A multi-analytical platform based on pressurized-liquid extraction, *in vitro* assays and liquid chromatography/gas chromatography coupled to high resolution mass spectrometry for food by-products valorisation. Part 2: Characterization of bioactive compounds from goldenberry (*Physalis peruviana L.*) calyx extracts using hyphenated techniques.

Diego Ballesteros-Vivas[1,2a], Gerardo Alvarez-Rivera[2a], Elena Ibáñez[2], Fabián Parada-Alfonso[1], Alejandro Cifuentes[2]*

[1] High Pressure Laboratory, Department of Chemistry, Faculty of Science, Universidad Nacional de Colombia, Carrera 30 #45-03, Bogotá D.C., 111321, Colombia.

[2] Laboratory of Foodomics, Institute of Food Science Research, CIAL, CSIC, Nicolás Cabrera 9, 28049 Madrid, Spain.

[a] These two authors have contributed equally to this work.

*Corresponding author:

Prof. Dr. Alejandro Cifuentes, Laboratory of Foodomics, Institute of Food Science Research, CIAL (CSIC), Nicolás Cabrera 9, 28049 Madrid, Spain, e-mail: a.cifuentes@csic.es, Tel.: +34 910017955; fax: +34 910017905.

Keywords:

Goldenberry calyx; Withanolides; Valorization; GC-Q-TOF; LC-Q-TOF; Phytochemical profiling

Abbreviations:

ACN: acetonitrile

EI: electronic impact

EIC: extracted ion chromatogram

ESI: Electrospray ionization

EtOAc: ethyl acetate

EtOH: ethanol

GC: gas chromatography

HRMS: high resolution mass spectrometry

PLE: pressurized liquid extraction

q-TOF-MS/MS: quadrupole time-of-flight tandem mass spectrometry

UPLC: ultra-high-performance liquid chromatography

ABSTRACT

A multi-analytical strategy for the valorization of goldenberry calyx, a promising source of health-promoting compounds, is presented in this work. A comprehensive characterization of *P. peruviana* calyx extracts, obtained by an optimized pressurized liquid extraction (PLE) procedure, is developed applying first an ultra-high-performance liquid chromatography coupled to quadrupole time-of-flight tandem mass spectrometry (UPLC-ESI-q-TOF-MS/MS) method in positive and negative electrospray ionization (ESI) mode. A total of fifty-nine phytochemicals, including major phenolic components, several withanolides (C_{28}-isoprenoids) with a variety of biological activities, and a large family of anti-inflammatory sucrose esters were tentatively identified using this methodology. An integrated identification strategy based on accurate mass data obtained by high resolution mass spectrometry (HRMS), ion source fragmentation, MS/MS fragmentation patterns, generated molecular formulae and subsequent unsaturation degree calculation, along with database and bibliographic search is proposed. Isobaric withanolides-type compounds were tentatively identified or classified according to their different hydroxy and epoxy positions, on the basis of the complementary information provided by MS/MS product ion spectra obtained in both ESI+ and ESI- mode. The proposed structural elucidation approach provides a valuable contribution to the limited information available regarding the MS/MS structural analysis of withanolides in ESI(-) mode. Moreover, an alternative elucidation strategy based on deconvolution and database search was successfully applied for the phytochemical profiling analysis of the volatile fraction of *P. peruviana* calyx extracts by gas chromatography quadrupole time-of-flight mass spectrometry (GC-q-TOF-MS), which reveals the presence of relevant terpenoids, including phytosterols and tocopherols (Vitamin E). The results of the phytochemical characterization obtained herein demonstrates the great potential of applying integrated identification strategies to HRMS data obtained from complementary LC- and GC-q-TOF-MS(/MS) platforms, as powerful

identification tools for improving our understanding on the phytochemical composition of natural extracts intended to be used in functional foods or in traditional medicine preparations.

1. INTRODUCTION

Goldenberry or cape gooseberry is an exotic fruit produced by the plant species *Physalis peruviana L.* (Solanaceae family), commonly commercialized as fresh fruit or as derived processed products such as juices, sauces, syrups, marmalades, and snacks [1]. Industrial processing of goldenberry generates a significant amount of by-products, mainly juice pomace (seeds and skins) and calyx. The goldenberry pomace (seeds and skins) represents a large portion of the waste generated during juice processing (ca. 27% of fruit weight) [2] and the nutritional properties of this waste have been well described [3]. However, limited information is available about the composition of goldenberry calyx, which represents 5% of the raw fruit, accounting for around 33 tons of generated waste per cultivated hectare of *P. peruviana* [2].

Several medicinal properties such as antispasmodic, diuretic, antiseptic, sedative, analgesic, throat trouble relief, elimination of intestinal parasites and amoeba are attributed to *P. peruviana L.* Antidiabetic properties have also been attributed to goldenberries, recommending the consumption of five fruits a day [4]. In addition, the calyces of *P. peruviana L.* are widely used in folk medicine for its properties as anticancer, antimicrobial, antipyretic, diuretic, and anti-inflammatory immunomodulator [5].

Most of reported studies in literature about the phytochemical composition of *P. peruviana* are focused on the fruit [6–8]. Previous research works on the genus *Physalis* reported the isolation of steroids (especially withanolides), flavonoids, alkaloids, terpenoids and sucrose esters, among others [9–12]. In particular, withanolides are a family of C_{28} ergostane-type steroids of great interest from the pharmacological point of view, as they were reported to have anti-inflammatory, antitumor, cytotoxic, hepatotoxic and antimicrobial activities [13].

In order to tackle the challenge of analyzing complex phytochemical extracts from natural sources and traditional medicine preparations, advanced hyphenated techniques such as liquid and gas chromatography (GC) coupled to high-resolution mass spectrometry (HRMS) have emerged as powerful tools for this purpose. Citar aquí los siguientes papers:

- Recent developments and emerging trends of mass spectrometry for herbal ingredients analysis
- Recent development in mass spectrometry and its hyphenated techniques for the analysis of medicinal plants

. Despite the wide application of these techniques, HRMS-based methodologies generate complex and huge datasets containing thousands of MS features, making the post-acquisition data processing a laborious and time consuming task [14]. Therefore, an integrated identification and elucidation strategy must be applied to raw HRMS data in order to accurately identify the phytochemical (and potentially bioactive) compounds. In this regard, several works have recently proposed valuable approaches to facilitate the post-acquisition data processing, including mass defect filtering, diagnostic fragment ion filtering and neutral loss filtering among others [15,16].

In view of the potential of goldenberries calyx as promising source of bioactive phytochemicals, a multi-analytical platform based on HRMS is presented in this work as part of an integrated valorisation strategy for this by-product. Thus, a comprehensive phytochemical profiling analysis of the compounds extracted from goldenberries calyx (using and optimized PLE process, as described in our previous work [17], was carried out by LC and GC coupled to quadrupole time-of-flight tandem mass spectrometry (q-TOF-MS/MS), applying integrated identification approaches for the confident identification and structural elucidation of bioactive phytochemicals. The proposed strategy can be readily implemented for the complete

116 characterization of preparations with health-promoting effects, intended to be used as
117 functional foods and in traditional medicine.
118
119 **2. MATERIAL AND METHODS**
120
121 *2.1 Reagents and materials*
122 Gallic acid, protocatechuic acid, 4-hydroxybenzoic acid, vanillic acid, caffeic acid, benzoic
123 acid, p-coumaric acid, ferulic acid, quercetin, kaempferol, quercetin rutinoside, trolox,
124 withanolide A, ammonium acetate and formic acid were purchased from Sigma-Aldrich
125 (Madrid, Spain). The solvents employed were HPLC-grade. Acetonitrile, ethanol and methanol
126 were acquired from VWR Chemicals (Barcelona, Spain), whereas dichloromethane was
127 provided from Fluka AG (Buchs, Switzerland) and ethyl acetate from Scharlau (Barcelona,
128 Spain). Ultrapure water was obtained from a Millipore system (Billerica, MA, USA). For the
129 UPLC-q-TOF-MS analyses, MS grade ACN and water from LabScan (Dublin, Ireland) were
130 employed.
131
132 *2.2 Calyx extracts*
133 Calyx extracts from goldenberry fruit (*Physalis peruviana*) were obtained as described in our
134 previous work [17], where PLE conditions (temperature and extraction solvent) were optimized
135 to maximize extraction yield, withanolides recoveries, total phenolic content, total flavonoids
136 content and antioxidant activity. In brieve, dried sample of goldenberry calyces (~1 g) was
137 mixed with sea sand (~2 g). The mixture was loaded into an 11 mL stainless steel extraction
138 cell. PLE experiments were developed at static extraction mode for 20 min and 100 bar. After
139 the extraction, the solvent was evaporated under a stream of nitrogen at 25 °C (TurboVap® LV

Biotage, Uppsala, Sweden). The selected PLE conditions were 125 °C and 75/25 of EtOH/EtOAc (v/v) as extraction solvent.

2.3 Phytochemical profiling of P. peruviana extracts

2.3.1 Liquid chromatography-tandem mass spectrometry (UPLC-q-TOF-MS/MS)

Liquid chromatography coupled to a high-resolution mass spectrometer was employed to characterize the phytochemical compounds extracted from goldenberry calyx. Analyses were performed using an ultrahigh performance liquid chromatography (UPLC) system 1290 from Agilent (Agilent Technologies, Santa Clara, CA, USA) coupled to a quadrupole-time-of-flight mass spectrometer (q-TOF-MS) Agilent 6540 that was equipped with an orthogonal ESI source (Agilent Jet Stream, AJS, Santa Clara, CA, USA), and controlled by a PC running the Mass Hunter Workstation software 4.0 (MH) from Agilent. Two chromatographic methods were carried out using a Zorbax Eclipse Plus C18 column (2.1 × 100mm, 1.8 μm particle diameter, Agilent Technologies, Santa Clara, CA) at 30 °C. Mobile phase composition was water with ammonium acetate (5 mM at pH 3.0 adjusted with formic acid, solvent A) and acetonitrile (0.1% formic acid, solvent B) for acquisition in positive ionization mode (ESI+), whereas water (0.01% formic acid, solvent A) and acetonitrile (0.01% formic acid, solvent B) were used for acquisition in negative ionization mode (ESI-). In both methods, the gradient program was as follows: 0 min, 0% B; 12 min, 80% B; 14 min, 100% B; 16 min, 100% B; 17 min, 0% B. A flow rate of 0.5 mL/min and an injection volume of 20 μL were employed. The mass spectrometer was operated in MS and MS/MS modes for the structural analysis of all compounds. MS parameters were the following: capillary voltage, 4000 V; nebulizer pressure, 40 psi; drying gas flow rate, 10 L/min; gas temperature, 350 ºC; skimmer voltage, 45 V; fragmentor voltage, 110 V. The MS and Auto MS/MS modes were set to acquire m/z ranging between 50-1100 and 50-800 amu, respectively, operating at a resolving power of 40,000 at

m/z values around 1000, scan rate of 5 spectra per second, and scan time of 200 ms per spectrum. Auto MS/MS mode was operated by selecting 2 precursor ions per cycle at an abolute threshold of 200 count. Withanolides were tentatively identified based on their structural analogy with the standard Withanolide A.

2.3.2 Gas chromatography-mass spectrometry (GC-q-TOF-MS)

The main volatile compounds in the extracts were analysed using GC-q-TOF-MS. The analysis was performed employing a 7890B Agilent system (Agilent Technologies, Santa Clara, CA, USA) coupled to a quadrupole time-of-fight mass spectrometer (q-TOF-MS) 7200 (Agilent Technologies, Santa Clara, CA, USA) equipped with an electronic ionization (EI) source. One microlitle of each extract was injected with a split ratio of 10:1 and a split flow of 8.4 mL min^{-1} with the injector at a temperature of 250 °C. The separation of compounds was achieved using an Agilent Zorbax DB5- MS + 10 m Duragard Capillary Column (30 m × 250 μm x 0.25 μm). Helium was used as carrier gas at a constant flow rate of 0.8 mL min^{-1}. The column temperature was maintained at 60 °C for 1 min, then increased at a rate of 10 °C/min to 325 °C, and held at this temperature for 10 min. MS parameters were the following: electron impact ionization at 70 eV, filament source temperature of 250 °C, quadrupole temperature of 150 °C, *m*/*z* scan range 50–600 amu at a rate of 5 spectra per second.

2.4 Identification strategy and structural elucidation

Structural elucidation of the detected compounds was performed based on accurate mass data obtained by HRMS, ion source fragmentation, MS/MS fragmentation patterns, generated molecular formulae and subsequent unsaturation degree calculation, along with MS database, on-line databases (e.g., Massbank, Pubmed, Google Scholar), and bibliographic search. Isobaric forms were elucidated or classified on the basis of the complementary information

provided by MS/MS product ion spectra obtained in both positive and negative ESI mode. Unambiguous identifications were achieved by comparing retention time, MS/MS diagnostic ions and accurately measured mass with that of commercial standards, when available.

For GC-MS analysis, systematic mass spectra deconvolution of chromatographic signals and tentative identification of unknown compounds was carried out using the Agilent Mass Hunter Unknown Analysis tool and the NIST Mass Spectral database (NIST MS Search 2.0).

3. RESULTS AND DISCUSSION

3.1 Phytochemical profiling by UPLC-q-TOF-MS/MS of P. peruviana calyx compounds

The PLE extract from *P. peruviana* calyx obtained under the optimal conditions (125 °C and 75/25 of EtOH/EtOAc (v/v) as extraction solvent) reported in our previous work [17], was analysed by UPLC-q-TOF-MS/MS. Untargeted analysis was performed to determine the profile of the main compounds present in the *P. peruviana* calyx extract. To cover a broad number of chemical structures and to obtain complementary structural information, MS data were acquired in both positive and negative ionization mode (ESI +/-)

The tentative identification of the phytochemicals was carried out based on the information provided by MS data (accurate mass, isotopic distribution and fragmentation pattern), the use of commercial standards, and data found in the literature. The identified compounds together with their retention times, observed molecular ions, exact mass error and the main fragments obtained by MS/MS fragment ions are summarized in Table 1. As can be seen, a total of 59 metabolites belonging to 4 different families, including phenolic acids, flavonoids and glycosylated flavonoids, withasteroids and sucrose esters were identified.

<**Table 1**. Tentatively identified compounds from *Physalis peruviana* calyx by LC-q-TOF-MS/MS analysis>

3.1.1- Phenolic acids and flavonoids

Since phenolic compounds contain hydroxyl and/or carboxylic acid groups, these molecules were detected in their deprotonated form in ESI(-) ionization mode. Sixteen major phenolic compounds were identified, including 8 phenolic acids and 8 flavonoids and glycosylated derivatives. The identity of gallic acid (P1), protocatechuic acid (P2), 4-hydroxybenzoic acid (P3), vanillic acid (P4), caffeic acid (P5), benzoic acid (P6), p-coumaric acid (P8) and ferulic acid (P11) was confirmed comparing retention time and MS data with those of commercial standards. In agreement with data reported in literature regarding cape gooseberry components [18], caffeic, ferulic, p-coumaric, gallic, and protocatechuic acid were the main phenolic acids detected. These results reveal the capacity of *P.peruviana* to accumulate a significant content of phenolic acids in the calyx, which had not been previously described. These compounds might have a preventive effect on colon cancer, as they are believed to participate in the inhibition of tumour promotion and progression [19]. Further research work is being carried out by our group in order to provide a foodomics evaluation of this activity.

Besides phenolic acids, flavonoids such as myricetin (P7), quercetin (P14), isorhamnetin (P15), kaempferol (P16), and the hexoside (most probably glucoside) and rutinoside derivatives of quercetin and kaempferol (P9-15) were the major phenolic compounds detected (see Figure 1). Quercetin, kaempferol, and quercetin rutinoside (rutin) were confirmed using reference standards, whereas the remaining glycoconjugates were tentatively identified based on the fragmentation data from MS/MS spectra and the reference mass spectrum of its corresponding aglycone moiety. Thus, compounds P10, P12 and P13 showed deprotonated molecular ions $[M-H]^-$ at m/z 463.0882 ($C_{21}H_{19}O_{12}^-$, Δm/z = 0.2 ppm), m/z 593.1512 ($C_{27}H_{29}O_{15}^-$, Δm/z =

-2.5 ppm), and m/z 447.0933 ($C_{21}H_{19}O_{11}^-$, Δm/z = -1.6 ppm), as determined by HRMS. These exact masses were attributed to the glycoconjugates quercetin-O-hexoside, kaempferol rutinoside and kaempferol-O-hexoside, respectively. This assignation was confirmed by the characteristic MS/MS fragmentation behaviour of flavonoid O-glycosides, whose MS/MS spectra showed $[M-162-H]^-$ and $[M-308-H]^-$ product ions, corresponding to the loss of hexose (most probably glucose) and rutinose moieties, respectively.

To our knowledge, these results reveal for the first time that *Physalis peruviana* calyx are important sources of phenolics, in agreement with data reported in literature for other species of *Physalis* [20]. The presence of relevant phenolic components in the calyx extracts are in accordance with the results obtained in our previous work on the total phenolic content and in vitro antioxidant activity assays [17].

3.1.2.- Withanolides

Although withanolides (C_{28}-steroidal lactones) are reported to be determined in positive ionization mode by generating the protonated molecular ion $[M+H]^+$, and $[M+NH_4]^+$ or $[M+Na]^+$ molecular adducts [21], comparable responses were obtained in negative ionization mode at the operation conditions employed in this work. Considering that mechanisms of fragmentation in ESI(+) and ESI(-) are expected to be different, both ionization modes were used in order to obtain complementary structural information. Thus, two chromatographic runs were conducted with different mobile phases. Deprotonated $[M-H]^-$ molecular ions were obtained in negative ionization mode using low concentration of formic acid (0.01%) in the aqueous phase, whereas protonated molecular ion $[M+H]^+$ and ammonium adducts $[M+NH_4]^+$were obtained in positive ion mode using ammonium acetate (5 mM at pH 3.0 adjusted with formic acid).

A total of 17 withanolide-type compounds (W1-W17) were tentatively identified within the retention time interval from 4 to 9 min (see Figure 2A). Withanolides W1, W4 and W17 showed deprotonated molecular ions at m/z 517.2443 $[C_{28}H_{37}O_9]^-$, 521.2756 $[C_{28}H_{44}O_9]^-$, and 469.2596 $[C_{28}H_{37}O_6]^-$, respectively, whereas the remaining withasteroids exhibited isobaric forms, as summarized in Table 1. Thus, W12, W13 and W14 shared the same molecular ion at m/z 485.2545 $[C_{28}H_{37}O_7]^-$; in the same way as W15 and W16 shared m/z 487.2701 $[C_{28}H_{37}O_7]^-$; W6, W8 and W10 the same m/z 501.2494 $[C_{28}H_{37}O_8]^-$; W7, W9 and W11 the same m/z 503.2650 $[C_{28}H_{39}O_8]^-$; and W2 and W3 the same m/z 519.2600 $[C_{28}H_{39}O_9]^-$ as their respective shared molecular ions. As determined by HRMS, molecular formulae of detected withanolides were confidently assigned with $\Delta m/z < 5$ppm mass accuracy, corresponding to C28-isoprenoids molecular structures with 8, 9 or 10 unsaturation degrees.

Next, MS/MS fragmentation data were analysed to tentatively elucidate the chemical structure of detected withanolides. As exemplified in Figures 2B-D, fragmentation pattern of withanolides is mainly characterized by the loss of water molecules (-18 Da) and subsequent cleavage/rearrangement of the lactone (Lac) moiety from the deprotonated molecular ion [M-H]-. Further dehydration fragments are also observed in MS/MS spectra, generated by several loses of water molecules from the ergostane moiety [M-H-Lac]-. This fragmentation pattern obtained in negative ionization mode is consistent with withanolide fragmentation pathways proposed in literature in positive ionization mode, although with subtle differences [22]. For instance, removal of lactone moiety from the deprotonated molecular ion was shown to occurs after C20-C21 bond cleavage, whereas lactone part cleaves between C17-C20 in protonated withanolides. Main MS/MS product ions obtained from ESI(-)- and ESI(+)-Q-TOF analysis of withanolides are summarized in Tables 2 and 3, respectively.

288 <**Table 2**. Assignation of $[M-H]^-$ precursor and ESI(-)-MS/MS product ions of tentatively
289 identified withanolides.>
290 <**Table 3**. Assignation of $[M+NH_4]^+$ precursor and ESI(+)-MS/MS product ions of tentatively
291 identified withanolides.>
292
293 As shown in Tables 2 and 3, small variations within the same type of product ions are observed
294 between different withanolides due to minor differences in oxygen functionalities on the
295 lactone and the ergostane moiety. As observed for most of the identified compounds, [M-H-
296 Lac]- and $[M+H-Lac]^+$ ions showed, at least, two successive loses of water molecules with
297 significant abundance in the MS/MS spectra. This behaviour is explained in literature due to
298 the presence of 4/5-hydroxyl and 5,6/6,7-epoxide groups in the ergostane moiety, whose
299 removal might be favoured by the generation of extended conjugation with the C1-αβ-
300 unsaturated keto group [22]. Additional structural information about hydroxylation of the
301 lactone moiety can be obtained in negative ionization mode, as this part is removed as a neutral
302 moiety while the ergostane retain the charge. A neutral loss of 140.0473 Da corresponds to a
303 hydroxylated lactone (see Figure 2B-C), most probably at C27, whereas loss of 124.0524 Da
304 correspond to non-hydroxylated lactones (see Figure 2D).
305 Thus, compound W1, which shows m/z 377.1968 as [M-H-Lac]- in ESI(-) MS/MS spectrum
306 (Figure 2B), is expected to contain a 4-hydroxy-5,6-epoxy group in the ergostane moiety
307 according to the proposed elucidation strategy proposed by Ghulam [22], as this compound
308 exhibits m/z 299 $[M+H-Lac-H_20]^+$ as the most intense peak of the water removal cluster ions
309 from the ergostane moiety in ESI(+) MS/MS spectra (see fragment ions in Table 3). In addition,
310 a 140 Da neutral loss in ESI(-) MS/MS spectrum indicates C27 hydroxylation in the lactone
311 part. Therefore, this compound has shown to present a withaferin A-based structure with 3
312 additional OH groups, most probably at positions C14, C17 and C20. Hydroxylation at C17

and C20 is supported by MS/MS fragmentation in positive ionization mode, as they are contained in the lactone part after the C17-C20 bond cleavage and subsequent neutral loss, whereas 14-OH is proposed as the most probable option for the third OH group. Hence, W1 was tentatively identified as 27-hydroxylated 4β-hydroxywithanolide E isomer.

Some similarities were observed between the fragmentation patterns of W1, W2, W3 and W4, which also share the same number of carbon and oxygen atoms, although minor differences in the degree of unsaturation are observed. W2 was tentatively identified as 2,3-dehydro-27-hydroxy-4β-hydroxywithanolide E isomer, as it contains one unsaturation less than W1 in the ergostane. This assumption is supported by the m/z 299 $[M+H-Lac]^+$ ion as the most intense peak of the ergostane moiety in ESI(+) MS/MS spectrum, which indicates that the removal of the OH group in C4 might not be favoured due to the absence of the unsaturation in C2-C3. C27 hydroxylation in W2 is again supported by the 140 Da neutral loss in ESI(-) MS/MS spectrum.

Similar fragmentation was observed in the ergostane between W1-W3 and W2-W4, respectively. Thus, W3 was tentatively identified as a hydroxylated-4β-hydroxywithanolide derivative, containing one unsaturation less that W1 in the lactone moiety, whereas W4 was identified as a 2,3-dehydro-27-hydroxylated withanolide derivative, showing one degrees of unsaturation less compared to W2 in the lactone part.

Following this elucidation strategy and according to MS/MS data, withanolides W6, W7, W8, W9, W15 and W16 exhibited a similar base structure corresponding to mono or di-hydroxylated withanolide derivatives, with different degrees of unsaturation. As a representative example, Figure 2C illustrates the MS/MS spectra of W8 (di-hydroxylated withaferin A), showing the neutral loss of a C27-hydroxylated lactone (-140 Da) to yield a di-hydroxylated ergostane moiety (m/z 361.2023). On the other hand, withanolides W10, W11, W12, W13 and W14, shared a common structure based on withanolide E, presenting a non-

hydroxylated lactone moiety unlike the abovementioned compounds. Figure 2D shows the ESI(-)-MS/MS spectrum of W10 as representative example of the MS/MS fragmentation pattern of a withanolide E derivative. Compared to W12, W13 and W14, withanolides W9, W10 and W11 contain an additional OH group in the ergostane part, evidenced by the 16/18 Da mass difference in $[M+H-Lac]^+$ moiety (see Table 3).

3.1.3- Sucrose esters

The product ion chromatogram obtained by ESI(-)-Q-TOF analysis of the target *P. peruviana* calyx extracts revealed the presence of a large group of compounds, which represented the major contribution in the chromatogram in terms of peak areas (Figure 3A). Despite their presence all throughout the chromatogram, they are mostly abundant within the retention time range from 9 to 15 min showing MS/MS fragmentation spectra characterized by the successive loss of acyl groups attached to a sucrose moiety (see Figure 3B-C). These compounds, previously described in literature as sucrose esters, are considered the main protective constituents of the resin covering the inner parts of the calyx of several *Physalis* species, exhibiting aphicidal, molluscidal, and antifeedant activities [23]. A total of 23 acylsucroses were tentatively identified in the calyx extracts obtained under the optimized PLE conditions. Deprotonated molecular ions and diagnostic product ion corresponding to successive loses of acyl residues of the tentatively identified sucrose esters are summarized in Table 4. Although hydroxyl groups at C6, C1', and C6' positions of the sucrose moiety are expected to be more reactive due to lower steric hindrance, C2, C3, C1’ and C3’ are described in literature as the most favoured positions to be esterified by saturated fatty acids in *P. peruviana* [12]. Hence, the identified ester residues and the positions of all substituents in the disaccharide structure were tentatively assigned according to MS/MS data and the most plausible structure according to data reported in literature. Figure 3B-C illustrates the MS/MS product ion assignation for

363 structural elucidation of di-O-isobutanoyl-O-decanoylsucrose and di-O-isobutanoyl-O-
364 dodecanoyl-O-(2-methylbutanoyl)sucrose, respectively, as representative congeners of the
365 identified sucrose esters. In accordance to previous references, hydroxyl group at C2 is
366 frequently esterified by the largest fatty acid, whereas other favoured positions of the sucrose
367 moiety can be O-acylated by isobutanoyl or 2-methylbutanoyl groups. As determined by
368 HRMS, product ions at m/z 481.1927 and 411.1508 ($\Delta m/z < 3ppm$) are commonly observed
369 as major peaks in MS/MS fragmentation spectra for most of detected sucrose esters,
370 corresponding to the mono- and di-O-isobutanoylated sucrose moiety, respectively. This
371 assignation supports the identification of S1 as O-isobutanoyl sucrose and S2 and S3 as di-O-
372 isobutanoyl sucrose esters showing deprotonated molecular ions at m/z 411.1510 and
373 481.1936, respectively (see tentative assignation of MS/MS product ions in Table 4). Fragment
374 ion m/z 481 is generated after removal of the long chain acyl group in compounds S6, S7, S8,
375 S11, S13, S14, and after the successive loss of the long chain fatty acid ester and an isobutanoyl
376 group in tetra esterified sugars S17, S18 and S19. Further removal of another O-isobutanoyl
377 group gives rise to fragment ion m/z 411 for S6, S7, S8, S11, S13, S14. Similarly, sucrose
378 tetraester S20, S21, S22 exhibit m/z 481 product ion losing the long chain acyl group and the
379 2-methylbutanoyl ester. However, triesters S9, S12, S15, and S16 show m/z 411 rather than
380 m/z 481, as they can only generate mono-isobutanoylated fragments.
381 Long chain fatty acid esters at C2 were elucidated as octanoic acid, nonanoic acid, decanoic
382 acid and dodecanoic acid, as deduced form the detected fragment ions $[R_2]$- at m/z 143.1072
383 $[C_8H_{15}O_2]^-$, 157.1229 $[C_9H_{17}O_2]^-$, 171.1385 $[C_{10}H_{19}O_2]^-$ and 199.1698 $[C_{12}H_{23}O_2]^-$,
384 respectively. Additional structural information was obtained from the neutral loss generated
385 from the deprotonated molecular ion yielding ion fragments above m/z 431. Thus, 70
386 Da neutral loses were assigned to the removal of an isobutanoyl group, whereas the loss of 84
387 or 82 Da was accounted to the removal of 2-methylbutanoyl fragments.

<Table 4. Assignation of $[M-H]^-$ precursor and ESI(-)-MS/MS product ions of tentatively identified sucrose esters.>

3.2 Phytochemical profiling by GC-q-TOF-MS of P. peruviana calyx compounds

The analysis of unknowns based on the GC-q-TOF data corresponding to the volatile fraction of the aforementioned PLE extract obtained under optimal conditions led to the tentative identification of fifty-three compounds, classified in different families of compounds, mainly mono-, sesqui-, di- and triterpenes, ionones, phenolics, steroids and phytol derivatives. These phytochemicals were identified based on the positive match of experimental mass spectra with theoretical MS data in the NIST MS database, calculated mass accuracy for the $[M]^+$ and data reported in literature. Table 5 summarizes the complete list of GC identified phytochemicals, including their corresponding characteristic GC-MS parameters (e.g. retention time, reverse match (R. Match) values given by NIST database, molecular formula, m/z $[M]^+$ exact mas, calculated mass error and MS/MS fragments), that confirm their unambiguous identification. Most of tentatively identified metabolites (45 out of 53) showed a match factor value above 70 and most of them with satisfactory mass accuracy ($\Delta m/z < 5$ ppm) for the molecular ion. The use of EI, as hard ionization source, may led to undetectable molecular ions for some compounds due to the high fragmentation. In this case, the use of softer ionization alternatives (e.g. chemical ionization) is strongly recommendable for improving mass accuracy of the molecular ion estimation.

In terms of relative abundance and bioactive properties, the most relevant identified phytoconstituents were diterpenes such as copalol, sclareol oxide, phytol, dihydromanoyl oxide; phytol derivatives such as tocopherols; as well as phytosteroids such as stigmastadienol, ergostenol, sitosterol and cholestane derivatives. Terpenic compounds have been associated with plant protection mechanisms against oxidative stress [24]. In particular, the scientific

414 literature describes phytosterols as bioactive compounds of great interest because of its
415 antioxidant capacity and impact on health. They are reported as anti-inflammatory, antitumor,
416 antibacterial and antifungal and hypocholesterolemic compounds. Their presence at significant
417 levels in the oil extracted from the skin and pulp of *P. peruviana* L. has been reported [25].
418 Tocopherols (vitamin E) properties are attributed mainly to its ability to prevent cell membrane
419 damage by free radicals, by reducing the levels of lipid peroxides [18]. In this regard, the
420 glycosylated form of α-tocopherol, the most efficient of these components, was identified in
421 the obtained PLE extracts of *P. peruviana* calyx. β-, δ-, and γ- tocopherol were also detected,
422 being the latter the most abundant form of vitamin E in the analyzed extract.
423 Other minor compounds including a broad group of sesquiterpenes as well as monoterpenes
424 such as limonene and terpineol, along with ionone derivatives were also determined, which
425 contribute not only to the floral flavor but may also provide antiseptic, anti-microbial and anti-
426 inflammatory properties to the calyx extract of *P. peruviana* [26].
427
428 <**Table 5**. Tentatively identified compounds from Physalus Peruviana calyx by GC-q-TOF-
429 MS analysis.>
430
431 **4. CONCLUSIONS**
432 A multi-analytical platform based on pressurized-liquid extraction, in vitro assays and LC/GC
433 coupled to q-TOF mass spectrometry for food by-products valorization was successfully
434 developed in this work, demonstrating the great potential of the proposed strategy to obtain and
435 characterize potential bioactive compounds from *P. peruviana* calyx as case study. The results
436 obtained from the phytochemical characterization by LC and GC coupled to q-TOF- MS/MS
437 reveal that *P. peruviana* calyx is an important source of a broad variety of health-promoting
438 compounds such as withanolides, phenolic acids, flavonoids, anti-inflammatory sucrose esters,

terpenoids, phytosterols, and phytol derivatives (vitamin E). Complementary identification strategies were applied in this work, including comparative evaluation of MS/MS product ion spectra obtained in both positive and negative ESI mode. Based on the differential fragmentation of negative [M-H]- and positive $[M+NH_4]^+$ molecular ions, relevant structural information was obtained for tentative identification and structural classification of withanolides compounds, which is considered a valuable contribution to the limited information in literature about the MS/MS structural analysis of withanolides in ESI(-) mode. In this paper, analytical strategies based on LC and GC coupled to HRMS are developed for the untargeted analysis and structural elucidation of compounds of interest in food by-products from a qualitative point of view. Quantitative information can be obtained with the proposed LC and GC approaches after appropriate method validation, making use of the qualitative parameters obtained from the current profiling analysis (e.g., diagnostic product ions, retention time, ionization mode).

The obtained results highlight the importance of using complementary analytical platforms, operating in multiple ionization modes and applying an integrated elucidation strategy to unravel complex phytochemical samples such as those found in natural extracts and/or traditional medicine preparations. Considering the potential bioactivity of the obtained extract, Foodomics studies are now being carried out in our lab in order to better understand the promising benefits of *P. peruviana* calyx extract on human health.

Acknowledgements

This research was supported by the COOPA20145 project from CSIC (Programa de Cooperación Científica para el Desarrollo "i-COOP+"). G.A.-R. would like to acknowledge the Ministry of Economy and Competitiveness for a "Juan de la Cierva" postdoctoral grant. The authors also thank the support from the AGL2017-89417-R project.

References

[1] M.L. Olivares-Tenorio, M. Dekker, R. Verkerk, M.A.J.S. van Boekel, Health-promoting compounds in cape gooseberry (Physalis peruviana L.): Review from a supply chain perspective, Trends Food Sci. Technol. 57 (2016) 83–92. doi:10.1016/j.tifs.2016.09.009.

[2] M.F. Ramadan, Bioactive phytochemicals, nutritional value, and functional properties of cape gooseberry (Physalis peruviana): An overview, Food Res. Int. 44 (2011) 1830–1836. doi:10.1016/j.foodres.2010.12.042.

[3] S.M. Mokhtar, H.M. Swailam, H.E.S. Embaby, Physicochemical properties, nutritional value and techno-functional properties of goldenberry (Physalis peruviana) waste powder concise title: Composition of goldenberry juice waste, Food Chem. 248 (2018) 1–7. doi:10.1016/j.foodchem.2017.11.117.

[4] S. Rodríguez, E. Rodríguez, Efecto de la ingesta de Physalis peruviana (aguaymanto) sobre la glicemia postprandial en adultos jóvenes, Rev. Médica Vallejiana. 1 (2007) 43–52. doi:10.1590/0100-2945-441/13.

[5] L.A. Franco, G.E. Matiz, J. Calle, R. Pinzón, L.F. Ospina, Actividad antinflamatoria de extractos y fracciones obtenidas de cálices de Physalis peruviana L., Biomédica. 27 (2007) 110–115.

[6] S.M. Llano, A.M. Muñoz-jiménez, C. Jiménez-cartagena, J. Londoño-londoño, S. Medina, Untargeted metabolomics reveals speci fi c withanolides and fatty acyl glycoside as tentative metabolites to differentiate organic and conventional Physalis peruviana fruits, Food Chem. 244 (2018) 120–127. doi:10.1016/j.foodchem.2017.10.026.

[7] O. Rop, J. Mlcek, T. Jurikova, M. Valsikova, Bioactive content and antioxidant capacity of Cape gooseberry fruit, Cent. Eur. J. Biol. 7 (2012) 672–679.

490 doi:10.2478/s11535-012-0063-y.
491 [8] G. Yıldız, N. İzli, H. Ünal, V. Uylaşer, Physical and chemical characteristics of
492 goldenberry fruit (Physalis peruviana L.), J. Food Sci. Technol. 52 (2015) 2320–2327.
493 doi:10.1007/s13197-014-1280-3.
494 [9] J.M. Calderón, N. Ruiz, L. Castellanos, Within and between plant variation of 4β-
495 hydroxiwithanolide E in cape gooseberry (Physalis peruviana; Solanaceae), Biochem.
496 Syst. Ecol. 41 (2012) 21–25. doi:10.1016/j.bse.2011.12.009.
497 [10] L.A. Franco, Y.C. Ocampo, H.A. Gómez, R. De La Puerta, J.L. Espartero, L.F.
498 Ospina, Sucrose esters from Physalis peruviana calyces with anti-inflammatory
499 activity, Planta Med. 80 (2014) 1605–1614. doi:10.1055/s-0034-1383192.
500 [11] L.X. Chen, G.Y. Xia, Q.Y. Liu, Y.Y. Xie, F. Qiu, Chemical constituents from the
501 calyces of Physalis alkekengi var. franchetii, Biochem. Syst. Ecol. 54 (2014) 31–35.
502 doi:10.1016/j.bse.2013.12.030.
503 [12] C.R. Zhang, W. Khan, J. Bakht, M.G. Nair, New antiinflammatory sucrose esters in
504 the natural sticky coating of tomatillo (Physalis philadelphica), an important culinary
505 fruit, Food Chem. 196 (2016) 726–732. doi:10.1016/j.foodchem.2015.10.007.
506 [13] M. Sang-Ngern, U.J. Youn, E.J. Park, T.P. Kondratyuk, C.J. Simmons, M.M. Wall, M.
507 Ruf, S.E. Lorch, E. Leong, J.M. Pezzuto, L.C. Chang, Withanolides derived from
508 Physalis peruviana (Poha) with potential anti-inflammatory activity, Bioorganic Med.
509 Chem. Lett. 26 (2016) 2755–2759. doi:10.1016/j.bmcl.2016.04.077.
510 [14] D. Ren, L. Ran, C. Yang, M. Xu, L. Yi, Integrated strategy for identifying minor
511 components in complex samples combining mass defect, diagnostic ions and neutral
512 loss information based on ultra-performance liquid chromatography-high resolution
513 mass spectrometry platform: Folium Artemisiae Argy, J. Chromatogr. A. 1550 (2018)
514 35–44. doi:10.1016/j.chroma.2018.03.044.

[15] W. Zhao, Z. Shang, Q. Li, M. Huang, W. He, Z. Wang, J. Zhang, Rapid screening and identification of daidzein metabolites in rats based on uhplc-ltq-orbitrap mass spectrometry coupled with data-mining technologies, Molecules. 23 (2018). doi:10.3390/molecules23010151.

[16] Z. Shang, W. Cai, Y. Cao, F. Wang, Z. Wang, J. Lu, J. Zhang, An integrated strategy for rapid discovery and identification of the sequential piperine metabolites in rats using ultra high-performance liquid chromatography/high resolution mass spectrometery, J. Pharm. Biomed. Anal. 146 (2017) 387–401. doi:10.1016/j.jpba.2017.09.012.

[17] D. Ballesteros-Vivas, G. Alvarez-Rivera, A. Sánchez-Camargo, E. Ibáñez, F. Parada-Alfonso, A. Cifuentes, A multi-analytical platform based on pressurized-liquid extraction, in vitro assays and liquid chromatography/gas chromatography coupled to q-TOF mass spectrometry for food by-products revalorization. Part 1: Withanolide-rich extracts from goldenberry (Physalis peruviana L) calyces as case study. J. Chromatogr. A. (n.d.).

[18] M.L. Olivares-Tenorio, M. Dekker, R. Verkerk, M.A.J.S. van Boekel, Health-promoting compounds in cape gooseberry (Physalis peruviana L.): Review from a supply chain perspective, Trends Food Sci. Technol. 57 (2016) 83–92. doi:10.1016/j.tifs.2016.09.009.

[19] L. Rosa, N. Silva, N. Soares, M. Monteiro, A. Teodoro, Anticancer Properties of Phenolic Acids in Colon Cancer – A Review, J. Nutr. Food Sci. 06 (2016) 1–7. doi:10.4172/2155-9600.1000468.

[20] J.R. Medina-Medrano, N. Almaraz-Abarca, M. Socorro González-Elizondo, J.N. Uribe-Soto, L.S. González-Valdez, Y. Herrera-Arrieta, Phenolic constituents and antioxidant properties of five wild species of Physalis (Solanaceae), Bot. Stud. 56

(2015). doi:10.1186/s40529-015-0101-y.

[21] D. Patil, M. Gautam, S. Mishra, S. Karupothula, S. Gairola, S. Jadhav, S. Pawar, B. Patwardhan, Determination of withaferin A and withanolide A in mice plasma using high-performance liquid chromatography-tandem mass spectrometry: Application to pharmacokinetics after oral administration of Withania somnifera aqueous extract, J. Pharm. Biomed. Anal. 80 (2013) 203–212. doi:10.1016/j.jpba.2013.03.001.

[22] S. Ghulam Musharraf, A. Ali, R. Azher Ali, S. Yousuf, A.U. Rahman, M. Iqbal Choudhary, Analysis and development of structure-fragmentation relationships in withanolides using an electrospray ionization quadropole time-of-flight tandem mass spectrometry hybrid instrument, Rapid Commun. Mass Spectrom. 25 (2011) 104–114. doi:10.1002/rcm.4835.

[23] C.A. Bernal, L. Castellanos, D.M. Aragón, D. Martínez-Matamoros, C. Jiménez, Y. Baena, F.A. Ramos, Peruvioses A to F, sucrose esters from the exudate of Physalis peruviana fruit as α-amylase inhibitors, Carbohydr. Res. 461 (2018) 4–10. doi:10.1016/j.carres.2018.03.003.

[24] K.A. Wojtunik, L.M. Ciesla, M. Waksmundzka-Hajnos, Model studies on the antioxidant activity of common terpenoid constituents of essential oils by means of the 2,2-Diphenyl-1-picrylhydrazyl method, J. Agric. Food Chem. 62 (2014) 9088–9094. doi:10.1021/jf502857s.

[25] L.A. Puente, C.A. Pinto-Muñoz, E.S. Castro, M. Cortés, Physalis peruviana Linnaeus, the multiple properties of a highly functional fruit: A review, Food Res. Int. 44 (2011) 1733–1740. doi:10.1016/j.foodres.2010.09.034.

[26] D.L. and K.-W. Wang, Natural New Sesquiterpenes: Structural Diversity and Bioactivity, Curr. Org. Chem. 20 (2016) 994–1042. doi:http://dx.doi.org/10.2174/1385272819666151008014405.

565

566

567

568

569

570

571

572

573 **Figure captions**

574

575 **Figure 1.** Most abundant phenolic compounds identified in *P. peruviana* calyx extracts by
576 LC- ESI(-)-q-TOF analysis. Overlapped extracted ion chromatograms of most abundant
577 phenolic acids and flavonoids (A), including for clarity in a second chromatogram the EIC of
578 rutin and quercetin as the major flavonoids found (B).

579

580 **Figure 2.** LC- ESI(-)-q-TOF extracted ion chromatogram of detected withanolides (A) and
581 MS/MS fragmentation spectra of W1 (B), W8 (C) and W10 (D).

582

583 **Figure 3.** LC- ESI(-)-q-TOF extracted ion chromatogram of detected sucrose esters (upper
584 chromatogram) and MS/MS fragmentation spectra of SU14 (middle) and SU22 (below).

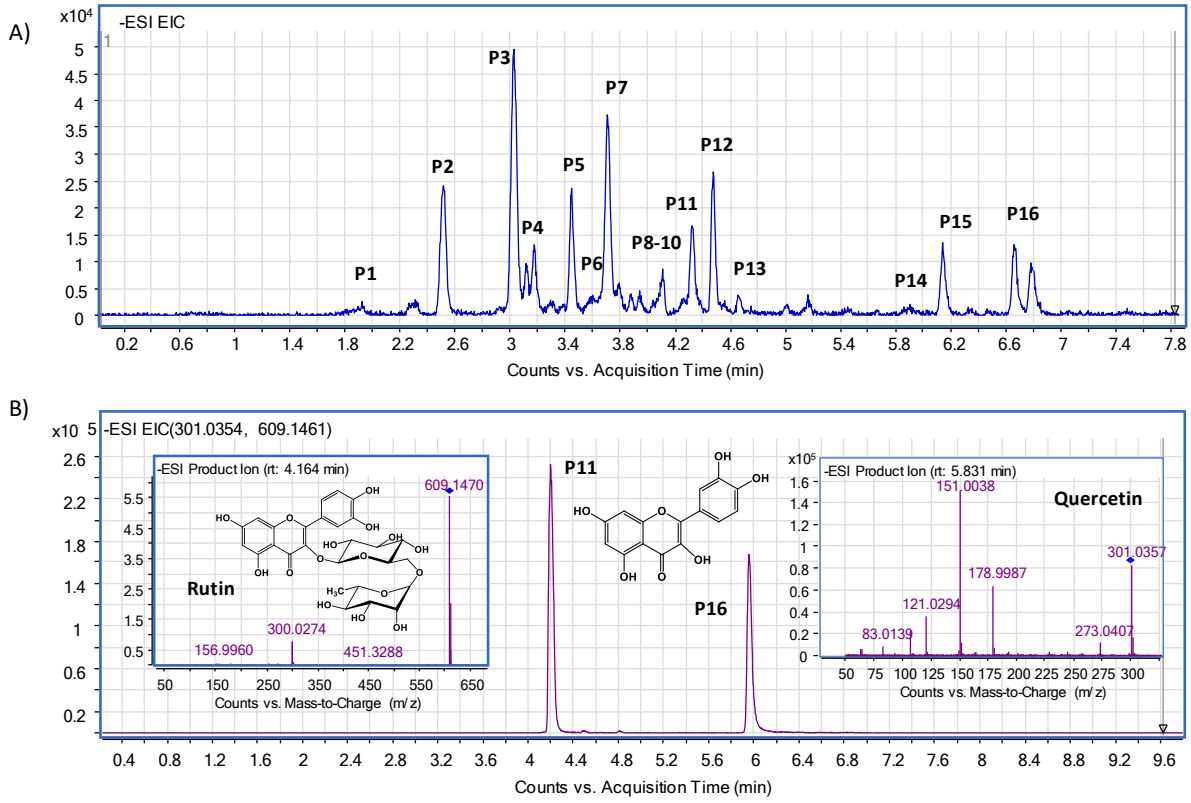

Figure 1. Most abundant phenolic compounds identified in *P. peruviana* calyx extracts by LC-ESI(-)-q-TOF analysis. Overlapped extracted ion chromatograms of most abundant phenolic acids and flavonoids (A), including for clarity in a second chromatogram the EIC of rutin and quercetin as the major flavonoids found (B).

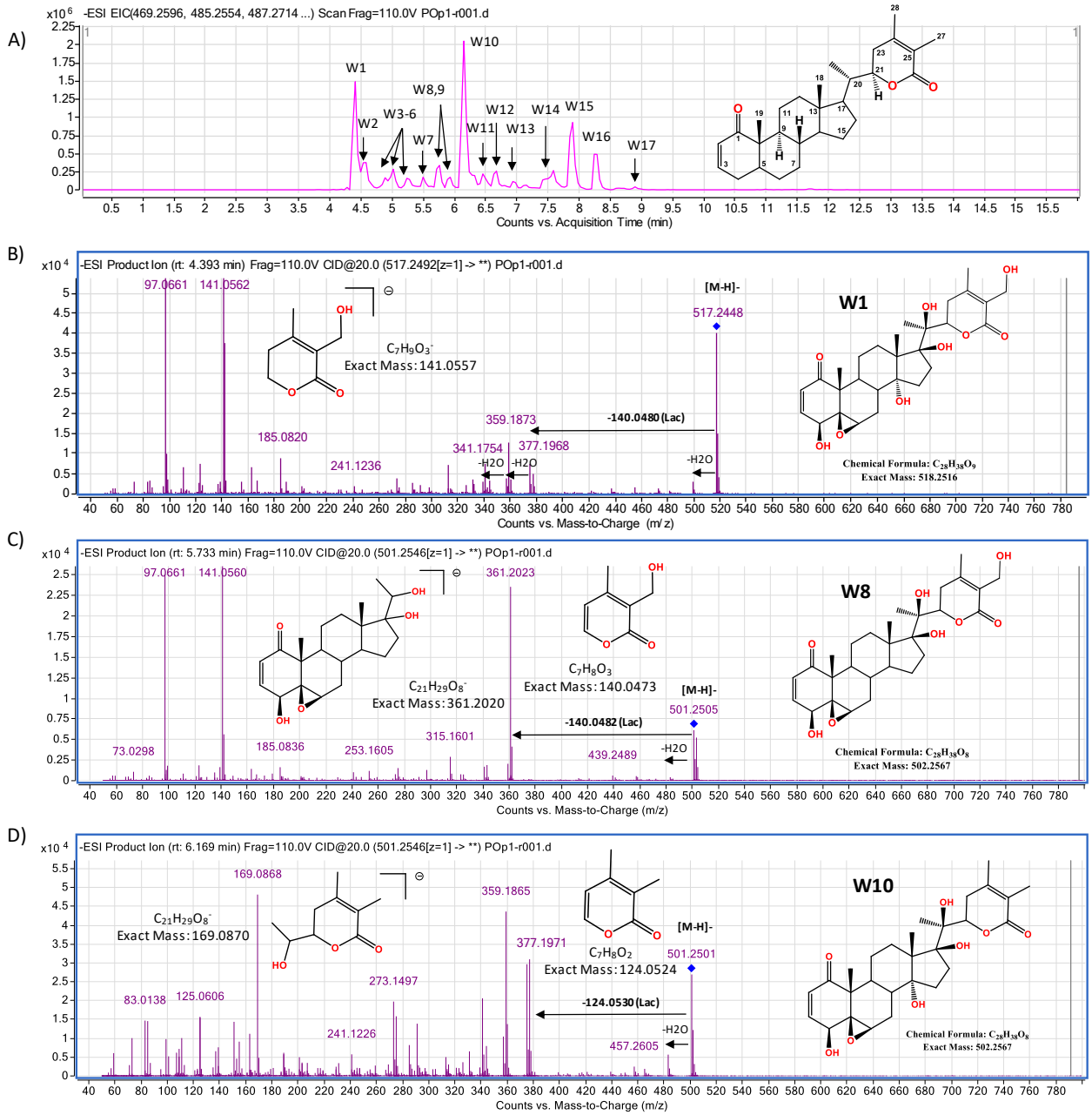

Figure 2. LC-ESI(-)-q-TOF extracted ion chromatogram of detected withanolides (A) and MS/MS fragmentation spectra of W1 (B), W8 (C) and W10 (D).

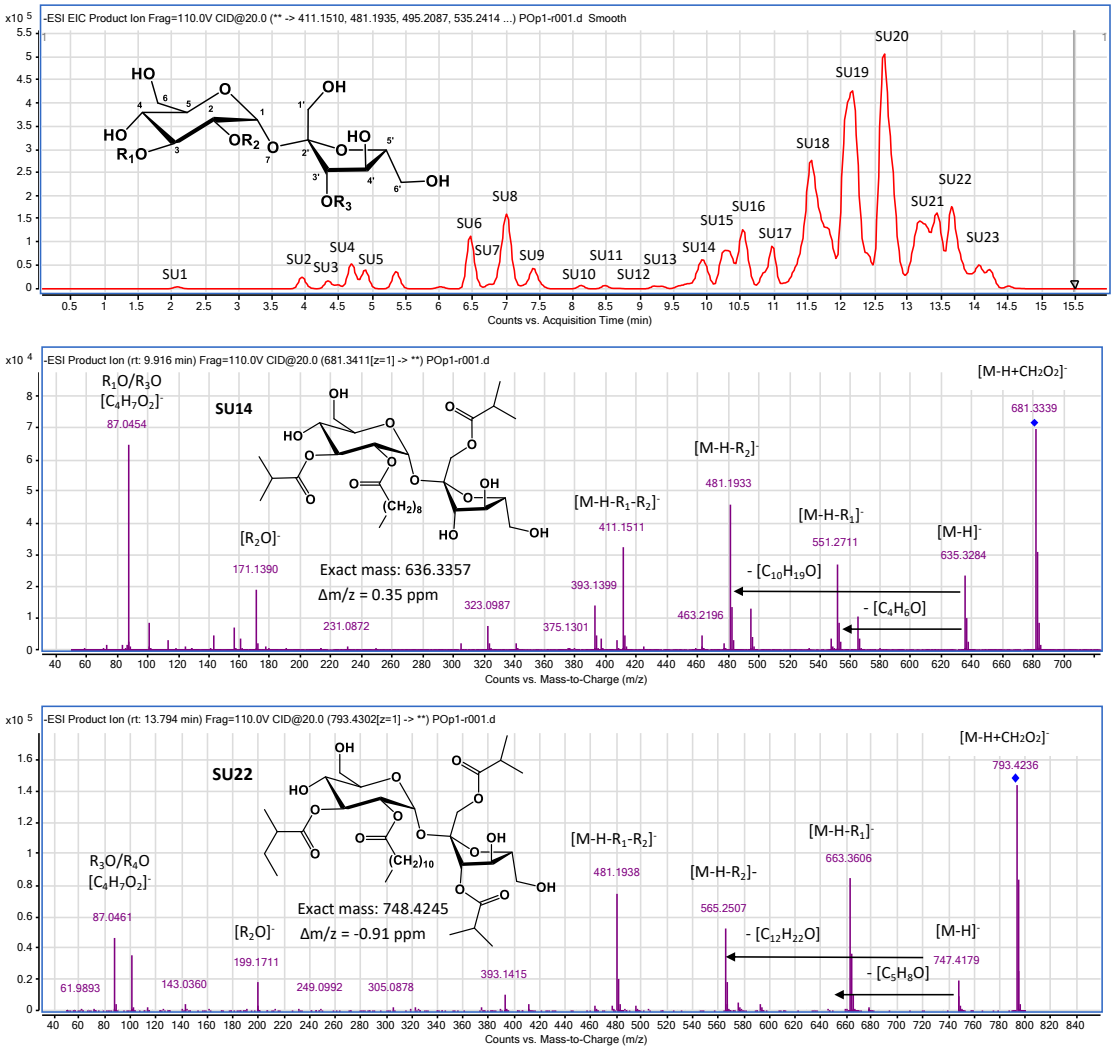

Figure 3. LC-ESI(-)-q-TOF extracted ion chromatogram of detected sucrose esters (upper chromatogram) and MS/MS fragmentation spectra of SU14 (middle) and SU22 (below).

Table 1. Tentatively identified compounds from *Physalis peruviana* calyces extract by LC-q-TOF-MS/MS analysis

Family	Peak Number	Ret. time (min)	Tentative identification	Formula	Monoisotopic mass	$[M-H]^-$ (m/z)	Error (ppm)	MS/MS product ions (m/z)
Phenolic acids	P1	1.966	Gallic acid*	C7H6O5	170.0215	169.0142	0.9	125, 107, 97, 79
	P2	2.484	Protocatechuic acid*	C7H6O4	154.0266	153.0193	-2.4	119, 109, 92, 81
	P3	3.048	4-HBA*	C7H6O3	138.0317	137.0244	-1.3	123, 108, 92, 81
	P4	3.360	Vanillic acid*	C8H8O4	168.0423	167.0350	-1.3	152, 124, 110
	P5	3.401	Caffeic acid*	C9H8O4	180.0423	179.0350	-1.2	151, 135, 122, 107
	P6	3.665	Benzoic acid*	C7H6O2	122.0368	121.0295	-0.8	111, 94, 67
	P8	4.014	p-Coumaric acid*	C9H8O3	164.0473	163.0401	-3.3	147, 119, 110
	P11	4.319	Ferulic acid*	C10H10O4	194.0579	193.0506	-2.4	178, 134, 106
Flavonoids	P7	3.752	Myricetin	C15H10O8	318.0376	317.0303	-1.6	179, 153, 113
	P9	4.101	Rutin*	C27H30O16	610.1534	609.1461	-2.8	300, 271, 151
	P10	4.275	Quercetin-hexoside	C21H20O12	464.0955	463.0882	0.2	301, 179, 137
	P12	4.449	Kaempferol-rutinoside	C27H30O15	594.1585	593.1512	-2.5	285, 255, 227, 125
	P13	4.624	Kaempferol-hexoside	C21H20O11	448.1006	447.0933	-1.6	284, 255, 227, 151
	P14	5.844	Quercetin	C15H10O7	302.0427	301.0354	-3.4	179, 151, 121
	P15	6.105	Isorhamnetin	C16H12O7	316.0583	315.0510	-1.8	300, 271, 152
	P16	6.585	Kaempferol*	C15H10O6	286.0477	285.0405	-1.2	257, 229, 151
Withanolides	W1	4.362	27-hydroxy-4β-hydroxywithanolide E isomer	C28H38O9	518.2516	517.2443	-2.9	91, 141, 341, 359, 375, 411, 481
	W2	4.580	2,3-Dihydro-27-hydroxy-4β-hydroxywithanolide E isomer	C28H40O9	520.2672	519.2600	-1.6	97, 141, 343, 361, 377, 501
	W3	4.972	Hydroxylated 4β-hydroxywithanolide E derivative	C28H40O9	520.2672	519.2600	0.7	139, 189, 341, 359, 377, 501

Table 1 (Cont.)

Family	Peak Number	Ret. time (min)	Tentative identification	Formula	Monoisotopic mass	$[M-H]^-$ (m/z)	Error (ppm)	MS/MS product ions (m/z)
	W4	5.120	2,3-Dihydro-hydroxylated 4β-hydroxywithanolide E derivative	C28H42O9	522.2829	521.2756	3.1	379, 361, 343, 315
	W6	5.303	17,27-Dihydroxylated withanolide D isomer 1	C28H38O8	502.2567	501.2494	0.4	97, 141, 343,2, 359, 483
	W7	5.495	2,3-Dihydro-17,27-hydroxylated withanolide D derivative	C28H40O8	504.2723	503.2650	-0.5	73, 175, 325, 343, 361
	W8	5.757	17,27-Dihydroxylated withanolide D isomer 2	C28H38O8	502.2567	501.2494	-2.2	97, 141, 343,2, 361, 485
	W9	5.910	Dihydro-4β-hydroxywithanolide E	C28H40O8	504.2723	503.2650	-0.9	169, 125, 325, 343, 361, 379, 485
	W10	6.192	4β-Hydroxywithanolide E	C28H38O8	502.2567	501.2494	-2.0	169, 273, 341, 359, 377, 483
	W11	6.323	2,3-Dihydro-4β-hydroxywithanolide E	C28H40O8	504.2723	503.2650	2.7	169, 343, 361, 377, 485
	W12	6.628	Withanolide E isomer 1	C28H38O7	486.2618	485.2545	-1.9	169, 307, 325, 343, 361, 449, 467
	W13	6.933	Withanolide E isomer 2	C28H38O7	486.2618	485.2545	0.0	169, 325,343,361, 467
	W14	7.587	Withanolide E isomer 3	C28H38O7	486.2618	485.2545	-0.9	
	W15	7.935	2,3-Dihydro-27-hydroxylated withanolide D isomer 1	C28H40O7	488.2774	487.2701	-2.6	141, 283, 327, 345
	W16	8.284	2,3-Dihydro-27-hydroxylated withanolide D isomer 2	C28H40O7	488.2774	487.2701	-1.6	169, 243, 327, 343, 469
	W17	8.981	Withanolide D isomer	C28H38O6	470.2668	469.2596	-0.1	125, 159, 345

Table 1 (Cont.)

Family	Peak Number	Ret. time (min)	Tentative identification	Formula	Monoisotopic mass	$[M-H]^-$ (m/z)	Error (ppm)	MS/MS product ions (m/z)
Sucrose esters	S1	2.096	O-isobutanoylsucrose	C16H28O12	412.1581	411.1508	-0.5	87, 323
	S2	3.990	Di-O-isobutanoylsucrose	C20H34O13	482.1999	481.1927	-1.7	87, 143, 323, 393, 411
	S3	4.338	Di-O-isobutanoylsucrose	C20H34O13	482.1999	481.1927	-1.9	87, 143, 323, 393, 411
	S4	4.731	O-isobutanoyl-O-(2-methylbutanoyl)sucrose	C21H36O13	496.2156	495.2083	-0.8	87, 161, 323, 411
	S5	4.905	O-isobutanoyl-O-octenoylsucrose	C24H40O13	536.2469	535.2396	-3.3	87, 143, 323, 393, 481
	S6	6.517	O-butanoyl-di-O-isobutanoylsucrose	C24H40O14	552.2418	551.2345	-0.3	87, 143, 393, 481,
	S7	6.779	Di-O-isobutanoyl-O-pentenoylsucrose	C25H40O4	564.2418	563.2345	-2.1	87, 143, 393, 481,
	S8	7.127	Di-O-isobutanoyl-O-(2-methylbutanoyl)-O-pentenoylsucrose	C25H42O14	566.2575	565.2502	0.1	87, 143, 393, 481,
	S9	7.280	O-isobutanoyl-O-(2-methylbutanoyl)-O-pentenoylsucrose	C26H42O14	578.2575	577.2502	-0.5	87, 143, 323, 411, 493,
	S10	8.130	O-decanoyl-O-isobutanoylsucrose	C26H46O13	566.2938	565.2865	0.2	87, 171, 323, 411, 477,
	S11	8.652	Di-O-isobutanoyl-O-octanoylsucrose	C28H48O14	608.3044	607.2971	-2.8	87, 143, 393, 481, 537
	S12	9.119	O-isobutanoyl-O-(2-methylbutanoyl)-O-octanoylsucrose	C29H50O14	622.3201	621.3128	0.0	87, 143, 323, 411, 495, 537,
	S13	9.350	Di-O-isobutanoyl-O-nonanoylsucrose	C29H50O14	622.3201	621.3129	-2.3	87, 157, 343, 411, 481, 551
	S14	9.853	Di-O-isobutanoyl-O-decanoylsucrose	C30H52O14	636.3357	635.3284	0.4	87, 171, 323, 411, 481, 551

Table 1 (Contd.)

Family	Peak Number	Ret. time (min)	Tentative identification.	Formula	Monoisotopic mass	$[M-H]^-$ (m/z)	Error (ppm)	MS/MS product ions (m/z)
	S15	10.288	O-decanoyl-O-isobutanoyl-O-(2-methylbutenoyl)sucrose	C31H52O14	648.3357	647.3284	-1.0	87, 171, 323, 411, 493, 565
	S16	10.593	O-decanoyl-O-isobutanoyl-O-(2-methylbutanoyl)sucrose	C31H54O14	650.3514	649.3441	1.7	87, 171, 323, 411, 495, 565
	S17	10.942	O-octanoyl-tri-O-isobutanoyl-sucrose	C32H54O15	678.3463	677.3390	-0.6	87, 143, 393, 481, 551, 607
	S18	11.509	O-nonanoyl-tri-O-isobutanoyl-sucrose	C33H56O15	692.3619	691.3546	-0.4	87, 157, 393, 481, 551, 621
	S19	12.119	O-decanoyl-tri-O-isobutanoyl-sucrose	C34H58O15	706.3776	705.3703	0.3	87, 171, 393, 481, 551, 635
	S20	12.729	Di-O-isobutanoyl-O-decanoyl-O-(2-methylbutanoyl)sucrose	C35H60O15	720.3932	719.3859	-0.6	87, 171, 393, 481, 565, 635
	S21	13.382	Di-O-isobutanoyl-O-decanoyl-O-(2-methylbutanoyl)sucrose	C36H64O14	720.4296	719.4223	-0.6	87, 171, 323, 411, 565
	S22	13.774	Di-O-isobutanoyl-O-dodecanoyl-O-(2-methylbutanoyl)sucrose	C37H64O15	748.4245	747.4172	-0.9	87, 199, 393, 481, 565, 663
	S23	13.968	O-dodecanoyl-O-isobutanoyl-O-nonanoylsucrose	C37H66O14	734.4453	733.4380	0.3	87, 157, 199, 323, 411, 593

* Identification confirmed by commercial standard

Table 2. Assignation of $[M-H]^-$ precursor and ESI(-) MS/MS product ions of tentatively identified withanolides.

Peak Number	Ret. Time (min)	$[M-H]^-$	$[M-H-H_2O]^-$	$[M-H-Lac]^-$	$[M-H-Lac-H_2O]^-$	$[M-H-Lac-2H_2O]^-$	$[M-H-Lac-3H_2O]^-$	Other MS/MS fragments		
W1	4.362	517	499	377	359	341	-	185	141	91
W2	4.580	519	501	379	361	343	-	185	141	97
W3	4.972	519	501	377	359	341	-	189	139	-
W4	5.120	521	503	379	361	343	-	187	135	83
W6	5.303	501	483	361	343	-	-	185	141	97
W7	5.495	503	485	361	343	325	307	187	135	97
W8	5.757	501	483	361	343	-	297	185	141	97
W9	5.910	503	485	379	361	343	325	169	125	87
W10	6.192	501	483	377	359	341	-	189	169	83
W11	6.323	503	485	379	361	343	-	169	125	85
W12	6.628	485	467	361	343	325	307	169	125	85
W13	6.933	485	467	361	343	325	-	169	141	97
W14	7.587	485	467	361	343	325	-	169	125	85
W15	7.935	487	469	345	327	309	-	187	141	83
W16	8.284	487	469	345	327	-	-	169	141	101
W17	8.981	469	-	345	-	-	-	159	125	83

Table 3. Assignation of $[M+NH_4]^+$ precursor and ESI(+) MS/MS product ions of tentatively identified withanolides.

Peak Number	$[M+NH_4]^+$	$[M+H]^+$	$[M+H-H_2O]^+$	$[M+H-2H_2O]^+$	$[M+H-3H_2O]^+$	$[M+H-4H_2O]^+$	$[M+H-Lac]^+$	$[M+H-Lac-H_2O]^+$	$[M+H-Lac-2H_2O]^+$	$[M+H-Lac-3H_2O]^+$	Other MS/MS fragments
W1	536	519	501	483	465	447	317	299	281	263	185, 167, 139
W2	538	521	503	485	467	449	319	301	283	265	185, 167, 139
W3	538	521	503	485	467	449	317	299	281	263	187, 169, 143
W4	540	523	505	487	469	451	319	301	283	265	187, 169, 143
W6	520	503	485	467	449	431	301	283	265	247	185, 171, 155, 139
W7	522	505	487	469	451	433	301	283	265	247	169, 123, 69
W8	520	503	485	467	449	431	301	283	265	247	185, 167, 139, 123
W9	522	505	487	469	451	433	319	301	283	265	169, 125, 107
W10	520	503	485	467	449	431	317	299	281	263	169, 125
W11	522	505	487	469	451	433	319	301	283	265	169, 125
W12	504	487	469	451	433	415	301	283	265	247	169, 125
W13	504	487	469	451	433	415	301	283	265	247	185, 169, 139
W14	504	487	469	451	433	415	301	283	265	247	169, 125
W15	506	489	471	453	435	417	285	267	249	-	187, 169, 155, 123
W16	506	489	471	453	435	-	-	-	-	-	-
W17	488	471	453	435	417	399	285	267	249	-	169, 155, 125

Table 4. Assignation of $[M-H]^-$ precursor and ESI(-) MS/MS product ions of tentatively identified sucrose esters.

Peak Number	$[M-H]^-$	$[M-H-R_1]^-$	$[M-H-R_2]^-$	$[M-H-R_1-R_2]^-$	$[481-C_4H_8O_2]^-$	$[Sucr-H-H_2O]^-$	$[R_2O]^-$
S1	411	341	-	-	-	323	-
S2	481	411	-	-	393	323	-
S3	481	411	-	-	393	323	-
S4	495	411	-	-	-	323	-
S5	535	-	481	411	393	323	-
S6	551	481	-	411	393	-	-
S7	563	481	-	411	393	-	-
S8	565	481	-	411	393	-	-
S9	577	493	-	411	-	323	-
S10	565	495	411	-	-	323	171
S11	607	537	481	-	393	-	143
S13	621	551	481	411	-	-	157
S12	621	537	495	411	-	323	143
S14	635	551	481	411	-	323	171
S15	647	565	493	411	-	323	171
S16	649	565	495	411	-	323	171
S17	677	607	551	481	393	-	143
S18	691	621	551	481	393	-	157
S19	705	635	551	481	393	-	171
S20	719	635	551	481	393	-	171
S21	719		565	411	-	323	171
S22	747	663	565	481	393	-	199
S23	733	593	551	411	-	-	199

Table 5. Tentatively identified compounds from Physalus Peruviana calyces by GC-q-TOF-MS analysis.

Peak Number	Ret. Time (min)	Family	Tentative identification	Match factor	Formula	m/z $[M]^{\cdot}$ (measured)	Monoisotopic mass	Error (ppm)	Main fragments (m/z)
1	6.691	Monoterpene	D-Limonene	91	C10H16	136.1239	136.1252	5.5	136, 121, 107, 93
2	11.267	Phenolic	Phenol, 2-propyl-	77	C9H12O	n.d.	178.0477	-	136, 107, 77
3	11.700	Phenolic	Vanillin	91	C8H8O3	152.0459	152.0473	5.9	152, 123, 109, 81
4	11.998	Phenolic	Tyrosol	92	C8H10O2	138.0670	138.0681	3.8	138, 107, 77
5	12.144	Sesquiterpene	Sesquichamene	73	C15H24	n.d.	204.1878	-	189, 133, 121, 105, 91
6	13.179	Monoterpene	δ-Terpineol	77	C10H18O	n.d.	154.1358	-	136, 121, 93, 71
7	13.431	Sesquiterpene	Eudesmadienol	79	C15H24O	220.1822	220.1827	-0.1	159, 131, 105, 93
8	13.687	Sesquiterpene	α-Elemol	92	C15H26O	n.d.	222.1984	-	189, 161, 107, 93
9	13.884	Sesquiterpene	Maalialcohol	68	C15H26O	222.1972	222.1984	2.7	204, 189, 161, 109
10	14.193	Sesquiterpene	Germacratrienol isomer 1	75	C15H24O	220.1796	220.1827	11.6	187, 159, 109, 91
11	14.817	Sesquiterpene	δ-Cadinol	67	C15H26O	-	222.1984	-	204, 161, 119, 105
12	15.026	Sesquiterpene	Sesquiterpeneol isomer	96	C15H26O	-	222.1984	-	204, 189, 161, 149, 93
13	15.084	Ionone	3-Oxo-7,8-dihydro-α-ionone	84	C13H20O2	208.1464	208.1463	-2.9	208, 151, 135, 109
14	15.344	Sesquiterpene	Germacratrienol isomer 2	95	C15H24O	220.1816	220.1827	2.5	220, 159, 109, 91
15	15.432	Iononol	3-Oxo-7,8-dihydro-α-ionol	86	C13H22O2	210.1614	210.1620	0.1	210, 177, 135, 108
16	15.799	Phenolic	Coniferol	80	C10H12O3	180.0779	180.0786	1.0	180, 137, 124, 91
17	16.634	Sesquiterpene	Diepicedrene-1-oxide	87	C15H24O	220.1816	220.1827	2.5	177, 159, 109, 95
18	16.693	Sesquiterpene	Ambrial	84	C16H26O	234.197	234.1984	3.4	190, 137, 123, 95
19	16.822	Sesquiterpene	Cryptomeridiol	92	C15H28O2	n.d.	240.2089	-	164, 149, 123, 109

Table 5 (Contd.)

Peak Number	Ret. Time (min)	Family	Tentative identification	Match factor	Formula	m/z $[M]^{+\cdot}$ (measured)	Monoisotopic mass	Error (ppm)	Main fragments (m/z)
20	17.323	Sesquiterpene	Germacratrienol isomer 3	78	C15H24O	220.1822	220.1827	-0.1	220, 159, 107, 91
21	17.376	Sesquiterpene	Isoaromadendrene epoxide	90	C15H24O	220.1814	220.1827	3.4	220, 149, 119, 105
22	17.601	Diterpene	Sclareol oxide	92	C18H30O	262.2302	262.2297	-4.1	262, 191, 123, 109
23	17.752	Sesquiterpene	α-Copaeneol	56	C15H24O	n.d.	220.1827	-	163, 147, 119, 105
24	17.962	Diterpene	Dihydromanoyl oxide 1	85	C20H36O	n.d.	292.2766	-	263, 245, 177, 137
25	18.440	Sesquiterpene	Isoaromadendrene epoxide	82	C15H24O	n.d.	220.1827	-	135, 121, 107, 93
26	18.901	Diterpene	Epimanoyl oxide	92	C20H34O	n.d.	290.2610	-	275, 257, 177, 137
27	19.549	Diterpene	Phytol	95	C20H40O	n.d.	296.3079	-	123, 111, 95, 81
28	19.663	Diterpene	Dihydromanoyl oxide 2	70	C20H36O	n.d.	292.2766	-	263, 245, 191, 137
29	20.151	Sesquiterpene	Farnesol, acetate	90	C17H28O2	n.d.	264.2089	-	136, 121, 107, 93
30	20.242	Diterpene	trans-Geranylgeraniol	96	C20H34O	290.2596	290.2610	2.8	121, 107, 93, 81
31	20.511	Diterpene	Dihydromanoyl oxide 3	74	C20H36O	n.d.	292.2766	-	263, 245, 137, 95
32	20.548	Sesquiterpene	Khusiol	70	C15H26O	222.1971	222.1984	3.2	222, 177, 123, 107
33	20.589	Diterpene	Copalol isomer 1	70	C20H34O	290.2599	290.2610	1.7	177, 137, 109, 95
34	20.675	Diterpene	13-Epimanool	75	C20H34O	n.d.	290.2610	-	257, 137, 121, 95
35	20.861	Diterpene	Copalol isomer 2	90	C20H34O	290.2596	290.2610	2.8	275, 257, 137, 95
36	21.192	Triterpene	Friedelan-3-one	66	C30H50O	n.d.	426.3862	-	274, 177, 137, 97
37	21.405	Diterpene	Sclareol	88	C20H36O2	308.2708	308.2715	0.5	177, 137, 123, 109
38	21.799	Diterpene	Copalol isomer 3	71	C20H34O	290.2623	290.2610	-6.4	275, 177, 137, 123

39	21.963	Diterpene	Dihydromanoyl oxide 3	78	C20H36O	n.d.	292.2766	-	245, 137, 109, 95

Table 5 (Contd.)

Peak Number	Ret. Time (min)	Family	Tentative identification	Match factor	Formula	m/z $[M]^{\cdot}$ (measured)	Monoisotopic mass	Error (ppm)	Main fragments (m/z)
40	22.409	Diterpene	5-(7a-Isopropenyl-4,5-dimethyl-octahydroinden-4-yl)-3-methyl-pent-2-en-1-ol	94	C20H34O	290.2604	290.2610	0.1	192, 177, 135, 122
41	22.789	Diterpene	Dihydromanoyl oxide 4	60	C20H36O	n.d.	292.2766	-	263, 245, 137, 121
42	23.035	Diterpene	Dihydromanoyl oxide-7 carboxilic acid methyl ester	70	C21H36O3	n.d.	336.2664	-	307, 289, 245, 121
43	24.013	Diterpene	13,13-Dimethylpodocarp-7-en-3α-ol	56	C19H32O	n.d.	276.2453	-	243, 187, 135, 121
44	24.849	Steroid	16α-Methylpregnenolone	69	C22H34O2	n.d.	330.2559	-	330, 297, 245, 145, 105
45	26.220	Phytol derivative	δ -Tocopherol	86	C27H46O2	402.3497	402.3498	-1.2	402, 177, 137, 121
46	26.941	Phytol derivative	β -Tocopherol	90	C28H48O2	416.3657	416.3654	-1.9	416, 191, 151, 121
47	27.461	Phytol derivative	γ-Tocopherol	92	C29H50O2	430.3824	430.3811	-4.3	430, 205, 165, 121
48	28.275	Steroid	7δ-Ergostenol	71	C28H48O	400.3708	400.3705	-2.0	400, 255, 214, 105
49	28.456	Steroid	Methyl 3,7-bis(acetyloxy)cholestan-26-oate	57	C32H52O6	n.d.	532.3764	-	412, 255, 159, 105
50	28.929	Steroid	γ-Sitosterol	90	C29H50O	414.3868	414.3862	-2.8	414, 329, 213, 145
51	29.047	Steroid	(Z)-Stigmasta-5,24(28)-dien-3β-ol	93	C29H48O	412.3687	412.3705	3.0	314, 281, 299, 105
52	29.754	Phytol derivative	α-Tocopherol-β-D-mannoside	83	C35H60O7	n.d.	592.4339	-	430, 205, 165, 71
53	30.700	Steroid	4,4-Dimethyl-5-α-cholestane-3-one	62	C29H50O	n.d.	414.3862	-	414, 287, 123, 95