Chemical Profiling of Ginseng Species and Ginseng Herbal Products Using UPLC/QTOF-MS

Jimmy Yuk,*a Dhavalkumar N. Patel,a Giorgis Isaac,a Kerri Smith,a Mark Wrona,a Hernando J. Olivos,a and Kate Yu*a

aWaters Corporation, 34 Maple Street, 01757 Milford, MA, USA
bWaters Pacific Pte Ltd., 1 Science Park Road, 02-01/06, The Capricorn, Singapore Science Park II, 117528 Singapore

The chemical profiles of four ginseng roots samples from three species of ginseng (Panax quinquefolius, Panax ginseng and Panax notoginseng) and two commercial ginseng products containing P. quinquefolius and red P. ginseng were compared using ultra-performance liquid chromatography coupled with quadrupole time of flight mass spectrometry (UPLC/QTOF-MS). Principal component analysis allowed a holistic approach in showing distinct chemical differences between the three ginseng species and correct classification of the two commercial products to their respective species. Further investigation of the chemical profile variations yielded ten main markers that were distinct for the three species. This study shows the potential of chemical profiling for the classification of complex natural product samples, such as ginseng, and application to commercial products sold in the market. This methodology can assist the industry in authenticating the various species of ginseng and providing a quick assessment of the quality of commercial ginseng products.

Keywords: metabolic profiling, natural product, mass spectrometry, authentication, LC-MS

Introduction

Ginseng (Panax species, Araliaceae family) has been well-used and studied in traditional Chinese medicine and has been widely adapted in North America as a herbal supplement with high growth in sales every year.1 Ginseng has been known for its therapeutic effects on the endocrine, cardiovascular, immune, central nervous system and also reported to possess anti-cancer properties.2,3 Panax ginseng C. A. Meyer (Asian or Korean ginseng), Panax quinquefolius L. (American ginseng) and Panax notoginseng (Burk) F. H. Chen (Tienchi or Sanchi, notoginseng) are the most commonly used and studied ginseng herbs.4 Due to the continual increases in sales, there has been an increase in concern of adulteration of herbal products sold in the market.4 Practices including misidentification of herbal species and the addition of unlabeled pharmaceuticals5 can cause harm to the consumer and potential mistrust in the herbal and traditional medicine industry. Specifically for ginseng, intentional substitution of the wrong species in ginseng products in the market is a known issue.6 To address this growing concern, many analytical methods have been developed to authenticate and classify the various species of ginseng.

Morphological examinations7 to genetic testing8 has been commonly used to authenticate different species of ginseng. However, certain products in the market can be packaged ground into powder or prepared differently (extracts) which can decrease the effectiveness of these methods. Chemical methods using analytical instrumentation such as high-performance liquid chromatography (HPLC) with diode array,9,11 nuclear magnetic resonance (NMR) spectroscopy,2,12,13 and liquid chromatography-mass spectrometry (LC-MS)14 have also been used to profile ginseng species. Many of these studies have focused on identifying key chemical components in the various ginseng species for authentication. However, spiking in specific markers to represent a natural product in commercial products could be done and can mislead regulators.15 A newer approach has been the development of untargeted methods to compare whole chemical profiles through the sample’s spectrum or chromatogram using authenticated references.16 This chemical profiling methodology offers many benefits as a targeted list of
key compounds typically found in the natural product are not needed for authentication. Instead, the data analysis using chemometrics takes into consideration the complete chemical signature for comparison. This makes adulterations much more difficult as the assessment for quality is based on the entire chemical constituents of the natural product rather than just a set list of markers. Chemical profiling allows researchers to have a holistic comparison of the samples (authentic vs. commercial) and through multivariate modeling, detect specific chemical signals to focus on which could be indicative of adulteration (i.e., spiking, species substitution, etc.) or quality controls concerns (i.e., growth conditions, environment contaminants, sample preparation, etc.). Chemical profiling is a growing trend and is becoming a popular method for the quality control and standardization of traditional Chinese herbs. One important requirement for untargeted methods is obtaining a well-known source for the authenticated natural products (i.e., arboretum, botanical garden, etc.) which will allow an appropriate reference for comparison. For ginseng, chemical profiling studies have been investigated for various types of species, farms, cultivation ages including effects of processing conditions. Each of these studies utilize an analytical platform (HPLC, LC-MS or NMR) but require multiple informatics platforms to acquire, process, and analyze the data. These approaches of data analysis and results interpretation using multiple software packages is extremely laborious, time consuming and can be harder to replicate in other laboratories.

In this study, four types of ginseng (P. quinquefolius, P. notoginseng, P. ginseng and red P. ginseng) and two commercial ginseng products were analyzed and profiled using ultra-performance liquid chromatography (UPLC) coupled with a quadrupole time-of-flight (QTOF)-MS. UPLC is becoming a common analytical platform for chemical profiling as the higher pressure (12,000-15,000 psi compared to ca. 6,000 psi for HPLC) and smaller particle size (ca. sub 2 µm) provides a much more rapid and efficient separation for complex natural product mixtures. A chemical profiling study by Nordstrom et al. compared a UPLC/QTOF-MS study with HPLC-MS and found a 20% increase in chemical components for a complex biological mixture. A single informatics platform called UNIFI was used for data acquisition, process and analysis (chemometrics and identification of markers). UNIFI has been recently used for the chemical profiling of Tinospora crispa for α-glucosidase and α-amylase inhibitory constituents and for the discrimination of two species of Cynanchum. The purpose of this work is to expand the chemical profiling approach utilizing modern analytical instrumentation such as ultra-performance liquid chromatography quadrupole time of flight mass spectrometry (UPLC/QTOF-MS) and to evaluate its applicability for natural product authentication such as ginseng.

**Experimental**

**Ginseng samples and reagents**

Four authentic ginseng roots (P. quinquefolius, P. ginseng, red P. ginseng and P. notoginseng) were obtained from the American Herbal Pharmacopoeia. Each sample had a certificate of authenticity to confirm their identity: P. quinquefolius (certificate No. 1762), P. ginseng (certificate No. 1764), red P. ginseng (certificate No. 1710) and P. notoginseng (certificate No. 1763). Two commercial ginseng products were purchased from a local nutritional store (Whole Foods, Boston, Massachusetts). The first ginseng product, Pro1, contained ground P. quinquefolius root in a capsule. The North American ginseng root was cultivated in Sandy City, Oregon. Three capsules containing powder Pro1 were randomly selected and combined together to minimize any product variation. The capsule shell was removed before sample extraction. The second ginseng product, Pro2, contained red P. ginseng root extract in a liquid vial. The Asian ginseng root was cultivated from Changbai, China. There were other miscellaneous ingredients that were in Pro2 such as honey, sodium benzoate and potassium sorbate.

For the UPLC-MS analysis, LC-MS-grade acetonitrile, methanol, formic acid and water was purchased from Sigma-Aldrich (St. Louis, MO, USA).

**Ginseng sample preparation**

The authentic ginseng root raw samples were cut into smaller pieces and grounded using pestle and mortar before extraction. Pro1 was in a capsule form and only the content was used for extraction. 0.5 g of Pro1 was dissolved in 10 mL of 70% methanol. Pro2 was in a liquid form and 0.5 g (by weight) was added in 10 mL of 70% methanol. The samples were vortexed for 1 minute and then sonicated for 30 minutes. The samples were vortexed again for 1 minute and then centrifuged for 20 minutes at 13,000 × g. The supernatant was collected and was centrifuged again for 5 minutes to remove any residual particulate matter. The final supernatant was diluted 20 times with 70% methanol before LC-MS acquisition. Blanks were also prepared in the same way which contained only 70% methanol.
UPLC-MS data acquisition

UPLC was performed on a Waters ACQUITY I-Class UPLC (Waters Corporation, Milford, USA). A reverse phase HSS T3 C18 column (2.1 × 100 mm, 1.8 µm particle size) was used and maintained at 45 °C. The mobile phases consisted of A (water in 0.1% formic acid) and B (acetonitrile in 0.1% formic acid). A gradient elution was done at a flow rate of 0.6 mL min⁻¹ with an injection volume of 1 µL. The gradient was as follows: 10 to 25% B (0-2 min), 25 to 35% B (2-4 min), 35% B (4-9 min), 35 to 60% B (9-11 min), 60 to 90% B (11-15 min), 90% B (15-18 min) and then 2 minute equilibration to 2% B (18-20 min).

The MS acquisition was performed on Waters Xevo G2-XS QTOF (Waters Corporation, Milford, USA) using electrospray ionization (ESI) in both positive and negative modes. In this study, only the positive mode data will be used for data analysis. Data acquisition was controlled with UNIFI 1.8 informatics platform (Waters Corporation, Milford, USA). The data range was from 100-1500 Da. Leucine enkephalin (200 pg µL⁻¹) was used as the lock mass (m/z 556.2766). The source temperature applied was 120 °C and the desolvation gas flow was at 1000 L h⁻¹ with a desolvation temperature of 550 °C. The capillary voltage was at 2.5 kV and the cone voltage was 30 V. Alternative low and high collision energy scans were used to obtain data independent acquisition (known as MS⁸ mode). The low collision energy scan was set at 6 eV and the high collision energy scan was set at a ramp energy scan from 25 to 55 eV. The scan time for each function was 0.10 s.

Three injections were performed for each sample including blanks and were randomized before data acquisition.

Data and statistical analysis

All LC-MS data was processed, peak picked and analyzed using UNIFI informatics platform. A 3 dimensional (3D) peak detection algorithm was used to detect the peak apexes of all the ion responses based on their 3D shapes to obtain cleaner spectra and more accurate peak volumes than 2D extracted ion chromatograms. The total intensity of each ion was normalized to the total ion count to generate a matrix consisting of the m/z value, retention time (RT) and normalized peak area. The data matrix was used for multivariate statistical analysis through the EZinfo software 3.03 (Umetrics, Umea, Sweden). The data was mean-centered and Pareto-scaled for principal component analysis (PCA). The variables of interest were then identified using the discovery tools within UNIFI software. The discovery tool in UNIFI connects directly to Waters traditional medicine library.

Results and Discussion

The LC-MS chromatograms of all the authentic and commercial ginseng samples are shown in Figure 1.

Using a single extraction procedure of 70% methanol for all samples combined with an optimal UPLC method, many of the peaks were well separated in a 14 minute run. From the authentic ginseng species (red P. ginseng, P. notoginseng, P. quinquefolius), each chromatogram (Figures 1A-D) shows a unique chemical profile with similarities and differences across their peak profiles. The commercial samples (Pro1 and Pro2) had characteristic chemical profiles (Figures 1E-F) with more peaks than the authentic samples. With only visual comparison of the samples, it is difficult to fully understand the variation in their chromatograms. Instead, multivariate statistical modeling such as principal component analysis (PCA) was used to understand the similarities between the samples and identify the potential chemical markers of interest that classify them from each other.

In Figure 2, a PCA scores plot was created to classify all the authentic and commercial ginseng samples.

In a PCA scores plot, each sample is represented by a symbol and their distances from each other reflect how closely related their chemical profiles are. The variables used for comparison are the exact mass to retention time pair (EMRT) and normalized peak areas. PCA is considered an unbiased model, as it does not use any class discrimination during the calculation and results can show unique groupings and interesting outliers of complex datasets. From the scores plot (Figure 2), the first two principal components were used which encompassed 71.6% of the total explained variance. There was excellent reproducibility for all the samples with tight groupings of the authentic and commercial samples. For the authentic ginseng samples, there was a clear separation between P. quinquefolius (PQ), P. notoginseng (PN) and the two P. ginseng types, red (RPG) and normal (PG). Each species group was situated in their own quadrant in the scores plot. The RPG and PG samples were closely associated with each other in the PCA scores which show similarities between the two types of samples. RPG and PG have the same plant origins but are different due to sample processing. RPG is prepared by using fresh PG roots and then steamed for a period of time before drying. Overall, this process changes the RPG’s chemical profile and biological activity. The tight association of the two preparations of ginseng is mainly due to the vast profile differences from the other two species, PQ and PN, which forces the two
P. ginseng groups to be closer together. When the two PG groups were modeled separately (not shown), there was a clear separation between the two preparations of ginseng. For this study, the further classification of RPG and PG was not considered, as the objective was to understand their distinction with the other species, PQ and PN and also the commercial products. If readers are interested in their classification, a previous study by Zhang et al. described the differences in chemical profiles of RPG and PG. For Pro1, which contains PQ roots, the samples were closely grouped to the authentic PQ samples. Pro2, which contains red PG root extract, was also closely associated with the authentic PG groups. The correct classification of the commercial ginseng products with the authentic ginseng species show the high potential of using chemical profiling as a powerful verification of complex natural products when compared to authentic samples.

To understand the specific chemical variations causing the profile differences of the ginseng samples in the PCA scores plot, a loadings plot (Figure 3) was generated.

In the loadings plot, each point represents the EMRT pair and complements the information in the scores plot by showing which variables were important in differentiating the sample groupings. In Figure 3, the three distinct quadrants that classified the three ginseng species, PQ, PN and PG, in the scores plot have many variables (m/z) associated in that area. The other variables which were mainly located in the center of the plot are considered ones that are not significantly different in all the samples and thus were not analyzed further.

The next step to further understand the chemical profile differences, the key differentiating variables (EMRT) in each of the quadrants were tagged and linked into the discovery tools in UNIFI which connects to an in-house
traditional medicine library. In Figure 4, all the specific variables in the PCA loadings plot that corresponded to the three different ginseng species were screened. From Waters traditional medicine library which consists of 600 herbs and ca. 6400 compounds, the variables were screened to identify any potential markers. One of the markers, gypenoside XVII, a triterpene glycoside, was identified with a sodium adduct at \( m/z \) 969.5372 at 10.05 minutes. This marker was clearly seen as a distinct variable for PQ in the loadings plot (Figure 3). Gypenoside XVII is a protopanaxadiol-type triterpene saponin in which the hydroxyl groups at positions 3 and 20(S) have been converted to the corresponding \( \beta-D\)-glucopyranoside and \( \beta-D\)-glucopyranosyl-(1\( \rightarrow \)6)\-\( \beta-D\)-glucopyranoside respectively. Gypenoside XVII has been studied as a characteristic marker for the North American ginseng.\textsuperscript{27,28} To further confirm the identification of gypenoside XVII, the discovery tools also creates theoretical fragments for gypenoside XVII based on the susceptibility of bond breakage from the molecular structure. Using the fragment matching function of the discovery tools, twenty fragments for gypenoside XVII were matched in the high energy MS\textsuperscript{3} spectrum. The major fragments observed for gypenoside XVII are formed from the cleavage of glycosidic linkages resulting in losses of sugar residues thus yielding an aglycone moiety and subsequent series of dehydrated ions. Similar pattern of fragmentation is also observed for other identified marker ginsenosides. For illustration purposes, in the mass spectra

Figure 2. The principal component analysis (PCA) scores plot for the authenticated ginseng and commercial ginseng products. The first two principal components are shown. The ellipse represents the 95% confidence interval of the PCA model. The abbreviations in the PCA scores are as follows: RPG: red Panax ginseng; PG: Panax ginseng; PN: Panax notoginseng; PQ: Panax quinquefolius; Pro1: Product 1 and Pro2: Product 2.

Figure 3. The PCA loadings plot for the authenticated ginseng and commercial ginseng products. Each dot represents a \( m/z\)-RT variable.
of gypenoside XVII (Figure 4), structures and likely bond cleavages are shown for five fragment ions, 767.492 \([M + H−\text{Glc}−\text{H}_2\text{O}]\), 605.440 \([M + H−\text{Glc}−\text{Glc}−\text{H}_2\text{O}]\), 587.430 \([M + H−\text{Glc}−\text{Glc}−2\text{H}_2\text{O}]\), 425.377 \([M + H−\text{Glc}−\text{Glc}−\text{Glc}−2\text{H}_2\text{O}]\) and 325.113 (dehydrated \(\beta\)-D-glucopyranosyl-(1→6)-\(\beta\)-D-glucopyranoside residue). Two selected \(m/z\) variables in the marker matrix list at retention time of 10.05 min, 425.3774 and 587.4301 \(m/z\) were detected as confirmed fragments for gypenoside XVII. In the PCA loadings plot, these two fragments were seen for PQ.

The confirmation of the parent and fragment ions from the loadings plot provides a higher confidence for the identification of the marker. Using the workflow described above, ten main compounds were identified that separate the three ginseng species (Table 1). From the table, protopanaxatriol ginsenoside, \(R_f\), protopanaxdiol ginsenosides \(Rb_2\), \(Ra_1\) and \(Ra_2\) were detected as the distinct markers for PG. For PQ, the characteristic markers, psuedoginsenoside F11, ginsenoside \(Rd\) and the previously described gypsenoside XVII were identified. Three related saponins, notoginsenoside \(R1\), \(R4\) and \(Fa\) were detected for PN. Dong et al. conducted an extensive chemical assessment of PN in various cultivations areas of China and described notoginsenoside \(R1\) to be a distinct marker compared to the other two species, PQ and PG.

The European Medicines Agency (EMEA) defines chemical markers as appropriate indicators for the presence of the herbal medicines even if therapeutic activity is not present. The results of this study showed that chemical profiling can be an effective methodology for classifying herbal materials such as ginseng through a non-targeted approach. With appropriate herbal reference materials for comparison, chemical profiling can provide an accurate solution for quick verification of natural products sold in the market. The ginseng sample preparation, data acquisition (14 minutes) and analysis using UPLC/QTOF-MS with a fast, streamlined informatics workflow enabled high sample throughput for large sample batches. However, one caveat for chemical profiling of natural products is the standard operating procedure (SOP) for every sample. The slight changes in sample preparation to variations in parameters for data acquisition in a study can play a large role in the accuracy of the results. The difference in extraction solvents and techniques can cause a vast difference in the chemical profile of herbal materials and should be carefully assessed before starting any profiling study. To ensure proper comparison of sample batches, it is recommended to standardize the SOP. This will allow studies to be compared over time and the sharing of results to other laboratories that follow the same SOP.

Figure 4. The discovery tools in UNIFI showing an identified marker, gypenoside XVII from the list of selected variables from the PCA loadings plot. The chemical information and molecular structure of gypenoside XVII from an in-house traditional medicine library is also shown. A mass spectrum of ginseng is also displayed with all the theoretical fragments for gypenoside XVII identified on the spectrum.
Table 1. The list of chemical markers that were identified from the PCA loadings plot using UNIFI. The standard deviation was calculated to show variation in the average

<table>
<thead>
<tr>
<th>No.</th>
<th>RT / min</th>
<th>Compound name</th>
<th>Molecular formula</th>
<th>Expected neutral mass / Da</th>
<th>Observed neutral mass / Da</th>
<th>QTOF-MS (ESI+) m/z</th>
<th>Mass error / mDa</th>
<th>Theoretical fragment detected</th>
<th>Adduct</th>
<th>Species detected</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.23</td>
<td>ginsenoside Rf</td>
<td>C_{42}H_{72}O_{34}</td>
<td>800.49221</td>
<td>800.49077 ± 0.00059</td>
<td>823.47996 ± 0.00059</td>
<td>1.4 ± 0.6</td>
<td>22</td>
<td>+Na</td>
<td>P. ginseng</td>
</tr>
<tr>
<td>2</td>
<td>6.41</td>
<td>ginsenoside Rb2</td>
<td>C_{42}H_{72}O_{34}</td>
<td>1078.59237</td>
<td>1078.59031 ± 0.00091</td>
<td>1101.57956 ± 0.0009</td>
<td>2.0 ± 0.9</td>
<td>13</td>
<td>+Na</td>
<td>P. ginseng</td>
</tr>
<tr>
<td>3</td>
<td>5.78</td>
<td>ginsenoside Ra1/Ra2</td>
<td>C_{42}H_{72}O_{34}</td>
<td>1210.63643</td>
<td>1210.6313 ± 0.0015</td>
<td>1233.6205 ± 0.0015</td>
<td>3.4 ± 1.5</td>
<td>17</td>
<td>+Na</td>
<td>P. ginseng</td>
</tr>
<tr>
<td>4</td>
<td>3.20</td>
<td>ginsenoside Re</td>
<td>C_{42}H_{72}O_{34}</td>
<td>946.55012</td>
<td>946.54823 ± 0.00068</td>
<td>969.5374 ± 0.00068</td>
<td>1.9 ± 0.7</td>
<td>18</td>
<td>+Na</td>
<td>P. quinquefolius</td>
</tr>
<tr>
<td>5</td>
<td>3.60</td>
<td>pseudoginsenoside F11</td>
<td>C_{42}H_{72}O_{34}</td>
<td>800.49221</td>
<td>800.49013 ± 0.00071</td>
<td>801.4974 ± 0.00070</td>
<td>2.1 ± 0.7</td>
<td>4</td>
<td>+H</td>
<td>P. quinquefolius</td>
</tr>
<tr>
<td>6</td>
<td>5.35</td>
<td>ginsenoside Rd</td>
<td>C_{42}H_{72}O_{34}</td>
<td>946.55012</td>
<td>946.54897 ± 0.00095</td>
<td>947.5652 ± 0.00068</td>
<td>1.1 ± 0.5</td>
<td>23</td>
<td>+H, +Na</td>
<td>P. quinquefolius</td>
</tr>
<tr>
<td>7</td>
<td>10.05</td>
<td>gypenoside X VII</td>
<td>C_{42}H_{72}O_{34}</td>
<td>946.55012</td>
<td>946.54787 ± 0.00025</td>
<td>969.5371 ± 0.00027</td>
<td>2.2 ± 0.2</td>
<td>20</td>
<td>+H, +Na</td>
<td>P. quinquefolius</td>
</tr>
<tr>
<td>8</td>
<td>2.98</td>
<td>notoginsenoside R1</td>
<td>C_{42}H_{72}O_{34}</td>
<td>932.53447</td>
<td>932.53257 ± 0.00093</td>
<td>955.5218 ± 0.00099</td>
<td>1.9 ± 1.0</td>
<td>8</td>
<td>+Na</td>
<td>P. notoginseng</td>
</tr>
<tr>
<td>9</td>
<td>4.62</td>
<td>notoginsenoside R4</td>
<td>C_{42}H_{72}O_{34}</td>
<td>1240.64520</td>
<td>1240.64237 ± 0.00100</td>
<td>1263.632 ± 0.00095</td>
<td>2.8 ± 1.0</td>
<td>11</td>
<td>+H, +Na</td>
<td>P. notoginseng</td>
</tr>
<tr>
<td>10</td>
<td>4.89</td>
<td>notoginsenoside Fa</td>
<td>C_{42}H_{72}O_{34}</td>
<td>1240.64520</td>
<td>1240.64273 ± 0.00038</td>
<td>1263.632 ± 0.00032</td>
<td>2.5 ± 0.4</td>
<td>15</td>
<td>+Na</td>
<td>P. notoginseng</td>
</tr>
</tbody>
</table>

The average value of all the samples for the specific ginseng samples; RT: retention time; ESI: electrospray ionization; QTOF-MS: quadrupole time-of-flight mass spectrometry.

Conclusions

The application of chemical profiling using UPLC/QTOF-MS on three ginseng species (Panax quinquefolius, Panax ginseng and Panax notoginseng) and two commercial ginseng products containing P. quinquefolius and red P. ginseng was examined. The multivariate statistical analysis was successful in discriminating among three ginsen species and correctly classifying the commercial products to the proper standard reference. Ten main compounds were identified when the chemical signals from the model were screened using an in-house traditional medicine library. All the compounds were considered distinct markers for each of the ginseng species. Overall, this non-targeted approach with an integrated and rapid workflow shows high potential for the assessment of complex natural products such as ginseng and their commercial products.

Acknowledgments

The authors thank Sukhdev Bangar for technical assistance and valuable discussions.

References


Submitted: January 25, 2016
Published online: June 21, 2016