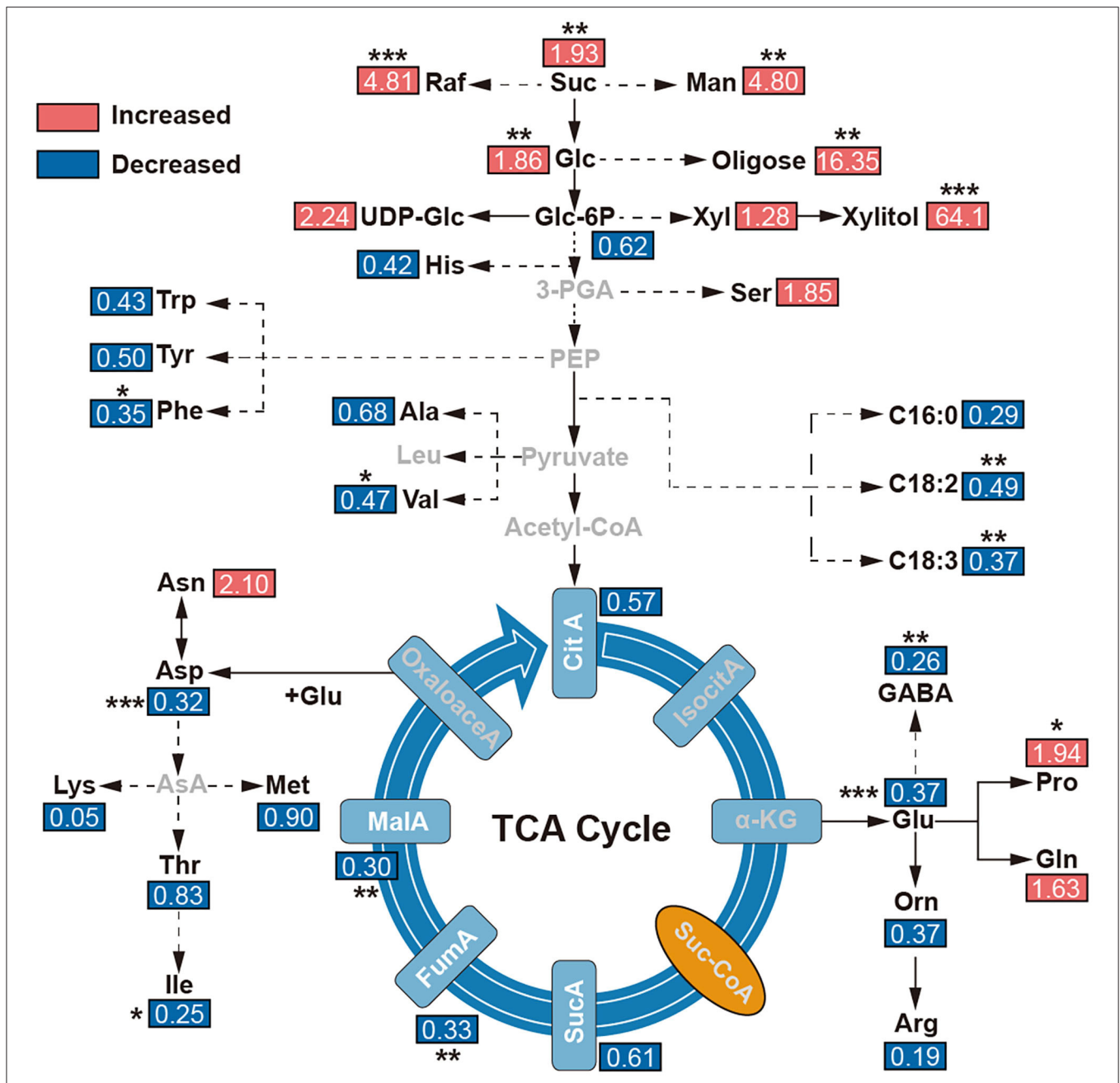


FIGURE 6 | Levels of metabolites in main metabolic maps. The comparisons of metabolite contents were calculated by dividing the metabolite level in *E. fischeriana* with that in *E. ehracteolata*. 3-PGA, glycerate 3-phosphate; α -KG, alpha-ketoglutarate; Arg, arginine; ASA, aspartic acid- β -hemialdehyde, Asn, asparagine; Asp, aspartic acid; Ala, beta-alanine; CitA, Citric acid; PalA, palmitic acid; OctA, octadecanedioic acid; LinA, linoleic acid; α -Lin, α -linolenic acid; FumA, fumarate; GABA, γ -aminobutyric acid; Glc, glucose; Glc3P, glucose-3-phosphate; Gln, glutamine; Glu, glutamate; Gly, glycine; Ile, isoleucine; IsocitA, isocitric acid; Leu, leucine; Lys, lysine; MalA, malate; Man, mannitol; Met, methionine; Orn, ornithine; Pro, proline; Raf, raffinose; Ser, L-serine; Stigm, stigmaterol; Suc, sucrose; SucA, succinic acid; Suc-CoA, succinate coenzyme A; Thr, threonine; Trp, tryptophan; Val, valine; Xyl, xylulose. The value represents the ratio of *E. fischeriana* to *E. ehracteolata*. Asterisks denote Student's *t*-test significance: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

mg/g^{-1}) and slightly accumulated in the leaves (0.03 mg/g^{-1}), whereas the content of jolkinolide B in the roots of *E. ehracteolata* was 0.03 mg/g^{-1} . Jolkinolide E reached 0.16 mg/g^{-1} in the root of *E. fischeriana* and 0.06 mg/g^{-1} in *E. ehracteolata*. Conversely, bicyclic diterpenoid ingenol was only detected in the roots of

E. ehracteolata, about 0.04 mg/g^{-1} . Obviously, there were the differences in active ingredients biosynthesis genes and their expression levels between the two species during the evolution of environmental adaptation, which resulted in the preference of accumulation.

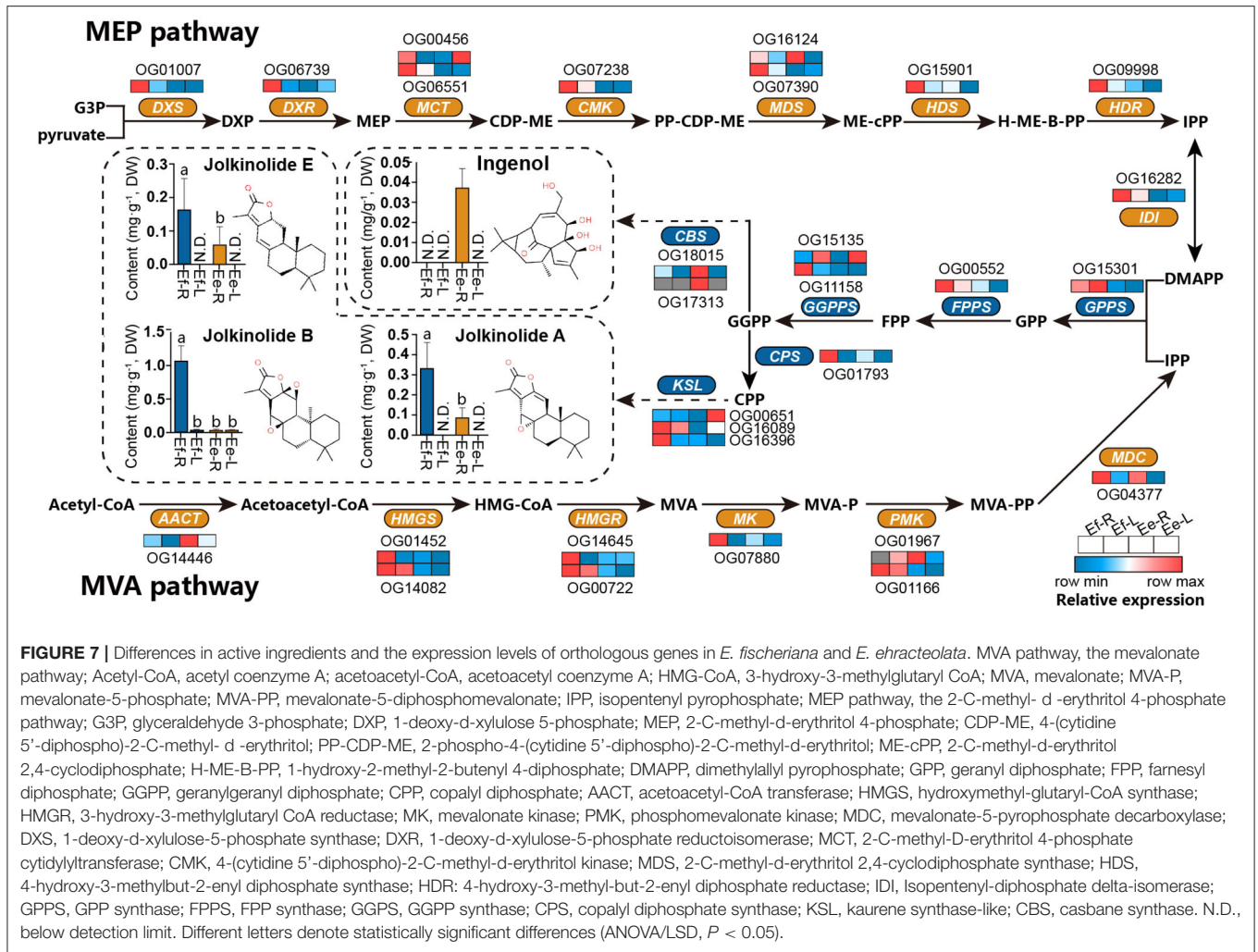


FIGURE 7 | Differences in active ingredients and the expression levels of orthologous genes in *E. fischeriana* and *E. ehracteolata*. MVA pathway, the mevalonate pathway; Acetyl-CoA, acetyl coenzyme A; acetoacetyl-CoA, acetoacetyl coenzyme A; HMG-CoA, 3-hydroxy-3-methylglutaryl CoA; MVA, mevalonate; MVA-P, mevalonate-5-phosphate; MVA-PP, mevalonate-5-diphosphomevalonate; IPP, isopentenyl pyrophosphate; MEP pathway, the 2-C-methyl- d -erythritol 4-phosphate pathway; G3P, glyceraldehyde 3-phosphate; DXP, 1-deoxy-d-xylulose 5-phosphate; MEP, 2-C-methyl-d-erythritol 4-phosphate; CDP-ME, 4-(cytidine 5'-diphospho)-2-C-methyl- d -erythritol; PP-CDP-ME, 2-phospho-4-(cytidine 5'-diphospho)-2-C-methyl-d-erythritol; ME-cPP, 2-C-methyl-d-erythritol 2,4-cyclodiphosphate; H-ME-B-PP, 1-hydroxy-2-methyl-2-butenyl 4-diphosphate; DMAPP, dimethylallyl pyrophosphate; GPP, geranyl diphosphate; FPP, farnesyl diphosphate; GGPP, geranylgeranyl diphosphate; CPP, copalyl diphosphate; AACT, acetoacetyl-CoA transferase; HMGS, hydroxymethyl-glutaryl-CoA synthase; HMGR, 3-hydroxy-3-methylglutaryl CoA reductase; MK, mevalonate kinase; PMK, phosphomevalonate kinase; MDC, mevalonate-5-pyrophosphate decarboxylase; DXS, 1-deoxy-d-xylulose-5-phosphate synthase; DXR, 1-deoxy-d-xylulose-5-phosphate reductoisomerase; MCT, 2-C-methyl-D-erythritol 4-phosphate cytidyltransferase; CMK, 4-(cytidine 5'-diphospho)-2-C-methyl-d-erythritol kinase; MDS, 2-C-methyl-d-erythritol 2,4-cyclodiphosphate synthase; HDS, 4-hydroxy-3-methylbut-2-enyl diphosphate synthase; HDR, 4-hydroxy-3-methyl-but-2-enyl diphosphate reductase; IDI, Isopentenyl-diphosphate delta-isomerase; GGPS, GPP synthase; FPPS, FPP synthase; GGPPS, GGPP synthase; CPS, copalyl diphosphate synthase; KSL, kaurene synthase-like; CBS, casbane synthase. N.D., below detection limit. Different letters denote statistically significant differences (ANOVA/LSD, $P < 0.05$).

Development of EST-SSR Markers in *Euphorbia* Species

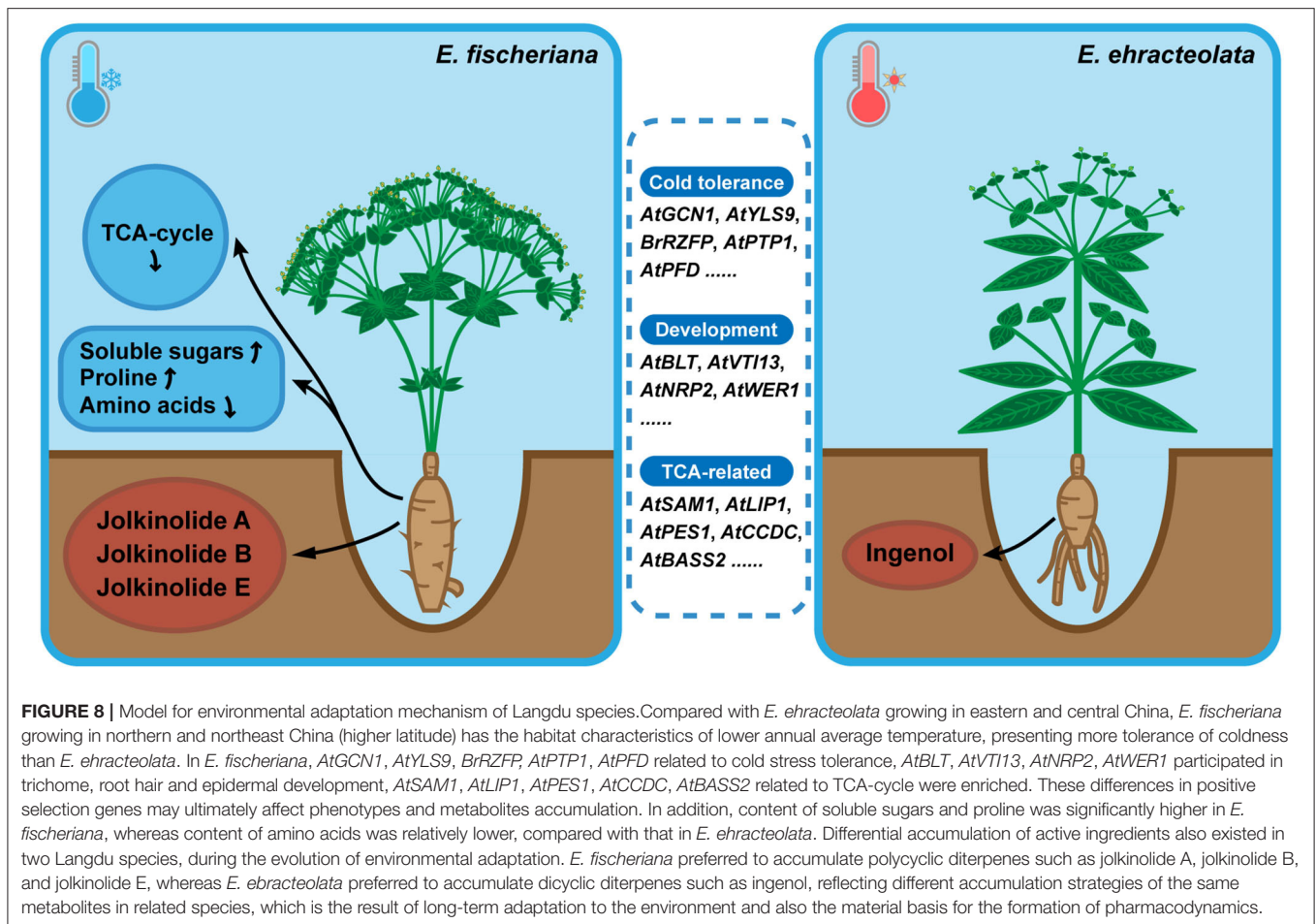
Based on the *E. fischeriana* and *E. ehracteolata* transcriptomes, 8,458 and 10,612 EST-SSRs were identified, respectively. The most common repeat types were dinucleotides (44.0 and 44.5%), followed by trinucleotides (42.8 and 43.3%), and the most common repeat motifs were AG/CT, AT/AT, and AC/GT, followed by AAG/CTT and AAT/ATT (Supplementary Table S5). To maximize the universal applicability of markers developed in this study and hence reduce their cost, we searched for EST-SSRs in the 6,153 pairs of putative orthologous unigenes and found 967 EST-SSRs distributed among 780 pairs of orthologous unigenes. Taking only those with a repeat unit length of at least 15 bp, 68 pairs of EST-SSRs contained in 532 pairs of orthologs were selected for primer design, and 68 pairs of sequences with conserved, sufficiently long flanking sites were used to design primers successfully. To evaluate the reliability and cross-species transferability of these primers, we tested all the 68 pairs of EST-SSR markers for 9 species of *Euphorbia*. The average commonality ratio of primers in species was 35.69%, 49 pairs

produced clear fragments with the expected sizes in both two Langdu and two pairs (OG00193 and OG07454) which produced clear fragments with the expected sizes in all nine species, respectively. A total of nine pairs produced fragments in eight species (Supplementary Table S6). Among them, OG07421, OG10880, and OG13687 could be amplified with obvious polymorphism bands in several *Euphorbia* species.

DISCUSSION

The Environmental Adaptation of Two Langdu Is Influenced by Genotype and Metabolites

E. fischeriana mostly grows in northern and northeast China, with high latitude, less rainfall, low annual average temperature. However, *E. ehracteolata* mostly grows in eastern and central China (Li et al., 2008). Temperature, as one of the main environmental variables, makes that the two species exhibit a typical pattern of adaptive evolution and explosive speciation, which may be influenced by genotype and metabolites. The positive selection orthologous genes were enriched to TCA



cycle, stress tolerance, development, amino acid biosynthesis and metabolism, fructose and mannose metabolism, starch and sucrose metabolism, and fatty acid biosynthesis pathways (Figure 8). Among them, OG06839 is homologous to *AtGCN1*, which mediated phosphorylation of eIF2 α and is essential for cold tolerance (Wang et al., 2017b); OG13992 is homologous to *AtYLS9*, which tolerates cold stress by activating an immune response (Griebel et al., 2022); OG06917 is homologous to *BrRZFP*, which regulates germination, fresh weight, and length of shoots and roots to tolerate stress (Jung et al., 2013); OG01773 is homologous to *AtPTP1* and encodes a protein with tyrosine phosphatase activity that is downregulated in response to cold stress (Liu et al., 2016); OG07015 is homologous to *AtPFD* and controls the levels of HY5, a key role in cold acclimation by activating anthocyanin biosynthesis, in response to low temperature (Perea-Resa et al., 2017). The positive selection of the plant physiology-related genes may be one of the reasons why the two species can survive in different cold environments.

Trichomes play important protective roles as against abiotic stressors such as cold, drought, heat, excess of light, and UV radiation (Wagner et al., 2004; Hauser, 2014). For example, OG07214, which is homologous with *AtBLT*, may be a key gene in trichome development (Kasili et al., 2011; Mazie and Baum,

2016). The non-synonymous substitution of OG07214 may cause formation of trichome in *E. fischeriana* fruits (Figures 1A,F). In addition, *AtVTI13* (OG07291) is essential for the maintenance of cell wall organization and root hair growth (Larson et al., 2014); *AtNRP2* moderates chromatin structure for proper root hair development (Zhu et al., 2017). *AtWER1* plays an important role in generating the proper balance of downstream transcriptional factors in the gene regulatory network that establishes root epidermal cell fate (Wang et al., 2019). These gene differences may explain the larger root crown and more root epidermis of *E. fischeriana* (Figures 1D,E).

The accumulation of pyruvate, polyamines, fatty acid, and starch metabolism helps plants to tolerate cold. *AtSAM1* is responsible for production of S-adenosylmethionine, which promotes the accumulation of polyamines and phytohormone ethylene (Sekula et al., 2020); *AtLIP1* is a triacylglycerol lipases, which negative regulates cold tolerance (Wang et al., 2022); *AtPES1* involves in fatty acid phytyl ester synthesis in chloroplasts (Lippold et al., 2012); *AtCCDC* promotes starch metabolism by interacting with several potential enzyme (Lohmeier-Vogel et al., 2008); *AtBASS2* is a plastidial sodium-dependent pyruvate transporter (Furumoto et al., 2011; Mueller et al., 2014; Lee et al., 2017). Positive selection in above genes leads to non-synonymous

substitutions of key amino acid sites that alter enzyme activity, thereby affecting the accumulation of metabolites and promoting environmental adaptation.

These genetic changes were reflected in the differences in the accumulation of components in the two Langdu species. First, *E. fischeriana* reduces the rate of the TCA cycle and maintains low respiration (Close, 1997). At the same time, most amino acid nutrients are reduced to ensure that carbon sources are mostly used to supply protective substances such as glucose, sucrose, raffinose, and xylulose (Figure 6). These soluble sugars are accumulated in large quantities, protecting cell membranes and proteins from freezing and dehydration under cold stress (Klotke et al., 2004). Proline and mannitol also accumulate in *E. fischeriana*, which is due to response to abiotic and biological stresses by affecting osmotic pressure tolerance within plants (Szabados and Savoure, 2010). The accumulation of different metabolites in the roots of two Langdu species may reflect their adaptation to different environments.

Phenotypic formation is the result of both genotype and environmental modification. The phenotype of genuine medicinal materials includes the characteristics, tissue structure, active ingredient composition, and efficacy. The existence of specific genes is the basis for producing specific phenotypes, and the suitable habitat is the driving force for producing specific phenotypes (Yuan and Huang, 2020). Overall, in this study, we detected positive selection genes and metabolites, and these findings will not only shed light on how differentiations between two Langdu species occurs, but also open the door to increased understanding of how plants living in cold environments adapt to different characteristics of high latitude.

Differences in Active Ingredients of Two Langdu Species Might Lead to Transition in Clinical Usage

Active ingredients, as a part of the phenotype, are also influenced by genotype and environmental modifications. Diterpenoids are regarded as the main bioactive constituents of *E. fischeriana* and *E. ebracteolata*, which can be classified into abietane, tiglane, atisane, pimarane, rosane, kaurane, ingenane, and lathyrane types (Li et al., 2021b; Yang et al., 2021). Among them, abietane type is the main diterpenoid type of the two species. A total of 53 compounds have been isolated from *E. fischeriana*, including jolkinolide B, 17-hydroxyjolkinolide B, jolkinolide A, and 26 compounds from *E. ebracteolata*. In this study, it was found that content of jolkinolide A, jolkinolide B, and jolkinolide E in *E. fischeriana* was 3.1, 36.7, and 2.7 times of that in *E. ebracteolata*, respectively. *DXR* (OG06739) and *MK* (OG07880), involved in diterpenoid biosynthesis, were found in the positive selection library ($Ka/Ks > 0.5$) mentioned in the previous study, indicating that these two genes had adaptive mutations. Such variation may lead to the changes in enzyme activity and thus becomes more efficient in substrate catalysis, which requires further research. We also found that the most gene expression level of *E. fischeriana* was higher than that of *E. ebracteolata*, which may promote the accumulation of Jolkinolide A/B/E (Figure 7). Moreover, casbane type is unique to *E. ebracteolata*, including

Yuexiandajisu A and Yuexiandajisu B (Xu et al., 1998). Casbene is thought to form bicyclic diterpenoids, which in turn forms ingenanes (Luo et al., 2016). A total of 20 ingenanes (ingenol and Esters) have been isolated from *E. ebracteolata* and only 6 from *E. fischeriana*. However, the expression level of *CBS* in the roots of *E. ebracteolata* was significantly higher than that of *E. fischeriana*, and OG17313 was not detected in the latter. We speculated that this was one of the reasons that the content of ingenol in *E. fischeriana* was higher than that in *E. ebracteolata*. However, more definitive evidence will need to be obtained through genome sequencing and biosynthesis pathway analysis. In conclusion, this reflects the different accumulation strategies of the same metabolites in related species, which is the result of long-term adaptation to the environment and also the material basis for the formation of pharmacodynamics.

Jolkinolide B and derivatives in *E. fischeriana* have been proved to have significant anticancer activity (Wang et al., 2009, 2017a). Pharmacological activity of ingenol in *E. ebracteolata* is not clear (Huang et al., 2019), but various esters of ingenol have shown the remarkable biological properties to mimic diacylglycerol and function as endogenous activators of protein kinase (PKC). Furthermore, they were found to have potential in treatment of pancreatic tumor, actinic keratosis (Siller et al., 2009; Parker et al., 2017), and HIV (Johnson et al., 2008). In conclusion, although *E. fischeriana* and *E. ebracteolata* have been used as the same traditional Chinese medicine for a long time, they should have different focuses on anticancer, anti-HIV, and actinic keratosis treatment in clinical use.

EST-SSRs Are Useful for Population Genetic Analysis Between *Euphorbia*

The development of EST-SSR primers is the further utilization of a large number of EST sequence information and has the characteristics of low cost. *E. fischeriana* and *E. ebracteolata* transcriptome data just provide this data resource. Because EST-SSRs are derived from the relatively conserved transcription part, EST-SSR has a higher translocation between species than genomic SSR, which provides a good tool for studying interspecies population inheritance (Varshney et al., 2005; Kalia et al., 2011). At present, few EST-SSR markers of Langdu have been reported involving *Euphorbia* species (Li et al., 2021a). In this study, 49 of 68 pairs of EST-SSR primers (72.1%) could amplify clear fragments in both Langdu species, 33 pairs (48.5%) were in at least 4 species, 9 were in 8 species, and 2 were in all 9 species (Supplementary Table S6). EST-SSR markers developed in *Euphorbia* showed mobility, indicating that these species may have close genetic relationship and that these markers may have new application value. EST-SSR markers developed based on two Langdu could provide genetic information for research on variety identification, genetic diversity analysis, and molecular marker-assisted breeding.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The two RNA-seq library datasets (root, stem,

and leaf) for this study can be found in the National Center for Biotechnology Information (NCBI) BioProject: PRJNA693977, PRJNA693983.

AUTHOR CONTRIBUTIONS

HZ coordinated the project, directed the work, and revised the manuscript. HZ, M-YY, and YH wrote the manuscript. S-FN, F-QL, and HX collected experimental samples. M-YY performed microsection and image processing. M-YY and C-JP performed an in-depth analysis of comparative transcriptome. X-TZ, BT, R-FJ, and KC performed metabolite determination and analysis. L-QH, X-TZ, and YS involved in funding. All authors contributed to the article and approved the submitted version.

REFERENCES

- Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J. H., Zhang, Z., Miller, W., et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402. doi: 10.1093/nar/25.17.3389
- Barrero, R. A., Chapman, B., Yang, Y., Moolhuijzen, P., Keeble-Gagnere, G., Zhang, N., et al. (2011). De novo assembly of *Euphorbia fischeriana* root transcriptome identifies prostratin pathway related genes. *BMC Genomics* 12, 600. doi: 10.1186/1471-2164-12-600
- Bluthgen, N., Brand, K., Cajavec, B., Swat, M., Herzel, H., and Beule, D. (2005). Biological profiling of gene groups utilizing Gene Ontology. *Genome Inform.* 16, 106–115. doi: 10.48550/arXiv.q-bio/0407034
- Bylesjo, M., Rantalainen, M., Cloarec, O., Nicholson, J. K., Holmes, E., and Trygg, J. (2006). OPLS discriminant analysis: combining the strengths of PLS-DA and SIMCA classification. *J. Chemometr.* 20, 341–351. doi: 10.1002/cem.1006
- Close, T. J. (1997). Dehydrins: a commonality in the response of plants to dehydration and low temperature. *Physiol. Plant.* 100, 291–296. doi: 10.1111/j.1399-3054.1997.tb04785.x
- Conesa, A., Gotz, S., Garcia-Gomez, J. M., Terol, J., Talon, M., and Robles, M. (2005). Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21, 3674–3676. doi: 10.1093/bioinformatics/bti610
- De Vos, R. C. H., Moco, S., Lommen, A., Keurentjes, J. J. B., Bino, R. J., and Hall, R. D. (2007). Untargeted large-scale plant metabolomics using liquid chromatography coupled to mass spectrometry. *Nat. Protoc.* 2, 778–791. doi: 10.1038/nprot.2007.95
- Dieringer, D., and Schlötterer, C. (2003). Microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite data sets. *Mol. Ecol. Notes* 3, 167–169. doi: 10.1046/j.1471-8286.2003.00351.x
- Duan, L.-X., Chen, T.-L., Li, M., Chen, M., Zhou, Y.-Q., Cui, G.-H., et al. (2012). Use of the metabolomics approach to characterize chinese medicinal material Huangqi. *Mol. Plant* 5, 376–386. doi: 10.1093/mp/sss093
- Fiehn, O. (2002). Metabolomics - the link between genotypes and phenotypes. *Plant Mol. Biol.* 48, 155–171. doi: 10.1023/A:1013713905833
- Finn, R. D., Mistry, J., Tate, J., Coghill, P., Heger, A., Pollington, J. E., et al. (2010). The Pfam protein families database. *Nucleic Acids Res.* 38, D211–222. doi: 10.1093/nar/gkp985
- Frajman, B., Zaveska, E., Gamisch, A., Moser, T., Schoenswetter, P., Arthofer, W., et al. (2019). Integrating phylogenomics, phylogenetics, morphometrics, relative genome size and ecological niche modelling disentangles the diversification of Eurasian *Euphorbia seguieriana* s.l. (Euphorbiaceae). *Mol. Phylogenet. Evol.* 134, 238–252. doi: 10.1016/j.ympev.2018.10.046
- Furumoto, T., Yamaguchi, T., Ohshima-Ichie, Y., Nakamura, M., Tsuchida-Iwata, Y., Shimamura, M., et al. (2011). A plastidial sodium-dependent pyruvate transporter. *Nature* 476, 472–U131. doi: 10.1038/nature10250
- Graur, D., and Li, W. H. (2000). *Fundamentals of Molecular Evolution. 2nd Edn.* Sunderland, MA: Sinauer Associates.

FUNDING

This research was funded by CACMS Innovation Fund (CI2021A041003), Natural Science Foundation of China (82104342), Fundamental Research Funds of CACMS (ZZ15-YQ-060), and Independent Studies Supported by CACMS (ZK2021001).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2022.905275/full#supplementary-material>

- Griebel, T., Schutte, D., Ebert, A., Nguyen, H. H., and Baier, M. (2022). Cold exposure memory reduces pathogen susceptibility in Arabidopsis based on a functional plastid peroxidase system. *Mol. Plant Microbe Interact.* doi: 10.1094/MPMI-11-21-0283-FI. [Epub ahead of print].
- Hauser, M.-T. (2014). Molecular basis of natural variation and environmental control of trichome patterning. *Front. Plant Sci.* 5, 320. doi: 10.3389/fpls.2014.00320
- Horn, J. W., van Ee, B. W., Morawetz, J. J., Riina, R., Steinmann, V. W., Berry, P. E., et al. (2012). Phylogenetics and the evolution of major structural characters in the giant genus *Euphorbia* L. (Euphorbiaceae). *Mol. Phylogenet. Evol.* 63, 305–326. doi: 10.1016/j.ympev.2011.12.022
- Huang, Y.-S., Lu, Y., Chen, C.-H., Lee, K.-H., and Chen, D.-F. (2019). Potent anti-HIV ingenane diterpenoids from *Euphorbia ebracteolata*. *J. Nat. Prod.* 82, 1587–1592. doi: 10.1021/acs.jnatprod.9b00088
- Johnson, H. E., Banack, S. A., and Cox, P. A. (2008). Variability in content of the anti-AIDS drug candidate prostratin in samoan populations of *Homalanthus nutans*. *J. Nat. Prod.* 71, 2041–2044. doi: 10.1021/np800295m
- Jung, Y. J., Lee, I. H., Nou, I. S., Lee, K. D., Rashotte, A. M., and Kang, K. K. (2013). BrRZFP1 a Brassica rapa C3HC4-type RING zinc finger protein involved in cold, salt and dehydration stress. *Plant Biol.* 15, 274–283. doi: 10.1111/j.1438-8677.2012.00631.x
- Kalia, R. K., Rai, M. K., Kalia, S., Singh, R., and Dhawan, A. K. (2011). Microsatellite markers: an overview of the recent progress in plants. *Euphytica* 177, 309–334. doi: 10.1007/s10681-010-0286-9
- Kanehisa, M., Araki, M., Goto, S., Hattori, M., Hirakawa, M., Itoh, M., et al. (2008). KEGG for linking genomes to life and the environment. *Nucleic Acids Res.* 36, D480–D484. doi: 10.1093/nar/gkm882
- Kasili, R., Huang, C.-C., Walker, J. D., Simmons, L. A., Zhou, J., Faulk, C., et al. (2011). BRANCHLESS TRICHOMES links cell shape and cell cycle control in *Arabidopsis* trichomes. *Development* 138, 2379–2388. doi: 10.1242/dev.058982
- Klotke, J., Kopka, J., Gatzke, N., and Heyer, A. G. (2004). Impact of soluble sugar concentrations on the acquisition of freezing tolerance in accessions of *Arabidopsis thaliana* with contrasting cold adaptation - evidence for a role of raffinose in cold acclimation. *Plant Cell Environ.* 27, 1395–1404. doi: 10.1111/j.1365-3040.2004.01242.x
- Koch, M. A., Haubold, B., and Mitchell-Olds, T. (2000). Comparative evolutionary analysis of chalcone synthase and alcohol dehydrogenase loci in *Arabidopsis*, *Arabis*, and related genera (Brassicaceae). *Mol. Biol. Evol.* 17, 1483–1498. doi: 10.1093/oxfordjournals.molbev.a026248
- Koenig, D., Jimenez-Gomez, J. M., Kimura, S., Fulop, D., Chitwood, D. H., Headland, L. R., et al. (2013). Comparative transcriptomics reveals patterns of selection in domesticated and wild tomato. *Proc. Natl. Acad. Sci. U. S. A.* 110, E2655–E2662. doi: 10.1073/pnas.1309606110
- Kong, L., Xie, Y., Hu, L., Si, J., and Wang, Z. (2017). Excessive nitrogen application dampens antioxidant capacity and grain filling in wheat as revealed by metabolic and physiological analyses. *Sci. Rep.* 7, 43363. doi: 10.1038/srep43363
- Larson, E. R., Domozych, D. S., and Tierney, M. L. (2014). SNARE VTI13 plays a unique role in endosomal trafficking pathways associated with the vacuole and

- is essential for cell wall organization and root hair growth in arabidopsis. *Ann. Bot.* 114, 1147–1159. doi: 10.1093/aob/mcu041
- Lee, E.-J., Oh, M., Hwang, J.-U., Li-Beisson, Y., Nishida, I., and Lee, Y. (2017). Seed-specific overexpression of the pyruvate transporter BASS2 increases oil content in arabidopsis seeds. *Front. Plant Sci.* 8, 194. doi: 10.3389/fpls.2017.00194
- Li, B., Ma, J., Zhu, H., Gilbert, M. G., Esser, H.-J., Dressler, S., et al. (2008). *Flora of China*. Beijing; St. Louis: Science Press and Missouri Botanical Garden.
- Li, H., Ma, Y., Pei, F., Zhang, H., Liu, J., and Jiang, M. (2021a). Large-scale advances in SSR markers with high-throughput sequencing in *Euphorbia fischeriana* Steud. *Electron. J. Biotechnol.* 49, 50–55. doi: 10.1016/j.ejbt.2020.11.004
- Li, L., Stoeckert, C. J., and Roos, D. S. (2003). OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res.* 13, 2178–2189. doi: 10.1101/gr.1224503
- Li, Y.-N., He, J., Zhang, J., Shi, Y.-X., Guo, L.-B., Peng, Z.-C., et al. (2021b). Existing knowledge on *Euphorbia fischeriana* Steud. (Euphorbiaceae): traditional uses, clinical applications, phytochemistry, pharmacology and toxicology. *J. Ethnopharmacol.* 275, 114095. doi: 10.1016/j.jep.2021.114095
- Lippold, F., vom Dorp, K., Abraham, M., Holzl, G., Wewer, V., Yilmaz, J. L., et al. (2012). Fatty acid phytyl ester synthesis in chloroplasts of *Arabidopsis*. *Plant Cell* 24, 2001–2014. doi: 10.1105/tpc.112.095588
- Liu, S., Chen, H., Li, X., and Zhang, W. (2016). A low-temperature-responsive element involved in the regulation of the *Arabidopsis thaliana* *At1g71850/At1g71860* divergent gene pair. *Plant Cell Rep.* 35, 1757–1767. doi: 10.1007/s00299-016-1994-y
- Lohmeier-Vogel, E. M., Kerk, D., Nimick, M., Wrobel, S., Vickerman, L., Muench, D. G., et al. (2008). Arabidopsis At5g39790 encodes a chloroplast-localized, carbohydrate-binding, coiled-coil domain-containing putative scaffold protein. *BMC Plant Biol.* 8, 120. doi: 10.1186/1471-2229-8-120
- Luo, D., Callari, R., Hamberger, B., Wubshet, S. G., Nielsen, M. T., Andersenranberg, J., et al. (2016). Oxidation and cyclization of casbene in the biosynthesis of Euphorbia factors from mature seeds of *Euphorbia lathyris* L. *Proc. Natl. Acad. Sci. U. S. A.* 16, 34–72. doi: 10.1073/pnas.1607504113
- Mao, Y., Zhang, Y., Xu, C., and Qiu, Y. (2016). Comparative transcriptome resources of two *Diosma* species (Berberidaceae) and molecular evolution of the CYP719A gene in Podophylloideae. *Mol. Ecol. Resour.* 16, 228–241. doi: 10.1111/1755-0998.12415
- Mazie, A. R., and Baum, D. A. (2016). Clade-specific positive selection on a developmental gene: *BRANCHLESS TRICHOME* and the evolution of stellate trichomes in *Physaria* (Brassicaceae). *Mol. Phylogenet. Evol.* 100, 31–40. doi: 10.1016/j.ympev.2016.03.027
- Miana, G. A., Riaz, M., Shahzad-ul-Hussan, S., Paracha, R. Z., and Paracha, U. Z. (2015). Prostratin: an overview. *Mini Rev. Med. Chem.* 15, 1122–1130. doi: 10.2174/1389557515666150511154108
- Mueller, M., Kunz, H.-H., Schroeder, J. I., Kemp, G., Young, H. S., and Neuhaus, H. E. (2014). Decreased capacity for sodium export out of Arabidopsis chloroplasts impairs salt tolerance, photosynthesis and plant performance. *Plant J.* 78, 646–658. doi: 10.1111/tj.12501
- Parker, C. G., Kutttruff, C. A., Galmozzi, A., Jorgensen, L., Yeh, C.-H., Hermanson, D. J., et al. (2017). Chemical proteomics identifies SLC25A20 as a functional target of the ingenol class of actinic keratosis drugs. *ACS Central Sci.* 3, 1276–1285. doi: 10.1021/acscentsci.7b00420
- Perea-Resa, C., Rodriguez-Milla, M. A., Iniesto, E., Rubio, V., and Salinas, J. (2017). Prefoldins negatively regulate cold acclimation in *Arabidopsis thaliana* by promoting nuclear proteasome-mediated HY5 degradation. *Mol. Plant* 10, 791–804. doi: 10.1016/j.molp.2017.03.012
- Prenner, G., and Rudall, P. J. (2007). Comparative ontogeny of the cyathium in *Euphorbia* (Euphorbiaceae) and its allies: exploring the organ-flower-inflorescence boundary. *Am. J. Bot.* 94, 1612–1629. doi: 10.3732/ajb.94.10.1612
- Schauer, N., Semel, Y., Roessner, U., Gur, A., Balbo, I., Carrari, F., et al. (2006). Comprehensive metabolic profiling and phenotyping of interspecific introgression lines for tomato improvement. *Nat. Biotechnol.* 24, 447–454. doi: 10.1038/nbt1192
- Sekula, B., Ruskowski, M., and Dauter, Z. (2020). S-adenosylmethionine synthases in plants: structural characterization of type I and II isoenzymes from *Arabidopsis thaliana* and *Medicago truncatula*. *Int. J. Biol. Macromol.* 151, 554–565. doi: 10.1016/j.ijbiomac.2020.02.100
- Siller, G., Gebauer, K., Welburn, P., Katsamas, J., and Ogbourne, S. M. (2009). PEP005 (ingenol mebutate) gel, a novel agent for the treatment of actinic keratosis: results of a randomized, double-blind, vehicle-controlled, multicentre, phase IIa study. *Australas. J. Dermatol.* 50, 16–22. doi: 10.1111/j.1440-0960.2008.00497.x
- Su, T., Tan, Y., Tsui, M.-S., Yi, H., Fu, X.-Q., Li, T., et al. (2016). Metabolomics reveals the mechanisms for the cardiotoxicity of Pinelliae Rhizoma and the toxicity-reducing effect of processing. *Sci. Rep.* 6, 34692. doi: 10.1038/srep34692
- Swanson, W. J., Wong, A., Wolfner, M. F., and Aquadro, C. F. (2004). Evolutionary expressed sequence tag analysis of *Drosophila* female reproductive tracts identifies genes subjected to positive selection. *Genetics* 168, 1457–1465. doi: 10.1534/genetics.104.030478
- Szabados, L., and Savoure, A. (2010). Proline: a multifunctional amino acid. *Trends Plant Sci.* 15, 89–97. doi: 10.1016/j.tplants.2009.11.009
- Varshney, R. K., Graner, A., and Sorrells, M. E. (2005). Genic microsatellite markers in plants: features and applications. *Trends Biotechnol.* 23, 48–55. doi: 10.1016/j.tibtech.2004.11.005
- Wagner, G. J., Wang, E., and Shepherd, R. W. (2004). New approaches for studying and exploiting an old protuberance, the plant trichome. *Ann. Bot.* 93, 3–11. doi: 10.1093/aob/mch011
- Wang, C.-J., Yan, Q.-L., Ma, Y.-F., Sun, C.-P., Chen, C.-M., Tian, X.-G., et al. (2017a). ent-Abietane and tiglane diterpenoids from the Roots of *Euphorbia fischeriana* and their inhibitory effects against *Mycobacterium smegmatis*. *J. Nat. Prod.* 80, 1248–1254. doi: 10.1021/acs.jnatprod.6b00786
- Wang, L., Qian, B., Zhao, L., Liang, M.-H., Zhan, X., and Zhu, J. (2022). Two triacylglycerol lipases are negative regulators of chilling stress tolerance in arabidopsis. *Int. J. Mol. Sci.* 23, 3380. doi: 10.3390/ijms23063380
- Wang, L. J., Li, H. H., Zhao, C. Z., Li, S. F., Kong, L. Y., Wu, W. W., et al. (2017b). The inhibition of protein translation mediated by AtGCN1 is essential for cold tolerance in *Arabidopsis thaliana*. *Plant Cell Environ.* 40, 56–68. doi: 10.1111/pce.12826
- Wang, W., Ryu, K. H., Barron, C., and Schiefelbein, J. (2019). Root epidermal cell patterning is modulated by a critical residue in the WEREWOLF transcription factor. *Plant Physiol.* 181, 1239–1256. doi: 10.1104/pp.19.00458
- Wang, Y., Ma, X., Yan, S., Shen, S., Zhu, H., Gu, Y., et al. (2009). 17-Hydroxy-jolkinolide B inhibits signal transducers and activators of transcription 3 signaling by covalently cross-linking janus kinases and induces apoptosis of human cancer cells. *Cancer Res.* 69, 7302–7310. doi: 10.1158/0008-5472.CAN-09-0462
- Wanner, L. A., and Junttila, O. (1999). Cold-induced freezing tolerance in Arabidopsis. *Plant Physiol.* 120, 391–399. doi: 10.1104/pp.120.2.391
- Wei, G., Dong, L., Yang, J., Zhang, L., Xu, J., Yang, F., et al. (2018). Integrated metabolomic and transcriptomic analyses revealed the distribution of saponins in *Panax notoginseng*. *Acta Pharmaceutica Sinica B* 8, 458–465. doi: 10.1016/j.apsb.2017.12.010
- Xu, Z. H., Sun, J., Xu, R. S., and Qin, G. W. (1998). Casbene diterpenoids from *Euphorbia ebracteolata*? *Phytochemistry* 49, 149–151. doi: 10.1016/S0031-9422(97)00900-X
- Yang, T., He, J., Yan, Y., Lian, W.-W., Xia, C.-Y., Xu, J.-K., et al. (2021). *Euphorbia ebracteolata* Hayata (Euphorbiaceae): a systematic review of its traditional uses, botany, phytochemistry, pharmacology, toxicology, and quality control. *Phytochemistry* 186, 112736. doi: 10.1016/j.phytochem.2021.112736
- Yang, Z. (2007). PAML 4: Phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* 24, 1586–1591. doi: 10.1093/molbev/msm088
- Ye, J., Fang, L., Zheng, H., Zhang, Y., Chen, J., Zhang, Z., et al. (2006). WEGO: a web tool for plotting GO annotations. *Nucleic Acids Res.* 34, W293–W297. doi: 10.1093/nar/gkl031
- Yuan, Y., and Huang, L. (2020). Molecular pharmacognosy in Daodi herbs. *Chin. Sci. Bull. Chin.* 65, 1093–1102. doi: 10.1360/TB-2020-0184
- Zhang, L., Yan, H.-F., Wu, W., Yu, H., and Ge, X.-J. (2013). Comparative transcriptome analysis and marker development of two closely related Primrose species (*Primula poissonii* and *Primula wilsonii*). *BMC Genomics* 14, 329. doi: 10.1186/1471-2164-14-329
- Zhao, K. J., X. G., Jin, R. L., and Xu, L. S. (1996). Herbalogical study on traditional chinese drug langdu and its confused species lvru. *Chin. Trad. Herbal Drugs* 27, 554–558. doi: 10.7501/j.issn.0253-2670.1996.9.302

- Zhao, S.-Y., Chen, L.-Y., Muchuku, J. K., Hu, G.-W., and Wang, Q.-F. (2016). Genetic adaptation of giant lobelias (*Lobelia aberdarica* and *Lobelia telekil*) to different altitudes in East African mountains. *Front. Plant Sci.* 7, 488. doi: 10.3389/fpls.2016.00488
- Zhao, Y., Zhang, J., Zhang, Z., and Xie, W. (2019). Elymus nutans genes for seed shattering and candidate gene-derived EST-SSR markers for germplasm evaluation. *BMC Plant Biol.* 19, 102. doi: 10.1186/s12870-019-1691-4
- Zhao, Y.-J., Cao, Y., Wang, J., and Xiong, Z. (2018). Transcriptome sequencing of *Pinus kesiya* var. *langbianensis* and comparative analysis in the *Pinus* phylogeny. *BMC Genom.* 19, 725. doi: 10.1186/s12864-018-5127-6
- Zheng, H., Jing, L., Jiang, X., Pu, C., Zhao, S., Yang, J., et al. (2021). The ERF-VII transcription factor SmERF73 coordinately regulates tanshinone biosynthesis in response to stress elicitors in *Salvia miltiorrhiza*. *New Phytol.* 231, 1940–1955. doi: 10.1111/nph.17463
- Zhu, Y., Rong, L., Luo, Q., Wang, B., Zhou, N., Yang, Y., et al. (2017). The histone chaperone NRP1 interacts with WEREWOLF to activate *GLBRA2* in arabidopsis root hair development. *Plant Cell* 29, 260–276. doi: 10.1105/tpc.16.00719

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Zheng, Yu, Han, Tai, Ni, Ji, Pu, Chen, Li, Xiao, Shen, Zhou and Huang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.