Published in final edited form as:

Planta Med. 2014 March ; 80(4): 321–329. doi:10.1055/s-0033-1360330.

Characterization of Steroidal Saponins from *Dioscorea villosa* and *D. cayenensis* Using Ultrahigh Performance Liquid Chromatography/Electrospray Ionization Quadrupole Time-of-Flight Mass Spectrometry

Bharathi Avula¹, Yan-Hong Wang¹, Mei Wang¹, Zulfiqar Ali¹, Troy J. Smillie¹, Jerry Zweigenbaum², and Ikhlas A. Khan^{1,3}

¹National Center for Natural Products Research, Research Institute of Pharmaceutical Sciences, The University of Mississippi, University, MS, USA

²Agilent Technologies, Wilmington, DE, USA

³Department of Pharmacognosy, School of Pharmacy, The University of Mississippi, University, USA

Abstract

Yam (*Dioscorea spp.*) is an important edible tuber plant used for medicinal purposes to promote health and longevity in Chinese tradition. Steroidal saponins were reported to be the major physiologically active constituents in yams. In this current work, the structural characteristics of steroidal saponins in methanolic extracts from dried rhizomes of two *Dioscorea* species (*D. villosa* L. and *D. cayenensis* Lam.) and dietary supplements have been identified and analyzed using UHPLC/QTOF-MS in both negative and positive ion modes. The fragmentation patterns of reference standards were determined and the steroidal saponins in the extracts were identified or tentatively characterized from their retention times and mass spectra. The fragments produced by collision-induced dissociation (CID) revealed the characteristic cleavage of glycosidic bonds, and the fragmentation pattern provided structural information about the sugars. Twenty-one saponins, including four tentatively identified compounds, were detected in the crude extracts of two *Dioscorea* species. These saponins can be used to distinguish *D. villosa* from *D. cayenensis*. For example, asperin and gracillin are found only in *D. cayenensis*, and dioscoreavilloside A and B and parvifloside are only found in *D. villosa*. This can be used to determine the presence or absence of *D. villosa* in commercial products, which may help determine the spiking of plant material, and/or prevent the use of potentially mislabeled or misidentified “*Dioscorea*” material. The analytical method also provided an alternative, fast method for quality control of *Dioscorea*

Correspondence, Dr. Ikhlas A. Khan, National Center for Natural Products Research, Research Institute of Pharmaceutical Sciences, The University of Mississippi, University, MS 38677, USA, Phone: + 1 66 29 15 78 21, ikhan@olemiss.edu.

Supporting information available online at <http://www.thieme-connect.de/ejournals/toc/plantamedica>

Supporting information

MS data of (–)-ESI and (+)-ESI modes and identification of steroidal saponins from *Dioscorea* using LC-QTOF-MS as well as the ESI-MSⁿ data of the steroidal saponins and steroidal saponins from crude extracts of *D. villosa* and *D. cayenensis* are available as Supporting Information.

Conflict of Interest

The authors declare no conflict of interest.

species in dietary supplements. Principal component analysis showed that *Dioscorea* species and commercial products were easily distinguished. From this a partial least squares model was constructed to determine what species are in different products.

Keywords

Dioscorea villosa; *Dioscorea cayenensis*; Dioscoreaceae; UHPLC-QTOF-MS; steroidal saponins; furostan saponins; spirostan saponins

Introduction

The rhizomes and roots of several species of the genus *Dioscorea* (Family Dioscoreaceae), also commonly known as yams, are staple foods for the populations of Africa, Asia, and tropical America. The genus *Dioscorea* L. consists of about 630 species and is distributed throughout the world. Most of this species contain steroidal saponins and were found to have a wide range of pharmacological activities [1–7]. The most well-known species is *Dioscorea villosa*, also called Wild Yam Root, and is native to North America. The wild yam *D. villosa* was shown not only to be an important source of diosgenin but also used to minimize postmenopausal symptoms and treat low progesterone levels [3].

Reports have indicated that steroidal saponins yield diosgenin (the aglycon of the yam steroidal saponin) from hydrolysis of yam saponins. Dios-genin is used as an industrial starting material for the partial synthesis of steroidal-based drugs, e.g., progesterone and testosterone [8,9]. Over 50 steroid saponins of furostan-, spirostan-, and pregnane-type skeletons have been isolated and characterized from various *Dioscorea* species, and these compounds have been reported to be the major physiologically active constituents in yams [3,10]. Sautour and coworkers [10,11] identified six saponins from *D. villosa*: protodioscin, methyl protodioscin, parrisaponin, dioscin, pro-genin III (= prosapogenin A of dioscin), and proge-nin II. Hayes et al. [12] identified four major and three minor steroidal saponins from *D. villosa* using proton, carbon, and 2D NMR spectroscopy. The major saponins (none of which were reported by Sautour et al.) included two furanostane types, methyl parvifloside and protodeltonin, as well as two spirostane types, deltonin and glucosidodeltonin (zingiberensis I). The minor saponins [10,11] reported by Sautour et al. were methylprotodioscin, disoscin, and prosapogenin A of diosgenin [12]. Morgan [9] noted that some commercial extracts were reported to contain diosgenin only and these extracts may have been derived from other species (e.g., Chinese yam, *D. oppositifolia* L.), which may lack the full chemical profile of *D. villosa* [13]. Biologically active steroidal saponins have also been reported by Sautour et al. [14]. The common sugars present in steroidal saponins from Dioscoreaceae are hexose (glucose) and 6-deoxyhexose (rhamnose). Glucosylated saponins are generally connected through an oxygen atom at the C-3 and/or C-26 position of steroidal aglycone [15]. Most saponin aglycones possess a furostan-type skeleton and differ by the substituent configurations and by the nature of the E ring.

Characterization of steroidal saponins from *D. villosa* have been determined by NMR spectra [14,16–18]; however, LC–MS has also proven to be a very convenient and efficient technique for the identification of steroidal saponins from plant extracts [19–23]. Recently,

ultrahigh performance liquid chromatography (UHPLC) with MS detection has been used as an efficient tool for complex sample analysis and to facilitate structural elucidation [24–29]. One paper [15] characterized the steroidal saponins from *D. zingiberensis*, which is used in TCM. This work used five reference standards for the characterization of the plant material.

Utilizing the high chromatographic resolution and separation capabilities of UHPLC with QTOF-MS provides the structural characterization from accurate mass measurement for both MS and MS-MS experiments. These techniques offer a significant advantage for rapid screening of target compounds in complex matrices. In our previous studies, the quantification of steroidal saponins from *D. villosa* and *Dioscorea* species has been reported [30]. In the present study, the structural characteristics of steroidal saponins from methanolic extracts of dried rhizomes, the roots of *Dioscorea* species (*D. villosa* and *D. cayenensis*), and dietary supplements have been studied using UHPLC-QTOF-MS in both negative and positive ion modes. Herein, we also describe an extensive investigation of the fragmentation pathways leading to the identification of key diagnostic fragment ions for 21 steroidal saponins. The identification of diagnostic fragment ions is important to the data mining of new saponins (that is the ones that have not yet been identified). MS-MS experiments lessen the ambiguities in the identification of steroidal saponins. Because of the many saponin isomers, even accurate mass MS/MS is insufficient for complete identification, requiring chromatographic separation and comparison of standards. Some compounds with similar masses were further characterized and confirmed from their retention times obtained from standard compounds. This technique also provides fast screening of steroidal compounds from *Dioscorea* species and dietary supplements.

Results and Discussion

Steroidal saponins lack a specific chromophore for UV/Vis detection, therefore mass specific detection was employed. In addition, a Poroshell 120-EC8 column was selected to provide sufficient resolution of the complex constituents of the *Dioscorea* samples. This packing has a solid core of 1.7 μm with a 0.5- μm thick porous outer layer and a particle size of 2.7 μm , providing high efficiency. In addition, formic acid at 0.1% was introduced into the mobile phase to alleviate peak tailing and improve ionization. Optimized chromatographic conditions were achieved after several trials with acetonitrile and water in different proportions as the mobile phase. In order to determine the optimal ionization method for analytes, positive and negative ion modes were investigated with the same UHPLC mobile phase at a flow rate of 0.25 mL/min. All compounds could be detected in both negative and positive ion modes. Identification or elucidation of these compounds using UHPLC-QTOF-MS can be performed with the assistance of reference standards and taking into consideration previous knowledge about chemical constituents in these plants. Some chromatographic peaks contained the same deprotonated or protonated molecules in the MS spectra and similar product ions in MS/MS spectra, but had different retention behaviors which were helpful for their identification. Under the developed chromatographic conditions, the retention times of compounds **1** and **2** (cholestane-type compounds) were found to be 3.01 and 4.3 min, compounds **3–9** (furostan-type saponins) were found to be 3.61, 3.96, 4.15, 4.66, 7.56, 7.65, and 7.83 min, and compounds **10–17** (spirostan-type saponins) were 9.21, 10.05, 10.39, 10.42, 10.59, 10.67, 10.97, and 14.83 min, respectively

(Figs. 1 and 2). For the spirostanol saponins, all of their structures are derived from diosgenin (compound **15**) as the major fragment ion.

MassHunter Workstation software, including Qualitative Analysis (version B.05.00), was used for processing both raw MS and MS/MS data, including molecular feature extraction, background subtraction, data filtering, and molecular formula estimation (using the accurate mass of all isotopes, their relative abundances, and all detected adducts with their measured isotopes). To perform subtraction of the molecular features (MFs) originating from the background, analysis of a blank sample (methanol) was carried out under identical instrument settings, and background MFs were removed. In (–)-ESI-MS, the mass spectra of the chromatographic peaks showed deprotonated molecules $[M - H]^-$ and because of the presence of formic acid in the mobile phase, the standards also produced abundant $[M + HCOO]^-$ ions (Tables 1 and 1S). This method also involved the use of $[M + H]^+$ and $[M + Na]^+$ ions in the positive ion mode found in the extracted ion chromatogram (EIC). Other significant peaks in the positive ion mode conformed to the loss of one to two hexoses (Fig. 1S).

In (–)-ESI-MS, the mass spectrum of compound **1** showed the deprotonated molecule $[M - H]^-$ (m/z 757.4371), and the standard also gave an abundant $[M + HCOO]^-$ ion at m/z 803.4424 (Tables 1 and 1S, Fig. 1S). The tentative molecular formula $C_{39}O_{66}H_{14}$ was determined from the (+)-ESI-MS spectrum: m/z 781.4348 corresponding to the $[M + Na]^+$ with a mass accuracy of 0.38 ppm, isotopic spacing, and abundance matching this formula. Other significant peaks at 579.3842 $[M - H_2O - 162]^+$, 561.3774 $[M - 2H_2O - 162]^+$, 543.3661 $[M - 3H_2O - 162]^+$, 383.3300 $[(M + H) - 3H_2O - 2 \times 162]^+$, and 365.3190 $[(M + H) - 4H_2O - 2 \times 162]^+$ resulted from losses of one or two hexoses. A similar type of pattern was observed for compound **2** (Tables 1 and 1S, Fig. 1S). Compounds **1** and **2** were shown to be present only in *D. villosa* [31].

Based on MS and MS/MS data (Tables 1 and 1S, Fig. 1S) of compounds **3–9**, it was determined that these compounds all have the same aglycone skeleton (Fig. 1). Compounds **5** and **7** were chosen as examples to elucidate the structures of these steroidal saponins. Compound **5** showed a deprotonated molecule $[M - H]^-$ (m/z 1047.5376) in the (–)-ESI-MS spectra. In (+)-ESI-MS, four fragments (m/z 885.48, 739.43, 577.37, and 415.32) from the precursor ion $[M - H_2O + H]^+$ (m/z 1031.5426) could be attributed to consecutive losses of two rhamnosyl and two glucosyls sugars. The MS-MS of the precursor ion m/z 1031.5422 gave fragment ions at m/z 885.4842, 739.4409, and 577.3733, indicating the presence of two terminal rhamnoses, one terminal glucose at the C-3 position, and one terminal C-26 glycosylated-diosgenin fragment [15, 31]. These assignments are all in agreement with the structure presented in Figs. 1 and 1S. Compound **7** produced a strong deprotonated molecule $[M - H]^-$ (m/z 1045.5221) in (–)-ESI-MS and also gave an abundant $[M + HCOO]^-$ ion at m/z 1091.527. In (+)-ESI-MS, six fragment ions at m/z 885.4827, 723.4289, 577.3735, 415.3200, and 397.3099 could be attributed to the consecutive loss of one molecule of water, one deoxyhexose, and three hexoses from the ion $[M + H]^+$ (m/z 1047.5376). So, compound **7** was identified as huangjiangsu A, which had been isolated from *D. villosa* [31]. The fragment ion at m/z 885.48 resulted from the loss of a hexose as the terminal unit $[M - Glu + H]^+$. MS-MS fragment peaks corresponded to the loss of hexoses as terminal units at the

C-3 and C-26 positions, which suggested that compound 7 was a bisdesmosidic steroidal saponin containing two hexoses and one deoxy hexose constituent at the C-3 position and one hexose unit at the C-26 position. Considering the retention times of the chromatographic peaks that described the same fragmentation pattern, the structural difference was only the location of the sugar units [15]. The fragmentation of compounds **3–9** suggested the presence of a glucosyl unit connected to the hydroxyl group at the C-26 position and a saccharide chain connected to the hydroxyl group at the C-3 position of the aglycone. For compounds **3–6** (Fig. 1), the mass spectrum was dominated by the fragment ion $[M - H_2O + H]^+$, suggesting the presence of a hydroxy group at the C-22 position of the aglycone [15]. From the MS and MS/MS spectra (Tables 1 and 1S), and the retention times, compounds **10–17** could be deduced to have the same aglycone skeleton. Compounds **10** and **14** were chosen as an example to discuss the fragmentation patterns for these saponins in detail. Compound **10** produced a deprotonated molecule $[M - H]^-$ (m/z 883.4683) and an abundant $[M + HCOO]^-$ ion at m/z 929.4743. In (+)-ESI-MS, six fragments (m/z 739.43, 577.37, 431.31, 413.31, 395.29, and 377.28) from the precursor ion $[M + H]^+$ (m/z 885.4844) were obtained by a consecutive loss of one glucosyl, and two rhamnosyl and hydrated moieties. In (+)-ESI-MS-MS, the mass spectrum was dominated by the main fragment ion $[M - Glu - 2Rha - 2H_2O + H]^+$ (m/z 395.2944). In the positive ion mode, the compound has a characteristic fragment ion $[aglycone + H]^+$ (m/z 431.31). Thus, compound **10** was identified as 25*R*-dracaenoside G. Compound **14** produced a strong deprotonated molecule $[M - H]^-$ (m/z 867.4706) in (–)-ESI-MS. In (+)-ESI-MS, the precursor ion $[M + H]^+$ (m/z 869.4807) produced four fragments (m/z 723.43, 577.37, 415.32, and 397.31) by the consecutive loss of two rhamnosyls, one glucosyl, and water. The fragment ions at m/z 271.20 and 253.20 originated from the consecutive loss of 144 and 18 Da from the fragment ion at m/z 415.32 and have been reported to be the diagnostic fragment ions of this type of steroidal saponins [15, 31]. The elimination of 144 Da (formula $C_8H_{16}O_2$) might be produced by cleavage of the E-ring of the aglycone [23,31], and the 18 Da unit was derived from the loss of a molecule of water. Compounds **11–16** underwent fragmentation pathways similar to dioscin, suggesting that they shared the same aglycone skeleton. According to the literature and structural analysis [15,31], it was deduced that compounds **11–16** were zingiberensis saponin I, deltonin, gracillin, dioscin, asperin, and progenin III, respectively. Compounds **10–16** are monodesmosidic steroidal saponins containing hexose and deoxy-hexose constituents at the C-3 position. The mass spectra were dominated by the fragment ion at $[aglycone + H]^+$ (m/z 415.32), except for compound **10** where the aglycone was different from the others by having an additional hydroxyl group at the C-14 position. Compounds **9, 10**, and **15** displayed the same protonated molecule $[M + H]^+$ (m/z 885.48) in their mass spectra with different sequences of sugars and four main fragments (m/z 723.43, 577.37, 415.32, and 397.31) corresponding to a consecutive loss of two glucosyl and two rhamnosyl units from the protonated molecule. These compounds were differentiated based on their chromatographic retention. Compound **17** produced a strong protonated molecule $[M + H]^+$ (m/z 415.32) in (+)-ESI-MS. The fragment ions at m/z 271.20 and 253.20 originated from the consecutive loss of 144 and 18 Da.

The elimination of 144 Da (formula $C_8H_{16}O_2$) might be produced by cleavage of the E-ring of the aglycone [19,23], and the 18 Da unit was derived from loss of a water molecule.

Compounds **18–21** produced both strong deprotonated and protonated molecules. The sugar sequences, connected at the C-3 position, might have three glucosyl units, one rhamnose unit, and one glucosyl moiety at the C-26 position indicating a furastane-type structure for compounds **18** and **19**. Considering the fragmentation patterns and reported structures in the literature [11, 14–16], compounds **18–21** were tentatively identified as asperoside, zingiberenine G, diosgenin diglucoside, and progenin II, respectively.

Based on the aglycon moiety, steroidal saponins have been classified into cholestane, spirostan, and furastan types. *Dioscorea* species (*D. villosa* and *D. cayenensis*) have been found to contain mostly furastan- and spirostan-type compounds. Cholestane saponins (**1–2**), having a tetracyclic ABCD-ring system, and furostan saponins (**3–9**, **18**, **19**), having a pentacyclic ABCDE-ring system with the sixth open F ring, have two points of sugar attachment, whereas spirostan saponins (**10–17**, **20**, **21**), having a hexa-cyclic ABCDEF-ring system, have one point of sugar linkage. In all compounds except diosgenin, glycone attachment at C-3 is common with a sugar unit range from 1–4. Only glucose was found to be linked as a second sugar connection in cholestane saponins at C-16 and in furostan saponins at C-26. Diosgenin is an aglycone of spirostan saponins.

An MS-MS spectral library was created by analyzing 21 reference standards using the proposed chromatographic method. The data were obtained in both positive and negative electrospray ionization modes at collision energies 0, 10, 20, 30, and 40 eV. Multiple collision energies were necessary because the fragmentation behavior was different for all 21 compounds. Only singly charged positive ($[M + H]^+$ or $[M + H - H_2O]^+$) and negative ions ($[M - H]^-$ or $[M + HCOO]^-$) were used to produce targeted MS-MS spectra. Chromatographic peaks were found and spectra were generated by averaging the peak, and then this was presented to the library building tool (MassHunter PCDL Manager, Version B. 04.00). The utility of the MS-MS library was tested on all samples of *Dioscorea* species and five commercial products. Spectral matching was done by searching the corresponding peaks in the library and unknown spectra within a set mass tolerance. When a corresponding peak was found, a dot product of the library peak intensity and unknown peak intensity was calculated. A matching score was then generated by summing the dot products for all the peaks in a given spectrum, normalized to produce a score of 0–100, with 100 indicating a perfect match. This spectral library is a fast, reliable tool for confirming the identity of unknown compounds.

Nine samples of *D. villosa* and three samples of *D. cayenensis* were used for analysis. Once a fingerprint method had been developed through the use of the standardized plant material, other samples of *Dioscorea* were tested to ensure the usefulness of the library. All samples of *Dioscorea* are listed in the “Materials and Methods” section. The method was validated by testing a number of populations within a single species of *Dioscorea*. Compounds **1–17** present in the samples were identified by comparison of their retention times and MS spectra with those of the reference standards. Samples #5365, 9412, 9800, 9413, 7558, and 7557 indicated the presence of compounds **1–12**, **14**, and **16–17**. Sample #9161 showed the presence of **1–6**, **9–12**, **14**, and **16–17**. Samples #7559 and 9336 contained only one compound (diosgenin). All these samples were labeled as *D. villosa*. Five dietary supplements (#10210, 10211, 10212, 10213, and 10215) claiming to contain *D. villosa* were

also analyzed. Unlike the raw plant material, commercial supplements (#10210 and 10212) are a mixture of wild yam root or wild yam root extract powder with other botanicals (black pepper fruit extract, guarana seed, green tea leaf, yerba mate leaf, ashwaganda, ginseng root, schisandra fruit, ginger root, cayenne fruit, rice flour, or vegetable cellulose) along with dextrin, vitamins, minerals, excipients, and others. Three of these products (#10211, 10213, and 10215) contained only *Dioscorea* as the botanical ingredient (wild yam root powder or extract). Four samples (#10210, 10211, 10213, and 10215) showed similar profiles compared to *D. villosa*. Sample #10212 contained only one compound (diosgenin, **17**). All three samples (#9462, 9166, and 7600) of *D. cayenensis* showed the presence of compounds **4–6**, **12–16**, and **18**. All peaks displayed the characteristic fragmentation patterns of steroidal saponins. Data obtained from the samples showed that 18 compounds were found in *D. villosa* and nine compounds in *D. cayenensis*. Three compounds (**13**, **15**, and **18**) were found only in *D. cayenensis*, compounds **1–3**, **7–11**, **17**, and **19–21** were found in the majority of *D. villosa* samples, and compounds **4–6**, **12**, **14**, and **16** were common to both *Dioscorea* species.

The application of the LC-ESI-MS/MS technique in the current study provided useful information to characterize 18 compounds from *D. villosa* and nine compounds from *D. cayenensis* through the use of authenticated plant materials of both species. The samples #7559 and 9336, which were “labeled” as *D. villosa*, did not show the complete profile compared to authentic samples. With this method, the ground plant material could be tested to confirm or deny the presence of *D. villosa*, which may help detect spiking of plant materials or prevent the use of potentially mislabeled or misidentified “*Dioscorea*” material.

Principal component analysis (PCA) was used to interpret nine samples of *D. villosa*, three samples of *D. cayenensis*, and five dietary products claiming to contain *D. villosa* on the chemical parameters in order to investigate and visualize the homogeneity of the samples (Fig. 3). PCA was performed using the Agilent Mass Profiler Professional (MPP Version 12.5) software. Although it was possible to visually differentiate the various chromatograms, the process was, however, subjective and not quantitative. In addition, minor differences between very similar chromatograms might be missed. The principal components can be displayed graphically as a score plot. This plot is useful for observing any groupings in the data set. The numerical value of a loading of a given variable of a PC shows how much the variable has in common with that component. Generally, this separation was observed in the first two principal components (PC-1 and PC-2 where 91 % of the variance was explained). Examination of the scores and loading plots for PC-1 versus PC-2 showed that compounds **3** and **7–11** were clearly separated from the compounds **13** and **15**. The separation is due mainly to PC-1. PC-2 did not have much effect on the separation of compounds because the same samples showed a less reproducible value when compared to the PC-1 value. The samples of *D. villosa* (#5365, 9412, 9800, 9413, 7558, 9161, and 7557) with distinct saponins were distinguished from *D. cayenensis* samples (#9462, 9166, and 7600). Samples #5365, 9412, 9800, 9413, 7558, and 7557 were closely related, and the other sample (#9161) was positioned apart due to specific patterns of compounds **8** and **9**. Samples #9462, 9166, and 7600 were located away from the samples #5365, 9412, 9800, 9413, 7558, and 7557. Products #10210 and 10213 were closely related to sample #9161. Samples #9336, 7559,

and 10212 were excluded from PCA as these were outliers and they were not found to contain the majority of compounds present in *D. villosa* samples. The contributions of the first and second principal components obtained, using the peak area data of compounds **3**, **7**–**11**, **13**, and **15**, accounted for 91% of the variation in the data. This indicates the significance of the ability of these compounds to differentiate between the two species. This suggests that a partial least squares discriminant analysis (PLS-DA) model can be constructed to predict these two species. The issue of quality control and standardization can also be addressed with the use of PCA and PLS-DA models to observe the variations in the chemical constituents and use of substitutes or adulterants.

In conclusion, UHPLC-QTOF-MS has been applied to the characterization of steroidal saponins from *Dioscorea* species. A total of 21 steroidal saponins with six aglycone skeletons were detected from the methanolic extracts of *D. villosa* and *D. cayenensis*. Saponins were identified by comparing the retention times and the MS and MS/MS data with reference standards. For most compounds, MS in the negative mode gave clear $[M - H]^-$ or $[M + HCOO]^-$ ions. The MS and MS-MS in the positive mode gave sufficient information for structural elucidation. These results showed that the current method could be used to identify steroidal saponins from methanolic extracts and to differentiate two *Dioscorea* species with high sensitivity and reproducibility. PCA was also used for the determination of chemical variation and percentage similarity between the two species. Finally, because of the simple sample preparation, short analysis time, and high reproducibility, the developed analytical method involving PCA would be valuable not only for herbal identification and quality control, but would also enhance understanding of the biological activity of *Dioscorea* roots and their benefits to human health.

Materials and Methods

Instrumentation and analytical conditions: ultrahigh performance liquid chromatography-mass spectrometry (UHPLC/QTOF-MS)

The liquid chromatographic system was an Agilent Series 1290 comprised of the following modular components: binary pump, a vacuum solvent microdegasser, an autosampler with a 100-well tray, and a thermostatically controlled column compartment. Separation was achieved on an Agilent Poroshell 120 EC-8 (2.1×100 mm, 2.7μ) column. The mobile phase consisted of water with 0.1% formic acid (A) and acetonitrile with 0.1% formic acid (B) at a flow rate of 0.25 mL/min, with gradient elution according to Ist Table 2.

Each run was followed by a 5-min wash with 100% B and an equilibration period of 3 min with 75% A/25% B. Two microliters of the sample were injected. The column temperature was set at 35°C and UV spectra were recorded from 190–300 nm.

The mass spectrometric analysis was performed with a QTOF-MS/MS (Model #G6530A; Agilent Technologies) equipped with an ESI source with Jet Stream technology using the following parameters: drying gas (N_2); flow rate 11.0 L/min; drying gas temperature 250°C; nebulizer 35 psig; sheath gas temperature 325°C; sheath gas flow 10 L/min; capillary 3500 V; skimmer 65 V; Oct RF V 750 V; and fragmentor voltage 100 V. The sample collision energy was set at 30 eV. All operations, acquisitions, and analyses of the data were

controlled by Agilent MassHunter Acquisition Software Ver. A.05.00 and were processed with Mass-Hunter Qualitative Analysis software Ver. B.05.00. Each sample was analyzed in both positive and negative modes in the range of $m/z = 100$ – 1700 . Accurate mass measurements were obtained by means of ion correction techniques using reference masses at m/z 121.0509 (protonated purine) and 922.0098 [protonated hexakis (1H, 1H, 3H–tetrafluoropropoxy) phosphazine or HP-921] in the positive ion mode, while m/z 112.9856 (deprotonated trifluoroacetic acid-TFA) and 1033.9881 (TFA adducted HP-921) were used in the negative ion mode. The compounds were confirmed in each spectrum. For this purpose, the reference solution was introduced into the ESI source via a T-junction using an Agilent Series 1200 isocratic pump (Agilent Technologies) with a 100:1 splitter set at a flow rate of 20 $\mu\text{L}/\text{min}$.

For recording TOF-MS spectra, the quadrupole was set to pass all ions (Rf only mode) and all ions were transmitted into the pusher region of the time-of-flight analyzer where they were mass-analyzed with a 1-s integration time up to m/z 1700. For the ESI-MS/MS CID experiments, ions of interest were mass-selected by the quadrupole mass filter. The selected ions were then collided with nitrogen in a high-pressure collision cell and the collision energy was optimized to afford good product ion signals that were subsequently mass measured with the TOF analyzer. Analysis was performed in the reflectron mode with a resolving power of about 10 000 at m/z 922. The instrument was set to an extended dynamic range (up to 10^5 with lower resolving power). MS/MS spectra were recorded simultaneously at a rate of 2.0 spectra s^{-1} . In order to filter selected precursor ions and their isotopes for MS/MS, an isolation window of 1.3 m/z was set for the quadrupole.

Chemicals and plant materials

Chemicals—Seventeen standards were used as reference compounds. The steroidal saponins dioscoreavilloside A (**1**), dioscoreavilloside B (**2**), parvifloside (**3**), protodeltonin (**4**), protodioscin (**5**), protobioside (**6**), huangjiangsu A (**7**), pseudoprotodioscin (**8**), 26-*O*- β -D-glucopyranosyl-3 β ,26-diol-25(*R*)-furost-5,20(22)-dien-3-*O*- α -L-rhamnopyranosyl(1 $\rightarrow$ 2)-*O*- β -D-glucopyranoside (**9**), 25(*R*)-dracaenoside G (**10**), zingiberensis saponin I (**11**), deltonin (**12**), asperin (**13**), dioscin (**14**), gracillin (**15**), progenin III (**16**), and diosgenin (**17**) were isolated at the National Center for Natural Products Research (NCNPR, University of Mississippi, Mississippi, USA). Their identity and purity were confirmed by chromatographic (TLC, HPLC) methods, by the analysis of the spectral data (IR, 1D- and 2D-NMR, ESI-HRMS), and comparison with published spectral data [16,17,31]. The purity of these isolated compounds was greater than 90%. Acetonitrile and formic acid were of HPLC grade purchased from Fisher Scientific. Water for the HPLC mobile phase was purified using a Milli-Q system (Millipore).

Plant materials—Dried rhizomes and roots of *D. villosa* (NCNPR accession #9800) and *D. cayenensis* (NCNPR accessions #9166 and 9462) were obtained from the cultivated, living collection of the NCNPR Maynard W Quimby Medicinal Plant Garden, University of Mississippi. All species were identified by Dr. Aruna Weerasooriya of the Maynard W Quimby Medicinal Plant Garden at the University of Mississippi. The roots of *D. villosa*

(NCNPR accession #5365) and *D. cayenensis* (NCNPR accession #7600) were obtained from the Missouri Botanical Garden (Missouri, USA).

Root powders of NCNPR accessions #9412 and 9413 were procured from Starwest Botanicals. Roots of NCNPR accessions #7557 and 7558 were obtained from Frontier Natural Products Co-op and the roots of NCNPR accession #7559 from Xi'an, China. Roots of NCNPR accessions #161 and 9336 were obtained online (amazon.com and Kalyx.com, respectively). All the samples obtained were labeled as *D. villosa*. Dietary supplements (#10210, 10211, 10212, 10213, 10215) claiming to contain *D. villosa* were purchased online. Specimens of all samples are deposited at the NCNPR's botanical repository at the University of Mississippi.

Standards and plant sample preparation—All standards were prepared in a concentration range from 0.5 to 10 µg/mL Dry plant samples (0.5 g) or an adequate amount of capsule contents (average weight of dosage form) were sonicated in 2.5 mL of methanol for 30 min followed by centrifugation for 15 min at 959 ×g. The supernatant was transferred to a 10-mL volumetric flask. The procedure was repeated three more times and the respective supernatants were combined. The final volume was adjusted to 10 mL with methanol and mixed thoroughly. Prior to injection, an adequate volume (ca. 2 mL) was passed through a 0.45-µm PTFE membrane filter. The first 1.0 mL was discarded and the remaining volume was collected in an LC sample vial. For plant samples where a few constituents were highly concentrated, the sample solutions were further diluted 5 to 10 times.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research is supported in part by “Botanical Identification, Characterization, Quality Assurance and Quality Control” (NIH Prime award number 1P50AT006268-02) and the United States Department of Agriculture, Agricultural Research Service, Specific Cooperative Agreement No. 58-6408-2-0009. The authors would like to thank Annette Ford for the extractions of plant samples. The authors would also like to thank Dr. Jon F. Parcher for his help in correcting grammar and support.

References

1. Lacaille-Dubois MA, Wagner H. Bioactive saponins from plants: an update. *Stud Nat Prod.Chem.* 2000; 21:633–687.
2. Lacaille-Dubois MA. Bioactive saponins with cancer related and immunomodulatory activity: recent developments. *Stud Nat Prod Chem.* 2005; 32:209–246.
3. Benghuzzi H, Tucci M, Eckie R, Hughes J. The effects of sustained delivery of diosgenin on the adrenal gland of female rats. *Biomed Sci Instrum.* 2003; 39:335–340. [PubMed: 12724916]
4. Hu K, Dong AJ, Yao XS, Kobayashi H, Iwasaki S. Antineoplastic agents. Part 2. Four furostanol glycosides from rhizomes of *Dioscorea collettii* var. *hypoglauca*. *Planta Med.* 1997; 63:161–165. [PubMed: 17252340]
5. Zhang J, Meng Z, Zhang M, Ma D, Xu S, Kodama H. Effect of six steroidal saponins isolated from *Anemarrhenae* rhizoma on platelet aggregation and hemolysis in human blood. *Clin Chim Acta.* 1999; 289:79–88. [PubMed: 10556655]

6. Sauvaire Y, Ribes G, Baccou JC, Loubatieres-Marian MM. Implication of steroid saponins and sapogenins in the hypocholesterolemic effect of fenugreek. *Lipids*. 1991; 26:191–197. [PubMed: 2046485]
7. Liu SY, Wang JY, Shyu YT, Song LM. Studies on yams (*Dioscorea* spp.) in Taiwan. *J Chin Med*. 1995; 6:111–126.
8. Chen Y, Wu Y. Progress in research and manufacturing of steroidal sapogenins in China. *J Herb Spices Med Plants*. 1994; 2:59–70.
9. Morgan, M. Steroids. Kirk-Othmer Encyclopedia of Chemical Technology. Hoboken: John Wiley and Sons; 1997.
10. Sautour M, Mitaine-Offer AC, Lacaille-Dubois MA. The *Dioscorea* genus: a review of bioactive steroid saponins. *J Nat Med*. 2007; 61:91–101.
11. Sautour M, Miyamoto T, Lacaille-Dubois MA. Steroidal saponins and fla-van-3-ol glycosides from *Dioscorea villosa*. *Biochem Syst Ecol*. 2006; 34:60–63.
12. Hayes PY, Lambert LK, Lehmann R, Penman K, Kitching W, De VJJ. Complete (1)H and (13)C assignments of the four major saponins from *Dioscorea villosa* (wild yam). *Magn Reson Chem*. 2007; 45:1001–1005. [PubMed: 17887511]
13. Agbor-Egbe T, Treche S. Evaluation of the chemical composition of Cameroonian yam germplasm. *J Food Comp Anal*. 1995; 8:274–283.
14. Sautour M, Mitaine-Offer AC, Miyamoto T, Dongmo A, Lacaille-Dubois MA. Antifungal steroid saponins from *Dioscorea cayenensis*. *Planta Med*. 2004; 70:90–92. [PubMed: 14765305]
15. Zhu J, Guo X, Fu S, Zhang X, Liang X. 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*. 2010; 53:462–474. [PubMed: 20561762]
16. Sautour M, Mitaine-Offer AC, Miyamoto T, Dongmo A, Lacaille-Dubois MA. A new steroidal saponin from *Dioscorea cayenensis*. *Chem Pharm Bull*. 2004; 52:1353–1355. [PubMed: 15516762]
17. Hayes PY, Lambert LK, Lehmann R, Penman K, Kitching W, De VJJ. Spectral assignments and reference data complete 1H and 13C assignments of the four major saponins from *Dioscorea villosa* (wild yam). *Magn Reson Chem*. 2007; 45:1001–1005. [PubMed: 17887511]
18. Hu CC, Lin JT, Liu SC, Yang DJ. A spirostanol glycoside from wild yam (*Dioscorea villosa*) extract and its cytostatic activity on three cancer cells. *Yaowu Shipin Fenxi*. 2007; 15:310–315.
19. Li R, Zhou Y, Wu Z, Ding L. ESI-QqTOF-MS/MS and APCI-T-MS/MS analysis of steroid saponins from the rhizomes of *Dioscorea panthaica*. *J Mass Spectrom*. 2006; 41:1–22. [PubMed: 16402411]
20. Liang M, Zheng Z, Yuan Y, Kong L, Shen Y, Liu R, Zhang C, Zhang W. Identification and quantification of C21 steroidal saponins from Radix Cynanchi Atrati by high-performance liquid chromatography with evaporative light scattering detection and electrospray mass spectrometric detection. *Phytochem Anal*. 2007; 18:428–435. [PubMed: 17624893]
21. Lin S, Wang D, Yang D, Yao J, Tong Y, Chen J. Characterization of steroidal saponins in crude extract from *Dioscorea nipponica* Makino by liquid chromatography tandem multi-stage mass spectrometry. *Anal Chim Acta*. 2007; 599:98–106. [PubMed: 17765069]
22. Zheng Z, Zhang W, Kong L, Liang M, Li H, Lin M, Liu R, Zhang C. Rapid identification of C21 steroidal saponins in *Cynanchum versicolor* Bunge by electrospray ionization multi-stage tandem mass spectrometry and liquid chromatography/tandem mass spectrometry. *Rapid Commun Mass Spectrom*. 2007; 21:279–285. [PubMed: 17200975]
23. Lu Y, Luo J, Xu D, Huang X, Kong L. Characterization of spirostanol saponins in *Solanum torvum* by high-performance liquid chromatography/ evaporative light scattering detector/electrospray ionization with multi-stage tandem mass spectrometry. *Rapid Commun Mass Spectrom*. 2008; 22:2447–2452. [PubMed: 18634112]
24. Zhou DY, Zhang XL, Xu Q, Xue XY, Zhang FF, Liang XM. UPLC/Q-TOFMS/ MS as a powerful technique for rapid identification of polymethoxylated flavones in *Fructus aurantii*. *J Pharm Biomed Anal*. 2009; 50:2–8. [PubMed: 19428213]

25. Novak TJ, Grinberg N, Hartman B, Marcinko S, DiMichele L, Mao B. LCMS using a hybrid quadrupole time of flight mass spectrometer for impurity identification during process chemical development of a novel integrase inhibitor. *J Pharm Biomed Anal.* 2010; 51:78–83. [PubMed: 19733026]
26. Li CM, Zhang XL, Xue XY, Zhang FF, Xu Q, Liang XM. Structural characterization of iridoid glucosides by ultra-performance liquid chromatography/electrospray ionization quadrupole time-of-flight tandem mass spectrometry. *Rapid Commun Mass Spectrom.* 2008; 22:1941–1954. [PubMed: 18491284]
27. Zhang J, Jin Y, Dong J, Xiao Y, Feng J, Xue X, Zhang X, Liang X. Systematic screening and characterization of tertiary and quaternary alkaloids from *Corydalis yanhusuo* W.T. Wang using ultra-performance liquid chromatography-quadrupole-time-of-flight mass spectrometry. *Talanta.* 2009; 78:513–522. [PubMed: 19203617]
28. Wang X, Sun W, Sun H, Lv H, Wu Z, Wang P, Liu L, Cao H. Analysis of the constituents in the rat plasma after oral administration of Yin Chen Hao Tang by UPLC/Q-TOF-MS/MS. *J Pharm Biomed Anal.* 2008; 46:477–490. [PubMed: 18164893]
29. Li SL, Song JZ, Choi FFK, Qiao CF, Zhou Y, Han QB, Xu HX. Chemical profiling of Radix *Paenoniae* evaluated by ultra-performance liquid chromatography/photo-diode-array/quadrupole time-of-flight mass spectrometry. *J Pharm Biomed Anal.* 2009; 49:253–266. [PubMed: 19097720]
30. Avula B, Wang YH, Ali Z, Smillie TJ, Khan IA. Chemical fingerprint analysis and quantitative determination of steroidal compounds from *Dioscorea villosa*, *Dioscorea species*, and dietary supplements using UHPLC-ELSD. *Biomed Chromatogr.* 2014; 28:281–294. [PubMed: 24019066]
31. Ali Z, Smillie TJ, Khan IA. Cholestane steroid glycosides from the rhizomes of *Dioscorea villosa* (Wild Yam). *Carbohydr Res.* 2013; 370:86–91. [PubMed: 23454141]

Fig. 1.
Structure of standard compounds.

Fig. 2.
Base peak chromatograms of the *D. villosa* and *D. cayenensis* extracts analyzed by UHPLC-QTOF-MS in the positive ESI mode.

Fig. 3.

Dioscorea samples; scatterplot of first and second principal components obtained by PCA on the basis of UHPLC-S peak area data of parvifloside, Huangjiangsu A, pseudoprotodioscin, 26-*O*- β -D-glucopyranosyl-3 β ,26-diol-25(*R*)-furost-5,20 (22)-dien-3-*O*- α -L-rhamnopyranosyl(1 $\rightarrow$ 2)-*O*- β -D-glucopyranoside, 25(*R*)-dracaenoside G, zingiberensis saponin I, asperin, and gracillin. Samples #9336, 7559, and 10212 were excluded from PCA.

Table 1
Accurate mass data for 17 steroidal saponin standards determined using UHPLC/QTOF-MS.

Peak	Formula	Selected ions	m/z calculated	m/z experimental	Identities	Error (mDa)
1	C ₃₉ H ₆₆ O ₁₄	[M + Na] ⁺	781.4350	781.4348	Dioscoreavilloside A	0.2
		[M + HCOO] ⁻	803.4429	803.4431		- 0.2
2	C ₃₉ H ₆₄ O ₁₄	[M + Na] ⁺	779.4194	779.4194	Dioscoreavilloside B	0.0
		[M + H - H ₂ O] ⁺	739.4269	739.4268		0.1
		[M + HCOO] ⁻	801.4273	801.4273		0.0
3	C ₃₇ H ₆₄ O ₂₈	[M + H - H ₂ O] ⁺	1209.5904	1209.5903	Parvifloside	0.1
		[M - H] ⁻	1225.5853	1225.5852		0.1
4	C ₅₁ H ₈₄ O ₂₃	[M + H - H ₂ O] ⁺	1047.5376	1047.5376	Protodeltonin	0.0
		[M - H] ⁻	1063.5325	1063.5329		- 0.4
5	C ₅₁ H ₈₄ O ₂₂	[M + H - H ₂ O] ⁺	1031.5427	1031.5427	Protodioscin	0.0
		[M - H] ⁻	1047.5376	1047.5371		0.5
6	C ₄₅ H ₇₄ O ₁₈	[M + H - H ₂ O] ⁺	885.4848	885.4850	Protobioside	- 0.2
		[M - H] ⁻	901.4797	901.4799		- 0.2
7	C ₅₁ H ₈₂ O ₂₂	[M + H] ⁺	1047.5376	1047.5376	Huangjiangsu A	0.0
		[M - H] ⁻	1045.5219	1045.5220		- 0.1
8	C ₅₁ H ₈₂ O ₂₁	[M + H] ⁺	1031.5427	1031.5428	Pseudoprotodioscin	- 0.1

Peak	Formula	Selected ions	<i>m/z</i> calculated	<i>m/z</i> experimental	Identities	Error (mDa)
9	$C_{45}H_{72}O_{17}$	$[M - H]^-$	1029.5270	1029.5272		-0.2
		$[M + H]^+$	885.4848	885.4846	26- <i>O</i> - β -D-glucopyranosyl-3,6,26-diol-25(<i>R</i>)-furost-5,20(22)-dien-3- <i>O</i> - α -L-rhamnopyranosyl(1 $\rightarrow$ 2)- <i>O</i> - β -D-glucopyranoside	0.2
		$[M - H]^-$	883.4691	883.4694		-0.3
10	$C_{45}H_{72}O_{17}$	$[M + H]^+$	885.4848	885.4841	25 <i>R</i> -dracaenoside G	0.7
		$[M - H]^-$	883.4691	883.4672		1.9
		$[M + HCOO]^-$	929.4746	929.4749		-0.3
11	$C_{51}H_{82}O_{22}$	$[M + Na]^+$	1069.5195	1069.5194	Zingiberensis saponin I	0.1
		$[M - H]^+$	1045.5220	1045.5207		1.3
12	$C_{45}H_{72}O_{17}$	$[M + H]^+$	885.4848	885.4850	Deltonin	-0.2
		$[M + Na]^+$	907.4667	907.4664		0.3
		$[M - H]^-$	883.4691	883.4691		0.0
13	$C_{51}H_{82}O_{20}$	$[M + Na]^+$	1037.5297	1037.5295	Parrisaponin or asperin	0.2
		$[M - H]^-$	1013.5321	1013.5323		-0.2
14	$C_{45}H_{72}O_{16}$	$[M + H]^+$	869.4899	869.4895	Dioscin	0.4
		$[M + HCOO]^-$	913.4797	913.4794		0.3
15	$C_{45}H_{72}O_{17}$	$[M + H]^+$	885.4848	885.4847	Gracillin	0.1
		$[M + Na]^+$	907.4667	907.4669		-0.2
		$[M + HCOO]^-$	929.4746	929.4746		0.0
16	$C_{39}H_{62}O_{12}$	$[M + H]^+$	723.4320	723.4319	Progenin III	0.1
		$[M - H]^-$	721.4163	721.4146		1.9

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Peak	Formula	Selected ions	m/z calculated	m/z experimental	Identities	Error (mDa)
17	C ₂₇ H ₄₂ O ₃	[M + H] ⁺	415.3212	415.3214	Diosgenin	- 0.2
		[M + HCOO] ⁻	767.4218	767.4217		0.1

Table 2

Gradient elution table.

Time (minutes)	%A	%B
0.00	75	25
5.00	70	30
7.00	55	45
10.00	40	60
15.00	0	100
20.00	0	100