

ORIGINAL ARTICLE

Open Access



Geographical origin classification of *Phoebe zhennan* and *Phoebe bournei* by solid phase micro-extraction and gas chromatography-mass spectrometry

Zhenan Chen^{1,2†} , Xiaoming Xue^{1*†}, Ruoyi Cheng^{1,2†}, Haoqi Wu², Handng Gao² and Zexun Gao¹

Abstract

Phoebe zhennan and *Phoebe bournei* are favored for their fine-grain, insect resistance, special fragrance, not easily cracked and deformed, which had been widely used in furniture and construction. Wood authenticity regarding different varieties and geographical origins is increasingly becoming a concern for consumers. Due to the difference in wood quality characteristics and economic value of genus *Phoebe* species from different geographical origins, it is important to establish a fast, efficient, and reliable method to discriminate the geographical origin of *Phoebe zhennan* and *Phoebe bournei*. Solid phase micro-extraction (SPME) was used as sample preparation approach and gas chromatography-mass spectrometry (GC–MS) technique was employed to analyze the volatile compounds. As a result, a total of 40 volatiles were identified by SPME–GC–MS among *Phoebe zhennan* samples from four origins, while 34 substances identified among *Phoebe bournei* samples. The wood samples of *Phoebe zhennan* and *Phoebe bournei* from different major production areas in China were able to be discriminated by GC–MS spectroscopy. This finding indicated that genus *Phoebe* species from different climate zones had different chemical composition, which verified that the SPME–GC–MS technique can be used to classify different species of wood with characteristic compounds.

Keywords SPME, GC–MS, Geographical origin, Fingerprint analysis, Wood

Introduction

Phoebe zhennan and *Phoebe bournei* are trees belonging to the genus *Phoebe* Nees of the Lauraceae family [1]. *Phoebe zhennan* is found mainly in western Hubei, northwest Guizhou, and Sichuan, while *Phoebe bournei* is usually cultivated in Guangdong, Fujian, Guangxi, and Guizhou. These two kinds of wood are favored for their fine-grain, insect resistance, special fragrance, not easily

cracked and deformed, which had been widely used in furniture and construction [2, 3]. The wood properties of *Phoebe zhennan* and *Phoebe bournei* depend on the geographical origin, growing environment and cultivation method. It is commonly believed that the *Phoebe zhennan* population from Sichuan Province showed the best wood quality traits and have the highest price in comparison to all origins [4]. With the rapid globalization of wood trade, the issue of wood authenticity has become an urgent and worldwide problem, such as the confusion of varieties and origins.

According to our findings, the macrostructural and microstructural differences of genus *Phoebe* species from different geographical origins are not obvious. It is very difficult for non-professional identifiers to distinguish among them. However, both *Phoebe zhennan* and *Phoebe*

[†]Zhenan Chen, Xiaoming Xue and Ruoyi Cheng contributed equally to this work and co-first author.

*Correspondence:

Xiaoming Xue
xuexm@nfpc.edu.cn

¹ Nanjing Forest Police College, Nanjing, Jiangsu, China

² Nanjing Forestry University, Nanjing, Jiangsu, China

bournei have a special odor, making them suitable for the analysis by gas chromatography–mass spectrometry (GC–MS) approach [5]. The technique of GC–MS has been widely employed to analyze complex and unknown samples because of its rapid, sensitivity, strong qualitative and other superior traits. Besides, solid phase microextraction (SPME) technique has been employed to extract volatile compounds. SPME was considered as a fast economical sensitive and solvent free volatile extraction method before the GC–MS analysis. In recent years, the SPME–GC–MS technique has been used analyzing the volatile profiles of wood samples [6]. Chemometric has also been used to predict some wood qualities and properties, for example authenticity of oak [7–9], and quality of eaglewood [10].

In this study, the technique of SPME–GC–MS was used to identify volatile compounds in *Phoebe zhennan* and *Phoebe bournei* samples. Meanwhile, the datasets of GC–MS were subjected to chemometric analyses to identify the geographical origins of *Phoebe zhennan*. The proposed SPME–GC–MS will provide an accurate and quick way for the geographical origin identification of genus *Phoebe* species.

Materials and methods

Samples

The Latin names, Chinese names, and information on sampling location are presented in Table 1. The samples were obtained from China National Forestry and Grassland Administration Wildlife Criminal Evidence Identification Center (Nanjing Forest police college), collected from natural plant populations. *Phoebe zhennan* samples (130 samples) were collected from Hubei Province, Sichuan Province, Yunnan Province, and Guizhou

Province, respectively, in China. *Phoebe bournei* samples (45 samples) were collected from Guizhou Province, Hubei Province, and Guangxi Province. The sampled trees had an average diameter at breast of 21.7 cm and were from 18 to 28 years old. From each tree, one cores was extracted at 1.3 m stem height by a 12-mm-diameter borer. Wood samples were chopped into pieces and finely powdered to pass a 40-mesh sieve. Then, 1.5 g of wood powder was added into a headspace bottle and extracted by SPME technique.

We declare that this study has the official permission to collect the plant sample and complies with the Chinese legislation. Wood species have been identified by Xiaoming Xue (based on phenotypic characteristics).

GC–MS instrument and conditions

GC–MS analysis was performed on an Agilent 7000B System (Agilent Technologies, USA). Separation was achieved using a DB-5MS column (length: 30 m, inner diameter: 0.25 mm, film thickness: 0.25 micro meter) with a temperature program shown in Table 2. Helium was used as the carrier gas with the gas flow rate through the column was 1 mL/min. The mass spectrometry program was set to electron impact ionization mode, energy 70 eV, mass scanning range 33–450 amu and full scanning

Table 2 Temperature program setting

Temperature (°C)	Rate (°C/min)	Hold (min)	Total (min)
60	–	1	1
120	10	0	7
160	5	6	21
210	10	3	29

Table 1 Listing of samples for testing

Latin name	Chinese name	Geographical origin	Sampling location	Elevation (m)	Soil type	Number of test samples
<i>Phoebe zhennan</i>	Zhennan	Hubei	Hefeng	1293	Yellow–brown soil	22
		Sichuan	Leshan	1347	Yellow–brown soil	12
			Gulan	1258	Purple soil	14
			Dazhu	1298	Yellow–brown soil	16
		Yunnan	Zhaotong	1840	Red soil	20
		Guizhou	Suiyang	1350	Red soil	12
			Fanjing Mountain	1476	Red soil	14
			Guandu	1229	Red soil	11
			Zunyi	1280	Red soil	9
<i>Phoebe bournei</i>	Mingnan	Guizhou	Fanjing Mountain	1576	Red soil	18
		Hubei	Laifeng	1353	Yellow–brown soil	10
		Guangxi	Baise	1120	Red soil	9
			Jinxiu	1571	Red soil	8

mode. Compounds were identified via the National Institute of Standards and Technology (NIST17; Gaithersburg, MD, USA) Mass Spectral Library. A mixture of 1 μ L n-alkane (C7–C25) was run under the same capillary column and GC–MS conditions to obtain the retention index (RI). The peak area was chosen as the analytical signal for the relative amount. The relative contents of each volatile were calculated by area normalization, and the average value of all tested wood samples was taken.

Data analysis

The Random Forest (RF) model and k-nearest neighbors (KNN) model were used to classify *Phoebe zhennan* samples by MATLAB R2018b (MathWorks Inc., Natick, MA, USA). The dataset was randomly divided into a training set (70%) and a test set (30%). Optimal parameters adjust and five-fold cross-validation was performed by Toolbox 802 (Eigenvector Research, Inc., Manson, WA).

The RF algorithm was proposed by Breiman in 2001. This method based on a set of decision trees formed using bootstrap samples from a learning data set and can solve multi-classification and regression problems [11]. The final classification is determined by calculating the predicted frequency of each group in the entire forest. RF has achieved high performance in the classification of woods [12], foods [13], pests [14], pesticide [15], etc. In this study, we explore the potential of RF for modeling GC–MS data of *Phoebe zhennan* samples for identifying.

Besides the RF algorithm, the KNN are also used in modeling the same GC–MS data for comparison. As a well-known supervised pattern recognition method, KNN is based on the principle that the majority of the most similar samples of one sample belong to a certain category, this sample also belongs to this category [16]. This method is not limited by the number of categories and has proven to be suitable for GC–MS data analysis [17].

Due to the insufficient number of *Phoebe bournei* samples, its GC–MS data were not used for model building.

Results and discussion

Results of GC–MS analysis

The volatile compounds extracted from *Phoebe zhennan* and *Phoebe bournei* were separated by GC–MS, in total 51 kinds of chemical composition (peak area above 0.5%) were scrutinized (Tables 3, 4), including alkanes, olefins, alcohols, naphthalene, and aldehydes compounds.

Figure 1 shows the total ion chromatograms of *Phoebe zhennan*, the results clearly exhibited the different characteristic peaks of the samples from four geographical origins. The characteristic peaks of the wood samples

from Hubei and Sichuan mainly appeared between 6 and 25 min while the samples from Yunnan and Guizhou mainly appeared between 9 and 25 min. We observed that the spectra of wood samples from Yunnan and Guizhou could not be used for geographical origin classification. We believe that the samples from Sichuan province can be well distinguished by the total ion chromatograms. The quality of *Phoebe zhennan* from Sichuan production areas is different from that produced in surrounding areas due to different geographical and climatic conditions [4]. Therefore, the accurate identification of samples from this region is the main focus of this study.

Figure 2 shows the total ion chromatogram of *Phoebe bournei* obtained from GC–MS, it can be seen that the characteristic peaks of all specimens mainly appeared between 9 and 24 min, but the intensity of the peaks differed among the three geographical origins. Results demonstrated that three geographical origins of *Phoebe bournei* can be readily distinguished by visual inspection. In total, 34 volatiles were identified and presented in Table 5.

The difference analysis among four geographical origins of *Phoebe zhennan*

The retention time, chemical name, relative content, RI, and CAS number are presented in Table 3. Wood samples of *Phoebe zhennan* from Hubei, Sichuan, Yunnan, and Guizhou Province were extracted by SPME technique. In this study, 40 kinds of chemical composition have been identified in wood samples.

Among all the identified compounds, verbenone was the constitute found with the highest proportion in the *Phoebe zhennan* samples of the total chromatogram area. The (E)-verbenol and α -copaene, were the second and the third most abundant volatiles, respectively. The relative content of α -cubebene, α -ylangene, β -elemene, arisotolene, γ -muurolene, δ -cadinene, and α -agarofuran was significantly different in *Phoebe zhennan* samples from the four geographical origins. Six compounds were identified only in wood samples from Yunnan and Guizhou: (E)-verbenol, Cedrol, β -eudesmol, γ -eudesmol, agarospirol and valerianol. It suggests that these substances may be thought of as characteristic components for the origin classification of *Phoebe zhennan* in Yunnan and Guizhou. Closer inspection of the Table 3 revealed that a small difference in the composition of volatile compounds between samples from Yunnan and Guizhou, whereas the difference between other geographical origins is more evident.

The full-spectrum GC–MS data of *Phoebe zhennan* samples were used to establish RF and KNN models.

Table 3 Chemical composition of extractives analyzed by GC–MS (*Phoebe zhennan*)

RT (min)	Chemical name	Average content (%)				RI-practical	RI-NIST 17 library	CAS No
		Hubei	Sichuan	Yunnan	Guizhou			
5.98	D- α -pinene	–	1.19	–	–	933	937	7785-70-8
6.3	2,4-thujadiene	3.93	4.17	0.79	1.01	960	960	36262-09-6
7.32	Ortho-cymene	2.09	3.35	–	–	1030	1027	527-84-4
8.4	Para- α -dimethyl styrene	1.44	1.57	–	–	1087	1091	1195-32-0
9.04	α -campholenal	2.37	–	–	–	1132	1127	4501-58-0
9.31	Nopinone	1.05	1.61	–	–	1148	1142	38651-65-9
9.43	(E)-verbenol	–	–	8.97	10.23	1154	1151	1820-09-3
9.75	Pinocarvone	2.07	2.06	0.73	0.81	1173	1164	30460-92-5
10.39	Myrtenol	2.52	2.46	1.83	1.85	1208	1202	515-00-4
10.67	Verbenone	17.24	22.74	14.76	14.67	1221	1217	80-57-9
13.46	α -cubebene	6.78	3.02	4.50	3.34	1357	1354	17699-14-8
13.97	α -ylangene	2.18	1.68	3.82	1.14	1382	1375	14912-44-8
14.1	α -copaene	7.27	6.97	5.73	7.23	1387	1380	3856-25-5
14.39	β -elemene	6.61	3.48	3.89	3.00	1401	1394	515-13-9
14.85	cis- α -bergamotene	1.50	1.19	0.82	1.23	1420	1415	18252-46-5
15.01	α -cedrene	2.69	2.77	2.06	2.89	1428	1422	469-61-4
15.11	β -caryophyllene	1.52	1.94	1.06	1.20	1432	1429	87-44-5
15.41	Thujopsene	4.08	5.71	4.65	5.89	1444	1441	470-40-6
15.55	Aristolene	3.12	3.96	2.41	3.32	1455	1455	6831-16-9
15.91	α -humulene	1.04	1.04	–	0.95	1466	1462	6753-98-6
16.4	γ -muurolene	2.76	3.11	2.52	2.85	1492	1502	30021-74-0
16.75	β -guaiane	–	0.88	0.74	0.72	1505	1504	88-84-6
16.89	Valencene	2.00	1.49	–	–	1504	1496	4630-07-3
16.96	4-epi-cubebol	0.96	–	2.07	2.37	1507	1499	38230-60-3
17.19	β -dihydroagarofuran	–	1.30	0.81	–	1515	1504	5956-09-2
17.27	Cuparene	0.90	1.30	1.28	1.82	1516	1515	16982-00-6
17.65	δ -cadinene	6.98	6.41	3.72	5.27	1528	1526	483-76-1
17.98	cadinadiene	1.59	–	–	–	1539	1533	29837-12-5
18.35	α -calacorene	0.92	1.21	–	0.69	1551	1546	21391-99-1
18.54	α -agarofuran	1.65	2.35	3.18	2.61	1558	1553	5956-12-7
19.71	Furopolargone A	2.6	2.98	–	–	1594	1588	1143-45-9
19.82	β -caryophyllene oxide	1.35	–	3.27	3.01	1597	1594	1139-30-6
20.68	Cedrol	–	–	2.13	2.27	1622	1615	77-53-2
20.82	Humulene oxide II	2.58	–	0.75	0.82	1626	1620	19888-34-7
21.12	β -eudesmol	–	–	2.60	1.80	1644	1649	473-15-4
21.29	γ -eudesmol	–	–	1.47	1.70	1649	1652	1209-71-8
21.64	Longiverbenone	1.60	1.26	0.90	–	1650	1649	64180-68-3
21.84	Agarospinol	–	–	6.48	6.83	1654	1643	1460-73-7
22.39	Valerianol	–	–	5.55	5.79	1661	1661	20489-45-6
23.1	Mustakone	4.62	6.81	6.51	2.69	1691	1687	1209-91-2

Relative content deviations lower than 0.1% were not listed in the table

Optimal parameters were selected according to the principle of the cross-validation and minimum discrimination rate. The highest discrimination rate of KNN was obtained when $k=1$ (single nearest neighbor). The

classification results of RF model and KNN model are shown in Table 4. In the RF model and KNN model, the classification rates of the training set and the test set for the Hubei and Sichuan samples were both 100%. The RF

Table 4 Prediction accuracy of RF and KNN models for *Phoebe zhennan* samples based on their GC–MS full-spectrum data

Geographical origins	RF		KNN	
	Training set	Test set	Training set	Test set
Hubei	100 (15/15)	100 (7/7)	100 (15/15)	100 (7/7)
Sichuan	100 (29/29)	100 (13/13)	100 (29/29)	100 (13/13)
Yunnan	71 (10/14)	50 (3/6)	57 (8/14)	50 (3/6)
Guizhou	88 (28/32)	79 (11/14)	81 (26/32)	79 (11/14)

The number of correct/total classifications appears in parentheses (in percent)

model training set classification rate for Yunnan samples was 71%, for Guizhou samples was 88%. While the KNN model training set classification rate for Yunnan samples was 57%, for Guizhou samples was 81%. These results showed that the *Phoebe zhennan* samples from Hubei and Sichuan can be accurately classified using RF and KNN models.

The sampling sites in Guizhou and Yunnan Province are located in Yunnan-Guizhou Plateau and exhibit typical karst landforms [18]. Being located in the southwestern of China, Yunnan-Guizhou Plateau possesses a short sunshine duration and low sunshine intensity. This region is located in a typical monsoon climate zone, with the highest moisture content in summer and the lowest moisture in winter. The monsoon's strength or weakness, and advancement or retreat, often result in the occurrence of flooding or droughts [19]. The climatic characteristics of sampling sites in Yunnan and Guizhou are similar [20]. Meanwhile, the main soil type at Yunnan and Guizhou is red soil, mainly derived from limestone parent material [21], which is different from other sampling sites.

Therefore, it is unreasonable to divide *Phoebe zhennan* provenance through administrative regions. Because of the close geographical proximity and the similar climatic conditions, it seems reasonable to combine wood samples from Yunnan and Guizhou into a single category.

The difference analysis among three geographical origins of *Phoebe bournei*

The relative content of the same ingredient in *Phoebe bournei* samples from the three geographical origins were different. There were 25 common volatile compounds among the three kinds of wood specimens (Table 5). The main components of *Phoebe bournei* samples from Guangxi were verbenone (13.07%), (E)-verbenol (8.42%), α -copaene (6.97%), valerianol (6.85%), mustakone (6.80%), and agarospirol (6.04%), while Guizhou origins were verbenone (13.09%), (E)-verbenol (10.24%) and agarospirol (6.40%). Analysis also revealed the presence

of compounds in Hubei origin with verbenone (19.63%), (E)-verbenol (13.52%) and thujopsene (6.49%) being the most predominant.

In addition to these volatiles, it is worth noting the presence of borneol, Cubenol, Ylangenal, since these unique compounds can be used as evidence for the classification of the tree origins. cubenol was found exclusively in *Phoebe bournei* of Guangxi origin and borneol was only found in Guizhou origin, a geographical classification could be achieved using a detect of these compounds.

Unlike *Phoebe zhennan*, the chemical composition of *Phoebe bournei* samples taken from different sampling locations showed significant variation, probably due to climate factors and genetic diversity existing within plant populations. Overall, GC–MS analyzing and comparing characteristic components of fingerprint can provide theory basis for the geographical origin classification of *Phoebe bournei*.

The difference analysis between *Phoebe zhennan* and *Phoebe bournei*

Phoebe zhennan was considered to be the most authentic Gold Phoebe, which has a much greater economic value than *Phoebe bournei*. As shown in Table 3 and Table 5, the relative contents in *Phoebe zhennan* samples were characterized by higher levels of α -cubebene and cedrol compared to *Phoebe bournei*, which allowed the differentiation between these two species. Interestingly, a high content of compounds such as verbenone and α -copaene could not be used as discriminating variables, since not all origins of genus *Phoebe* contained these substances.

To distinguish *Phoebe zhennan* and *Phoebe bournei*, the volatiles only detected in one species were defined as characteristic compounds. As shown in Table 6 and Table 7, 17 kinds of specific phytoconstituents were identified in *Phoebe zhennan* wood samples, while 12 kinds of volatiles were only identified and treated as characteristic compounds in *Phoebe bournei*. In summary, *Phoebe zhennan* and *Phoebe bournei* can be classified successfully according to chemical composition.

Conclusions

In this study, SPME–GC–MS technique has been used to analyze *Phoebe zhennan* and *Phoebe bournei* wood samples. In total, 40 volatiles were identified in *Phoebe zhennan* and 34 substances identified in *Phoebe bournei*, these phytoconstituents were classified into alkanes, olefins, alcohols, naphthalene, and aldehydes compounds, with verbenone showed the highest relative content in both species.

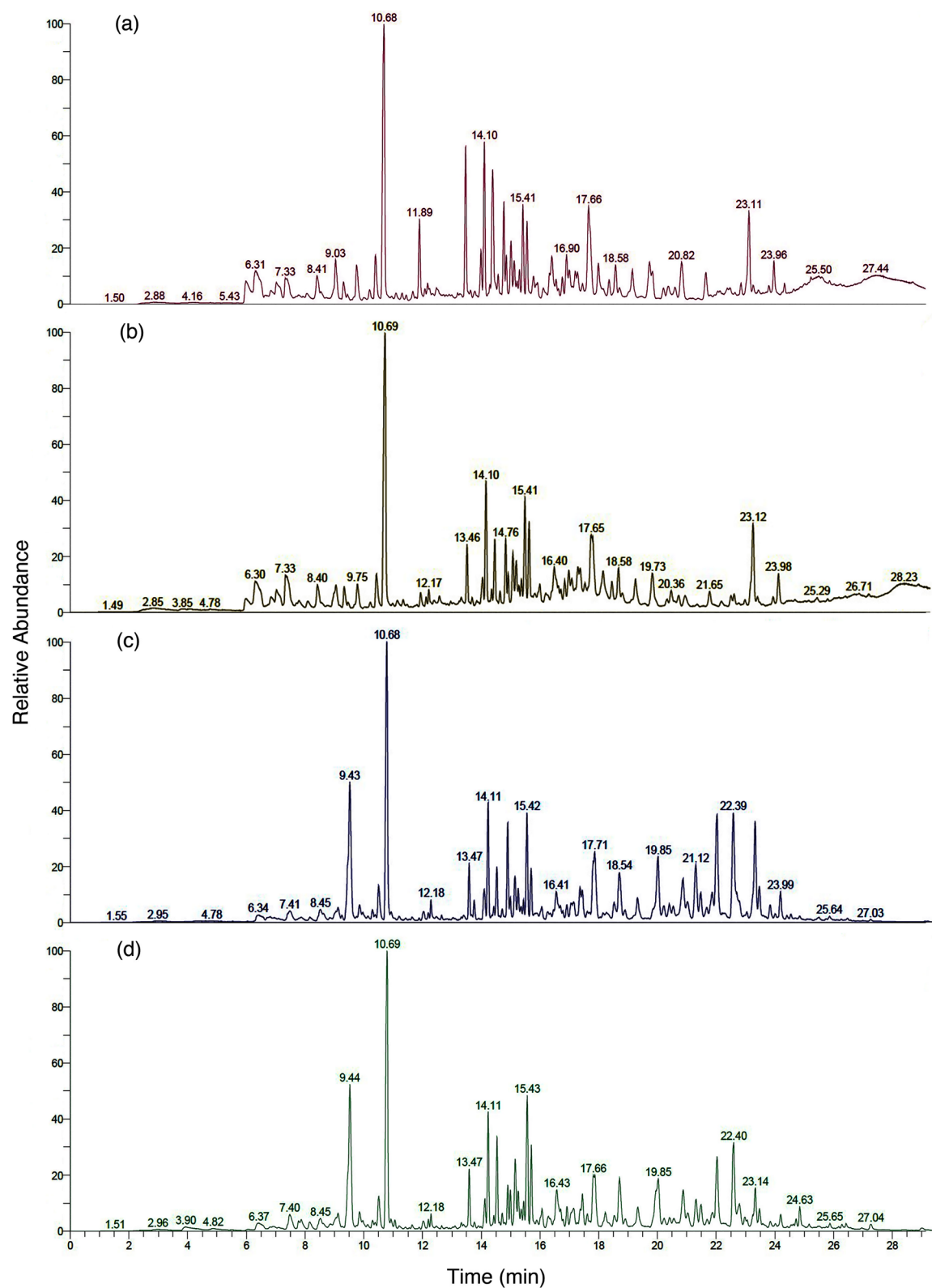


Fig. 1 GC–MS total ion chromatogram of *Phoebe zhennan*. **a** GC–MS total ion chromatogram of specimens from Hubei. **b** GC–MS total ion chromatogram of specimens from Sichuan. **c** GC–MS total ion chromatogram of specimens from Yunnan. **d** GC–MS total ion chromatogram of specimens from Guizhou

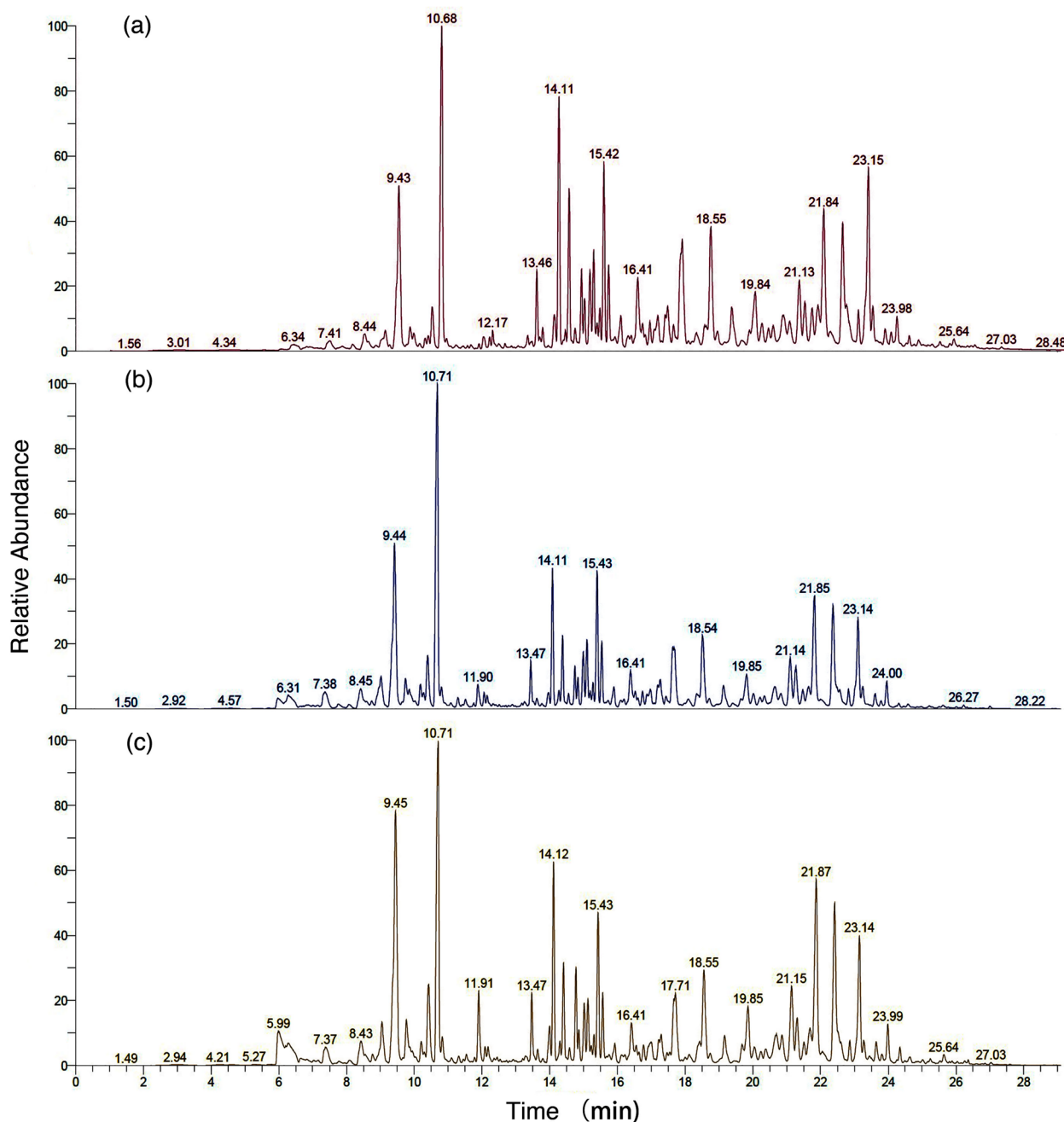


Fig. 2 GC-MS total ion chromatogram of *Phoebe bournei*. **a** GC-MS total ion chromatogram of specimens from Guizhou. **b** GC-MS total ion chromatogram of specimens from Hubei. **c** GC-MS total ion chromatogram of specimens from Guangxi

The geographical origins of *Phoebe zhennan* and *Phoebe bournei* were well-separated based on GC-MS analysis, which indicated that volatile compounds play an important role in the geographical discrimination of wood samples. Meanwhile, our experimental data suggested that the chemical composition of *Phoebe zhennan*

from Yunnan and Guizhou are highly similar. It seems reasonable to combine Yunnan origin and Guizhou origin into a single category. Using the content of characteristic compounds combined with the RF or KNN models is an accurate and effective method to identify the geographic origin of *Phoebe zhennan*.

Table 5 Chemical composition of extractives analyzed by GC–MS (*Phoebe bournei*)

RT (min)	Chemical name	Average content (%)			RI-practical	RI-NIST 17 library	CAS No
		Guangxi	Guizhou	Hubei			
6.3	2,4-thujadiene	–	1.18	2.02	960	960	36262-09-6
7.38	M-cymene	–	1.30	1.58	1033	1026	535-77-3
9.43	(E)-verbenol	8.42	10.24	13.52	1154	1151	1820-09-3
9.75	Pinocarvone	–	1.29	2.05	1173	1164	30460-92-5
9.87	Borneol	–	0.77	–	1179	1177	507-70-0
10.21	Cherry propanol	–	0.86	1.06	1198	1193	1197-01-9
10.39	Myrtenol	1.67	2.54	3.15	1208	1202	515-00-4
10.67	Verbenone	13.07	18.90	19.63	1221	1217	80-57-9
13.46	α -cubebene	1.70	1.70	1.84	1357	1354	17699-14-8
13.96	Cyclosativene	1.33	0.72	1.42	1381	1373	22469-52-9
14.1	α -copaene	6.97	5.61	5.19	1387	1380	3856-25-5
14.39	β -elemene	4.74	2.76	2.60	1401	1394	515-13-9
14.85	Cis- α -bergamotene	1.01	0.94	0.84	1420	1415	18252-46-5
15.01	α -cedrene	1.86	2.11	2.38	1427	1422	469-61-4
15.11	β -caryophyllene	2.19	2.65	1.91	1432	1429	87-44-5
15.41	Thujopsene	4.68	5.37	6.49	1444	1441	470-40-6
15.55	Aristolene	2.12	2.30	2.26	1455	1455	6831-16-9
15.91	α -humulene	0.98	1.00	0.73	1466	1462	6753-98-6
16.39	γ -amorphene	2.26	1.55	1.5	1492	1494	6980-46-7
16.99	α -amorphene	0.98	0.81	0.59	1513	1519	483-75-0
17.19	β -dihydroagarofuran	1.18	0.79	–	1515	1504	5956-09-2
17.27	Cuparene	1.10	1.27	1.97	1517	1515	16982-00-6
17.7	Calamenene	5.94	4.67	3.20	1535	1537	483-77-2
18.54	α -elemol	5.22	3.95	2.77	1557	1551	639-99-6
19.82	β -caryophyllene oxide	2.08	1.71	1.36	1597	1594	1139-30-6
20.03	Isospathulenol	0.90	–	0.59	1608	1620	88,395-46-4
20.68	Cedrol	1.24	1.00	0.85	1627	1625	77-53-2
21.12	β -eudesmol	2.99	2.56	2.92	1639	1649	473-15-4
21.29	γ -eudesmol	1.70	2.00	1.28	1644	1652	1209-71-8
21.5	Cubenol	1.12	–	–	1650	1649	21284-22-0
21.84	Agarospirol	6.04	6.40	4.37	1659	1654	1460-73-7
22.39	Valerianol	6.85	5.41	4.38	1667	1661	20489-45-6
22.85	Elemoyl acetate	1.18	0.64	–	1681	1679	60031-93-8
23.1	Mustakone	6.80	4.35	4.19	1691	1687	1209-91-2

Relative content deviations lower than 0.1% were not listed in the table

Because *Phoebe zhennan* and *Phoebe bournei* wood are expensive, it's valuable to establish a rapid and micro-destructive classification method. Although the prediction of *Phoebe zhennan* and *Phoebe bournei* by GC–MS technique and statistical analysis obtained high accuracy, given our limited number of samples, more validation studies are still necessary. Meanwhile, the composition of chemical substances in plants is determined by a combination of intrinsic and extrinsic factors, and how to define geographical boundaries

between different plant populations is still a question worthy of further study.

We also believe that the wood geographical origin classification would be much more accuracy if more detailed analyses are performed. Therefore, it is necessary to combine chemical constituents of woods with other wood properties (such as vessels distribution, ring width or density, etc.) to make sure the classification results are accurate and satisfactorily reliable.

Table 6 Unique components extracted from *Phoebe zhennan*

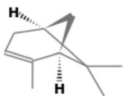
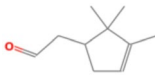
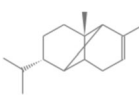
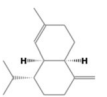
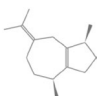
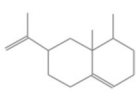
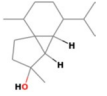
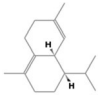
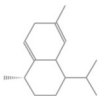
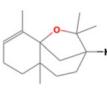
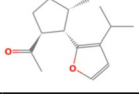
ID	Chemical name	Chemical structure	Formula	Molecular weight
1	d- α -pinene		C ₁₀ H ₁₆	136
2	Ortho-cymene		C ₁₀ H ₁₄	134
3	Para- α -dimethyl styrene		C ₁₀ H ₁₂	132
4	α -campholenal		C ₁₀ H ₁₆ O	152
5	Nopinone		C ₉ H ₁₄ O	138
6	α -ylangene		C ₁₅ H ₂₄	204
7	γ -muurolene		C ₁₅ H ₂₄	204
8	β -guaiene		C ₁₅ H ₂₄	204
9	Valencene		C ₁₅ H ₂₄	204
10	4-epi-cubebol		C ₁₅ H ₂₆ O	222
11	δ -cadinene		C ₁₅ H ₂₄	204
12	Cadinadiene		C ₁₅ H ₂₄	204
13	α -calacorene		C ₁₅ H ₂₀	200
14	α -agarofuran		C ₁₅ H ₂₄ O	220
15	Europelargone A		C ₁₅ H ₂₂ O ₂	234

Table 6 (continued)

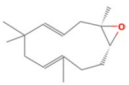
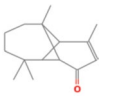
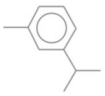
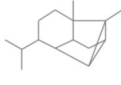
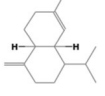
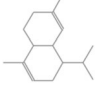
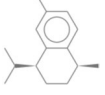
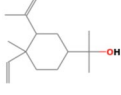
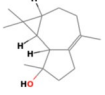
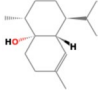
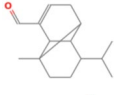
ID	Chemical name	Chemical structure	Formula	Molecular weight
16	Humulene oxide II		C ₁₅ H ₂₄ O	220
17	Longiverbenone		C ₁₅ H ₂₂ O	218

Table 7 Unique components extracted from *Phoebe bournei*

ID	Chemical name	Chemical structure	Formula	Molecular weight
1	M-cymene		C ₁₀ H ₁₄	134
2	Borneol		C ₁₀ H ₁₈ O	154
3	Cherry propanol		C ₁₀ H ₁₄ O	150
4	Cyclosativene		C ₁₅ H ₂₄	204
5	γ -amorphene		C ₁₅ H ₂₄	204
6	α -amorphene		C ₁₅ H ₂₄	204
7	Calamenene		C ₁₅ H ₂₂	202
8	α -elemol		C ₁₅ H ₂₆ O	222
9	Isospathulenol		C ₁₅ H ₂₄ O	220
10	Cubenol		C ₁₅ H ₂₆ O	222
11	Ylangenal		C ₁₅ H ₂₂ O	218
12	Elemoyl acetate		C ₁₇ H ₂₈ O ₂	264

Abbreviations

SPME	Solid phase micro-extraction
GC–MS	Gas chromatography-mass spectrometry
RI	Retention index
RF	Random Forest
KNN	K-nearest neighbors

Acknowledgements

Not applicable.

Author contributions

RC and ZG completed the experiments for the text; ZC and HW wrote the main manuscript text and prepared all figures; XX presented the idea of the text; HG revised the text.

Funding

Jiangsu Province Forestry Science and Technology Innovation and Extension Project: Research on propagation and efficient cultivation of precious green wood species of *Phoebe* (LYKJ 【2019】 46); Excellent scientific and technological innovation team of higher education in Jiangsu Province (2019–29); National College Student Innovation Training Programme (202112213019Z); Key Laboratory of State Forest and Grassland Administration on Wildlife Evidence Technology.

Availability of data and materials

The datasets generated and analyzed during the current study are available in the Figshare repository, [<https://doi.org/10.6084/m9.figshare.20410650.v1>].

Declarations

Consent for publication

Not applicable.

Competing interests

On behalf of all authors, I declare that our research has no conflict of interest.

Received: 16 November 2022 Accepted: 2 June 2023

Published online: 20 June 2023

References

- Zhang Q, Ning L, Ding W, Xing H, Zhou Q, Wei Z (2019) Chemical constituents of the essential oil extracted from *Phoebe bournei*. *Chem Nat Compd* 55:754–755. <https://doi.org/10.1007/s10600-019-02801-9>
- Pan F, Chen L, Jiang Y, Xiong L, Min L, Xie J, Qi J, Xiao H, Chen Y, De Hoop CF (2018) Bio-based UV protective films prepared with polylactic acid (PLA) and *Phoebe zhennan* extractives. *Int J Biol Macromol* 119:582–587. <https://doi.org/10.1016/j.ijbiomac.2018.07.189>
- Wang J, Ning L, Gao Q, Zhang S, Chen Q (2019) The identification and characteristics of *Phoebe zhennan* buried wood. *IAWA J* 40:804–819. <https://doi.org/10.1163/22941932-40190249>
- Xiao J-H, Ding X, Li L, Ma H, Ci X-Q, van der Merwe M, Conran JG, Li J (2020) Miocene diversification of a golden-thread nanmu tree species (*Phoebe zhennan*, Lauraceae) around the Sichuan Basin shaped by the East Asian monsoon. *Ecol Evol* 10:10543–10557
- Xie J, Qi J, Huang X, Zhou N, Hu Y (2015) Comparative analysis of modern and ancient buried *Phoebe zhennan* wood: surface color, chemical components, infrared spectroscopy, and essential oil composition. *J For Res* 26:501–507. <https://doi.org/10.1007/s11676-015-0034-z>
- Xu B, Zhu T, Li J, Liu S (2013) Identification of wood between *Phoebe zhennan* and *Machilus pingii* using the gas chromatography-mass spectrometry direct injection technique. *Eur J Mass Spectrom* (Chichester) 19:187–193. <https://doi.org/10.1255/ejms.1226>
- Duval CJ, Sok N, Laroche J, Gourrat K, Prida A, Lequin S, Chassagne D, Gougeon RD (2013) Dry vs soaked wood: modulating the volatile extractible fraction of oak wood by heat treatments. *Food Chem* 138:270–277. <https://doi.org/10.1016/j.foodchem.2012.09.117>
- Duval CJ, Gourrat K, Perre P, Prida A, Gougeon RD (2013) A HS–SPME–GC–MS analysis of IR heated wood: Impact of the water content on the depth profile of oak wood aromas extractability. *Food Res Int* 54:277–284
- Vichi S, Santini C, Natali N, Riponi C, Lopez-Tamames E, Buxaderas S (2007) Volatile and semi-volatile components of oak wood chips analysed by accelerated solvent extraction (ASE) coupled to gas chromatography–mass spectrometry (GC–MS). *Food Chem* 102:1260–1269
- He ZZ, Fan YM, Xie XM, Yuan GF, Chen ZX (2012) GC–MS analysis of fragrance constituents from offcuts of eaglewood and Chinese Eaglewood in Guangdong. In: *Advanced Materials Research*. Trans Tech Publ, pp 1904–1907
- Breiman L (2001) Random forests. *Mach Learn* 45:5–32. <https://doi.org/10.1023/A:1010933404324>
- Jiang Y, Zhang X, Zhang J, Chhin S (2022) Comparing the random forest algorithm with other modelling approaches to capture the complex patterns of intra-annual wood formation of Chinese fir with different ages. *Dendrochronologia* 77:126043. <https://doi.org/10.1016/j.dendro.2022.126043>
- Mascellani A, Hoca G, Babisz M, Krška P, Klouček P, Havlík J (2021) 1H NMR chemometric models for classification of Czech wine type and variety. *Food Chem* 339:127852
- Li Y, Ercisli S (2023) Data-efficient crop pest recognition based on KNN distance entropy. *Sustain Comput* 38:100860
- Zhu J, Sharma AS, Xu J, Xu Y, Jiao T, Ouyang Q, Li H, Chen Q (2021) Rapid on-site identification of pesticide residues in tea by one-dimensional convolutional neural network coupled with surface-enhanced Raman scattering. *Spectrochim Acta Part A Mol Biomol Spectrosc* 246:118994
- Wang F, Zhao H, Yu C, Tang J, Wu W, Yang Q (2020) Determination of the geographical origin of maize (*Zea mays* L.) using mineral element fingerprints. *J Sci Food Agric* 100:1294–1300
- Yu D, Zhang X, Guo S, Yan H, Wang J, Zhou J, Yang J, Duan J-A (2022) Headspace GC/MS and fast GC e-nose combined with chemometric analysis to identify the varieties and geographical origins of ginger (*Zingiber officinale* Roscoe). *Food Chem* 396:133672
- Cheng Q, Gao L, Zhong F, Zuo X, Ma M (2020) Spatiotemporal variations of drought in the Yunnan-Guizhou Plateau, southwest China, during 1960–2013 and their association with large-scale circulations and historical records. *Ecol Ind* 112:106041
- Zheng JM, Zhang WC, Chen Y, Ma T (2015) Analysis on climatologic characteristics and causes of the severe drought during 2009–2010 in Yunnan Province. *J Meteorol Sci* 35:488–496
- Zhijia G, Xingwu D, Bing L, Jinming H, Jiaonan H (2016) The spatial distribution and temporal variation of rainfall erosivity in the Yunnan Plateau, Southwest China: 1960–2012. *CATENA* 145:291–300
- Liu W-J, Liu C-Q, Zhao Z-Q, Xu Z-F, Liang C-S, Li L, Feng J-Y (2013) Elemental and strontium isotopic geochemistry of the soil profiles developed on limestone and sandstone in karstic terrain on Yunnan-Guizhou Plateau, China: implications for chemical weathering and parent materials. *J Asian Earth Sci* 67:138–152

Publisher's Note

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