Contents lists available at ScienceDirect

Synthetic and Systems Biotechnology

Comprehensive relative quantitative metabolomics analysis of lycopodium alkaloids in different tissues of *Huperzia serrata*

Shiwen Wu [a], Zhen Fan [a, b], Youli Xiao [a, b, c, *]

[a] *CAS Key Laboratory of Synthetic Biology, CAS Center for Excellence in Molecular Plant Sciences, Institute of Plant Physiology and Ecology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, 200032, China*
[b] *University of Chinese Academy of Sciences, Beijing, 100039, China*
[c] *CAS-JIC Centre of Excellence in Plant and Microbial Sciences, Shanghai, 200032, China*

ARTICLE INFO

ABSTRACT

Qian ceng Ta, the whole plant of *Huperzia serrata*, is an important landscape and medicinal herbs and contains abundant bioactive lycopodium alkaloids. Although the structures of more than 100 lycopodium alkaloids in *Huperzia serrata* have been isolated and identified, the content and distribution of these alkaloids in different tissues are still unclear. In current study, an ultra-performance liquid chromatography-mass spectrometry based comprehensive metabolomics strategy was developed, including the extraction, separation, identification, and statistical analysis. The results showed that different types lycopodium alkaloids could be separated at different time-windows, which was helpful for further metabolite identification. Peak4388 and peak3954 were metabolite biomarkers for the different tissues according to the principle component analysis and partial least squares-discriminant analysis model. A computational tool based in-house database was also built up and used for putative identification. Of the 2354 true peaks after four-step filtration, 118 peaks were putatively identified as lycopodium alkaloids by using in-house database, and four of which was identified by authentic standards. Alternatively, another computational software was used to predict the fragmentation pattern, to dereplicate the structure of identified peaks, and identified the peak3585 to N-methylhuperzine A. The integration of both computational tools could be used for more metabolites identification.

1. Introduction

Qian Ceng Ta, the whole plant of *Huperzia serrata* (*H. serrata*), was an important Traditional Chinese Herbs (TCH) for the treatment of contusions, strains, swellings, schizophrenia, myasthenia gavis, and organophosphate poisoning since Tang Dynasty [1,2]. Huperzine A (HupA), a representative lycodine-type lycopodium alkaloid, was first isolated and identified by Chinese scientist in 1986 [2] and was found as a powerful, highly specific, and reversible acetylcholinesterase inhibitors [3,4]. This natural product has attracted widespread attention because of its unique pharmacological activities and low toxicity. Owing to the low content of HupA, slow growth of the plant, and no plant culture method for this plant, the supply for this compound is not sustainable [5]. Producing the lycopodium alkaloids in heterologous host by using synthetic biology is an alternative method, because this method has been successful to produce ginsenoside in yeast [6] and artemisinin in tobacco [7].

The biosynthetic pathway of HupA has been proposed based on labeled precursors feeding experiment by Ma and Gang [4]. Lysine/Ornithine decarboxylase (L/ODC) was proposed as the first enzyme

for HupA biosynthesis. Indeed, L/ODC could catalyze the first-step in the biosynthesis pathway of lysine-derived alkaloids in different lycopodium plants [8−10]. After the decarboxylation of lysine to produce cadaverine, it would be catalyzed by copper amine oxidase (CAO) to yield 5-aminopentanal, which is spontaneously cyclized to the first intermediate for lysine-derived alkaloids production, Δ1-piperideine [11]. However, the genes involved in the skeleton formation and modification are still unclear. Although some candidate genes for HupA biosynthesis were proposed based on the EST data and the global RNA-Seq data from *H. serrata*, none was characterized because of the lack knowledge of substrate and intermediates involved in HupA biosynthesis.

Metabolomics, focusing on the high-throughput qualitative and quantitative profiling of small molecular metabolites, has been used for the candidate genes mining and characterization by integrating with other omics [12]. In the past years, targeted metabolomics has been used for the detection of HupA in different lycopodium plants base on high-performance liquid chromatography (HPLC) to find the better sources than *H. serrata* [13]. To improve the efficiency of extraction and detection, an high performance liquid chromatography coupled with mass spectrometry (HPLC-MS) based method was developed for HupA detection, identification and quantification [14]. However, these works only focused on the content of HupA. Subsequently, another ultra-performance liquid chromatography-mass spectrometry (UPLC-MS) based metabolomics method was developed and 27 lycopodium alkaloids were identified, of which only six metabolites were identified based on comparison of accurate MS and MS/MS data compared to the authentic standards [5]. Currently, more than 200 lycopodium alkaloids have been isolated and identified by NMR [4]. Given the limited authentic standards and recruited lycopodium alkaloids in the public databases, and limited reference MS/MS spectra, only few could be identified by commonly used tools. In addition, many lycopodium alkaloids had the same formula and skeleton, thus the MS and MS/MS spectra are similar, making it more difficult to identify. Combination of the computational tools will be more useful for the identification because the identification based on these tools were independent on the public databases.

In this study, a comprehensive relative quantitative metabolomics method for the extraction, detection and identification of lycopodium alkaloids in different tissues of *H. serrata* was firstly developed. For identification, a computational tool based lycopodium alkaloids specific in-house database was developed and identified118 putative chemicals. In addition, other computational software was also employed to dereplicate the structure of lycopodium alkaloids and identified the peak3585 to N-methyl-huperzine A based on the in-silico fragmentation pattern.

2. Materials and methods

2.1. Plant materials

The plants were collected from Xiangxi, Hunan, China, in January 2017. The plants were carefully rinsed in running tap water, and soil was removed by hand. Root, stem, and leaf were separated from the plant and ground to fine power in liquid nitrogen and ice-dried, and then stored at-80 °C until further using.

2.2. Authentic standards

The authentic standards of huperzine A, huperzine B (HupB), Lycoposerramine C (Lycop C) and lycoposerramine D (Lycop D) were separated from the crude alkaloids extracted from *H. serrata*

and identified by NMR (Supplementary method) according to previous reports [2,15]. All purified authentic standards (>95% purity, Supplementary Figs. S1, S2) were dissolved in methanol with 100 ng/mL, filtered through 0.22 μM PTEE filters and stored at −20 °C until for using.

2.3. Lycopodium alkaloids extraction

The extraction of lycopodium alkaloids was carried out according to previous report with minor modification [16]. Around twenty mg dry power was weighted and extracted by 1 mL methanol including 0.1% (v/v) formic acid. After adding 400 ng caffeine as internal standard, the sample was sonicated for 15 min at room temperature. Centrifugation at 17000g for 10 min, the supernatants were filtered through 0.22 μM PTEE filters and then analyzed by UPLC-MS. Meanwhile, the blank samples were also treated with the same procedure. All samples including the blank were repeated three times.

2.4. Analytical platform

A waters Ultra performance liquid chromatography system (Waters Corporation, Milford, MA, USA) was used. That system was coupled to a Thermo Scientific Q-Exactive mass spectrometer equipped with a HESI interface (Thermo Scientific, Hemel Hempstead, U.K.). Thermo Xcalibur Tune software (version 3.0) was used for instrument control and data acquisition.

The separation was performed with a Waters Acquity HSS T3 C18 column (150 mm × 2.1 mm, i.d. 1.8 μm) in positive mode ESI. A linear biphasic LC gradient was conducted from 2% to 60% B over 12 min, and then kept at this phase for 2 min, followed by 3 min wash with linear biphasic gradient from 60% B to 2% B over 1 min, and 3 min re-equilibration with 2% B, where solvent A was 0.1% formic acid in water and solvent B was acetonitrile. The flow rate was 0.25 mL/min, column temperature was maintained at 30 °C, injection volume was 2 μL, and samples were maintained at 4 °C in the autosampler.

2.5. Data processing

The raw data files were firstly converted to mzXML format files and mzML format files using MSConvert [17]. After peak deconvolution and alignment processing by XCMS [18,19] and MS-DIAL [20], the given peak alignment table was further used to choose the true peaks before putative identification based on four-step filtration (Fig. 1) according to the previous method [21]. The filtered peaks were used for putative identification in different databases by using precursor ions (*m/z*) and/or MS/MS spectrum based on different software, xMSannotator [22], Metabosearch [23], and MS-DIAL [20]. Here, an in-house database for lycopodium alkaloids in *H. serrata* was also developed based on the computational tool, plantMAT [24]. Furthermore, MS-FINDER [25] and CSI:FingerID [26], were used for the prediction of in-silico fragmentations.

After processing by MS-DIAL [20], a plot view was shown and each plot was a detected peak. The information of spectrum, retention time, *m/z* value, and so on could be shown by left click. The interested peak (plot) could be linked to MS-FINDER [25] for exact identification by in-silico fragmentation pattern by right click. More details could refer to the manual tutorial (http://prime.psc.riken.jp/Metabolomics_Software/MS-DIAL/MS-DIAL%20tutorial_vs9.pdf).
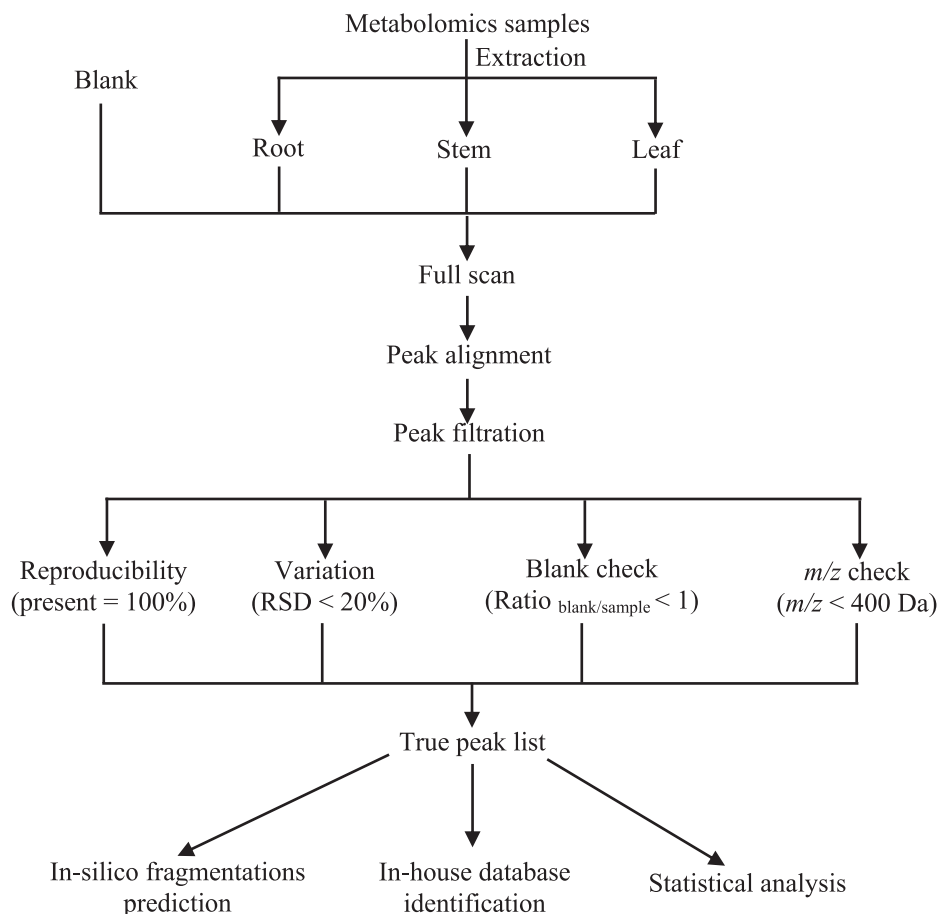
Fig. 1. The strategy of LC-MS-based comprehensive metabolomics analysis. The four-step filtration was carried out as described [21].

2.6. Chemicals identification

All peaks processed by XCMS were directly submitted to our in-house database for putative identification. After the identification by this in-house database, we could get the possible chemotypes and related constituents for each peak, and then the putatively identified peaks were submitted to MS-FINDER for exact identification by in-silico fragmentation pattern. Using these two steps analysis, the peak could be exactly identified including the chemotype and constituents.

2.7. Statistical analysis and visualization

The putatively identified peaks was listed in a csv file and then submitted to the web-based tool, MetaboAnalyst [27], for statistical analysis and visualization by using the default parameters.

3. Results and discussion

Owing to the complex of the lycopodium alkaloids biosynthesis network and limited genomic data, HupA biosynthetic pathway was still unclear although the proposed pathway had been published in 2004 based on feeding experiment [4]. To further elucidating HupA biosynthetic pathway, the transcriptomes were determined and mined candidate genes. However, these candidate genes were still not characterized with the exception of L/ODC

[8–10] and *CAO* [11]. Metabolomics coupled with transcriptomics was an efficient way to mine and characterize the candidate genes. For example, glycyrrhizin biosynthetic pathway was almost fully elucidated by using the metabolomics-based analysis [28–30]. Although more than 100 lycopodium alkaloids in *H. serrata* had been identified by NMR [31], the contents and distribution in different tissues were still unknown. For the comparison between different tissues or different individuals, absolutely quantitative metabolomics analysis is not necessary [32]. Hence, the comprehensive relative quantitative metabolomics analysis of the lycopodium alkaloids in different tissues of *H. serrata* was firstly performed in current study.

3.1. Analysis of the lycopodium alkaloid profiling in H. serrata

All samples from different tissues (root, stem, and leaf) were first extracted by the buffer of 0.1% formic acid in methanol and then analyzed in positive mode ESI [16]. In this case, the main peaks extracted from three tissues were mainly present at the retention time from 3 min to 8 min (Fig. 2A). After the data processing by MS-DIAL [20], the plot viewer also supported this point (Fig. 2B). The plot labeled with the blank circle ($m/z = 243.14919$, $t_R = 4.79$ min) which was identified as the HupA by authentic standard (Fig. 3A) was also located in these region (Fig. 2C). It suggested the lycopodium alkaloids were mainly separated in the time-window from 3 min to 8 min by using our LC separation method, which could give
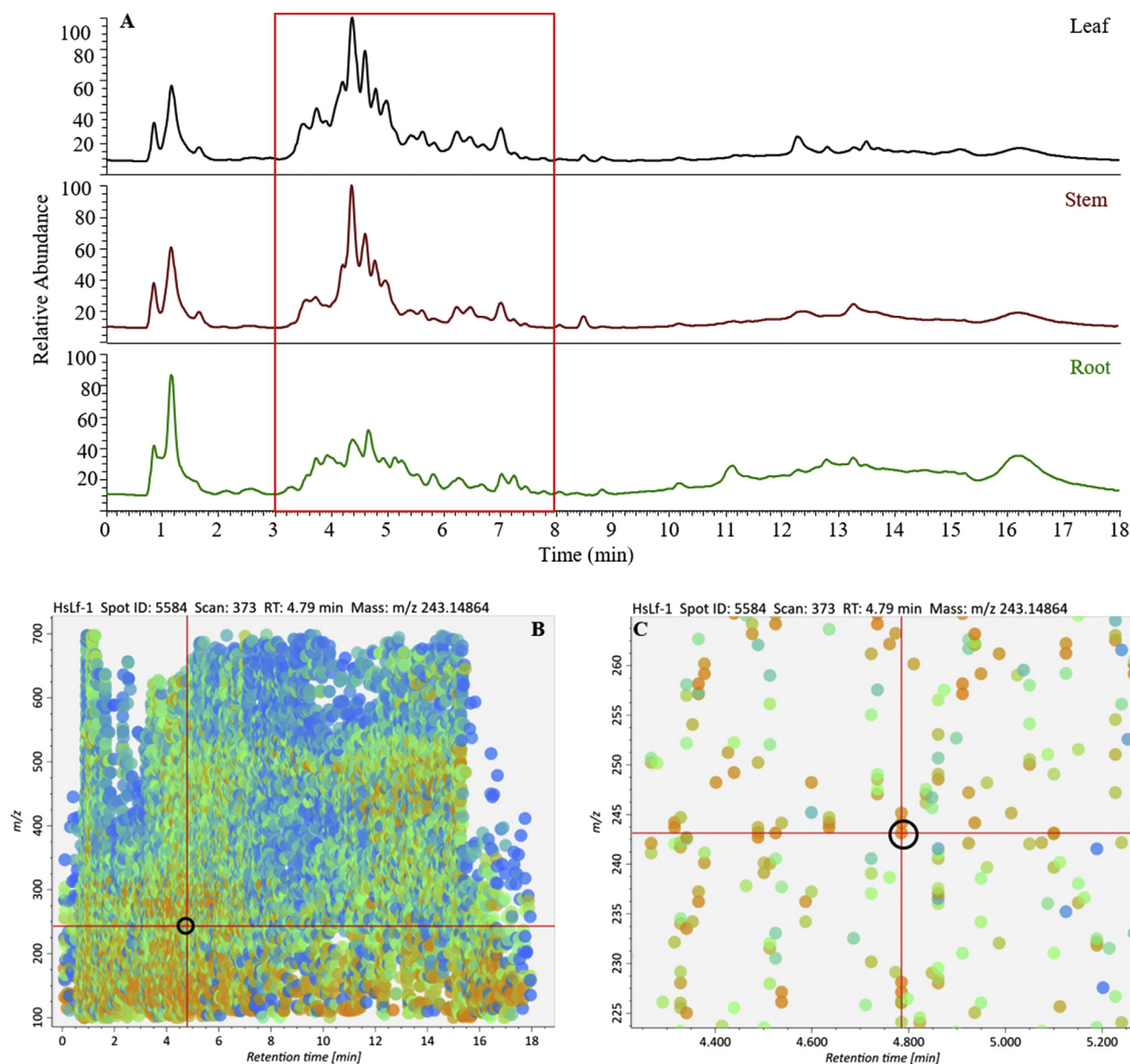
Fig. 2. Metabolite profiling of lycopodium alkaloids in leaf, stem, and root. (A) Total ion chromatography map of the lycopodium alkaloids in different tissues. (B) plot view of the peaks was processed by the software of MS-DIAL [20]. (C) Magnified plot view of (B) to showed the labeled plot. t_R ranged from 4.2 min to 5.3 min, and m/z value ranged from 220 Da to 270 Da. The plot labeled with black circle was HupA that was subsequently identified by chemical standard as shown in Fig. 3A.

a suggestion how to optimize the LC separation method to efficiently separate the lycopodium alkaloids.

Here, the four authentic standards and the internal standard were all identified in the tissues of leaf and stem. For the tissue of root, Lycop D was likely not present or the content of the compound in this tissue was lower than the determination limitation in this instrument. The lycodine-type lycopodium alkaloids (HupA and Hup B) could be shown earlier than fawcettimine-type lycopodium alkaloids (Lycop C and Lycop D) in our LC separation system. Similarly, Lycopodine, the lycopodine-type lycopodium alkaloid, had the later retention time than HupA [5]. The internal standard (IS), caffeine ($t_R = 5.34$), was totally different type alkaloid with the lycopodium alkaloids, and separated later than all the four lycopodium alkaloids, indicating that different types lycopodium alkaloids could be separated at different time-windows during the separation.

3.2. Data processing and putative metabolite identification

The file formats of mzXML and mzML are standardization initiatives [22,24], and could be submitted to the two commonly used metabolomics data analysis software of XCMS [18,19] and mzMine 2 [33] for deconvolution and alignment analysis. XCMS Online [34] was a useful web-based tool for metabolomics data analysis automatically, including peak deconvolution, alignment, putative identification, statistical analysis, and pathway enrichment after submitting the raw data less than 25 Gb size. In addition, mzMine 2 [33] was another useful tool for metabolomics analysis and putative identification based on different databases, especially the ADAP algorithm that could reduce the false positive and false negative
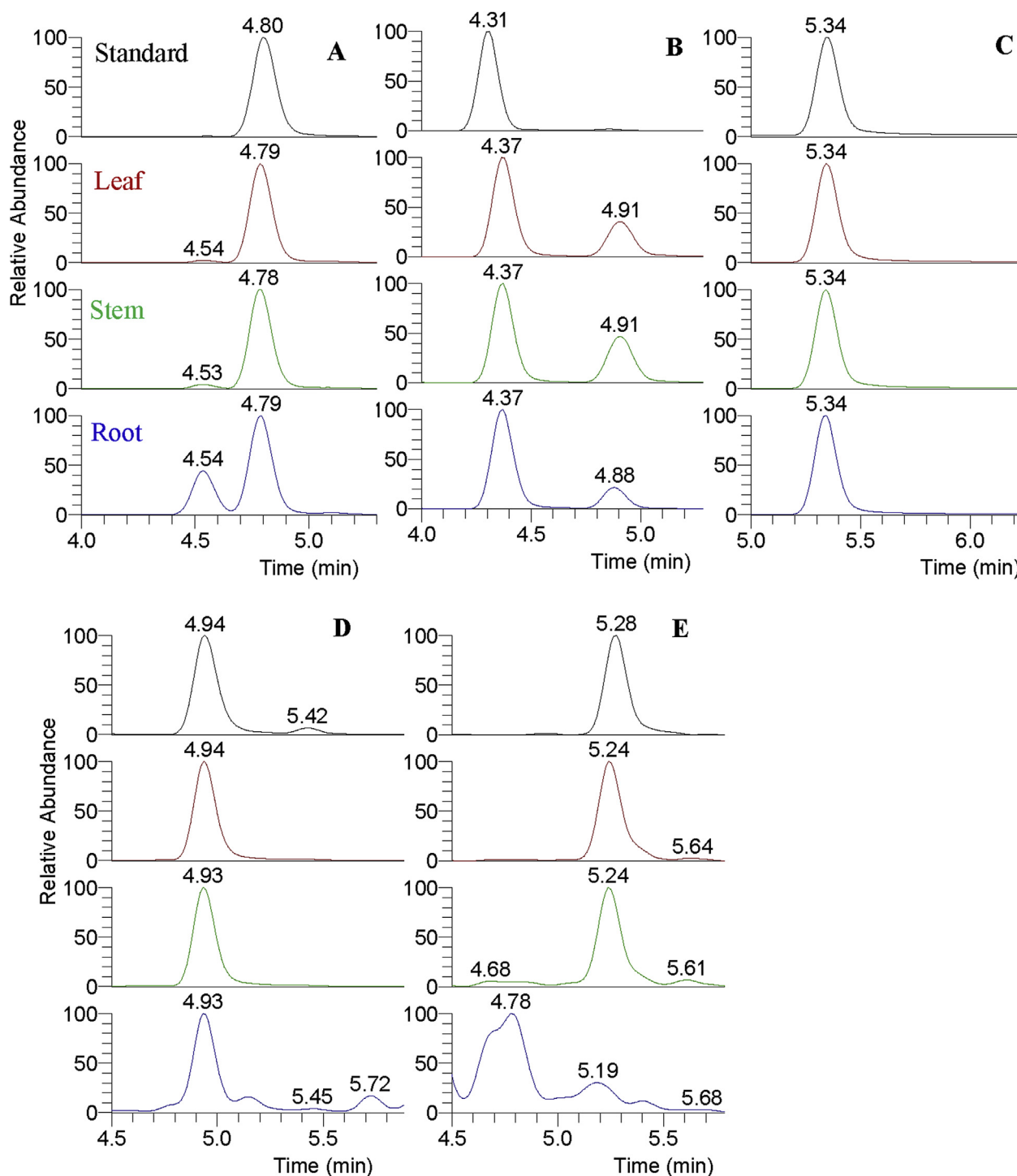
Fig. 3. Extracted ion maps (EIC) of identified peaks by authentic standards in different tissues of *H. serrata*. (A) HupA at $t_R = 4.79$ min extracted with $m/z = 243.14919$, peak4297 at $t_R = 4.54$ min was also extracted with $m/z = 243.14919$; (B) HupB at $t_R = 4.37$min extracted with $m/z = 257.16484$, peak4360 at $t_R = 4.91$ min was also extracted with $m/z = 257.16484$; (C) The internal standard of caffeine at $t_R = 5.34$ extracted with $m/z = 195.08765$; (D) Lycop C at $t_R = 4.94$ min extracted with $m/z = 262.18016$; (E) Lycop D at $t_R = 5.28$ min extracted with $m/z = 292.19072$, this metabolite perhaps not present in root or the content was lower than the limit of determination.

peak picking from XCMS [35]. However, these two tools were not convenient for our study because the former one was not allowed to download the identified result, and the later one was need higher computer configuration and required additional Java and R language platform to perform the ADAP algorithm. Here, we used two R packages, XCMS and CEMARE, to process the metabolomics data and resulted more than 15000 peaks as the described parameters [36]. To pick up the true peaks, four-step filtration was carried out as described [21] and resulted 2354 true peaks including the five chemicals identified by authentic standards, indicating that the four-step filtration was useful to get the true peaks.
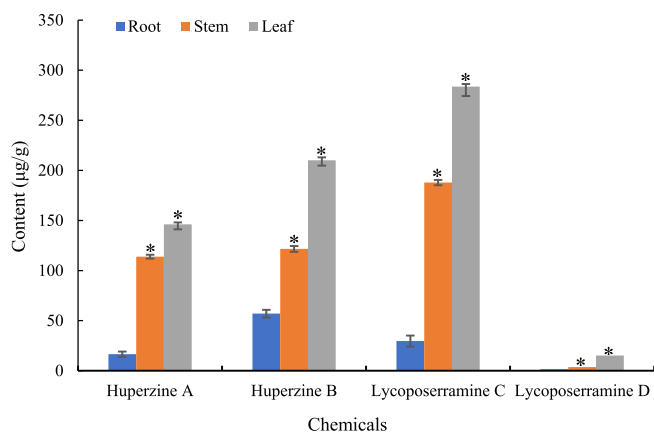
Fig. 4. The contents of identified chemicals in different tissues. *means that significant differences between leaf and stem comparing with root (*P* < .05). Three biological duplicates were performed.

The contents of the identified lycopodium alkaloids were firstly calculated based on the internal standard (Fig. 4). Comparing with root, the content of the four chemicals was significantly higher in stem and leaf, and the four chemicals had same tendency. The distribution of HupA was consistent with the previous study [13]. In addition, the content of HupA was not the most abundant one even in the four identified chemicals, suggesting the complex of the HupA biosynthetic pathway. Meanwhile, the true peaks were used for putative identification by different databases, KEGG, HMDB, LipidMAPS, T3DB, and CCMD, based on the software of MetaboSearch [23] and xMSannotaor [22] (Table 1). Among these databases, caffeine could be identified by KEGG, HMDB, and T3DB, and HupB was only identified by KEGG. Interestingly, HupB could be identified by CCMD using exact *m/z* value but the retention time is not matched. Other three chemicals could not be identified by these databases. It suggested that most lycopodium alkaloids in *H. serrata* are not recruited in these commonly used databases. MS-DIAL [20] could also be used for data processing and putative identification based on multiple MS/MS banks. Thus, we tried to putatively identify the lycopodium alkaloids based on the MS/MS bank directly using this software. Similarly, only caffeine could be identified by the MS/MS banks of MassBank, GNPS and Respect. It was reasonable that most the lycopodium alkaloids were not recruited by public databases and MS/MS banks. In another word, the current commonly used public MS/MS banks and databases were not suitable for the putative identification of lycopodium alkaloids in *H. serrata*.

As mentioned above, the lycopodium alkaloids specific database is pivotal for the putative identification of lycopodium alkaloids.

PlantMAT [24], a computational tool based on combinatorial enumeration of metabolic building blocks, was developed to further identifying the specialized metabolites that there was no authentic standards or not recruited in public databases. Using this tool, we built a lycopodium alkaloids specific in-house database for putative identification. All the five identified chemicals could be identified by this in-house database (Table 1), indicating that this in-house database was helpful for our study. Totally, 118 lycopodium alkaloids were putatively identified by this in-house database (Table 2). Nevertheless, most of the peaks could be identified to multiple metabolite names even using both MS spectrum and MS/MS spectrum. Thus, the peak number derived from XCMS alignment table was still kept for further analysis.

3.3. Statistical analysis and in-silico fragmentations prediction

The contents of the 118 putatively identified peaks were calculated by the content of internal standard, and then submitted to MetaboAnalyst [27] for statistical analysis and visualization. Principle component analysis (PCA) and partial least squares−discriminant analysis (PLS-DA) were two commonly used algorithms to analyze the metabolic profiling derived from different samples. PCA and PLS-DA plot showed the same result: metabolite profiling of three tissues were separated obviously by the first component (Fig. 5A, C), and peak4388 and peak3954 were metabolite biomarkers for the three tissues (Fig. 5B, D, Supplementary Dataset S1, S2). The content of these two peaks were significantly higher in leaf and stem than in root (Fig. 6), and the contents of these two peaks were also significantly higher than HupA content (Fig. 3). Among the 118 identified peaks, 72 peaks were significantly different among three different tissues (*P* < .05). Subsequently, the Pearson correlation analysis of the identified peaks was performed and all these peaks could be separated to 2 totally different groups (Fig. 7). Of which, the four identified lycopodium alkaloids and the two metabolite biomarkers of peak4388 and peak3954 were the same group with HupA. Nearly 70% of the peak had positive correlation, and these peaks were almost positively with HupA, suggesting that these peaks might participate in the biosynthesis and modification of HupA.

Given the limited authentic standards for lycopodium alkaloids, the predictive computational tools could be helpful for more efficient and accurate structure dereplication and prediction. CSI:-FingerID [26], a tool to identify chemical structures without search spectral libraries, enabled to predict unknown structures by machine learning of MS/MS data. There are four skeleton type lycopodium alkaloids in lycopodium plants, thus the fragments pattern will be similar for the chemicals with the same skeleton. Currently, we know nothing about the fragmentation patterns of lycopodium alkaloids even the identified HupA. Hence, HupA was first used to

Table 1
The value of different databases for lycopodium alkaloids identification.

Name	KEGG	HMDB	LipidMAPS	T3DB	CCMD	MS-DIAL	in-house database
Huperzine A	No	No	No	No	No	No	Yes
Huperzine B	Yes	No	No	No	*	No	Yes
Caffeine (IS)	Yes	Yes	No	Yes	No	Yes	Yes
Lycoposerramine C	No	No	No	No	No	No	Yes
Lycoposerramine D	No	No	No	No	No	No	Yes

No, means the metabolite was not identified by this database using *m/z* value and retention time.
Yes, means the metabolite could be identified by this database using *m/z* value and retention time;
*, means that the metabolite could be identified by *m/z* value but not be identified by retention time.

Table 2
Putatively identified peaks by in-house database.

Peaks	m/z	t_R (min)	Peaks	m/z	t_R (min)	Peaks	m/z	t_R (min)
peak0410	130.0860776	0.90	peak4302	278.1745	3.99	peak4356	279.242544	5.13
peak5825	335.1976279	0.94	peak3913	265.226918	4.00	peak3842	263.24764	5.14
peak3664	259.1434914	1.19	peak3160	243.185066	4.01	peak3864	264.195198	5.15
peak0409	130.0861463	1.19	peak3870	264.195199	4.06	peak4878	295.237311	5.18
peak4824	294.1692148	1.24	peak4128	273.159142	4.07	peak3305	247.216448	5.23
peak0411	130.086147	1.65	peak4797	293.221687	4.12	peak4798	293.221611	5.23
peak4389	280.1902164	2.19	peak4898	296.184946	4.14	**peak4764**	**292.190069**	**5.23**
peak3659	259.1436636	2.63	peak3954	266.210889	4.20	peak5132	305.22169	5.25
peak4821	294.1694482	2.70	peak3773	262.17974	4.20	peak3668	259.180031	5.26
peak4129	273.1592313	2.90	peak5823	335.195651	4.22	peak4358	279.242582	5.32
peak4192	275.1749012	3.22	peak4383	280.190039	4.22	peak3920	265.22687	5.32
peak4766	292.1901053	3.26	peak5134	305.221647	4.26	**peak1904**	**195.08742**	**5.33**
peak3778	262.179677	3.27	peak3869	264.195336	4.27	peak3776	262.179623	5.41
peak4904	296.1850784	3.27	peak3658	259.143606	4.27	peak3334	248.20036	5.44
peak4225	276.1591354	3.27	peak4300	278.174535	4.28	peak5547	322.200617	5.47
peak3158	243.1486756	3.28	**peak3586**	**257.164199**	**4.37**	peak3912	265.226887	5.53
peak4296	278.1746205	3.29	peak3671	259.179879	4.38	peak4876	295.237215	5.56
peak4825	294.1693143	3.29	peak3437	251.24757	4.46	peak3956	266.210667	5.60
peak3669	259.1800797	3.29	peak4072	271.179803	4.49	peak3865	264.195264	5.62
peak4392	280.1901437	3.31	peak3156	243.14865	4.53	peak4228	276.19535	5.66
peak4127	273.1592432	3.37	peak4297	278.174518	4.54	peak5135	305.221683	5.70
peak3657	259.1436043	3.46	peak3777	262.179552	4.61	peak4365	279.242585	5.78
peak3957	266.2108828	3.48	peak4361	279.242533	4.66	peak4386	280.190274	5.84
peak4822	294.1693111	3.56	peak4385	280.190123	4.66	peak5510	320.232512	5.92
peak4298	278.1745387	3.58	peak3863	264.195179	4.72	peak3915	265.2269	5.98
peak3780	262.179616	3.59	**peak3155**	**243.148617**	**4.79**	peak3841	263.247602	6.19
peak5133	305.2216587	3.60	peak4900	296.184949	4.80	peak4230	276.19525	6.21
peak4190	275.1748371	3.61	peak4229	276.195037	4.80	peak3866	264.195235	6.27
peak3868	264.1952289	3.66	peak4877	295.237442	4.86	peak4299	278.174504	6.38
peak4765	292.1899733	3.73	peak4071	271.179715	4.89	peak6323	364.211006	6.58
peak3589	257.2006295	3.75	peak3585	257.164321	4.90	peak3774	262.1796	6.59
peak3666	259.1799926	3.75	peak4360	279.242394	4.91	peak3840	263.247614	6.74
peak5822	335.1956965	3.85	peak3303	247.216299	4.92	peak4303	278.174424	6.92
peak5550	322.2005207	3.87	**peak3779**	**262.179556**	**4.93**	peak3919	265.226833	7.18
peak4126	273.1591185	3.88	peak4301	278.174479	4.93	peak6321	364.211022	7.29
peak3772	262.1795936	3.88	peak5824	335.195522	4.93	peak3304	247.216393	7.81
peak4191	275.1747109	3.88	peak3672	259.180011	4.95	peak4879	295.237444	8.37
peak3656	259.143585	3.92	peak4388	280.190078	5.00	peak3775	262.179593	9.36
peak4875	295.2373523	3.96	peak5027	301.154019	5.08	peak4267	277.179252	12.29
peak4589	287.1746869	3.99	peak5546	322.200566	5.11			

The peaks labeled with red color and bold font were identified by authentic standards.

peak3155, Huperzine A; peak3586, Huperzine B; peak1904, caffeine (IS);

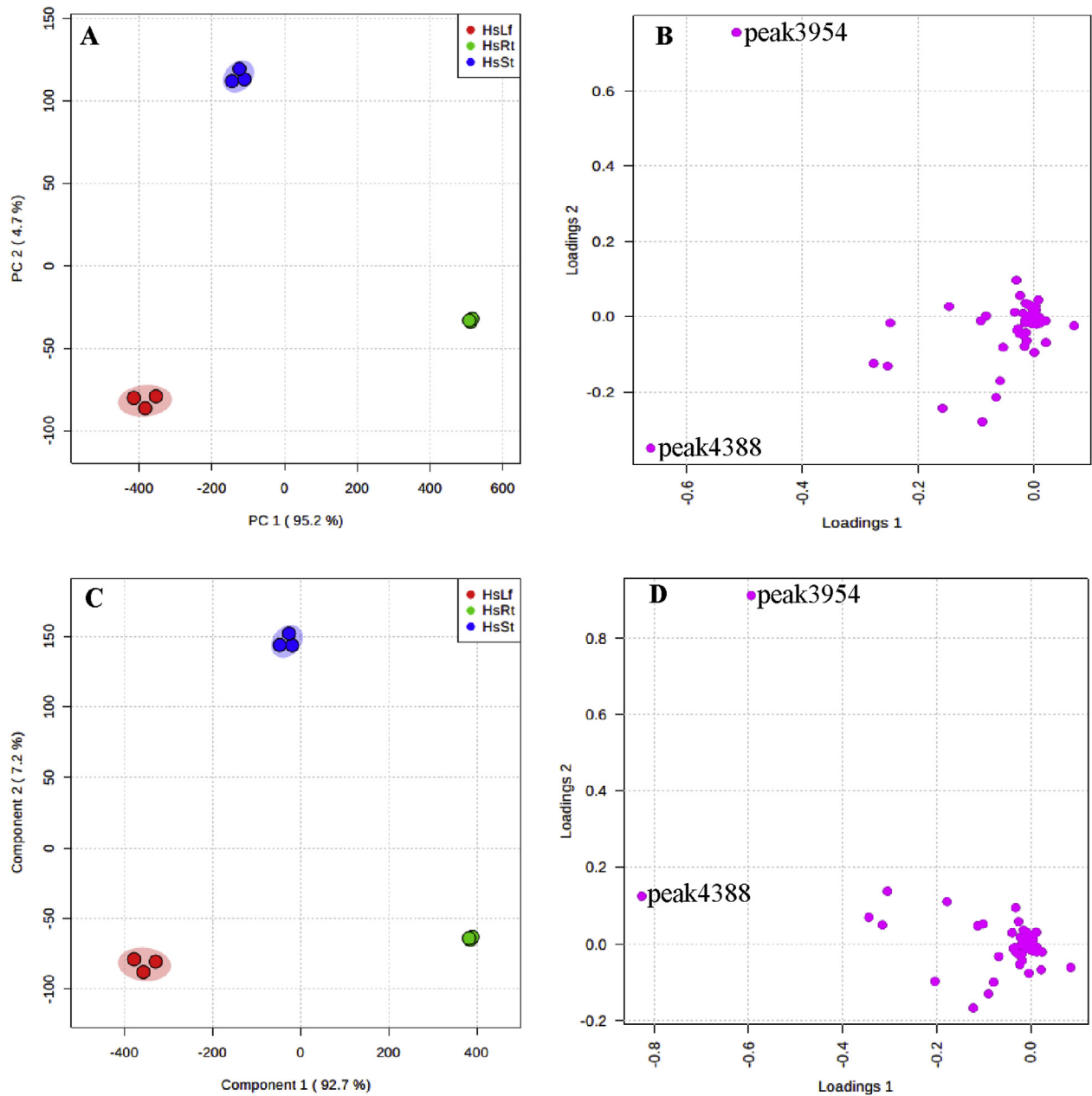peak3779, Lycoposerramine C; and peak4764, Lycoposerramine D.

predict the structure of unknown metabolites based on exactly analysis the MS/MS data and considering the hydrogen rearrangement rule. In addition, this tool could process the data derived from MS-DIAL. Namely, this tool could automatically recognize the experimental MS/MS spectrum and produce the in-silico fragmentations. Thus, we processed the data with MS-DIAL and then submitted HupA to MS-FINDER as the tutorial for fragmentation prediction. Similarly, the fragmentation pattern was the same as the former predicted one (Fig. 9A). Using this in-silico fragmentation pattern, it was easy to find which function group was lost. The fragmentation pattern (Fig. 9C) and experimental spectrum or in-silico spectrum (Fig. 9D) of peak3585 were similar with HupA, indicating this compound is HupA related. According the in-silico

spectra, precursor of peak3585 had a more group of [CH3] than HupA, thus it predicted to be methyl-huperzine A. When identifying using our in-house database, N-methylhuperzine A was a candidate, thus peak3585 should be N-methylhuperzine A. It indicated integrating our in-house database and MS-FINDER is more efficient for the lycopodium alkaloids identification.

In summary, a comprehensive metabolomics method was developed and could be used to separate different types of lycopodium alkaloids at different time-windows. Meanwhile, the developed in-house database was a useful tool to identify lycopodium alkaloids. Integration of the in-house database and the computational tool MS-FINDER could be used for identification of lycopodium alkaloids.
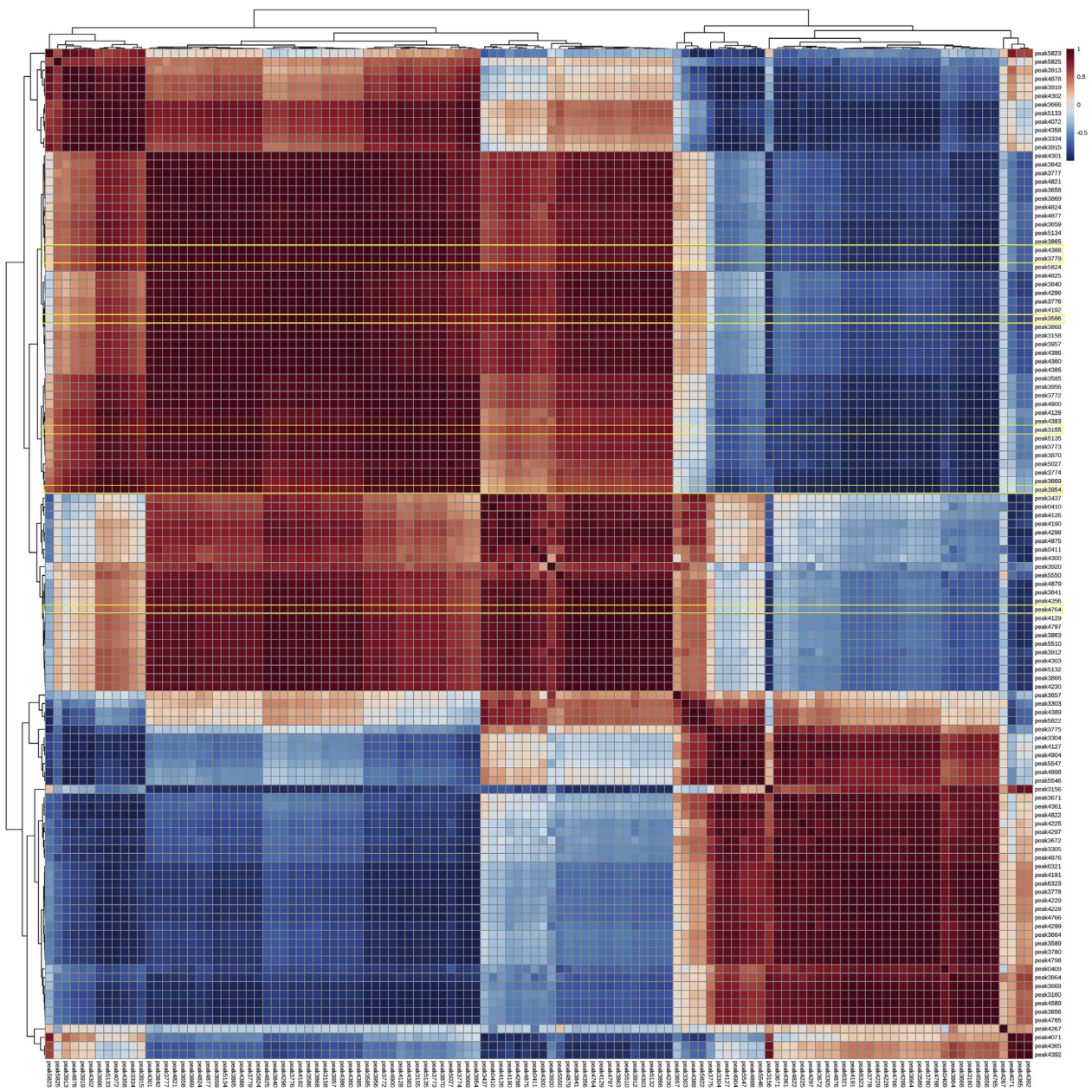

Fig. 7. The Pearson correlation of identified peaks. The peaks labeled with yellow square were the identified peaks by authentic standards and two metabolite biomarkers.
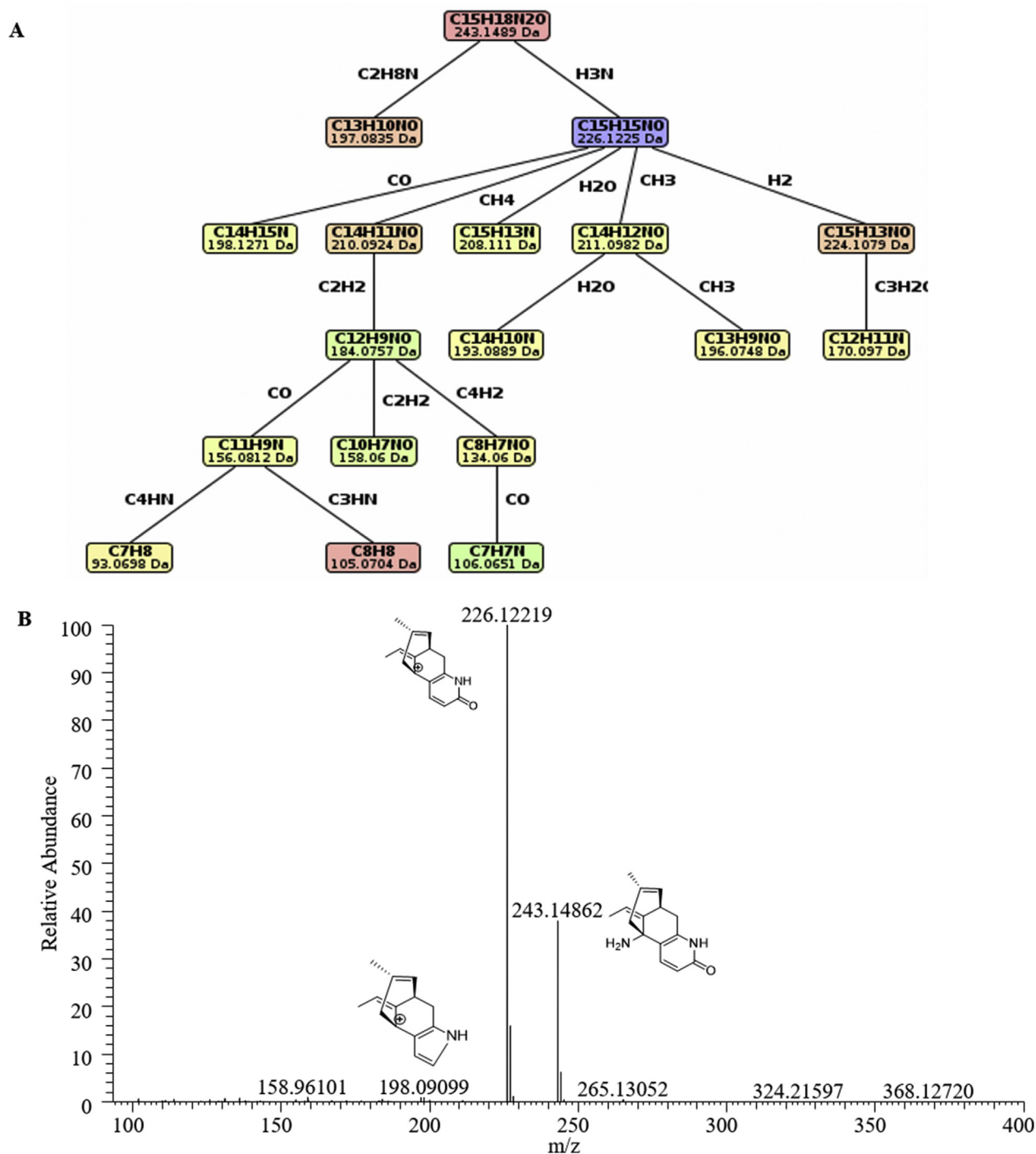
Fig. 8. Possible fragmentation pattern of HupA. (A) Fragmentation pattern predicted by CSI:FingerID; (B) Spectra and three predicted fragmentations of HupA.

4. Conclusion

In current study, we built a UPLC-MS strategy for the comprehensive metabolomics analysis for lycopodium alkaloids in different tissues of *H. serrata*. Here, a computational method based in-house database was developed and putatively identified 118 lycopodium alkaloids. Integrating our in-house database and another computational tool, MS-FINDER, could dereplicate the structure of the peak3585 to N-methylhuperzine A without the authentic standard. This strategy could help us to identify more lycopodium alkaloids.

Author contributions

YX and SW designed the experiment. SW performed most experiments and processed metabolomics data. ZF isolated and identified the authentic standards. SW and YX wrote the manuscript and revised the manuscript. All authors provided helpful
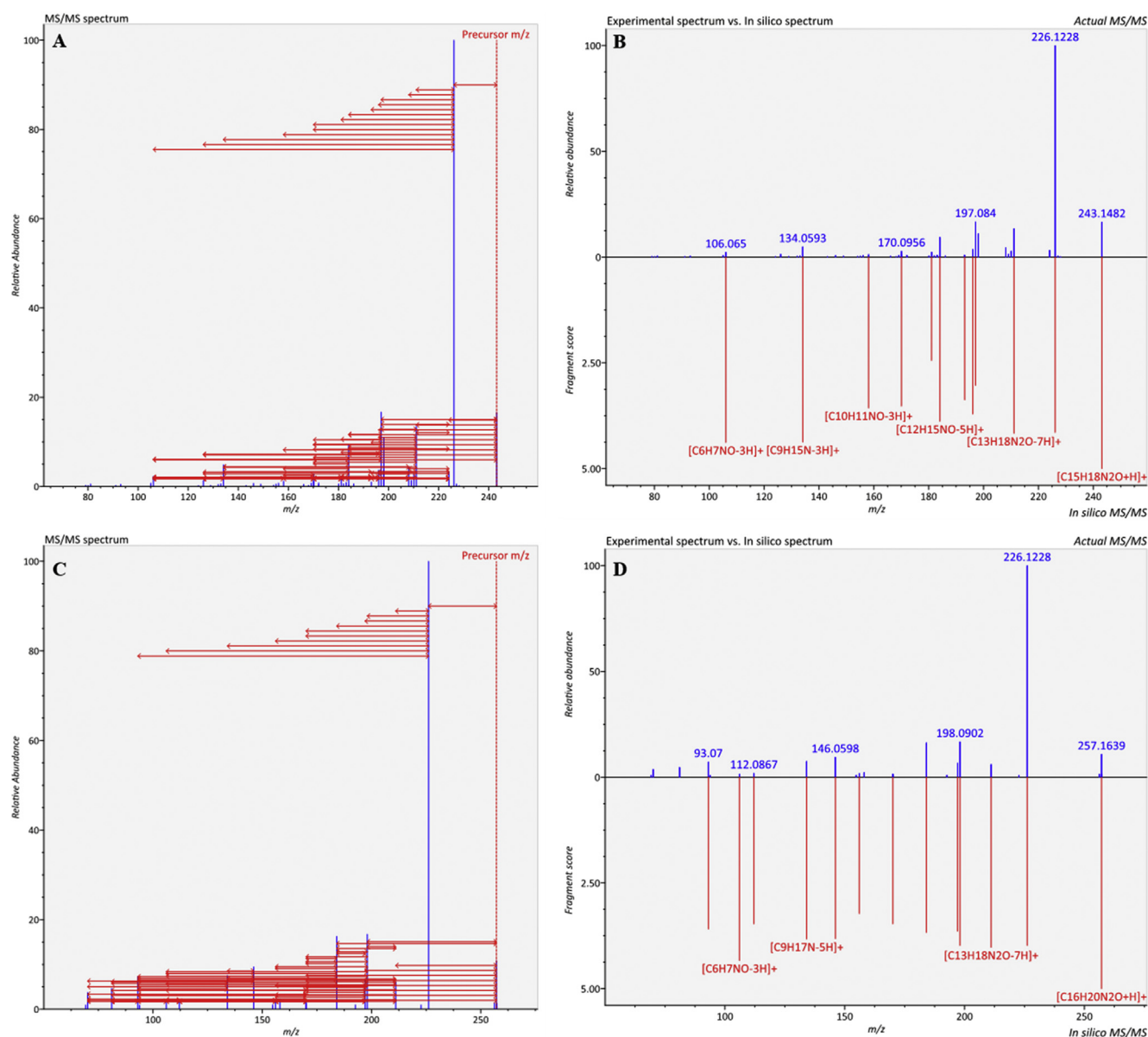
Fig. 9. Possible fragmentation pattern of HupA and peak3585. (A) Fragmentation pattern of HupA predicted by MS-FINDER; (B) Comparison between experimental spectrum and in-silico spectrum of HupA; (C) Fragmentation pattern of peak3585 predicted by MS-FINDER; (D) Comparison between experimental spectrum and in-silico spectrum of peak3585.

discussions and approved its final version.

Acknowledgments

This work was financially supported by the Science and Technology Commission of Shanghai Municipality (Grant 15JC1400402), CAS-JIC Centre of Excellence in Plant and Microbial Sciences (CEPAMS) funding, and the Strategic Priority Research Program "Molecular mechanism of Plant Growth and Development" of CAS (Grant XDPB0402). We thank Dr. Yuanhong Shan in the Core Facility Centre of the Institute of Plant Physiology and Ecology for mass spectrometry assistance.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.synbio.2017.12.003.

References

[1] Ma X, Tan C, Zhu D, Gang DR, Xiao P. Huperzine a from *huperzia* species—an ethnopharmacolgical review. J Ethnopharmacol 2007;113:15–34.
[2] Liu J, Zhu Y, Yu C, Zhou Y, Han Y, Wu F, et al. The structures of huperzine A and B, two new alkaloids exhibiting marked anticholinesterase activity. Can J Chem 1986;64:837–9.
[3] Tang XC, Desarno P, Sugaya K, Giacobini E. Effect of huperzine A, a new cholinesterase inhibitor, on the central cholinergic system of the rat. J Neurosci Res 1989;24:276–85.
[4] Ma X, Gang DR. The lycopodium alkaloids. Nat Prod Rep 2004;21:752–72.
[5] Ishiuchi K, Park J-j, Long RM, Gang DR. Production of huperzine A and other lycopodium alkaloids in *huperzia* species grown under controlled conditions and in vitro. Phytochemistry 2013;91:208–19.
[6] Yan X, Fan Y, Wei W, Wang P, Liu Q, Wei Y, et al. Production of bioactive ginsenoside compound K in metabolically engineered yeast. Cell Res 2014;24: 770–3.
[7] Farhi M, Marhevka E, Ben-Ari J, Algamas-Dimantov A, Liang Z, Zeevi V, et al. Generation of the potent anti-malarial drug artemisinin in tobacco. Nat Biotechnol 2011;29:1072–4.
[8] Bunsupa S, Katayama K, Ikeura E, Oikawa A, Toyooka K, Saito K, et al. Lysine decarboxylase catalyzes the first step of quinolizidine alkaloid biosynthesis

and coevolved with alkaloid production in leguminosae. Plant Cell 2012;24: 1202—16.

[9] Bunsupa S, Hanada K, Maruyama A, Aoyagi K, Komatsu K, Ueno H, et al. Molecular evolution and functional characterization of a bifunctional decarboxylase involved in lycopodium alkaloid biosynthesis. Plant Physiol 2016;171:2432—44.

[10] Xu B, Lei L, Zhu X, Zhou Y, Xiao Y. Identification and characterization of L-lysine decarboxylase from *Huperzia serrata* and its role in the metabolic pathway of lycopodium alkaloid. Phytochemistry 2017;136:23—30.

[11] Sun J, Morita H, Chen G, Noguchi H, Abe I. Molecular cloning and characterization of copper amine oxidase from *Huperzia serrata*. Bioorg Med Chem Lett 2012;22:5784—90.

[12] Luo J. Metabolite-based genome-wide association studies in plants. Curr Opin Plant Biol 2015;24:31—8.

[13] Ma X, Tan C, Zhu D, Gang DR. Is there a better source of huperzine A than *Huperzia serrata*? Huperzine a content of *Huperziaceae* species in China. J Agric Food Chem 2005;53:1393—8.

[14] Cuthbertson D, Piljac-Žegarac J, Lange BM. Validation of a microscale extraction and high-throughput UPLC-qTOF-MS analysis method for huperzine a in *Huperzia*. Biomed Chromatogr 2012;26:1191—5.

[15] Takayama H, Katakawa K, Kitajima M, Yamaguchi K, Aimi N. Seven new lycopodium alkaloids, lycoposerramines-C, -D, -E, -P, -Q, -S, and -U, from *Lycopodium serratum* thunb. Tetrahedron Lett 2002;43:8307—11.

[16] Tatsis EC, Ines M, Teto S, De TD, Stavrinides AK, Oudin A, et al. A three enzyme system to generate the strychnos alkaloid scaffold from a central biosynthetic intermediate. Nat Commun 2017;8:316.

[17] Kessner D, Chambers M, Burke R, Agus D, Mallick P. Proteowizard: open source software for rapid proteomics tools development. Bioinformatics 2008;24:2534—6.

[18] Smith CA, Want EJ, O'Maille G, Abagyan R, Siuzdak G, O'maille G, et al. XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. Anal Chem 2006;78:779—87.

[19] Benton HP, Wong DM, Trauger SA, Siuzdak G. XCMS2: processing tandem mass spectrometry data for metabolite identification and structural characterization. Anal Chem 2008;80:6382—9.

[20] Tsugawa H, Cajka T, Kind T, Ma Y, Higgins B, Ikeda K, et al. MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis. Nat. Methods 2015;12:523—6.

[21] Duan L, Molnár I, Snyder JH, Shen GA, Qi X. Discrimination and quantification of true biological signals in metabolomics analysis based on liquid chromatography-mass spectrometry. Mol Plant 2016;16:1217—20.

[22] Uppal K, Walker DI, Jones DP. xMSannotator: an R package for network-based annotation of high-resolution metabolomics data. Anal Chem 2017;89:

[23] Zhou B, Wang J, Ressom HW. Metabosearch: tool for mass-based metabolite identification using multiple databases. PLos One 2012;7.

[24] Qiu F, Fine DD, Wherritt DJ, Lei Z, Sumner LW. PlantMAT: a metabolomics tool for predicting the specialized metabolic potential of a system and for large-scale metabolite identifications. Anal Chem 2016;88:11373—83.

[25] Tsugawa H, Kind T, Nakabayashi R, Yukihira D, Tanaka W, Cajka T, et al. Hydrogen rearrangement rules: computational MS/MS fragmentation and structure elucidation using MS-FINDER software. Anal Chem 2016;88: 7946—58.

[26] Dührkop K, Shen H, Meusel M, Rousu J, Böcker S. Searching molecular structure databases with tandem mass spectra using CSI: FingerID. PNAS 2015;112:12580—5.

[27] Xia J, Wishart DS. Web-based inference of biological patterns, functions and pathways from metabolomic data using MetaboAnalyst. Nat Protoc 2011;6: 743—60.

[28] Seki H, Ohyama K, Sawai S, Mizutani M, Ohnishi T, Sudo H, et al. Licorice β-amyrin 11-oxidase, a cytochrome P450 with a key role in the biosynthesis of the triterpene sweetener glycyrrhizin. PNAS 2008;105:14204—9.

[29] Seki H, Sawai S, Ohyama K, Mizutani M, Ohnishi T, Sudo H, et al. Triterpene functional genomics in licorice for identification of CYP72A154 involved in the biosynthesis of glycyrrhizin. Plant Cell 2011;23:4112—23.

[30] Xu G, Cai W, Gao W, Liu C. A novel glucuronosyltransferase has an unprecedented ability to catalyse continuous two-step glucuronosylation of glycyrrhetinic acid to yield glycyrrhizin. New Phytol 2016;212:123—35.

[31] Maclean DB. Lycopodium alkaloids, in alkaloids: chemistry and pharmacology. 1985. p. 241—98.

[32] Lisec J, Schauer N, Kopka J, Willmitzer L, Fernie AR. Gas chromatography mass spectrometry-based metabolite profiling in plants. Nat Protoc 2006;1: 387—96.

[33] Pluskal TT, Castillo S, Villar-Briones A, Oresic M, Orešič M. mzMine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. BMC Bioinf 2010;11:395.

[34] Huan T, Forsberg EM, Rinehart D, Johnson CH, Ivanisevic J, Benton HP, et al. Systems biology guided by XCMS Online metabolomics. Nat. Methods 2017;14:461—2.

[35] Myers OD, Sumner SJ, Li S, Barnes S, Du X. One step forward for reducing false positive and false negative compound identifications from mass spectrometry metabolomics data: new algorithms for constructing extracted ion chromatograms and detecting chromatographic peaks. Anal Chem 2017;89: 8696—703.

[36] Patti GJ, Tautenhahn R, Siuzdak G. Meta-analysis of untargeted metabolomic data from multiple profiling experiments. Nat Protoc 2012;7:508—16.