Analyst

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/analyst

Characterization of the herb-derived components in rats following oral administration of *Carthamus tinctorius* extract by extracting diagnostic fragment ions (DFIs) in the MS^n chromatograms

Jin-Feng Chen,[a#] Yue-Lin Song,[b#] Xiao-Yu Guo,[a] Peng-Fei Tu,[ab] and Yong Jiang*[a]

[a] *State Key Laboratory of Natural and Biomimetic Drugs, School of Pharmaceutical Sciences, Peking University, Beijing 100191, China. E-mail: yongjiang@bjmu.edu.cn; Tel/Fax: +86-10-82802719*

[b] *Modern Research Center for Traditional Chinese Medicine, Beijing University of Chinese Medicine, Beijing 100029, China*

[#]: These two authors contribute equally.

ABSTRACT

In this study, a new extracting diagnostic fragment ions (DFIs) in the MS^n chromatograms [E(DFI)MS^nCs]-based strategy was proposed to rapidly detect and identify the *in vivo* components derived from the extract of *Carthamus tinctorius* (ECT), using high performance liquid chromatography hyphenated with hybrid ion trap-time of flight mass spectrometry. In order to comprehensively summarize the DFIs for the global identification of *in vivo* constituents of ECT, the chemical profiling was carried out, and then the typical metabolic pathways of the primary components were proposed according to their chemical categories, by orally administering representative reference compounds. Based on the proposed metabolic pathways and the fragmentation rules, a DFIs schedule was constructed and adopted to differentiate and identify the metabolites from the endogenous substances in the MS^n chromatograms of ECT-treated biological samples, in combination with the neutral loss scan mode as a supplement. As a result, a total of 156 compounds were tentatively assigned *in vivo*, including 63, 73, 50, and 17 components from the rat plasma, urine, bile and feces, respectively, following oral administration of ECT. Deglycosylation, oxidation, methylation, sulfonation and glucuronidation were observed as the major metabolic pathways for the chemical constituents of ECT, and dehydroxylation was detected at the A-ring of flavones for the first time. The findings suggested that the E(DFI)MS^nCs-based strategy which integrated ideas from single compounds to herbal extracts and from extract chemical profiling to *in vivo* metabolite profiling, could be used as a reliable tool for rapid discovering and identifying the herb-related constituents *in vivo*.

Keywords: Herb-derived metabolite profile; Diagnostic fragment ions; MS^n chromatograms; Fragmentation pathways; *Carthamus tinctorius*

1. Introduction

Traditional Chinese medicines (TCMs) have drawn increasing worldwide interests due to the change of human disease spectrum, particularly, the prevalence of chronic and systematic diseases [1,2]. As multi-component and multi-target agents, TCMs exert holistic therapeutic actions; however, the characterization of the effective material basis that plays the therapeutic role cannot be easily achieved. One feasible method is to identify the components in the circulatory system since only the exposed xenobiotics could contribute to the therapeutic outcomes in most cases. However, the ability to profile the herb-derived components *in vivo* remains a significant challenge [1]. The difficulties are caused by not only the interference from the endogenous substances and the trace concentrations of most herb-related compounds, but also the diverse structures of the ingredients and their unpredictable metabolites.

Currently, dozens of reports are available concerning the post-acquisition data mining workflow to extract the metabolite information from a complex high resolution mass spectrometric (HR-MS) dataset. These techniques can be generally categorized as follows: 1) HR-MS extracted ion chromatography [2] has been used to detect the common metabolites formed *via* known or predictable metabolic pathways because the molecular weights of those metabolites are predictable; 2) the adoption of mass defect filter [3,4], isotope pattern filter [5] and background subtraction [3] could facilitate metabolite detection; 3) product ion filter and neutral loss filter have been employed for metabolites mining [6-10]; and 4) pattern recognition approaches, such as principal component analysis (PCA), partial least squares-discriminant analysis (PLS-DA) [11], as well as orthogonal partial least squares-discriminant analysis (OPLS-DA) [12], have been introduced to characterize the

xenobiotic metabolome. However, most of these techniques were only applied for the detection and identification of metabolites from chemical drugs or single ingredient of TCMs. The insufficient information about the herbal chemolome and/or metabolic information about the representative compounds tremendously hindered the metabolome characterization of TCMs.

Safflower (*Honghua* in Chinese), consisting of the dried flowers of *Carthamus tinctorius* L., is one of the most important medicinal materials in a number of prescriptions for the treatment of cardiovascular disorders. Modern pharmacological evaluations have demonstrated that the extract of *C. tinctorius* (ECT), which was mainly composed of the flavonoid constituents, such as quinochalcone *C*-glycosides, flavonol glycosides, and flavanone glycosides, exhibits promising antioxidant and cardioprotective effects [13-15]. Contrary to the extensive application of this herbal medicine or its extract, currently, there is no report available to address its metabolite profile *in vivo*.

In the current study, a new extracting diagnostic fragment ions (DFIs) in the MS^n chromatograms [E(DFI)MS^nCs]-based strategy was proposed to detect and characterize the ECT-derived components in rat. The strategy is illustrated in Fig. 1, including the construction and the application of DFIs. The DFIs schedule was constructed by: 1), proposal of the fragmentation patterns of its primary chemical homologues and characterization of the chemical profile of ECT; 2), identification of the metabolites and clarification of the metabolic pathways of three representative compounds *in vivo*. Afterwards, the herb-derived components *in vivo* were rapidly picked out according to E(DFI)MS^nCs, and identified using the proposed fragmentation rules. The findings obtained are expected to advance our understanding of the potentially active forms being

responsible for the health benefits of safflower and this study is expected to act as a model case for characterizing the *in vivo* metabolites of complex herbal medicines by applying this flexible strategy.

2. Materials and methods

2.1 Chemical and reagents

Acetonitrile (ACN) and methanol were of HPLC grade and purchased from Merck (Darmstadt, Germany). Deionized water was prepared by a Milli-Q water purification system (Millipore, MA, USA). Analytical grade formic acid and ammonia were obtained from Beijing Chemical Works (Beijing, China).

The safflower materials were collected from the Xinjiang province in China. The botanical origin was authenticated as the flowers of *C. tinctorius* L. by one of the authors, Prof. Peng-Fei Tu, and the voucher specimen (No. 20110301) was deposited in the herbarium of the Modern Research Center for Traditional Chinese Medicine, Peking University (Beijing, China). The ECT was prepared following the protocol described in our previous report [15]. The reference compounds, including hydroxysafflor yellow A (HSYA), kaempferol-3-*O*-rutinoside, 6-hydoxykaempferol-3-*O*-rutinoside, kaempferol-3-*O*-*β*-D-glucoside, rutin, quercetin-3-*O*-*β*-D-glucoside, 6-hydroxykaempferol-3,6,7-*tri*-*O*-*β*-D-glucoside, 6-hydroxykaempferol-3-*O*-*β*-D-glucoside, 6-hydroxykaempferol-6,7-*di*-*O*-*β*-D-glucoside, 6-hydroxykaempferol-3,6-*di*-*O*-*β*-D-glucosyl-7-*O*-*β*-D-glucuronide, isosafflomin C and safflomin C were previously identified from ECT in our group, and their structures were identified *via* analysis of their spectroscopic data (UV, MS and NMR) [15]. Their purities were determined to be greater than 98% by

the normalization of the peak areas detected by HPLC-UV, and confirmed by ^{1}H NMR analyses.

2.2 Preparation of the extract and references samples

ECT solution was prepared at a concentration of 8 mg·mL^{-1} using deionized water. All reference samples were prepared by dissolving each accurately weighed reference compound in an appropriate volume of 50% aqueous MeOH (with a final concentration of each reference of approximately 1 mg·mL^{-1}). All solutions were maintained at –20°C until use.

2.3 Preparation of the biological samples

Male Sprague-Dawley rats (12–14 weeks; 200–240 g) were provided by the Experimental Animal Center, Peking University Health Science Center, and all animal experimental protocols (LA2012-45) were approved by the Biomedical Ethical Committee of Peking University Health Science Center. The animals were acclimated at a temperature of 23 ± 1°C with a 12-h light/dark cycle and a relative humidity of 50% in an animal breeding room for three days before oral treatment. Standard chow and Milli-Q water were provided *ad libitum*. All rats fasted overnight but had free access to water prior to treatment. Rats were randomly divided into five groups. HSYA, 6-hydroxykaempferol-3-*O*-rutinoside, kaempferol-3-*O*-rutinoside and ECT freeze-dried powders were dissolved using saline and orally dosed at 100, 40, 80 and 4000 mg·kg^{-1}, respectively, whereas an equivalent amount of saline was administered to the vehicle group.

Blood sampling was performed by decapitation at 1.5 h (n = 5) following oral administration. Each blood sample was centrifuged (4000 rpm) for 10 min at 4°C, and the plasma was then pooled within a group and transferred to another clean tube. Bile, urine and fecal samples were collected over 0–12 h and 12–24 h and pooled within a group (n = 3). All of the samples were stored at –70°C until analysis.

Oasis® HLB solid phase extraction (SPE) columns (3 cc/60 mg, Waters, Milford, MA), which were successively preconditioned with 5 mL of methanol and 5 mL of deionized water, were used to process all biological samples. The feces were extracted with 10 volumes of 50% methanol under ultrasonification for 30 min. Then the supernatant (0.5 mL) was transferred to a clean test tube and evaporated to dryness under a gentle flow of nitrogen at 35°C. The residue was reconstituted in 0.5 mL deionized water to prepare fecal extract. Afterwards, all biofluids were mixed with 5% aqueous formic acid (v/v), vortexed for 1 min, and centrifuged (9600 rpm) for 10 min at 4°C. The supernatant was diluted with deionized water (1:1, v/v) and subsequently loaded onto a SPE column. Gradient elution for the plasma sample (loading volume of 1 mL) was performed using 2 mL of 2% aqueous formic acid, 2 mL of 60% aqueous methanol containing 2% ammonia, and 2 mL of methanol, successively; whereas the elution program for the bile, urine and fecal samples (loading volumes of 1 mL for bile and urine, and 0.5 mL for fecal extract) was 2 mL of 2% aqueous formic acid, 1 mL of 5% aqueous methanol containing 2% formic acid, 2 mL of 60% aqueous methanol containing 2% ammonia, and 2 mL of methanol, sequentially. All 60% aqueous methanol eluents were evaporated to dryness with nitrogen. The residues were reconstituted using 500 μL of 2% ACN and centrifuged at 12000 rpm for 10 min before an aliquot of supernatant was

subjected to LC-MS/MS analysis. The injection volumes of the single compound-treated samples and the ECT-treated samples were 15 μL and 5 μL, respectively. The preparation and measurement of the drug-free samples were performed in parallel with those of the treated samples.

2.4 LC-IT-TOF-MS^n analysis

HPLC-IT-TOF-MS^n analysis was performed on a Shimadzu Prominence HPLC system (CBM-20A controller, two LC-20AD binary pumps, an SPD-M20A diode array detector, an SIL-20AC auto-sampler, a CTO-20A column oven and a DGU-20A5 degasser) coupled to a IT-TOF-MS instrument (Shimadzu, Kyoto, Japan) through an ESI interface.

The chromatographic separation was performed on an AichromBond-AQ C_{18} column (250 mm × 4.6 mm i.d., particle size 5 μm, Abel Industries Ltd., Canada), which was protected by a Phenomenex C_{18} guard cartridge (4 mm × 2 mm i.d., particle size 5 μm, Torrance, CA, USA). The mobile phase consisted of ACN (A) and 0.1% aqueous formic acid (B), and was delivered in gradient as follows: 0–40 min, 2%–19%A; 40–60 min, 19%–22%A; 60–70 min, 22%–30%A; 70–85 min, 30–35%A; 85–86 min, 35%–95%A; flow rate, 1.0 $mL \cdot min^{-1}$. At the end of each run, 100%A was allowed to flush the column for 10 min and 2%A was delivered for the subsequent 10 min to re-equilibrate the entire system. The column temperature was maintained at 30°C.

The LC eluent was roughly split in a ratio of 5:1 (v/v) before entering the ion source from 3 min to 85 min for each sample. The optimized MS parameters were set as follows: alternative ion mode; electrospray voltage, +/−3.5 kV; detector voltage, 1.7 kV; the

temperature of curved desolvation line (CDL) temperature and heat block temperature, 200 °C; nebulizing gas (N_2), 1.5 L/min; drying gas (N_2) pressure, 100 kPa; scan ranges, *m*/*z* 100–1500 for MS^1, *m*/*z* 100–1200 for MS^2, *m*/*z* 50–800 for MS^3 and *m*/*z* 50–600 for MS^4; collision energy, 50% for MS^2, MS^3 and MS^4 with a region pressure of 1.4×10^{-4} Pa; ion trap pressure, 1.8×10^{-2} Pa; ion accumulation time, 30 ms. The accurate mass determination was calibrated using the sodium trifluoroacetate. Ultra-high purity argon was used as the collision gas for the collision-induced dissociation (CID) experiments. Moreover, the neutral loss-dependent acquisition mode was employed as a complementary tool to detect the metabolites. Data acquisition and analysis were achieved using LCMS Solution software package (Shimadzu).

3. Results and discussion

3.1 Biological sample preparation

It is feasible to reveal the therapeutic material basis of herbal medicines by the comprehensive characterization of the *in vivo* components because the exposed components usually play the determinant role for the efficacy of xenobiotics. However, one of the difficulties in profiling the absorbed constituents and the metabolites of herbal medicines is the interference from the endogenous substances and the biological macromolecules in biofluids. A desired biological sample preparation method for LC-MS^n analysis should exhibit efficient recovery for analyte, overcome drug-protein binding, and avoid matrix related ion-suppression in the ion source [16]. Solvent protein precipitation (SPP), liquid–liquid extraction (LLE) and solid phase extraction (SPE) have been reported as the most commonly used biological sample processing techniques. In

particular, SPE has been widely demonstrated as a convenient and efficient tool for complex sample preparation. After comparison, Oasis® HLB was found superior to Alltech C_{18} (Geneva, IL, USA) and Oasis® MCX (Waters) columns. Subsequently, the factors including the loading volume, the mobile phase composition, and the elution program, was systematically optimized for various biological samples. Afterwards, the findings from our preliminary study suggested that high sensitivity was achieved for this sample preparation protocol using the reference compounds (data not shown).

3.2 Construction of DFIs schedule for E(DFI)MS^nCs -based strategy

In general, compounds that share the same aglycone exhibit similar fragmentation behaviors in collision induced dissociation mode, thus generating certain common fragments, namely DFIs, due to their identical skeleton [17-22]. Given that most metabolites maintain the skeleton of their corresponding parent compounds, we hypothesized that DFIs could be also used for the rapid detection and identification of the prototypes and their corresponding metabolites *in vivo*. In comparison with the time-consuming and tedious workflow to identify metabolites using the total ion chromatograms (TICs) or base peak chromatograms (BPCs), MS^n chromatograms analysis would be advantageous at sensitivity, convenience, and reliability. However, the metabolites of herbal medicines are diverse due to their complex composition and the variety of metabolic pathways. Thus, the selection of appropriate DFIs is the critical step to achieve the comprehensive E(DFI)MS^nCs-based strategy. In the present study, the DFIs were preliminarily constructed by profiling the chemical composition of ECT and summarizing the mass fragmentation rules of the major chemical types. Subsequently, the representative

components were orally administered to determine the typical metabolic pathways of the primary chemicals of ECT, and the DFIs schedule was replenished by the metabolic pathways and the mass fragmentation pathways of the metabolites.

3.2.1. DFIs determined from the chemical components in ECT

Chemical profiling of ECT was carried out to screen out the primary chemical components in ECT and to summarize their mass fragmentation pathways. As aforementioned, quinochalcone *C*-glycosides, flavonol glycosides, and flavanone glycosides have been revealed as the dominant chemical components in ECT; thus, they were the primary targets in this study. The constituents were tentatively identified by comparing with the reference compounds, analyzing the chromatographic and spectrometric data, and referring to the proposed fragmentation rules summarized from the references (Supplemental Information A, Fig. S1) and archived in the literature [23,24]. Overall, 51 compounds were plausibly assigned (Fig. S2), including 15 chalcone derivatives, 9 flavanone glycosides (most were carthamidin or isocarthamidin analogues), 24 flavonol glycosides (5 quercetin glycosides, 8 kaempferol glycosides and 11 6-hydroxykaempferol glycosides), 1 phenylpropionic acid-4-*O*-glucoside, 1 cinnamic acid-4-*O*-glucoside and its isomer, along with 2 unknown components. The retention times, molecular weights, parent ions and identities of those compounds are presented in Table S1 and detailed descriptions are elucidated in Supplemental Information A.

The results showed that most flavonol glycosides and flavanone glycosides are *O*-glycosides, and few are quinochalcone *C*-glycosides. Step-wise neutral cleavages of the sugar substituents occurred for those *O*-glycosides generated the deprotonated

aglycone ions ($[A–H]^-$) as the most abundant fragments. $[A–H]^-$ ions were thereby chosen as the DFIs for filtering the *O*-glycosides and their metabolites. For example, the significant fragment ions at *m*/*z* 285.05, 287.04, 301.03 and 317.03 were considered as the DFIs for mining the kaemferol-, carthamidin-/isocarthamidin-, 6-hydroxykaempferol-/quercetin- and 6-hydroxyquercetin-type compounds *in vivo*, respectively. Subsequently, the other fragment ions, such as $[A–H–H_2O]^-$, $[A–H–H_2O–CO]^-$, and $[^{0,3}A]^-$ yielded by the cleavages at the *C*-ring [24] (Fig. S1B), were also adopted as DFIs to further unearth the possible metabolites.

3.2.2 DFIs proposed from the metabolites of the representative compounds

Most of the components identified in the ECT could be categorized into the derivatives of quinochalcones, flavonols and flavanones. Considering that the metabolic pathways of the carthamidin/isocarthamidin analogues (flavanone glycosides) have been well summarized in a previous report [25], thus, only the metabolic profiles of HSYA, 6-hydroxykaempferol-3-*O*-rutinoside and kaempferol-3-*O*-rutinoside, the delegates of quinochalcone glycosides and flavonol glycosides, were studied here to offer meaningful information for the metabolome clarification of ECT *in vivo*.

After oral administration, 37, 23, and 14 metabolites were observed in the biofluids of the kaempferol-3-*O*-rutinoside, 6-hydroxykaempferol-3-*O*-rutinoside and HSYA, respectively. The retention times, molecular weights, precursor ions, fragment ions and identities of their metabolites are illustrated in Supplementary Information A and B (Figs. S3-S5 and Tables S2-S4). Overall, similar metabolic patterns (Fig. 2) were revealed for these three representative compounds. Hydrolysis of the glycosidic bonds occurred

initially to afford aglycone, which subsequently conjugated with glucuronyl and/or sulfonyl groups according to phase II conjugation. In addition, ring-fission metabolism, which was speculated to be mediated by the intestinal bacteria, was observed for the aglycones, and the products were further conjugated with glucuronyl, sulfonyl and methyl groups like above. Furthermore, it is interesting to note that dehydroxylation of the flavonol aglycones was observed at the A-ring for the first time, which might be catalyzed by the gut bacteria [26].

Based on the mass spectral information from the metabolites of representative compounds, various deprotonated aglycones ($[A–H]^-$) were observed as the significant fragment ions in the multistage MS data. As a consequence, the deprotonated aglycone ($[A–H]^-$) that can be predicted from the metabolic pathways were chosen as the major DFIs to screen the metabolites *in vivo* of ECT. In addition, the fragment ions such as 137.02, 151.04 and 165.04, were also taken as DFIs for the ring-fission metabolites of quinochalcone derivatives. The DFIs are summarized in Table 1.

3.2.3 Supplemental neutral loss scan for MS^nE(DFI)Cs-based strategy

Generally, the MS^n information of the compounds was collected by dissociating the most abundant precursor ions in step-wise using automatic tandem mass spectrometry. With the assistance of neutral loss scan, the multi-stage mass spectrometry will be achieved by dissociating the ions resulted from the pre-defined neutral losses, which could enhance the sensitivity and selectivity of data acquisition. Thus, in order to globally characterize the herb-related components *in vivo*, the neutral loss scan was adopted as a supplement of the aforementioned DFIs extraction. The mass spectra of parent compounds suggest that

the dissociation of glucosyl (162.05 Da), rhamnosyl (146.06 Da), carbon monoxide (27.99 Da) and hydrogen oxide (18.01 Da) were the prominent neutral cleavages. The metabolic study of the three representative compounds revealed that the characteristic neutral losses of the glucuronsyl (176.05 Da), and sulfonyl (79.95 Da) residues, or radical cleavage of methyl (15.02 Da) group widely occurred for those phase II metabolites in the MS^n chromatograms. Therefore, those cleavages (176.05, 162.05, 146.06, 79.95, 27.99, 18.01, and 15.02 Da) were introduced to the neutral list for the detection of the ECT-derived components in rat.

3.2.4 Accomplishment of the E(DFI)MS^nCs-based strategy

The E(DFI)MS^nCs were employed to act as an efficient filter to detect the metabolites in the complex biological samples, and their structures were then deduced based on the accurate mass and the proposed fragmentation rules. Because the herb-related components may exist *in vivo* as the prototypes and the metabolites, and most are present at a trace concentration, the following workflow was adopted to identify the ECT-derived compounds in the drug-treated biological samples.

First, the prototypes were mined by extracting DFIs from the biological samples, and then matching the retention times and the mass spectral profiles with ECT. Second, extraction of the DFIs was performed in the MS^2 and MS^3 chromatograms of the ECT-treated and vehicle samples. Except the prototypes, the additional peaks in ECT-treated samples were regarded as the metabolites compared with the drug-free samples (Fig. S6). Moreover, comparison was also carried out between ECT-treated and single compound-administrated samples to map the identical metabolites. The

corresponding MS spectral information was employed to characterize the metabolites with the assistance of the proposed fragmentation and metabolic pathways. The additional signals detected using the neutral loss scan were also regarded as the metabolites, which could be subsequently identified based on their mass spectral profiles.

3.3 Identification of the ECT-derived compounds in the biological samples

A total of 156 compounds were tentatively detected as ECT-derived substances *in vivo*, including 63, 73, 50 and 17 components from rat plasma, urine, bile and feces, respectively (Fig. 3), in which the identities of 13 metabolites (U1–U13) could not be determined due to insufficient information. The detected metabolites were categorized into four types (Ⅰ–Ⅳ) based on their aglycones and generation pathways. The type Ⅰ components could be separated into the chalcone/flavanonol glycosides (Ⅰa) and the dihydrochalcone glycosides (Ⅰb). Owing that the flavonols can be metabolized to chalcones and flavanonols by the intestinal bacteria, and it is difficult to definitely differentiate chalcones from flavanonols using mass spectral information. Therefore, chalcone and flavanonol derivatives were grouped into the same subtype Ⅰa. The dihydrochalcone analogues, which are the reduced products of chalcones or flavonols, were assigned to subtype Ⅰb. The flavanone (carthamidin/isocarthamidin) analogues, which are the unique flavanones in the ECT, were sorted into type Ⅱ due to their characteristic fragmentation pathways (Fig. S1). The flavonoids in type Ⅲ can be further divided into five subtypes due to their different aglycones, namely kaempferol-(Ⅲa),

6-hydroxykaempferol-/quercetin-(IIIb), 6-hydroxyquercetin-(IIIc), 6-hydroxymyricetin-(IIId) and other flavonol-(primarily trihydroxyflavonol, dihydroxyflavonol and monohydroxyflavonol, IIIe) derivatives, respectively. The ring-fission products, primarily phenolic acids, which were derived from the chalcones, carthamidins or flavonols by the intestinal bacteria, were sorted into type Ⅳ (Table 1).

The identification of certain representative compounds is described as follows, whereas the detailed information of other compounds is given in Table 2 and Table S5.

3.3.1 Flavanonol- or chalcone-type metabolites

Twenty-two metabolites (A1–A22) were deduced as flavanonol or chalcone derivatives, of which 10 were confirmed as prototypes in comparison with the chemical profile of ECT, and the other 12 metabolites were mined by E(DFI)MS^nCs using *m/z* 239.07, 255.07, 271.06, 287.05, 303.05, and 319.05. Among those 12 metabolites, A13, A18 and A19 were also observed as the metabolites of HSYA (A6, A8 and A9 in Table S4), and A7 and A15 were identical to the two metabolites of kaempferol-3-*O*-rutinoside (A3 and A11 in Table S2). The other metabolites (A4, A8, A9, A12, A14, A17 and A20) were tentatively assigned using their HR-MS data, most of which were glucuronides and sulfonates (Supplemental Information A).

3.3.2 Dihydrochalcone-type metabolites

With the assistance of the DFIs at *m/z* 225.098, 241.09, 257.09, 273.08, 289.07, 305.07, and 321.07, which are a serial of $[A–H]^-$ ions of dihydrochalcones, a total of 11 dihydrochalcone-type metabolites (H1–H11) were detected in the MS^n chromatograms.

The elemental composition of H1 was determined as $C_{21}H_{22}O_8$ based on its mass spectral data. With the observation of the neutral loss of 176.05 Da, H1 was assigned as dihydrochalcone-2-*O*-glucuronide. Because H2 and H3 showed the successive neutral losses of 176.05 Da and 79.95 Da, they were deduced as the glucuronidative sulfonated conjugates of 2,6-dihydroxydihydrochalcone. In addition, H4 and H5 were assigned as the glucuronidated conjugates of α,2,4,6-tetrahydroxydihydrochalcone. H6, H7, H8, H9 and H10 exhibited the same $[A–H]^-$ ion at *m/z* 257.0802, which is consistent with the molecular composition of α,2,6-trihydroxydihydrochalcone. H7 and H8 were proposed as the glucuronidated conjugates with the neutral loss of 176.05 Da. H6, H9 and H10 were proposed as glucuronidated products of the α,2,6-trihydroxydihydrochalcone sulfonates because they exhibited the neutral loss of 176.05 Da in addition to the cleavage of 79.95 Da. Accordingly, H11 was deduced as α,2,4,4′,6-pentahydroxydihydrochalcone glucuronide owing to the observation of the neutral loss of 176.05 Da.

3.3.3 Carthamidin- or isocarthamidin-type (flavanone-type) metabolites

Twenty-two metabolites (B1–B22) were identified as carthamidin or isocarthamidin derivatives according to the DFIs of $[A–H]^-$ at *m/z* 287.05, $[M–H–B–ring]^-$ at *m/z* 193.01, $[^{1,2}A]^-$ at *m/z* 181.01, $[^{1,2}A–CO]^-$ at *m/z* 153.03 and $[^{1,3}A]^-$ at *m/z* 167.00 [25]. By comparing with the chemical components in ECT, B1 and B6 were identified to be prototypes (B3 and B6 in Table S1); whereas, B10 was assigned as carthamidin or isocarthamidin using the MS^n spectra. The other 17 metabolites were assigned as the glucuronides and sulfonates of the carthamidin or isocarthamidin analogues due to the detection of the diagnostic neutral loss of 176.05 or 79.95 Da. Detailed descriptions about the 17

metabolites can be found in Supplemental Information A.

3.3.4 Kaempferol-type metabolites

When extracting the DFI at *m/z* 285.04 ($[A–H]^-$), 17 metabolites (D1–D17) were found. These metabolites were further identified using their molecular formula ions, and the corresponding products ions of $[A–H–CO]^-$, $[A–H–H_2O]^-$ and $[^{1,3}A]^-$. D3, D5, D6, D7, D11, D12, and D13 were plausibly assigned as the metabolites of kaempferol-3-*O*-rutinoside due to their identical retention times and fragmentation behaviors with those (D1, D5, D6, D7, D11, D12, and D13 in Table S2) observed in the kaempferol-3-*O*-rutinoside-treated samples. In addition, D9 was identified as kaempferol-*O*-sulfo-*O*-glucuronide by a comparison with the metabolite profile of 6-hydroxykaempferol. The other 10 metabolites were plausibly characterized based on observation of the neutral losses of 176.05 Da or 79.95 Da. Detailed information about those 10 metabolites can be found in Supplemental Information A.

3.3.5 6-Hydroxykaempferol- or quercetin-type compounds

Nineteen metabolites (E1–E19) were categorized as 6-hydroxykaempferol- or quercetin-derived compounds, including 1 prototype (E1) and 18 metabolites, which were captured using DFI of *m/z* 301.03 ($[A–H]^-$). E2, E4, E7, E8, E9, E10, E13, E16 and E19 were revealed to be identical to the metabolites of 6-hydroxykaempferol-3-*O*-rutinoside, namely E1, E2, E3, E4, E5, E6, E7, E8 and E10 in Table S3, respectively. E3, E11 and E12 were assigned as 6-hydroxykaempferol-*di*-*O*-glucuronide due to the presence of the characteristic fragment ions at *m/z* 151.0419 and 179.0018, whereas E5, E6, E14, E15,

E17 and E18 were tentatively deduced as the methylated and glucuronidated products of 6-hydroxykaempferol and/or quercetin (see Supplemental Information A).

3.3.6 6-Hydroxyquercetin analogues

One prototype (C1) and 1 metabolite (C2) were identified as the 6-hydroxyquercetin analogues by extracting DFI at *m/z* 317.03 ($[A–H]^-$). Due to the occurrence of the neutral cleavage of 176.05 Da, C2 was proposed as 6-hydroxyquercetin-*O*-glucuronide.

3.3.7 Other flavone-type metabolites

In addition to the previously mentioned components, 14 metabolites (G1–G14) were assigned to the flavone-type metabolites, including five dihydroxyflavone-type metabolites (G1–G5) and nine trihydroxyflavone derivatives (G6–G14), which were detected using DFIs ($[A–H]^-$ ions) of *m/z* 269.04 and 253.05, respectively. The nine trihydroxyflavone (G6–G14) metabolites were speculated as galangin regio-isomers, the generation of which was mediated by the intestinal bacteria from kaempferol *via* dehydroxylation [26,27]. At the meanwhile, the five dihydroxyflavone-type metabolites were deduced as the regio-isomers of 3,5-dihydroxyflavone, which were afforded by the step-wise hydroxylation of kaempferol. Detailed information about the 14 metabolites can be found in Supplemental Information A.

3.3.8 The phenolic acid-type metabolites

Using the DFIs filter, 36 phenolic acids (O1–O36) were found, most of which were generated from the flavonoids by the ring-fission metabolism. Based on the proposed

metabolic pathway and the mass spectral profiles, these metabolites were tentatively characterized as the analogues of (4-hydroxy-phenyl)-acetic acid, 2,3,4,6-tetrahydroxy-benzoic acid, 2,4,6-trihydroxy-benzoic acid, 4-hydroxyphenylpropionic acid, (3-hydroxy-phenyl)-acetic acid, benzene-1,2,3,5-tetraol, coumaric acid and phloroglucinol. Detailed information about the 36 metabolites can be found in Supplemental Information A.

Deglycosylation, oxidation, methylation, sulfonation, glucuronidation and ring-fission were detected as the dominant metabolic pathways for the components in the flavones-enriched extract. It is interesting to observe that dehydroxylation occurred for poly-hydroxylated flavones. It was found that 6-hydoxykaempferol-3-*O*-rutinoside was the C-6 hydroxylated product of kaempferol-3-*O*-rutinoside, and certain metabolites of 6-hydoxykaempferol-3-*O*-rutinoside (Table S3) were found to be identical to A1, A2, A3, A4, A5, A7, A9, A10, and D2 of kaempferol-3-*O*-rutinoside (Table S2), unambiguously indicating that dehydroxylation occurred at the C-6 position of 6-hydoxykaempferol-3-*O*-rutinoside. In previous reports, dehydroxylation could occur at the B-ring of the poly-hydroxylated flavones (for example catechin [28] and hyperoside [29]). For the first time, the occurrence of dehydroxylation at the A-ring of polyhydroxyflavones was observed. Generally, most of the metabolic pathways were catalyzed by the enzymes in the liver, such as CYP450s, glucuronyltransferases and sulfotransferases, corresponding to oxidation, glucuronidation and sulfonation, respectively; however, a growing number of reports have revealed the key role of intestinal bacteria for the metabolism of xenobiotics. In the present study, ring-fission metabolism and deglycosylation are attributed to the enzymes from intestinal bacteria [30].

In addition, an increasing number of articles have suggested that dehydroxylation should be also catalyzed by intestinal bacteria [28, 31]. Therefore, it is reasonable to hypothesize that the metabolic profile of ECT *in vivo* was constructed by the combined roles of intestinal bacteria and the enzymes in the liver and intestine tissues.

A holistic operation and a synergetic effect have been proved for TCMs by an increasing number of studies, and a vast number of reports suggested that the components in the blood stream provide the primary therapeutic effects. However, it is difficult to globally and precisely characterize the metabolome of TCMs *in vivo*, despite the crucial necessity to reveal the effective material basis. To date, seldom reports have been available concerning the metabolic profiles of herbal drugs based on a systematic procedure, which suffered from tedious procedures and a high possibility of missing detection [32,33]. Generally speaking, the qualitative analysis of the plant extract-derived components by LC-MS relies on the detection of the analytes using a full-scan analysis. However, this attempt is challenging because the ions of interest, especially those at trace levels, are typically masked by the background interferences or by endogenous components. Moreover, herbal medicines contain hundreds of components; a single component can give rise to several quasi-molecular ions, and each quasi-molecular ion can produce a series of fragment ions in CID mode. Fortunately, the compounds presented in herbal medicine could be structurally classified into several families, and the chemical homologues that share the same carbon skeletons or substructures could yield the same fragment ions through characteristic fragmentation patterns by tandem mass spectrometry [22]. Therefore, these characteristic fragment ions could be adopted as the diagnostic foundation for the identification of different chemical types. For example, the

DFIs of the carthamidin and isocarthamidin derivatives including $[^{1,2}A]^-$, $[^{1,2}A–CO]^-$, and $[^{1,3}A]^-$ ions, could be detected at *m/z* 181.01, 153.03 and 167.00, respectively, corresponding to characteristic cross-ring cleavages, and the $[A–H]^-$ and $[M–H–B\text{-ring}]^-$ ions could be observed at *m/z* 287.05 and 193.01, respectively, attributing to the dehydrogenation and B-ring neutral loss of the aglycone. To date, several applications of the DFIs filtering approach have been reported in the literature [17,18,22], which has been proved to be an efficient method for the extraction of the related compounds. However, there is none report for adopting this method for the metabolite identification. Therefore, an attempt to universally characterize the metabolites of TCM *in vivo* was performed in this case using E(DFI)MS^nCs-based strategy. Chemical profiling was performed to reveal the major structure types in ECT and the MS fragmentation pathways to preliminarily construct DFIs list. Simultaneously, the metabolic features for the major structure types in ECT were proposed by orally administrating the representative compounds, and the DFIs from the metabolic pathways were replenished to the preliminary list. The DFIs list was subsequently employed to act as an efficient filter to rapidly detect the metabolites in the complex biological samples by extracting the DFIs in the MS^n chromatograms.

4. Conclusions

The metabolic studies of TCMs have been challenging due to their complicated chemical composition and the relatively low concentrations in biological samples. In this study, we proposed a practical E(DFI)MS^nCs-based strategy to comprehensively profile the metabolome of ECT in rats following oral administration. First, 51 components were identified from ECT using the proposed fragmentation pathways. Second, the *in vivo*

metabolism of three representative compounds was studied to facilitate the understanding of the ECT metabolite profile. Finally, 156 compounds were assigned to the ECT-derived analytes using the DFIs filter, including 63 from plasma, 73 from rat urine, 50 from bile and 17 from feces, after oral administration of ECT. The results obtained in this study globally revealed the potential efficacy material basis of ECT *in vivo*, and also suggested that E(DFI)MS^nCs-based strategy which integrated ideas from single compounds to herbal extracts and from extract chemical profiling to *in vivo* metabolite profiling, could be used as a reliable tool for rapid discovering and identifying the herb-related constituents *in vivo*.

Acknowledgement

This work was financially supported by National Natural Science Foundation for Excellent Young Scholars of China (No. 81222051), and National Key Technology R&D Program "New Drug Innovation" of China (Nos. 2012ZX09103201-036, 2012ZX09301002-002-002 and 2012ZX09304-005).

Appendix A&B Supplementary data

Supplementary data (Supplemental information A and B) associated with this article can be found, inthe online version, at http://dx.doi.org/.....

References:

1. M. Baker, *Nat. Methods*, 2011, **8**, 117–122.
2. Q. Ruan, Q.C. Ji, M.E. Arnold, W.G. Humphreys, and M. Zhu, *Anal. Chem*., 2011,

83, 8937–8944.

3. M. Liu, S. Zhao, Z. Wang, Y. Wang, T. Liu, S. Li, C. Wang, H. Wang, and P. Tu, *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.*, 2014, **949–950**, 115–126.

4. J.L. Geng, Y. Dai, Z. Yao, Z. Qin, X. Yao, and W.L. Qin, *J. Pharm. Biomed. Anal.*, 2014, **96**, 90–103.

5. J. Guo, M. Zhang, C.S. Elmore, and K. Vishwanathan, *Anal. Chim. Acta*, 2013, **780**, 55–64.

6. J. Zhang, Z.H. Huang, X.H. Qiu, Y.M. Yang, D.Y. Zhu, and W. Xu, *PloS one,* 2012, **7**, e52352.

7. Y.Q. Huang, Q.Y. Wang, J.Q. Liu, Y.H. Hao, B.F. Yuan, and Y.Q. Feng, *Analyst,* 2014, **139**, 3446-3454.

8. S.G. Ma, and M.S. Zhu, *Chem. Biol. Interact.*, 2009, **179**, 25–37.

9. M.S. Zhu, D.L. Zhang, H.Y. Zhang, and W.C. Shyu, *Biopharm. Drug Dispos.*, 2009, **30**, 163–184.

10. Q. Ruan, S. Peterman, M.A. Szewc, L. Ma, D. Cui, W.G. Humphreys, and M.S. Zhu, *J. Mass Spectrom.*, 2008, **43**, 251–261.

11. G. Tan, M. Liu, X. Dong, S. Wu, L. Fan, Y. Qiao, Y. Chai, and H. Wu, *J. Pharm. Biomed. Anal.*, 2014, **96**, 187–196.

12. M.X. Guo, L. Zhang, H.Y. Liu, L.L. Qin, Z.X. Zhang, X. Bai, and X. Y. Gao, *J. Sep. Sci*., In press, DOI: 10.1002/jssc.201400339.

13. X. Wei, H. Liu, X. Sun, F. Fu, X. Zhang, J. Wang, J. An, and H. Ding, *Neurosci. Lett.*, 2005, **386**, 58–62.

14. J.F. Chen, P.F. Tu, Y. Jiang, *J. Chin. Pharm. Sci.*, 2014, **23**, 490-495.

15. H.X. Li, S.Y. Han, X.W. Wang, X. Ma, K. Zhang, L. Wang, Z.Z. Ma, and P.F. Tu, *Food Chem. Toxicol.*, 2009, **47**, 1797–1802.

16. Y. Song, W. Jing, F. Yang, Z. Shi, M. Yao, R. Yan, and Y. Wang, *J. Pharm. Biomed. Anal.,* 2014, **88**, 269–277.

17. J.Y. Zhang, Q. Zhang, N. Li, Z.J. Wang, J.Q. Lu, and Y.J. Qiao, *Talanta*, 2013, **104**, 1–9.

18. S. Wang, L. Chen, J. Leng, P. Chen, X. Fan, and Y. Cheng, *J. Pharm. Biomed. Anal.,* 2014, **98**, 22-35.

19. W. Xu, J. Zhang, D. Zhu, J. Huang, Z. Huang, J. Bai, and X. Qiu, *J. Sep. Sci.*, 2014, in press. doi: 10.1002/jssc.201400365.

20. J. Sultan, and W. Gabryelski, *Anal. Chem.*, 2006, **78**, 2905–2917.

21. R. Li, Y. Zhou, Z. Wu, and L. Ding, *J. Mass Spectrom.*, 2006, **41**, 1–22.

22. C. Zheng, H. Hao, X. Wang, X. Wu, G. Wang, G. Sang, Y. Liang, L. Xie, C. Xia, and X. Yao, *J. Mass Spectrom.*, 2009, **44**, 230–244.

23. Y. Jin, X.L. Zhang, H. Shi, Y.S. Xiao, Y.X. Ke, X.Y. Xue, F.F. Zhang, and X.M. Liang, *Rapid Commun. Mass Spectrom.*, 2008, **22**, 1275–1287.

24. Y. Jin, Y.S. Xiao, F.F. Zhang, X.Y. Xue, Q. Xu, and X.M. Liang, *J. Pharm. Biomed. Anal.*, 2008, **46**, 418–430.

25. Y. Miyake, K. Shimoi, S. Kumazawa, K. Yamamoto, N. Kinae, and T. Osawa, *J. Agric. Food Chem.*, 2000, **48**, 3217–3224.

26. L.A. Griffiths, and G.E. Smith, *Biochem. J.*, 1972, **130**, 141–151.

27. X. Meng, S. Sang, N. Zhu, H. Lu, S. Sheng, M.J. Lee, C.T. Ho, and C.S. Yang, *Chem. Res. Toxicol.*, 2002, **15**, 1042–1050.

28. L.Q. Wang, M.R. Meselhy, Y. Li, N. Nakamura, B.S. Min, G.W. Qin, and M. Hattori, *Chem. Pharm. Bull.*, 2011, **49**, 1640–1643.

29. J. Yang, D. Qian, J. Guo, S. Jiang, E.X. Shang, J.A. Duan, and J. Xu, *J. Ethnopharmacol.*, 2013, **147**, 174–179.

30. Y. Song, X. Yang, Y. Jiang, and P. Tu, *J. Pharm. Biomed. Anal.*, 2012, **70**, 700–707.

31. J. Yang, D. Qian, S. Jiang, E.X. Shang, J. Guo, and J.A. Duan, *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.*, 2012, **898**, 95–100.

32. C. Xiang, X. Qiao, Q. Wang, R. Li, W. Miao, D. Guo, and M. Ye, *Drug Metab. Dispos.,* 2011, **39**, 1597–1608.

33. P. Gong, N. Cui, L. Wu, Y. Liang, K. Hao, X. Xu, W. Tang, G. Wang, and H. Hao, Anal. Chem., 2012, **84**, 2995–3002.

Figure Legends

Fig. 1 Flow chart of the diagnostic fragment ions filter-based strategy.

Fig. 2 Proposed metabolic pathways of three representative compounds in rats. A, proposed metabolic pathways of kaempferol-3-*O*-rutinoside; B, proposed metabolic pathways of 6-hydroxykaempferol-3-*O*-rutinoside; and C, proposed metabolic pathways of HSYA.

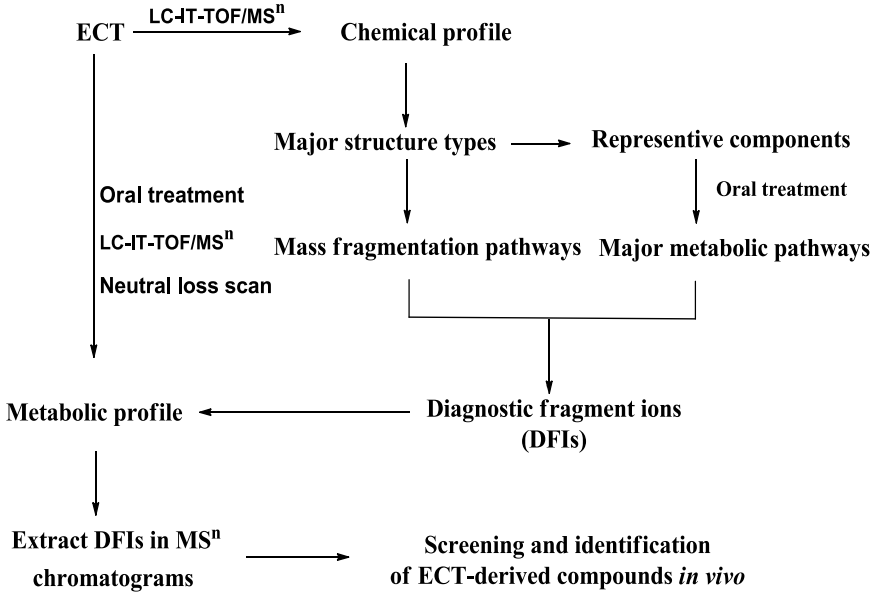
Fig. 3 The extracted ion current chromatograms (EICs) of the vehicle and the biological samples in rats after oral administration of ECT. A, plasma; B, bile; C, urine; D, feces. 1, vehicle sample; 2, ECT-treated sample.

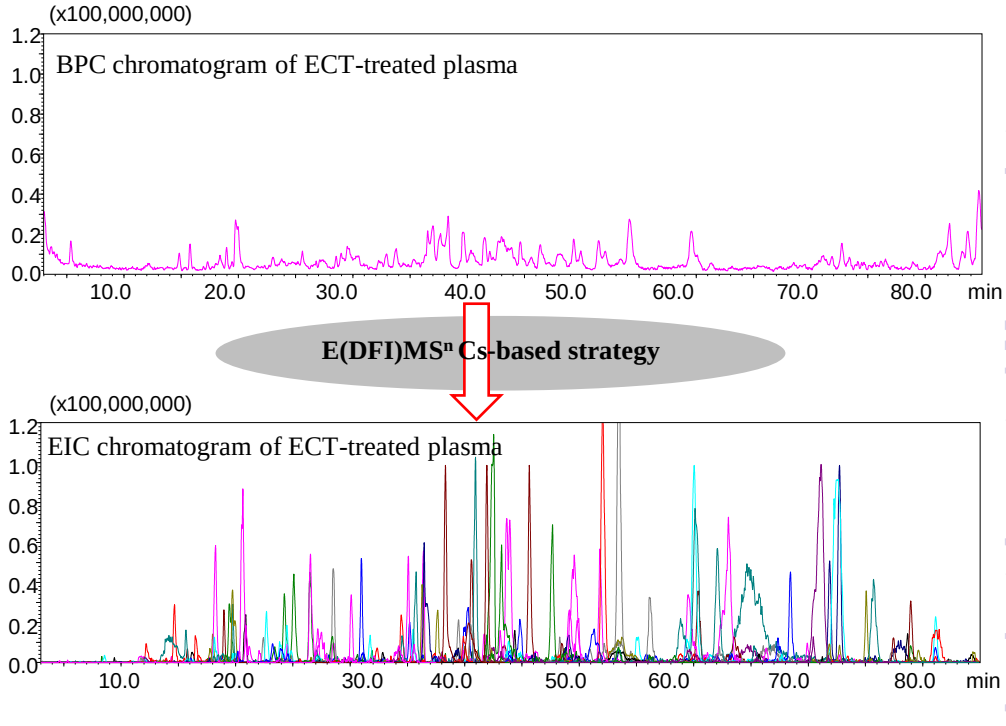
Table captions

Table 1 The DFIs schedule for the detection and identification of the metabolites in rats after oral administration of ECT by E(DFI)MS^{n}Cs-based strategy

Table 2 Identification of the metabolites in rats after oral administration of ECT

An E(DFI)MS^{n}Cs-based strategy was proposed to rapidly detect and identify the *in vivo* components derived from the extract of *Carthamus tinctorius* using LC-IT-TOF-MS^n.





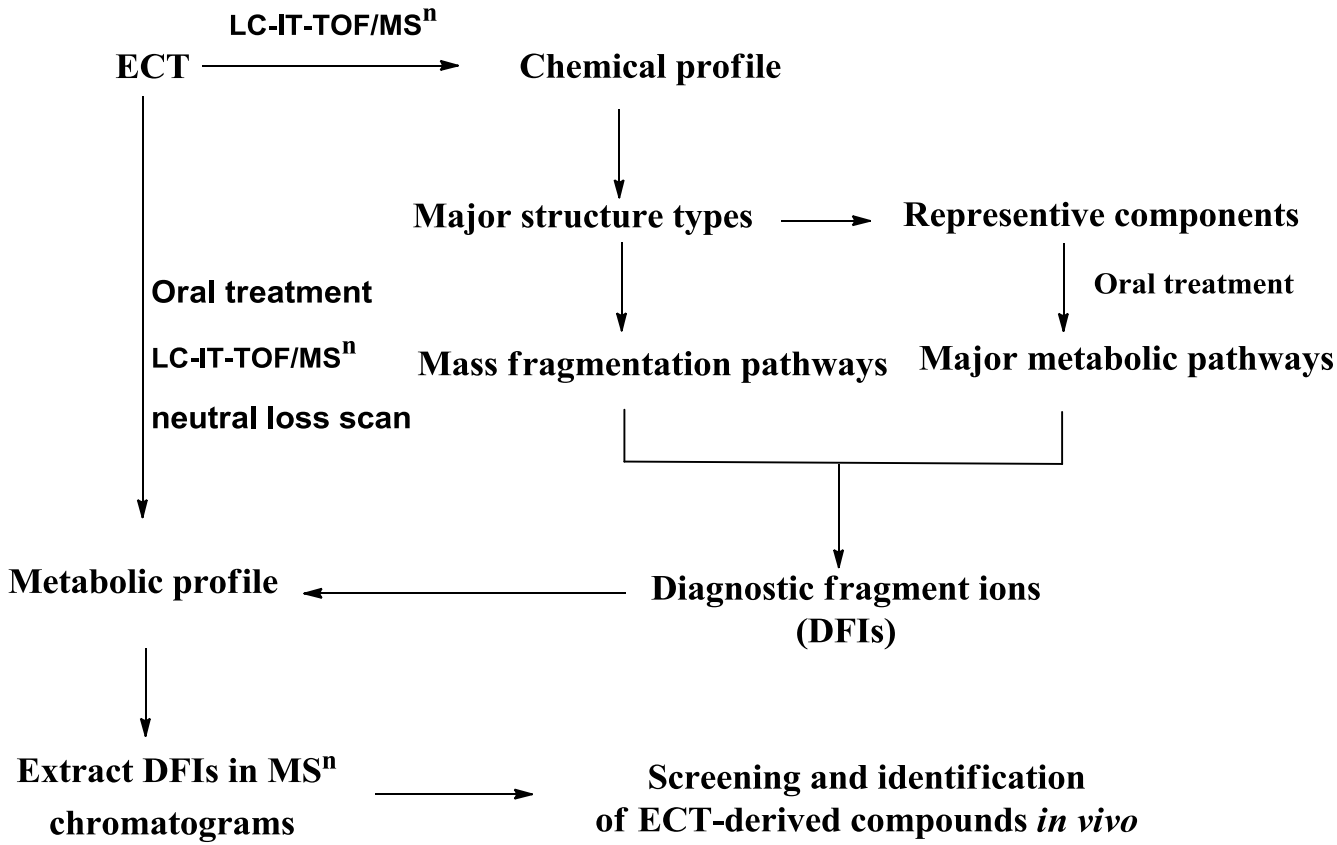


Fig. 1

A

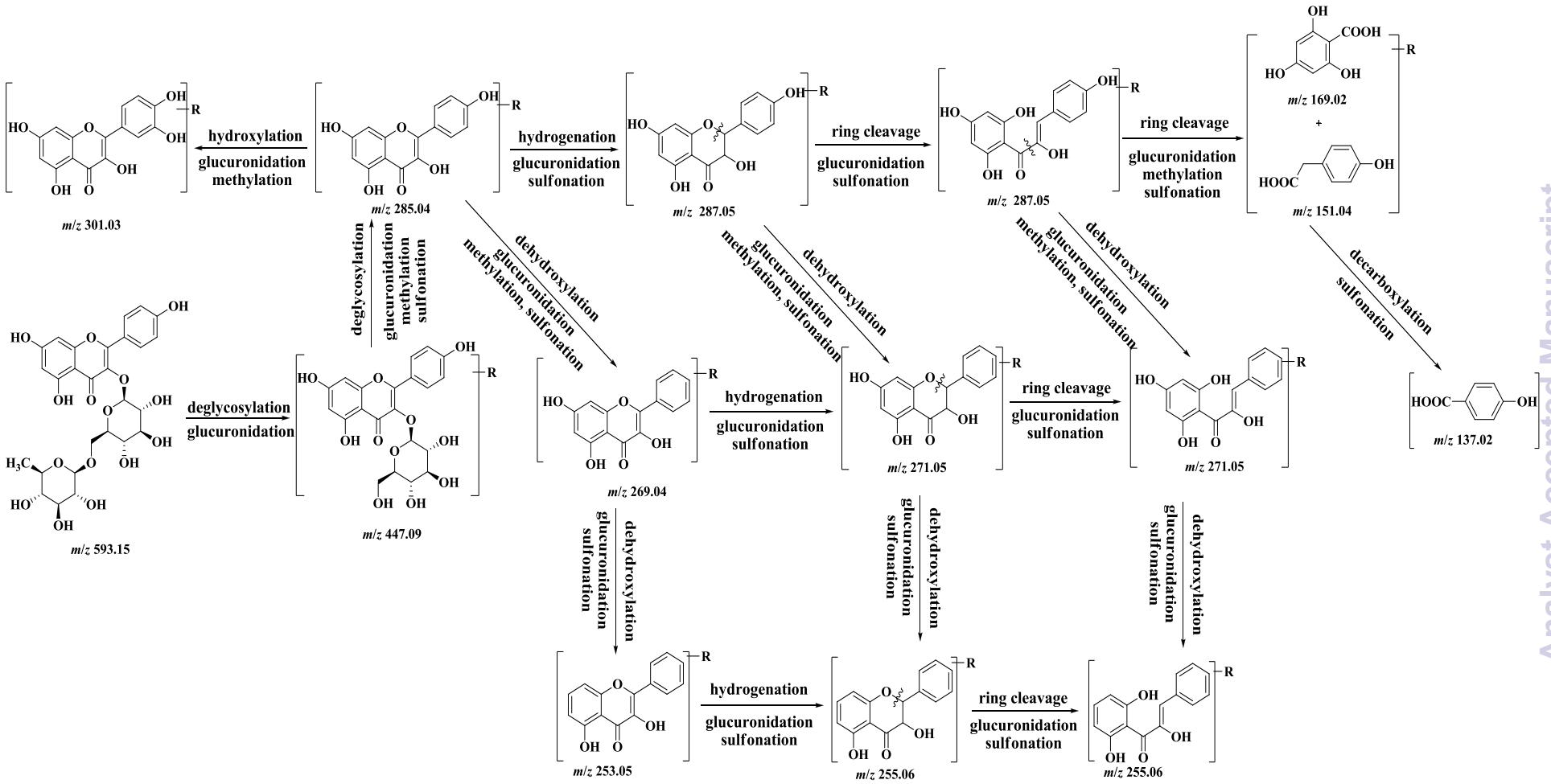


Fig. 2A

B

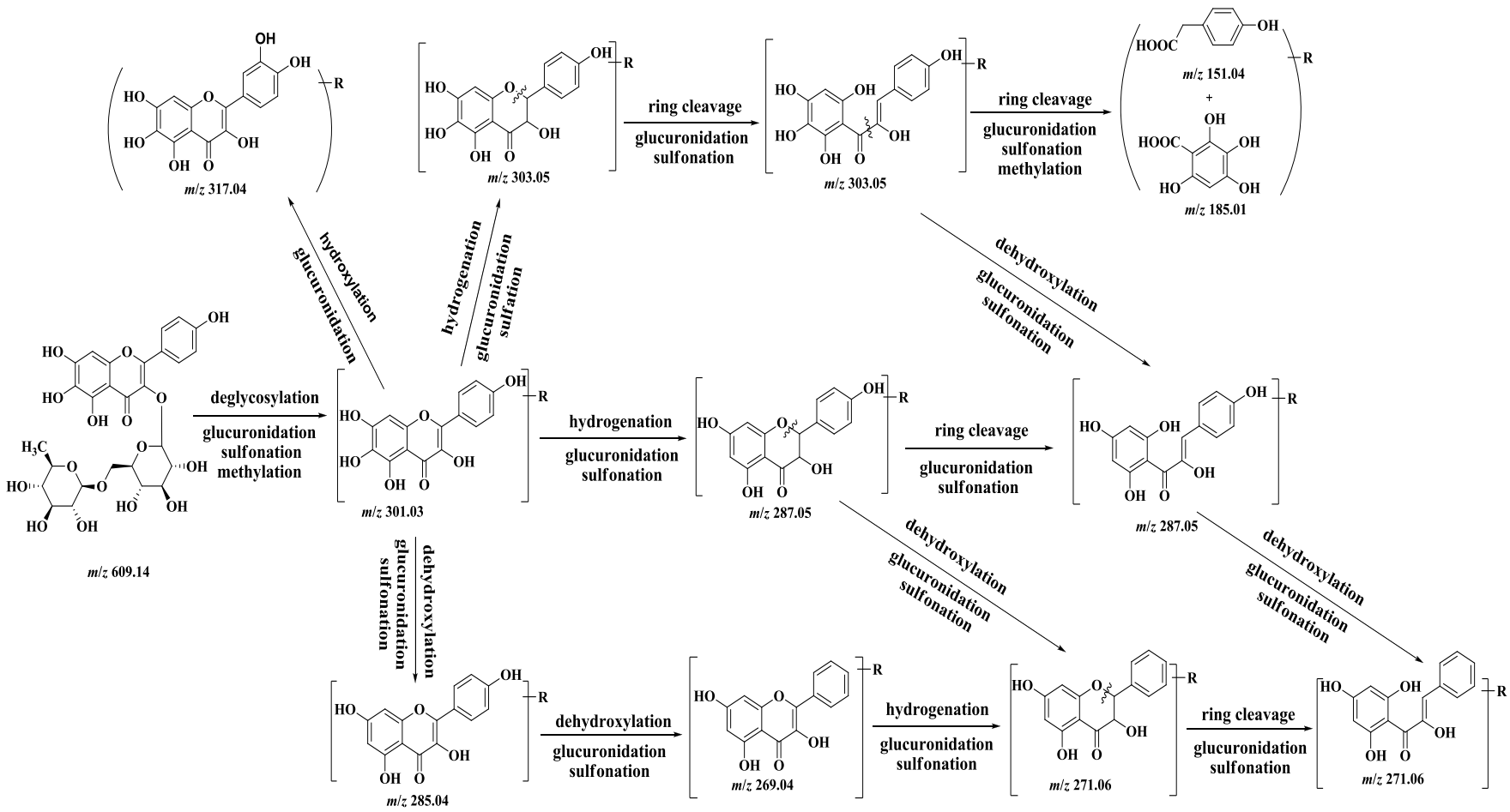


Fig. 2B

C

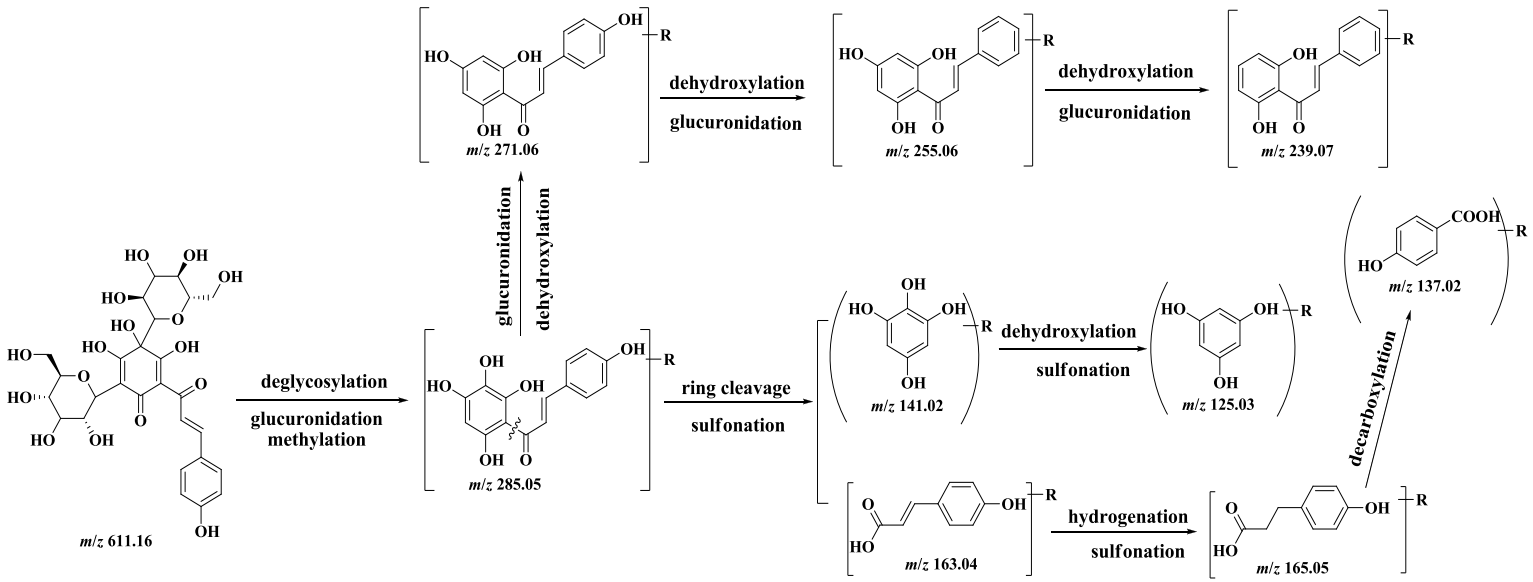


Fig. 2C

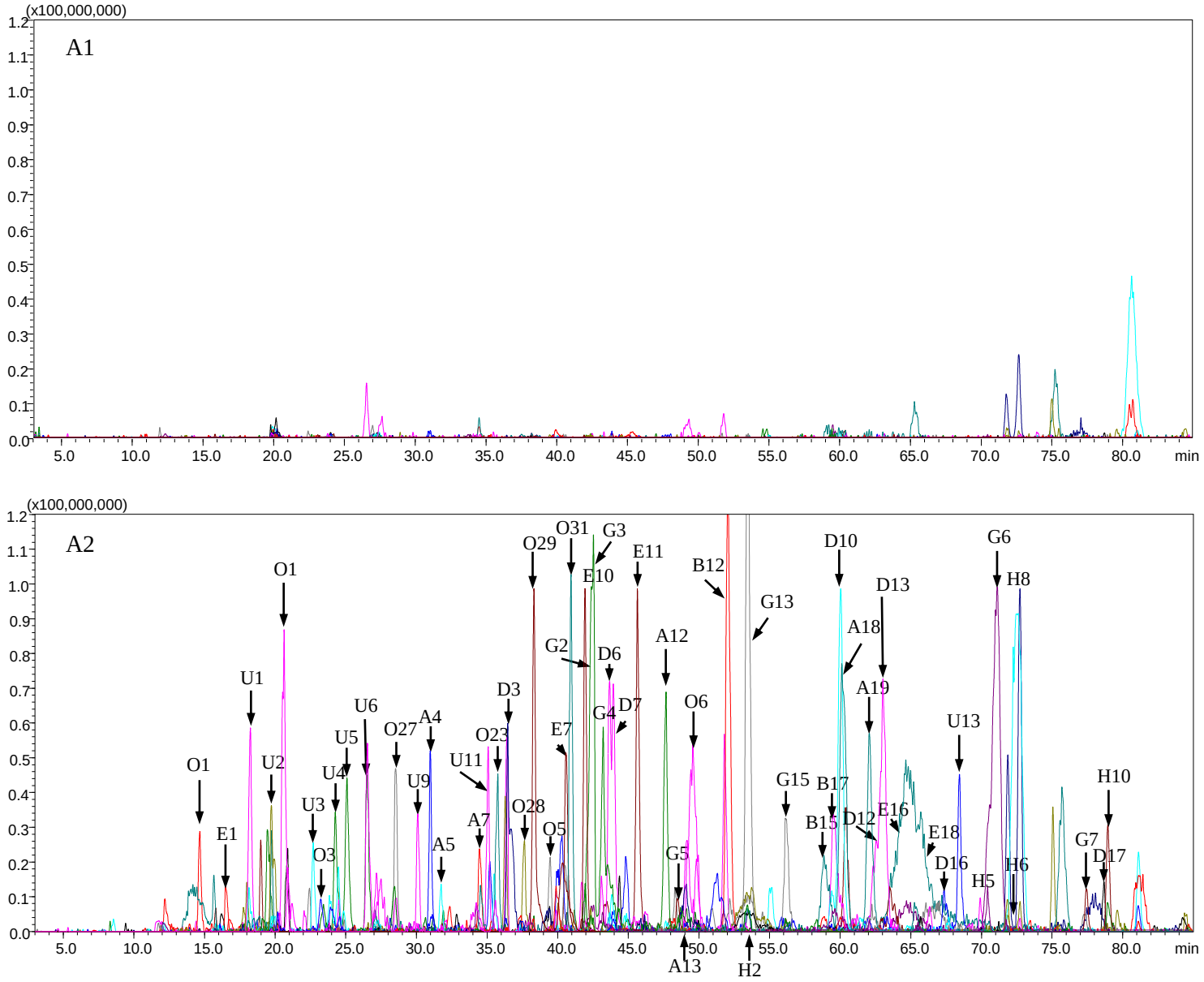


Fig. 3A

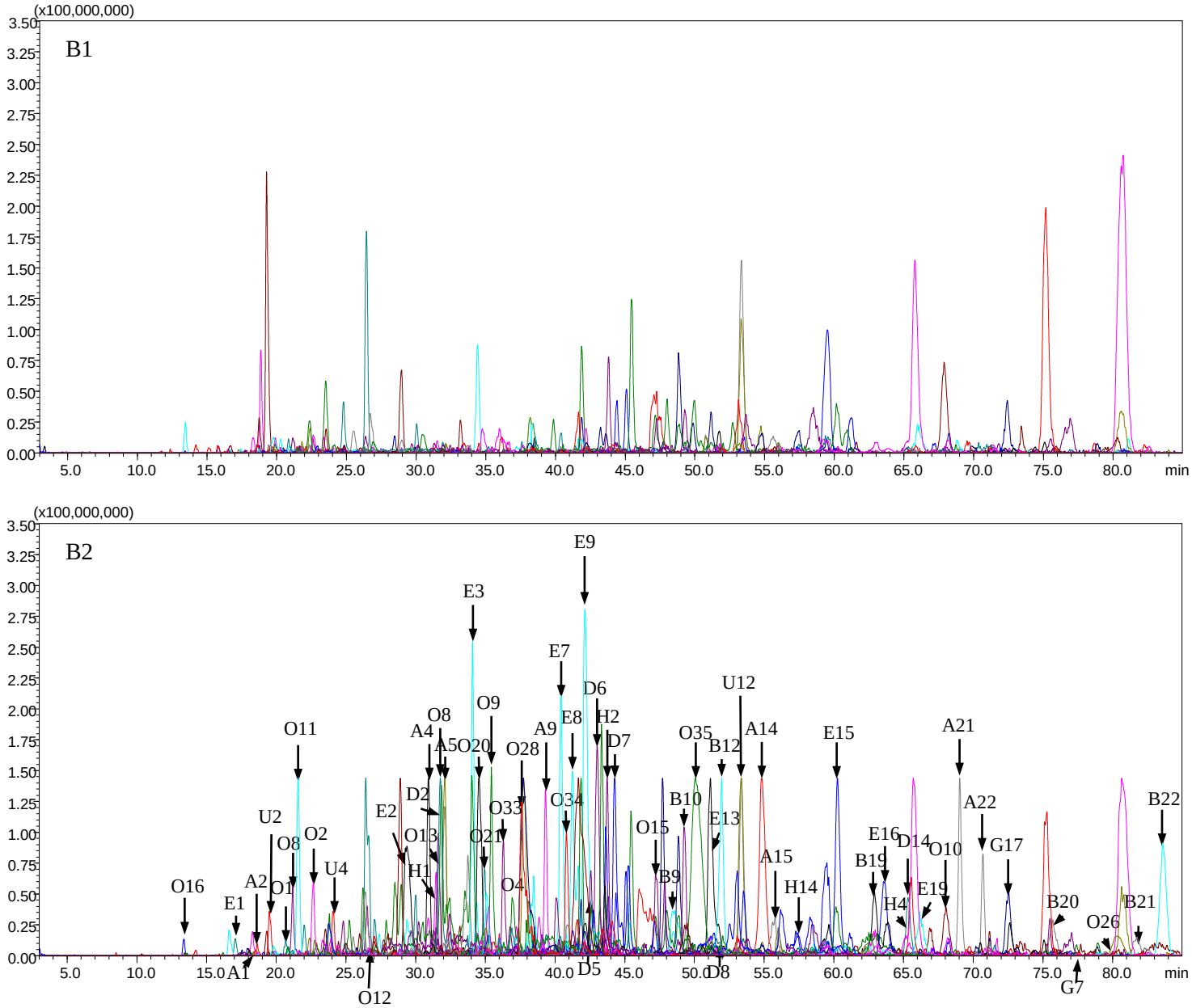


Fig. 3B

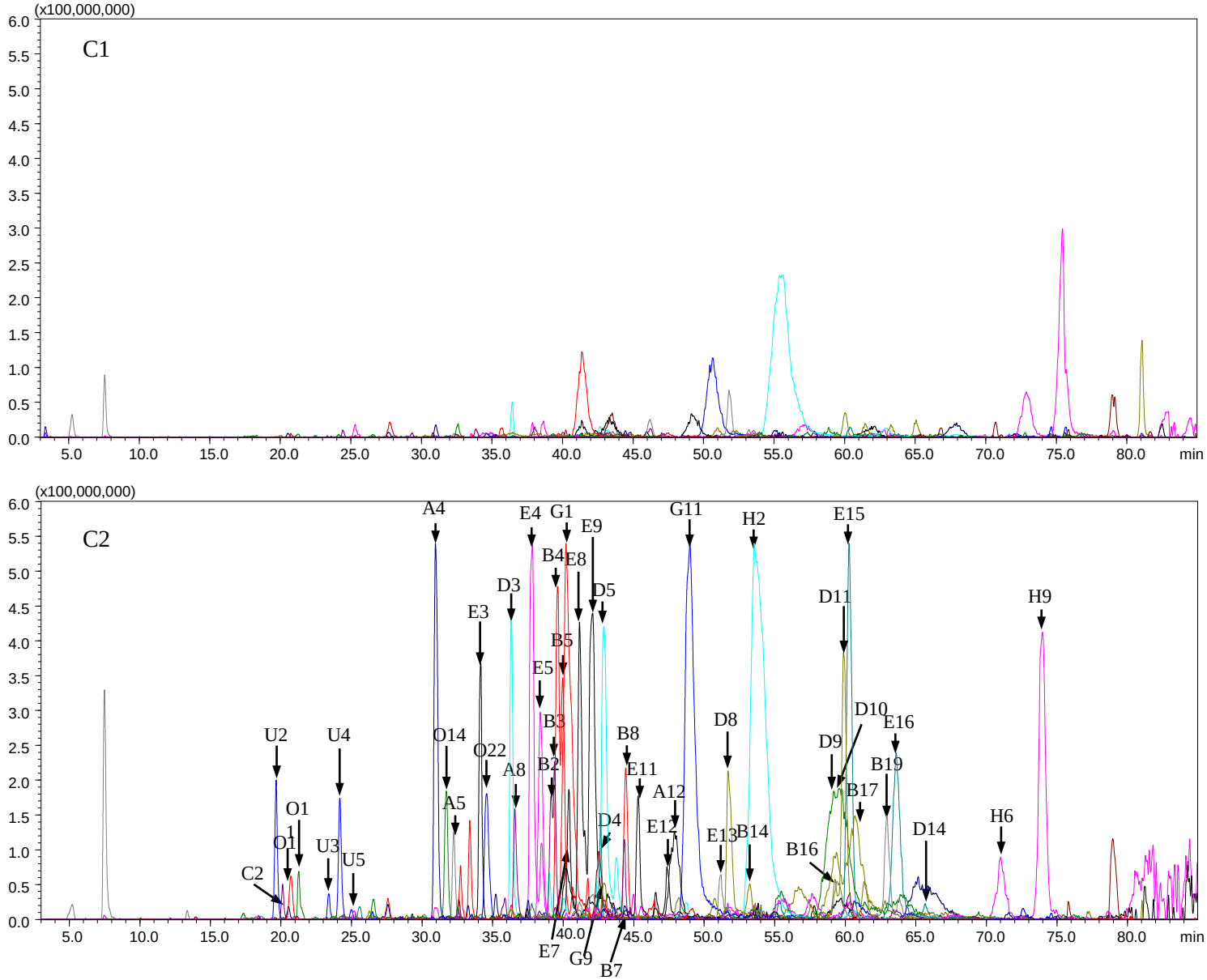


Fig. 3C

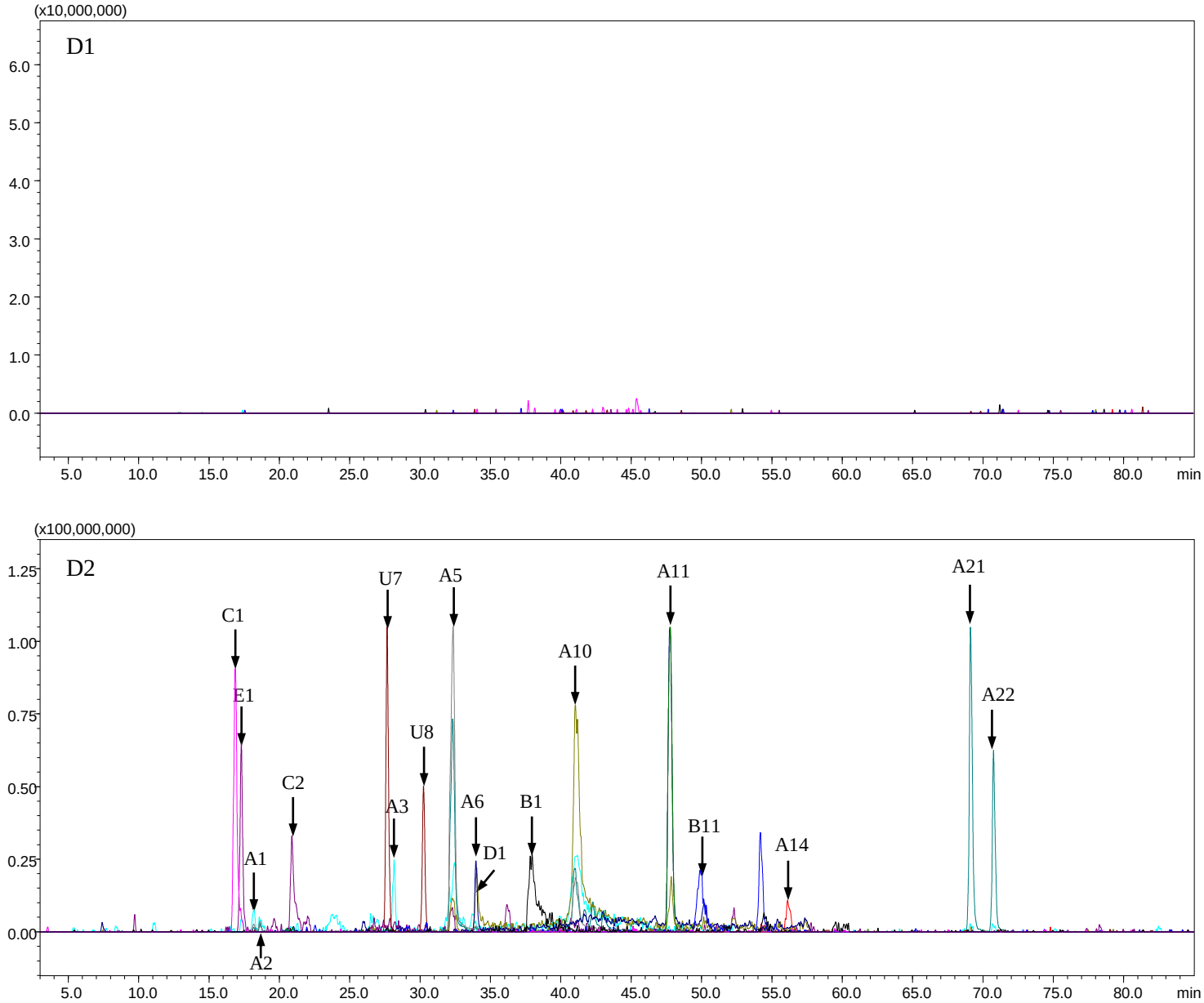


Fig. 3D

Table 1 The DFIs schedule for the detection and identification of the metabolites in rats after oral administration of ECT by E(DFI)MS^{n}Cs-based strategy

Metabolites types	DFIs [c]
Ⅰ: Chalcones [b]	
Ⅰa: Chalcones/ flavanonols (A)	**319.05** $[A–H]^–$; **303.05** $[A–H]^–$, 285.04 $[A–H–H_2O]^–$, 275.05 $[A–H–CO]^–$; **287.05** $[A–H]^–$, 259.06 $[A–H–CO]^–$; **271.06** $[A–H]^–$; **255.07** $[A–H]^–$; **239.07** $[A–H]^–$
Ⅰb: Dihydrochalcones (H)	**321.07** $[A–H]^–$; **305.07** $[A–H]^–$; **289.07** $[A–H]^–$; **273.08** $[A–H]^–$; **257.09** $[A–H]^–$; **241.09** $[A–H]^–$; **225.09** $[A–H]^–$
Ⅱ: Flavanones (B) [a]	**287.05** $[A–H]^–$, 269.05 $[A–H–H_2O]^–$, 259.06 $[A–H–CO]^–$, 193.01 $[M–H–B–ring]^–$, 181.01$[^{1,2}A]^–$, 167.00 $[^{1,3}A]^–$, 153.03 $[^{1,2}A–CO]^–$
Ⅲ: Flavonols	
Ⅲa: Kaempferols (D) [a]	**285.04** $[A–H]^–$, 267.02 $[A–H–H_2O]^–$, 257.04 $[A–H–CO]^–$, 239.04 $[A–H–H_2O–CO]^–$, 151.00 $[^{1,3}A]^–$
Ⅲb: 6-Hydroxykaempferols or Quercetins (E) [a]	**301.03** $[A–H]^–$, 283.02 $[A–H–H_2O]^–$, 255.03 $[A–H–H_2O–CO]^–$, 245.05$[A–H–CO–CO]^–$
Ⅲc: 6-Hydroxyquercetins (C) [a]	**317.03** $[A–H]^–$, 299.11 $[A–H–H_2O]^–$, 269.09 $[A–H–H_2O–CH_2O]^–$
Ⅲd: 6-Hydroxymyricetins (I) [b]	**333.02** $[A–H]^–$, 315.01 $[A–H–H_2O]^–$, 297.00 $[A–H–H_2O–H_2O]^–$
Ⅲe: Other flavonols (G) [b]	**269.04** $[A–H]^–$, 241.05 $[A–H–CO]^–$; **253.05** $[A–H]^–$, 235.04 $[A–H–H_2O]^–$, 225.05 $[A–H–CO]^–$; **237.05** $[A–H]^–$
Ⅳ: Other phenolic acid compounds (O) [b]	**185.01** $[A–H]^–$, 167.00 $[A–H–H_2O]^–$, 141.02 $[A–H–CO_2]^–$; **169.01** $[A–H]^–$, 151.00 $[A–H–H_2O]^–$, 125.02 $[A–H–CO_2]^–$; **167.03** $[A–H]^–$, 149.02 $[A–H–H_2O]^–$, 123.04 $[A–H–CO_2]^–$; **165.04** $[A–H]^–$, 147.02 $[A–H–H_2O]^–$; **163.04** $[A–H]^–$, 145.03 $[A–H–H_2O]^–$, 119.06 $[A–H–CO_2]^–$; **151.04** $[A–H]^–$, 107.05 $[A–H–CO_2]^–$; **141.02** $[A–H]^–$, 123.01 $[A–H–H_2O]^–$; **137.02** $[A–H]^–$, 93.03 $[A–H–CO_2]^–$; **125.02** $[A–H]^–$

[a]: DFIs were mainly determined by chemical profiling of ECT; [b]: DFIs were determined by metabolic profiling of the representative compounds; [c]: the DFIs in bold were adopted as the superior fragment

ions for the rapid screen the ECT-related compounds *in vivo*, whereas the others were mainly used for the structure identification; and the ion nomenclatures were elucidated in Fig. S1.

Table 2 Identification of the metabolites in rats after oral administration of ECT

Meta	t_R (min)	Formula	$[M+H]^+/[M-H]^-$				Distribution*				Type#
			mode	Meas. (Da)	Calcd. (Da)	Error (ppm)	P	U	B	F	
Chalcones and flavanonols											
A1[b]	18.038	$C_{27}H_{32}O_{16}$	Neg	611.1617	611.1617	0.00	-	+	-	+	P
A2[b]	18.462	$C_{27}H_{32}O_{16}$	Neg	611.1616	611.1617	-0.82	-	+	-	+	P
A3[b]	28.163	$C_{27}H_{32}O_{17}$	Neg	627.1565	627.1567	-0.32	-	-	-	+	P
A4	30.850	$C_{21}H_{20}O_8$	Neg	399.1088	399.1085	0.75	+	+	+	-	G
A5[a,b]	32.380	$C_{27}H_{32}O_{16}$	Neg	611.1605	611.1617	-1.96	+	+	+	+	P
A6[b]	34.028	$C_{27}H_{32}O_{16}$	Neg	611.1625	611.1617	1.31	-	-	-	+	P
A7[c]	34.348	$C_{21}H_{20}O_{12}$	Neg	463.0886	463.0882	0.86	+	-	-	-	G
A8	36.448	$C_{22}H_{22}O_{13}$	Neg	493.0983	493.0987	-0.81	-	-	+	-	G+M
A9	39.225	$C_{21}H_{20}O_{13}$	Neg	479.0834	479.0831	0.63	-	+	-	-	G
A10[b]	40.907	$C_{27}H_{30}O_{15}$	Pos	595.1659	595.1658	0.17	-	-	-	+	P
A11[b]	47.698	$C_{27}H_{31}NO_{14}$	Neg	592.1677	592.1672	0.84	-	-	-	+	P
A12	48.017	$C_{21}H_{20}O_{13}S$	Neg	511.0552	511.0552	0.00	+	-	-	-	G+S
A13[e]	48.022	$C_{21}H_{20}O_{10}$	Neg	431.0986	431.0983	0.70	+	-	-	-	G
A14	55.105	$C_{16}H_{14}O_{10}S$	Neg	397.0235	397.0235	0.00	-	-	-	-	P
A15[c]	55.435	$C_{15}H_{12}O_9S$	Neg	367.0130	367.0129	0.27	-	+	-	-	S
A16[b]	56.120	$C_{48}H_{52}O_{26}$	Neg	1043.2683	1043.2674	0.86	-	-	-	+	P
A17	58.640	$C_{21}H_{20}O_{11}$	Neg	447.0931	447.0933	-0.45	+	-	-	-	G
A18[e]	60.102	$C_{21}H_{20}O_{11}$	Neg	447.0933	447.0933	0.00	+	-	-	-	G
A19[e]	62.020	$C_{21}H_{20}O_{11}$	Neg	447.0930	447.0933	-0.67	+	-	-	-	G
A20	66.915	$C_{21}H_{20}O_{14}S$	Neg	527.0507	527.0501	1.14	+	-	-	-	G+S
A21[a]	69.027	$C_{30}H_{30}O_{14}$	Neg	613.1556	613.1563	-1.14	-	+	-	+	P
A22[a]	70.627	$C_{30}H_{30}O_{14}$	Neg	613.1572	613.1563	1.47	-	+	-	+	P
Carthamidins/isocarthamidins											
B1[b]	37.820	$C_{21}H_{22}O_{11}$	Neg	449.1084	449.1089	-1.11	-	-	-	+	P
B2	39.035	$C_{27}H_{30}O_{17}$	Neg	625.1410	625.1410	0.00	-	-	+	-	G
B3	39.307	$C_{27}H_{30}O_{17}$	Neg	625.1408	625.1410	-0.32	-	-	+	-	G
B4	39.448	$C_{27}H_{28}O_{18}$	Neg	639.1200	639.1203	-0.47	-	-	+	-	G
B5	39.978	$C_{27}H_{28}O_{18}$	Neg	639.1204	639.1203	0.16	-	-	+	-	G
B6	40.795	$C_{27}H_{32}O_{16}$	Neg	611.1620	611.1617	0.49	-	+	-	-	P
B7	44.242	$C_{27}H_{30}O_{17}$	Neg	625.1410	625.1410	0.00	-	-	+	-	G

B8	44.367	$C_{27}H_{28}O_{18}$	Neg	639.1208	639.1203	0.78	-	-	+	-	G
B9	48.467	$C_{21}H_{20}O_{12}$	Neg	463.0884	463.0882	0.43	-	+	-	-	G
B10	49.357	$C_{15}H_{12}O_6$	Neg	287.0568	287.0561	2.44	-	+	-	-	H
B11	50.032	$C_{22}H_{22}O_{12}$	Neg	477.1039	477.1038	0.21	-	-	-	+	G+M
B12	51.900	$C_{21}H_{20}O_{12}$	Neg	463.0883	463.0882	0.22	+	+	-	-	G
B13	52.133	$C_{21}H_{20}O_{15}S$	Neg	543.0448	543.0450	-0.37	-	-	+	-	G+S
B14	52.625	$C_{21}H_{20}O_{15}S$	Neg	543.0449	543.0450	-0.18	-	-	+	-	G+S
B15	58.837	$C_{21}H_{20}O_{12}$	Neg	463.0883	463.0882	0.22	+	-	-	-	G
B16	58.845	$C_{22}H_{22}O_{15}S$	Neg	557.0603	557.0606	-0.54	-	-	+	-	G+M+S
B17	59.108	$C_{22}H_{22}O_{12}$	Neg	477.1021	477.1038	-3.56	+	+	-	-	G+M
B18	62.053	$C_{22}H_{22}O_{15}S$	Neg	557.0603	557.0606	-0.54	-	+	-	-	G+M+S
B19	62.823	$C_{22}H_{22}O_{12}$	Neg	477.1038	477.1038	0.00	-	+	+	-	G+M
B20	75.695	$C_{15}H_{12}O_9S$	Neg	367.0122	367.0129	-1.91	-	+	-	-	S
B21	81.830	$C_{15}H_{12}O_9S$	Neg	367.0120	367.0129	-2.45	-	+	-	-	S
B22	83.280	$C_{16}H_{14}O_9S$	Neg	381.0288	381.0286	0.52	-	+	-	-	S+M
6-Hydroxyquercetins											
C1	16.997	$C_{21}H_{20}O_{12}$	Neg	463.0886	463.0882	0.86	-	-	-	+	P
C2	20.063	$C_{21}H_{18}O_{14}$	Neg	493.0642	493.0624	3.65	-	-	+	-	G
Kaempferols											
D1[b]	33.853	$C_{27}H_{30}O_{15}$	Neg	593.1504	593.1512	-1.35	-	-	-	+	P
D2	34.367	$C_{27}H_{26}O_{18}$	Neg	637.1043	637.1046	-0.47	-	+	-	-	G
D3[c]	36.242	$C_{27}H_{26}O_{18}$	Neg	637.1053	637.1046	1.10	+	-	+	-	G
D4	42.597	$C_{27}H_{28}O_{17}$	Neg	623.1257	623.1254	0.48	-	-	+	-	G
D5[c]	42.975	$C_{27}H_{26}O_{18}$	Neg	637.1048	637.1046	0.31	+	+	+	-	G
D6	43.598	$C_{27}H_{26}O_{18}$	Neg	637.1041	637.1046	-0.78	+	+	-	-	G
D7[c]	44.005	$C_{27}H_{26}O_{18}$	Neg	637.1051	637.1046	0.78	+	+	-	-	G
D8	51.505	$C_{21}H_{18}O_{12}$	Neg	461.0725	461.0725	0.00	-	+	+	-	G
D9	58.845	$C_{21}H_{18}O_{15}S$	Neg	541.0294	541.0293	0.18	-	-	+	-	G+S
D10	59.582	$C_{21}H_{18}O_{15}S$	Neg	541.0291	541.0293	-0.18	-	-	+	-	G+S
D11[c]	59.707	$C_{21}H_{18}O_{12}$	Neg	461.0725	461.0725	0.00	+	-	+	-	G
D12[c]	61.635	$C_{21}H_{18}O_{12}$	Neg	461.0732	461.0725	1.52	+	-	-	-	G
D13[c]	63.307	$C_{21}H_{18}O_{15}S$	Neg	541.0291	541.0293	-0.37	+	-	+	-	G+S
D14	64.753	$C_{22}H_{20}O_{15}S$	Neg	555.0451	555.0450	0.18	-	-	+	-	G+S+M
D15	65.437	$C_{22}H_{20}O_{12}$	Neg	475.0884	475.0882	0.42	-	+	-	-	G+M
D16	67.657	$C_{21}H_{18}O_{12}$	Neg	461.0732	461.0725	1.52	+	-	-	-	G

D17	78.613	$C_{22}H_{20}O_{12}$	Neg	475.0888	475.0882	1.26	+	-	-	-	G+M
6-Hydroxykaempferols/Quercetins											
E1	17.110	$C_{27}H_{30}O_{17}$	Neg	625.1413	625.1410	0.48	+	+	-	+	P
E2[d]	29.662	$C_{33}H_{34}O_{25}$	Neg	829.1323	829.1316	0.84	-	-	+	-	G
E3	34.005	$C_{27}H_{26}O_{19}$	Neg	653.1001	653.0995	0.92	-	+	+	-	G
E4	37.628	$C_{28}H_{28}O_{19}$	Neg	667.1145	667.1152	-1.05	-	-	+	-	G+M
E5[d]	38.580	$C_{28}H_{28}O_{19}$	Neg	667.1153	667.1152	0.15	-	-	+	-	G+M
E6	40.212	$C_{27}H_{28}O_{18}$	Neg	639.1194	639.1203	-1.41	+	-	-	-	G
E7[d]	40.442	$C_{27}H_{26}O_{19}$	Neg	653.0994	653.0995	-0.15	+	+	+	-	G
E8[d]	41.082	$C_{27}H_{26}O_{19}$	Neg	653.0995	653.0995	0.00	-	+	+	-	G
E9[d]	41.612	$C_{27}H_{26}O_{19}$	Neg	653.0996	653.0995	0.15	-	+	+	-	G
E10[d]	42.060	$C_{27}H_{26}O_{19}$	Neg	653.0994	653.0995	-0.15	-	+	-	-	G
E11	45.318	$C_{27}H_{26}O_{19}$	Neg	653.0989	653.0995	-0.92	+	+	+	-	G
E12	47.340	$C_{27}H_{26}O_{19}$	Neg	653.0987	653.0995	-1.22	-	-	+	-	G
E13[d]	51.203	$C_{22}H_{22}O_{12}$	Neg	477.0649	477.0674	-5.24	-	+	+	-	G
E14	55.315	$C_{22}H_{20}O_{13}$	Neg	491.0836	491.0831	1.02	-	+	-	-	G+M
E15	60.060	$C_{22}H_{20}O_{13}$	Neg	491.0831	491.0831	0.00	-	+	+	-	G+M
E16[d]	62.253	$C_{22}H_{20}O_{13}$	Neg	491.0829	491.0831	-0.41	+	+	+	-	G+M
E17	63.697	$C_{22}H_{22}O_{12}$	Neg	477.0672	477.0674	-0.42	-	+	-	-	G
E18	64.535	$C_{22}H_{20}O_{16}S$	Neg	571.0396	571.0399	-0.53	+	-	-	-	G+S
E19[d]	66.115	$C_{21}H_{18}O_{13}$	Neg	477.0699	477.0674	5.24	-	+	-	-	G
Other flavonols											
G1	40.722	$C_{21}H_{18}O_{13}S$	Neg	509.0394	509.0395	-0.20	-	-	+	-	G+S
G2	41.765	$C_{21}H_{18}O_{10}$	Pos	431.0956	431.0973	-3.94	+	+	-	-	G
G3	43.157	$C_{21}H_{18}O_{13}S$	Neg	509.0392	509.0395	-0.58	+	-	-	-	G+S
G4	71.392	$C_{15}H_{10}O_{7}S$	Neg	333.0071	333.0074	-0.90	+	-	-	-	G
G5	78.873	$C_{21}H_{18}O_{10}$	Neg	429.0833	429.0827	1.40	-	+	-	-	G
G6	42.385	$C_{22}H_{20}O_{14}S$	Neg	539.0505	539.0501	0.74	-	-	+	-	G+S+M
G7	43.827	$C_{22}H_{20}O_{11}$	Pos	461.1079	461.1079	0.00	-	+	-	-	G+M
G8	49.222	$C_{21}H_{18}O_{14}S$	Neg	525.0343	525.0344	0.19	-	-	+	-	G+S
G9	52.408	$C_{21}H_{18}O_{14}S$	Neg	525.0343	525.0344	-0.19	+	-	-	-	G+M+S
G10	53.384	$C_{21}H_{18}O_{11}$	Neg	525.0343	525.0343	0.00	+	-	-	-	G+M+S
G11	55.817	$C_{21}H_{18}O_{11}$	Pos	447.0916	447.0922	-1.34	-	+	-	-	G
G12	56.152	$C_{21}H_{18}O_{11}$	Neg	525.0343	525.0343	0.00	+	-	-	-	G+M+S
G13	62.278	$C_{21}H_{18}O_{11}$	Neg	445.0773	445.0776	-0.67	-	+	-	-	G

G14	72.393	$C_{16}H_{12}O_8S$	Neg	363.0179	363.0180	-0.28	-	+	-	-	S
H1	31.538	$C_{21}H_{22}O_8$	Neg	401.1242	401.1231	2.74	-	+	-	-	G
H2	54.687	$C_{21}H_{22}O_{12}S$	Neg	497.0758	497.0759	-0.20	-	-	+	-	G+S
H3	58.415	$C_{21}H_{22}O_{12}S$	Neg	497.0762	497.0759	0.60	+	-	-	-	G+S
H4	65.322	$C_{21}H_{22}O_{11}$	Neg	449.1087	449.1089	-0.45	-	+	-	-	G
H5	70.435	$C_{21}H_{22}O_{11}$	Neg	449.1089	449.1089	0.00	+	-	-	-	G
H6	70.908	$C_{21}H_{22}O_{13}S$	Neg	513.0710	513.0708	3.89	-	-	+	-	G+S
H7	71.960	$C_{21}H_{22}O_{10}$	Neg	433.1126	433.1140	-0.92	+	-	-	-	G
H8	72.445	$C_{21}H_{22}O_{10}$	Neg	433.1133	433.1140	-1.62	+	-	-	-	G
H9	74.168	$C_{21}H_{22}O_{13}S$	Neg	513.0709	513.0708	0.19	-	-	+	-	G+S
H10	78.283	$C_{21}H_{22}O_{13}S$	Neg	513.0711	513.0708	0.58	+	-	-	-	G+S
H11	83.343	$C_{21}H_{22}O_{15}S$	Neg	545.0602	545.0606	-0.73	-	+	-	-	G
Other phenolic acid compounds											
O1	20.355	$C_8H_8O_6S$	Neg	230.9967	230.9969	-0.85	+	+	+	-	S
O2	22.667	$C_{14}H_{16}O_{10}$	Neg	343.0667	343.0670	-0.87	+	+	-	-	G
O3	23.002	$C_9H_{10}O_4$	Neg	181.0500	181.0506	-3.31	+	-	-	-	H
O4	37.110	$C_{14}H_{16}O_{10}$	Neg	343.0674	343.0670	1.17	-	+	-	-	G+M
O5	39.882	$C_9H_{10}O_3$	Neg	165.0555	165.0557	-1.21	+	-	-	-	H
O6	49.665	$C_8H_8O_6S$	Neg	230.9967	230.9969	-0.87	+	-	-	-	S
O7	23.368	$C_{13}H_{14}O_{12}$	Neg	361.0410	361.0412	-0.55	-	-	+	-	G
O8	31.947	$C_{13}H_{14}O_{12}$	Neg	361.0427	361.0412	4.15	-	+	-	-	G
O9	35.433	$C_{13}H_{14}O_{12}$	Neg	361.0387	361.0412	-6.92	-	+	-	-	G
O10	68.097	$C_8H_8O_6$	Neg	199.0243	199.0248	-2.51	-	+	-	-	M
O11	21.087	$C_{15}H_{18}O_{11}$	Neg	373.0781	373.0776	1.34	-	+	+	-	G
O12	26.322	$C_8H_8O_8S$	Neg	262.9867	262.9867	0.00	-	+	-	-	S+M
O13	31.585	$C_9H_{10}O_8S$	Neg	277.0024	277.0023	0.36	-	+	-	-	S+M
O14	31.598	$C_{15}H_{18}O_{11}S$	Neg	373.0769	373.0776	-1.88	-	-	+	-	G
O15	47.453	$C_9H_{10}O_8S$	Neg	277.0013	277.0023	-3.61	-	+	-	-	S+M
O16	14.627	$C_{13}H_{16}O_8$	Neg	299.0772	299.0767	1.67	+	-	-	-	G
O17	27.912	$C_8H_8O_7S$	Neg	246.9918	246.9918	0.00	+	-	-	-	S+M
O18	21.940	$C_{16}H_{20}O_{10}$	Neg	371.0983	371.0986	-0.80	-	+	-	-	G+M
O19	28.478	$C_{15}H_{18}O_8$	Neg	325.0927	325.0929	-0.62	+	-	-	-	G
O20	34.342	$C_{15}H_{18}O_9$	Neg	341.0878	341.0878	0.00	-	+	-	-	G
O21	34.912	$C_{16}H_{20}O_{10}$	Neg	371.0985	371.0986	-0.54	+	+	-	-	G
O22[e]	34.915	$C_9H_{10}O_6S$	Neg	245.0129	245.0125	1.63	-	+	-	-	S

O23	35.660	$C_9H_8O_6S$	Neg	245.0123	245.0125	-0.82	+	-	-	-	S
O24	37.050	$C_{15}H_{18}O_9$	Neg	341.0885	341.0878	2.05	-	+	-	-	G
O25	38.160	$C_9H_{10}O_6S$	Neg	245.0128	245.0125	1.22	-	+	-	-	S
O26	80.512	$C_9H_{10}O_6S$	Neg	245.0123	245.0125	-0.82	-	+	-	-	S
O27	28.298	$C_9H_{12}O_7S$	Neg	263.0231	263.0230	0.38	+	-	-	-	S+M
O28	37.653	$C_9H_8O_3$	Neg	163.0401	163.0400	0.61	+	+	-	-	H
O29[e]	38.163	$C_9H_8O_6S$	Neg	242.9967	242.9969	-0.82	+	+	+	-	S
O30	39.987	$C_9H_8O_6S$	Neg	242.9967	242.9969	-0.82	+	-	-	-	S
O31	40.262	$C_9H_8O_6S$	Neg	242.9966	242.9969	-1.23	+	-	-	-	S
O32	23.650	$C_6H_6O_6S$	Neg	204.9812	204.9812	0.00	-	+	-	-	S
O33	35.707	$C_{14}H_{18}O_9$	Neg	329.0877	329.0878	-0.30	-	+	-	-	G+M
O34	40.675	$C_{14}H_{18}O_9$	Neg	329.0875	329.0878	-0.91	-	+	-	-	G+M
O35	49.237	$C_7H_8O_6S$	Neg	218.9970	218.9969	0.46	-	+	-	-	S+M
O36	49.657	$C_8H_{10}O_6S$	Neg	233.0124	233.0125	-0.43	-	+	-	-	S+M
Unknown compounds											
U1	17.972	$C_{21}H_{24}O_{10}$	Neg	435.1423	435.1417	1.38	+	-	-	-	G
U2	19.493	$C_{28}H_{30}O_{21}$	Neg	701.1198	701.1201	-0.71	-	-	+	-	G
U3	23.445	$C_{28}H_{30}O_{21}$	Neg	701.1203	701.1201	0.27	-	-	+	-	G
U4	24.028	$C_{28}H_{30}O_{21}$	Neg	701.1201	701.1201	0.00	+	-	+	-	G
U5	24.960	$C_{28}H_{30}O_{21}$	Neg	701.1207	701.1201	0.86	+	+	+	-	G
U6	26.363	$C_{28}H_{30}O_{21}$	Neg	701.1209	701.1201	1.14	+	-	-	-	G
U7	27.538	$C_{22}H_{26}O_{12}$	Pos	483.1507	483.1497	2.07	-	-	-	+	P
U8	30.180	$C_{22}H_{26}O_{12}$	Pos	483.1493	483.1497	-0.83	-	-	-	+	P
U9	30.513	$C_{22}H_{22}O_{14}$	Neg	509.0937	509.0937	0.00	+	-	-	-	G
U10	31.818	$C_{22}H_{22}O_{14}$	Neg	509.0934	509.0937	-0.59	+	-	-	-	G
U11	36.227	$C_{22}H_{22}O_{15}$	Neg	525.0851	525.0880	-5.52	+	-	-	-	G
U12	53.570	$C_{21}H_{24}O_9$	Neg	419.1367	419.1347	4.77	-	+	-	-	G
U13	68.383	$C_{21}H_{24}O_{13}$	Neg	483.1134	483.1144	-2.07	+	-	-	-	G

[a]: Identified by comparison with the reference compounds; [b]: Identified by comparison with ECT; [c]: Identified by comparison with the metabolites in rats after oral administration of kaempferol-3-*O*-rutinoside; [d]: Identified by comparison with the metabolites in rats after oral administration of 6-hydroxykaempferol-3-*O*-rutinoside; [e]: Identified by comparison with the metabolites in rats after oral administration of HSYA; [*]: Distribution in: P, plasma; U, urine; F, feces; B, bile; +, detected in screening; −, not detected in screening; [#]: Metabolic type: M, methylation; O, oxidation; H, hydrolysis; S, sulfonation; G, glucuronidation; P, prototype.