7-2013

Quantitative comparison of multiple components in dioscorea nipponica and D. panthaica by ultra-high performance liquid chromatography coupled with quadrupole time-of-flight mass spectrometry

Yina Tang
Hong Kong Baptist University

Tao Yi
Hong Kong Baptist University, yitao@hkbu.edu.hk

Homing Chen
Hong Kong Baptist University

Zhongzhen Zhao
Hong Kong Baptist University, zzzhao@hkbu.edu.hk

Zhitao Liang
Hong Kong Baptist University

Quantitative Comparison of Multiple Components in *Dioscorea nipponica* and *D. panthaica* by Ultra Performance Liquid Chromatography Coupled with Quadrupole Time-of-Flight Mass Spectrometry

Yina Tang, Tao Yi*, Homing Chen, Zhongzhen Zhao, Zhitao Liang and Hubiao Chen*

* Correspondence to:

Hubiao Chen, School of Chinese Medicine, Hong Kong Baptist University, Kowloon, Hong Kong Special Administrative Region, People's Republic of China. E-Mail: hbchen@hkbu.edu.hk

Tao Yi, School of Chinese Medicine, Hong Kong Baptist University, Kowloon, Hong Kong Special Administrative Region, People's Republic of China. E-Mail: yitao@hkbu.edu.hk

ABSTRACT

Introduction – *Dioscorea nipponica* (DN) and *D. panthaica* (DP) have been uniquely prepared as herbal medicinal products for treating coronary heart disease (CHD) in China. However, so far there has been little discussion and no work comparing the qualitative and quantitative differences between the two herbs nor assessing whether they have similarities in chemical composition that would support their common application for treating CHD.

Objective – To develop an efficient and reliable method based on UPLC-qTOF-MS for quantitative comparison of saponins in both DN and DP.

Methodology – Using electrospray ionization and atmospheric-pressure chemical ionization respectively, six steroidal glycosides and one aglycone were determined in 13 DN samples and 13 DP samples. The comparative analysis of chemical components in DN and DP was carried out by chromatographic fingerprint similarity evaluation, test of significance (*t*-test) and principle component analysis (PCA).

Results – The UPLC-qTOF-MS method showed limit of detection and quantitation within the range 0.02-0.2 ng and 0.08-0.5 ng, respectively, for the 7 studied saponins. The intra- and inter-day precision (RSD) was below 5%. The recoveries for the quantified compounds were within the range of 72.79-118.31%.

Conclusion – This UPLC-qTOF-MS assay provides a suitable method for the identification and determination of major bioactive constituents both in DN and DP. The chemical composition of all DN and DP samples studied exhibited a high level of global similarity. This chemical similarity validates their common application in the pharmaceutical industry as anti-CHD herbal drugs.

Keywords:

Dioscorea nipponica, Dioscorea panthaica, steroidal saponins, principal component analysis, chromatographic fingerprint similarity, UPLC-qTOF-MS

Introduction

Plants of the *Dioscorea* genus (family Dioscoreaceae) with significant edible or medicinal values have long been the research focus of both the food and pharmaceutical industries. Known as yam, the roots and tubers of several edible species of *Dioscorea*, such as *D. opposita*, *D. alata*, and *D. japonica*, are consumed as a staple food by the populations of Africa, Asia and tropical America (Adedayo *et al.*, 2011). Medicinal species of *Dioscorea* have a wide spectrum of biological activities, for instance, anti-inflammatory, hormone-like, antitumor, immunoregulatory, hypoglycemic, and cardiovascular properties. As such, they have received a great deal of attention for the past several decades (Sautour *et al.*, 2007). In the modern pharmaceutical industry, it is well known that a few species of *Dioscorea*, such as *D. zingiberensis*, *D. nipponica* (DN), *D. panthaica* (DP) (Editorial Board of Flora of China, 1985), are important sources of diosgenin, which is used as the precursor of synthetic steroidal anti-inflammatory and hormone-based drugs. In addition, investigations into the cardiovascular activity of components of *Dioscorea* species have led to the successful development of effective pharmaceutical preparations that have been used for decades for treating cardiovascular diseases. As early as in 1975, the preparation named 'polysponin' obtained from DN was approved by the Soviet Union's Ministry of Health to be used as an anti-atherosclerosis agent and cholesterol-lowering drug in clinical practice (Leskov *et al.,* 1976; Solomonova, 1968). In China, several patented medicines indicated for myocardial ischemia or angina pectoris, such as Dioscornin Tablet (Sichuan Biology Research Institute. 1977) and Di'ao Xinxuekang Capsule (Chinese Pharmacopeia Committee, 2010), were developed during the 1970s and 1980s, and have been on the market since then; the bioactive components of these medicines are the total saponins extracted from DN or DP.

However, as far as we know, among the 49 species of the genus *Dioscorea* distributed in China (Editorial Board of Flora of China, 1985), only two have been recorded as crude drugs in the Chinese Pharmacopoeia and particularly prepared as

medications for treating CHD (Chinese Pharmacopeia Committee, 2010). These are DN and DP, the rhizomes of which are named Chuanshanlong and Huangshanyao, respectively, in folk medicine. It is also noteworthy that DN and DP are equally used as starting material for Di'ao Xinxuekang Capsule, as documented in the Chinese Pharmacopoeia. Interestingly, the capsule was the first Chinese patented medicine authorized as a traditional Herbal Medicinal Product (tHMP) by the Dutch Medicine Evaluation Board in March, 2012 (Dutch Medicines Evaluation Board, 2012). Some individual research, mainly focusing on discovery of bioactive constituents contained in DN or DP, has been done (Dong *et al.* 2001, 2004; Kang *et al.* 2005) with the result that, up to now, a single or simultaneously two to four bioactive markers have been found in quantitative assays of DN or DP using HPLC-UV (Qin *et al.* 2007) and HPLC-ELSD (Liu *et al.* 2008; Shen *et al.* 2011) methods. However, the UV detection wavelength for steroidal saponins is susceptible to interference from the cut-off absorption wavelength of solvents in the mobile phase, such as methanol and acetonitrile, and this could affect the results of quantitative analysis. Although ELSD can avoid the interference caused by the wavelength of end absorption, its lower sensitivity means it might be unable to quantify trace saponins. Thus, attempts to establish a more sensitive and accurate method for simultaneous quantification of multiple saponins in DN or DP have met with little success. Equally important, there have been no attempts to make detailed comparisons of the qualitative and quantitative differences between the two herbs nor to determine whether they have similarities in chemical composition that support their common application in anti-CHD drugs. In recent years, chemometrics coupled with liquid chromatography and high resolution mass spectrum fingerprinting has become one of the most frequently applied approaches in evaluation of chemical profiles of different phytomedicines (Yi *et al.* 2009; Avula *et al.* 2011). This is a promising approach to clarifying this unsolved problem. Compared to conventional HPLC-UV and HPLC-ELSD, ultra high performance liquid chromatography coupled with quadrupole time-of-flight mass spectrometry (UPLC-qTOF-MS) with its superior performances in sensitivity and selectivity has been utilized in research on some

Dioscorea species, such as *D. zingiberensis*. These studies, however, focused on structural elucidation and qualitative identification of steroidal saponins and provided little quantitative information that could be used in comparison of detected components in different herbs (Li *et al.* 2006; Zhu *et al.* 2010).

This paper aimed to develop a reliable and effective protocol for the quantification of the major constituents, including six glycosides and one aglycone, contained in DN and DP. The protocol entailed UPLC-qTOF-MS followed by comparisons of chromatographic fingerprint profiles and evaluation of the contents of determined components using fingerprint similarity evaluation, test of significance and principal component analysis. In the present study, the chemical similarity between DN and DP is evaluated and clarified so as to support their rational use in pharmacology and lay the foundation for future research on the mechanism(s) of their effects on cardiovascular activities.

Experimental

Plant materials

Dioscorea nipponica (DN) samples and *D. panthaica* (DP) samples were collected from different regions of China, the sources of which are listed in Table 1. The plant materials were identified by Dr. Hubiao Chen, School of Chinese Medicine, Hong Kong Baptist University. Corresponding voucher specimens were deposited in the Herbarium Centre, Hong Kong Baptist University.

Insert Table 1 here

Chemicals and reagents

Analytical grade methanol purchased from the RCI Lab-Scan Limited (Bangkok, Thailand) was used for the extraction and preparation of samples. Hydrochloric acid (37% concentration) purchased from the RCI Lab-Scan Limited (Bangkok, Thailand) and chloroform (with ethanol) purchased from the Anaqua Chemical Supply (Houston,

USA) were used for the acid hydrolysis and partition of samples. Water purified using a Milli-Q water system (Millipore; Bedford, MA, USA) and acetonitrile (Fisher Scientific, New Jersey, USA) were used as the mobile phase for analysis. Formic acid (Sigma-Aldrich, USA) was added to the mobile phase for determination of glycosides.

Seven reference standards were used for qualitative and quantitative analysis, namely six glycosides and one aglycone of diosgenin (Figure 1). The authentic standards (purity > 98%) of protodioscin, dioscin, gracillin and diosgenin were purchased from Phytomarker Ltd., Tianjin. Protogracillin and polyphyllin V (purity > 98%) used in this study were purchased from Emilion Technology Ltd., Beijing. Pseudoprotodioscin was provided by National Institute for Food and Drug Control (Beijing, China).

Insert Figure 1 here

UPLC-qTOF-MS analysis

Ultra high performance liquid chromatography with ultra-high definition accurate mass quadrupole time-of-flight mass spectrometry (Agilent Technologies, G6540A) was used for qualitative and quantitative analyses. Chromatographic separation was conducted on a Waters ACQUITY^{TM} UPLC BEH C_{18} column (1.7 μm, 2.1 × 100 mm). The MS spectra were acquired in negative ion mode for detecting glycosides and positive ion mode for detecting aglycones. High-purity nitrogen was used as nebulizer and auxiliary gas.

For the determination of glycosides, separation was carried out in a linear gradient program, in which the mobile phase consisted of 0.1% (v/v) formic acid in water (A) and 0.1% (v/v) formic acid in acetonitrile (B). The mobile phase was programed as follows: 0-2 min, 20-20% B; 2-12min, 20-28% B; 12-20 min, 28-45% B; 20-35min, 45-48% B. The flow rate was 0.4 mL/min. The Injection volume was 1 μL and the column temperature was set at 25^{o}C. The mass scan was over the range of *m/z* 50-1700 for negative ion mode with electrospray ionization (ESI) interface. The ESI

capillary voltage was set at -3.5kV for negative mode.

For the determination of aglycones, water (A) and methanol (B) were used as the mobile phases. A linear gradient program was developed as follows: 0-10 min, 55-95% B; 10-18 min, 95-95% B. The flow rate was 0.45 mL/min. The injected volume was 1 μL and the column temperature was set at 50^{o}C. The mass scan was over the range of *m/z* 100-1000 for positive ion mode with an atmospheric-pressure chemical ionization (APCI) source. The capillary voltage was set at +4.5kV. All the results were analyzed by Aglient MassHunter Workstation Qualitative Analysis B.04.00 and Quantitative Analysis (Q-TOF) B.04.00.

Sample preparation

Plant materials were cut into small pieces and mixed thoroughly. A representative portion of each of the samples was ground into a powder that passed through a 250 mesh (55μm) sieve. For the preparation of test solutions used to determinate steroidal glycosides, an accurately weighed portion of 0.5 g was ultrasonicated with 15 mL methanol at room temperature for 30 min. This step was then repeated twice for complete extraction. The residue was washed by an additional 5 mL of methanol. The total extracts were combined in a 50 mL volumetric flask and the volume was made up to the calibration mark with methanol. The extracts were then filtered through a syringe filter (0.2 μm). An aliquot of 1 μL solution was injected for UPLC-MS analysis. Three replicates were prepared for each sample.

For preparation of sample solutions used to determinate steroidal aglycones, the assay was primarily based on the procedures described in the monograph Nippon Yam Rhizome of Chinese Pharmacopeia (Chinese Pharmacopeia Committee, 2005) with minor modification. The sample solution of 25 mL for glycoside assay was transferred into a round-bottomed flask and evaporated to dryness by rotary evaporation under vacuum at 60^{o}C. 20 mL hydrochloric acid solution (3 mol/L) was added to the residue, heated to hydrolyze on a water bath for 30 minutes, cooled, and washed with 10 mL of chloroform each time for three times. After the combined mixture was extracted

and partitioned, the lower layer (chloroform layer) was collected and the upper layer was extracted with an additional 30 mL of chloroform once. The combined chloroform layer was evaporated to dryness and the residue was dissolved by an appropriate amount of methanol, transferred to a 25 mL volumetric flask and the volume was made up to the calibration mark with methanol. The extracts were then filtered through a syringe filter (0.2 μm). An aliquot of 1 μL solution was injected for UPLC-MS analysis.

Preparation of reference standard solutions

Stock solutions of reference standards containing either one aglycone (diosgenin) or six mixed glycosides, namely protodioscin, protogracillin, pseudoprotodioscin, dioscin, gracillin, and polyphyllin V, was prepared separately. The stock solutions of standards (0.5 mg/mL) were prepared in methanol and stored in the refrigerator. The working solutions were prepared by appropriate dilution of the stock solutions with methanol to yield seven concentrations, from 0.5 mg/L to 40 mg/L.

Assay validation

Linearity, precision and recovery were all carried out to validate the method, according to the International Conference on Harmonization (ICH) Guidelines (ICH, 2005). Prepared reference standard solutions with different concentrations were analyzed, and then the calibration curves were constructed by plotting the peak areas versus the concentrations of each analyte. Linearity was evaluated by squared linear correlation coefficient (R^2) of the calibration curves of each reference standard (Table S1 in the online supporting material). Precision and recovery studies were performed on the previously quantified sample of DP-02 with a representative chromatographic profile, and the relative standard deviation (RSD) was taken as a measure of repeatability and precision. Instrument precision was determined by analyzing known concentrations of the seven analytes in the sample solution of DP-02 by consecutive

injection seven times in a single day (Table S2 in the online supporting material). To confirm repeatability, triplicate samples prepared from DP-02 were analyzed once daily for intra-day RSD and on three successive days for inter-day RSD (Table S2). For the recovery test (Table S3 in the online supporting material), three spike levels were set as 50%, 100% and 200% of each reference standards in DP-02. The standards were spiked into 0.5 g of plant material as appropriate concentrations of standard solutions. The samples were then extracted and analyzed with the described method. The average percentage recoveries were evaluated by calculating the ratio of detected amount versus the added amount.

Results and Discussion

Optimization of extraction conditions

Various extraction methods (e.g reflux, sonication), solvents (e.g. different concentrations of methanol or ethanol), and times of extraction were evaluated to obtain optimized extraction efficiency. The results demonstrated that there was no significant difference in the yield of analytes between sonication and reflux, and exhaustive extraction was achieved by sonication with methanol three times for 30 min each time.

Optimization of UPLC-MS condition for detecting glycosides and aglycones

The analytical condition was optimized mainly based on peak resolution, baseline, and elution time. The mobile phase (methanol-water and acetonitrile-water) and column temperature (25 ^{o}C, 40 ^{o}C or 50 ^{o}C) and column mode (BEH C_{18} and HSS C_{18}), ion source (ESI and APCI) and ion mode (negative and positive) were examined and compared. Acetonitrile was preferred over methanol as the mobile phase because its use resulted in an improved separation and a significantly reduced column pressure. The column temperature was optimized at 25 ^{o}C for detecting glycosides and 50 ^{o}C for

detecting aglycones. The column BEH C_{18} was preferred to HSS C_{18} as the former gave better peak resolution. For the selection of ion source, as it was observed that the few aglycones were transformed to ion fragments under the ESI source in our pre-test. Thus APCI, which provided higher fragmentor voltage, was selected for detecting aglycones. Both negative and positive ion mode were tried under the ESI and APCI conditions. Comparing the peak purity and number of characteristic peaks acquired under different ion modes, the negative mode was chosen for testing glycosides while the positive mode was chosen for testing aglycones.

Assay validation

The calibration curve of each analyte showed good linearity with $R^2 > 0.99$ in a wide range of concentrations. The relative standard deviation (RSD) values for precision were in the range of 0.54-4.96% for intra-day assays and 1.33-3.73% for inter-day assays, while the RSD value for instrument precision was within 3% (Table S2). These values showed that the system was suitable for the chemical analysis of the two herbs. The average recovery at three spiked levels was calculated for the evaluation of method accuracy. The recovery of the method was in the range of 74.14-118.31%, with RSD < 6% as shown in Table S3.

Limit of detection (LOD) and limit of quantification (LOQ) were defined as a signal-to-noise ratio (SNR) equal to 3 and 10, respectively. Based on stepwise dilution of standard solutions and calculation of the corresponding SNR, LOD and LOQ were determined within the range 0.02- 0.2ng and 0.08-0.5ng, respectively.

Qualitative analysis

Identification of glycosides and aglycones by known reference standards

By comparing the retention time and accurate mass data between samples and reference standard solutions (Table 2) and the widely accepted accurate threshold for confirmation of elemental composition established at 5 ppm (Zhu *et al.,* 2010), six

characteristic peaks of glycosides were identified in both DN and DP sample solutions as protodioscin (1), protogracillin (2), pseudoprotodioscin (3), dioscin (5), gracillin (6) and polyphyllin V (7) (Figure 2), and one peak of aglycone was identified as diosgenin (10) in both hydrolyzed DN and DP sample solutions (Figure 3). Base peak chromatograms (BPC) of all unhydrolyzed samples of DN and DP are shown in Figure 4 and Figure 5, respectively.

Insert Table 2 here

Insert Figure 2, Figure 3, Figure 4 and Figure 5 here

In addition, as trillin was previously discovered in DN (Wang *et al.*, 2012) and other aglycones, such as ruscogenin, tigogenin and sarsasapogenin, were found in other plants rich in sapogenins (Kostova and Dinchev, 2005), four additional reference standards were used to identify whether these constituents were found in the samples of DN and DP collected in our study. As seen from Figure 2 and Figure 3, no corresponding peaks of trillin (peak 8), ruscogenin (peak 9), tigogenin and sarsasapogenin (peak 11) were observed in all samples of DN and DP. Thus, diosgenin was found to be the sole aglycone in DN and DP. It was noted that as tigogenin and sarsasapogenin were diastereoisomers, as they had the same retention time and overlapped at peak 11.

Qualitative identification of two unknown compounds

The fragments of steroidal saponins revealed the characteristic cleavage of glycosidic bonds, and the fragmentation pattern provided detailed structural information about the sequence of sugars. Similar MS/MS behavior of this type of saponin in UPLC-qTOF-MS/MS has been reported in the literature (Zhu *et al.*, 2010). To our knowledge, the common hexose and 6-deoxyhexose present in steroidal saponins from Dioscoreaceae are glucose and rhamnose, respectively, and generally, glucosyls are connected with the hydroxyl groups at C-3 and/or C-26 positions of steroidal aglycones.

Peak 4 in the total ion chromatogram (TIC) of DP-8 was highlighted by extracted ion chromatogram (EIC). Peak 4 readily produced a strong formate molecular adduct $[M+HCOO]^-$ (*m/z* 1091.5280) and a weak deprotonated molecular ion $[M-H]^-$ (*m/z* 1045.4865) in negative mode of ESI-MS. The ion *m/z* 1091.5280 was further targeted for MS/MS analysis under collision voltages of 40 eV and 60 eV. Figure 6 explains the cleavage path of the sugar chain from the parent ion (*m/z* 1091.5280), in which *m/z* differences in the experimental values and calculated values of the glucosyl and rhamnosyl are shown. Results suggest that the unknown compound is a steroidal saponin containing three glucosyls and one rhamnosyl.

Insert Figure 6 here

According to the formula calculator of Agilent MassHunter Workstation Qualitative Analysis software (B.04.00), the most likely elemental composition formula of the unknown compound of peak 4 is $C_{51}H_{82}O_{22}$, the calculated molecular weight of which would be 1046.5298. Refering to the "substances exploration" in the SciFinder® Database, based on the retention time and reported structures in the literature (Liu *et al.*, 2007), peak 4 was tentatively identified as pseudoprotogracillin.

For identification of the unknown peak 12 in Figure 3, the mass spectrum of peak 12 (Figure 7) indicates that the fragment ion at *m/z* 397.3076 was produced by the loss of one water molecule from the protonated molecular ion of diosgenin $[M+H]^+$ (*m/z* 415.3201). Another fragment ion at *m/z* 253.1914 could be explained as loss of one neutral fragment $C_8H_{16}O_2$ from the ion at *m/z* 397.3056 in the positive ion mode of APCI-MS (Zhu *et al.,* 2010). Taking into account the condition of acid hydrolysis, peak 12 was tentatively identified as 3,5-deoxytigogenin, most likely an artifact formed during the hydrolytic cleavage of diosgenin glycosides (Tsukamoto *et al.,* 1957).

Insert Figure 7 here

Quantitative analysis for glycosides and aglycones

Altogether 13 DN samples and 13 DP samples collected from various regions were analyzed by the developed method. From these samples, six glycosides, namely, protodioscin, protogracillin, pseudoprotodioscin, dioscin, gracillin, polyphyllin V, and one aglycone, namely diosgenin, were quantified. The results are shown in Table 3.

Insert Table 3 here

The results show that the highest and the second highest amounts of components were dioscin and protodioscin, respectively, in 10 of 13 batches of DN. For all DP samples, the amounts of protodioscin and protogracillin were in close proportion, ranking either the first or the second highest content in 6 of 13 batches of DP. The content of polyphyllin V was significantly lower than other determined components both in DN and DP samples.

Comparative analyses for chemical composition of DN and DP

Chemical composition of DN and DP was compared in terms of overall chromatographic profile and in terms of quantified constituents. Chromatographic fingerprint similarity was qualitative and quantitative evaluated on the basis of both major and minor peaks detected in chromatograms of DN and DP samples using a Similarity Evaluation Systems for Chromatographic Fingerprint software (Zhejiang University, Version 1.0.1). Test of significance performed by SPSS 16.0 software was used to infer whether the mean contents of individual components could be considered statistically significant between two groups of independent DN and DP samples. Principal component analysis was useful in regression analysis to mitigate the problem of multicollinearity and to explore the relationships among the independent variables (seven determined constituents). This allowed the identification of the primary predictors with minimal multicollinearity, which was accomplished using SIMCA-P 12.0 software.

Evaluation of chromatographic fingerprint similarity

Similarity tests were performed based on the relative retention times and relative peak areas shown in Figure 4 and Figure 5 (Tistaert *et al.*, 2011). The average chromatogram from the 13 batches of DN samples was chosen as the consensus fingerprint. Peak 1 (protodioscin) and peak 5 (dioscin) were selected as the markers by which the peaks and retention times of the other chromatograms were adjusted. In comparison with the consensus fingerprint, all DN samples showed a similarity of at least ≥ 0.95 except the sample DN-13 for which the similarity was 0.89, while a majority of DP samples showed a similarity higher than 0.85, among which 9 samples demonstrated a similarity higher than 0.95, and only the samples of DP-04 and DP-12 showed an exceptionally low similarity of 0.50 and 0.57, respectively (Table 1). Examining the BPC data of samples DP-04 and DP-12 in detail, it was noted that their peak areas of dioscin and gracillin were approximately 1/20 and 1/10 of the corresponding average value of DN, while, moreover, neither of them contained polyphllin V. These anomalies were likely the source of dramatically low similarity with the consensus chromatogram. In general, the two sets of fingerprints revealed a high level of similarity for all main components of DN and DP.

Tests of significance

In addition to comparison of chromatographic profiles of all DN and DP samples, the amounts of seven specific constituents were also compared in both herbs. By independent-samples *t*-test (Table 3), the average content of protodioscin ($P < 0.05$), dioscin ($P < 0.05$) and diosgenin ($P < 0.05$) in DN and DP were significantly different, except protogracillin ($P > 0.05$) and gracillin ($P > 0.05$). The determined values concerning pseudoprotodioscin and polyphyllin V in all samples of DN and DP were subjected to the nonparametric test of two-independent samples due to their inconformity of normality and homogeneity of variance. Results indicated that the average contents of pseudoprotodioscin ($P > 0.05$) in DN and in DP did not differ significantly, while the average contents of polyphyllin V ($P < 0.05$) did differ signficantly in these same samples. Overall, the contents of the three major

constituents, namely protogracillin, gracillin and pseudoprotodioscin, were of similar levels while the contents of another three main constituents, namely, protodioscin, dioscin and diosgenin, varied among all samples of DN and DP.

Principal component analysis

Principal component analysis (PCA) was employed here to compare DN and DP samples from different sources. Seven determined constituents were set as seven variables, while 26 batches of samples were set as observations. According to factor loading matrix calculations, the PCA loading plot (Figure 8) indicated that protodioscin, dioscin and gracillin showed the greatest influence on Component 1, Component 2 and Component 3, respectively. The three-dimensional graphics of PCA scores (Figure 9) showed that all samples of DN and DP were clustered in their own small regions except DN-13, but they were both within a larger sphere. DN-13, with the lower chromatographic similarity of 0.89, was distributed outside the sphere due to its large score on t3 caused by its very high content of protodioscin. The t1 scores of the DP-04 sample and DP-12 sample (marked by the arrow in Figure 9) with significantly low chromatographic similarity (0.50 and 0.57) were much lower than other samples; this is mainly caused by their dramatically low content of gracillin. The PCA scores of DN samples from different regions did not vary significantly; similarly, the PCA scores of DP samples did not show significant differences. Thus, the PCA results also verified the findings of evaluation of chromatographic fingerprint similarity that DN-13, DP-04 and DP12 varied noticeably c
ompared with the rest of samples.

Insert Figure 8 and Figure 9 here

Based on the comparative results described above, as for the intra-species difference in chemical composition, it was shown in our present study that the major composition of either DN or DP fluctuated more or less among 13 DN or 13 DP

samples, respectively. As it is well-known that various factors, such as the environment, growth period, date of collection, stress, the status of being cultivated or collected in the wild might influence the variation of the secondary metabolites accumulate in the body of plants, it is presumed that the big chemical differences for sample DN-13, DP-04 and DP-12 in comparison with the rest of the collections were caused probably due to very complex factors as mentioned above. However, these anomalies only account for the minority of total 26 samples. As for the inter-species similarity in chemical composition between DN and DP, it might be inferred from their classification in phytotaxonomy and their growing environment and ecosystem. The genus of *Dioscorea* is sub-classified into five sections, such as Sect. Stenophora Uline, Sect. Combilium Prain et Burkill and Sect. Shannicorea Prain et Burkill. DN and DP both are classified in Sect. Stenophora Uline, which indicates their close genetic relationship. And the growing environment of DN and DP is fairly similar. Both grow in the hillside bush, at the forest edge, or by the roadside on the hill. DN and DP distribute at an altitude of 100-1800 m and 1000-3500 m, respectively (Editorial Board of Flora of China, 1985). Therefore, it is observed that the global similarity in chemical composition is demonstrated both at intra-species and inter-species, and the variations to some extent do exist at intra-species and inter-species.

In conclusion, an efficient and reliable method of UPLC-qTOF-MS was developed for quantitative comparison of six steroidal glycosides and one aglycone of diosgenin in 13 batches of *Dioscorea nipponica* (DN) and 13 batches of *D. panthaica* (DP). The linearity, precision and recovery of this method were fully validated and satisfactorily achieved. Based on the base peak chromatograms acquired and assay results, the chemical profiles of DN and DP were clarified and compared by statistical analyses. Results demonstrate that DN and DP, even from different sources, have a highly similar chemical composition. Thus, it is reasonable to think that, from the aspect of chemical evaluation, DN and DP both can be used as starting material for anti-CHD drugs in the pharmaceutical industry.

SUPPORTING INFORMATION

Supporting information can be found in the online version of this article.

ACKNOWLEDGEMENTS

This research was funded by the Faculty Research Grant of Hong Kong Baptist University. The authors would like to thank Mr. Alan Ho and Mr. Chi Leung Chan (Research and Development Division, School of Chinese Medicine, Hong Kong Baptist University) for technical support with the UPLC-qTOF-MS experiments.

References

Adedayo CB, Oboh G, Ademiluyi AO and Akindaahunsi AA. 2011. Comparative studies on antioxidant properties of some tropical Nigerian yam varieties (*Dioscorea* spp.). *Adv Food Sci* **33**: 28-33.

Avula B, Wang YH, Ali Z, Smillie TJ and Khan IA. 2011. Quantitative determination of triperpene saponins and alkenated-phenolics from *Labisia pumila* using an LC-UV/ELSD method and confirmation by LC-ESI-TOF. *Planta Med* **77**: 1742-1748.

Chinese Pharmacopeia Committee. 2005. *Pharmacopoeia of the People's Republic of China (English version).* Vol. 1. People's Medical Publishing House: Beijing; 262-263.

Chinese Pharmacopeia Committee. 2010. *Pharmacopoeia of the People's Republic of China (English version).* Vol. 1. China Medical Science and Technology Press: Beijing; 157-158, 627-628.

Dong M, Feng XZ, Wang BX, Wu LJ and Ikejima T. 2001. Two novel furostanol saponins from the rhizomes of *Dioscorea panthaica* Prain et Burkill and their cytotoxic activity. *Tetrahedron* **57**: 501-506.

Dong M, Feng XZ, Wang BX, Ikejima T and Wu LJ. 2004. Steroidal saponins from *Dioscorea panthaica* and their cytotoxic activity. *Pharmazie* **59**: 294-296.

Dutch Medicines Evaluation Board. 2012. First Authorisation of Traditional Herbal Medicine from outside the European Union. European Community: The Hague. URL[http://www.cbg-meb.nl/CBG/en/human-medicines/actueel/First_Authorisation_of_Traditional_Herbal_Medicine_from_outside_the_European_Union/default.htm]; accessed May 2012.

Editorial Board of Flora of China, Chinese Academy of Sciences. 1985. *Flora of China.* Vol. 16. Science Press: Beijing; 54-55, 60, 68.

ICH. 2005. *Guideline Q2 (R1)–Validation of Analytical Procedues: Text and Methodology.* ICH: Geneva; 1-13.

Kang LP, Ma BP, Wang Y, Zhang J, Xiong CQ and Tan DW. 2005. Study on separation and identification of steroidal saponins of *Dioscorea nipponica* Makino. *Chin Pharm J* **40**: 1539-1541.

Kostova I and Dinchev D. 2005. Saponins in *Tribulus terrestris* – chemistry and bioactivity. *Phytochem Rev* **4**: 111-137.

Leskov AI, Martynova RG and Sokolov SYa. 1976. Polysponin, a new medicinal preparation of antisclerotic effect, *Pharm Chem J* **10**:147-150

Li R, Zhou Y, Wu ZJ and Ding LS. 2006. ESI-QqTOF-MS/MS and APCI-IT-MS/MS analysis of steroid saponins from the rhizomes of *Dioscorea panthaica*. *J Mass Spectrom* **41**: 1-22.

Liu ZB, W TJ, Lu ZQ, Chen XH and Bi KS. 2008. Simultaneous determination dioscin and protoneodioscin in *Dioscorea nipponica* by HPLC. *Chin Tradit Herb Drugs* **39**: 774-775.

Liu ZR, Qi W, Fu TJ, Zou WJ, Ji YQ, Li BG and Huang Y. 2007. Steroidal saponin pharmaceutical composition, its preparation method and use. EP 1800685 A1.

Qin LY, Jia LY, Sun QS and Huang SW. 2007. Determination of contents of dioscin and gracillin in *Dioscorea nipponica* Makino by RP-HPLC. *J Shenyang Pharm Univ* **24**: 593, 628-630.

Sautour M, Mitaine-Offer AC and Lacaille-Dubois MA. 2007. The *Dioscorea* genus: a review of bioactive steroid saponins. *J Nat Med* **61**: 91-101.

Shen Z, Zhang WT and Zhao WL. 2011. Simultaneous determination of four major steroidal saponins in seven species of *Dioscorea* L. by HPLC-ELSD. *Chin Herb Med* **3**: 70-74.

Sichuan Biology Research Institute. 1977. The extract of *Dioscorea nipponica* in the treatment of 216 cases of coronary heart diseases. *Chin Med J* **57**: 520-522.

Solomonova LN. 1968. The use of polysponin in patients with coronary atherosclerosis. *Kardiologiia* **8**: 71-74.

Tistaert C, Dejaegher B and Vander Heyden Y. 2011. Chromatographic separation techniques and data handling methods for herbal fingerprints: A review. *Anal Chim Acta* **690**: 148-161.

Tsukamoto T, Kawasaki T, Yamauchi T and Shimauchi Y. 1957. Isolation of Δ3,5-Desoxytigogenin (25D-spirosta-3,5-diene) from some Dioscoreaceae plants. *Chem Pharm Bull* **5**: 492-493.

Wang TJ, Choi RC, Li J, Bi CW, Ran W, Chen XH, Dong TT, Bi KS and Tsim KW. 2012. Trillin, a steroidal saponin isolated from the rhizomes of *Dioscorea nipponica*, exerts protective effects against hyperlipidemia and oxidative stress. *J Ethnopharmacol* **139**: 214-220.

Yi T, Chen HB, Zhao ZZ, Jiang ZH, Cai SQ and Wang TM. 2009. Comparative analysis of the major constituents in the traditional Tibetan medicinal plants *Saussurea laniceps* and *S. medusa* by LC-DAD-MS. *Chromatographia* **70**:

957-962.

Zhu JB, Guo XJ, Fu SP, Zhang XL and Liang XM. 2010. Characterization of steroidal saponins in crude extracts from *Dioscorea zingiberensis* C. H. Wright by ultra-performance liquid chromatography/electrospray ionization quadrupole time-of-flight tandem mass spectrometry. *J Pharm Biomed Anal* **53**: 462-474.

Legends for Tables

Table 1. List of samples of DP and DN

Table 2. Peaks of reference standards identified in mass spectra

Table 3. Contents of glycosides and aglycone in 13 DN and 13 DP samples (mg/g)

Table 1 List of samples of DN and DP

No.	Date of collection	Source	Growth period (year)	Similarity[a]
DN-01	7/2011	Lingbao, Henan	>2y, cultivated	0.98
DN-02	7/2011	Chengdu, Sichuan	1-2y, cultivated	0.99
DN-03	7/2011	Northeast China	1-2y, wild	0.99
DN-04	7/2011	Northeast China	1-2y, wild	0.97
DN-05	7/2011	Hebei	1-2y, wild	1.00
DN-06	2/2012	Hebei	1-2y, wild	0.97
DN-07	2/2012	Sichuan	1-2y, cultivated	0.99
DN-08	6/2007	Northeast China	1-2y, wild	0.95
DN-09	8/2007	Northeast China	1-2y, wild	0.96
DN-10	9/2005	Northeast China	1-2y, cultivated	0.98
DN-11	6/2005	Northeast China	1-2y, cultivated	0.97
DN-12	2/2010	Sichuan	1-2y, cultivated	0.99
DN-13	3/2005	Lingbao, Henan	2 y, cultivated	0.89
DP-01	3/2010	Xichang, Sichuan	2 y, cultivated	0.95
DP-02	3/2008	Ganzi, Sichuan	2 y, wild	0.99
DP-03	5/2006	Shaoyang, Hunan	1-2 y, wild	0.99
DP-04	4/2008	Maoxian, Sichuan	2 y, cultivated	0.50
DP-05	2/2008	Xichang, Sichuan	2 y, cultivated	0.86
DP-06	3/2008	Luding, Sichuan	>2y, wild	0.95
DP-07	12/2006	Ganzi, Sichuan	>2y, wild	0.94
DP-08	3/2008	Ganzi, Sichuan	>2y, wild	0.95
DP-09	3/2008	Maoxian, Sichuan	>3 y cultivated	0.99
DP-10	3/2008	Xichang, Sichuan	2y, cultivated	0.90
DP-11	3/2008	Lijiang, Yunnan	2y, cultivated	0.95
DP-12	12/2005	Lijiang, Yunnan	1y, cultivated	0.57
DP-13	12/2007	Zhaotong, Yunnan	1-2y, wild	0.91

[a] Similarity was analyzed as described in the section "Comparative analyses for chemical composition of DN and DP" of this paper.

Table 2. Peaks of reference standards identified in mass spectra

Peak	Retention time (min)	Identification	Formula	Selected ion	*m/z* calculated	*m/z* observed	Error (ppm)
1	10.208	Protodioscin	$C_{51}H_{83}O_{22}$	$[M-H]^-$	1047.5381	1047.5422 [e]	3.87
2	11.076	Protogracillin	$C_{51}H_{83}O_{23}$	$[M-H]^-$	1063.5331	1063.5370	3.7
3	15.836	Pseudoprotodioscin	$C_{51}H_{81}O_{21}$	$[M-H]^-$	1029.5276	1029.5280	0.4
5	26.325	Dioscin	$C_{45}H_{71}O_{16}$	$[M-H]^-$	867.4748	867.4777	3.39
6	26.742	Gracillin	$C_{45}H_{71}O_{17}$	$[M-H]^-$	883.4697	883.4722	2.86
7	28.496	Polyphyllin V	$C_{39}H_{61}O_{12}$	$[M-H]^-$	721.4169	721.4182	1.87
10	9.407	Diosgenin	$C_{27}H_{43}O_3$	$[M+H]^+$	415.3207	415.3190	-4.02

[e] Values based on sample of DN-07 as an example, and errors of observed values from all DN and DP samples were within 5 ppm.

Table 3. Contents of glycosides and aglycone in 13 DN and 13 DP samples (mg/g)

No.	Protodioscin	Protogracillin	Pseudoprotodioscin	Dioscin	Gracillin	Polyphyllin V	Diosgenin
DN-01	20.88 ± 0.12	6.71 ± 0.07	0.98 ± 0.01	11.71 ± 0.10	2.50 ± 0.01	$0.10 \pm 2.32 \times 10^{-3}$	9.41 ± 0.26
DN-02	14.41 ± 0.53	8.24 ± 0.39	2.33 ± 0.01	16.48 ± 0.35	7.12 ± 0.31	$0.22 \pm 1.71 \times 10^{-3}$	14.50 ± 0.43
DN-03	17.19 ± 0.28	10.92 ± 0.20	0.44 ± 0.01	20.45 ± 0.45	8.45 ± 0.19	$0.26 \pm 2.34 \times 10^{-3}$	19.39 ± 0.59
DN-04	22.79 ± 0.71	15.57 ± 0.21	3.87 ± 0.09	12.18 ± 0.36	6.19 ± 0.16	$0.18 \pm 3.63 \times 10^{-3}$	13.63 ± 0.81
DN-05	16.10 ± 0.09	7.64 ± 0.11	2.56 ± 0.03	19.32 ± 0.55	6.67 ± 0.17	$0.26 \pm 4.08 \times 10^{-3}$	16.78 ± 0.42
DN-06	13.95 ± 0.85	2.58 ± 0.07	1.08 ± 0.02	27.66 ± 0.80	2.70 ± 0.14	$0.20 \pm 4.74 \times 10^{-3}$	9.62 ± 0.15
DN-07	18.12 ± 0.16	10.67 ± 0.17	0.59 ± 0.03	23.13 ± 0.48	9.19 ± 0.13	$0.33 \pm 15.28 \times 10^{-3}$	10.79 ± 0.81
DN-08	14.08 ± 0.33	1.72 ± 0.04	11.01 ± 0.40	19.98 ± 0.12	1.14 ± 0.03	$0.14 \pm 5.42 \times 10^{-3}$	6.31 ± 0.15
DN-09	12.00 ± 0.25	1.97 ± 0.08	13.52 ± 0.25	22.86 ± 0.41	1.56 ± 0.05	$0.18 \pm 3.55 \times 10^{-3}$	11.00 ± 0.66
DN-10	14.59 ± 1.11	3.78 ± 0.26	4.76 ± 0.30	27.50 ± 1.10	3.36 ± 0.14	$0.24 \pm 7.85 \times 10^{-3}$	14.17 ± 0.11
DN-11	20.42 ± 0.67	3.71 ± 0.15	6.56 ± 0.24	29.28 ± 1.01	2.21 ± 0.06	$0.21 \pm 3.61 \times 10^{-3}$	15.14 ± 0.61
DN-12	19.51 ± 0.10	10.88 ± 0.29	5.40 ± 0.10	20.15 ± 0.47	6.49 ± 0.12	$0.26 \pm 2.14 \times 10^{-3}$	12.46 ± 0.72
DN-13	38.81 ± 0.75	6.07 ± 0.13	16.59 ± 0.47	11.82 ± 0.23	0.88 ± 0.02	N/A	9.55 ± 0.62
Average DN	18.68 ± 6.85	6.96 ± 4.22	5.36 ± 5.25	20.19 ± 5.97	4.50 ± 2.93	0.20 ± 0.08	12.52 ± 3.55

Table 3. (continued)

No.	Protodioscin	Protogracillin	Pseudoprotodioscin	Dioscin	Gracillin	Polyphyllin V	Diosgenin
DP-01	11.79 ± 0.08	10.90 ± 0.26	0.36 ± 0.02	7.43 ± 0.08	5.44 ± 0.05	$0.36 \pm 15.34 \times 10^{-3}$	5.63 ± 0.08
DP-02	6.01 ± 0.20	5.07 ± 0.16	5.66 ± 0.14	8.42 ± 0.35	3.87 ± 0.13	$0.39 \pm 13.96 \times 10^{-3}$	4.26 ± 0.15
DP-03	7.43 ± 0.26	4.87 ± 0.14	0.58 ± 0.01	11.26 ± 0.40	6.00 ± 0.29	$0.51 \pm 5.87 \times 10^{-3}$	6.77 ± 0.21
DP-04	4.62 ± 0.15	5.20 ± 0.15	0.81 ± 0.01	0.65 ± 0.02	0.42 ± 0.01	N/A	0.44 ± 0.01
DP-05	13.75 ± 0.66	12.65 ± 0.58	2.75 ± 0.03	5.36 ± 0.25	3.41 ± 0.16	$0.30 \pm 1.64 \times 10^{-3}$	3.32 ± 0.04
DP-06	6.66 ± 0.19	5.49 ± 0.19	22.55 ± 0.53	6.84 ± 0.05	3.68 ± 0.06	$0.33 \pm 4.68 \times 10^{-3}$	4.26 ± 0.26
DP-07	16.09 ± 0.99	13.20 ± 0.79	2.68 ± 0.18	10.15 ± 0.63	5.47 ± 0.29	$0.68 \pm 5.64 \times 10^{-3}$	6.53 ± 0.44
DP-08	7.07 ± 0.54	5.49 ± 0.25	21.15 ± 0.19	6.31 ± 0.19	3.32 ± 0.08	$0.27 \pm 8.82 \times 10^{-3}$	3.29 ± 0.23
DP-09	10.15 ± 0.47	7.80 ± 0.36	2.38 ± 0.12	16.29 ± 0.53	9.08 ± 0.43	$0.76 \pm 15.52 \times 10^{-3}$	13.80 ± 0.93
DP-10	18.46 ± 1.05	19.39 ± 1.09	2.62 ± 0.01	8.29 ± 0.38	6.03 ± 0.41	$0.59 \pm 15.74 \times 10^{-3}$	10.11 ± 0.70
DP-11	17.39 ± 0.44	12.03 ± 0.47	4.41 ± 0.24	9.20 ± 0.32	4.19 ± 0.23	$0.42 \pm 10.89 \times 10^{-3}$	6.89 ± 0.37
DP-12	4.00 ± 0.20	4.44 ± 0.10	5.90 ± 0.28	0.72 ± 0.03	0.52 ± 0.03	N/A	0.45 ± 0.02
DP-13	11.40 ± 0.59	7.72 ± 0.23	12.11 ± 0.37	5.09 ± 0.20	5.17 ± 0.25	$0.18 \pm 8.06 \times 10^{-3}$	3.07 ± 0.20
Average DP	10.37 ± 4.90*	8.79 ± 4.53	6.46 ± 7.50	7.39 ± 4.16*	4.35 ± 2.31	0.37 ± 0.23*	5.29 ± 3.70*

All values are expressed as mean ± standard deviation (n = 3);

N/A = Not detected; Average DN and Average DP refer to 13 batches of DN and DP, respectively. $^{*}P < 0.05$

Legends for Figures

Figure 1. Chemical structures of constituents identified in DN and DP.

Figure 2. Typical Base peak chromatogram (BPC) of DN and DP and mixed reference standard solutions for identifying glycosides.

Figure 3. Typical BPC of hydrolyzed samples of DN and DP, and of mixed standard solution of four aglycones.

Figure 4. BPC fingerprint of 13 DN samples.

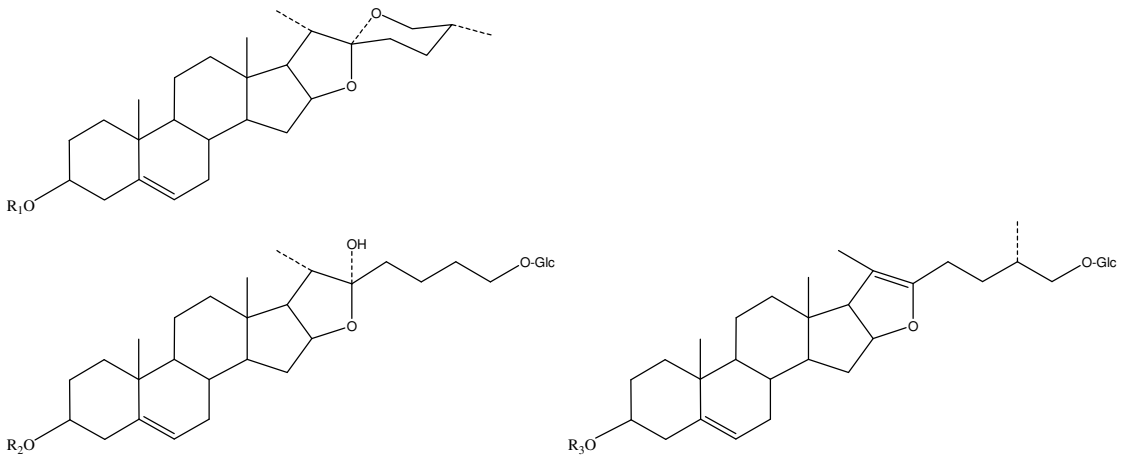
Figure 5. BPC fingerprint of 13 DP samples.

Figure 6. Elucidation of the unknown compound of Peak 4 by ESI-MS/MS data in negative ion mode.

Figure 7. Mass spectrum of peak 12.

Figure 8. Loading scatter plot of 7 variables on three principal components.

Figure 9. Three-dimensional graphics of PCA scores of 13 batches of DN and DP respectively.



Type of chemical structure	Chemical name	Molecular formula	R_1
Spirostanol sapogenin /saponins	Diosgenin	$C_{39}H_{42}O_3$	- H
	Dioscin	$C_{45}H_{72}O_{16}$	- β - D - Glc(4←1) - α - L - Rha (2←1) α - L - Rha
	Gracillin	$C_{45}H_{72}O_{17}$	- β - D - Glc(3←1) - β - D - Glc (2←1) α - L - Rha
	Polyphyllin V	$C_{39}H_{62}O_{12}$	- β - D - Glc(2←1) - α - L - Rha
Furosta-5-en-type saponins	Protodioscin	$C_{51}H_{84}O_{22}$	R_2 - β - D - Glc(4←1) - α - L - Rha (2←1) α - L - Rha
Furosta-5,20 (22)-dien-type saponins	Pseudoprotodioscin	$C_{51}H_{82}O_{21}$	R_3 - β - D - Glc(4←1) - α - L - Rha (2←1) α - L - Rha
	Pseudoprotogracillin	$C_{51}H_{82}O_{22}$	- β - D - Glc(2←1) - α - L - Rha (3←1) β - D - Glc

Figure 1. Chemical structures of constituents identified in DN and DP.

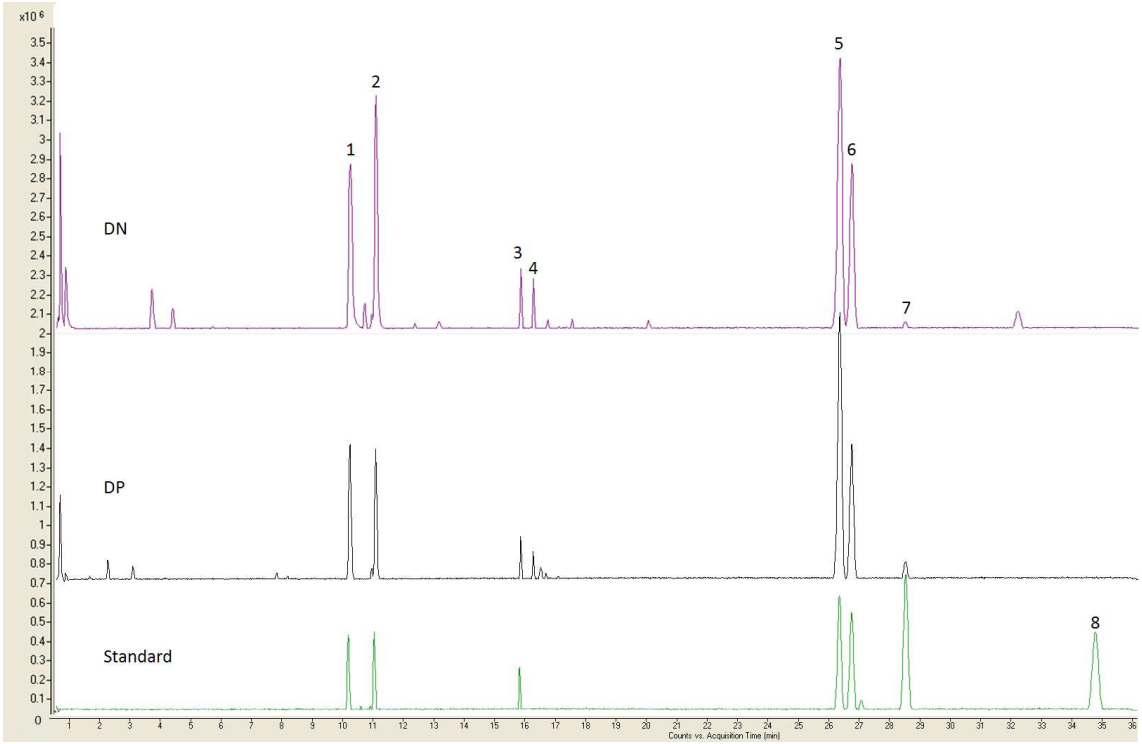


Figure 2. Typical Base peak chromatogram (BPC) of DN, DP and mixed reference standard solutions for identifying glycosides. (1. Protodioscin, 2. Protogracillin, 3. Pseudoprotodioscin, 4. Pseudoprotogracillin, 5. Dioscin, 6. Gracillin, 7. Polyphyllin V)

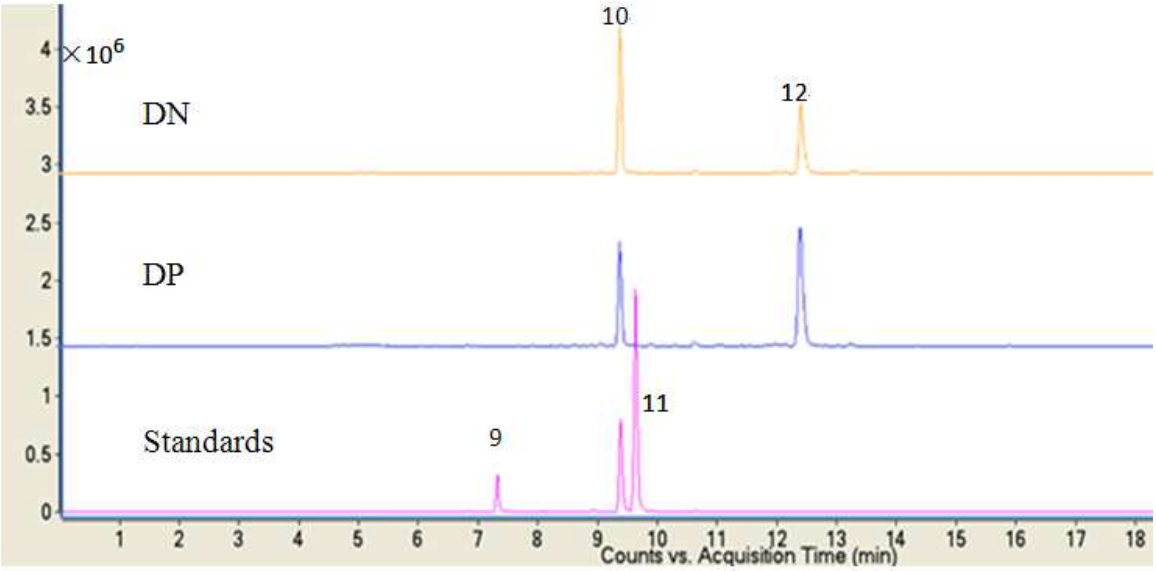


Figure 3. Typical BPC of hydrolyzed samples of DN and DP, and of mixed standard solution of four aglycones (9. Ruscogenin, 10. Diosgenin, 11. Tigogenin and sarsasapogenin, 12. 3,5-deoxytigogenin).

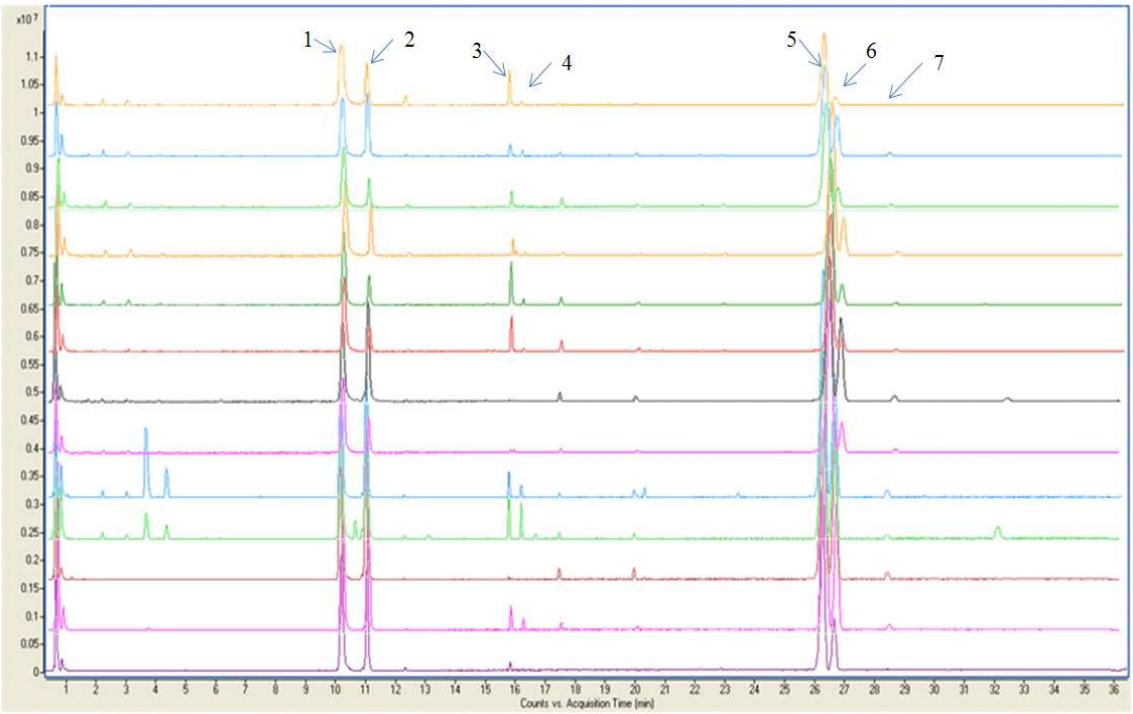


Figure 4. BPC fingerprint of 13 DN samples (the numbered peaks are the same as the ones indicated in Figure 2).

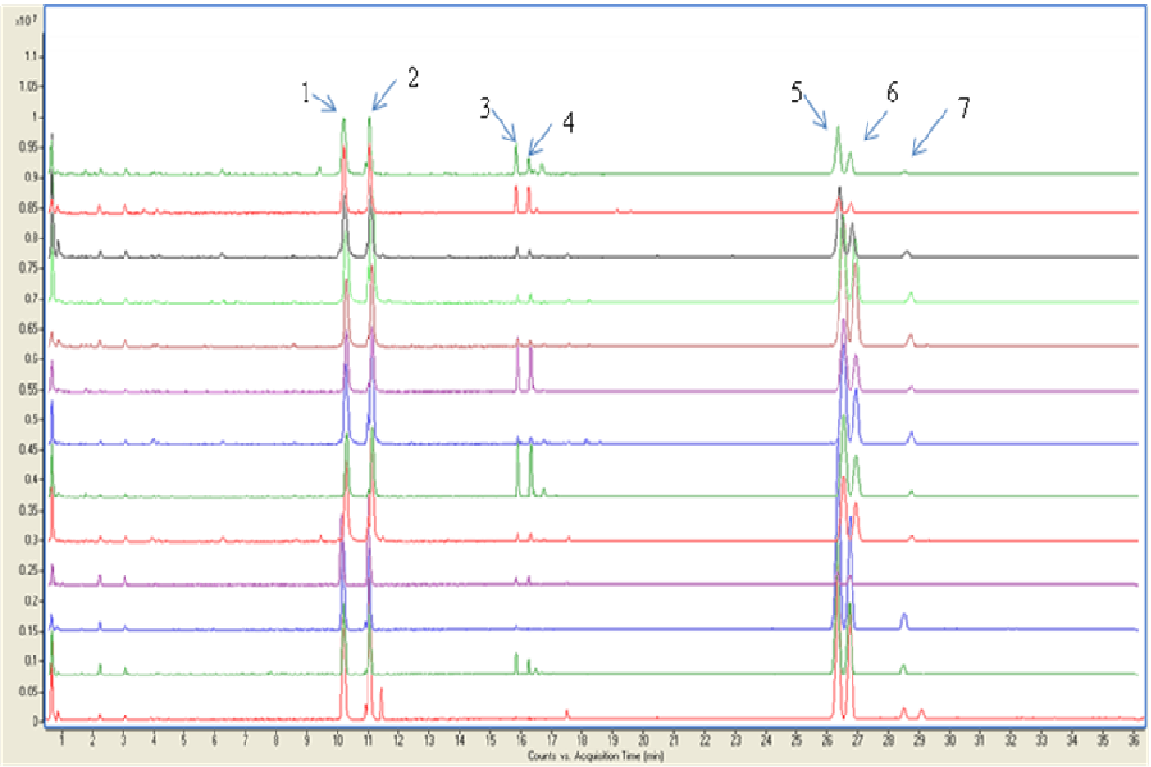
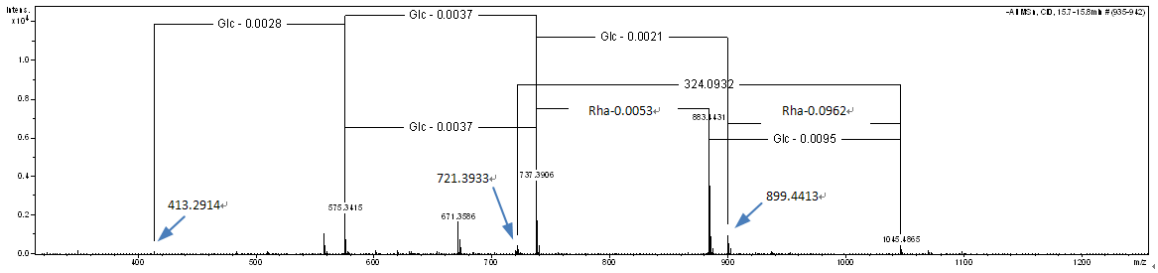


Figure 5. BPC fingerprint of 13 DP samples (the numbered peaks are the same as the ones indicated in Figure 2).



Ion fragment (*m/z* experimental)	(-)ESI-MS/MS *m/z*
1045.4865	$[M-H]^-$
899.4413	$[M-H-Rha]^-$
883.4431	$[M-H-Glc]^-$
737.3906	$[M-H-Glc-Rha]^-$
721.3933	$[M-H-Glc-Glc]^-$
575.3415	$[M-H-Rha-Glc-Glc]^-$
413.2914	$[M-H-Rha-Glc-Glc-Glc]^-$

Figure 6. Elucidation of the unknown compound of Peak 4 by ESI-MS/MS Data in negative ion mode.

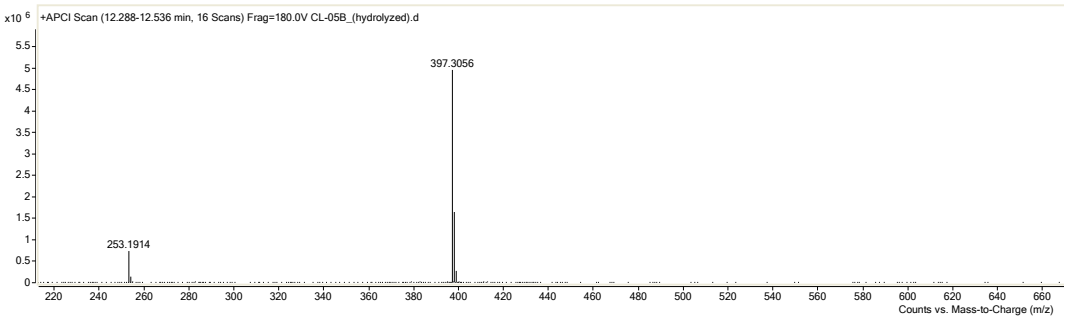


Figure 7. Mass spectrum of peak 12.

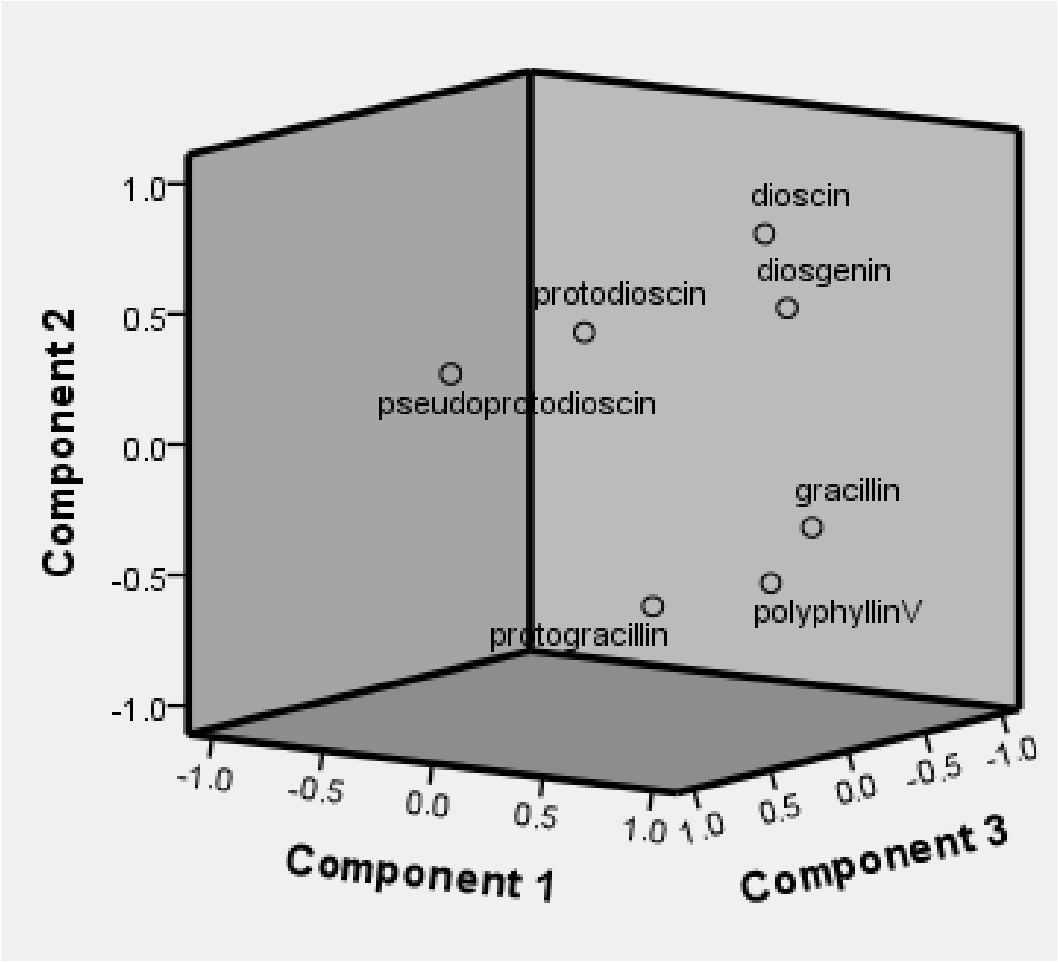


Figure 8. Loading scatter plot of 7 variables on three principal components.

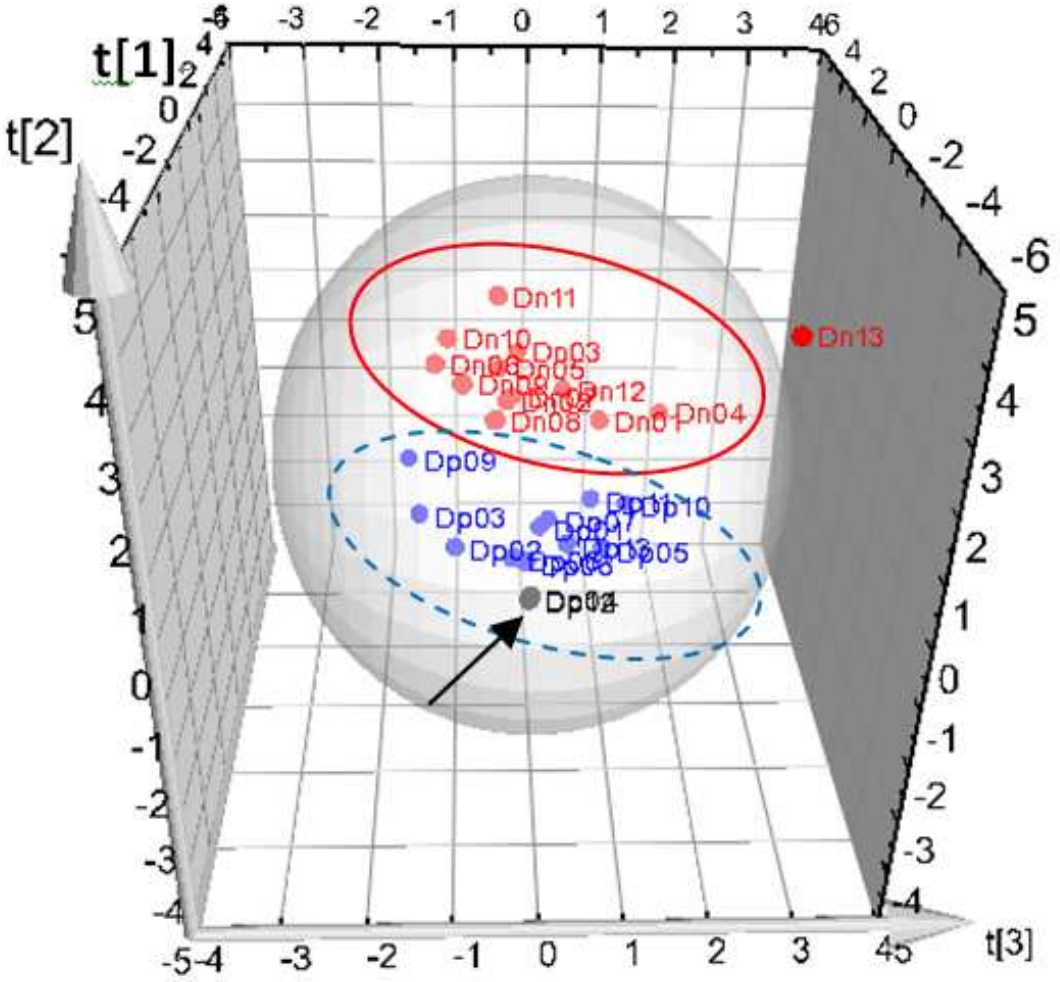


Figure 9. Three-dimensional graphics of PCA scores of 13 batches of DN and DP respectively.