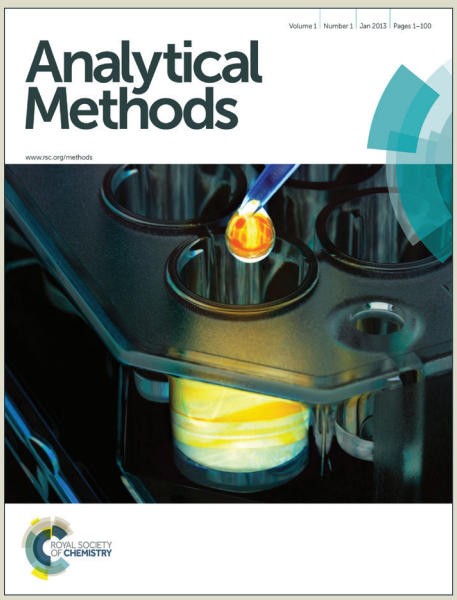
Analytical Methods

Accepted Manuscript



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/methods

Strategy for rapid structural characterization of saponins and flavonoids from the testa of *Camellia oleifera* Abel. seed by ultra-high-pressure liquid chromatography combined with electrospray ionization liner ion trap-Orbitrap mass spectrometry

Mengxue Feng[a,b], Zhiling Zhu[a,b], Limin Zuo[b], Long Chen[a], Qipeng Yuan[a], Guangzhi Shan[b,*] Shi-Zhong Luo[a,*]

[a] *Beijing Key Laboratory of Bioprocess, College of Life Science and Technology, Beijing University of Chemical Technology, Beijing10029, PR China*

[b] *Institute of Medicinal Biotechnology, Chinese Academy of Medical Science and Peking Union Medical College, Beijing 10050, PR China*

* **Corresponding authors:**

Shi-Zhong Luo, Beijing Key Laboratory of Bioprocess, College of Life Science and Technology, Beijing University of Chemical Technology, Beijing 100029, China Tel: +8610 64438015; Fax: +8610 64438015. E-mail address: luosz@mail.buct.edu.cn

Guangzhi Shan, Institute of Medicinal Biotechnology, Chinese Academy of Medical Science and Peking Union Medical College, Beijing 100050, China.

Abstract: Saponins and flavonoids are the main bioactive ingredients in *Camellia oleifera* Abel.. In this study, the fragmentation pathways of triterpene saponins in *Camellia oleifera* were investigated and confirmed by electrospray ionization liner ion trap-Orbitrap (LTQ-orbitrap), and a new strategy for rapid characterization of saponins by ultra-high-pressure liquid chromatography with LTQ-orbitrap mass spectrometry (UHPLC-LTQ-orbitrap-MS^n) was developed. Based on the summarized fragmentation rules, 36 triterpene saponins from testa of *C. oleifera* extracted by n-butanol were found，23 of which were completely and tentatively characterized. Besides, 35 sapoinins were identified as novel saponins. Additionally, 8 flavonol-*O*-glycosides were separated from *C. oleifera*, with 7 being tentatively characterized but one unambiguously identified, and one of the 7 flavonol-*O*-glycoside was first reported in genus *Camellia*. This study provided a systematic strategy in identification of saponins by UPLC-MS^n method in genus *Camellia* and summarized essential data for phytochemical studies of *C. oleifera*.

Keywords: Camellia oleifera; UHPLC-LTQ-orbitrap-MS^n; Flavonoids; Triterpenoid saponins; Orbitrap

1. Introduction

Camellia oleifera Abel, a traditional oil crop in south China, has been widely planted as an economic crop. The seeds of *C. oleifera* Abel. are the raw materials for the extraction of edible oil containing its unsaturated fatty acids contents and nutrients [1]. However, as its byproduct, the testa of seeds is commonly used as a detergent or an organic fertilizer with low economic value, resulting in the waste of many bioactive compounds，such as flavonoids and saponins. Flavonoids is a kind of important bioactive compounds with antioxidant activity [2, 3] and anticancer activity [4], which has been isolated and structurally identified in literatures [2-6] . Saponins are the major active ingredients in *C. oleifera* Abel, Which have various bioactivities, such as antimicrobial activity [7], antioxidant activity [8, 9] and cytotoxic activity [10]. However, few studies elucidated the structural of saponins from *C. oleifera*. As far as we are concerned, only 9 types of triterpenoid saponins [3, 9-11], which were the same (2 types) as or similar to the saponins from genus *Camellia* [12], have been isolated and structurally identified by nuclear magnetic resonance spectroscopy (NMR).

It is available to isolate and identify the high-content components by NMR with，but not appropriate for the minor compounds [13]. Multi-stage mass spectrometry (MS^n) is a valuable method in characterization of compounds [14-17], when combined with liquid chromatography, especially ultra-high-pressure Liquid chromatography (UHPLC), it is suitable for rapid structural analysis in plant extracts [18, 19] such as flavonoids [20, 21] and triterpene saponins [22, 23], due to its high sensitivity, short time and low consumption [24-26], Better performance can be obtained when the high resolution mass spectrumetric analyzer are (HR-MS) applied in the UHPLC-MS^n system by providing exact mass and possible

chemical formula[13, 27], which can overcome the difficultes in distinction of isobaric compounds[28]. Recently, the LTQ-orbitrap, an emerging high resolution mass spectrum detector which can provide high mass accuracy measurements for precursor and product ions rapidly and simultaneously, has been developed as an useful technology in detection and identification of plant extracts [29, 30].

The saponins characterized by LC-MS^2 in some genera have been reported, such as *Glycyrhiza* [23], *Symplocos* [26], *Momordica* [31] and *Albizia* [32]. Their typical fragment pathway is loss of monosaccharide residues and acyl groups, but different structures exhibit difference in fragment behaviors. Few literatures studied on genus *Camellia* [8, 33]. Scoparo have identified about 20 different saponins derived from tea leaves without describing the process of identification [33]. Hu deduced 9 saponins from *C. oleifera* in positive ion mode [8] and gave a brief explanation for the identification. However, the fragmentation behavior and pathway of triterpene saponins in genus *Camellia* has not been clearly and systematically studied yet.

In this paper, the fragmentation pathways of triterpene saponins in *C. oleifera* were systematically studied and confirmed by LTQ orbitrap. Based on the summarized fragmentation rules, we developed a new strategy for rapid characterization of saponins in n-butanol extract by UPLC-LTQ-orbitrap-MS^n. Also, flavonol-*O*-glycosides were separated and identified. This study contributed to the identification of saponins in genus *Camellia* and provided essential data for further phytochemical study of *C. oleifera*.

2. Materials and methods

2.1 Chemicals and materials

The testa of *C. oleifera* Abel seed was purchased from Realpark Company (Hunan, China) and collected from Hunan province, China. Camelliaside A was purchased from Biopurify Phytochemicals Ltd (Chengdu, China), and total saponins (purity > 96%) of *C. oleifera* were purchased from Solarbio Co. (Beijing, China).

Analytical grade formic acid, HPLC-grade methanol (MeOH) and acetonitrile (ACN) were purchased from Merck (Darmstadt, Germany). Analytical grade ethanol, ethyl acetate (EtOAc), n-butanol (BuOH) and trichloromethane ($CHCl_3$) were purchased from Beijing Chemical Works (Beijing,China). Water was purified by a Milli-Q system (Milford，MA，USA).

2.2 Sample preparation

The testa of *C. oleifera* seeds (5 g) were powdered and then refluxed at 55 ^{o}C with 70% ethanol (3×150 ml, 1 h) to get the crude extract. The extract was evaporated to dry, and then partitioned in $CHCl_3$, EtOAc, BuOH, and H_2O. The BuOH extract was filtered, evaporated and freeze dried, then dissolved with 50% MeOH to a concentration of about 2mg/ml. The solution was filtered through a 0.22 μm membrane, and a volume of 5 μL was injected for UPLC-MS^3analysis.

2.3 UHPLC-MS^3 Method

UHPLC-MS^3 analysis was carried out with a Thermo UltiMate 3000 UHPLC system and an LTQ Orbitrap XL mass detector (ThermoFisher, CA, USA). The UHPLC system consisted with a quaternary pump, a diode-array-detector, an autosampler and a column compartment. An Agilent ZORBAX SB-C18 column (2.1×100 mm, 1.8 μm; Agilent, Palo Alto, USA) was used at the flow rate of 0.2 ml/min. The column temperature was 35 ^{o}C,

and the mobile phase was water containing 0.1% formic acid (A) and the mixture of MeOH and ACN (50:50, v/v) containing 0.1% formic acid (B). A linear gradient elution program was as follows: 20% B (v/v) in 0-5 min, 20-80% B in 5 min-12 min, 80% B in 12-18 min, 100% B in 18.1-30 min, and 20% B in 30.1-40 min. The ESI source was operated both in positive and negative ion mode. High purity nitrogen (N_2) was used as sheath gas (30arb) and auxiliary gas (10arb). High purity helium (He) was used as collision gas. In positive ion mode, the other conditions were as follows: source temperature, 300 ^{o}C; Source voltage, 4.5 kv; Capillary voltage, 35 V; Tube lens, 110 V. For negative ion mode, source temperature was 350 ^{o}C and tube lens was 110 V, the other conditions were the same as in the positive mode. Full scan data acquisition (mass range: m/z 100-2000) and data dependant MS^3 were acquired. The resolution was 30,000 for full mass scan and 15,000 for data dependant MS scan. The CID collision energy was adjusted to 35% of the maximum and the isolation width of precursor ions was set at m/z 2.0.

3. Results and discussion

3.1 Extraction and Partition

To remove unnecessary compounds, liquid-liquid partition was used in separation based on the polarity. Each of the four extracts ($CHCl_3$, EtOAc, BuOH, and H_2O) was detected by UHPLC-MS, and the results showed the main bioactive compounds, flavonol-*O*-glycosides and triterpene saponins mostly existed in BuOH extract. Thus, BuOH extract were used for analysis. The base peak chromatograms of BuOH extract were shown in Fig. 1. However, saponins were hard to be separated by UHPLC because of their variability and similar structures [33], so extracted ion chromatogram (EIC) was applied to identify their structures.

3.2 Characterization of flavonol-*O*-glycosides

Totally, 8 flavonoids were identified in *C. oleifera* Abel., and all of them were flavonol-*O*-glycosides, including one found in this genus for the first time. The fragmentation pathway for flavonol-*O*-glycosides has been studied in detail [20], in which loss of monosaccharide residues are a typical way to obtain the aglycone structure and then the carbohydrate sequence. The structures of flavonol-*O*-glycosides in *C. oleifera* Abel. are usually characterized by the linkage of α-$_L$-rhamnopyranose, β-$_D$-glucopyranose, acetyl-α-$_L$-rhamnopyranose and xylopyranose to the flavonoid skeleton through the C-3 group.

Both positive and negative ion modes were collected to identify the flavonol-*O*-glycosides in *C. oleifera* Abel (Table 1). All 8 flavonol-*O*-glycosides gave an identical fragment ions in the MS^2 spectra at m/z 287.05637 ($C_{15}H_{11}O_6$; $[Aglycone+H]^+$) in positive ion mode and m/z 285.03961 ($C_{15}H_9O_6$; $[Aglycone-H]^-$) in negative ion mode (compound **5** showed a much higher relative abundance of $[Aglycone-H]^{-\bullet}$ ion at *m/z* 284 than that of $[Aglycone-H]^-$ ion [34], which was formed by the hemolytic cleavage probably due to the character of the sugar substitution[35]. Although it could not be observed in all flavonol-3-*O*-glycosides, the existence of remarkably higher relative abundance of $[Aglycone-H]^{-\bullet}$ ion was very pronounced for flavonol-3-*O*-glycosides[36]). Moreover, the MS^3 spectra of m/z 287 and m/z 285 were consistent (fig.2), which proved the same aglycone.

Compared with standard compound, as well as the literature data [33] of UV spectra data λmax at 266nm and 348nm and high resolution accurate molecular weight (ppm < 5), Compound **1** was unequivocally identified as camelliaside A. Since all flavonol-*O*-glycosides showed similar fragmentation pathway by loss of monosaccharide residues, camelliaside A was used as a standard. In positive ion mode, camelliaside A

displayed a $[M+H]^+$ ion at m/z 757.22092 ($C_{33}H_{41}O_{20}$) and MS^2 fragment at m/z 595.16770 ($C_{27}H_{31}O_{15}$; $[M+H-162]^+$), m/z 449.10884 ($C_{21}H_{21}O_{11}$; $[M+H-162-146]^+$) and m/z 287.05641 ($C_{15}H_{11}O_6$; $[M+H-162-146-162]^+$), corresponding to the cleavage of glucose (Glc), rhamnose (Rha) and glucose unit successively. Aglycone was further confirmed to be kaempferol by product ions of m/z 287.05641 ($C_{15}H_{11}O_6$), including neutral loss of CO (m/z 259, the product ion of kaempferol; *m/z* 231, the product ion of m/z 259), H_2O (*m/z* 269, the product ion of kaempferol; *m/z* 213, the product ion of m/z 231), retro-Diels-Alder fragmentation (*m/z* 153), some other rearrangements (*m/z* 259, *m/z* 258 and *m/z* 241), and the cleavage of the C ring from kaempferol protonated at C-3 position (*m/z* 165) which were consistent with former reports [37] (fig. **S3**C, shown in supporting materials). Moreover, the MS^3 spectra could be used to distinguish kaempferol and its isomers, 2'-hydroxygenistein and luteolin (fig. **S4**, shown in supporting materials). Though they showed the same *m/z* value at 287 in MS^2 spectra, the product ions in MS^3 spectra were different. The methoxy unit at C-3 position in the C ring for keampferol had more fragmentation pathways than others. The typical fragment ion $^{0,2}A^+$ ion at *m/z* 165, from protonation at C-3 position, could only be yielded by keampferol (Scheme S**1**, shown in supporting materials) [28, 38]. Thus, MS^n strategy was necessary in confirmation of aglycone.

In negative ion mode, the fragmentation pathway of aglycone at m/z 285.03961 was simaliar to that in positive ion mode (fig. **S3**D, shown in supporting materials), but MS^2 spectrum was different. Loss of disaccharide residues made it hard to deduce the sequence of glycosides. Meanwhile, flavonol-*O*-diglycosides, rutinose (L-rhamnopyranosyl-(1→6)-$_D$-glucopyranose) and neohesperidose (L-rhamnopyranosyl-(1→2)-$_D$-glucopyranose) could be distinguished by MS^2 spectrum in negative ion mode. Rutinose mainly caused aglycone fragment (for compound **7**), while

neohesperidose yielded more product ions at higher m/z values, corresponding to the [M-H-Rha]$^-$ ion, [M-H-Rha-H_2O]$^-$ ion, [M-H-Rha-120]$^-$ ion and [M-H-120]$^-$ ion[39]. Besides, for flavonoids with two hexoses, the 1→2 and 1→6 interglycosidic linkages could be distinguished by the [M-H-180]$^-$ ion[40]. Therefore, positive ion mode was generally used to identify the possible constituents of other flavonol-*O*-glycosides and negative ion mode was used to distinguish the linkage type.

The carbohydrate sequence could be easily determined by the loss of monosaccharide residues in positive ion modes: the loss of 162Da (C6H10O5), 146Da (C6H10O4), 132Da (C5H8O4) and 188Da (C8H12O5) corresponding to the cleavage of hexose, deoxyhexose, pentose and acetyl-deoxyhexose respectively. However, for most compounds, systematically structural identification of the sugar chain, including the linkage type and steric structure, was unavailable from MS data. Their full structures were tentatively deduced by literature (Table 1). Therefore, the hexose could be determined as glucose (for compounds **1, 2, 4, 5, 6, 7** and **8**) or galactose (Gal) (for compounds **2** and **3**) and the deoxyhexose was supposed to be rhamnose and the pentose was identified as xylose (Xyl) (Fig. 2). Additionally, the linkage type could be tentatively deduced. It was also used for the identification of saponins.

By investigating the reported compounds**,** compounds **2**-**8** were tentatively identified. All of them were reported in *C. oleifera*, except compounds **2** and **3**.

Compound **2** was comprised of a kaempferol unit and two hexose units. The high relative abundance of [M-H-180]$^-$ ion at *m/z* 429 indicated the (1→2) interglycosidic linkage[40]. Compared with literature data [41], the MS spectra data was the same as for camelliaside C (kaempferol-3-*O*-β-$_D$-galacopyranosyl-(1→2)-β-$_D$-glucopyranoside) in *C.*

sinesis. However, the stereoisomer, Kaempferol-3-*O*-sophoroside (kaempferol-3-*O*-β-$_D$-glucopyranosyl-(1→2)-β-$_D$-glucopyranoside), has also been reported in *C. oleifera* Abel[2]. Since the stereochemical structure of the hexose could not be identified only by MS^n spectra, compound **2** was tentatively identified as Kaempferol-3-*O*-sophoroside or camelliaside C.

Compound **3** was comprised of a kaempferol unit, a hexose unit and two methylpentose units, which has not been reported for this genus before. Based on the Chemical Abstracts database (https://scifinder.cas.org/) , Only one compound , kaempferol-3-*O*-α-$_L$-rhamnopyranosyl-(1→4)-[α-$_L$-rhamnopyranosyl-(1→6)-β-$_D$-galactopyranoside] [42], fits the MS spectra data. Thus, compound **3** was presumed to be this compound.

Of all the flavonol-*O*-glycosides, half of them (compounds **1**, **4**, **5** and **7**) have been reported with bioactivity. Camelliaside A and B have proved to have inhibition on the growth of hepatocellular carcinoma cells significantly [4]. Besides, compounds **5** and **7** have been reported to have moderate DPPH radical scavenging activity [2]. They showed potential use in medicine.

3.3 Characterization of triterpenoid saponins

To date, over 100 kinds of saponins in this genus, including acylated pentacyclic triterpenoid saponins, have been reported, and most of them are oleanane-type triterpenoid saponins. They have been characterized by glycosides at C-3 and one or more acyl groups at C-16, C-21, C-22 and C-28 in aglycones [9, 10, 12, 43-46]. Since there was no previous study, we observed the mass spectrometric behaviors of the saponins in C. *oleifera* and summarized their fragmentation rules systemically.

3.3.1 Fragmentation behavior of triterpenoid saponins

Compared with total saponin in fragmentation behavior and literature data [9], 36 compounds in *C. oleifera* Abel were tentatively identified as triterpenoid saponins. Their mass spectrometric behavior were summarized and the element composition of the fragment ions were confirmed by Orbitrap.

In positive ion mode, most saponin gave a $[M+NH_4]^+$ ion but sometimes a $[M+Na]^+$ ion as base peak in MS spectra. In MS^2 spectra, $[M+Na]^+$ ion yielded produce ions [Glycoside+Na]$^+$ at m/z 655.17120 ($C_{23}H_{36}O_{20}Na$; [Hexose+hexose+pentose+gluconic acid (GlcA)+Na]$^+$) or m/z 685.17975 ($C_{24}H_{38}O_{21}Na$; [Glc+Gal+Gal+GlcA+Na]$^+$) , but for $[M+NH_4]^+$ ion, the data was messy in the MS^2 spectrum.

In negative ion mode, most components had similar fragmentation behavior by losing acyl groups and sugar residues. Moreover, most saponins, except the desacyl-saponins, displayed a sugar residue ion in low abundance at m/z 451.10842 ($C_{17}H_{23}O_{14}$; [Pentose+hexose+GlcA-H]$^-$) or m/z 481.11891 ($C_{18}H_{25}O_{15}$; [Gal+Gal+GlcA-H]$^-$) by loss of a hexose or pentose in glycosides, while the desacyl-saponins showed [Aglycone-H]$^-$ ion instead of sugar residue ion due to lack of acyl groups. It could be used to distinguish normal saponins and desacyl-saponins, and the ion at m/z 451 or m/z 481 could be considered as a diagnose ion for normal saponins.

From low resolution mass spectra data, the sugar residue ion at m/z 451 ($C_{17}H_{23}O_{14}$) was hard to be distinguished from the [Aglycone-H]$^-$ ion ($C_{30}H_{43}O_3$; m/z 451) (the aglycone of Camelliasaponin B_1[47]), corresponding to two different fragmentation pathways, which were reported in triterpenoid saponins [26, 31]. However, with LC-LTQ-orbitrap, the ion at m/z 451.10842 could be confirmed ($C_{17}H_{23}O_{14}$; [Pentose+hexose+GlcA-H]$^-$). The MS^3

spectra could also help to support this point. For the saponins which glycosides consisted of 2 hexose, a pentose and a GlcA, the [M-H-hexose]$^-$ ion could yield this ion in MS^3 spectra, while the [M-H-pentose]$^-$ ion could not. Therefore, the ion at m/z 451 could not be the [Aglycone-H]$^-$ ion. In conclusion, UHPLC-LTQ-orbitrap-MS^3 was critical by providing the high accuracy mass measurements for the product ions.

3.3.2 Strategy for identification of triterpenoid saponins

Due to the various kinds of structure, the separation and identification of saponins were challenging. Thus, a strategy for rapid identification of triterpenoid saponins was proposed as shown in Fig. 3. The identification of saponins could be divided into three parts: acyl groups, glycosides, and aglycones. Firstly, for the acyl groups, whether the sugar residue ion at m/z 451 or m/z 481 existed or not could distinguish normal saponins and desacyl-saponins. For desacyl-saponins, the acyl group did not exist, and for normal saponins, it could be inferred from [M-H-NL]$^-$ ion by neutral loss (NL) of acyl group. The second part was glycoside. They could be deduced from the loss of monosaccharide residues successively, which were similar to flavonol-*O*-glycosides. The last part was aglycones. For desacyl-saponins, the formula for aglycones could be deduced from the [Aglycone-H]$^-$ ion, while for normal saponins, it could be calculated from the [Aglycone+acyl-H]$^-$ ion or [Aglycone+acyl+H_2O-H]$^-$ ion after inferred the acyl group. As a result, their aglycones could tentively identify based on the same formula as previous reports. This method is useful to identify the aglycones in saponins, such as in Glycyrrhiza *yunnanensis* [23].

The fragmentation pathway of compound **29** was shown in Fig. 4**,** as an example for the identification of triterpenoid saponins. The MS^2 spectra gave a [M-H-102]$^-$ ion at m/z

1101.50884 ($C_{53}H_{81}O_{24}$), indicating a MBOH (methylbutyryl-OH) unit or IsovOH (isovaleryl-OH) unit as an acyl group. The sequence constitution of sugar chain could also be deduced as S_1 (shown in Fig. 5) by the $[M-H-C_5H_8O_4]^-$ ion at m/z 1071.53418, $[M-H-C_6H_{10}O_5]^-$ ion at m/z 1023.51263, $[M-H-C_5H_8O_4-C_6H_{12}O_6]^-$ ion at m/z 891.47162 and $[M-H-C_5H_8O_4-C_6H_{12}O_6 -C_6H_{12}O_6-CO_2]^-$ ion at m/z 667.41888, corresponding to the cleavages of a pentose unit, two hexose unit and CO_2. The molecular weight of aglycones could be obtained by $[Aglycone+MBOH-H]^-$ ion at m/z 553.38806 ($C_{35}H_{53}O_5$), and the result ($C_{30}H_{44}O_3$; 452Da) indicated the formula of its aglycone was the same as Camelliasaponin B_1 [48]. Thus, its structure was tentatively inferred as shown in Fig. 5.

Following the above procedure, 23 saponins were tentatively deduced, while 13 saponins could not be completely inferred. Based on the formula of aglycones, the triterpenoid saponins in *C. oleifera* could be classified into 9 different types (Fig **5**). Type I: Compounds **17, 24, 25, 29, 30, 35, 36** and **37**, showed the same formula in aglycones as Camelliasaponin B_1 ($C_{30}H_{44}O_3$; 452Da) [48]. Type II: Compounds **18, 19**, **31, 32** and **38**, gave the $[aglycone+acyl+H_2O-H]^-$ ion instead of $[aglycone+acyl-H]^-$ ion, and their aglycones resembled those of Camelliasaponin A_1 ($C_{30}H_{46}O_2$; 438Da) [49]. Type III: Compounds **9** and **10**, which are desacyl-saponins, had the same aglycones as desacyl-assamsaponin F [50] ($C_{30}H_{46}O_5$; 486Da). Type IV: Compounds **13** and **14** (desacyl-saponins)**,** corresponded to desacyl-camelliasaponin B [48] ($C_{30}H_{46}O_4$; 470Da). Type V: Compounds **15** and **16** ($C_{30}H_{48}O_3$; 456 Da), whose behaviors were similar to type II, received a $[aglycone+H_2O-H]^-$ ion instead of $[aglycone-H]^-$ ion, and their aglycones were tentatively identified the same as desacyl-Camelliasaponin A [49]. Type VI: Compound **11** ($C_{30}H_{44}O_4$; 468Da), just like the aglycone of Assamsaponin F [50]. Type VII: Compound **12**, was the isomer of Theasaponin A4 [51] (different in carbohydrate sequence) ($C_{30}H_{46}O_4$; 470Da),

and can be distinguished from Type IV by the different fragmentation behaviors of normal saponins and desacyl-saponins. Type VIII: Compounds **20** and **21**, the desacyl-saponins, had the same aglycone as Camallidin II [52] ($C_{30}H_{46}O_3$; 454Da). Type IX: Other 13 compounds that could not be calculated in their aglycones for the cleavage of acyl groups cannot be observed in those compounds, which was probably because of their special structure.

Totally, 36 saponins were separated and characterized in the BuOH extract, 23 of which could completely inferred but 13 of them could not. All the saponins, except compound **37**, were identified as novel saponin. About half of them are the isomer of reported saponins, for they are different in structures. In addition, a new kind of glycosides (S_1) containing many compounds was found. S_1 consisted of the same monosaccharides with a known glycoside (S_4), but the sequence of sugar chain (S_1) was different. In MS^2 spectra, the compounds which had this kind of glycoside, such as compound **29,** showed high abundance of $[M\text{-}H\text{-}C_5H_8O_4]^-$ ions, which could not be produced by S_4. Such fragment behavior could also be observed in TOF-MS [8, 53]. Though this kind of glycoside (S_1) was not reported in genus *Camellia*, it was reported in other triterpenoid saponins, such as in *Antonia ovate* [54].

4. Conclusion

In conclusion, the main bioactive compounds in the BuOH extract from the testa of *C.oleifera* Abel. seeds were first characterized by UHPLC-LTQ-orbitrap-MS^n. To the best of our knowledge, it was the first time to summarize the available fragmentation pathways of triterpenoid saponins and developed a strategy for characterizing the saponins in *C.oleifera*. In total, one flavonol-*O*-glycoside was unambiguously identified and 7

flavonol-*O*-glycosides and 23 triterpene saponins were tentatively characterized, as well as 13 saponins could not be completely inferred. In all these 44 compounds, 35 saponins were characterized as novel compounds and one flavonol-*O*-glycoside was first reported in genus *Camellia*.

The UHPLC-LTQ-orbitrap-MS^n showed advantages in identification of flavonoids and saponins. The high accuracy mass measurements provided by orbitrap could help to propose the elemental composition for ions, and make it possible to distinguish the sugar residue ion and the $[Aglycone-H]^-$ ion, while the MS^n spectra could be used to confirm the aglycones of flavonoid and find the possible fragmentation pathways of saponins with HR-MS^n.

This work was helpful for the identification of other triterpenoid saponins in this genus by LC-MS^n for their similar structure, and provided essential data for further photochemical studies of C. oleifera. However, the UHPLC-LTQ-orbitrap-MS^n could be hardly used in distinguishing the linkage positions and stereoisomers. NMR data was necessary for completely structural characterization.

Funding

This work was supported by the National Basic Res earch Program of China (973 program) (2013CB910700), the State Key Laboratory of NBC Protection for Civilian. No. SKLNBC2013-02K, the National Natural Science Foundation of China (21372026), Beijing NOVA Programme (Z131102000413010) and the Fundamental Research Funds for the Central Universities (YS1407).

References:

1. H. Zhong, D. R. Bedgood Jr., A. G. Bishop, P. D. Prenzler and K. Robards, *Food Chem*, 2007, **100**, 1544-1551.
2. D. Gao, M. Xu, P. Zhao, X. Zhang, Y. Wang, C. Yang and Y. Zhang, *Food Chem*, 2011, **124**, 432-436.
3. J. Chen, H. Wu, B. Liau, C. J. Chang, T. Jong and L. Wu, *Food Chem.*, 2010, **121**, 1246-1254.
4. L. C. Du, B. L. Wu and J. M. Chen, *Chinese Chem. Lett.*, 2008, **19**, 1315-1318.
5. J. Chen, B. Liau, T. Jong and C. J. Chang, *Se.p Purif. Technol.*, 2009, **67**, 31-37.
6. Y. Ye, Y. Guo and Y. Luo, *Int. J. mol. sci.*, 2012, **13**, 12401-12411.
7. J. Hu, S. Nie, D. Huang, C. Li, M. Xie and Y. Wan, *J. Sci. Food Agr.*, 2012, **92**, 2443-2449.
8. J. Hu, S. Nie, D. Huang, C. Li and M. Xie, *Int. J. Food Sci. Tech*, 2012, **47**, 1676-1687.
9. X. Zhang, Y. Han, G. Bao, T. Ling, L. Zhang, L. Gao and T. Xia, *Molecules*, 2012, **17**, 11721-11728.
10. H. Zhou, C. Z. Wang, J. Z. Ye and H. X. Chen, *Phytochem. Lett.*, 2014, **8**, 46-51.
11. S. Sugimoto, G. Chi, Y. Kato, S. Nakamura, H. Matsuda and M. Yoshikawa, *Chem. Pharm. Bull.*, 2009, **57**, 269-275.
12. P. Zhao, D. Gao, M. Xu, Z. Shi, D. Wang, C. Yang and Y. Zhang, *Chem. Biodivers.*, 2011, **8**, 1931-1942.
13. S. Wang, L. Chen, J. Leng, P. Chen, X. Fan and Y. Cheng, *J. Pharmaceut. Biomed.*, 2014, **98**, 22-35.
14. Y. Gao and S. A. McLuckey, *Rapid Commun. mass Sp.*., 2013, **27**, 249-257.
15. Y. Gao and S. A. McLuckey, *J. Mass Spectrom.*, 2012, **47**, 364-369.
16. Y. Gao, J. Yang, M. T. Cancilla, F. Meng and S. A. McLuckey, *Anal. Chem.*, 2013, **85**, 4713-4720.
17. I. K. Webb, Y. Gao, F. A. Londry and S. A. McLuckey, *J. Mass Spectrom*, 2013, **48**, 1059-1065.
18. G. Zheng, P. Zhou, H. Yang, Y. Li, P. Li and E. H. Liu, *Food Chem*, 2013, **136**, 604-611.
19. X. Chen, L. Hu, X. Su, L. Kong, M. Ye and H. Zou, *J Pharmaceut Biomed*, 2006, **40**, 559-570.
20. F. Cuyckens and M. Claeys, *J Mass Spectrom*, 2004, **39**, 1-15.
21. G. B. Gonzales, K. Raes, S. Coelus, K. Struijs, G. Smagghe and J. Van Camp, *J Chromatogr A*, 2014, **1323**, 39-48.
22. L. Qi, H. Wang, H. Zhang, C. Wang, P. Li and C. Yuan, *J Chromatogr A*, 2012, **1230**, 93-99.
23. S. Ji, Q. Wang, X. Qiao, H. Guo, Y. Yang, T. Bo, C. Xiang, D. Guo and M. Ye, *J Pharmaceut Biomed*, 2014, **90**, 15-26.
24. M. Cui, W. Sun, F. Song, Z. Liu and S. Liu, *Rapid Commun Mass Sp*, 1999, **13**, 873-879.
25. B. Li, Z. Abliz, G. Fu, M. Tang and S. Yu, *Rapid Commun Mass Sp*, 2005, **19**, 381-390.
26. B. Li, Z. Abliz, M. Tang, G. Fu and S. Yu, *J Chromatogr A*, 2006, **1101**, 53-62.
27. S. Wang, P. Chen, Y. Xu, X. Li and X. Fan, *J. Sep. Sci.*, 2014, **37**, 1748-1761.
28. M. Stobiecki, P. Kachlicki, A. Wojakowska and Ł. Marczak, *Phytochem. Lett.*, 2015, **11**, 358-367.

29. X. Qiu, J. Zhang, Z. Huang, D. Zhu and W. Xu, *J. Chromatogr. A*, 2013, **1292**, 121-131.
30. W. Songsong, X. Haiyu, M. Yan, W. Xuguang, S. Yang, H. Bin, T. Shihuan, Z. Yi, L. Defeng, L. Rixin and Y. Hongjun, *J. Pharmaceut. Biomed.*, 2015, **111**, 104-118.
31. Z. Lin, X. Liu, F. Yang and Y. Yu, *Int. J. Mass Spectrom.*, 2012, **328-329**, 43-66.
32. L. Han, G. Pan, Y. Wang, X. Song, X. Gao, B. Ma and L. Kang, *J. Pharmaceut. Biomed*., 2011, **55**, 996-1009.
33. C. T. Scoparo, L. M. de Souza, N. Dartora, G. L. Sassaki, P. A. J. Gorin and M. Iacomini, *J. Chromatogr. A*, 2012, **1222**, 29-37.
34. F. Ferreres, D. Gomes, P. Valentão, R. Gonçalves, R. Pio, E. A. Chagas, R. M. Seabra and P. B. Andrade, *Food Chem.*, 2009, **114**, 1019-1027.
35. E. Hvattum and D. Ekeberg, *J Mass Spectrom.*, 2003, **38**, 43-49.
36. F. Cuyckens and M. Claeys, *J Mass Spectrom.*, 2005, **40**, 364-372.
37. R. E. March and X. Miao, *Int J Mass Spectrom.*, 2004, **231**, 157-167.
38. Y. L. Ma, Q. M. Li, H. Van den Heuvel and M. Claeys, *Rapid Commun. Mass Sp.*, 1997, **11**, 1357-1364.
39. F. Cuyckens, R. Rozenberg, E. de Hoffmann and M. Claeys, *J. Mass Spectrom.*, 2001, **36**, 1203-1210.
40. F. Ferreres, R. Llorach and A. Gil-Izquierdo, *J. Mass Spectrom.*, 2004, **39**, 312-321.
41. T. Sekine, Y. Arai, F. Ikegami, Y. Fujii, S. Shindo, T. Yanagisawa, Y. Ishida, S. Okonogi and I. Murakoshi, *Chem. Pharm. Bull.*, 1993, **41**, 1185-1187.
42. Y. Murai, R. Takahashi, F. R. Rodas, J. Kitajima and T. Iwashina, *Nat. Prod. Commun.*, 2013, **8**, 453-456.
43. H. Matsuda, M. Hamao, S. Nakamura, H. Kon'I, M. Murata and M. Yoshikawa, *Chem. Pharm. Bull.*, 2012, **60**, 674-680.
44. K. Fujimoto, S. Nakamura, S. Nakashima, T. Matsumoto, K. Uno, T. Ohta, T. Miura, H. Matsuda and M. Yoshikawa, *Chem. Pharm. Bull.*, 2012, **60**, 1188-1194.
45. N. Li, Z. Ma, Y. Chu, Y. Wang and X. Li, *Fitoterapia*, 2013, **84**, 321-325.
46. M. Myose, T. Warashina and T. Miyase, *Chem. Pharm. Bull.*, 2012, **60**, 612-623.
47. M. Yoshikawa, E. Harada, T. Murakami, H. Matsuda, J. Yamahara and N. Murakami, *Chem. Pharm. Bull.*, 1994, **42**, 742-744.
48. M. Yoshikawa, T. Murakami, S. Yoshizumi, N. Murakami, J. Yamahara and H. Matsuda, *Chem. Pharm. Bull.*, 1996, **44**, 1899-1907.
49. M. Yoshikawa, T. Murakami, S. Yoshizumi, N. Murakami, J. Yamahara and H. Matsuda, *Chem. Pharm. Bull.*, 1996, **44**, 1899-1907.
50. T. Murakami, J. Nakamura, T. Kageura, H. Matsuda and M. Yoshikawa, *Chem. Pharm. Bull.*, 2000, **48**, 1720-1725.
51. M. Yoshikawa, T. Morikawa, S. Nakamura, N. Li, X. Li and H. Matsuda, *Chem. Pharm. Bull.*, 2007, **55**, 57.
52. C. Nishino, S. Manabe, N. Enoki, T. Nagata, T. Tsushida and E. Hamaya, *J. Chem. Soc., Chem. Commun.*, 1986, 720-723.
53. X. Zhang, X. Zhu, Y. Liu, T. Xia, X. Du, L. Gao and D. Zhang, *Shipin Kexue (Beijing, China)*,

2012, **33**, 7-11.

54. A. A. Magid, H. Bobichon, N. Borie, N. Lalun, C. Long, C. Moretti and C. Lavaud, *Phytochemistry*, 2010, **71**, 429-434.

55. B. Li and Y. Luo, *Chemical Journal on Internet*, 2003, **5**, No pp. given.

56. C. Ma, K. Dastmalchi, G. Flores, S. Wu, P. Pedraza-Peñalosa, C. Long and E. J. Kennelly, *J. Agr. Food Chem.*, 2013, **61**, 3548-3559.

57. M. P. Rodríguez-Rivera, E. Lugo-Cervantes, P. Winterhalter and G. Jerz, *Food Chem.*, 2014, **158**, 139-152.

58. H. Fu, C. Li, J. Yang, X. Chen and D. Zhang, *Phytochemistry*, 2013, **85**, 167-174.

59. T. Nagata, T. Tsushida, E. Hamaya, N. Enoki, S. Manabe and C. Nishino, *Agr. Bio. Chem.*, 1985, **49**, 1181-1186.

60. H. Matsuda, S. Nakamura, K. Fujimoto, R. Moriuchi, Y. Kimura, N. Ikoma, Y. Hata, O. Muraoka and M. Yoshikawa, *Chem. Pharm. Bull.*, 2010, **58**, 1617-1621.

Figure Captions

Fig.1 Base peak chromatograms of the following fractions: (A)BuOH extract in negative ion mode(T70Bu); (B) Partial enlarged drawing of BuOH extract in negative ion mode.

Fig. 2 The procedure to identify flavonol-O-glycosides. Glu: glucose; Gal: galactose; Rha: rhamnose; Xyl: xylose; Ac: acetyl.

Fig. 3 The procedure to identify saponins. NL: neutral loss; MB: methylbutyryl; MBOH: methylbutyryl-OH; Ang: angeloyl; Tig: tigeloyl; Ac: acetyl; Hex: hexenoyl; Isov: isovaleryl.

Fig. 4 Proposed fragmentation pathway for compound **29**.

Fig. 5 Probable structure of triterpenoid saponins identified in T70Bu. S_1, S_2, S_3 and S_4 are the sugar residues in saponins.

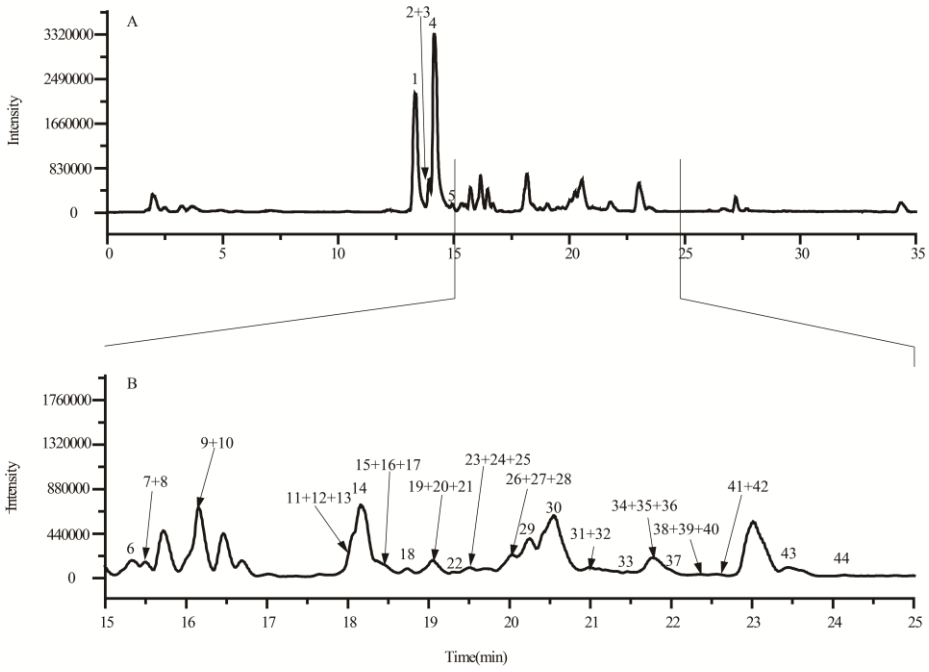


Fig.1 Base peak chromatograms of the following fractions: (A)BuOH extract in negative ion mode(T70Bu); (B) Partial enlarged drawing of BuOH extract in negative ion mode.

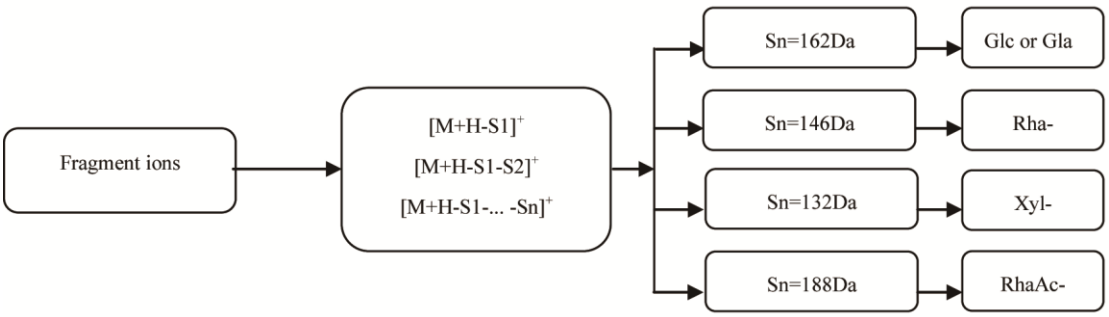


Fig. 2 The procedure to identify flavonol-O-glycosides. Glu: glucose; Gal: galactose; Rha: rhamnose; Xyl: xylose; Ac: acetyl.

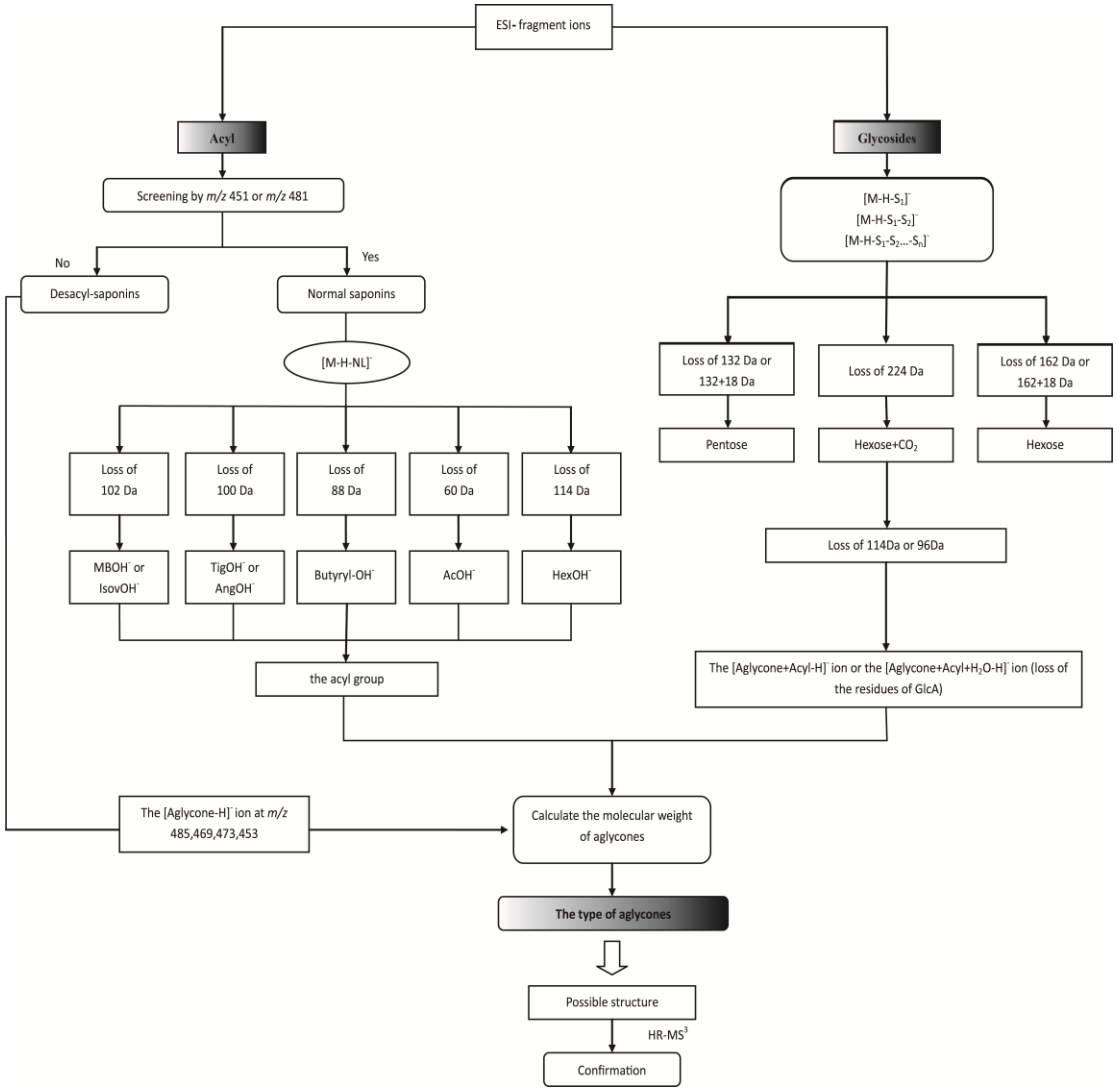


Fig. 3 The procedure to identify saponins. NL: neutral loss; MB: methylbutyryl; MBOH: methylbutyryl-OH; Ang: angeloyl; Tig: tigeloyl; Ac: acetyl; Hex: hexenoyl; Isov: isovaleryl.

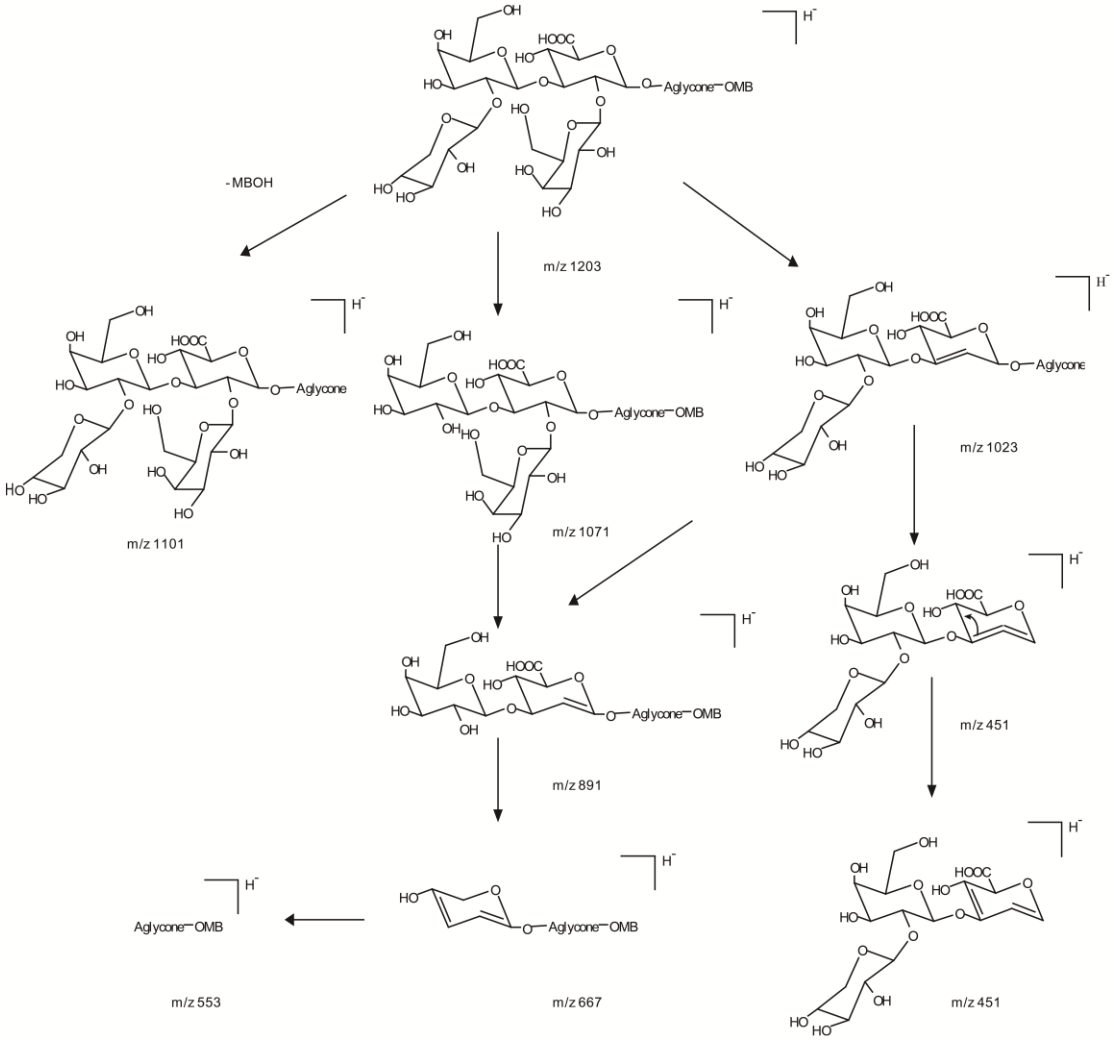


Fig. 4 Proposed fragmentation pathway for compound **29**.

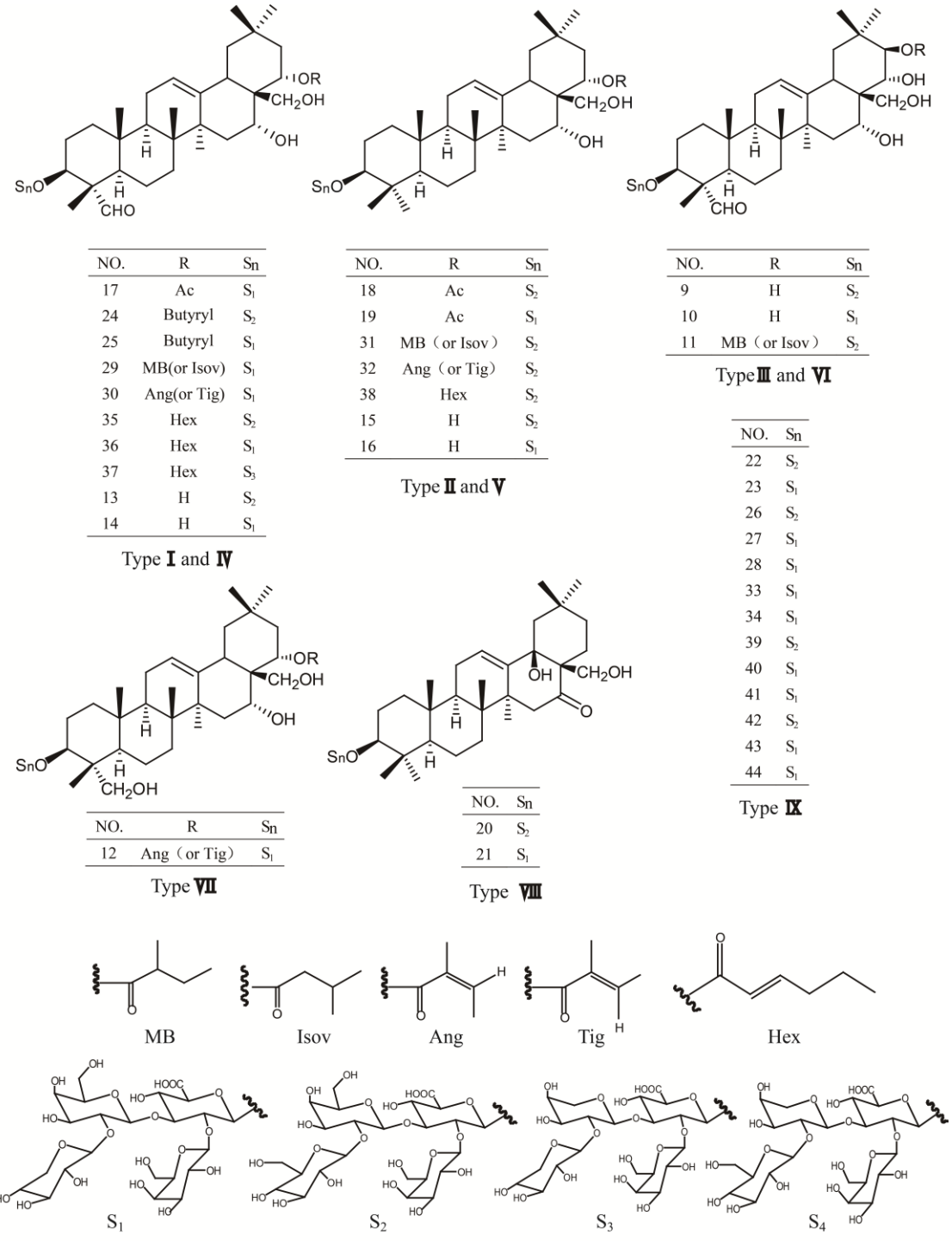


Fig. 5 Probable structure of triterpenoid saponins identified in BuOH extract. S_1, S_2, S_3 and S_4 are the sugar residues in saponins.

Table 1 Characterization of flavonol-O-glycosides.

No.	tR(+/-)	Identification	Formula	ESI-[M-H]-			ESI+[M+H]+			reference
				Mean measured mass	Major MS2 ions (m/z)%	Error (ppm)	Mean measured mass	Major MS2 ions (m/z)%	Error (ppm)	
1	13.44/ 13.31	camelliaside A[a]	$C_{33}H_{40}O_{20}$	755.20292	593 (27),575 (100),285 (39)	1.36	757.21938	287 (100),449 (29),595 (28)	1.07	4, 33
2	13.98/ 13.88	kaempferol-3-O-sophoroside or camelliaside C	$C_{27}H_{30}O_{16}$	609.14656	447 (13),429 (100),285 (99)	2.55	611.16066	287(100),449(39)	0.53	2, 41
3	14.03/ 13.94	Kaempferol-3-*O*-α-L-rhamnopyranosyl-(1→4)-[α-L-rhamnopyranosyl-(1→6)-β-D-galactopyranoside].[b]	$C_{33}H_{40}O_{19}$	739.20959	593 (23),575 (100),285 (27),284 (39)	2.15	741.22437	287(100),449(10),595(32)	0.97	42
4	14.28/ 14.15	camelliaside B	$C_{32}H_{38}O_{19}$	725.19236	593(28),575(100),285(39)	1.22	727.20801	287(100),449(20),595(25)	0.90	4, 55
5	14.73/ 14.63	leucoside	$C_{26}H_{29}O_{15}$	579.13445	447(28),429(86),285(88),284(100)	2.21	581.15079	287(100),449(29)	1.19	2, 34
6	15.03/ 14.93	kaempferol-3-O-[4''''-O-acetyl-α-L-rhamnopyranosyl-(1→6)]-[β-D-glucopyranosyl-(1→2)]-β-D-glucopyranoside	$C_{35}H_{42}O_{21}$	797.21477	755(100),617(31),285(16)	1.61	799.22979	287(100),449(12), 637(15)	0.82	2
7	15.42/ 15.32	kaempferol-3-O-rutinoside	$C_{27}H_{31}O_{15}$	593.15131	285(100)	2.04	595.16629	287(100),449(19)	0.92	2, 56, 57
8	15.59/ 15.50	kaempferol-3-O-[4''''-O-acetyl-α-L-rhamnopyranosyl-(1→6)]-[β-D-xylopyranosyl-(1→2)]-β-D-glucopyranoside	$C_{34}H_{41}O_{20}$	767.20411	725(100),617(20),285(12)	1.56	769.21937	287(100),449(22),581(25),637(16)	1.04	2

a Compared with standard compounds

b Found in this genus for the first time.

Table 2 Characterization of triterpenoid saponins.

No.	tR(-)	Identification	Formula	ESI-			reference
				Mean measured mass[M-H]-	Major MS2 ions (m/z)%	Error（ppm)	
9	16.03	novel saponin	$C_{54}H_{86}O_{27}$	1165.52825	1103 (46), 1003 (46), 985 (100), 823 (16), 599 (12), 485 (12)	0.84	
10	16.14	Isomer of desacyl-assamsaponin F	$C_{53}H_{84}O_{26}$	1135.51769	1003 (46), 955 (100), 823 (22), 599 (12), 485 (11)	0.87	50
11	17.94	novel saponin	$C_{59}H_{94}O_{28}$	1249.58608	1147 (23), 1087 (50), 1069 (100), 907 (19), 683 (8), 569 (8), 481 (2)	1.03	
12	18.03	Isomer of Theasaponin A4	$C_{58}H_{92}O_{27}$	1219.57522	1119 (11), 1087 (16), 1039 (100), 907 (21), 683 (8), 587 (7) ,569 (4), 451 (2)	0.81	51
13	18.07	novel saponin	$C_{54}H_{86}O_{26}$	1149.53381	987 (46), 969 (100), 807 (18), 583 (11), 469 (15)	1.26	
14	18.15	Isomer of desacyl-camelliasaponin B	$C_{53}H_{84}O_{25}$	1119.52192	987 (46), 939 (100), 807 (23), 583 (11), 469 (14)	0.11	47
15	18.35	novel saponin	$C_{54}H_{88}O_{25}$	1135.55450	973 (52), 955 (100), 793 (14), 569 (8), 473 (6)	1.24	
16	18.43	Isomer of desacyl- Gordonsaponin F	$C_{53}H_{86}O_{24}$	1105.54358	973 (53), 925 (100), 793 (15), 569 (7), 473 (6)	0.95	58
17	18.45	novel saponin	$C_{55}H_{86}O_{26}$	1161.53236	1101 (33), 1029 (41), 981 (100), 849 (17), 625 (8), 511 (7), 451 (1)	0.50	
18	18.73	novel saponin	$C_{56}H_{90}O_{26}$	1177.56398	1117 (30), 1015 (52), 997 (100), 835 (15), 611 (7), 515 (4), 481 (4)	0.27	
19	18.83	Isomer of Gordonsaponin F	C55H88O25	1147.55348	1087 (20), 1015 (33), 967 (100), 835 (22), 611 (8), 515 (5), 451 (4)	0.27	58
20	18.95	novel saponin	C54H86O25	1133.53761	953 (100), 791 (14), 567 (10), 453 (13)	0.14	

Table 2 Continued

No.	tR(-)	Identification	Formula	ESI-			reference
				Mean measured mass[M-H]-	Major MS^2 ions (m/z)%	Error (ppm)	
21	19.03	isomer of Camallidin Ⅱ	$C_{53}H_{84}O_{24}$	1103.52722	971 (50), 923 (100), 791(19),567(10),453 (14)	0.31	52, 59
22	19.23	novel saponin	$C_{60}H_{96}O_{29}$	1279.59535	1243 (5), 1191 (4), 1131 (3), 1117 (50), 1099 (100), 937 (15), 713 (4), 585 (19), 481(4)	1.60	
23	19.36	novel saponin	$C_{59}H_{94}O_{28}$	1249.58338	1117 (52), 1069 (100), 937 (18), 713 (8), 617 (12), 451 (4)	1.12	
24	19.48	Isomer of Theasaponin A4	$C_{58}H_{92}O_{27}$	1219.57397	1131 (45), 1057 (45), 1039 (100), 877 (17), 653 (7), 539 (5), 481 (6)	0.20	51
25	19.50	Isomer of Theasaponin A1	$C_{57}H_{90}O_{26}$	1189.56396	1101 (30), 1057 (47), 1039 (6), 1009 (100), 877 (20), 653 (10), 539 (7), 451 (2)	0.26	26
26	19.95	novel saponin	$C_{61}H_{96}O_{29}$	1291.59361	1129 (42), 1111 (100), 949 (16), 725 (7), 597 (9), 481 (3)	1.35	
27	20.04	Isomer of TheasaponinA5	$C_{60}H_{94}O_{28}$	1261.58366	1129 (42), 1081 (100), 949 (20), 725 (9), 629 (6), 597(9), 451 (2)	1.13	51
28	20.07	novel saponin	$C_{60}H_{96}O_{28}$	1263.59705	1161 (6), 1131 (46), 1083 (100), 951 (23), 727 (5), 631 (11), 599 (18), 451 (2)	2.68	
29	20.48	Isomer of CamelliasaponinC1	$C_{58}H_{92}O_{26}$	1203.57786	1101 (8), 1071 (43), 1023 (100), 891 (21), 667 (8), 553 (6), 451(5)	1.20	48
30	20.56	Isomer of CamelliasaponinB1	$C_{58}H_{90}O_{26}$	1201.56368	1101 (18), 1069 (49), 1021 (100), 889 (19), 665 (11), 551 (8), 451 (3)	0.02	48
31	20.92	novel saponin	$C_{59}H_{96}O_{26}$	1219.60913	1117 (17), 1057 (59), 1039 (100), 895 (12), 877 (11), 653 (2), 557 (5), 481 (8)	1.21	
32	20.98	Isomer of Sasanquasaponins II	$C_{59}H_{94}O_{26}$	1217.59530	1117 (13), 1037 (100), 875 (13), 651 (7), 555 (4), 481 (3)	0.28	60

Table 2 Continued

No.	tR(-)	Identification	Formula	ESI-			reference
				Mean measured mass[M-H]-	Major MS^2 ions (m/z)%	Error (ppm)	
33	21.44	novel saponin	$C_{55}H_{94}O_{31}$	1249.56348	1135 (1), 1117 (39), 1069 (100), 937 (19), 713 (9), 599(3), 451 (3)	4.85	
34	21.51	novel saponin	$C_{59}H_{92}O_{28}$	1247.56934	1115 (44), 1067 (100), 935 (16), 711 (9), 597 (6), 451 (3)	0.17	
35	21.65	novel saponin	$C_{60}H_{94}O_{27}$	1245.59045	1131 (19), 1083 (46), 1065 (100), 921 (17), 903 (20), 679 (14), 565 (16), 481 (5)	0.47	
36	21.77	Isomer of Oleifera A	$C_{59}H_{92}O_{26}$	1215.57931	1101 (20), 1083 (48), 1035 (100), 903 (19), 679 (13), 565(9), 451 (3)	1.04	9
37	22.12	TeaseedsaponinG or its isomer	$C_{58}H_{90}O_{25}$	1185.56874	1071 (29), 1053 (56), 1005 (100), 903 (12), 679 (12), 565 (6), 451 (2)	0.80	46
38	22.27	Isomer of sasanquasaponin I	$C_{60}H_{96}O_{26}$	1231.57422	1117 (9), 1069 (39), 1051 (100), 907 (15), 889 (13), 665 (6), 569 (4), 481 (6)	0.88	60
39	22.32	Isomer of Teaseedsaponin A	$C_{62}H_{96}O_{28}$	1287.60099	1201 (2), 1155 (58), 1107 (100), 975 (19), 751 (16), 637 (12), 451 (6)	0.43	46
40	22.35	novel saponin	$C_{62}H_{98}O_{28}$	1289.61523	1203 (3), 1157 (50), 1109 (100), 977 (25), 753 (38), 639 (4), 451 (2)	0.66	
41	22.55	novel saponin	$C_{63}H_{96}O_{28}$	1299.60114	1167 (53), 1119 (100), 987 (22), 763 (13), 649 (10), 451 (2)	0.54	
42	22.79	novel saponin	$C_{64}H_{98}O_{29}$	1329.61200	1167 (57), 1149 (100), 1005 (20), 987 (10), 763 (10), 649 (7), 481 (5)	0.75	
43	23.45	Isomer of TeaseedsaponinD	$C_{63}H_{98}O_{28}$	1301.61589	1169 (51), 1121 (100), 989 (22), 765 (12), 651 (9), 451 (2)	0.15	46
44	24.14	novel saponin	$C_{67}H_{96}O_{29}$	1347.60107	1215 (43) ,1185 (12), 1167 (100), 1035 (23), 811 (9), 697 (5), 451 (1)	0.47	