RSC Advances

PAPER

View Article Online
View Journal | View Issue

Cite this: *RSC Adv.*, 2019, **9**, 20557

GC-MS-based identification and statistical analysis of liposoluble components in the rhizosphere soils of *Panax notoginseng*†

Yi-Jun Qiao,[ab] Jia-Jiao Zhang,[c] Jia-Huan Shang,[ab] Hong-Tao Zhu,[a] Dong Wang,[a] Chong-Ren Yang[a] and Ying-Jun Zhang *[a]

Continuous cropping obstacle, mainly caused by microorganisms and organic components in soil, has become a serious problem for the plantation of *Panax notoginseng* (Araliaceae) due to the rapidly increased demands of this famous herbal medicine in recent decades. The rhizosphere soils cultivated with 3-year-old healthy and ill notoginseng were chemically investigated by gas chromatography-mass spectrometry (GC-MS) and compared with the corresponding soils without the plantation of notoginseng. Totally 47 liposoluble components were identified. Furthermore, the multiple statistical analysis showed that these constituents were qualitatively and quantitatively associated with the differences between the cultivated soil with *P. notoginseng* and the uncultivated soil. Among them, neophytadiene (**4**), D-α-tocopherol (**38**), (3β,22*E*,24*S*)-ergosta-5,22-dien-3-ol (**39**), (3β,24*R*)-ergost-5-en-3-ol (**40**), stigmasta-5,22-dien-3-ol (**41**), stigmast-4-en-3-one (**44**) and (5α)-stigmastane-3,6-dione (**47**) contributed most to the significant differences between the cultivated and uncultivated soils, whereas cyclopentadecane (**3**), octadecanoic acid methyl ester (**16**), docosanoic acid ethyl ester (**31**), nonacosane (**34**), **38** and **39** were found in much higher amount in the soils with ill *P. notoginseng* as compared to the case of those with the healthy *P. notoginseng*. On the other hand, liposoluble components in different cultivation areas were of great diversity; however, they were able to remain relatively consistent across the overall trend of differential substances.

Received 19th March 2019
Accepted 14th June 2019

DOI: 10.1039/c9ra02110h

rsc.li/rsc-advances

Introduction

Continuous cropping obstacle (CCO), mainly caused by the changes in the microbial community and organic components in the cultivated soil, is a common problem in the cultivation of many crops and herbs. *Panax notoginseng* (Burk.) F. H. Chen (Araliaceae), a famous herbal medicine in the Ginseng family with various pharmaceutical and health beneficial effects, has been historically cultivated in Wenshan district of Yunnan province and its adjacent areas of Guangxi province, China, for over 400 years.[1–4] Due to the rapidly increased demands and planting scales in recent decades, CCO has become a serious problem for notoginseng cultivation, resulting in the lack of suitable land plots for the plantation of *P. notoginseng*.[5]

Previous studies have shown that allelochemicals contribute significantly to the CCO problem.[6] Phenolic acids and organic acids in root exudates could accelerate the formation of CCO. Benzoic acid, 2,2-di(4-hydroxyphenyl)propane and palmitic acid showed stronger allelopathic effects on the radicle or hypocotyl growth of the *P. ginseng* seeds.[7] Ferulic acid, total saponin and root extract of *P. notoginseng* could inhibit the growth of the plant itself.[8] Moreover, the seed germination rate of *P. notoginseng* was reduced by the degradation of its leaf residues in the soil.[9] The soil extracts of notoginseng also displayed different allelopathic effects on its root length, seedling height, fresh weight and nitrate reductase activity, as well as on radish and lettuce seedlings.[10–12] Recently, 29 liposoluble components have also been identified from the rhizosphere soil of continuously cultivated *P. notoginseng*; among these, *p*-hydroxybenzoic acid and phthalic acid diisobutylester have an allelopathic effect on the growth of the *P. notoginseng* pathogen.[13] Another study has reported 26 volatiles originated from the cultivated soil of *P. notoginseng*, with high content of allelochemicals such as diisobutyl phthalate.[14]

[a]*State Key Laboratory of Phytochemistry and Plant Resources of West China, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650201, People's Republic of China. E-mail: zhangyj@mail.kib.ac.cn*

[b]*University of Chinese Academy of Sciences, Beijing 100049, People's Republic of China*

[c]*State Key Laboratory of Hybrid Rice, College of Life Sciences, Wuhan University, Wuhan 430072, People's Republic of China*

† Electronic supplementary information (ESI) available: GC-MS total ion chromatograms (TIC) of the soil sample L-3I, relative peak area percentage of total (%) for 47 identified compounds in the soil samples, the sources of identified compounds in the collected soils of *P. notoginseng*, contribution of variances identified in the cultivated and uncultivated soils and compounds responsible for separation of cultivated and uncultivated soils. See DOI: 10.1039/c9ra02110h

As a part of our research to reveal the formation mechanism of CCO for *P. notoginseng*, in this study, the liposoluble components in the rhizosphere soils cultivated with 3-year-old healthy or ill notoginseng obtained from five different plantations were studied by GC-MS analysis, combining with multiple statistical analysis. The innovation of this research is that for the first time, the combination method of instrumental analysis and multivariate statistical analysis was used to study the liposoluble components in different soil samples of *Panax notoginseng*, by which the different key substances were more accurately and efficiently identified.

Results and discussion

Herein, ten rhizosphere soil samples obtained from five different plantations [A-San-Long (A), Lao-Mu-Shao (L), Ba-Tang-Chong (W-1), Zhai-Tou (W-2), and Ba-Zi (W-3)] in Yunnan province, China, each with one healthy (H) and one illness (I) growing status of 3-year-old *P. notoginseng*, were obtained for study; moreover, one soil sample (M) without the cultivation of notoginseng was obtained from the adjacent field of each plantation.

A total of 47 different liposoluble components with more than 80% matching value were identified from the soil samples based on the GC-MS analysis, and the results are listed in Table 1. The identified components were of 10 types: 15 esters (**6**, **8**, **10–13**, **16–20**, **25**, **29**, **31**, and **32**), 9 steroids (**35–37**, **39–41**, **43**, **44**, and **47**), 7 alkanes (**1**, **3**, **15**, **22**, **24**, **33**, and **34**), 4 benzene derivatives (**2**, **9**, **30**, and **38**), 3 alkones (**5**, **26**, and **28**), 3 alkenes (**4**, **7**, and **27**), 3 terpenoids (**42**, **45**, and **46**), 1 phenanthrene (**21**), 1 alcohol (**14**) and 1 aldehyde (**23**). The relative peak area percentage of identified compounds and their distributions normalized based on the total ion current (TIC) in the soil samples are provided in Tables S1 and S2,† respectively.

The 10 types of liposoluble components were detected from all five plantations, with a slight difference in proportions, whereas a noticeable difference in contents. Esters, steroids and alkanes composed more than half of the total components (Fig. 1), with 31.81% of esters ranking the most. Compared to the case of the soils without the cultivation of *P. notoginseng*, the steroids **39–41** and **47** were present in higher contents in the cultivated soils. These results are slightly different from those obtained in a previous study, in which alkanes, alkenes, aldehydes, organic acids, esters and acetylenic alcohols have been reported as major liposoluble components in the cultivated soil.[16] Note that some dominant components, such as cyclopentadecane (**3**), nonacosane (**34**), D-α-tocopherol (**38**), (3β,22*E*,24*S*)-ergosta-5,22-dien-3-ol (**39**), stigmasta-5,22-dien-3-ol (**41**), and (5α)-stigmastane-3,6-dione (**47**), present in the cultivated soils were not detected in the uncultivated soils.

The GC-MS results were analysed by two different algorithms: the PCA and the PLS-DA models. As shown in Fig. 2 and 3, the soil samples obtained from three different plantations (W-1, W-2 and W-3) in Wenshan county were nearer to each other than the other two sites (A and L). Moreover, it was found that only in Wenshan plantations (W-1, W-2 and W-3), the concentrations of some components, such as octadecanoic acid methyl ester (**16**), cycloeicosane (**22**) and (3β,22*E*,24*S*)-ergosta-5,22-dien-3-ol (**39**), were changed by the growing status and the cultivation of *P. notoginseng*. However, this was not observed in the soil samples obtained from the other two regions (Tables 2 and 3).

Moreover, as shown in the PCA score plot (Fig. 2a), the clusters of the cultivated soils were separated from those of the uncultivated soils by combining PC1 (23.3%) and PC2 (17.8%), and the uncultivated soils located relatively closer. Among the samples W-1, W-2, and W-3 obtained from Wenshan county, the cultivated and uncultivated soils were separated from each other by PC1. By combining these results with those of the loading plot (Fig. 2b), the top ten variables with greatest impact on the classification were found (Table S3†) to be octadecanoic acid methyl ester (**16**), cycloeicosane (**22**), stigmasta-5,22-dien-3-ol (**40**), 15-methyl-heptadecanoic acid ethyl ester (**17**), 1-octadecanol (**14**), dibutyl phthalate (**9**), (3β,24*R*)-ergost-5-en-3-ol (**41**), 4,8,12,16-tetra-methylheptadecan-4-olide (**26**), bis(2-ethylhexyl)phthalate (**30**) and heneicosane (**15**). Among these, the contents of the compounds **40** and **41** in the cultivated soils (both healthy and ill samples) were more than 2-fold higher than those in the uncultivated soils.

However, in the other two plantations A and L, the cultivated soils were separated from the uncultivated soils by PC2, and the top ten variables responsible for the discrimination were found in the loading plot (Fig. 2b): 2-pentadecanone (**5**), 6,10,14-trimethyl-cholesterol (**37**), 10-methylnonadecane (**1**), neophytadiene isomer I (**7**), (3β,22*E*,24*S*)-ergosta-5,22-dien-3-ol (**39**), phenol 2,4-bis(1,1-dimethyl-ethyl)- (**2**), (5α)-stigmastane-3,6-dione (**47**), stigmast-4-en-3-one (**44**), neophytadiene (**4**) and docosanoic acid methyl ester (**29**). Overall, in the cultivated soils, the contents of the compounds **44** and **47** in the soil obtained from the A-San-Long plantation (A) as well as those of **2**, **4**, **5**, **7**, **37**, **39**, **44** and **47** in the soil obtained from the Lao-Mu-Shao plantation (L) were found to increase when compared with the case of their corresponding uncultivated samples.

The PLS-DA analysis is the most commonly used classification method in metabolomics data analysis and has been conducted to investigate the discriminatory components. It combines the regression models with dimensionality reduction and discriminates the regression results with certain discriminant thresholds. As shown in Fig. 3a, the cultivated and the uncultivated soils were clearly separated by PC1 and PC2 in the score plot. Based on the VIP value (Table S4†) and the loading plot (Fig. 3b), 20 compounds, including esters, steroids, alkanes, aldehyde, the derivative of benzene and terpenoid, were found to play key roles in the classification.

Moreover, one-way ANOVA was applied to investigate lipids in different soil samples, and the lipid levels changed significantly in the cultivated soils as compared to those of the uncultivated soils. Variables with $p < 0.05$ were considered of remarkable difference between every two groups (Tables 2 and 4). After planting *P. notoginseng* for 3 years, it was observed that the contents of neophytadiene (**4**), D-α-tocopherol (**38**), (3β,22*E*,24*S*)-ergosta-5,22-dien-3-ol (**39**), (3β,24*R*)-ergost-5-en-3-ol (**40**), stigmasta-5,22-dien-3-ol (**41**), stigmast-4-en-3-one (**44**), and (5α)-stigmastane-3,6-dione (**47**) increased significantly in

Table 1 The list of the identified compounds in the obtained soils of *P. notoginseng*[a]

Peaks	Compounds	Mol. formula	RT (min)	Qualifier ions (*m*/*z*)	Classification
1	10-Methylnonadecane	$C_{20}H_{64}$	27.920	57, 71, 85	Alkane
2	Phenol, 2,4-bis(1,1-dimethylethyl)-	$C_{14}H_{22}O$	28.448	191, 206, 57	Benzene derivative
3	Cyclopentadecane	$C_{15}H_{30}$	32.995	149, 223, 104	Alkane
4	Neophytadiene	$C_{20}H_{38}$	35.618	95, 68, 82	Alkene
5	2-Pentadecanone, 6,10,14-trimethyl-	$C_{18}H_{36}O$	35.751	58, 71, 85	Alkone
6	1,2-Benzenedicarboxylic acid, bis(2-methylpropyl)ester	$C_{16}H_{22}O_4$	36.275	149, 57, 223	Ester
7	Neophytadiene, isomer I	$C_{20}H_{38}$	36.494	81, 82, 95	Alkene
8	Pentadecanoic acid, ethyl ester	$C_{17}H_{34}O_2$	36.724	88, 101, 157	Ester
9	Dibutyl phthalate	$C_{16}H_{22}O_{25}$	38.145	149, 223, 104	Benzene derivative
10	Ethyl 9-hexadecenoate	$C_{18}H_{34}O_2$	38.294	55, 69, 88	Ester
11	*E*-11-Hexadecenoic acid, ethyl ester	$C_{18}H_{34}O_3$	38.487	55, 69, 88	Ester
12	Hexadecanoic acid, ethyl ester	$C_{18}H_{36}O_2$	38.700	88, 101, 55	Ester
13	Heptadecanoic acid, ethyl ester	$C_{19}H_{38}O_2$	39.892	88, 101, 57	Ester
14	1-Octadecanol	$C_{18}H_{38}O$	40.431	82, 57, 96	Alcohol
15	Heneicosane	$C_{21}H_{44}$	40.661	57, 71, 85	Alkane
16	Octadecanoic acid, methyl ester	$C_{19}H_{38}O_2$	41.104	74, 87, 143	Ester
17	Heptadecanoic acid, 15-methyl-, ethyl ester	$C_{20}H_{40}O_2$	41.350	88, 101, 57	Ester
18	Linoleic acid, ethyl ester	$C_{20}H_{36}O_2$	41.804	81, 67, 55	Ester
19	Ethyl oleate	$C_{20}H_{38}O_2$	41.911	55, 69, 41	Ester
20	Octadecanoic acid, ethyl ester	$C_{20}H_{40}O_2$	42.370	88, 101, 312	Ester
21	Phenanthrene, 1-methyl-7-(1-methylethyl)-	$C_{18}H_{18}$	42.931	219, 234, 204	Phenanthrene
22	Cycloeicosane	$C_{20}H_{40}$	43.888	55, 69, 83	Alkane
23	Tetradecanal	$C_{14}H_{28}O$	44.032	57, 82, 96	Aldehyde
24	Tricosane	$C_{23}H_{48}$	44.165	57, 71, 85	Alkane
25	2-Propenoic acid, 3-(4-methoxyphenyl)-, 2-ethylhexyl ester	$C_{18}H_{26}O_3$	44.598	178, 161, 133	Ester
26	4,8,12,16-Tetramethylheptadecan-4-olide	$C_{21}H_{40}O_2$	45.122	99, 43, 55	Alkone
27	1,21-Docosadiene	$C_{22}H_{42}$	46.313	95, 68, 55	Alkene
28	1-Docosanal	$C_{22}H_{44}O$	47.344	82, 97, 111	Alkone
29	Docosanoic acid, methyl ester	$C_{23}H_{46}O_2$	47.916	41, 96, 82	Ester
30	Bis(2-ethylhexyl)phthalate	$C_{24}H_{38}O_4$	48.209	149, 167, 279	Benzene derivative
31	Docosanoic acid, ethyl ester	$C_{24}H_{48}O_2$	48.904	88, 97, 157	Ester
32	Tricosanoic acid, methyl ester	$C_{24}H_{48}O_2$	49.457	88, 101, 55	Ester
33	Heptacosane	$C_{27}H_{56}$	50.416	57, 71, 85	Alkane
34	Nonacosane	$C_{29}H_{60}$	53.231	57, 71, 85	Alkane
35	Stigmasta-3,5,22-trien	$C_{29}H_{46}$	56.212	394, 145, 69	Steroid
36	Stigmasta-3,5-dien	$C_{29}H_{48}$	56.602	396, 147, 81	Steroid
37	Cholesterol	$C_{27}H_{46}O$	57.094	386, 275, 107	Steroid
38	D-α-Tocopherol	$C_{29}H_{50}O_2$	57.382	430, 165, 69	Benzene derivative
39	(3β,22*E*,24*S*)-Ergosta-5,22-dien-3-ol	$C_{28}H_{46}O$	58.093	398, 69, 207	Steroid
40	(3β,24*R*)-Ergost-5-en-3-ol	$C_{28}H_{48}O$	59.487	81, 95, 400	Steroid
41	Stigmasta-5,22-dien-3-ol	$C_{29}H_{48}O$	60.374	55, 271, 412	Steroid
42	Hop-22(29)-en-3β-ol	$C_{30}H_{50}O$	61.517	189, 205, 175	Terpenoid
43	γ-Sitosterol	$C_{29}H_{50}O$	61.880	414, 329, 303	Steroid
44	Stigmast-4-en-3-one	$C_{29}H_{48}O$	66.095	124, 55, 229	Steroid
45	(3α)-D:A-Friedooleanan-3-ol	$C_{36}H_{62}O_2$	67.917	95, 169, 65	Terpenoid
46	Friedelin	$C_{30}H_{50}O$	68.830	69, 109, 95	Terpenoid
47	(5α)-Stigmastane-3,6-dione	$C_{29}H_{48}O_2$	72.816	428, 245, 207	Steroid

[a] The identification criteria based on mass spectra matching with the Wiley7n.l library and NIST14 library which resulted in level 2 identifications.[15]

the soils regardless of the growing status (healthy or ill) of *P. notoginseng*.

Comparison of the liposoluble components in the cultivated soils with different growing statuses

Usually, *P. notoginseng* grows for 3 years before harvest. During the planting process, some of the plants might get affected by diseases, resulting in changes in their chemical compositions in the rhizosphere soils. The differences between the liposoluble components of soils with different growing statuses (healthy or ill) of *P. notoginseng* were analyzed by one-way ANOVA. As shown in Table 3, the concentrations of six liposoluble components, *i.e.* cyclopentadecane (**3**), octadecanoic acid methyl ester (**16**), docosanoic acid ethyl ester (**31**), nonacosane (**34**), D-α-tocopherol (**38**) and (3β,22*E*,24*S*)-ergosta-5,22-dien-3-ol (**39**), were found to be changed significantly and much higher in the soils with ill notoginseng. Although other compounds with values of VIP > 1 did not change consistently,

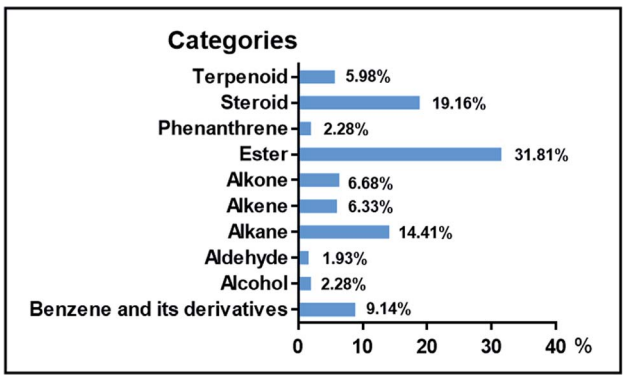

Fig. 1 Categories of the liposoluble components identified from the soil samples by GC-MS analysis.

they also contributed to the three clusters of variances in the PLS-DA analysis (Fig. 3a).

These compounds may be produced by the pathogenic microorganisms after the plants get infested by diseases. Since the differences between the cultivation areas and plant statuses are not obvious and the changes in the contents of these compounds are basically consistent in the same planting area, attention should be paid mainly to the differences between the liposoluble components of the cultivated soils with the healthy status of plant and the uncultivated soils.

Origins and biosynthetic pathways of liposoluble compounds identified in the soils

The liposoluble components that could significantly affect the classification in the PCA and PLS-DA analysis were of different types. Among them, alkanes, esters, steroids and terpenoids account for more than half of all the compounds identified in the cultivated soils.

The origins of alkanes in soils and sediments are very complex. The input paths mainly include direct input of mineral oil, sedimentation of atmospheric particulate matter, input through an aqueous medium, industrial solid, municipal and domestic waste, as well as biochemical degradation products of natural organics. There are some differences in the composition of hydrocarbon pollutants obtained from different sources. Based on these differences, the sources of hydrocarbon pollutants in the environment can be identified.[17] Unlike the phthalate esters that mainly originate from chemical fertilizers and pesticides,[18] the esters identified in this study are fatty acid esters. They are more likely to originate from industrial chemicals, which may be used as solvents in skin care products. The steroids can be isolated, extracted and purified from natural resources. *In vivo*, acetic acid can form squalene by head-to-head contact of farnesyl pyrophosphate under the action of an enzyme. The cyclization of the 2,3-epoxide of squalene produces lanosterol. Lanosterol undergoes a series of transformations in the body to form steroidal substances such as cholesterol and hormones. Terpenoids widely exist in advanced plants in the form of volatile oils. They can be divided into monoterpene, diterpene, sesquiterpenes, triterpene and polyterpenoid according to their structures. As the main type of allelochemicals, monoterpene and sesquiterpene possess strong bioactivities. After synthesis, most of the terpenoids enter the soil through plant volatilization or root secretion, which can affect the growth and development of the host plant and the neighboring plants.[19,20]

Experimental

Materials and methods

Soil samples collection. Using a five-point sampling method, the soil samples were obtained in November 2015, from five plantations while harvesting the 3-year-old *P. notoginseng* in Yunnan province, China (Table 5). According to the sample collecting areas and growth status of *P. notoginseng*, these samples were divided into healthy (H) and ill (I) groups. The soil sample without the cultivation of notoginseng was obtained from the adjacent fields at each site as the uncultivated soil group (M). After removing impurities, each soil sample was

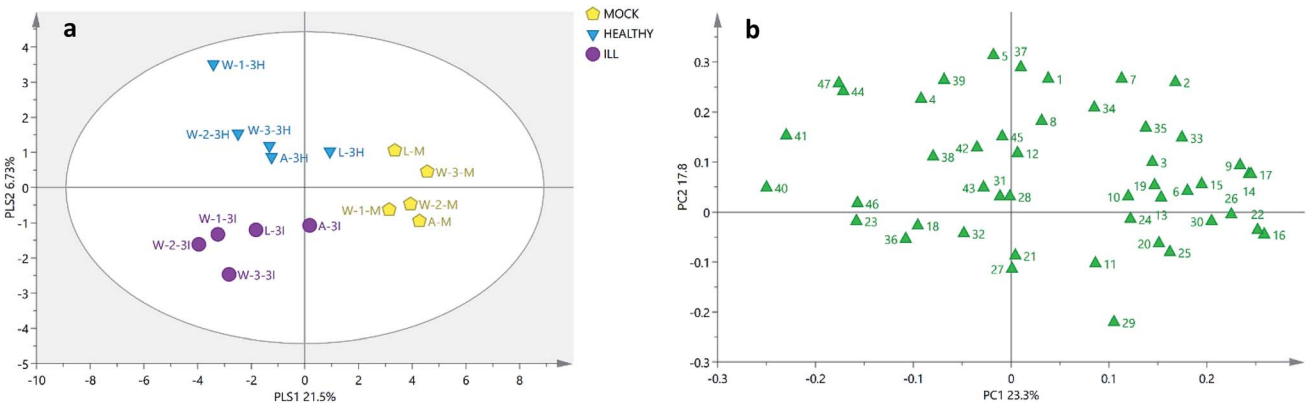

Fig. 2 PCA score plots and loading plots of 15 soil samples. (a) PCA score plot of the soil samples. (b) PCA loading plot of the soil samples. A, L and W: the sites of sample collection (A: A-San-Long; L: Lao-Mu-Shao; W: Wenshan); three sites for sample collection in Wenshan county were numbered as W-1, W-2 and W-3.; M: mock group, used as a blank control. I and H: the growing status of 3-year-old *P. notoginseng* (I: ill, H: healthy); the numbers in the loading plot of all the soil samples correspond to the values provided in Table 1.

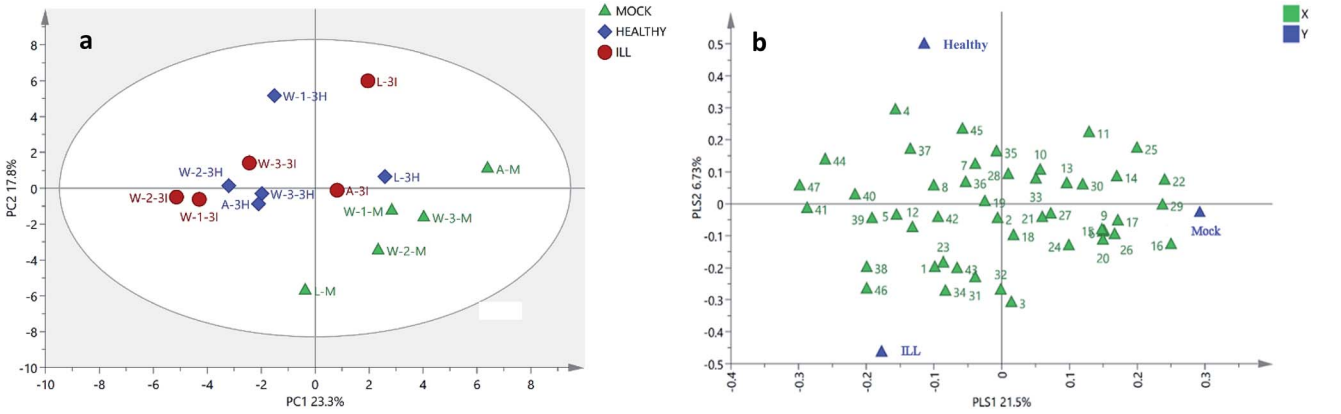

Fig. 3 PLS-DA score plots and loading plots of 15 soil samples. (a) PLS-DA score plot of the soil samples. (b) PLS-DA loading plot of the soil samples. A, L and W: the sites of sample collection (A: A-San-Long; L: Lao-Mu-Shao; W: Wenshan); three sites for sample collection in Wenshan county were numbered as W-1, W-2 and W-3.; M: mock group, used as a blank control. I and H: the growing status of 3-year-old *P. notoginseng* (I: ill, H: healthy); and the numbers in the loading plot of all the soil samples correspond to the values presented in Table 1.

comminuted and sieved through a 40-mesh screen and stored in a refrigerator at −80 °C before pre-treatment.

Sample preparation. The soil sample (each of 5.0 kg) was separately soaked overnight in methanol at room temperature, and this soaking process was repeated thrice. The yielded methanol extract was fractionated with *n*-hexane. After removing the organic solvent by rotary evaporation under reduced pressure at 45 °C, the *n*-hexane extract was obtained and re-dissolved in 1 mL *n*-hexane to prepare a 20 mg mL^{-1} solution for the subsequent GC-MS analysis. Moreover, three samples were prepared in parallel for each sample.

Table 2 Comparison between the liposoluble components of the uncultivated and cultivated soils with ill *P. notoginseng*

Components	AM/AI	LM/LI	W-1M/W-1I	W-2M/W-2I	W-3M/W-3I
Benzene derivative					
D-α-Tocopherol (**38**)	0.012↑	0.002↑	0.087	0.003↑	0.001↑
Aldehyde					
Tetradecanal (**23**)	1.000	0.827	0.001↑	0.001↑	0.012↓
Alkanes					
10-Methylnonadecane (**1**)	0.951	0.001↑	0.045↑	1.000	0.047↑
Cyclopentadecane (**3**)	0.058	0.013↑	0.605	0.002↓	0.059
Neophytadiene (**4**)	0.223	0.003↑	0.524	0.008↑	0.067
Cycloeicosane (**22**)	0.037↓	0.210	0.006↓	0.001↓	0.001↓
Nonacosane (**34**)	0.894	0.007↑	0.392	0.373	0.002↑
Esters					
E-11-Hexadecenoic acid, ethyl ester (**11**)	0.240	0.964	0.004↓	0.008↓	0.135
Octadecanoic acid, methyl ester (**16**)	0.172	0.001↑	0.001↓	0.002↓	0.018↓
2-Propenoic acid, 3-(4-methoxy-phenyl)-, 2-ethylhexyl ester (**25**)	0.004↓	0.017↓	0.003↓	0.045↓	1.000
Docosanoic acid, methyl ester (**29**)	0.002↓	0.002↓	0.140	0.000↓	0.625
Docosanoic acid, ethyl ester (**31**)	0.143	0.006↓	0.091	0.001↑	0.002↑
Tricosanoic acid, methyl ester (**32**)	0.200	0.001↑	0.392	0.005↓	0.002↑
Steroids					
(3β,22*E*,24*S*)-Ergosta-5,22-dien-3-ol (**39**)	0.001↓	0.001↑	0.002↑	0.004↑	0.001↑
(3β,24*R*)-Ergost-5-en-3-ol (**40**)	1.000	0.023↑	0.001↑	0.001↑	0.002↑
Stigmasta-5,22-dien-3-ol (**41**)	0.015↑	0.001↑	0.001↑	0.005↑	0.002↑
Stigmast-4-en-3-one (**44**)	0.003↑	0.001↑	0.058	0.008↑	0.001↑
(5α)-Stigmastane-3,6-dione (**47**)	0.125	0.002↑	0.010↑	0.002↑	0.001↑
Terpenoids					
(3α)-D:A-Friedooleanan-3-ol (**45**)	0.003↑	0.005↑	0.005↓	1.000	0.189
Friedelin (**46**)	0.012↑	0.048↑	0.226	0.001↑	0.001↑

Table 3 Comparison between the liposoluble components of the cultivated soils with healthy and ill *P. notoginseng*

Metabolites	AH/AI	LH/LI	W-1H/W-1I	W-2H/W-2I	W-3H/W-3I
Benzene derivatives					
D-α-Tocopherol (**38**)	0.012↑	0.234	0.003↑	0.160	0.474
Aldehyde					
Tetradecanal (**23**)	0.005↓	0.003↑	0.150	0.001↑	0.044↓
Alkane					
10-Methylnonadecane (**1**)	0.045↑	0.001↑	0.001↓	1.000	0.164
Cyclopentadecane (**3**)	0.454	0.028↑	0.007↑	1.000	0.001↑
Neophytadiene (**4**)	0.002↓	0.064	0.008↓	0.212	0.001↓
Cycloeicosane (**22**)	0.026↑	0.001↓	0.002↓	0.062	0.809
Nonacosane (**34**)	0.026↑	0.021↑	0.001↓	0.004↑	0.003↑
Ester					
E-11-Hexadecenoic acid, ethyl ester (**11**)	0.001↓	0.008↓	1.000	0.212	0.003↓
Octadecanoic acid, methyl ester (**16**)	0.011↑	0.028↑	0.063	1.000	1.000
2-Propenoic acid, 3-(4-methoxy-phenyl)-, 2-ethylhexyl ester (**25**)	0.840	0.045↑	0.160	0.042↓	0.001↓
Docosanoic acid, methyl ester (**29**)	0.953	0.609	0.004↑	0.003↓	0.988
Docosanoic acid, ethyl ester (**31**)	0.007↑	0.307	0.006↑	0.074	0.695
Tricosanoic acid, methyl ester (**32**)	0.016↓	0.006↑	0.001↑	0.715	0.002
Steroid					
(3β,22*E*,24*S*)-Ergosta-5,22-dien-3-ol (**39**)	0.205	0.093	0.986	0.175	0.010↑
(3β,24*R*)-Ergost-5-en-3-ol (**40**)	0.012↓	0.048↑	0.013↑	0.789	0.929
Stigmasta-5,22-dien-3-ol (**41**)	0.014↓	0.005↑	0.010↑	0.922	0.999
Stigmast-4-en-3-one (**44**)	0.449	0.119	0.005↓	0.128	0.015↑
(5α)-Stigmastane-3,6-dione (**47**)	0.995	0.027↑	0.037↓	0.339	0.208
Terpenoid					
(3α)-D:A-Friedooleanan-3-ol (**45**)	0.002↑	0.270	0.001↓	0.003↓	0.090
Friedelin (**46**)	0.704	0.046↑	0.001↑	0.185	0.409

GC-MS conditions. Identification of liposoluble components in soil samples was carried out using a GC-MS system. The GC-MS analysis was performed *via* Agilent HP 6890 GC equipped with Agilent 5973 MS (Agilent Technologies) using a 5 : 1 split injection ratio. The ionization mode was electron impact at 70 eV. An HP-5MS capillary column (0.25 mm × 30 m, 0.25 μm) with helium as a carrier gas at 1.0 mL min^{-1} was used to analyze the samples. The injected volume was 2.0 μL for each sample. The oven temperature was initially held at 40 °C, firstly ramped to 80 °C at 3 °C min^{-1}, and then to the target temperature of 280 °C at 5 °C min^{-1} with the duration of 30 min. The temperatures of injector, the ion source and the quadrupole were maintained at 250 °C, 230 °C and 150 °C, respectively. The obtained mass range *m*/*z* 35–500 was acquired using the full scan monitoring mode. The solvent delay time was set at 2.4 min.[21–26]

GC-MS data acquisition. Data acquisition was carried out by the ChemStation software (Agilent Technologies). The peak area in a GC-MS chromatogram was automatically integrated and corrected through the ChemStation software. Peaks with area lower than 100 000 were rejected. Peak width was set at 0.1 s, and the threshold was set at 14.0. The compounds were identified by searching NIST98 and Wiley7n.l library, with the assistance of their qualifier ions. Further, to minimize the number of missing values, compounds with less than 80% matching value were discarded from the total peak list. Peak alignment was performed by manually comparing the retention times with the values present in a reference chromatogram, which had the most of the peaks among all samples. In the end, the relative content (%) of each compound in a sample was normalized based on total ion current (TIC) and subjected to further statistical analysis.[27–29]

Statistical analysis. Multivariate statistical analysis was carried out by SIMCA software (version 14.1, Umetrics AB, Umea, Sweden) and SPSS Statistics software (version 20, SPSS Inc., Chicago, IL, USA). All variables were Pareto scaled prior to principal component analysis (PCA) and partial least squares-discriminant analysis (PLS-DA) for classification. The null value was replaced by 1/2 of the minimum value of all analysis variances. PCA is an unsupervised mode of recognition, which relatively reflects the overall situation more objectively,[14,30] whereas PLS-DA is the identification of the supervised mode, and the data is processed in advance by artificial grouping.[31] At first, the unsupervised PCA was employed to describe an overview of all the soil samples. Furthermore, PLS-DA was applied to pick out variables with VIP (Variable Importance in the Projection) > 1. Next, we analyzed each liposoluble component with the value of VIP > 1 using one-way ANOVA (analysis of variance) with the Fisher's LSD (least significant difference) post hoc test to reveal the differences that existed in various states of the soil

Table 4 Comparison between the 0 of the uncultivated and cultivated soils with healthy *P. notoginseng*[a]

Metabolites	AM/AH	LM/LH	W-1M/W-1H	W-2M/W-2H	W-3M/W-3H
Benzene derivatives					
D-α-Tocopherol (**38**)	1.000	0.006↑	0.001	0.001↑	0.002↑
Aldehyde					
Tetradecanal (**23**)	0.005↑	0.003	0.006↑	0.872	0.357
Alkane					
10-Methylnonadecane (**1**)	0.007	0.426	0.023↑	1.000	0.396
Cyclopentadecane (**3**)	0.177	0.015↑	0.013	0.002	0.004
Neophytadiene (**4**)	0.001↑	0.049↑	0.004↑	0.002↑	0.006↑
Cycloeicosane (**22**)	0.029	0.002↑	0.235	0.009	0.001
Nonacosane (**34**)	0.031	0.003↑	0.002↑	0.002	0.630
Ester					
E-11-Hexadecenoic acid, ethyl ester (**11**)	0.002↑	0.007↑	0.004	0.002	0.025↑
Octadecanoic acid, methyl ester (**16**)	0.002	0.011↑	0.011	0.002	0.018
2-Propenoic acid, 3-(4-methoxy-phenyl)-, 2-ethylhexyl ester (**25**)	0.003	0.001	0.023	0.202	0.001↑
Docosanoic acid, methyl ester (**29**)	0.002	0.003	0.001	0.049	0.615
Docosanoic acid, ethyl ester (**31**)	0.001	0.023	0.001	0.006↑	0.003↑
Tricosanoic acid, methyl ester (**32**)	0.023↑	0.135	0.002	0.004	1.000
Steroid					
(3β,22*E*,24*S*)-Ergosta-5,22-dien-3-ol (**39**)	0.003	0.005↑	0.002↑	0.001↑	0.043↑
(3β,24*R*)-Ergost-5-en-3-ol (**40**)	0.011↑	0.587	0.029↑	0.002↑	0.002↑
Stigmasta-5,22-dien-3-ol (**41**)	0.001↑	0.140	0.023↑	0.005↑	0.002↑
Stigmast-4-en-3-one (**44**)	0.001↑	0.004↑	0.001↑	0.001↑	0.060
(5α)-Stigmastane-3,6-dione (**47**)	0.126	0.060	0.001↑	0.001↑	0.002↑
Terpenoid					
(3α)-D:A-Friedooleanan-3-ol (**45**)	0.701	0.020↑	0.079	0.003↑	0.607
Friedelin (**46**)	0.008↑	0.980	0.002↓	0.004↑	0.002↑

[a] The numbers are the *p*-value calculated by ANOVA with Fisher's LSD; ↑, significant increase compared with control ($p < 0.05$); ↓, significant decrease compared with control ($p < 0.05$); A, L and W: the sites of sample collection; A: A-San-Long; L: Lao-Mu-Shao; W: Wenshan; three sites in Wenshan county were numbered W-1, W-2 and W-3, resp.; M: mock group, used as blank control; I and H: the growing status of *P. notoginseng*; I: ill, H: healthy.

Table 5 The locations and sources of the soil sample used for the study[a]

No.	Longitude	Latitude	Altitude (*m*)	Sources
A-M	103°37′40.19″	23°47′59.34″	1466	Uncultivated
A-3H	103°37′37.91″	23°47′57.34″	1462	3H
A-3I	103°37′39.89″	23°47′57.41″	1465	3I
L-M	103°26′13.57″	24°54′34.12″	1968	Uncultivated
L-3H	103°26′14.84″	24°54′36.59″	1968	3H
L-3I	103°26′16.15″	24°54′34.67″	1968	3I
W-1-M	104°17′22″	23°29′53″	1598	Uncultivated
W-1-3H	104°17′22″	23°29′54″	1601	3H
W-1-3I	104°17′21″	23°29′55″	1597	3I
W-2-M	104°18′16″	23°29′47″	1522	Uncultivated
W-2-3H	104°18′18″	23°29′47″	1521	3H
W-2-3I	104°18′19″	23°29′46″	1521	3I
W-3-M	104°18′18″	23°30′11″	1517	Uncultivated
W-3-3H	104°18′22″	23°29′56″	1519	3H
W-3-3I	104°18′19″	23°30′00″	1522	3I

[a] All the samples were collected in Yunnan Province, China; the samples (each 5.0 kg) were obtained from five separate plantations at A-San-Long village of Honghe county (A), Lao-Mu-Shao village of Shilin county (L), and Ba-Tang-Chong (W-1), Zhai-Tou (W-2), and Ba-Zi (W-3) in Bai-Yun-Yan village of Wenshan county, resp.; 3H: 3-year-old healthy *P. notoginseng*; 3I: 3-year-old ill *P. notoginseng*.

samples. The variables with significant differences ($p < 0.05$) between the control and compared groups were marked ↑ or ↓.[32]

Conclusions

In conclusion, the liposoluble components in the soils obtained from 5 different plantations and the growing statuses of *P. notoginseng* were analyzed and compared with the case of the uncultivated soils by GC-MS analysis combined with the multivariate statistical analysis (PCA and PLS-DA). In total, 47 liposoluble components belonging to 10 different types were identified and put together for comprehensive analysis. Alkanes, *i.e.* cyclopentadecane (**3**) and 10-methylnonadecane (**1**), alkene, *i.e.* neophytadiene (**4**), esters, *i.e.* *E*-11-hexadecenoic acid ethyl ester (**11**) and octadecanoic acid methyl ester (**16**), benzene derivatives, *i.e.* D-α-tocopherol (**38**), and steroids, *i.e.* (3β,22*E*,24*S*)-ergosta-5,22-dien-3-ol (**39**) and (3β,24*R*)-ergost-5-en-3-ol (**40**), were revealed as discriminatory components by statistical analysis, which contributed significantly to the discrimination of the cultivated and uncultivated soils. Moreover, a comparison of liposoluble components in the soils

obtained from different planting areas and plant growing statuses were carried out, resulting in the demonstration of the components that were of notable differences.

Studies on the chemical substances in rhizosphere soil of *P. notoginseng* were mainly focused on the water-soluble components and the bioactivities of different solvent extracts.[33,34] Liposoluble components in the cultivated soil of *P. notoginseng* have been identified and reported in only two studies in recent years, and only some of the common allelochemicals, which have been found to exert an allelopathic effect on other plants, have been considered in their studies.[13,14] The differences between the cultivated and uncultivated soils had not been discussed to date. Therefore, by decently and statistically comparing the liposoluble components of different soils, this study provides a reference value for the subsequent research on the changes of the chemical composition of the soil and the key chemical substances leading to continuous cropping obstacle of *P. notoginseng*. Furthermore, by analyzing and discussing the possible origins and transformation pathways of these compounds, this study explains their potential role as allelochemicals and provides a theoretical basis for the further study of their allelopathic effects on *P. notoginseng* and solutions to solve the problem of continuous cropping obstacle.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

The work was supported by the Major Science and Technique Programs in Yunnan Province (2016ZF001-001), the Science and Technology Planning Project of Yunnan Province (2013FC008) and the Yung-Chi Cheng Academician Workstation of Yunnan Provincial Academy of Science and Technology (2015IC017).

Notes and references

1 Y.-J. Qiao, J.-H. Shang, D. Wang, H.-T. Zhu, C.-R. Yang and Y.-J. Zhang, *Nat. Prod. Bioprospect.*, 2018, **8**, 245–263.
2 C.-R. Yang, *Res. Pract. Chin. Med.*, 2015, **29**, 83–86.
3 T. Wang, R.-X. Guo, G.-H. Zhou, X. Zhou, Z.-Z. Kou, S. Feng, C. Li, L.-Y. Tang and Z.-J. Wang, *J. Ethnopharmacol.*, 2016, **188**, 234–258.
4 J.-W. Lu, *Yunnan Agriculture*, 2016, **11**, 28–30.
5 Z.-J. Chen, J. Zeng, Y. Wang, Y.-Q. Sun and G.-Q. Feng, *J. Chin. Med. Mater.*, 2002, **25**, 387–389.
6 E.-L. Rice, *Allelopathy*, Academic Press, New York, 2nd edn, 1984.
7 X.-F. Huang, Y. Li and W.-L. Ding, *Seed*, 2009, **28**, 4–7.
8 M.-L. Wei, Y.-Q. Sun, T.-W. Huang, X.-M. Cui, Z.-J. Chen and B.-Y. Wang, *Special Wild Economic Animal and Plant Research*, 2010, **32**, 32–34.
9 Y. Zhu, L. Yang, X.-M. Cui, Y.-Q. Sun, L.-P. Guo and L.-Q. Huang, *Spec. Wild Econ. Anim. Plant Res.*, 2013, **2**, 40–42.
10 P.-J. You, W.-Q. Wang, Y. Zhang, Z.-L. Zhang, Y.-X. Pang and X.-M. Cui, *Southwest China Journal of Agricultural Sciences*, 2009, **22**, 308–310.
11 P.-J. You, Y. Zhang, W.-Q. Wang, Z.-L. Zhang, J.-Z. Yang and L.-M. Yin, *Mod. Chin. Med.*, 2009, **11**, 12–13.
12 K.-M.-M. Dakshini, On laboratory bioassays in allelopathy, *Bot. Rev.*, 1995, **61**, 28–44.
13 J. Zhao, X.-D. Zhang, L.-C. Wang, L.-J. Zheng and Z.-Y. Wang, *Chin. J. Microecol.*, 2018, **30**, 146–149.
14 L. Zhu, N. Ma, X.-M. Cui, Y. Zhu, J.-M. Zhou, L.-P. Guo and L.-Q. Huang, *Res. Pract. Chin. Med.*, 2014, **28**, 3–5.
15 L.-W. Sumner, A. Amberg, D. Barrett, M.-H. Beale, R. Beger, C.-A. Daykin, T.-W. Fan, O. Fiehn, R. Goodacre and J.-L. Griffin, *Metabolomics*, 2007, **3**, 211–221.
16 J.-H. Liu, C.-S. Lee, K.-M. Leung, Z.-K. Yan, B.-H. Shen, Z.-Z. Zhao and Z.-H. Jiang, *J. Agric. Food Chem.*, 2007, **55**, 8830–8835.
17 J.-W. Wong, M.-K. Hennessy, D.-G. Hayward, A.-J. Krynitsky and I. Cassias andF.-J. Schenck, *J. Agric. Food Chem.*, 2007, **55**, 1117–1128.
18 S. Chan, M.-F. Kong, Y.-C. Wong, S.-K. Wong and D.-W.-M. Sin, *J. Agric. Food Chem.*, 2007, **55**, 3339–3345.
19 G.-X. Xie, Y.-P. Qiu, M.-F. Qiu, X.-F. Gao, Y.-M. Liu and W. Jia, *J. Pharm. Biomed. Anal.*, 2007, **43**, 920–925.
20 X.-G. Yang, T.-S. Liu, B. Chen, F.-Y. Li, S.-S. Chu and L.-M. Gong, *Chinese Agricultural Science Bulletin*, 2013, **29**, 173–177.
21 J. Ju, M.-Z. Zhu, J.-R. Zhang and L.-Z. Xia, *Chin. Hosp. Pharm. J.*, 2007, **27**, 1076–1078.
22 J. Zhang, D.-S. Yang, M.-X. Li and L.-X. Shi, *PLoS One*, 2016, **11**, e0159622.
23 R. Guo, L. Shi, C. Yang, C. Yan, X. Zhong, Q. Liu, X. Xia and H. Li, *Front. Plant Sci.*, 2016, **7**, 1–9.
24 H.-Z. Sun, B. Wang, J.-K. Wang, H.-Y. Liu and J.-X. Liu, *J. Anim. Sci. Biotechnol.*, 2016, **7**, 49.
25 M.-A. Nemeth, Multi- and megavariate data analysis, *Technometrics*, 2003, **45**, 362.
26 D.-L. Massart, B.-G.-M. Vandeginste, S.-N. Deming, Y. Michotte and L. Kaufman, *Chemometrics: a textbook*, ed. B.-G.-M. Vandeginste and L. Kaufman, Elsevier, 1988, pp. 104–105.
27 C.-F. Ke, T. Zhang, X.-Y. Wu and K. Li, *Chinese Journal of Health Statistics*, 2014, **31**, 357–365.
28 J.-Y. Aa, *Chin. J. Clin. Pharmacol. Ther.*, 2010, **15**, 481–489.
29 Z.-H. Zhang, S. Tao, B.-X. Ye, Z.-Q. Peng and J.-P. Yuan, *Chin. J. Soil Sci.*, 2004, **35**, 793–798.
30 B.-L. Zhou, F. Chen, N. Liu, Q. Wu and B. Lu, *Acta Agric. Boreali-Occident. Sin.*, 2010, **19**, 179–183.
31 Q.-J. Zhang, A.-H. Zhang, J.-B. Sun and L.-X. Zhang, *J. Ecol. Environ. Sci.*, 2012, **21**, 187–193.
32 X.-X. Sun, H.-B. Wang, H.-B. He, J.-C. Lu and W.-X. Lin, *Chin. J. Eco-Agric.*, 2007, **22**, 806–812.
33 J.-M. Zhou, W.-B. Zhang and J.-Z. Yang, *Res. Pract. Chin. Med.*, 2012, **26**, 14–16.
34 L.-J. Wu, J. Liu and W.-Y. Wang, *World Science and Technology-Modernization of Traditional Chinese Medicine and Materia Medica*, 2014, **4**, 825–829.

