



# Comparison of Pulegone and Estragole Chemotypes Provides New Insight Into Volatile Oil Biosynthesis of *Agastache rugosa*

Jingjie Dang<sup>1,2†</sup>, Guyin Lin<sup>1,2†</sup>, Licheng Liu<sup>1,2</sup>, Peina Zhou<sup>1,2</sup>, Yongfang Shao<sup>1,2</sup>, Shilin Dai<sup>1,2</sup>, Mengru Sang<sup>1,2</sup>, Zheng Jiang<sup>1,2</sup>, Chanchan Liu<sup>1,2\*</sup> and Qinan Wu<sup>1,2\*</sup>

<sup>1</sup> College of Pharmacy, Nanjing University of Chinese Medicine, Nanjing, China, <sup>2</sup> Collaborative Innovation Center of Chinese Medicinal Resources Industrialization, Nanjing, China

## OPEN ACCESS

### Edited by:

Dinesh A. Nagegowda,  
Central Institute of Medicinal  
and Aromatic Plants, Council  
of Scientific and Industrial Research  
(CSIR), India

### Reviewed by:

Debabrata Sircar,  
Indian Institute of Technology  
Roorkee, India  
Chenyi Li,  
University of Georgia, United States

### \*Correspondence:

Chanchan Liu  
liuchanchan@njucm.edu.cn  
Qinan Wu  
wuqn@njucm.edu.cn

† These authors have contributed  
equally to this work and share first  
authorship

### Specialty section:

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

Received: 07 January 2022

Accepted: 01 March 2022

Published: 06 April 2022

### Citation:

Dang J, Lin G, Liu L, Zhou P,  
Shao Y, Dai S, Sang M, Jiang Z, Liu C  
and Wu Q (2022) Comparison  
of Pulegone and Estragole  
Chemotypes Provides New Insight  
Into Volatile Oil Biosynthesis  
of *Agastache rugosa*.  
*Front. Plant Sci.* 13:850130.  
doi: 10.3389/fpls.2022.850130

The aerial parts of *Agastache rugosa* are rich in essential oils containing monoterpenoids, phenylpropanoids, and aromatic compounds. These are used as herbs, perfume plants, and ornamental plants. Based on the difference in the constituents of the essential oil, *A. rugosa* is divided into pulegone and estragole chemotypes, but the mechanism of key metabolite biosynthesis in these two *A. rugosa* chemotypes remains unclear. In this study, we compared the morphological differences, metabolite constituents, and transcriptomic data between the two chemotypes of *A. rugosa*. Monoterpenoid was the main compound in the pulegone chemotype, and phenylpropanoid was the main compound in the estragole chemotype; however, limonene was detected in both chemotypes. Furthermore, 46 genes related to pulegone and estragole biosynthesis were identified. *Limonene synthase*, *limonene-3-hydroxylase*, and *isopiperitenol dehydrogenase* were upregulated in the pulegone chemotype, while *phenylalanine ammonia-lyase*, *4-coumarate: CoA ligase*, *CYP73A*, *coumaroyl-aldehyde dehydrogenase*, and *eugenol synthase* were downregulated in the pulegone chemotype. We identified chavicol methyl transferase and limonene-3-hydroxylase in *A. rugosa*. This work not only provides the difference in morphology and metabolites in pulegone and estragole chemotypes, but also offers a molecular mechanism of volatile oil biosynthesis, which could be a basis for specialized metabolites in specialized chemotypes.

**Keywords:** *Agastache rugosa*, transcriptome sequencing, chemotypes, monoterpenoid biosynthesis, phenylpropanoids biosynthesis

## INTRODUCTION

*Agastache rugosa* (Fisch. & C.A.Mey.) Kuntze (*A. rugosa*), belonging to the family Lamiaceae is one of the best-known genera for its medicinal properties and economically important aromatic oils. *A. rugosa* has traditionally been used to treat naupathia, emesis, and to dispel dampness (Cao et al., 2017). Studies have indicated that *A. rugosa* has a variety of pharmacological actions, such as antifungal, anti-oxygenation, and anti-inflammatory activities. The aerial parts of *A. rugosa* and their constituent compounds have the potential for the development of natural nematicides (Li et al., 2013).

Based on the differences in chemical compounds within the same species, the plants were divided into different chemotypes. Chemotypes are common in plants, and most of them are concentrated in the Lamiaceae and Asteraceae families, such as *Artemisia annua* and *Mentha canadensis* Linnaeus (Šarić-Kundalić et al., 2009; Czechowski et al., 2018). The morphological differences between a variety of chemotypes were mainly in leaf type, trichome density, plant height, and internode length (Czechowski et al., 2018; Węglarz et al., 2020). The existence of a chemotype has been widely reported, but the mechanism of its generation is not clear. Research on the chemotypes of *A. rugosa* has mainly focused on differences in chemical composition. The chemical composition of the essential oils obtained from the aerial parts (stems, flowers, and leaves) of *A. rugosa* grown in different regions has been the subject of some studies (Skakovskii et al., 2010; Mo and Ma, 2011). Li et al. found that pulegone, estragole, and methyl eugenol were the principal components of *A. rugosa* aerial parts collected from Zhejiang, Hubei, and Henan province of China (Li et al., 2013) which were represented by chemotypes. However, studies on morphological and biosynthesis differences in the chemotypes of *A. rugosa* have not been conducted.

To better understand the differences in the chemotypes of *A. rugosa* and the molecular mechanism of their formation, we systematically compared morphological data and volatile oil components in the pulegone (ArP) and estragole (ArE) chemotypes of *A. rugosa*, the genes related to pulegone and estragole biosynthesis were identified using *de novo* transcriptome sequencing, and the functions of L3OH and CVOMT were verified *in vitro*. Quantitative real-time PCR (qRT-PCR) further validated the gene expression levels of the two chemotypes in *A. rugosa*.

## MATERIALS AND METHODS

### Plant Materials and Growth Condition

The seeds of *Agastache rugosa* (Fisch. et Mey.) O.Ktze. were collected from Jiangxi, China, and grown at Nanjing University of Chinese Medicine, Jiangsu, China. The growth condition was described as Liu L. et al. (2021).

### Trichome Density Measurements

Trichome density was quantified on the abaxial surface of the terminal leaflets of leaves 3–5 (counted from the apical meristem). Trichomes were visualized using a stereomicroscope (SetREO Discovery. V20, Zeiss, Germany). Images were recorded using Axio Vision 4.7 software (Carl Zeiss Ltd., Herts., United Kingdom). The trichome number was counted manually across a measured leaf area (automatically calculated by the software), and the average (mean) trichome density was then calculated for the whole leaf.

### Simultaneous Distillation Extraction of Essential Oils

Leaves, stems, and roots were harvested at the lush stage (60 day after germination). Five grams of fresh sample

material was transferred to a 500 mL round bottom flask with 100 mL of deionized water for subsequent simultaneous distillation extraction using a modified Likens-Nickerson apparatus and n-hexane as the carrier solvent. An aliquot of the n-hexane fraction, which contained volatile oil constituents, was transferred to a 2 mL glass vial for further analysis.

### Essential Oils Analysis

Gas chromatography-mass spectrometry (GC-MS) was performed using a 6890N GC interfaced with 5973 inert MS instrument (Agilent Technologies, Santa Clara, CA, United States) and chromatographic column (Agilent 19091S-433-HP-5ms, 30 m × 250 μm × 0.25 μm). GC/MS procedure: kept at 50°C for 3 min, elevated to 90°C at 3°C/min, and then raised to 220°C at 10°C/min, and maintained for 5 min. The injection volume was 1.0 μL and splitless. The temperature of the injection port was 220°C, the EI ion source was 230°C, the electron energy was 70 eV, and the connection line temperature was 150°C. The MS scan range (m/z) was 50–500 m/z. The scanning speed was 2 times/s, and the delay time was 2.5 min. For analysis, essential oils dissolved in n-hexane were directly injected. Metabolites were identified based on chromatographic and spectral comparisons against the NIST/EPA/NIH mass spectral library version 2.0 and standards. The percentage of each chemical constituent in the essential oil was determined by the normalization of the peak.

Gas chromatography (GC) was performed using an 8860 GC (Agilent Technologies, Santa Clara, CA, United States) and chromatographic column (Agilent 19091 J-413-HP-5, 30 m × 0.32 mm × 0.25 μm). The GC procedure was maintained at 50°C for 3 min, elevated to 90°C at 3°C/min, and then raised to 220°C at 10°C/min, and maintained for 5 min. The injection volume was 1.0 μL and splitless. The temperature of the injection port was 250°C, and the flame ionization detector temperature was 250°C.

Chiral-GC was performed using an 8860 GC (Agilent Technologies, Santa Clara, CA, United States) and chromatographic column (Agilent 112-5532-DB-5ms, 30 m × 0.25 mm × 0.25 μm). Chiral-GC procedure: kept at 40°C for 0 min, elevated to 90°C at 2°C/min, and then raised to 120°C at 1°C/min, and maintained for 2 min. The injection volume was 2.0 μL and splitless. The temperature of the injection port was 250°C, and the flame ionization detector temperature was 250°C.

### Total RNA Extraction, mRNA Library Construction, and Sequencing

Leaf samples with estragole or pulegone as the first principal component of essential oil were collected separately and quickly frozen in liquid nitrogen. Three biological replicates of the leaves of each chemotype were collected for RNA sequencing (RNA-Seq), each consisting of four young and two old leaves from the same plant. Total RNA was extracted using TRIzol Reagent (Sigma, United States). RNA samples of *A. rugosa* were separately sequenced using the BGISEQ-500 Transcriptome platform, in

which the mRNA was purified from total RNA using Oligo (dT)-attached magnetic beads, and the PCR products were purified using Ampure XP Beads and library quality was assessed on the Agilent Technologies 2100 bioanalyzer. The final library was amplified with phi29 to form a DNA nanoball (DNB), which had more than 300 copies of one molecule, DNBs were loaded into the patterned nanoarray, and pair-end 100 bases reads were generated on the BGISEQ500 platform (BGI-Shenzhen, China).

## Assembly, Annotation, and Differential Gene Expression Analysis

The raw data were cleaned by removing low-quality and adapter reads. Transcriptome *de novo* assembly of clean reads was conducted using Trinity (Love et al., 2014). After assembly, unigenes were obtained and annotation was conducted using the following public databases: GO, KEGG, KOG, NR, NT, Pfam, PRG, and Swissprot. Differential gene expression analyses of the two chemotypes were performed using genes with an adjusted *P*-value  $\leq 0.5$  and were considered to be differentially expressed. According to the results of differential gene expression, the PHYPER function in R software was used for the enrichment analysis.

## Quantitative Real-Time PCR Analysis

The qRT-PCR analysis was conducted to verify the expression of selected genes in *A. rugosa*. The Hiscript III 1st Strand cDNA Synthesis Kit (+gDNA wiper) (Vazyme Biotech, Nanjing, China) was used with 500 ng of total RNA to synthesize first-strand cDNA. The qRT-PCR reactions were conducted using the ChamQ Universal SYBR qPCR Master Mix Kit (Vazyme Biotech, Nanjing, China). The reaction system and steps were performed according to the manufacturer's instructions. The *A. rugosa* actin gene was used as a control to normalize the relative expression levels of the target genes. All results are representative of the three independent experiments. The primers used for qRT-PCR are listed in **Supplementary Table 1**.

## Isolation of Full-Length Limonene-3-Hydroxylase and Chavicol Methyl Transferase

Single-strand cDNA was used as the template for PCR amplification of the target cDNA with  $2 \times$  PrimeSTAR<sup>®</sup>Max DNA Polymerase (Takara, Japan) and gene-specific primers designed based on annotated results from the transcriptome database. The PCR products were separated using the GeneJET<sup>™</sup> Gel Extraction Kit (Thermo Scientific), and the purified DNA fragment was subcloned into the pCE2TA/Blunt-Zero vector (Vazyme Biotech, Nanjing, China) for sequencing. The primers used for cloning are listed in **Supplementary Table 2**.

## Expression of Limonene-3-Hydroxylase in *Saccharomyces cerevisiae* and Enzyme Assay

The limonene-3-hydroxylase open read frame (ORF) was cloned into the vector pYeDP60 using the ClonExpress II

One Step Cloning Kit (Vazyme Biotech, Nanjing, China). The obtained pYeDP60-L3OH was introduced into *Saccharomyces cerevisiae* WAT11 using a transformation kit (Frozen-EZ Yeast Transformation II Kit, ZYMO RESEARCH CORP.). The empty vector was used as a control. Transformation was plated on SGI medium containing 20 g/L glucose, 6.7 g/L yeast nitrogen base without amino acids, 1 g/L bacto casamino acids (Difco), 15 g/L agar, and 40 mg/L DL-tryptophan.

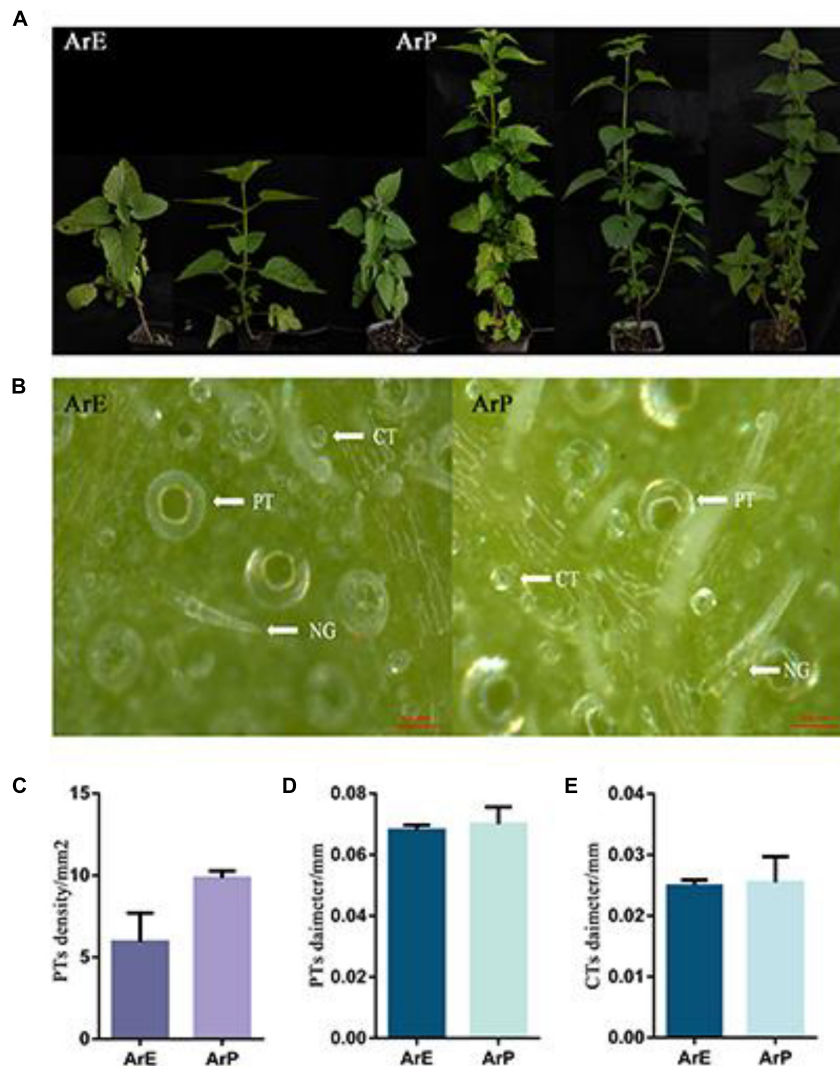
A single colony from SGI medium was grown in 50 mL of SGI liquid medium at 30°C for 12 h. The cells were centrifuged at  $1630 \times g$  for 5 min and resuspended in SGI medium containing 20 g/L galactose. The suspension was diluted OD600 of 0.4 and induced at 16°C for 12 h. Limonene solution was added to the culture to a final concentration of 0.2 mM. The reaction was stopped by sonication for 15 min for 12 h. The products from each reaction were extracted with 200  $\mu$ L n-hexane. After the enzyme reactions, the glass vial was placed in a freezer ( $-80^\circ\text{C}$ ) for 2 h (Zhou et al., 2016) the upper organic phase was transferred into a new 2 mL glass vial containing a conical glass insert (while the still frozen aqueous phase of the original glass vial was discarded) and analyzed by GC-MS. Negative controls were generated by heating the reconstitution mix for 5 min at 100°C. The GC-MS column is described above. The injection temperature was 250°C and the solvent was delayed for 3 min. The oven was set at a temperature program of 50°C for 2 min, increased at 12°C/min to 100°C, and held for 25 min (Liu L. et al., 2021).

## Expression of Chavicol Methyl Transferase in *Escherichia coli* and Enzyme Assay

The chavicol methyl transferase ORF was cloned into pET28a to express CVOMT-His recombinant proteins. The constructs were transformed into *Escherichia coli* BL21 cells (TIANGEN, Beijing) using the heat shock treatment method at 42°C for 60 s. The transformation was plated on LB medium containing 100 ng/L kanamycin.

Single colonies were used to inoculate 5 mL of LB medium containing 100 ng/L kanamycin. One milliliter of the culture was transferred into 100 mL of the same medium and continued to grow at 37°C to reach an OD600 of approximately 1.0. Protein expression was induced by adding isopropyl  $\beta$ -thiogalactopyranoside to a final concentration of 1 mM. After 24 h incubation at 16°C, the cells were harvested by centrifugation (15 min at 4°C and 5,000 rpm). The cells were resuspended in 5 mL lysis buffer containing 10 mM/L Tris-HCl (pH 8.0), 200 mM NaCl, and 5% (v/v) glycerine. The suspension was transferred to an ultrasonic cell disruptor system and ultrasonically disrupted for 25 s each with 35 s on ice between disruptions. The mixture was centrifuged for 30 min at 4°C and 12,000 rpm and the supernatant was carefully removed (Liu C. et al., 2021).

Enzyme assays were performed using 30  $\mu$ L protein extraction, 10 mM chavicol and 10 mM S-adenosylmethionine in buffer (50 mM Tris-HCl, pH 7.0, containing 10% glycerol, 14 mM 2-mercaptoethanol, 1 mM EDTA, and 10 mM NaCl) to final



**FIGURE 1 |** Morphological characterization of ArP and ArE of *A. rugosa*. **(A)** Photographs show three representative 12-week-old plants from the two chemotypes of *A. rugosa*. **(B)** Morphological features of trichomes on leaf surfaces, peltate trichome (PT), capitate trichome (CT), and non-glandular hairs (NG). **(C)** Peltate trichome density of ArE and ArP. Statistical significance was determined using the *t*-test ( $P < 0.05$ ). **(D)** Peltate trichome diameter of ArE and ArP. Statistical significance was determined using the *t*-test ( $P > 0.05$ ). **(E)** Capitate trichome diameter of ArE and ArP.

volume of 400  $\mu$ L in a 2 mL glass vial. The mixture was incubated at 31°C for 1 h and agitated by slowly rotating the glass vials (Gang et al., 2002). After the enzyme reactions, the glass vial was placed in a freezer (−80°C) for 2 h. The reaction was stopped by sonication for 15 min. The products from each reaction were extracted with 200  $\mu$ L n-hexane. After the enzyme reactions, the glass vial was placed in a freezer (−80°C) for 2 h (Zhou et al., 2016) the upper organic phase was transformed into a new 2 mL glass vial containing a conical glass insert (while the still frozen aqueous phase of the original glass vial was discarded) and analyzed by GC-MS. Negative controls were generated by heating the reconstitution medium n-hexane for 5 min at 100°C. The GC-MS column is described above. The injection temperature was 250°C and the solvent was delayed for 3 min. The oven was set at a temperature program of 50°C for 3 min, then increased at

3°C/min to 90°C, held for 0 min, increased at 10°C/min to 220°C, and held for 5 min.

## RESULTS

### Morphological Difference of Two Chemotypes of *Agastache rugosa*

In addition to having a distinct chemical composition, the ArP and ArE of *A. rugosa* also have distinct morphological features. The morphological characteristics of ArP and ArE in *A. rugosa* are presented in **Figure 1**. The two chemotypes of *A. rugosa* are both perennial herbs in the Lamiaceae family, with four-prism stems and heart-ovate leaves with coarsely toothed edges. The plant height of the ArP was 1.79 times higher than that of the ArE,



**TABLE 1** | The content of the main component in volatile oil from different parts of *A. rugosa*.

ID	Compound	CAS	Formula	Retention Time/min	Retention Index	Matching Index	Relative content (%)					
							ArP			ArE		
							Leaf	Stem	Root	Leaf	Stem	Root
1	$\beta$ -Pinene	127-91-3	C <sub>10</sub> H <sub>16</sub>	10.675	979	854	0.67	0.61	–	–	–	–
2	(+)-Limonene	5989-27-5	C <sub>10</sub> H <sub>16</sub>	12.259	1030	921	9.04	22.27	–	2.68	5.83	–
3	(+)-Menthone	1196-31-2	C <sub>10</sub> H <sub>18</sub> O	17.769	1154	940	2.53	4.11	–	–	–	–
4	(-)-Isomenthone	18309-28-9	C <sub>10</sub> H <sub>18</sub> O	18.115	1164	911	21.44	35.99	–	–	–	–
5	(+)-Isopulegone	29606-79-9	C <sub>10</sub> H <sub>16</sub> O	18.483	1177	849	0.88	–	–	–	–	–
6	Estragole	140-67-0	C <sub>10</sub> H <sub>12</sub> O	19.140	1196	926	–	–	–	90.44	92.26	–
7	(-)-Pulegone	3391-90-0	C <sub>10</sub> H <sub>16</sub> O	20.137	1237	895	59.01	36.27	–	–	–	–
8	2-Cyclohexen-1-one, 3-methyl-6-(1-methylethylidene)-	491-09-8	C <sub>10</sub> H <sub>14</sub> O	22.160	1340	838	0.56	–	–	–	–	–
9	Methyleugenol	93-15-2	C <sub>11</sub> H <sub>14</sub> O <sub>2</sub>	23.143	1402	841	–	–	–	6.89	1.92	–
10	Caryophyllene	87-44-5	C <sub>15</sub> H <sub>24</sub>	23.463	1419	856	0.58	0.20	–	–	–	–
11	Germacrene D	23986-74-5	C <sub>15</sub> H <sub>24</sub>	24.360	1481	897	1.29	–	–	–	–	–
12	$\gamma$ -Elemene	29873-99-2	C <sub>15</sub> H <sub>24</sub>	24.574	1433	873	2.11	0.56	–	–	–	–
13	$\beta$ -Cadinene	523-47-7	C <sub>15</sub> H <sub>24</sub>	24.901	1518	807	0.40	–	–	–	–	–
14	$\gamma$ -Muurolene	30021-74-0	C <sub>15</sub> H <sub>24</sub>	25.624	1477	814	1.06	–	–	–	–	–
15	$\alpha$ -Cadinol	481-34-5	C <sub>15</sub> H <sub>26</sub> O	26.572	1653	832	0.42	–	–	–	–	–

and the internode length was up to 1.36 times longer (**Figure 1**). The spike of the ArE grows at the top of the plant, while the spike of ArP grows at the top and axil of the plant, like that of peppermint and spearmint. The morphology of the flowers in the two chemotypes were similar: the corolla was pale blue-purple, and 2-lipped, the upper lip was straight, the apex was slightly emarginate, the lower lip was 3-lobed, and four stamens extended out of the corolla. The styles are subequal to stamens, filiform, and the apex equally 2-lobed (**Supplementary Figure 1**). The lower surface of the leaves of the two chemotypes of *A. rugosa* was evenly distributed with trichomes, and there were three types: peltate trichomes, capitate trichomes, and non-glandular hairs. The density of peltate trichomes in young leaves was higher than that in old leaves, and the diameter of peltate trichomes and capitate trichomes in young leaves was smaller than that in old leaves. There was a significant difference in the density of peltate trichomes between two chemotypes. The differences between ArP and ArE were slight in the diameter of peltate trichomes and capitate trichomes (**Figure 1**).

### Composition of Volatile Component in Different Organs of Two Chemotypes

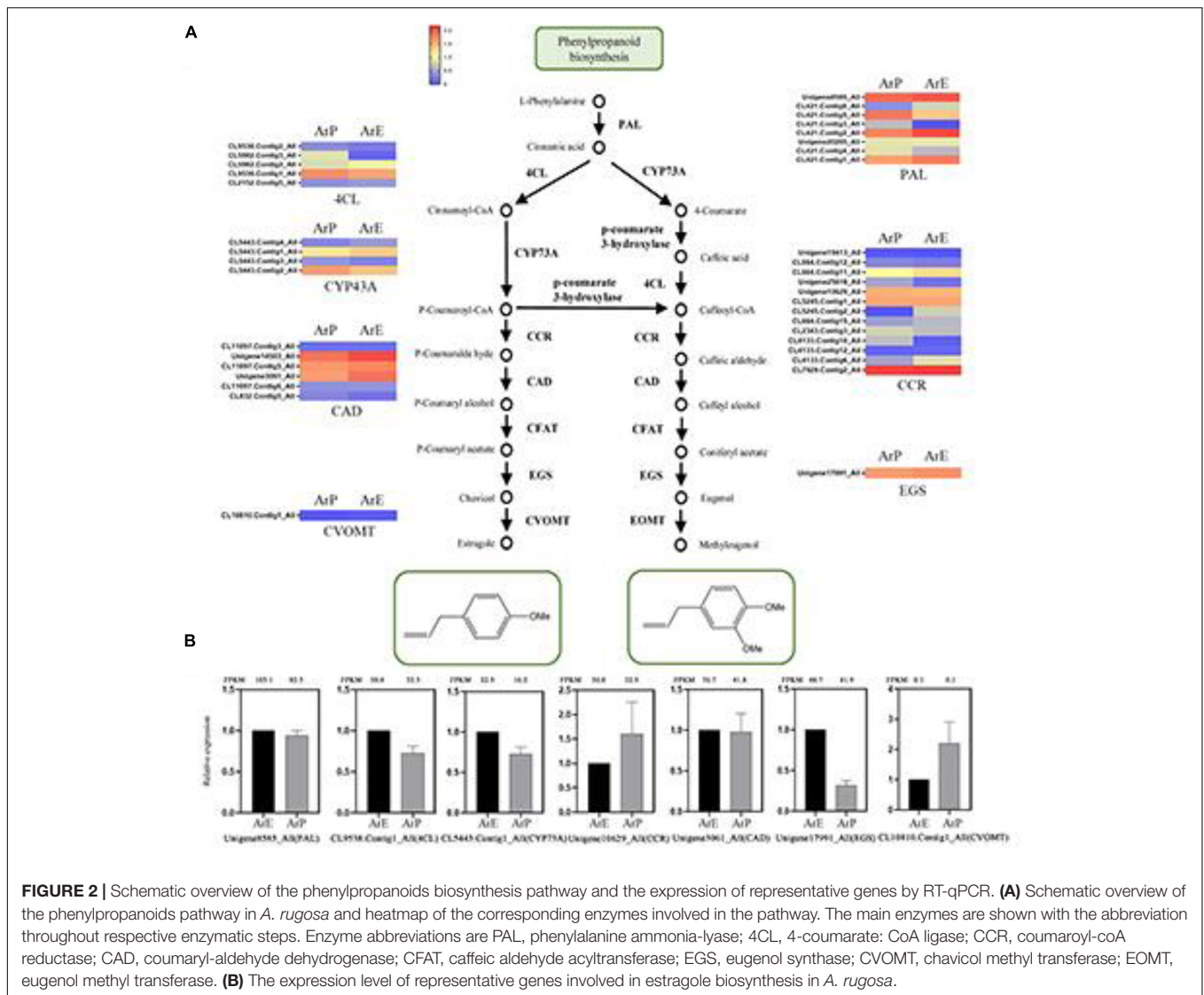
As shown in **Table 1**, a total of 15 volatile compounds were detected in the two chemotypes by GC-MS analysis. The essential oil of ArP was partially dominated by (-)-pulegone, (-)-isomenthone, and (+)-limonene, which are monoterpenoids (**Supplementary Figure 2**). Thirteen compounds were detected in leaves. The total composition of (-)-pulegone, (+)-menthone, (-)-isomenthone, and (+)-limonene was up to 89% in leaves. Leaves had the highest levels of (-)-pulegone (59.01%), with significant quantities of (+)-limonene (9.04%) and only a small amount of (+)-menthone (<3%). There were fewer varieties of

compounds in the stems than in the leaves. Compared to leaves, the composition of (+)-limonene in the stem was much higher than that in the leaves, which reached 22.27%. The composition of (+)-menthone in stems was also higher than that in the leaves. The content of (-)-pulegone was lower than that of the leaves. No essential oil was detected in the roots of ArP. Moreover, isopulegone was not detected in the essential oil of the stem. Estragole and methyl eugenol were not detected in ArP. In addition to monoterpenoids, caryophyllene, germacrene D,  $\gamma$ -elemene,  $\beta$ -cadinene,  $\gamma$ -muurolene, and  $\alpha$ -cadinol were detected in ArP, which were not present in the essential oil of ArE.

Only three compounds were detected in ArE; the main contents of the essential oil were estragole and caryophyllene. The composition of estragole was the highest in the stem and leaves of ArE, which was up to 90%. In addition, (+)-limonene was detected in the leaves and stems, while other monoterpenoids were not detected in ArE.

### RNA Sequencing and *de novo* Transcriptomic Assembly

We constructed separate cDNA libraries from the two chemotypes of the *A. rugosa* samples, and the cDNA was sequenced using the DNBSEQ platform (Mahmood et al., 2020). Six sequencing libraries were constructed, including three replicates for both estragole and pulegone chemotypes. Each sample produced at least 42 Mb clean reads and 6 GB of data. After filtering and removing the low quality, linker contamination, and high N content of unknown reads from the raw data, the percentage of clean reads was more than 91% and Q20 percentage was at least 95.3% for the six libraries (**Supplementary Table 3**). High-quality reads were selected for further analysis, and clean reads were assembled using the



**FIGURE 2 |** Schematic overview of the phenylpropanoids biosynthesis pathway and the expression of representative genes by RT-qPCR. **(A)** Schematic overview of the phenylpropanoids pathway in *A. rugosa* and heatmap of the corresponding enzymes involved in the pathway. The main enzymes are shown with the abbreviation throughout respective enzymatic steps. Enzyme abbreviations are PAL, phenylalanine ammonia-lyase; 4CL, 4-coumarate: CoA ligase; CCR, coumaroyl-coA reductase; CAD, coumaryl-aldehyde dehydrogenase; CFAT, caffeic aldehyde acyltransferase; EGS, eugenol synthase; CVOMT, chavicol methyl transferase; EOMT, eugenol methyl transferase. **(B)** The expression level of representative genes involved in estragole biosynthesis in *A. rugosa*.

Trinity program. We then used Tgicl to cluster transcripts to remove redundancy to obtain unigenes, named “All-Unigene” (Haas et al., 2013). Finally, 100852 unigenes were assembled with a GC percentage of 41.37%, and the average length and N50 lengths were 1530 and 2323 bp, respectively (Supplementary Table 4). The length distribution of the unigenes is indicated in Figure 2, and 85.17% of all unigenes had lengths longer than 300 bp. Benchmarking Universal Single-Copy Orthologs (BUSCO) was used to evaluate the quality of the assembled transcripts (Simão et al., 2015). By comparing them with the conserved genes, it shows the integrity of the transcriptome assembly to a certain extent. The assembled evaluation results are shown in Supplementary Figure 3.

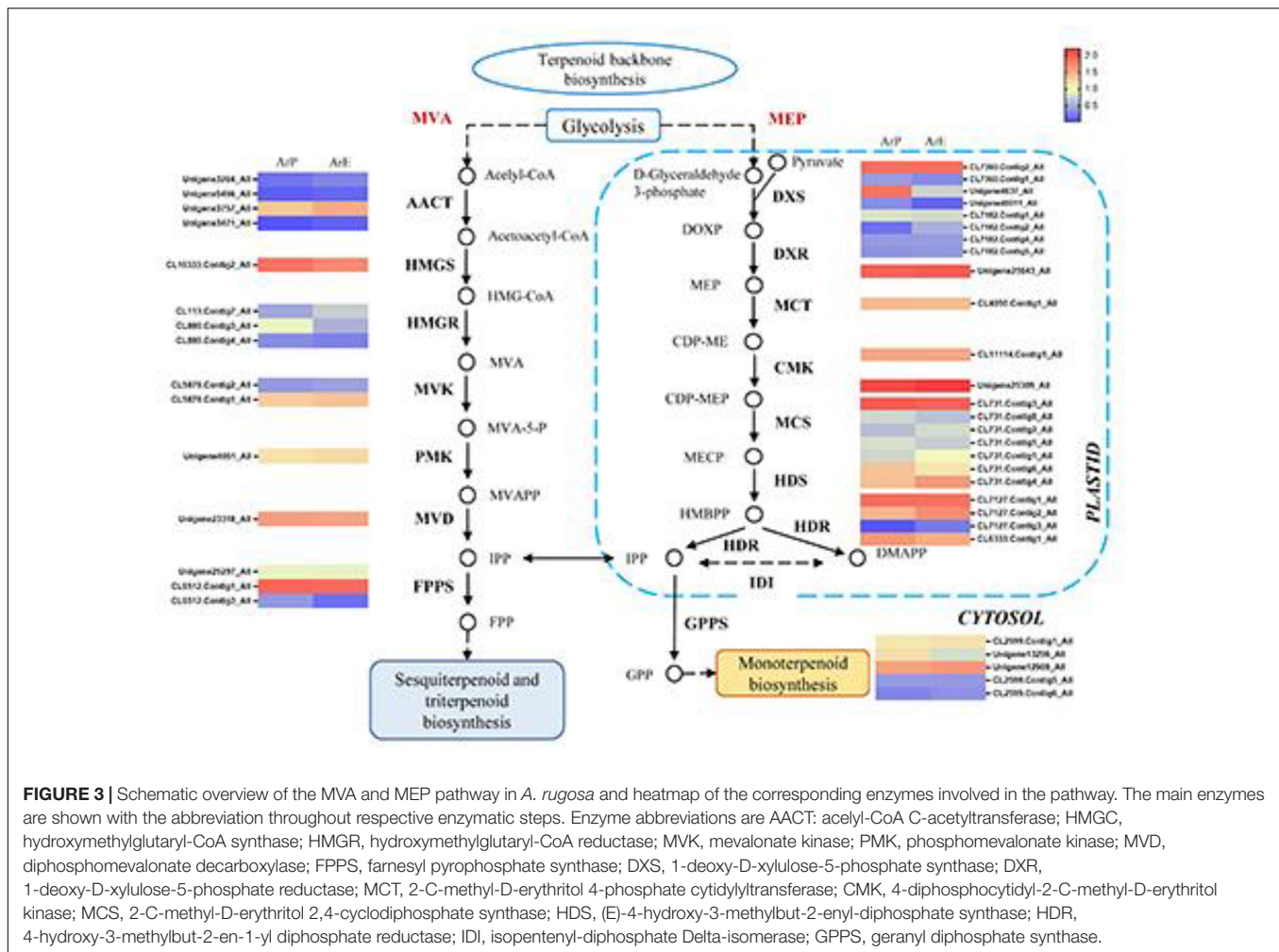
### Functional Annotation of Unigenes

The gene functions of the assembled transcripts were annotated using public databases (Supplementary Figure 4). The results showed that 39499 unigenes were annotated in five public

databases. GO, KOG, and KEGG assignments were used to classify the functions of the assembled unigenes of *A. rugosa* (Supplementary Table 5). For GO annotation, all unigenes were divided into 25 groups. Among them, general function prediction only (11691), signal transduction mechanisms (7663), and posttranslational modification, protein turnover, and chaperones (4697) were the largest groups. For KEGG pathway analysis, 32084 unigenes were mapped to metabolic pathways, among which “metabolism” was the largest category.

### Analysis of Differential Gene Expression

Using the DESeq2 method with Q-value (Adjusted P-value) ≤ 0.05, 4590 differentially expressed genes (DEGs) were detected in the two chemotypes of *A. rugosa*. In the ArE and ArP comparison groups, 2,635 and 1,955 genes were upregulated and downregulated, respectively (Supplementary Figure 5). To further explore the possible role of DEGs, we conducted GO and KEGG enrichment experiments. The enrichment degree



of the GO Term is displayed in the form of a histogram. By default, the top 20 GO Term with the smallest Q-value or the selected GO Term (sorted by Q-value, up to 60) are plotted. “ADP binding,” “auxin-activated signaling pathway,” and “beta-glucosidase activity” were the most significantly enriched terms in “molecular function.” The bubble chart of KEGG pathway enrichment results showed that 20 pathways were significantly enriched in “E vs. P.” Among them, “Plant-pathogen interaction” and “Plant hormone signal transduction” enriched the largest number of differentially expressed genes (Supplementary Figure 5). Using the heatmap function in R software to perform hierarchical cluster analysis on the FPKM values of differentially expressed genes, the three replicates of different chemotypes of samples are clustered into one type.

### Identification of Transcription Factors

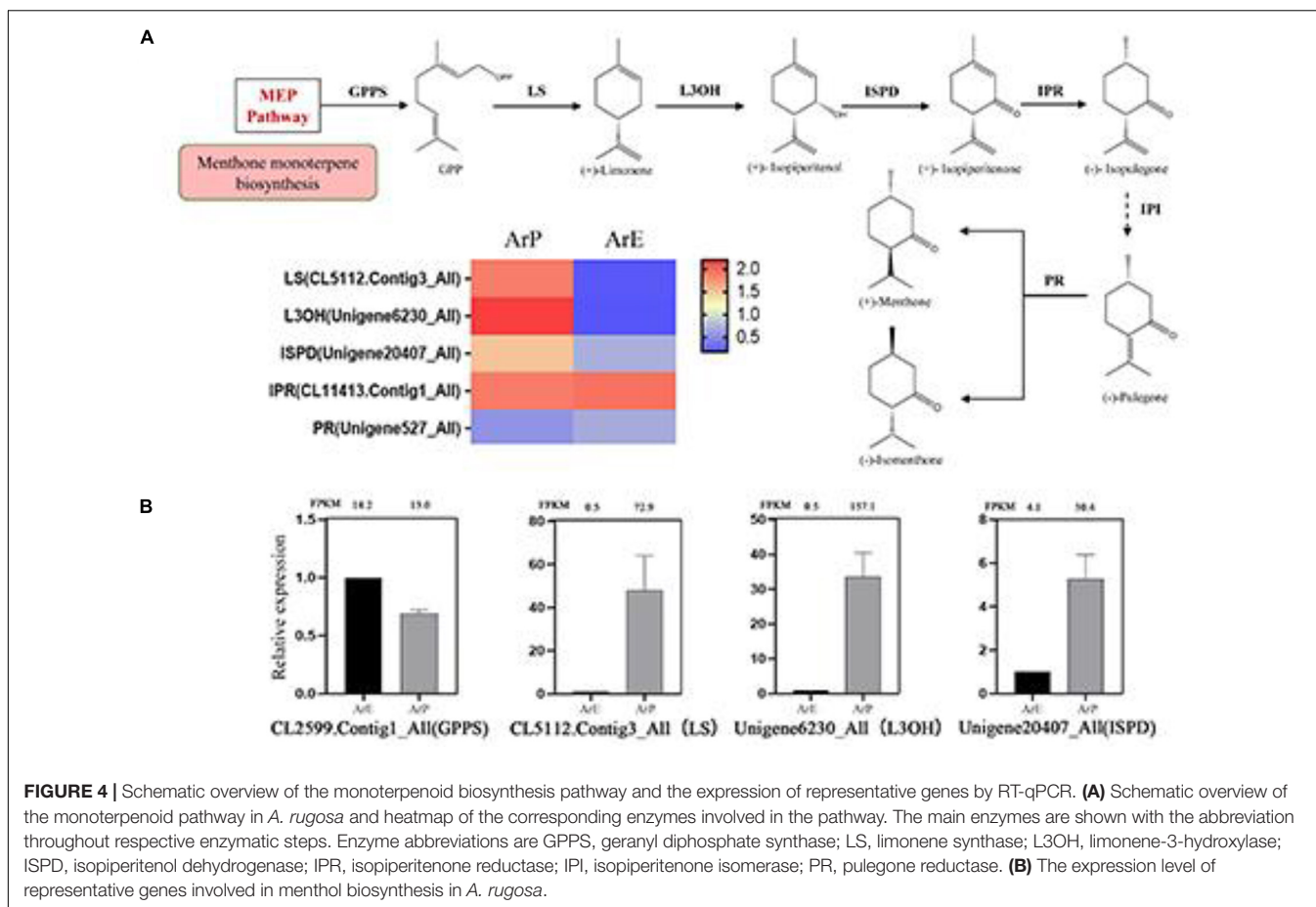
Plant transcription factor prediction uses the Plant TFDB database and compares it with Pfam23.0, using the hmmpfam of the HMMER program, and searches for protein domains to identify genes encoding transcription factors. A total of 2918 unigenes were annotated as transcription factors of different families. Among them, AP2/ARF, WRKY, bHLH, and bZip family

transcription factors are believed to be related to the synthesis of terpenoids, and AP2/ARF and MYB family transcription factors are considered to be related to the growth and development of glandular trichomes.

### Candidate Genes Involved in Monoterpenoids and Phenylpropanoids Biosynthesis

To explore the regulatory mechanisms for the accumulation patterns of different terpenoids in *A. rugosa*, the expression profiles of genes involved in terpenoid biosynthesis were analyzed. In the ArE and ArP comparison groups, 46 DEGs for the biosynthesis of terpenes were identified. Most of the genes exhibited high transcriptome expression levels, encoding key enzymes in the MEP and MVA pathways (KEGG entry ko00900). The expression levels of different unigenes between the two chemotypes were different, but the unigene expression level of the MEP pathway was higher than that of the MVA pathway (Figure 3).

Pulegone is the main constituent of monoterpenoids in the ArP of *A. rugosa*, which is synthesized by a series of



**FIGURE 4** | Schematic overview of the monoterpene biosynthesis pathway and the expression of representative genes by RT-qPCR. **(A)** Schematic overview of the monoterpene pathway in *A. rugosa* and heatmap of the corresponding enzymes involved in the pathway. The main enzymes are shown with the abbreviation throughout respective enzymatic steps. Enzyme abbreviations are GPPS, geranyl diphosphate synthase; LS, limonene synthase; L3OH, limonene-3-hydroxylase; ISPD, isopiperitenol dehydrogenase; IPR, isopiperitenone reductase; IP1, isopiperitenone isomerase; PR, pulegone reductase. **(B)** The expression level of representative genes involved in menthol biosynthesis in *A. rugosa*.

enzymatic reactions in the peltate trichomes on the surfaces of the plant. The biosynthetic pathway and enzyme catalysis mechanism of pulegone and menthone have been extensively studied in *Schizonepeta* (Liu et al., 2018). Using actin as the reference gene, four genes and their orthologous genes catalyze the biosynthesis of (-)-pulegone from (+)-limonene were identified in *A. rugosa* for the first time (Figure 4). Only one unigene encoding one enzyme was selected and identified based on sequence similarity, with the unigene: CL5112.Contig3\_All encoding limonene synthase (LS), Unigene6230\_All encoding limonene-3-hydroxylase (L3OH), Unigene20407\_All encoding isopiperitenone dehydrogenase (ISPD), and Unigene527\_All encoding pulegone reductase (PR). The expression levels of enzymes involved in monoterpene biosynthesis of ArP were higher than those of ArE, except for PR (Figure 4).

Estragole is the representative compound of ArE. The biosynthesis of estragole has also been reported (Yu and Jez, 2008). Based on sequence similarity, there were seven genes identified which catalyze the biosynthesis of phenylpropanoid from L-phenylalanine, with the unigene: Unigene8585\_All encoding phenylalanine ammonia-lyase (PAL), CL9538.Contig1\_All encoding 4-coumarate:CoA ligase (4CL), CL5443.Contig1\_All encoding CYP73A, Unigene10629\_All encoding coumaroyl-CoA reductase (CCR), Unigene3061\_All encoding coumaryl-aldehyde dehydrogenase

(CAD), Unigene17991\_All encoding eugenol synthase (EGS), and CL10810.Contig1\_All encoding chavicol methyl transferase (CVOMT). The expression levels of genes involved in estragole biosynthesis of ArE were higher than those of ArP, except for CCR and CVOMT (Figure 2).

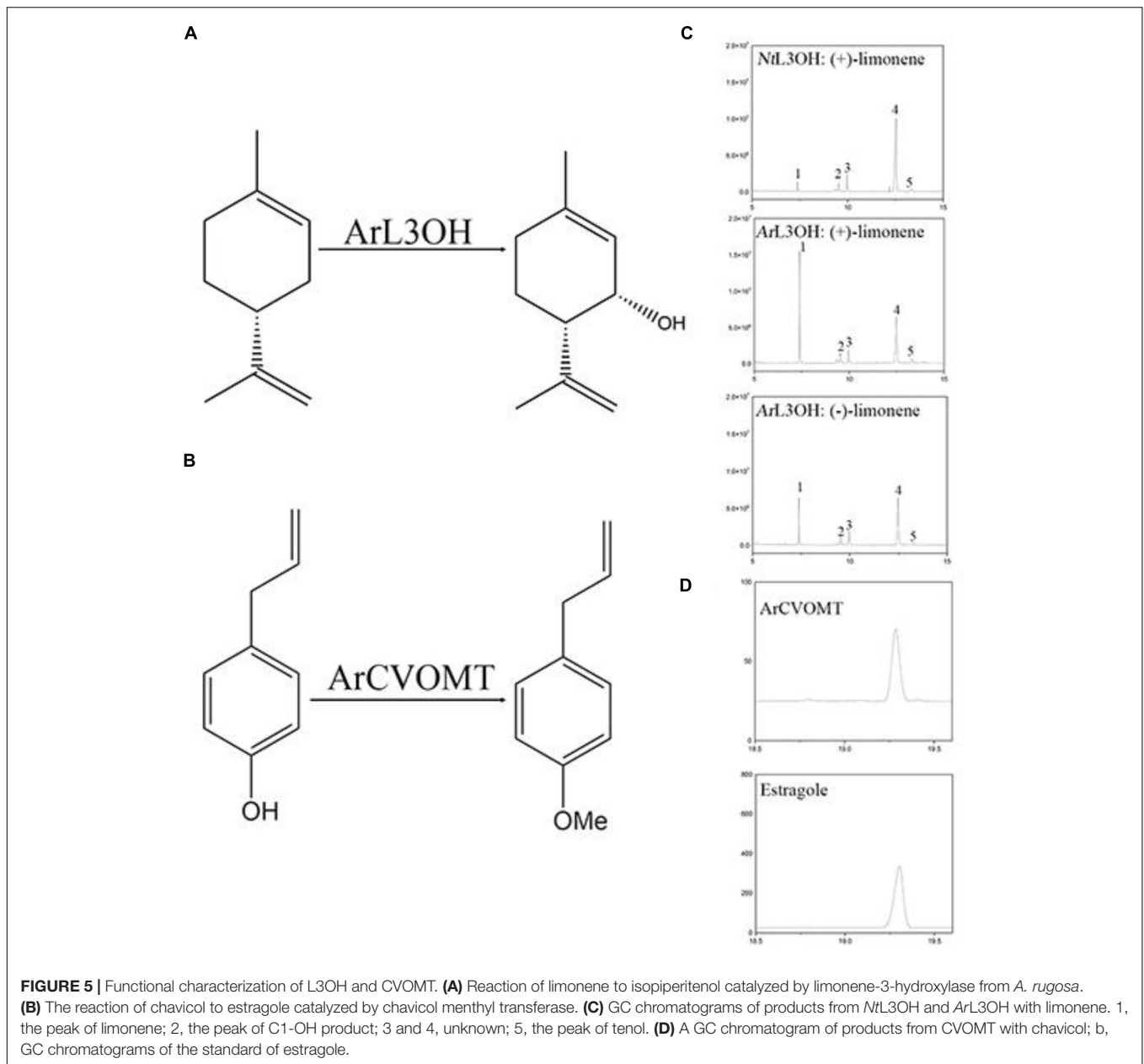
### Quantitative Real-Time PCR Validation of Differentially Expressed Genes

To validate the expression pattern of monoterpenes and phenylpropanoid biosynthesis, qRT-PCR was conducted to examine the expression levels of 14 unigenes in the two chemotypes. The expression levels of these selected genes from qRT-PCR analyses were generally consistent with the deduced from their fragments per kilobase per million mapped data from the RNA-seq. These results confirmed the reliability of the transcriptomic profiling data.

### Expression and Enzyme Assays of Limonene-3-Hydroxylase and Chavicol Methyl Transferase

The sequencing information extracted from the transcriptome data of L3OH was used to amplify the corresponding full-length cDNA (NCBI accession number: OK649356). To examine the function of the candidate gene in menthol biosynthesis





in *A. rugosa*, we expressed *AgL3OH* in *S. cerevisiae*. We conducted enzyme assays *in vitro*, in which the reaction was fed with (-)-limonene and (+)-limonene. The reaction was detected and analyzed using GC and chiral-GC. Since there was no authentic standard available, it can be speculated to be a product of limonene according to our previous report (Liu L. et al., 2021). The results showed that the *ArL3OH* enzyme converted (+)-limonene and (-)-limonene into (-)-isopiperitenol and (+)-isopiperitenol as the major products (Figure 5).

The full-length cDNA (NCBI accession number OK649357) of *CVOMT* was amplified by PCR. The recombinant *CVOMT* protein was transferred into the expression vector pET28a. We expressed *CVOMT* in *E. coli*. We performed enzyme assays

*in vitro*. Chavicol was added to the reaction and detected using GC. Compared with the retention time of the standard, the results showed that the *CVOMT* enzyme converted chavicol into estragole under the action of S-adenosylmethionine (Figure 5).

## DISCUSSION

There are different secondary metabolites among different chemotypes in Lamiaceae, Asteraceae, and Lauraceae, such as *Perilla frutescens* and *Artemisia annua* (Czechowski et al., 2018; Zhou et al., 2021), which generate multiple pharmacological activities, such as *Cinnamomum porrectum* and *Eugenia uniflora* (Figueiredo et al., 2019; Qiu et al., 2019). Therefore, chemotype

formation is the focus and complexity in research on plants, especially in medicinal plants.

At present, *A. rugosa* is classified as a pulegone chemotype, estragole chemotype, and methyl eugenol chemotype (ArM) based on their different oil constituents (Li et al., 2013), which may have different pharmacological effects. Pulegone of *A. rugosa* has antimicrobial activity, whereas the estragole of *A. rugosa* has strong nematocidal activity (Li et al., 2013; Gong et al., 2016). Therefore, studying the formation and mechanism of chemotypes plays an important role in the utilization of *A. rugosa*. However, only LS and PR in the pulegone biosynthesis pathway has been characterized in *A. rugosa* (Maruyama et al., 2002; Liu C. et al., 2021). In this study, we compared the morphological differences, metabolite constituents, and transcriptomic data between the two chemotypes of *A. rugosa*.

The chemotypes of one species are a general phenomenon in plants, leading to morphological differences (Czechowski et al., 2018). Plant height and internode length were the main morphological differences between the high- and low-artemisinin chemotypes in *Artemisia annua* (Czechowski et al., 2018). Different stem types exist in different chemotypes of *Thymus quinquecostatus* (Kim et al., 2020). The morphological differences of various chemotypes in mints mainly focused on inflorescence characteristics. Our study found that plant height and internode length were the main differences between the two chemotypes in *A. rugosa*.

For most plants, the main compounds of different chemotypes in the same species present similar skeletal structures. Terpenoids are the main compounds of different chemotypes in *Origanum vulgare* (Węglarz et al., 2020). There are also a few plants with different chemotypes of the principal components that have different skeletal structures, such as *Ocimum* (Anand et al., 2019). Compared with the metabolites of ArP and ArE by GC-MS, monoterpene is the main constituent in the essential oil of ArP, while phenylpropanoids are the main constituents of the essential oil of ArE, except for limonene, which is the only common component in the essential oil of these two chemotypes. We also detected methyl eugenol in ArE, which is a possible transition chemotype between ArP and ArM. However, there was no downstream product after limonene in ArE. This is consistent with the expression levels of genes involved in menthone biosynthesis of ArE. *LS*, *L3OH*, and *IPD* were highly expressed in ArP, and *GPPS* was highly expressed in both chemotypes of *A. rugosa*. The unigenes related to estragole biosynthesis in ArP were downregulated, except for *CCR* and *CVOMT*. We speculate that *CCR* and *CVOMT* are not only involved in estragole biosynthesis, but also participate in physiological biochemical processes (Wengenmayer et al., 1976; Gang et al., 2002). As expected, and we found that *L3OH* could catalyze both (+)-limonene and (-)-limonene to (+)-isopiperitenol and (-)-isopiperitenol. Apart from the main product, the byproducts were discovered as described (Liu L. et al., 2021). The functional characterization of *CVOMT* was identified. Under the catalytic action of *CVOMT* and S-adenosylmethionine, chavicol was transferred to estragole. Therefore, in addition to improving

our understanding of different chemotypes in the Lamiaceae, our results highlight targeted genes for volatile oil biosynthesis in *A. rugosa*, which could benefit the genetic improvement of *A. rugosa* and other Lamiaceae species. However, one of the key points in chemotype formation is the regulation of specialized gene expression in specialized chemotypes, which requires further research.

## DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: <https://www.ncbi.nlm.nih.gov/>, PRJNA774149; <https://www.ncbi.nlm.nih.gov/>, OK649356; and <https://www.ncbi.nlm.nih.gov/>, OK649357.

## AUTHOR CONTRIBUTIONS

JD performed functional characterization, content integration, and manuscript composition. GL performed data analysis, manuscript composition, qPCR experiments, and manuscript composition. CL assisted in the qPCR experiments. PZ and SD performed the morphological experiments. MS and ZJ conducted oil collection. YS performed GC-MS detection and analysis. QW and CL were responsible for the overall concept, design of the experiments, data integration, and manuscript revision. All authors contributed to the article and approved the submitted version.

## FUNDING

This research was supported by the National Natural Science Foundation of China (grant numbers 81973435 and 81473313); the National Natural Science Foundation for Young Scientists of China (grant number 81903756), and research on ecological planting and quality assurance of Jiangsu Dao-di herbs (2021).

## ACKNOWLEDGMENTS

We thank Qiyu Gao, Fuwei Li, Wukun Liu (Nanjing University of Chinese Medicine), and Mengjiao Yin for technical support. We also thank the Academician Xiaoya Chen and all the staff at the Shanghai Institute of Plant Physiology and Ecology (Chinese Academy of Sciences) for providing yeast cells and experimental guidance.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2022.850130/full#supplementary-material>

## REFERENCES

- Anand, A., Jayaramaiah, R. H., Beedkar, S. D., Dholakia, B. B., Lavhale, S. G., Punekar, S. A., et al. (2019). Terpene profiling, transcriptome analysis and characterization of cis- $\beta$ -terpineol synthase from *Ocimum*. *Physiol. Mol. Biol. Plants* 25, 47–57. doi: 10.1007/s12298-018-0612-6
- Cao, P., Xie, P., Wang, X., Wang, J., Wei, J., and Kang, W. Y. (2017). Chemical constituents and coagulation activity of *Agastache rugosa*. *BMC Compl. Altern. Med.* 17:93. doi: 10.1186/s12906-017-1592-8
- Czechowski, T., Larson, T. R., Catania, T. M., Harvey, D., Wei, C., Essome, M., et al. (2018). Detailed phytochemical analysis of high- and low artemisinin-producing chemotypes of *Artemisia annua*. *Front. Plant Sci.* 9:641. doi: 10.3389/fpls.2018.00641
- Figueredo, P. L. B., Pinto, L. C., da Costa, J. S., da Silva, A. R. C., Mourão, R. H. V., Montenegro, R. C., et al. (2019). Composition, antioxidant capacity and cytotoxic activity of *Eugenia uniflora* L. chemotype-oils from the Amazon. *J. Ethnopharmacol.* 232, 30–38. doi: 10.1016/j.jep.2018.12.011
- Gang, D. R., Lavid, N., Zubieta, C., Chen, F., Beuerle, T., Lewinsohn, E., et al. (2002). Characterization of phenylpropane O-Methyltransferases from sweet basil: Facile of substrate specificity and convergent evolution within a plant O-Methyltransferase family. *Plant Cell* 14, 505–519. doi: 10.1105/tpc.010327
- Gong, H., He, L., Li, S., Zhang, C., and Ashraf, M. (2016). Antimicrobial, antibiofilm and antitumor activities of essential oil of *Agastache rugosa* from Xinjiang, China. *Saoudi J. Biol. Sci.* 23, 524–530. doi: 10.1016/j.sjbs.2016.02.020
- Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., and Bowden, J. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* 8, 1494–1512. doi: 10.1038/nprot.2013.084
- Kim, M., Moon, J. C., Kim, S., and Sowndhararajan, K. (2020). Morphological, Chemical, and Genetic Characteristics of Korean Native Thyme Bak-Ri-Hyang (*Thymus quinquecostatus* Celak.). *Antibiotics* 9:289. doi: 10.3390/antibiotics9060289
- Li, H. Q., Liu, Q. Z., Liu, Z. L., Du, S. S., and Deng, Z. W. (2013). Chemical Composition and Nematicidal Activity of Essential Oil of *Agastache rugosa* against *incognita*. *Molecules* 18, 4170–4180. doi: 10.3390/molecules18044170
- Liu, C., Gao, Q., Shang, Z., Liu, J., Zhou, S., Dang, J., et al. (2021). Functional Characterization and Structural Insights Into Stereoselectivity of Pulegone Reductase in Menthol Biosynthesis. *Front. Plant Sci.* 30:780970. doi: 10.3389/fpls.2021.780970
- Liu, C., Srividya, N., Parrish, A. N., Yue, W., Shan, M., Wu, Q., et al. (2018). Morphology of glandular trichomes of Japanese catnip (*Schizonepeta tenuifolia* Briquet) and developmental dynamics of their secretory activity. *Phytochemistry* 150, 23–30. doi: 10.1016/j.phytochem.2018.02.018
- Liu, L., Yin, M., Lin, G., Wang, Q., Zhou, P., Dai, S., et al. (2021). Integrating RNA-seq with functional expression to analyze the regulation and characterization of genes involved in monoterpene biosynthesis in *Nepeta tenuifolia* Briq. *Plant Physiol. Biochem.* 167, 31–41. doi: 10.1016/j.plaphy.2021.07.026
- Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15:550. doi: 10.1186/s13059-014-0550-8
- Mahmood, K., Orabi, J., Kristensen, P. S., Sarup, P., Jørgensen, L. N., and Jahoor, A. (2020). De novo transcriptome assembly, functional annotation, and expression profiling of rye (*Secale cereale* L.) hybrids inoculated with ergot (*Claviceps purpurea*). *Sci. Rep.* 10:13475. doi: 10.1038/s41598-020-7406-2
- Maruyama, T., Saeki, D., Ito, M., and Honda, G. (2002). Molecular cloning, functional expression, and characterization of d-limonene synthase from *Agastache rugosa*. *Biol. Pharm. Bull.* 25, 661–665. doi: 10.1248/bpb.25.661
- Mo, J., and Ma, L. (2011). Volatile oil of *Herba Agastache* in various growth periods and different parts by GC-MS. *Chin. J. Pharm.* 42, 268–270. doi: 10.1080/10412905.1992.9698139
- Qiu, F., Wang, X., Zheng, Y., Wang, H., and Liu, X. (2019). Full-length transcriptome sequencing and different chemotype expression Analysis of genes related to monoterpene biosynthesis in *Cinnamomum porrectum*. *Int. J. Mol. Sci.* 20:24. doi: 10.3390/ijms20246230
- Šarić-Kundalić, B., Fialová, S., Dobeš, C., Ōlzant, S., Tekel'ová, D., Granèai, D, et al. (2009). Multivariate numerical taxonomy of *Mentha* species, hybrids, varieties, and cultivars. *Sci. Pharm.* 77, 851–876. doi: 10.3797/scipharm.0905-10
- Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., and Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31, 3210–3212. doi: 10.1093/bioinformatics/btv351
- Skakovskii, E. D., Kiselev, W. P., Tychinskaya, L. Y., Schutova, A. G., Gonsharova, L. W., Spiridowish, E. W., et al. (2010). Characterization of the essential oil of *Agastache rugosa* by NMR spectroscopy. *J. Appl. Spectrosc.* 77, 329–334. doi: 10.1007/s10812-010-9335-3
- Węglarz, Z., Kosakowska, O., Przybył, J. L., Pióro-Jabrucka, E., and Bączek, K. (2020). The quality of Greek oregano (*O. vulgare* L. subsp. *hirtum* (Link) Ietswaart) and common oregano (*O. vulgare* L. subsp. *vulgare*) Cultivated in the temperate climate of Central Europe. *Foods* 9:1671. doi: 10.3390/foods9111671
- Wengenmayer, H., Ebel, J., and Grisebach, H. (1976). Enzymic synthesis of lignin precursors. Purification and properties of a cinnamoyl-CoA: NADPH reductase from cell suspension cultures of soybean (*Glycinemax*). *Eur. J. Biochem.* 65, 529–536. doi: 10.1111/j.1432-1033.1976.tb10370.x
- Yu, O., and Jez, J. M. (2008). Nature's assembly line: Biosynthesis of simple phenylpropanoids and polyketides. *Plant J.* 54, 750–762. doi: 10.1111/j.1365-313X.2008.03436.x
- Zhou, P., Yin, M., Dai, S., Bao, K., Song, C., Liu, C., et al. (2021). Multi-omics analysis of the bioactive constituents biosynthesis of glandular trichome in *Perilla frutescens*. *BMC Plant Biol.* 21:277. doi: 10.1186/s12870-021-03069-4
- Zhou, T., Zhou, R., Yu, Y., Xiao, Y., Li, D., Xiao, B., et al. (2016). Cloning and characterization of a flavonoid 3'-hydroxylase gene from tea plant (*Camellia sinensis*). *Int. J. Mol. Sci.* 17:261. doi: 10.3390/ijms17020261

**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 Dang, Lin, Liu, Zhou, Shao, Dai, Sang, Jiang, Liu and Wu. 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.