Widely targeted metabolomics analysis reveals differences in volatile metabolites among four *Angelica* species

Jiaojiao Ji [1, 2, †], Lanlan Zang [1, 3, †], Tingting Lu [1, 5, 6], Cheng Li [1], Xiaoxu Han [1], Soo-Rang Lee [7], Li Wang [1, 4, *]

[1] Shenzhen Branch, Guangdong Laboratory for Lingnan Modern Agriculture, Genome Analysis Laboratory of the Ministry of Agriculture, Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen 518120, China

[2] College of Life Science and Technology, Huazhong Agricultural University, Wuhan 430070, China

[3] Zhengzhou Research Base, State Key Laboratory of Cotton Biology, School of Agricultural Sciences, Zhengzhou University, Zhengzhou 450001, China

[4] Kunpeng Institute of Modern Agriculture at Foshan, Foshan 528200, China

[5] State Key Laboratory of Crop Stress Adaptation and Improvement, School of Life Sciences, Henan University, Kaifeng 475004, China

[6] Shenzhen Research Institute of Henan university, Shenzhen 518000, China

[7] Department of Biology Education, College of Education, Chosun University, Gwangju 61452, South Korea

[*] Corresponding authors: wangli03@caas.cn (Li Wang)

[†] These authors contributed equally to this work.

Present address: NO 7, Peangfei Road, Dapeng District, Shenzhen 518120, China

Abstract

Angelica L. has attracted global interest for its traditional medicinal uses and commercial values. However, few studies have focused on the metabolomic differences among the *Angelica* species. In this study, employing the widely targeted metabolomics based on gas chromatography-tandem mass spectrometry, the metabolomes of four *Angelica* species were analyzed (*Angelica sinensis* (Oliv.) Diels (*A. sinensis*), *Angelica biserrata* (R.H.Shan & Yuan) C.Q.Yuan & R.H.Shan (*A. biserrata*), *Angelica dahurica* (Hoffm.) Benth. & Hook.f. ex Franch. & Sav. (*A. dahurica*), *Angelica keiskei* Koidz. (*A. keiskei*)). A total of 698 volatile metabolites were identified and classified into fifteen different categories. The metabolomic analysis indicated that 7-hydroxycoumarin and Z-ligustilide were accumulated at significantly higher levels in *A. sinensis*, whereas the opposite pattern was observed for bornyl acetate. In addition, a high correspondence between the dendrogram of metabolite contents and phylogenetic positions was detected in the four species. This study provides a biochemical map for the exploitation, application and development of the *Angelica* species as medicinal plants or health-related dietary supplements.

Keywords

Angelica; volatile metabolites; Chinese traditional medicine; phylogeny

1. Introduction

Angelica L., a genus in the family Apiaceae, is comprised of 90 species of herbs that are widespread in north-temperate regions, especially Eurasia [1,2]. Many plants in the genus have long been used in traditional Chinese medicine (TCM) [3], in particular, the dried roots of *Angelica* have been widely used for nourishing blood, regulating menstruation, and analgesic [1,4]. Various herbal preparations containing *Angelica* species are available over the counter, not only in China, but also in Europe and American countries [5,6]. Besides its medicinal value, *Angelica* is also highly appreciated in various industrial applications such as the dietary supplements, perfumery, and cosmetics [1,7,8]. A previous study demonstrated that the pharmacological activity of aromatic and medicinal plants is attributed to its effective volatile components [9]. Plants in *Angelica* are extremely rich in secondary metabolites, including coumarins, flavonoids, terpenoids, as well as volatiles oils (VOs) [1,3]. Modern medical research has revealed that the VOs composition is mainly responsible for the medicinal properties of the genus *Angelica* [10]. VOs are complex mixture of low molecular weight volatile compounds that are isolated from the raw plant material by distillation [11], which have been reported to treat serious health diseases, involving gynecological diseases, fever, and arthritis [1,12]. There are a couple of good examples showing the proven effects of VOs in *Angelica* species. Phthalides of *Angelica sinensis* (Oliv.) Diels (*A. sinensis*) are one of the highly effective VOs to analgesic and sedative activities [6,13]. *Angelica biserrata* (R.H.Shan & Yuan) C.Q.Yuan & R.H.Shan (*A. biserrata*) also contains active ingredients such as oxygenates, terpenoids, ketones and esters with analgesic and anti-inflammatory effects [14]. However, most of current studies only focused on several targeted compounds in a single *Angelica* species. There have been no comprehensive and comparative studies examining the volatile metabolites of

multiple *Angelica* species. It has posed a major obstacle to the application and exploitation of the medicinal plants in *Angelica* species.

With the development of metabolomics, high-throughput and high-resolution methods such as headspace solid phase micro-extraction gas chromatography-mass spectrometry (HS-SPME-GC-MS) have been widely used to identify metabolite profiles and detect differences in the biochemical compositions of aromatic and medicinal plants [10,15,16]. The four species *A. biserrata*, *Angelica dahurica* (Hoffm.) Benth. & Hook.f. ex Franch. & Sav. (*A. dahurica*), *Angelica keiskei* Koidz. (*A. keiskei*) and *A. sinensis* are the representative medicinal plants in *Angelica*, and it is noteworthy that roots of *A. sinensis* are one of the most widely prescribed medicine in China owing to its rich VOs [6]. In this study, volatile metabolites of four *Angelica* species were identified and quantified using widely targeted metabolomics. The aim was to reveal the differed accumulation of medicinally important metabolites among the four species. This study provides useful information for the chemical composition of *Angelica* plants and may help the identification of the biologically active substances responsible for the pharmacological activity of *Angelica* plants.

2. Materials and methods

2.1. Plant samples

Four species in genus *Angelica*, including *A. sinensis*, *A. dahurica*, *A. biserrata*, and *A. keiskei*, were analyzed in this study. The *A. sinensis* plants were collected from Minxian County, Gansu Province, China. The *A. dahurica*, *A. biserrata*, and *A. keiskei* plants were collected from Shenzhen City, Guangdong Province, China. The specimens of the four species were deposited at Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences and identified by Prof. Li Wang (*A. sinensis* (202011002), *A. dahurica* (202008001), *A. biserrata* (202106003) and *A. keiskei* (202109004))**.**

Roots of each species were sampled with three biological replicates. The collected roots were washed, naturally dried, frozen in liquid nitrogen, and then stored at -80℃ for further analysis.

2.2. Solid phase microextraction (SPEM) extraction

The samples were ground into powder in liquid nitrogen. Powdered samples (1 g) were weighed and transferred immediately to a 20 mL head-space vial (Agilent, Palo Alto, CA, USA), containing NaCl saturated solution to inhibit potential enzyme reactions. The headspace vials were sealed using crimp-top caps. As for SPME analysis, each vial was placed in 60℃ for 5 min, and then a 120 μm DVB/CWR/PDMS fiber (Agilent, Palo Alto, CA, USA) was exposed to the headspace of the sample for 15 min at 100℃. DVB/CWR/PDMS are three-phase fiber heads, which were confirmed to be able to extract more volatile metabolites than other fiber headers [17]. The quality control (mix) sample was prepared by mixing equal volumes of samples into a single tube, and during the instrumental analysis, a quality control sample was inserted into each of the 10 test samples to check the repeatability of the analysis process (Figure S1 and Table S1).

2.3. GC-MS analysis

After the extraction procedure, the fiber was transferred to the injection port of the GC-MS system (Model 8890; Agilent, Palo Alto, CA, USA). The SPME fiber was desorbed and maintained in the injection port at 250℃ for 5 min in the split-less mode. The identification and quantification of volatile metabolites was carried out using an Agilent Model 8890 GC and a 7000 D mass spectrometer (Agilent, Palo Alto, CA, USA), equipped with a 30 m × 0.25 mm × 0.25 μm DB-5MS (5% phenyl-polymethylsiloxane) capillary column. Helium was used as the carrier gas at a linear velocity of 1.2 mL/min. The injector temperature was kept at 250℃ and the detector at 280℃. The oven temperature was programmed as followings: 40℃ (3.5 min),

increasing at 10□ min^{-1} to 100□, 7□ min^{-1} to 180□, 25□ min^{-1} to 280□ and hold for 5 min. Mass spectra was recorded in electron impact ionization mode at 70 eV. The quadrupole mass detector, ion source and transfer line temperatures were set, respectively, at 150, 230 and 280□. For the identification and quantification of analytes, the MS was selected ion monitoring mode.

2.4. Qualitative and quantitative analysis

After the mass spectrometry analysis, all raw data were analyzed with the software Qualitative Analysis Workflows B.08.00 (Agilent, Palo Alto, CA, USA). The qualitative analysis of primary and secondary mass spectrometry data was annotated based on the self-built database MWDB (Metware Biotechnology Co., Ltd. Wuhan, China) and the publicly available metabolite databases.

2.5 Statistical analysis

After the metabolite data was transformed with Hellinger transformation, principal component analysis (PCA) was performed using the function rda in the R package vegan v 2.6-2 [18]. In addition, the data set that was log_{10} transform and mean centering were imported into the R package MetaboAnalystR v5.0 [19] to conduct orthogonal partial least squares-discriminant analysis (OPLS-DA) and extract the variable important in projection (VIP) value from the analysis results to evaluate the variance importance of compounds. The values of R2X, R2Y, and Q2 for OPLS-DA models were showed in Figure S2. In order to avoid overfitting, a permutation test (100 permutations) was performed. Based on Bray-Curtis's dissimilarity distances of the composition and abundance of all volatile metabolites after ln transform, which were calculated using the function vegdist built in vegan, hierarchical clustering was visualized with the R package factoextra v.1.0.7 [20].

The chloroplast sequence alignments of *A. sinensis*, *A. dahurica*, *A. biserrata*, and *A. keiskei* were generated using MAFFT v7.475 [21]. Phylogenetic trees were

constructed by maximum likelihood using IQ-TREE v 2.1.2 [22] with *Hydrocotyle sibthorpioides* Lam. as an outgroup.

All identified metabolites were annotated with KEGG database (http://www.kegg.jp/kegg/compound/) and further subjected to KEGG enrichment analyses with the R package clusterProfiler v. 4.4.4 [23].

3. Results

3.1. Metabolomics profiling of four Angelica species

Widely targeted metabolomics offered a promising way for the chemical screening of volatile metabolites and allowed the characterization of new metabolites in *Angelica* [10]. In this study, to get insight into differences of volatile metabolites among four *Angelica* species, the root metabolomics data were generated. A total of 698 non-redundant volatile metabolites were qualified and quantified based on GC-MS (Table S1). Among them, 616, 536, 576, and 545 volatile metabolites in *A. sinensis*, *A. dahurica*, *A. biserrata*, *A. keiskei*, respectively. Three hundred and ninety-one metabolites were commonly detected in the roots of all four species (Figure 1a).

PCA of the metabolome data, transformed with Hellinger transformation method, the samples were divided into four distinct groups corresponding to the four species. And the three biological replicates were clustered together, suggesting that the data are reproducible and reliable. Based on the PCA plot, where PC1 and PC2 explained 48.09% and 30.62% of the total variance, respectively. Of the four clusters, PC1 mainly differentiated *A. sinensis* from the other *Angelica* species, while PC2 primarily segregated *A. dahurica* from the other *Angelica* species (Figure 1b).

The abundances of volatile metabolites were transformed by *Z*-score and then subjected to hierarchical clustering analysis (Figure 1c). The results showed that showed significant differences among the four species. Of these, the abundance of volatile metabolites in *A. keiskei* was the highest.

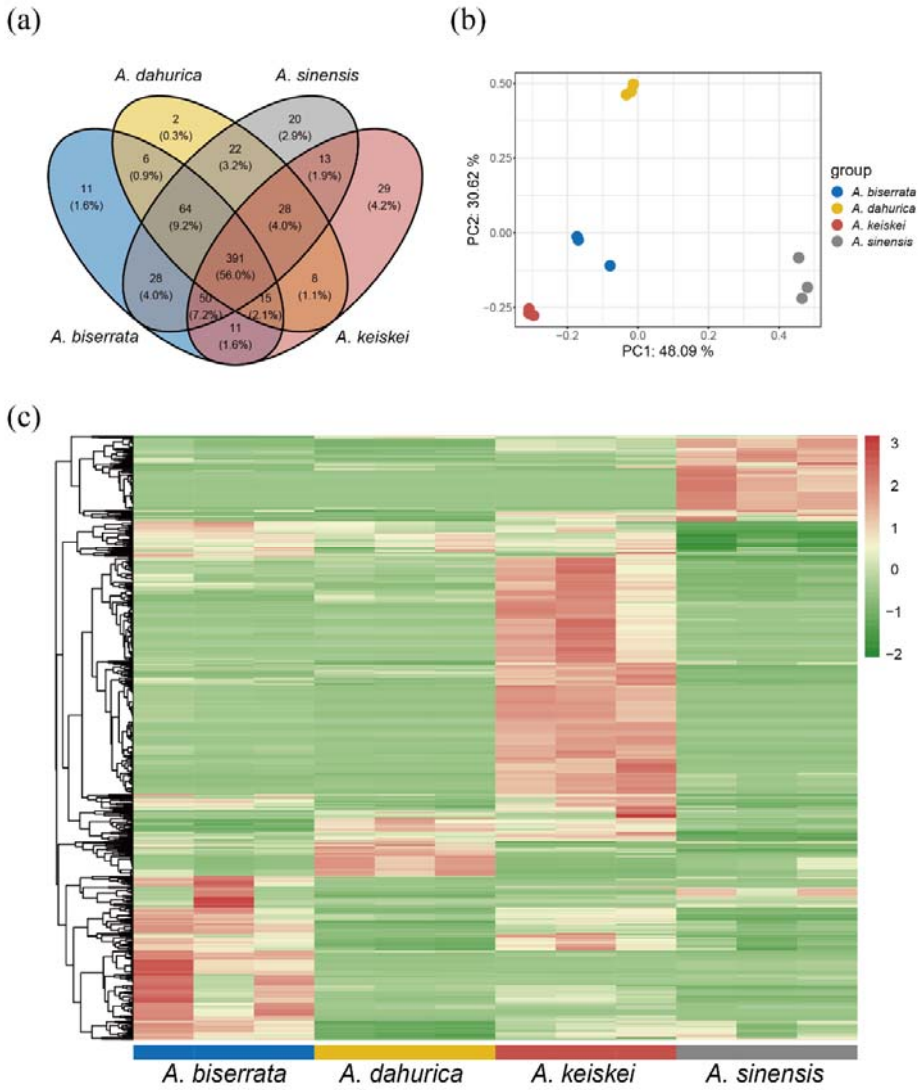

Figure 1. An overview of volatile metabolites among four *Angelica* species. (a) Venn diagram showing the number of common and specific metabolites in the four species. (b) PCA of volatile metabolites for the four species with three biological replicates. (c) Heatmap clustering of volatile metabolites identified from the four species. Volatile metabolite abundance was *Z*-score transformed. The color-coded scale grading from green to red corresponds to the content of volatile metabolites shifting from low to high.

3.2. Identification of differential metabolites in four Angelica species

To explore the metabolite composition of the four species, 698 volatile metabolites were classified into 15 different categories, including terpenoids, ester, heterocyclic, aromatics and 11 others (Figure 2). The terpenoids took up the highest proportion of all measured volatile metabolites in the four *Angelica* species, followed by heterocyclic compounds, eater, and aromatics. Notably, *A. sinensis* contained a relatively lower proportion (43.34%) of terpenoids than other *Angelica* species, but exhibited a more balanced metabolite composition in volatile metabolites. In contrast, the amount of terpenoids accounted for more than half of the total volatile metabolites in *A. dahurica*, *A. biserrata*, and *A. keiskei*, especially in *A. dahurica*, its proportion reached up to 75.64%.

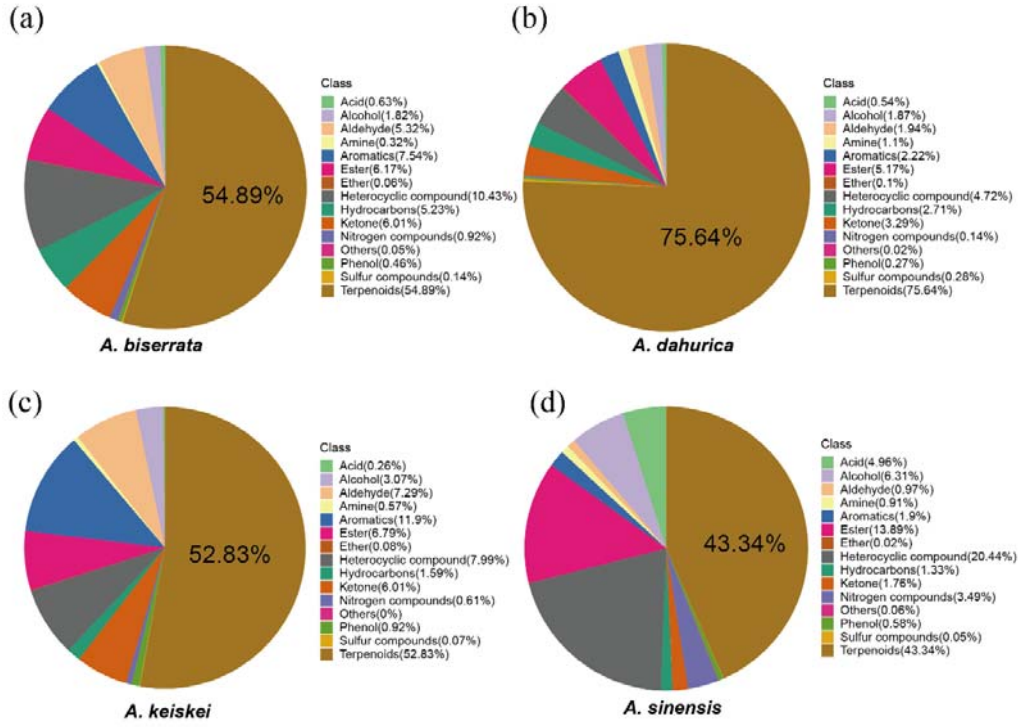

Figure 2. Classification and proportion of volatile metabolites detected in the four

Angelica species. (a) *A. biserrata*, (b) *A. dahurica*, (c) *A. keiskei*, (d) *A. sinensis*.

In addition, the relative abundance of each metabolite category in the four species were compared with Kruskal-Wallis test. The *p* values corrected by the bonferroni method showed significant difference in eight categories among the four species, including alcohol, aldehyde, aromatics, ester, heterocyclic compounds, hydrocarbons, ketone and terpenoids (Figure 3, Figure S3; *p*-values were shown in the Table S2). Terpenoids was the most different metabolites among the four species. In each of the eight categories, *A. sinensis* and *A. keiskei* were significantly different from at least one species. These results indicated that the difference between the presented *Angelica* species lies in the representation of individual components in the metabolomic profiles of the samples, while the qualitative composition is approximately the same.

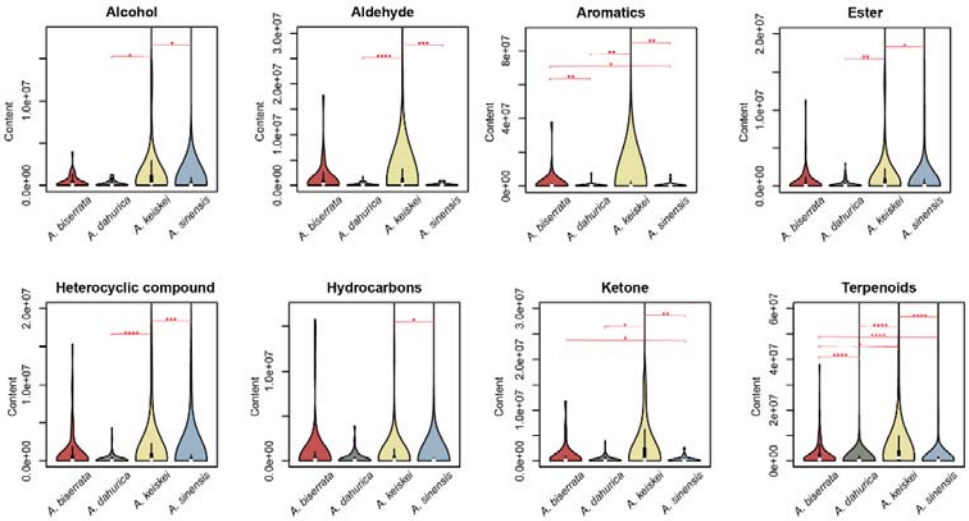

Figure 3. Comparison for the relative abundance of seven categories (alcohol, aromatics, aldehyde, ester, heterocyclic compounds, ketone and terpenoids) with significant differences in the four *Angelica* species.

Here, this study performed hierarchical clustering analysis based on Bray-Curtis's dissimilarity distances of the composition and abundance of volatile metabolites in the four *Angelica* species. The dendrogram (Figure 4a) showed high correspondence with the phylogenetic tree (Figure 4b) based on chloroplast sequences, suggesting a correlation relationship between the volatile metabolites and the phylogenetic relationships.

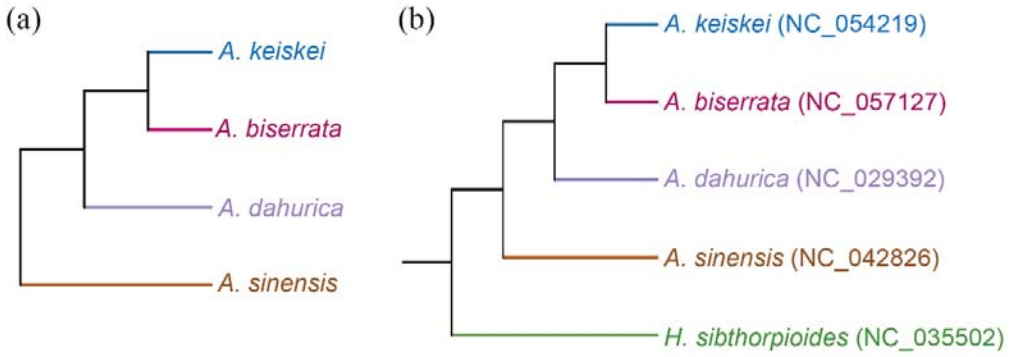

Figure 4. Hierarchical clustering based on the similarity of volatile metabolites (a) and phylogenetic tree of the four *Angelica* species and *H. sibthorpioides* (b). The chloroplast sequences above were available in GenBank of NCBI.

3.3. Differential metabolites between A. sinensis and the three other Angelica species

To further identify the metabolites responsible for differences among the four *Angelica* species, significantly different accumulated metabolites between groups were screened by |Log2FC|≥1 and VIP≥1. *A. sinensis* belongs to *Sinodielsia* clade in *Angelica* genus, that was phylogenetically distant from core *Angelica* group, including *A. biserrata*, *A. dahurica*, *A. keiskei* [2,24]. Moreover, PC1 mainly differentiated *A. sinensis* from the other three species (Figure 1b), therefore, *A. sinensis* was used to comparison to other three species. Interestingly, there were fewer up-regulated metabolites in *A. sinensis* when compared with the other species. And

no significantly enriched pathway was detected in the KEGG enrichment results of these differential metabolites, which could be a bias caused by the small dataset. Compared with *A. biserrata*, 446 significantly differential metabolites (123 up-regulated and 323 down-regulated) were screened in *A. sinensis* (Figure 5a), and the top 3 enrichment pathways of these substance were metabolic pathways (23 metabolites with p = 0.19), tyrosine metabolism (3 metabolites with p = 0.22) and limonene and pinene degradation (5 metabolites with p = 0.24) (Figure 5d). Compared with *A. dahurica*, 429 significantly differential metabolites (169 up-regulated and 260 down-regulated) were detected in *A. sinensis* (Figure 5b), and the top 3 enrichment pathways of these metabolites were tyrosine metabolism (3 metabolites with p = 0.21), limonene and pinene degradation (5 metabolites with p = 0.22) and metabolic pathways (22 metabolites with p = 0.26) (Figure 5e). When compared with *A. keiskei*, 502 significantly differential metabolites (105 up-regulated and 397 down-regulated) were identified in *A. sinensis* (Figure 5c), which were the most abundant compared with the other two group, and the top 3 enrichment pathways of these metabolites were metabolic pathways (25 metabolites with p = 0.09), biosynthesis of various plant secondary metabolites (5 metabolites with p = 0.10), and tyrosine metabolism (3 metabolites with p = 0.26) (Figure 5f).

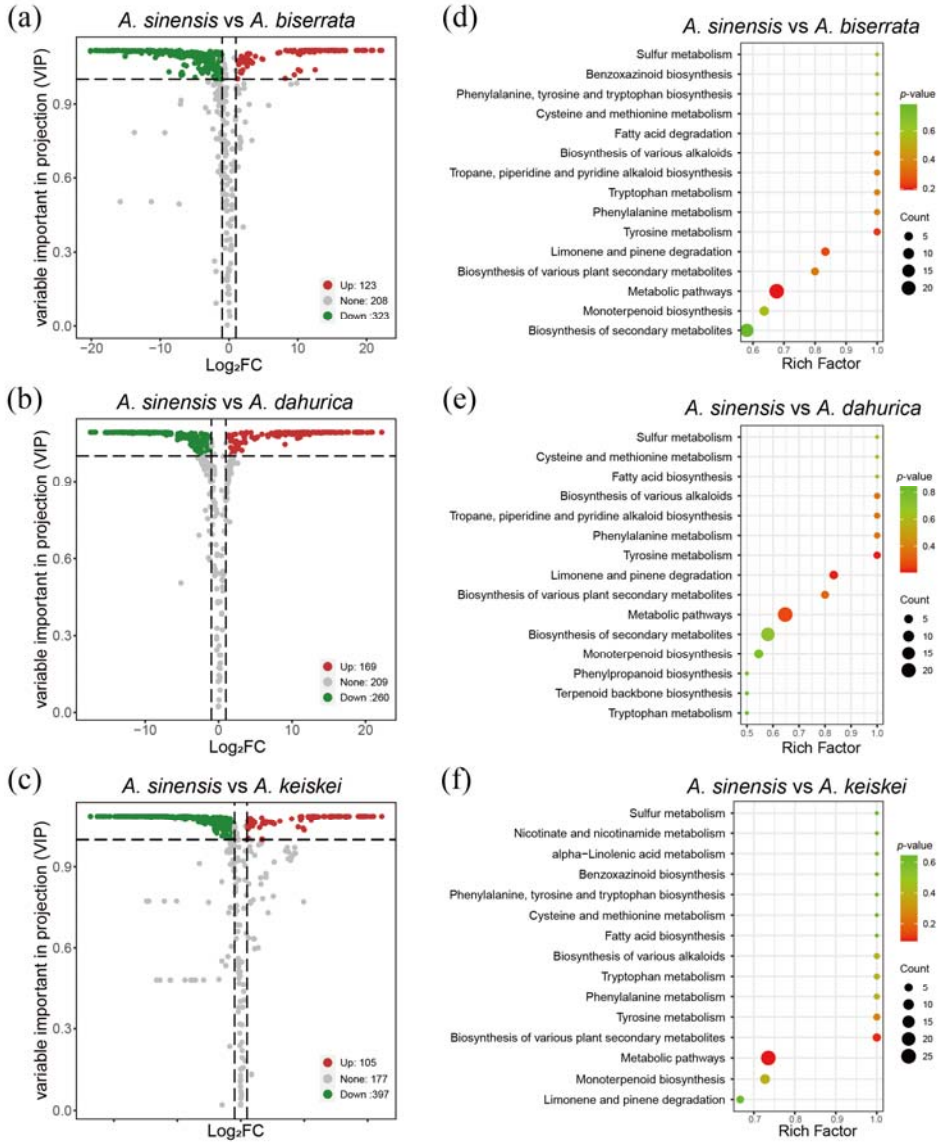

Figure 5. The overall distribution and KEGG enrichment analysis of differential metabolites between *A. sinensis* and the three other *Angelica* species. (a-c) Volcano plots for differential metabolites between *A. sinensis* and the three other *Angelica* species. (a) *A. sinensis* vs *A. biserrata*. (b) *A. sinensis* vs *A. dahurica*. (c) *A. sinensis* vs *A. keiskei*. Colors of metabolites indicated significant differences (red, upregulated; green, downregulated). (d-f) KEGG pathway enrichment analysis of differential

metabolites for *A. sinensis* vs *A. biserrata* (d), *A. sinensis* vs *A. dahurica* (e) and *A. sinensis* vs *A. keiskei* (f). Color of the bubbles represented statistical significance of the enriched terms, and the size of the bubbles represented number of differentially enriched metabolites. The pathway of "Biosynthesis of various plant secondary metabolites" including: crocin biosynthesis, cannabidiol biosynthesis, mugineic acid biosynthesis, pentagalloylglucose biosynthesis, benzoxazinoid biosynthesis, gramine biosynthesis, coumarin biosynthesis, furanocoumarin biosynthesis, hordatine biosynthesis, podophyllotoxin biosynthesis.

In order to delve into the details of the volatile metabolite difference between *A. sinensis* and the other three species, the most significantly twenty metabolites (the top 10 for up-regulation and down-regulation, respectively) were selected (Figure 6). It was discovered that hippuric acid, 7-hydroxycoumarin and 7-ethoxycoumarin were more enriched in *A. sinensis* than the three other *Angelica* species. In addition, the abundance of 3-butylisobenzofuran-1(3H)-one in *A. sinensis* was also substantially higher than that in *A. dahurica and A. keiskei* ($\log_2$FC > 19). Meanwhile, the metabolites γ-terpinene and bornyl acetate in *A. dahurica, A. keiskei* and *A. biserrata* were in high abundance, but the metabolites were lower in *A. sinensis*.

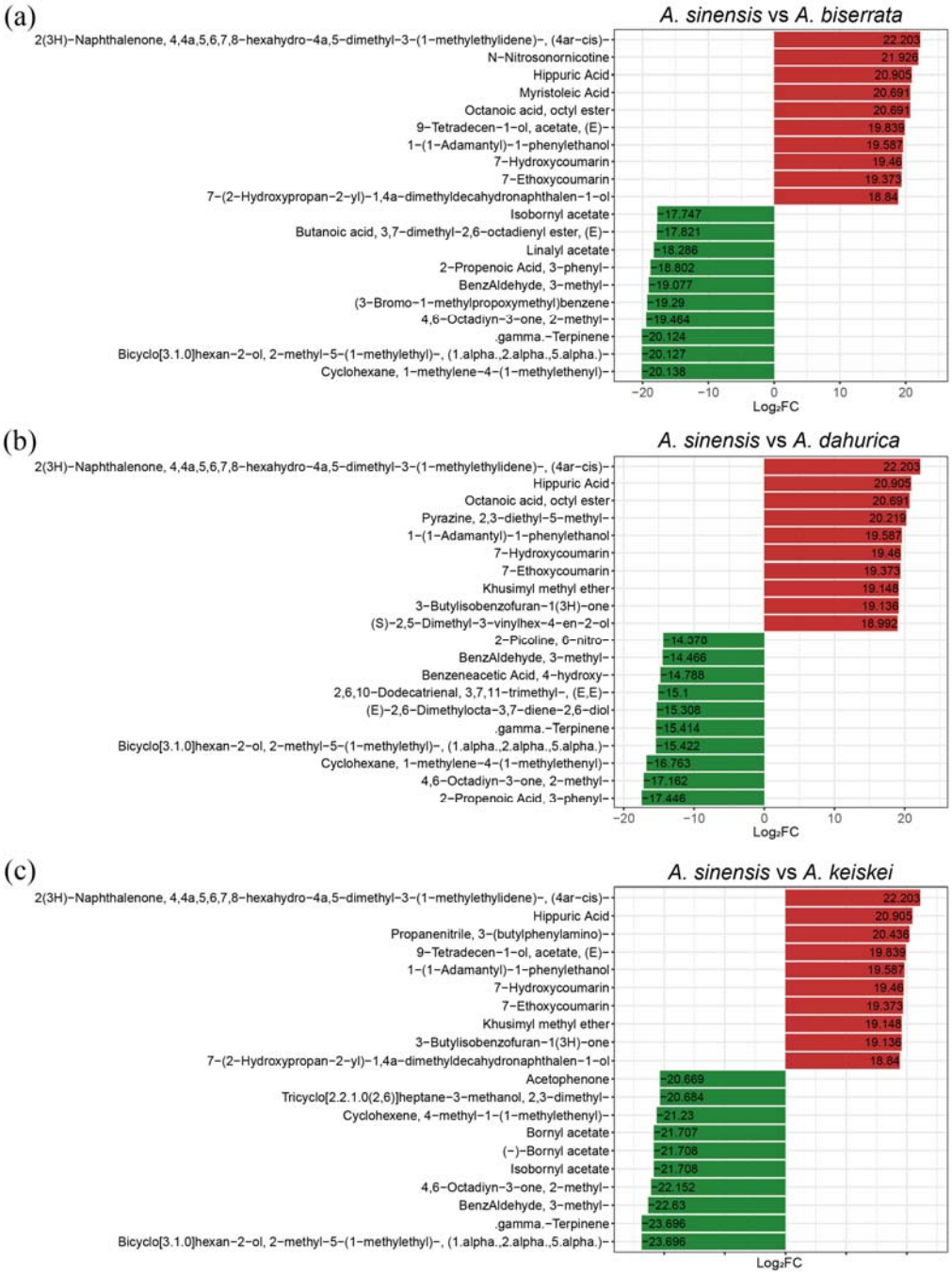

Figure 6. The top 20 metabolites of significantly differential volatiles between *A. sinensis* and three other *Angelica* species. Red indicates the more abundant

metabolites in *A. sinensis* compared to *A. biserrata* (a), *A. dahurica* (b), *A. keiskei* (c). Green indicates the lower levels of metabolites in *A. sinensis* than that in other species.

3.4. Differential metabolites between A. keiskei and the three other Angelica species

Given the abundance of volatile metabolites in the root of the *A. keiskei* was the highest among the four species. It showed that the non-medicinal parts of *A. keiskei* also have potential to be exploited for practical uses. Therefore, the differences of metabolites between *A. keiskei* and the other *Angelica* species were further compared. The volcanic map visually showed the overall distribution of differential metabolites in each comparison. Four hundred and one significantly different metabolites (308 up-regulated and 93 down-regulated) were detected in the comparison between *A. keiskei* and *A. biserrata* (Figure 7a), which were related to phenylpropanoid biosynthesis (2 metabolites with $p = 0.21$), metabolic pathways (17 metabolites with $p = 0.38$) and tyrosine metabolism (3 metabolites with $p = 0.45$) (Figure 7c). Four hundred seventy-three significantly different metabolites (421 up-regulated and 52 down-regulated) were detected in the comparison between *A. keiskei* and *A. dahurica* (Figure 7b), which were associated with sesquiterpenoid and triterpenoid biosynthesis (8 metabolites with $p = 0.14$), monoterpenoid biosynthesis (9 metabolites with $p = 0.23$) and biosynthesis of secondary metabolites (22 metabolites with $p = 0.39$) (Figure 7d).

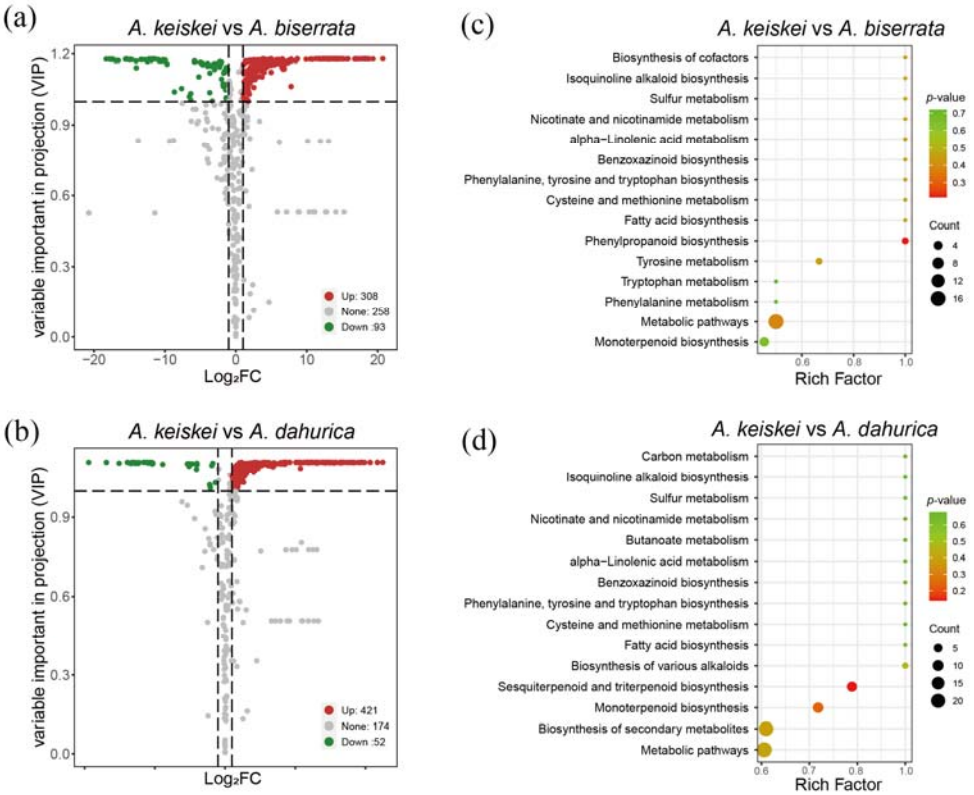

Figure 7. The overall distribution and KEGG enrichment analysis of differential metabolites between *A. keiskei* and *A. biserrata* (a, c), *A. keiskei* and *A. dahurica* (b, d). (a-b) Volcano plots for differential metabolites. The colors of metabolites indicated significant differences (red, upregulated; green, downregulated). (c-d) KEGG pathway enrichment analysis of differential metabolites. Color of the bubbles represented statistical significance of the enriched terms, and the size of the bubbles represented number of differential metabolites.

Moreover, to further investigate the differences of volatile metabolites in *A. keiskei* and other *Angelica* species, twenty metabolites that were differentiated the most between the two species were subsampled (Figure 8). From the comparison, the terpenoids metabolites, carvenone and cedrene were more abundant in *A. keiskei* than

that in *A. biserrata*; and carene, bornyl acetate and isobornyl acetate were the most enriched in *A. keiskei* compared with *A. dahurica*. Additionally, the β-pinene was more enriched in *A. dahurica* and *A. biserrata* than in *A. keiskei*.

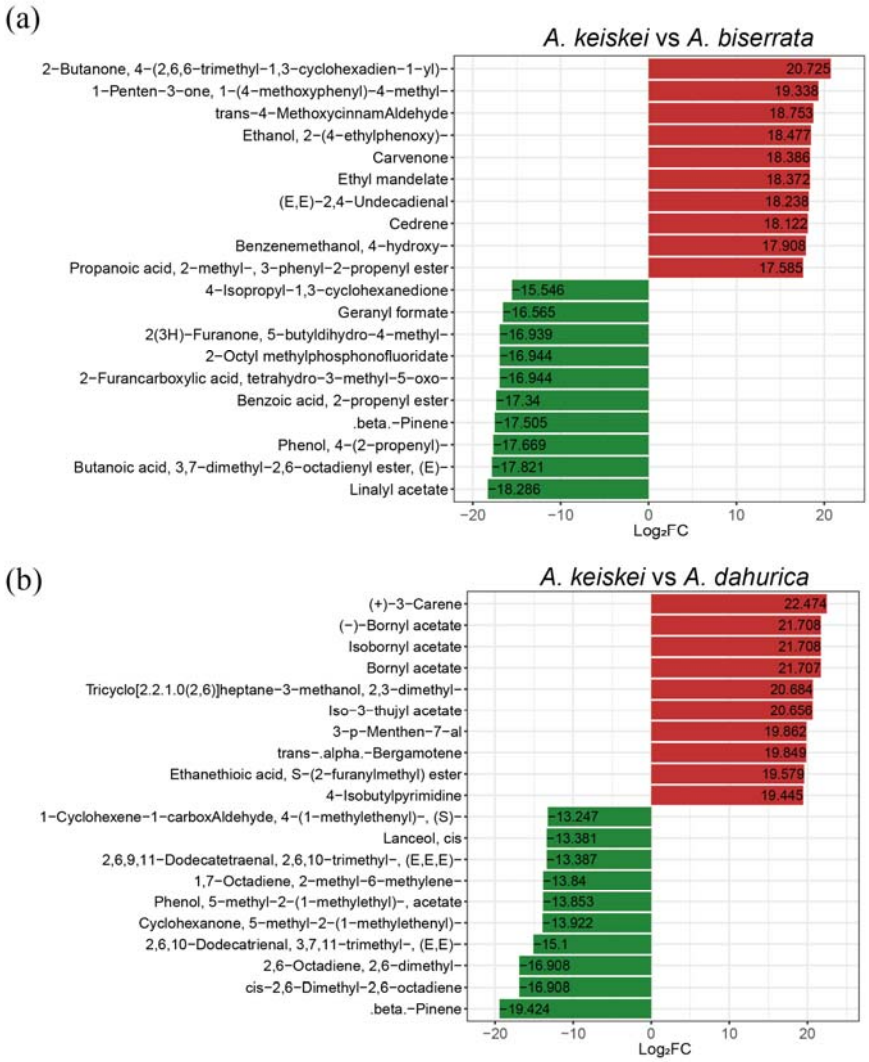

Figure 8. Top 20 metabolites with significant difference between *A. keiskei* and *A. biserrata* (a), *A. keiskei* and *A. dahurica* (b). Red and green represent up-regulated and down-regulated metabolites in *A. keiskei*, respectively.

4. Discussion

Widely targeted metabolomics offered a promising way for the chemical screening of volatile metabolites and allowed the characterization of new volatile metabolites in *Angelica* [10]. Using the method, a total of 698 volatile metabolites were identified and further classified into 15 different categories, including terpenoids, ester, heterocyclic, aromatics and 11 others (Figure 2). A clustering heat map of the metabolites showed significant difference among the four species. The number of types and abundance of volatile metabolites in *A. keiskei* was the highest. Consistent with the previous reports, terpenoids were the largest and most diverse class of volatile metabolites in the four *Angelica* species [1,2]. The pair wise comparisons between two species for the metabolite's differences revealed that there were fewer up-regulated metabolites in *A. sinensis* when compared to the other three species (*A. dahurica, A. keiskei, A. biserrata*) whereas, relative to *A. keiskei,* most differential metabolites were down-regulated in *A. dahurica and A. biserrata*. It demonstrates that the analysis of differential metabolites is useful for understanding the differences of chemical properties among the four species.

A. sinensis also known as "female ginseng" is a traditional herb, which has long been used to treat various gynecological conditions [5,6,25]. Phthalides is believed to be responsible for the bioactivities of *A. sinensis* [6,26]. This study shows that 3-butylisobenzofuran-1(3H)-one and Z-ligustilide were detected in the four species and its contents were relatively higher in *A. sinensis*, which is consistent with previous studies [5]. In addition, coumarin and its derivatives are one of the important heterocyclic metabolites [27], which is mainly used as anti-HIV, anticancer activity agents, and anticoagulant activities [28,29]. The results show that the contents of 7-hydroxycoumarin and 7-ethoxycoumarin in *A. sinensis* were significantly higher than *A. dahurica*, *A. biserrata*, and *A. keiskei* (Figure 2). By virtue of its structural

simplicity, 7-hydroxycoumarin has been generally accepted as the parent metabolites for the furocoumarins and pyranocoumarins and is widely used as a synthon for a wide variety of coumarin-heterocycles [30-32]. Its higher abundance in *A. sinensis* was probably associated with biosynthesis of furocoumarins and pyranocoumarins, which were reported as one of the main active components influencing the pharmaceutical activity of the herb [9]. Nevertheless, in ancient Chinese medical systems, the pharmacological effect of medicinal plants depends not only on the high abundance of a single compound, but also on the synergy of multiple active ingredients [33,34]. Furthermore, this study also found that the proportion of various components in volatile metabolites was more balanced in *A. sinensis* (Figure 2). This might explain the wide and common applications of *A. sinensis* in TCM.

Moreover, *A. keiskei* is called ashitaba in Japanese. Its leaves have been used as the medicinal part. And it is economically used as herbs, food and spices [35,36]. *A. keiskei* has been used as a medicine and food owing to its abundant pharmacological effects, including anti-cancer, lowering blood sugar and blood lipids, and improving human immunity [35,37]. However, these pharmacological effects have not been validated in scientific research. To date, it is only found in the form of raw materials in tea and cosmetics, which has limited its medicinal and clinical applications [36,38]. Interestingly, bornyl acetate, previously unmentioned terpenoid substances was detected with high expression levels in the root of *A. keiskei*, and it has been reported that bornyl acetate has antibacterial, insecticidal, and anesthetic effects symbiotically with other aromatic metabolites in the VOs [39]. This discovery provides a basis for the development and utilization of active ingredients in *A. keiskei* for health-related dietary supplements. Taken together, this study greatly enriches the database of chemical composition in *A. keiskei* and imply that *A. keiskei* exhibited benign

potential to be exploited as medicinal materials and health-related dietary supplements.

Previous studies have verified that plants with closer phylogenetic relationship are not only similar in morphology but also in chemical composition and curative effects [40-42]. This study indicated the high correspondence between the volatile metabolites and the phylogenetic relationships. Although more extensive sampling and deeper investigations would be necessary to reveal more reliable correlations, the study implied that phylogenetic relationships could serve as a window to coarsely apprehend the unknown biochemical diversity of some plants based on the known biochemical map of phylogenetically related species. This finding may offer a great tool for searching replacements of medicinal plant resources that are endangered with closely related non-endangered species.

5. Conclusions

This study investigated the metabolites of four *Angelica* species by using widely targeted metabolomics, and found the differed accumulation of medicinally important metabolites among species. For example, high levels of bornyl acetate metabolites accumulated in *A. keiskei*, whereas coumarins and phthalides were significantly lower in *A. keiskei* than in *A. sinensis*. Moreover, the high correspondence between the dendrogram of metabolite contents and the phylogenetic tree suggested a potential correlation between the volatile metabolites and the phylogenetic relationships. Taken all together, the present study provides a biochemical map for the exploitation, application, and development of the *Angelica* species as TCM or health-related dietary supplements.

Supplementary Materials

Figure S1: TIC chromatogram of quality control samples; Figure S2: Permutation test of OPLS-DA model; Figure S3: The violin plot of relative abundance of 15 classes in

the four *Angelica* species. Table S1: The volatile metabolites detected; Table S2: The *p*-values of the t-test for 15 classes in four species.

Funding sources

This work was supported by the National Key Research and Development Program of China, grant 2023YFA0915800; National Natural Science Foundation of China, grants 32300223, 32070242, and 82373837; Shenzhen Fundamental Research Program, grant 20220817165436004; Shenzhen Science and Technology Program, grant KQTD20161130104826 51; Key Project at Central Government Level (The ability establishment of sustainable use for valuable Chinese medicine resources), grant 2060302; Special Funds for Science Technology Innovation and Industrial Development of Shenzhen Dapeng New District, grants RC201901-05 and PT201901-19; China Postdoctoral Science Foundation, grant 2020 M672904; Basic and Applied Basic Research Fund of Guangdong, grant 2020A1515110912; Science, Technology, and Innovation Commission of Shenzhen Municipality of China, grant ZDSYS20200811142605017.

References

[1] Sowndhararajan, K., Deepa, P., Kim, M., Park, S. J., & Kim, S. (2017, Sep 20). A Review of the Composition of the Essential Oils and Biological Activities of Angelica Species. *Sci Pharm, 85*(3), 33. https://doi.org/10.3390/scipharm85030033

[2] Feng, T., Downie, S. R., Yu, Y., Zhang, X., Chen, W., He, X., & Liu, S. (2009, Jul). Molecular systematics of Angelica and allied genera (Apiaceae) from the Hengduan Mountains of China based on nrDNA ITS sequences: phylogenetic affinities and biogeographic implications. *J Plant Res, 122*(4), 403-414. https://doi.org/10.1007/s10265-009-0238-4

[3] Sarker, S., & Nahar, L. (2004). Natural medicine: the genus Angelica. *Current medicinal chemistry, 11*(11), 1479-1500.

[4] Dong, H., Li, M., Jin, L., Xie, X., Li, M., & Wei, J. (2022, Jan 5). Cool Temperature Enhances Growth, Ferulic Acid and Flavonoid Biosynthesis While Inhibiting Polysaccharide Biosynthesis in Angelica sinensis. *Molecules, 27*(1), 320. https://doi.org/10.3390/molecules27010320

[5] Hook, I. L. I. (2014, 2014/02/27/). Danggui to Angelica sinensis root: Are potential benefits to European women lost in translation? A review. *Journal of Ethnopharmacology, 152*(1), 1-13. https://doi.org/https://doi.org/10.1016/j.jep.2013.12.018

[6] Wei, W. L., Zeng, R., Gu, C. M., Qu, Y., & Huang, L. F. (2016, Aug 22). Angelica sinensis in China-A review of botanical profile, ethnopharmacology, phytochemistry and chemical analysis. *J Ethnopharmacol, 190*, 116-141. https://doi.org/10.1016/j.jep.2016.05.023

[7] Alkan Turkucar, S., Aktas Karacelik, A., & Karakose, M. (2021). Phenolic compounds, essential oil composition, and antioxidant activity of Angelica purpurascens (Ave-Lall.) Gill. *Turk J Chem, 45*(3), 956-966. https://doi.org/10.3906/kim-2101-28

[8] Zhang, H. Y., Bi, W. G., Yu, Y., & Liao, W. B. (2012). Angelica sinensis (Oliv.) Diels in China: distribution, cultivation, utilization and variation. *Genetic Resources and Crop Evolution, 59*(4), 607-613. https://doi.org/10.1007/s10722-012-9795-9

[9] Pandey, A. K., Kumar, P., Saxena, M., & Maurya, P. (2020). *Distribution of aromatic plants in the world and their properties*. Elsevier Science Publishing Co Inc.

[10] Kumar, P., Rana, V., & Singh, A. N. (2022). Angelica glauca Edgew. - A comprehensive review. *Journal of Applied Research on Medicinal and Aromatic Plants, 31*, 100397. https://doi.org/10.1016/j.jarmap.2022.100397

[11] Sadgrove, N. J., Padilla-Gonzalez, G. F., & Phumthum, M. (2022, Mar 16). Fundamental Chemistry of Essential Oils and Volatile Organic Compounds, Methods of Analysis and Authentication. *Plants (Basel), 11*(6), 789. https://doi.org/10.3390/plants11060789

[12] Perveen, A., Ijaz, S., & Ghaffar, N. (2020). Comparative phytochemical and

physicochemical study of seeds of the genus Angelica L. from Neelum valley Azad Kashmir, Pakistan. *Pakistan Journal of Botany, 52*(1), 257-260. https://doi.org/10.30848/pjb2020-1(44)

[13] Du, J., Bai, B., Kuang, X., Yu, Y., Wang, C., Ke, Y., Xu, Y., Tzang, A. H., & Qian, Z. M. (2006). Ligustilide inhibits spontaneous and agonists-or K+ depolarization-induced contraction of rat uterus. *Journal of ethnopharmacology, 108*(1), 54-58.

[14] Ma, J., Huang, J., Hua, S., Zhang, Y., Zhang, Y., Li, T., Dong, L., Gao, Q., & Fu, X. (2019, Mar 1). The ethnopharmacology, phytochemistry and pharmacology of Angelica biserrata - A review. *J Ethnopharmacol, 231*, 152-169. https://doi.org/10.1016/j.jep.2018.10.040

[15] Chen, Q., Hu, Y., Wen, R., Wang, Y., Qin, L., & Kong, B. (2021, 2021/02/01/). Characterisation of the flavour profile of dry fermented sausages with different NaCl substitutes using HS-SPME-GC-MS combined with electronic nose and electronic tongue. *Meat Science, 172*, 108338. https://doi.org/10.1016/j.meatsci.2020.108338

[16] Hua, Y. L., Ma, Q., Zhang, X. S., Yao, W. L., Ji, P., Hu, J. J., & Wei, Y. M. (2019, Feb). Urinary metabolomics analysis reveals the effect of volatile oil from Angelica sinensis on LPS-induced inflammation rats. *Biomed Chromatogr, 33*(2), e4402. https://doi.org/10.1002/bmc.4402

[17] Bertuzzi, A. S., McSweeney, P. L. H., Rea, M. C., & Kilcawley, K. N. (2018, Mar). Detection of Volatile Compounds of Cheese and Their Contribution to the Flavor Profile of Surface-Ripened Cheese. *Compr Rev Food Sci Food Saf, 17*(2), 371-390. https://doi.org/10.1111/1541-4337.12332

[18] Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., O'Hara, R. G., Simpson, G., Solymos, P., Stevens, H., & Wagner, H. (2013, 01/01). Multivariate analysis of ecological communities in R: vegan tutorial. R package version 1.7.

[19] Chong, J., & Xia, J. (2018, Dec 15). MetaboAnalystR: an R package for flexible and reproducible analysis of metabolomics data. *Bioinformatics, 34*(24), 4313-4314. https://doi.org/10.1093/bioinformatics/bty528

[20] Lê, S., Josse, J., & Husson, F. (2008, 03/18). FactoMineR: An R Package for

Multivariate Analysis. *Journal of statistical software*, *25*(1), 1-18. https://doi.org/10.18637/jss.v025.i01

[21] Katoh, K., & Standley, D. M. (2013, Apr). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*, *30*(4), 772-780. https://doi.org/10.1093/molbev/mst010

[22] Nguyen, L.-T., Schmidt, H. A., Von Haeseler, A., & Minh, B. Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular biology and evolution*, *32*(1), 268-274.

[23] Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., & Zhan, L. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *The Innovation*, *2*(3), 100141.

[24] Liao, C. Y., Gao, Q., Katz□Downie, D. S., & Downie, S. R. (2022). A systematic study of North American Angelica species (Apiaceae) based on nrDNA ITS and cpDNA sequences and fruit morphology. *Journal of Systematics and Evolution*, *60*(4), 789-808.

[25] Yeh, J. C., Cindrova-Davies, T., Belleri, M., Morbidelli, L., Miller, N., Cho, C. W., Chan, K., Wang, Y. T., Luo, G. A., Ziche, M., Presta, M., Charnock-Jones, D. S., & Fan, T. P. (2011, May). The natural compound n-butylidenephthalide derived from the volatile oil of Radix Angelica sinensis inhibits angiogenesis in vitro and in vivo. *Angiogenesis*, *14*(2), 187-197. https://doi.org/10.1007/s10456-011-9202-8

[26] Chen, X.-P., Li, W., Xiao, X.-F., Zhang, L.-L., & Liu, C.-X. (2013). Phytochemical and pharmacological studies on Radix Angelica sinensis. *Chinese Journal of Natural Medicines*, *11*(6), 577-587. https://doi.org/10.1016/s1875-5364(13)60067-9

[27] Wu, L., Wang, X., Xu, W., Farzaneh, F., & Xu, R. (2009). The structure and pharmacological functions of coumarins and their derivatives. *Curr Med Chem*, *16*(32), 4236-4260. https://doi.org/10.2174/092986709789578187

[28] Zhou, D., Li, N., Zhang, Y., Yan, C., Jiao, K., Sun, Y., Ni, H., Lin, B., & Hou, Y. (2016). Biotransformation of neuro-inflammation inhibitor kellerin using Angelica sinensis (Oliv.) Diels callus. *RSC Advances*, *6*(99), 97302-97312.

https://doi.org/10.1039/c6ra22502k

[29] Kim, C. Y., Mitchell, A. J., Kastner, D. W., Albright, C. E., Gutierrez, M. A., Glinkerman, C. M., Kulik, H. J., & Weng, J.-K. (2023, 2023/02/03). Emergence of a proton exchange-based isomerization and lactonization mechanism in the plant coumarin synthase COSY. *Nature Communications*, *14*(1), 597. https://doi.org/10.1038/s41467-023-36299-1

[30] Mazimba, O. (2017, 2017/12/01/). Umbelliferone: Sources, chemistry and bioactivities review. *Bulletin of Faculty of Pharmacy, Cairo University*, *55*(2), 223-232. https://doi.org/https://doi.org/10.1016/j.bfopcu.2017.05.001

[31] Vanholme, R., Sundin, L., Seetso, K. C., Kim, H., Liu, X., Li, J., De Meester, B., Hoengenaert, L., Goeminne, G., Morreel, K., Haustraete, J., Tsai, H. H., Schmidt, W., Vanholme, B., Ralph, J., & Boerjan, W. (2019, Oct). COSY catalyses trans-cis isomerization and lactonization in the biosynthesis of coumarins. *Nat Plants*, *5*(10), 1066-1075. https://doi.org/10.1038/s41477-019-0510-0

[32] Han, X., Li, C., Sun, S., Ji, J., Nie, B., Maker, G., Ren, Y., & Wang, L. (2022, Oct 18). The chromosome-level genome of female ginseng (Angelica sinensis) provides insights into molecular mechanisms and evolution of coumarin biosynthesis. *Plant J*. https://doi.org/10.1111/tpj.16007

[33] Liu, P., Yang, H., Long, F., Hao, H. P., Xu, X., Liu, Y., Shi, X. W., Zhang, D. D., Zheng, H. C., Wen, Q. Y., Li, W. W., Ji, H., Jiang, X. J., Zhang, B. L., Qi, L. W., & Li, P. (2014, Jul). Bioactive equivalence of combinatorial components identified in screening of an herbal medicine. *Pharm Res*, *31*(7), 1788-1800. https://doi.org/10.1007/s11095-013-1283-1

[34] Song, H. P., Wu, S. Q., Hao, H., Chen, J., Lu, J., Xu, X., Li, P., & Yang, H. (2016, Mar 30). A chemical family-based strategy for uncovering hidden bioactive molecules and multicomponent interactions in herbal medicines. *Sci Rep*, *6*, 23840. https://doi.org/10.1038/srep23840

[35] Kil, Y. S., Pham, S. T., Seo, E. K., & Jafari, M. (2017, Jun). Angelica keiskei, an emerging medicinal herb with various bioactive constituents and biological activities. *Arch Pharm Res*, *40*(6), 655-675. https://doi.org/10.1007/s12272-017-0892-3

[36] Rong, Y., Gu, X., Li, D., Chen, L., Zhang, Y., & Wang, Z. (2021, 2021/07/01). Characterization of aroma, sensory and taste properties of Angelica keiskei tea. *European Food Research and Technology*, *247*(7), 1665-1677. https://doi.org/10.1007/s00217-021-03737-7

[37] Guiné, R. P., & Gonçalves, F. J. (2016). Bioactive Compounds in Some Culinary Aromatic Herbs and Their Effects on Human Health. *Mini Rev Med Chem*, *16*(11), 855-866. https://doi.org/10.2174/1389557516666160211120540

[38] Kim, D. W., Curtis-Long, M. J., Yuk, H. J., Wang, Y., Song, Y. H., Jeong, S. H., & Park, K. H. (2014, Jun 15). Quantitative analysis of phenolic metabolites from different parts of Angelica keiskei by HPLC-ESI MS/MS and their xanthine oxidase inhibition. *Food Chem*, *153*, 20-27. https://doi.org/10.1016/j.foodchem.2013.12.026

[39] Liang, H., Lin, X., Yang, P., Sun, Y., Wu, Q., Alimujiang, S., Zhao, H., Ma, D., Zhan, R., & Yang, J. (2022). Genome-Wide Identification of BAHD Superfamily and Functional Characterization of Bornyl Acetyltransferases Involved in the Bornyl Acetate Biosynthesis in Wurfbainia villosa. *Frontiers in Plant Science*, *13*, 860152. https://doi.org/10.3389/fpls.2022.860152

[40] Hao, D.-c., & Xiao, P.-g. (2020). Pharmaceutical resource discovery from traditional medicinal plants: Pharmacophylogeny and pharmacophylogenomics. *Chinese Herbal Medicines*, *12*(2), 104-117. https://doi.org/10.1016/j.chmed.2020.03.002

[41] Kang, K. B., Ernst, M., van der Hooft, J. J., da Silva, R. R., Park, J., Medema, M. H., Sung, S. H., & Dorrestein, P. C. (2019). Comprehensive mass spectrometry□guided phenotyping of plant specialized metabolites reveals metabolic diversity in the cosmopolitan plant family Rhamnaceae. *The Plant Journal*, *98*(6), 1134-1144.

[42] Saslis-Lagoudakis, C. H., Klitgaard, B. B., Forest, F., Francis, L., Savolainen, V., Williamson, E. M., & Hawkins, J. A. (2011). The use of phylogeny to interpret cross-cultural patterns in plant use and guide medicinal plant discovery: an example from Pterocarpus (Leguminosae). *PLoS ONE*, *6*(7), e22275.