NJC

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal | View Issue

Cite this: *New J. Chem.*, 2022, **46**, 18426

Characterization of chemical components and the potential anti-influenza mechanism of Fructus Arctii by a strategy integrating pharmacological evaluations, chemical profiling, serum pharmacochemistry, and network pharmacology†

Min Li,‡[a] Zhong-Yu Huang,‡[c] Yu-Lin-Lan Yuan,[d] Shuang-Shuang Cui,[d] Hai-Jun Li*[ae] and Feng-Xiang Zhang*[bd]

The presence of chemical components in traditional Chinese medicine (TCM) is the functional basis for its therapy achievement, and the absorbed components under disease conditions mainly contribute to the therapeutic effects. Fructus Arctii (FA) was widely used in traditional Chinese medicine preparation (TCMP) to clear heat or treat influenza. To date, the chemical profile (*in vitro* and *in vivo*) and anti-influenza mechanism of FA are still unrevealed, limiting its clinical applications. In this work, an integrated strategy combining pharmacological evaluations, chemical profiling, serum pharmacochemistry, and network pharmacology (PCSN strategy) is proposed and applied to investigate chemical components (*in vitro* and *in vivo*), anti-influenza effects, and potential mechanisms of FA. First, the anti-influenza effects of FA were evaluated using an H1N1 virus-infected mouse model. Second, the chemical compounds of FA *in vitro* and in the disease model were profiled and characterized using liquid chromatography coupled with mass spectrometry. Lastly, the targets and potential mechanisms are predicted based on the absorbed components and a network pharmacology method. As a result, FA (20 g crude drug kg^{-1} day^{-1}) could significantly improve the survival rate and ameliorate lung inflammation caused by the H1N1 virus. Moreover, a total of 123 compounds (53 potential novel compounds) were identified or tentatively characterized in FA. Furthermore, 124 xenobiotics (38 prototypes and 86 metabolites) were tentatively characterized in the infected mice. The xenobiotics were unitized to predict targets and mechanisms in treating influenza. It was found that FA could target EGFR, CASP3, MAPK1, PTGS2, PIK3CA, ESR1, *etc.*, which are mainly involved in the PI3K-Akt signaling pathway, progesterone-mediated oocyte maturation, *etc.* Among them, the PI3K-Akt signaling pathway (correlated to PI3K complex) and TNF signalling pathway, are regarded as the core pathways for FA in the treatment of influenza. It is the first time that chemical information is obtained on FA *in vitro* and *in vivo*, and the anti-influenza effects and the potential mechanism of FA were also evaluated. Moreover, the potential novel compounds for further validation of FA are also provided.

Received 7th June 2022,
Accepted 24th August 2022

DOI: 10.1039/d2nj02799b

rsc.li/njc

[a] *Key Laboratory of Hainan Trauma and Disaster Rescue, The First Affiliated Hospital of Hainan Medical University, Emergency and Trauma College, Hainan Medical University, Haikou 571199, China. E-mail: liliankl@163.com; Tel: +860898-66962607*

[b] *State Key Laboratory for Chemistry and Molecular Engineering of Medicinal Resources, Collaborative Innovation Center for Guangxi Ethnic Medicine, School of Chemistry and Pharmaceutical Sciences, Guangxi Normal University, Guilin 541004, China. E-mail: zhangfengxiangjnu@163.com*

[c] *Laboratory for Neuroscience in Health and Disease, Guangzhou First People's Hospital School of Medicine, South China University of Technology, Guangzhou 510013, China*

[d] *Department of gynaecology and obstetrics, The First Affiliated Hospital of Jinan University, Jinan University, Guangzhou 510632, China*

[e] *General Medicine and Continuous Education College, Hainan Medical University, Haikou 571199, China*

† Electronic supplementary information (ESI) available. See DOI: **https://doi.org/10.1039/d2nj02799b**

‡ The author contributes equally to this work.

1. Introduction

The efficacy of traditional Chinese medicines (TCMs) has been confirmed by thousands of years of history as well as by modern clinical practice. Generally, the therapeutic effects of TCM were mainly achieved by the interaction of the absorbed components and targets in the host.[1] Further, the characterization of chemical information of TCM *in vitro* and *in vivo* is an effective way to reveal functional mechanisms, which is essential to clinical applications. Currently, the methods of pharmacological evaluation,[2] chemical profiling,[3] network pharmacology[4,5] *etc.* are widely used to explore the functional mechanism of TCM. However, there are still some limitations on using only one method to reveal the functional mechanism. Moreover, the chemical compounds of TCM in the disease model and health host are quite different.[6] Thus, the chemical information in a disease model under therapeutic dosage is the basis to explore the potential pharmacological mechanism of TCM. Moreover, the pharmacology method, with the ability to systematically reveal the pharmacological index and functional basis of the complex systems, has been widely used in traditional Chinese medicine to explore pharmacological mechanisms, such as the Lian-Hua-Qing-Wen capsule.[7] In consideration of the above condition, an integrated strategy by a combination of pharmacological evaluations, chemical profiling, serum pharmacochemistry, and network pharmacology (PCSN strategy) is proposed in our work and applied to characterize the functional mechanism of TCM, using Fructus Arctii as an example.

Fructus Arctii (FA), the dried ripe fruit of *Arctium lappa* L. (also called burdock), is widely used in Asian countries (China, Japan, and Korea, *etc*) due to its definite therapeutic effects in clinical or daily life, such as for heat removal, detoxification, and elimination of swelling.[8] *Arctium lappa* L. has always been used as a vegetable,[9] and its roots[10] are used as a tea in China or other Asian countries. Meanwhile, its seeds (Fructus Arctii) are treated as food homology Chinese medicine and a healthy food supplemental product recognized by the National Health Commission of the People's Republic of China. Previous pharmacological studies have shown that FA could present the inhibition of streptozotocin-induced diabetic retinopathy[8] and depression of alpha-glucosidase activity,[11] *etc.* In the Chinese Pharmacopeia 2015 edition, FA has been reported for use as a core herb in TCM preparations such as Yin Qiao powder[12] and Suhuang Zhike Capsules to treat influenzas.[13] Meanwhile, the anti-flu effects of some representative components in FA also had been reported, such as arctiin and arctigenin.[14] However, the anti-flu effects of FA, as a whole, have not been reported.

Until now, more than 200 compounds have been isolated for FA, and they are mainly classified into five chemical structure types, including caffeoylquinic acids,[15] lignans,[16–21] steroids,[16] essential oils,[22] and other types. Moreover, the chemical constituents in Fructus Arctii were also profiled and characterized by liquid chromatography coupled with mass spectrometry (LC-MS) or gas chromatography coupled with mass spectrometry (GC-MS) methods. In detail, Qin *et al.* characterized 13 constituents (caffeoylquinic acids and lignans) in FA[23] and Liu *et al.* profiled 12 lignans in FA.[24] The method to determine chlorogenic acid, isochlorogenic acid, arctiin and arctigenin was developed and applied to evaluate the quality of Fructus Arctii.[25] In the meantime, the metabolic fate of some representative compounds such as arctiin was also summarized.[26–28] It was noted that there were still no methods for comprehensively characterizing the chemical constituents of FA *in vitro* and *in vivo*, especially in a disease model.

In this work, a strategy integrating pharmacological evaluations, chemical profiling, serum pharmacochemistry, and network pharmacology (PCSN strategy) is proposed and applied to reveal the anti-flu effects and potential pharmacological mechanisms of FA (Fig. 1). It contained three parts. First, the anti-influenza effects of FA were evaluated using an H1N1 virus-infected mouse model. Second, the chemical compounds in FA were profiled, and subsequently, the xenobiotics in H1N1 infected mice after ingestion of the effective dosage of FA were characterized. Lastly, the potential targets and mechanisms are predicted by a network pharmacology method.

2. Materials and methods

2.1 Materials

Fructus Arctii was purchased from the market and further identified by Dr Chang Li at Harbin Medical University. 5-*O*-Caffeoylquinic acid (**2**), chlorogenic acid (**4**), 4-*O*-caffeoylquinic acid (**6**), 3,4-*O*-dicaffeoylquinic acid (**26**), 3,5-*O*-dicaffeoylquinic acid (**28**), and 4,5-*O*-dicaffeoylquinic acid (**32**), matairesinoside (**38**), arctiin (**54**) matairesinol (**86**), and arctigenin (**102**) were obtained from Chengdu Push Bio-Technology Co., Ltd (Chengdu, China), and their purity was more than 98%, HPLC grade. Water, methanol, and ethanol were all HPLC grade. LC-MS grade acetonitrile and water were obtained from Fisher Scientific (Fair Lawn, New Jersey, USA). LC-MS grade formic acid was purchased from Sigma-Aldrich (St. Louis, MO, USA). The influenza virus A/FM/1/47 (H1N1 2009) was donated by Prof. Jian-Xin Chen at South China Agriculture University. All experiments with the H1N1 virus were conducted in a Biosafety Level 3 (BSL-3) containment laboratory approved by the Ministry of Agriculture of China.

2.2 Preparation of Fructus Arctii extract

Fructus Arctii was weighed and crushed to powder. Then, the powder was extracted with boiling water using the reflux extraction method three times (2 h each). The solution was filtered through gauze and condensed at a final concentration of 1 g crude drug mL^{-1} using a rotary evaporator under 40 °C. Then, the condensed solution was chilled at room temperature and stored at 4 °C before the experiment.

2.3 Animal feeding conditions and pharmacological experiment design

2.3.1 Mouse feeding condition. Specific pathogen-free (SPF) male adult mice (15–18 g) were obtained from Jinan Pengyue

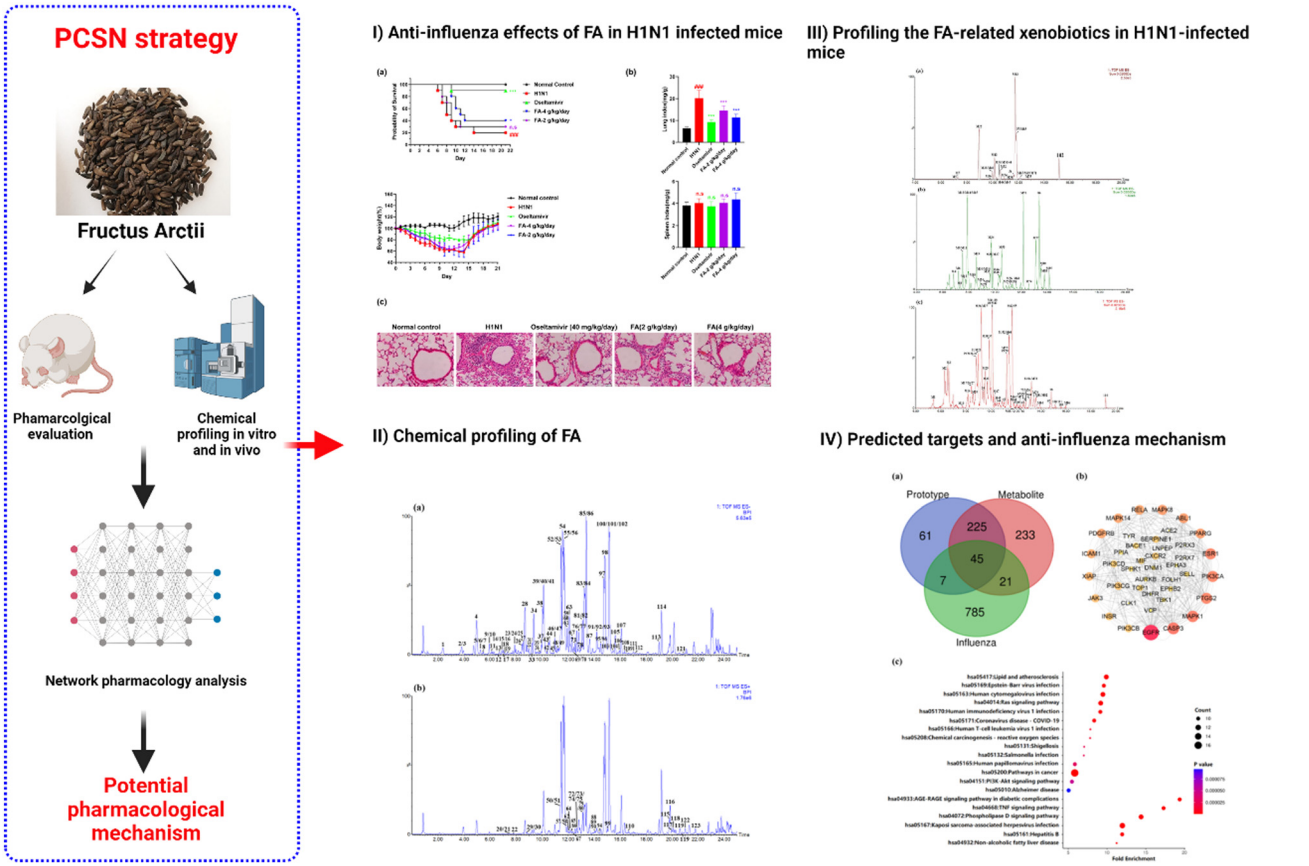

Fig. 1 The flowchart for revealing the potential anti-influenza mechanism of Fructus Arctii by a strategy integrating pharmacological evaluations, chemical profiling, serum pharmacochemistry, and network pharmacology.

Experimental Animal Breeding Co. Ltd (Jinan, China). The mice were housed at an ambient temperature of 20 ± 2 °C with 12 h light/dark cycles for one week, and in the meantime, a standard diet and water were made freely available to mice. The animal protocols were approved by the Guide for the Care and Use of Laboratory Animals of Jinan University, and procedures were following Guide for the Care and Use of Laboratory Animals (National Institutes of Health).

2.3.2 Evaluation of anti-influenza effects of FA. Influenza mouse model: before the experiment, the mice were adapted to the lab for one week. Then, mice were anesthetized with isoflurane (RWD life science, China) and intranasally infected with the 10-fold median lethal dose (LD50) H1N1 virus to establish the influenza mouse model. During the experimental period, mice were provided with free access to water and food.

2.3.2.1 Survival rate and body weight. Fifty BALB/c mice were selected and randomized into five groups, including normal control groups (NC), water-treated group (Placebo), FA groups (5 and 20 g crude drug kg^{-1}), and oseltamivir-treated groups (40 mg kg^{-1} day^{-1}) (n = 10 group). Before the viral infection, the mice were pro-ingested with drugs or water for three days, and treatment lasted for five days after viral infections. During this time, survival rate and body weight changes were observed and recorded daily until 21 days.

2.3.2.2 Lung histology. With the same design mentioned above, mice were sacrificed, and the lung and spleen were weighed to calculate the lung index (the ratio of lung weight to body weight) and spleen index (the ratio of spleen weight to body weight) on day 5 post-infection. Lungs were removed from euthanized mice and fixed in 4% formalin at room temperature for 48 h. Serial tissue sections 5 μm thick were obtained after embedding in paraffin. Each slide was stained with hematoxylin and eosin (H&E) and then examined under light microscopy (Olympus BX41, Olympus Optical Co., Tokyo, Japan). The lung score based on H&E was calculated based on the criterion reported in our previous work.[29]

2.4 Biological sample collection and pretreatment

The H1N1 infected mice (n = 10) were administered with FA (20 g crude drug kg^{-1} day^{-1}). During the period of drug administration, urine and feces were collected and stored at −80 °C before pre-treatment. On Day 5, mice were sacrificed at 0.5, 1, 2, 4, and 8 h after the last ingestion, and the blood samples were collected. Meanwhile, the organs (liver, heart, spleen, lung, kidneys, and brain) were obtained after collecting blood, and they were washed with normal saline (4 °C) until there was no blood. They were then stored at −80 °C before the pretreatment. The biosamples from mice were pretreated as follows.

2.4.1 Plasma. Blood samples were centrifugated at 13 000 rpm for 10 min (4 °C) to prepare the plasma. The plasma was treated with acetonitrile at a ratio of 1 : 4 to precipitate proteins. The supernatant was obtained after centrifugation at 13 000 rpm for 10 min (4 °C), and then, it was dried with nitrogen gas at room temperature. The residue was reconstituted in 300 μL 60% methanol–water (v/v), and further centrifugated at 13 000 rpm for 10 min (4 °C) to obtain supernatant for analysis.

2.4.2 Urine. For pre-treatment, urine-mixed samples (2 mL) were thawed and subsequently centrifuged at 13 000 rpm for 10 min (4 °C), and the supernatant was directly loaded on a pre-activity HLB column (6 cm^3, 200 mg, Waters Oasis, Ireland). Then, it was orderly eluted with 6 mL of 5% methanol and 6 mL of methanol. The methanol eluate was collected and dried using nitrogen gas at room temperature. The residue was reconstituted in 300 μL 60% methanol–water (v/v), and further centrifugated at 13 000 rpm for 10 min (4 °C) to obtain supernatant for analysis.

2.4.3 Feces. Then, samples (10 g) were extracted with 60% methanol–water (v/v) and centrifuged at 13 000 rpm for 10 min (4 °C), then the supernatant was dried with nitrogen and reconstituted using water. Then, the reconstituted samples were centrifuged by the same method, and 2 mL supernatants of them were loaded on a pre-activity HLB column (6 m^3, 200 mg, Waters Oasis, Ireland) directly, and then eluted with 6 mL of 5% methanol and 6 mL of methanol, successively. The methanol eluate was collected and dried under nitrogen gas at room temperature. The residue was reconstituted in 600 μL 60% methanol–water (v/v).

2.4.4 Heart, spleen, brain, lung, kidneys, and liver. Organs were thawed and weighed. Then, 2 g of tissues were homogenized by adding 2 mL of normal saline (4 °C) and subsequently treated with acetonitrile at a ratio of 1 : 4 to precipitate protein. After centrifugation at 13 000 rpm for 10 min (4 °C), the supernatant was obtained and dried with nitrogen gas at room temperature. The residue was reconstituted in 300 μL of 60% methanol–water (v/v), and further centrifuged at 13 000 rpm for 10 min (4 °C) to obtain the supernatant for analysis.

The above samples were detected using UHPLC/Q-TOF MS with an injection volume of 2 μL, and three replicates were conducted. The collected MS information was analyzed using Masslynx 4.1 (Waters, USA).

2.5 UHPLC/Q-TOF MS

Chromatographic separation was achieved using a Waters Acquity™ UHPLC system (Waters, Milford, USA) coupled with an Acquity UHPLC BEH C_{18} Column (2.1 × 100 mm, 1.8 μm, Waters, Milford, USA, held at 35 °C). Two solvent mobile phases systems consisting of eluent A (0.1% formic acid in water, v/v) and eluent B (0.1% formic acid in acetonitrile, v/v), were used for separation, and the flow rate was set as 0.4 mL min^{-1} by using a linear gradient program. Detailed information is listed as follows: 0–2.0 min, 5% B; 2.0–5.0 min, 5–15% B; 5.0–10.0 min, 15–25% B; 10.0–22.0 min, 25–60% B; 22.0–23.0 min, 60–100% B; 23.0–24.0 min, 100% B, 24.0–24.5 min, 100–5%, 24.5–25.0 min, 5%.

A Waters Xevo™ G2-XS QTOF (Waters, Manchester, UK) connected to the UPLC system *via* an ESI interface was used for profiling the substances. The optimal conditions were set as follows: capillary voltage was set as 3.0 kV under the positive ion mode and 2.0 kV under the negative ion mode. The sampling cone voltage was set as 45.0 V, and the extraction cone voltage as 4.0 V, respectively. The temperature was set at 100 °C, the desolvation gas temperature at 350 °C, and the desolvation gas flow rate at 600 L h^{-1}. The full scan MS^E data, in the centroid mode, were acquired across the mass range of 50–1500 Da. During the data acquisition, the mass was corrected using an external reference (Lock-Spray™) comprising a 200 pg mL^{-1} solution of leucine enkephalin *via* a lock-spray interface, generating a reference ion at *m*/*z* 554.262 Da ($[M-H]^-$) under the negative ion mode and *m*/*z* 556.277 Da under the positive ion mode.

2.6 Target network analysis

The targets (Homon species) were retrieved and collected from Swiss Target Prediction (**https://www.swisstargetprediction.ch**)[30] by importing the compounds, which were converted into the SMILES format. Meanwhile, the influenza-related targets were obtained from the DisGeNet database.[31] The protein–protein interactions (PPIs) were achieved using the STRING database (version 11.0, **https://string-db.org/**),[32] and protein interactions, with a confidence score > 0.7, were obtained after eliminating duplicates. The chemical constituent-target networks and protein–protein interaction (PPIs) networks were constructed and viewed by Cytoscape software (version 3.2.1),[33] and the network was analyzed by default setting with a ''degree'' value. All proteins/genes were subjected to pathway enrichment analysis (KEGG analysis) using the DAVID Bioinformatics resources 6.7 database.[34]

3. Results and discussion

3.1. Anti-flu effects of FA on H1N1 infected mice

Influenza virus infection could cause mortality and morbidity in the host, including decreased survival rate and worsened lung pathology.[35,36] Therefore, survival status and lung inflammation are two critical indices for evaluating the effectiveness of drugs. Generally, the body weight and survival rate of mice were significantly decreased after H1N1 infection. As shown in Fig. 2a, it was found that FA, with dosages of 20 g crude drug kg^{-1} day^{-1}, could significantly increase survival rates (40%, $p < 0.05$) and rescue body weight loss, as compared to those in the viral infection group. However, FA, with a dosage of 5 g kg^{-1} day^{-1}, did not display the protective effects. On day 5, lung index increased dramatically in the placebo group after H1N1 infection, while FA treatments significantly depressed this phenomenon (Fig. 2b). However, there were no effects of FA on the spleen index. Meanwhile, the lung inflammation caused by the H1N1 virus was ameliorated by FA (20 g crude drug kg^{-1} day^{-1}) (Fig. 2c). The lung score based on HE staining was also

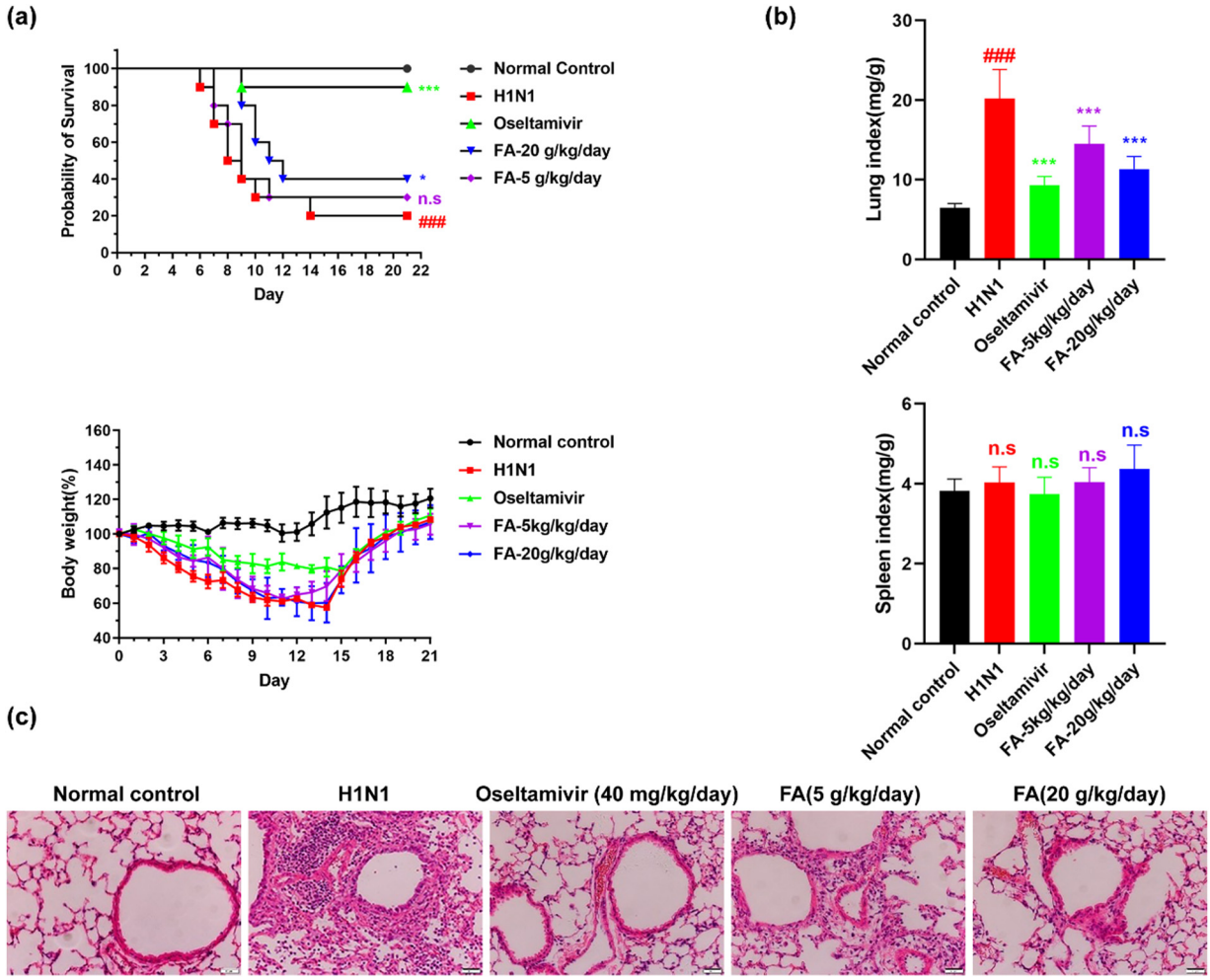

Fig. 2 The anti-H1N1 effects of Fructus Arctii. (a) 21 days survival rate and bodyweight; (b) lung index and spleen index at day 5 post H1N1 infections; (c) H&E staining of lungs at day 5 post H1N1 infections. Data were shown as the mean ± SD and further analyzed by one-way ANOVA Dunnett's multiple comparisons test, *vs.* Placebo $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, n.s means no significance as comparison with Placebo; *vs.* Normal control $^{\#}p < 0.05$, $^{\#\#}p < 0.01$, $^{\#\#\#}p < 0.001$.

calculated and shown in Fig. S1 (ESI†), reflecting the depressed lung inflammation after FA treatment.

3.2. Chemical profiling of Fructus Arctii by UHPLC/Q-TOF MS

The components of TCM were essential to their pharmacological mechanism and therapy achievement. In this part, the chemical constituents in FA were profiled by LC-MS, and further, they were systematically identified or tentatively characterized based on a target and non-target strategy reported in our previous work.[37] This strategy is described briefly. First, the formula and structure types of compounds isolated from FA were summarized (Table S2, ESI†), and the peaks with the same formula and matched fragmentation pathways were quickly screened out. Second, the mass fragmentation behaviors of representative compounds were sorted out and diagnostic ions were used for humongous compounds' screening. Last, the non-target compounds were predicted and characterized by a metabolism prediction platform and diagnostic ion extraction. As a result, a total of 123 components were tentatively characterized, including 111 lignans, 10 caffeoylquinic acid derivatives, 1 alkaloid, and 1 the other type. Among them, ten compounds were unambitiously identified by comparison with reference standards. Compared to the reported work,[23–25] fifty-three lignans were potential novel compounds.

3.2.1 Identification and characterization of the target compounds. Based on the target and non-target strategy, 62 matched compounds in an in-house library were screened out, and they were characterized in both positive and negative modes, including 52 lignans and 10 caffeoylquinic acids (Fig. 3, 4 and Table S2, ESI†).

3.2.1.1 Lignans. As reported in the previous work, lignans could lose hydroxyls, methyls, and phenyls to produce characteristic ions.[38] Interestingly, the eluted order and retention time were closely related to their chemical structural types, and the eluted order is listed as follows: glycosyl- and alcohol hydroxyl-containing lignans < alcohol hydroxyl-containing lignans [≥2*(-OH)] < benzofuran lignans < arctigenin-related lignans.

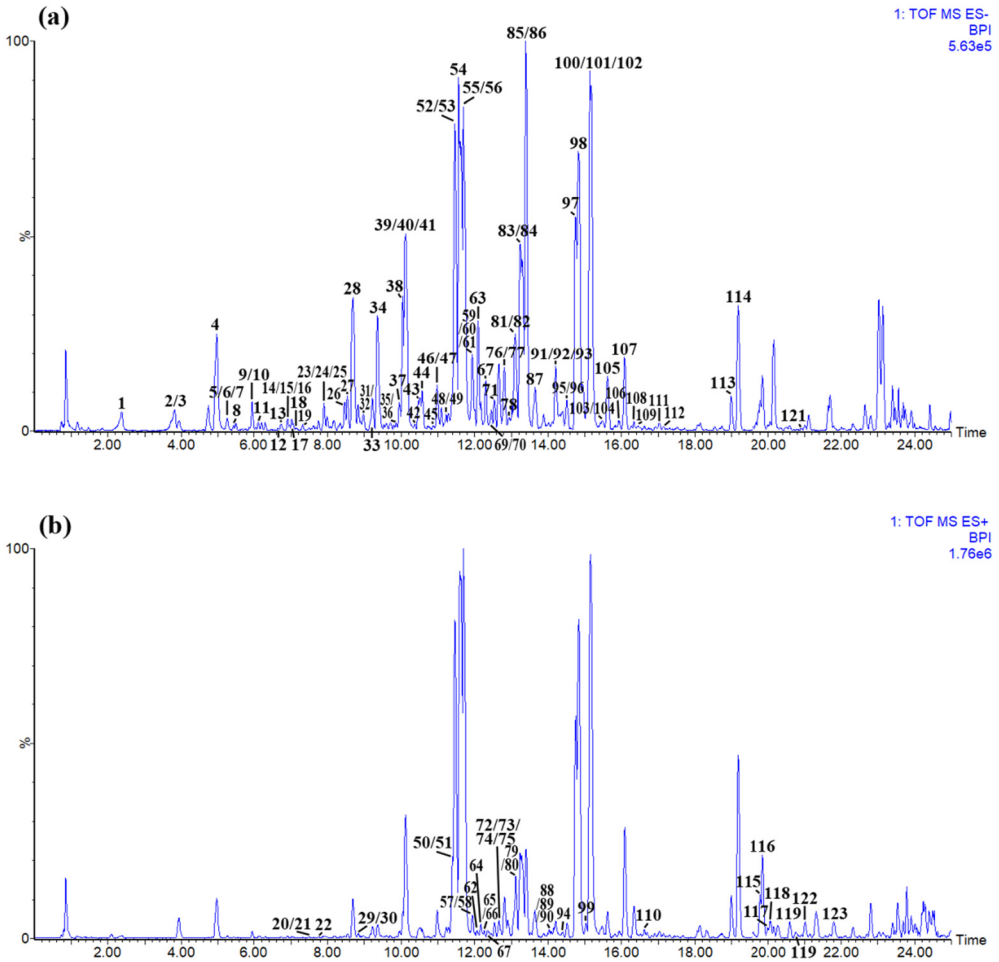

Fig. 3 The base peak ion chromatography of Fructus Arctii detected by UHPLC/Q-TOF MS. (a) negative ion mode; (b) positive ion mode.

3.2.1.1.1 Monolignans and their glycosides. Peaks **13**/**15**/**17**/**18**/**19**/**21** had the same deprotonated ion at *m*/*z* 535.18 with the formula of $C_{26}H_{31}O_{12}$. Peaks **13**/**15**/**18** exerted the same fragmentation behaviors, including the fragment ions at *m*/*z* 373.1283 and 355.1190. By searching data in the in-house library, only one butyrolactonelignan matched, and then, they were tentatively characterized as 4,7,4′-trihydroxy-3,3′-dimethoxyl-9-oxo-dibenzylbutyrolactonelignan-4-*O*-*β*-D-glucopyranoside or its isomers. Peaks **17**/**19** generated the fragment ion at *m*/*z* 313.1082 $[M\text{-}H\text{-}C_6H_{10}O_5\text{-}2CH_2O]^-$ ($C_{18}H_{17}O_5$), only 7,8-dihydrodehydrodiconiferyl alcohol-7′-oxo-4-*O*-*β*-D-glucopyranoside matched these fragmentation behaviors. Peak **21** produced a characteristic ion at *m*/*z* 151.0410 ($C_8H_7O_3$), and (7′*S*,8′*R*,8*S*)-4,4′,9′-trihydroxy-3,3′-dimethoxy-7′,9-epoxylignan-7-oxo-4-*O*-*β*-D-glucopyranoside matched. The MS spectrometry of **13**/**19**/**21** is shown in Fig. 5.

3.2.1.1.2 Sesquilignans. The sequilignans in Fructus Arctii could be classified into two sub-type based on their chemical structures, including two alcohol hydroxyl-containing lignans and benzofuran-lignans. Peaks **44**/**52**/**55**/**81** produced $[M\text{-}H]^-$ at *m*/*z* 553.20 with the same formula of $C_{30}H_{34}O_{10}$. Peak **81**, under the negative ion mode, produced the characteristic ion at *m*/*z* 357.1352 ($C_{20}H_{21}O_6$) (Fig. S2, ESI†). This fragment ion was formed by the cleavage of ether bone in lappaol E, and this has been reported.[39] The fragment ions of peaks **52**/**55** were the same, isolappaol C and lappaol C had the same formula in the database, and they could match these fragmentation behaviors. Using the *C* log *P*-value calculated in ChemDraw, peaks **52**/**55** were tentatively characterized as isolappaol C (*C* log *P* = 0.26) and lappaol C (*C* log *P* = 0.41). Lastly, peak **44** was characterized as arctignan A in the in-house library.

Peaks **97**/**98**/**99**/**100** had the same deprotonated ion at *m*/*z* 535.1975 with the formula of $C_{30}H_{32}O_9$. The peak areas of peaks **97**/**98** were higher than peaks **99**/**100**. Meanwhile, the characteristic fragment ions at *m*/*z* 505.1882 $[M\text{-}H\text{-}CH_2O]^-$ ($C_{29}H_{29}O_8$) and 283.0989 ($C_{17}H_{15}O_4$) were similar to lappaol F (Fig. S3, ESI†). In consideration of peak area and *C* log *P* value, peaks **97**and **98** were tentatively characterized as isolappaol A (*C* log *P* = 2.18) and lappaol A (*C* log *P* = 2.32) in the database. The other two peaks (**99**/**100**) were classified into non-target compounds.

3.2.1.1.3 Dilignans. Lappaol H was classified into dilignans with four alcohol hydroxyls. When extracting the calculated weight of lappaol H ($C_{40}H_{46}O_{14}$), peaks **40**/**51**/**53**/**56**/**60**/**72**/**76** were obviously found. Peak **40** yielded fragment ions at *m*/*z*

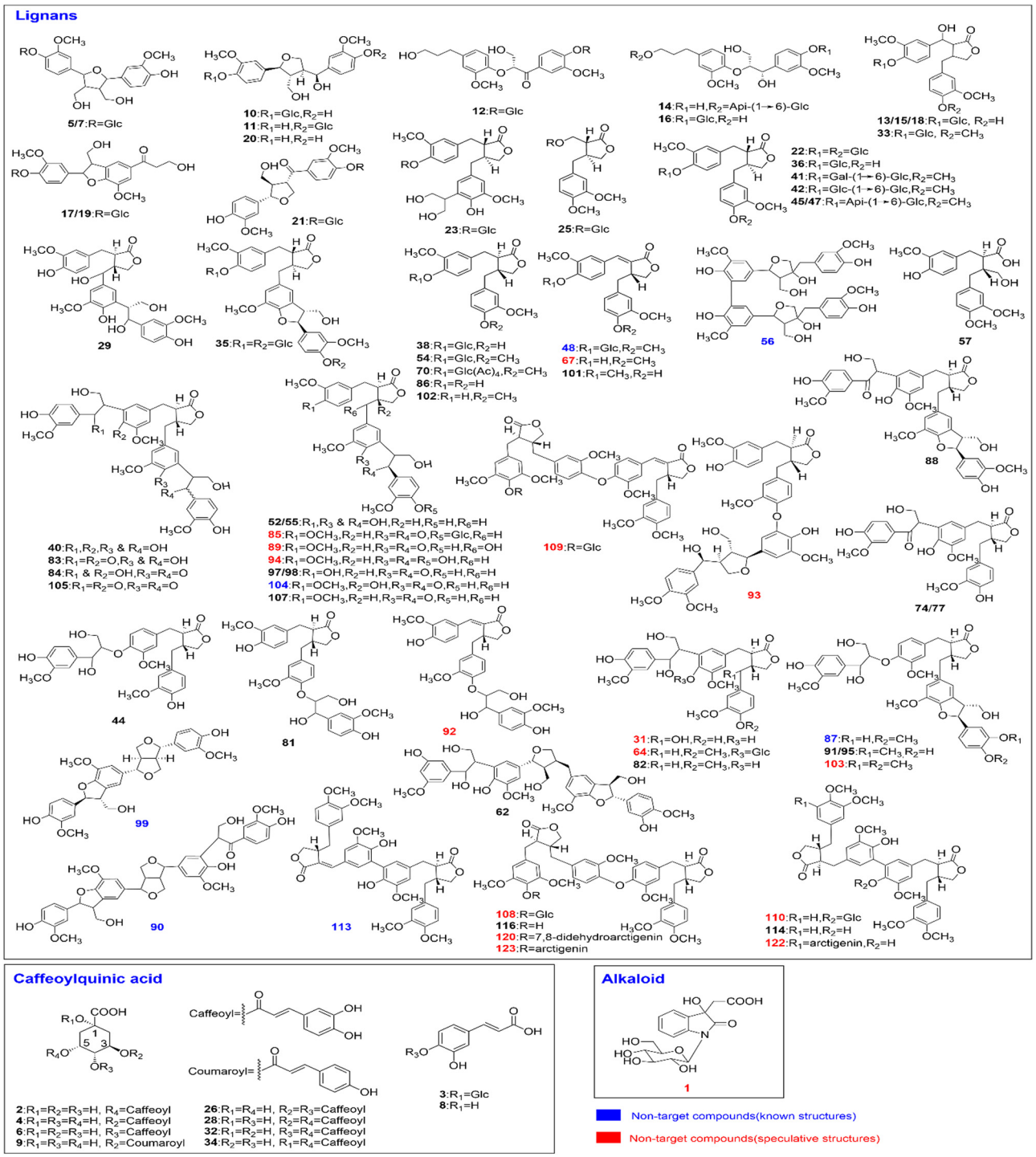

Fig. 4 The chemical structures of the identified or tentatively characterized compounds in Fructus Arctii.

731.2716, 713.2588 and 677.2382 by the successive losses of a series of H_2O. Peak **40** yielded a base ion at *m*/*z* 665.2383 by the loss of four hydroxyl moieties and a carbon atom. Meanwhile, lappaol H matched the above fragmentation behaviors since the structure of lappaol H includes four alcohol hydroxyl and it has characteristic fragmentation behavior of dehydration (Fig. S4, ESI†). Compared to the other six peaks, peak **40** had the largest peak area. Then, it was tentatively characterized as lappaol H. The others were classified as non-target compounds.

Meanwhile, dilignans containing benzofuran and alcohol hydroxyl groups were also found in Fructus Arctii. Six compounds (**80**, **83**, **84**, **87**, **91**, and **95**) had the same $[M\text{-}H]^-$ with the molecular formula of $C_{40}H_{44}O_{13}$. However, in the database, only four compounds had the same molecular formula. Among them, peaks **80**/**83**/**84** matched the fragmentation behaviors of

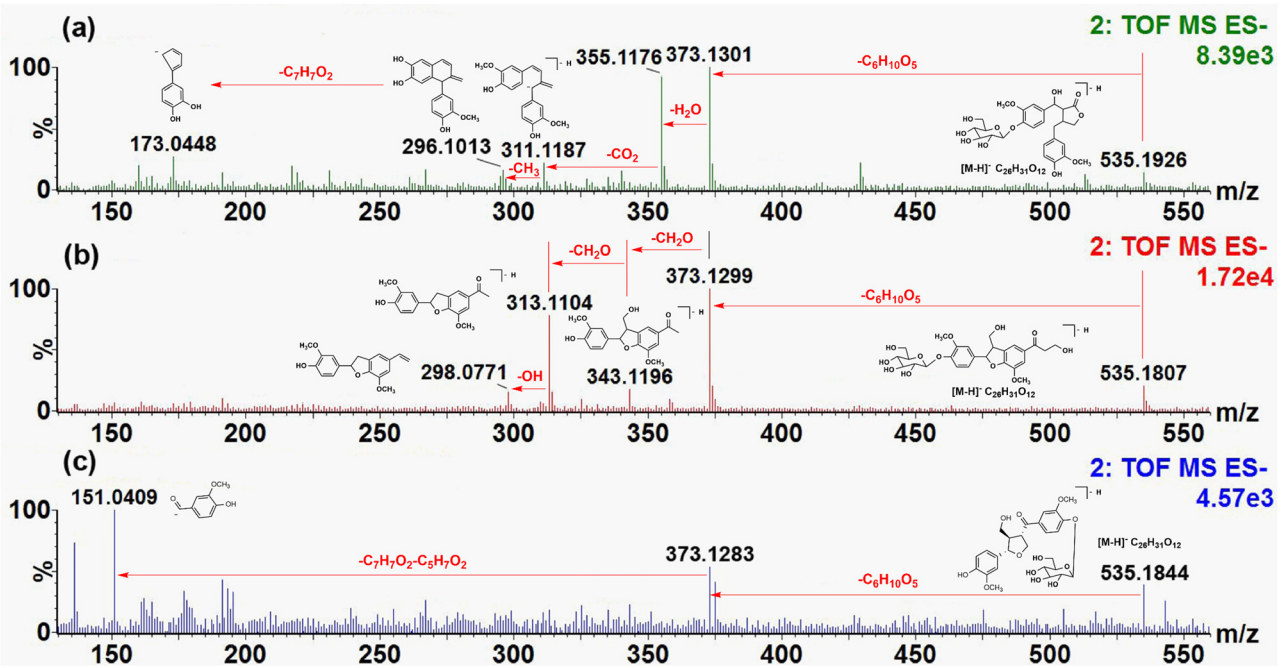

Fig. 5 Representative MS spectrometry of components in FA: (a) **13**, (b) **19**, and (c) **21**.

arctignan D/arctignan E (Table S2, ESI†); peaks **91**/**95** could produce the characteristic ion at *m*/*z* 505.1878 ($C_{31}H_{31}O_8$), indicating that they could match the arctignan G/arctignan H. Peaks **83**/**84**/**91**/**95** were tentatively characterized as arctignan D, arctignan E, arctignan G, and arctignan H. These four compounds had the benzofuran unit, resulting in the characteristic fragment ion at *m*/*z* 283.0970 ($C_{17}H_{15}O_4$) (Fig. S3, ESI†). The other compounds (**80**/**87**) were classified as non-target compounds.

Dibenzofuran dilignans could also be detected in Fructus Arctii. Peak **105** yielded the fragment ion $[M\text{-}H]^-$ at *m*/*z* 713.2612 ($C_{40}H_{41}O_{12}$) and predominant fragment ions at *m*/*z* 695.2497, 683.2524, 677.2398, 665.2401, and 653.2405 by the loss of H_2O and C. Moreover, it produced a characteristic ion at *m*/*z* 283.0976 by the benzofuran and phenol groups (Fig. S3, ESI†).

Further, it was tentatively characterized as lappaol F based on its mass fragment information (Table S2, ESI†).

3.2.1.1.4 Glycoside lignans. Peaks 5/7/10/11/12 had the same deprotonated ion at m/z 537.1983 ($C_{26}H_{33}O_{12}$). As the data shown in Table S2 (ESI†) peaks 5/7 had similar fragmentation behaviors, and both of them could produce the characteristic ion at *m*/*z* 297.1145 $[M\text{-}H\text{-}C_6H_{10}O_5\text{-}2CH_2OH\text{-}O]^-$ ($C_{18}H_{17}O_4$). Then, these were tentatively characterized as neo-olivil-4-*O*-*β*-D-glucopyranoside or its isomers. Differently, peaks **10**/**11** produced the characteristic ion at *m*/*z* 327.1248 $[M\text{-}H\text{-}C_6H_{10}O_5\text{-}2H_2O\text{-}C]^-$ ($C_{19}H_{19}O_5$). Then, they were tentatively characterized as 4,9,4′,7′-tetrahydroy-3,3′-dimethoxy-7,9′-epoxylignan-4-*O*-*β*-D-glucopyranoside or its isomers. Peak **12** produced a fragment ion at *m*/*z* 165.0536 ($C_9H_9O_3$). Only (8*R*)-4,9,9′-trihydroxy-3,3′-dimethoxy-7-oxo-8-*O*-4′-neolignan-4-*O*-*β*-D-glucopyranoside matched the fragmentation behaviors.

3.2.1.1.5 Others. Peaks **114**/**116** produced the deprotonated ion at *m*/*z* 741.29 ($C_{42}H_{45}O_{12}$). Under the positive ion mode, peak **116** exerted characteristic ions at *m*/*z* 387.1522 ($C_{21}H_{23}O_7$) and 369.1373 ($C_{21}H_{21}O_6$), which were the two groups formed by the cleavage of the ether bond (Fig. S5, ESI†). Then, peak **116** was characterized as neoarctin A. In addition, to neoarctin A, the other two compounds in the database had the same formula as $C_{42}H_{46}O_{12}$, including diarctigenin and neoarctin B. Peak **114**, under the positive ion mode, had a base peak at *m*/*z* 725.2980 due to the loss of H_2O, and diarctigenin matched these fragmentation behaviors (Fig. S6, ESI†). Then, peak **114** was tentatively characterized as diarctigenin.

3.2.1.2 Caffeoylquinic acid derivatives. In addition to the lignans, caffeoylquinic acid derivatives were the other main chemical types in Fructus Arctii. The structures of caffeoylquinic acid derivatives generally contain two main units (caffeic acid unit and quinic acid unit), and their fragment pathways have been reported previously.[40] The losses of caffeic acid moiety (179.0348, $C_9H_7O_4$), quinic acid (191.0554, $C_7H_{11}O_6$) dehydration, and CO_2 are the main characteristic fragment behaviors.[37] Caffeoylquinic acids (mono or dicaffeoylquinic acids) identified (peaks **2**, **3**, **4**, **6**, **8**, **9**, **26**, **28**, **32**, and **34**) in Fructus Arctii are illustrated in Table S2 (ESI†).

3.2.2 Characterization of non-target compounds in Fructus Arctii. Non-target compounds were screened out based on a metabolic platform in Masslynx. Based on the targeted compound, metabolic reactions (dehydrogenation, hydrogenation, glycosylation, methylation, *etc.*) were set up to explore the homologous compounds. As a result, 61 non-target compounds were screened out, including 59 lignans, 1 alkaloid, and 1 other. Among them, 53 were regarded as potential novel compounds by searching the results in SciFinder.

3.2.2.1 Lignans

3.2.2.1.1 Monolignans and their glycosides. Peak **48** had $[M-H + HCOOH]^-$ at *m/z* 577.1937 with the formula of $C_{28}H_{33}O_{13}$ and it produced the base ion at *m/z* 369.1364 ($C_{21}H_{21}O_6$) under the negative mode. Compared to the parent structure of arctiin (*m/z* 579.2119), the parent structure of peak **48** lacked two H. Further, it was tentatively characterized as hemislienoside by searching the results in SciFinder.

3.2.2.1.2 Sesquilignans. Peaks **89**/**92**/**104** had the same ion at *m/z* 565.20 $[M-H]^-$ with the formula of $C_{31}H_{34}O_{10}$. Compared to lappaol B ($C_{31}H_{34}O_9$), peaks **89**/**92**/**104** gained an additional O. Peak **89** had the characteristic ions at *m/z* 441.1544 ($C_{24}H_{25}O_8$), 423.1431 ($C_{24}H_{23}O_7$) and 411.1470 ($C_{23}H_{23}O_7$) under the negative ion mode. Peak **92** had a fragment ion at *m/z* 533.1855 ($C_{30}H_{29}O_9$) $[M-H-CH_2-H_2O]^-$, indicating that methyl and hydroxyls existed. Fragment ions at *m/z* 195.0680 ($C_{10}H_{11}O_4$), and 165.0573 ($C_9H_9O_3$) under the negative ion mode indicated that peak **92** had phenyl unit ($C_{10}H_{13}O_4$) in lappaol E (Fig. S2, ESI†). Peak **104** had the fragment ions at *m/z* 531.2008 ($C_{31}H_{31}O_8$) $[M-H-H_2O-O]^-$, 519.2096 ($C_{30}H_{31}O_8$) $[M-H-H_2O-O-C]^-$ and 397.1608 ($C_{23}H_{26}O_6$), and this fragmentation behavior matched lappaol B with hydroxyl at the C-2 position. Peak **104** was tentatively characterized as 2-hydroxylappaol B, and **89**/**92** found no suitable structures by searching the results in SciFinder. Peak **85** had $[M-H + HCOOH]^-$ at *m/z* 757.2730 ($C_{38}H_{45}O_{16}$) and the fragment ion at *m/z* 549.2117 ($C_{31}H_{33}O_9$) $[M-H-C_6H_{10}O_5]^-$. The other fragment ions under positive and negative ion modes were the same as lappaol B. The difference was that peak **85** gained an additional $C_6H_{10}O_5$ unit, and then peak **85** was tentatively characterized as lappaol B + Glc.

During the analysis of lappaol A/isolappaol A, their isomers were also found peaks **99**/**100**. Peaks **99** and **100** both produced fragment ions at *m/z* 517.18 ($C_{30}H_{29}O_8$) and 505.18 ($C_{29}H_{29}O_8$), indicating that they had a CH_2OH unit. Peak **99** had the fragment ion at *m/z* 283.0939 ($C_{17}H_{15}O_4$), suggesting that it had benzofuran or dihydroxyphenol unit. In the SciFinder, hedyotol A could have a similar fragmentation pathway as peak **99**. Further, peak **100** was characterized as lappaol A isomer.

3.2.2.1.3 Dilignans. Peaks **27**/**30** yielded a deprotonated ion at *m/z* 751.29 ($C_{40}H_{47}O_{14}$). Peaks **27**/**30** had the fragment ions at *m/z* 733.28, 715.27, 685.26, and 667.25, indicating that this compound had four alcoholic hydroxyls. The fragment ion at *m/z* 283.0987 indicated that it had a benzofuran or dihydroxyphenol unit. This compound had a similar fragmentation pathway as that of lappaol H (**40**), and the difference was that peaks **27**/**30** gained two additional H. Then, peaks **27**/**30** were tentatively characterized as lappaol H + 2H. When extracting lappaol H ($C_{40}H_{46}O_{14}$), seven peaks were found, including peaks **40**, **51**, **53**, **56**, **60**, **72**, and **76**. Among these, peak **40** was characterized as lappaol H. Peaks **51**/**53** had the fragment ions at *m/z* 731.27, 713.25, 683.24, and 665.23, indicating that they had four alcoholic hydroxyls. Meanwhile, both of them had the characteristic ion at *m/z* 701.26 ($C_{39}H_{41}O_{12}$) $[M-H-C-2H_2O]^-$, indicating that they could easily lose C and 2*OH. The fragment ion at *m/z* 283.0970 ($C_{17}H_{15}O_4$) indicated that they had benzofuran or dihydroxyphenol unit. Peak **56** also had four alcoholic hydroxyls by analyzing its fragment ions; however, one characteristic fragment ion at *m/z* 283.0970 was not found. It was tentatively characterized as cerberalignan A by searching the results in SciFinder. The fragment ions of peak **60** indicated that it had three alcoholic hydroxyls and a characteristic ion at *m/z* 401.1620 ($C_{22}H_{25}O_7$). Peaks **72** and **76** had four alcoholic hydroxyls, and they both had the characteristic ion at *m/z* 597.23 ($C_{32}H_{37}O_{11}$), indicating that they had a $C_8H_8O_3$ unit. **51**/**53**/**60**/**72**/**76** were potential novel compounds.

Peaks **93**/**103** had $[M + NH_4]^+$ *m/z* 764.3282 ($C_{41}H_{50}O_{13}N$), they exerted the same fragmentation behaviors. Peak **103** had the fragment ion at *m/z* 715.2682 $[M-H-H_2O]^-$, 697.2602 $[M-H-2H_2O]^-$, and 679.2489 $[M-H-3H_2O]^-$, indicating that it had three alcoholic hydroxyls. The characteristic ion at *m/z* 357.1341 ($C_{20}H_{21}O_6$) indicated that it had an ether bond. However, a suitable chemical structure could not be found in SciFinder. Its fragmentation behaviors were similar to arctignan G, and the difference was that peak **103** gained an additional CH_2. Then, peak **103** was characterized as arctignan G + CH_2.

Two dilignan glycosides (peaks **59**/**64**) were also found based on the metabolic prediction method. Peaks **59**/**64** had $[M-H + HCOOH]^-$ at *m/z* 775.28 with the formula $C_{37}H_{46}O_{15}$. They both had the fragment ions at *m/z* 567.2230 $[M-H-C_6H_{10}O_5]^-$ by losing 162.053 Da. Peak **59** had the characteristic ions at *m/z* 411.1429, 401.1627, 383.1520, and 371.1513, which were in line with peak **96**, and then it was characterized as lappaol D isomer 2 + Glc. Similarly, peak **64** had a similar fragmentation behavior as peak **82**, and then it was characterized as lappaol D + Glc. They were not found in SciFinder, and they are potential novel compounds.

3.2.2.1.4 New lignans (5 phenylpropanoids). Interestingly, compounds formed by five phenylpropanoids were found, and they were classified into new lignans, including peaks **37**/**39**/**43**/**46**/**49**/**50**/**69**/**75** and **58**/**61**/**66**/**68**/**71**/**73**/**78**/**79**. Peak **37** had the characteristic ions at *m/z* 927.3462 ($C_{50}H_{55}O_{17}$) $[M-H-H_2O]^-$, 909.3406 ($C_{50}H_{53}O_{16}$) $[M-H-2H_2O]^-$, 891.3221 ($C_{50}H_{51}O_{15}$) $[M-H-3H_2O]^-$, 873.3196($C_{50}H_{49}O_{14}$) $[M-H-4H_2O]^-$, 843.2952 ($C_{49}H_{47}O_{13}$) $[M-H-5H_2O-C]^-$ and 837.2925 ($C_{50}H_{45}O_{12}$) $[M-H-6H_2O]^-$, indicating that it had six alcoholic hydroxyls. Meanwhile, the other fragment ions at *m/z* 607.2224, 595.2129, 589.2062, 571.1993, 559.1963, and 547.2011 were the same as in lappaol H (Table S2, ESI†), indicating that it was the formed structure of lappaol H and the $C_{10}H_{12}O_4$ unit. Peak **39** had a similar fragmentation behavior as peak **37**. Peaks **43** and **46** had the fragment ions at *m/z* 927.34 $[M-H-H_2O]^-$, 909.33 $[M-H-2H_2O]^-$, indicating that they had two alcoholic hydroxyls. Meanwhile, the other fragment ions at *m/z* 531.20, 517.18, and 505.19 were the same as lappaol D, indicating that they had similar structures as lappaol D. Similarly, peaks **49**, **50**, **69**, and **75** had the same fragment ions as lappaol D (Table S2, ESI†). These eight peaks could not find matched structures in SciFinder, and they were regarded as potential novel compounds. Meanwhile, eight peaks **58**/**61**/**66**/**68**/**71**/**73**/**78**/**79** were found through

the metabolic reaction of dehydration. The fragmentation behaviors of these eight compounds were similar, such as losses of hydroxyls, methyl, phenol units. Moreover, they were also not found in SciFinder.

3.2.2.1.4 Others. Based on the metabolic platform, constituents correlated with neoarctin A and diarctigenin were found, including peaks **108**, **109**, and **110**. Peaks **108** and **110** both produced the $[M + Na]^+$ at *m/z* 927.34 and $[M + NH_4]^+$ at *m/z* 922.38, respectively. Their formula was calculated as $C_{48}H_{56}O_{17}$. Peaks **108**/**110** had a fragment ion at *m/z* 741.29 $[M\text{-}H\text{-}C_6H_{10}O_5]^-$ which was produced by losing $C_6H_{10}O_5$ (162.053 Da) from the parent structure. Furthermore, they had the same fragmentation behaviors as neoarctin A and diarctigenin. Then, peak **108** was tentatively characterized as neoarctin A + Glc, and peak **110** as diarctigenin + Glc, respectively. Similarly, peak **109** showed a fragment ion at *m/z* 925.32 $[M + Na]^+$ and 920.37 $[M + NH_4]^+$ with the formula of $C_{48}H_{54}O_{17}$. Meanwhile, the fragmentation ions of peak **109** were similar to those of neoarctin A. The difference between peak **109** and neoarctin A was that peak **109** lacked two H and gained one $C_6H_{10}O_5$ unit, and then peak **109** was characterized as neoarctin A-2H + Glc.

In the meantime, we also found the other peaks **113**/**115**/**117**/**118**/**119**/**121**. They all had the deprotonated ion at *m/z* 739.27 ($C_{42}H_{43}O_{12}$) and characteristic ions at *m/z* 247.10 ($C_{14}H_{15}O_4$), 203.10 ($C_{13}H_{15}O_2$), 151.0808 ($C_9H_{11}O_2$) under the positive ion mode, indicating that they had a characteristic structure ($C_{14}H_{15}O_4$) in neoarctin A (Fig. S5, ESI†). As compared to neoarctin A, peaks **113**/**115**/**117**/**118**/**119**/**121** lacked two H. In the SciFinder database, only viridissimaol E (diarctigenin-2H) had the same formula and matched the above fragmentation behaviors. Then, these six peaks were tentatively characterized as viridissimaol E or its isomers. This type of lignan was classified into butyrolactonelignan. Similarly, the other three peaks **120**/**122**/**123** were also classified into butyrolactonelignan. Peak **120** had the deprotonated ion at *m/z* 1109.4159 with the formula $C_{63}H_{65}O_{18}$, and as compared to peaks **113**/**115**/**117**/**118**/**119**/**121**, peak **120** gained an additional 7,8-didehydroarctigenin ($C_{21}H_{22}O_6$) unit. Notably, peak **120** produced a basic peak at *m/z* 739.2757 $[M\text{-}H\text{-}C_{21}H_{22}O_6]^-$, indicating there was an ether bond between diarctigenin-2H and 7,8-didehydroarctigenin units. It could not be found in SciFinder, and it was regarded as a potential novel compound. Peaks **122**/**123** produced $[M + H]^+$ at *m/z* 1113.45 ($C_{63}H_{69}O_{18}$), and as compared to neoarctin A (**116**), Peaks **122**/**123** gained an additional 7,8-didehydroarctigenin ($C_{21}H_{22}O_6$) unit. Peaks **122**/**123** had the primary fragments at *m/z* 727.31 $C_{42}H_{47}O_{11}$ $[M + H\text{-}C_{21}H_{22}O_6\text{-}O]^+$, indicating that peaks **122**/**123** had the $-C_{21}H_{22}O_6$ and -O. There were also no matched structures in SciFinder.

3.2.2.2 Alkaloids and others. Peak **1** was reported in our previous work, and it was classified into one type of alkaloids-indoleacetic acid derivatives.[41] Successive losses of CO and CH_2 molecules from the fragment ion at *m/z* 188.0717 $[^{0,2}X\text{-}COOH\text{-}OH + H]^+$ ($C_{11}H_{10}NO_2$) produced another obvious ion at *m/z* 146.0610 (C_9H_8NO).

Peak **24** had the fragment ions at *m/z* 451.1562 $[M\text{-}H\text{-}C_6H_{10}O_5]^-$ ($C_{22}H_{27}O_{10}$) and 179.0375 ($C_9H_7O_4$), indicating that it had the feruloyl moiety ($C_9H_7O_4$) and mono-saccharides ($C_6H_{10}O_5$). By searching the data in SciFinder, no compounds matched. Then, it was tentatively regarded as a potential novel compound.

3.3 Xenobiotics in H1N1 infected mice after ingestion of FA

Anti-influenza effects of FA were confirmed by an H1N1 virus-infected mouse model described in Section 3.1. Analyzing the xenobiotics in H1N1 infected mice after ingestion of FA, it was easier to elucidate its anti-influenza mechanism and therapeutic substances. Nine types of samples were collected and analyzed, including plasma, urine, feces, heart, liver, spleen, lung, kidneys, and brain. In this part, the biosamples of H1N1-infected mice treated with the effective dose of FA were obtained. As a result, a total of 124 xenobiotics were identified or tentatively characterized in biosamples, including 38 prototypes (their structures are shown in Fig. 4) and 86 metabolites. The EICs of FA-related xenobiotics in the drugged biosample are presented in Fig. 6.

3.3.1 Identification of prototype components of FA. According to the retention time and MS fragments of chemical components in the FA extract, 38 prototype components were detected in biosamples (shown in Table 1). In detail, these prototypes could be classified into 2 caffeoylquinic acids and 36 lignans, and all of them were detected in mice's feces.

3.3.2 Identification of metabolites of FA. In this study, 86 metabolites were screened in FA-related biosamples, and all of them were lignans (shown in Table 1). Among them, monolignans and their glycosides were the predominant metabolites, and their amount reached fifty-six.

3.3.2.1. Monolignans and their glycosides-related metabolites. During the process of characterization of metabolites, we found that the metabolites of monolignans and their glycosides were mainly original from four basic parent structures, including arctigenic acid, matairesinol, arctigenin, and 7,8-didehydroarctigenin. Considering this pheromone, the metabolic pathways of arctigenin-related in H1N1 infected mice (shown in Fig. 7) were proposed in this work.

3.3.2.1.1 Phase I reactions-related. Metabolite **M1**, which lacked three methyls than arctigenic acid, exerted $[M\text{-}H]^-$ ion at *m/z* 347.1150 ($C_{18}H_{19}O_7$) and the similar fragmentation behavior as arctigenic acid. Metabolites **M35**/**M48** were characterized as the demethylation products of arctigenic acid. The demethylation products (**M29**, **M80**, and **M83**) of arctigenin had been detected in biosamples. Metabolites **M6**/**M9**, with the same $[M\text{-}H]^-$ ion at *m/z* 373.12 ($C_{20}H_{21}O_7$), had one oxygen more than matairesinol. Similarly, metabolite **M77** was characterized as a hydroxylation product of arctigenin. The demethylation and dehydroxylation products (**M82** and **M84**) of arctigenin had been detected.

3.3.2.1.2 Phase II reactions-related. The phase II reactions mainly involved glucuronidation and sulfation. These two types

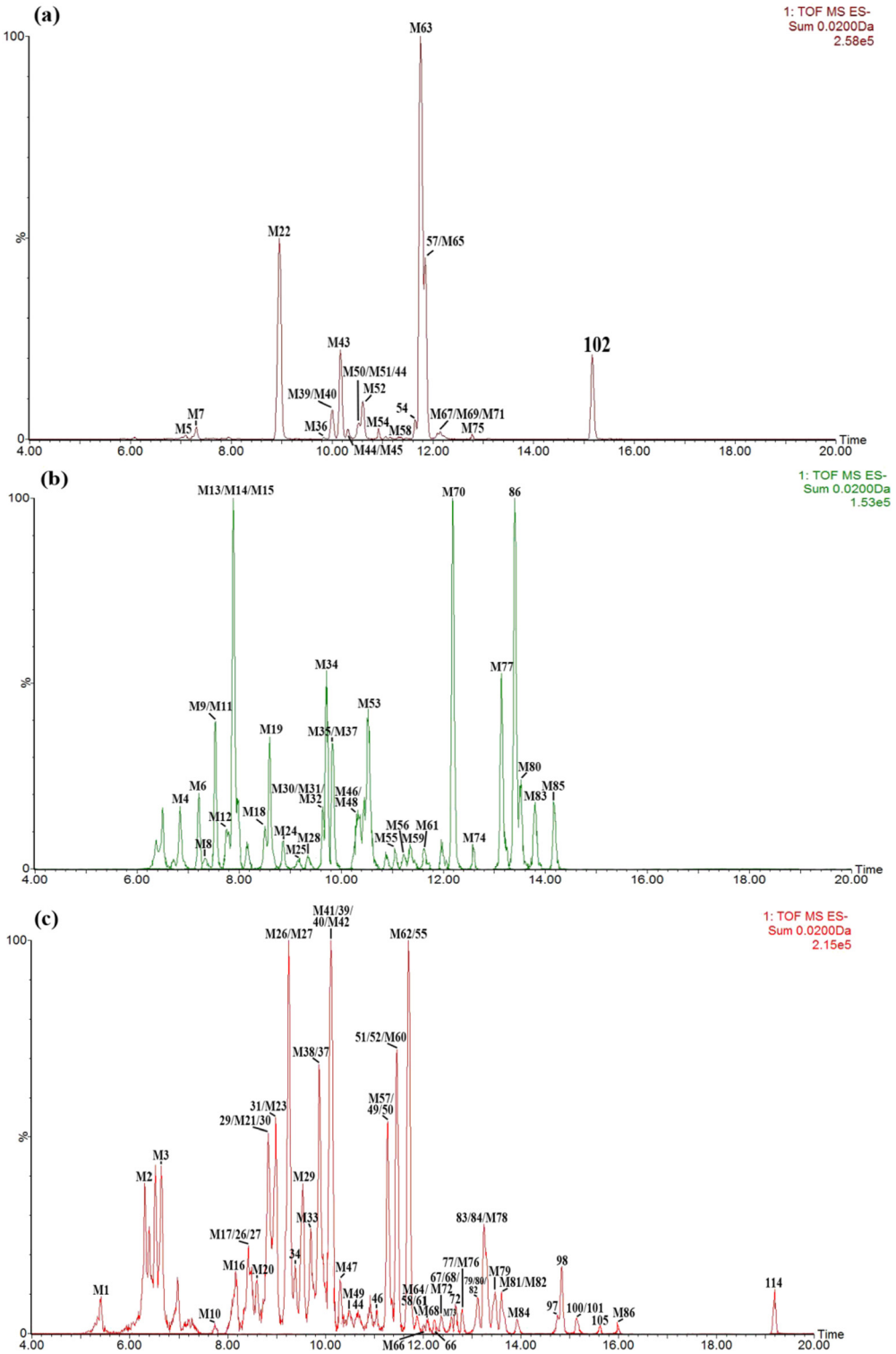

Fig. 6 Extracted in chromatograms (EICs) of FA-related xenobiotics in mice under negative ion mode. (a) plasma, (b) urine and (c) feces.

of metabolites have typical neutral losses, which were 176.032 Da ($C_6H_8O_6$) and 79.957 Da (SO_3). The glucuronidation products were identified, including **M22**, **M63**/**M67**, and **M58**/**M75**. In addition, the sulfation metabolites were detected, including **M30**, **M70**/**M74**, and **M78**.

3.3.2.1.3 Phases I and II reactions-related. Metabolites **M5**/**M7** gave the same $[M\text{-}H]^-$ ion at *m*/*z* 551.17 ($C_{26}H_{31}O_{13}$), which was 176.032 Da more than **M35**/**M48**. Four metabolites (**M11**, **M13**, **M19**, and **M21**) were the sulfation products of **M35**/**M48**. Moreover, one of them demethylated to produce metabolite

Table 1 Xenobiotics detected in H1N1 infected mice after oral administration of Fructus Arctii extract

NO.	t_R (min)	Molecular formula	Selected ion	Measured ion	Error (mDa)	(ESI−/ESI+) MS fragmentations	Identification or characterization	Source
M1	5.41	$C_{18}H_{20}O_7$	$[M-H]^-$	347.1150	1.9	299.0914 $[M-H-2H_2O-C]^-$, **281.0864 $[M-H-3H_2O-C]^-$**	Arctigenic acid-3CH_2	F
M2	6.31	$C_{26}H_{28}O_{14}$	$[M-H]^-$	563.1417	1.6	383.0804 $[M-H-Glc-H_2O]^-$, **353.0674 $[M-H-Glc-2H_2O-C]^-$**	Matairesinol + 3O-4H + Glc **1**	U,F
M3	6.63	$C_{26}H_{28}O_{14}$	$[M-H]^-$	563.1449	4.8	383.0842 $[M-H-Glc-H_2O]^-$, **353.0692 $[M-H-Glc-2H_2O-C]^-$**	Matairesinol + 3O-4H + Glc **2**	U,F
M4	6.84	$C_{20}H_{24}O_{11}S$	$[M-H]^-$	471.0985	2.4	391.1439 $[M-H-SO_3]^-$, **165.0556 $C_9H_9O_3$**, 150.0329 $C_8H_6O_3$	Arctigenic acid-CH_2 + O + SO_3	U,Li
M5	7.09	$C_{26}H_{32}O_{13}$	$[M-H]^-$	551.1782	1.7	399.1343 $[M + Na-GluA]^+$, **137.0618 $C_8H_9O_2$**	Arctigenic acid-CH_2 + GluA **1**	P,U
M6	7.21	$C_{20}H_{22}O_7$	$[M-H]^-$	373.1299	1.2	329.1412 $[M-H-CO_2]^-$, **314.1183 $[M-H-CO_2-CH_3]^-$**, 299.0964 $[M-H-CO_2-2CH_3]^-$	Matairesinol + O	U,F,K
M7	7.31	$C_{26}H_{32}O_{13}$	$[M-H]^-$	551.1790	2.5	399.1359 $[M + Na-GluA]^+$, **137.0591 $C_8H_9O_2$**	Arctigenic acid-CH_2 + GluA **2**	P,U
M8	7.35	$C_{26}H_{30}O_{13}$	$[M-H]^-$	549.1633	2.5	373.1273 $[M-H-GluA]^-$	Matairesinol + O + GluA	P,U
M9	7.53	$C_{20}H_{22}O_7$	$[M-H]^-$	373.1314	2.7	329.1418 $[M-H-CO_2]^-$, **314.1178 $[M-H-CO_2-CH_3]^-$**, 299.1003 $[M-H-CO_2-2CH_3]^-$	Arctigenin-CH_2 + O	U,F,Lu,K
M10	7.61	$C_{30}H_{36}O_{11}$	$[M-H]^-$	571.2222	4.3	505.1823 $[M-H-3H_2O-C]^-$, **445.1919 $[M-H-C_2H_6O_6]^-$**	Reported compound 1 + 2H	F
M11	7.75	$C_{20}H_{24}O_{10}S$	$[M-H]^-$	455.1035	2.3	**375.1470 $[M-H-SO_3]^-$**, 331.1566 $[M-H-SO_3-CO_2]^-$, 316.1363$[M-H-SO_3-CO_2-CH_3]^-$, 239.0914 $[M-H-C_8H_8O_2]^-$	Arctigenic acid-CH_2 + SO_3 **1**	P,U
M12	7.80	$C_{20}H_{22}O_{10}S$	$[M-H]^-$	453.0855	3.7	373.1301 $[M-H-SO_3]^-$, **313.1114 $[M-H-SO_3-2CH_2O]^-$**	Matairesinol + O + SO_3 **1**	P,U,F
M13	7.89	$C_{20}H_{24}O_{10}S$	$[M-H]^-$	455.1044	3.2	**375.1470 $[M-H-SO_3]^-$**, 331.1543 $[M-H-SO_3-CO_2]^-$, 316.1345 $[M-H-SO_3-CO_2-CH_3]^-$, 239.0948 $[M-H-C_8H_8O_2]^-$	Arctigenic acid-CH_2 + SO_3 **2**	P,U
M14	7.94	$C_{20}H_{22}O_{10}S$	$[M-H]^-$	453.0855	4.4	373.1344 $[M-H-SO_3]^-$, **313.1048 $[M-H-SO_3-2CH_2O]^-$**	Matairesinol + O + SO_3 **2**	U,F,Li
M15	7.97	$C_{19}H_{22}O_{10}S$	$[M-H]^-$	441.0889	3.4	361.1348 $[M-H-SO_3]^-$	Arctigenic acid-2CH_2 + SO_3	U
M16	8.16	$C_{40}H_{46}O_{15}$	$[M-H]^-$	765.2726	−3.2	747.2606 $[M-H-H_2O]^-$, 681.2353 $[M-H-4H_2O-C]^-$, **575.1934 $[M-H-C_7H_{10}O_4-2H_2O-C]^-$**	Lappaol H + O	F
M17	8.43	$C_{30}H_{34}O_{13}S$	$[M-H]^-$	633.1688	4.6	533.1777 $[M-H-SO_3]^-$, 517.1951 $[M-H-SO_3-2H_2O]^-$, **411.1502 $[M-H-SO_3-C_7H_{10}O_3]^-$**	Arctignan A + SO_3	F
26*	8.49	$C_{25}H_{24}O_{12}$	$[M-H]^-$	515.1224	3.4	**353.0889 [M-H-caffeoyl]$^-$**, 191.0526 $[M-H-2caffeoyl]^-$, 179.0387 $[M-H-caffeoyl-QA]^-$, 135.0460 $[M-H-caffeoyl-QA-2H_2O]^-$	3,4-*O*-Dicaffeoylquinic acid	F
27	8.50	$C_{40}H_{48}O_{14}$	$[M-H]^-$	751.3013	4.7	**679.2562 $[M-H-4H_2O]^-$**, 667.2417 $[M-H-4H_2O-C]^-$, 531.1984 $[M-H-C_9H_{16}O_7]^-$	Lappaol H + 2H **1**	F
M18	8.51	$C_{20}H_{22}O_{11}S$	$[M-H]^-$	469.0829	2.4	453.0940 $[M-H-O]^-$, **389.1260 $[M-H-SO_3]^-$**	Matairesinol + 2O + SO_3	U
M19	8.59	$C_{20}H_{24}O_{10}S$	$[M-H]^-$	455.1025	1.3	**375.1464 $[M-H-SO_3]^-$**, 331.1546 $[M-H-SO_3-CO_2]^-$, 316.1323 $[M-H-SO_3-CO_2-CH_3]^-$, 239.0952 $[M-H-C_8H_8O_2]^-$	Arctigenic acid-CH_2 + SO_3 **3**	P,U,F,Lu
M20	8.60	$C_{46}H_{56}O_{19}$	$[M-H]^-$	911.3374	3.6	731.2695 $[M-H-Glc-H_2O]^-$, 713.2505$[M-H-Glc-2H_2O]^-$, 677.2453 $[M-H-Glc-4H_2O]^-$, **665.2439 $[M-H-Glc-4H_2O-C]^-$**, 571.2051 $[M-H-Glc-C_7H_{10}O_4-2H_2O]^-$, 559.1996 $[M-H-C_7H_{10}O_4-2H_2O-C]^-$, 531.2050 $[M-H-C_7H_{10}O_4-2H_2O-2C-O]^-$	Lappaol H + Glc	U,F

Table 1 (*continued*)

NO.	t_R (min)	Molecular formula	Selected ion	Measured ion	Error (mDa)	(ESI−/ESI+) MS fragmentations	Identification or characterization	Source
29	8.82	$C_{30}H_{34}O_{11}$	$[M-H]^-$	569.2051	2.8	**533.1845 $[M-H-2H_2O]^-$**, 521.1876 $[M-H-2H_2O-C]^-$, 427.1404 $[M-H-C_7H_{10}O_3]^-$, 367.1187 $[M-H-C_8H_{10}O_3-2H_2O-C]^-$	Reported compound **1**	F
M21	8.87	$C_{20}H_{24}O_{10}S$	$[M-H]^-$	455.1010	−0.2	375.1455 $[M-H-SO_3]^-$	Arctigenic acid-CH_2 + SO_3 **4**	U,F
30	8.89	$C_{40}H_{48}O_{14}$	$[M-H]^-$	751.2972	0.6	733.2864 $[M-H-H_2O]^-$, 715.2651 $[M-H-2H_2O]^-$, 667.2532 $[M-H-4H_2O-C]^-$, 655.2513 $[M-H-4H_2O-2C]^-$, 561.2120 $[M-H-C_7H_{10}O_4-2H_2O-C]^-$, **371.1457 $C_{21}H_{23}O_6$,** 283.0979 $C_{17}H_{15}O_4$		F
M22	8.95	$C_{27}H_{34}O_{13}$	$[M-H]^-$	565.1940	1.9	413.1573 $[M + Na-GluA]^+$, **373.1691** $[M + H-GluA-H_2O]^+$, 355.1447 $[M + H-GluA-2H_2O]^+$, 305.1208 $[M + H-GluA-4H_2O-CH_2]^+$	Arctigenic acid + GluA	P,U,Li
31	8.96	$C_{30}H_{34}O_{11}$	$[M-H]^-$	569.2029	0.6	**521.1966 $[M-H-2H_2O-C]^-$**, 427.1333 $[M-H-C_7H_{10}O_3]^-$	Isomer of reported compound **1**	F
M23	8.98	$C_{30}H_{34}O_{13}S$	$[M-H]^-$	633.1685	4.3	517.1817 $[M-H-SO_3-2H_2O]^-$, **505.1947 $[M-H-SO_3-2H_2O-C]^-$**, 477.1942 $[M-H-SO_3-2H_2O-C-CO]^-$, 411.1448 $[M-H-SO_3-C_7H_{10}O_3]^-$	Isolappaol C + SO_3 **1**	F
M24	9.05	$C_{36}H_{42}O_{16}$	$[M-H]^-$	729.2399	0.4	**517.1881 $[M-H-GluA-2H_2O]^-$**, 505.1948 $[M-H-GluA-2H_2O-C]^-$	Arctignan A + GluA	P,U
M25	9.15	$C_{36}H_{42}O_{16}$	$[M-H]^-$	729.2432	3.7	517.1924 $[M-H-GluA-2H_2O]^-$, 505.1893 $[M-H-GluA-2H_2O-C]^-$, 411.1471 $[M-H-GluA-C_7H_{10}O_3]^-$, **283.0999 $C_{17}H_{15}O_4$**	Isolappaol C + GluA	P,U
M26	9.25	$C_{30}H_{34}O_{13}S$	$[M-H]^-$	633.1658	1.6	617.1753 $[M-H-O]^-$,517.1917 $[M-H-SO_3-2H_2O]^-$, **505.1908 $[M-H-SO_3-2H_2O-C]^-$**, 411.1523 $[M-H-SO_3-C_7H_{10}O_3]^-$	Lappaol C + SO_3 **1**	F
M27	9.25	$C_{18}H_{18}O_8S$	$[M-H]^-$	393.0690	4.6	313.1091 $[M-H-SO_3]^-$	Arctigenin-3CH_2-O + SO_3	F
M28	9.34	$C_{36}H_{42}O_{16}$	$[M-H]^-$	729.2444	4.9	517.1924 $[M-H-GluA-2H_2O]^-$, **505.1878 $[M-H-GluA-2H_2O-C]^-$**, 411.1393 $[M-H-GluA-C_7H_{10}O_3]^-$	Lappaol C + GluA	P,U
34	9.39	$C_{25}H_{24}O_{12}$	$[M-H]^-$	515.1234	4.4	353.0988 $[M-H-caffeoyl]^-$, 191.0625 $[M-H-2caffeoyl]^-$, 179.0389 $[M-H-caffeoyl-QA]^-$, **173.0524 $[M-H-C_{18}H_{14}O_7]^-$**	1,5-*O*-Dicaffeoylquinic acid	F
M29	9.53	$C_{18}H_{18}O_6$	$[M-H]^-$	329.1051	2.6	207.0689 $[M-H-C_7H_6O_2]^-$	Arctigenin-3CH_2	F
M30	9.64	$C_{21}H_{26}O_{10}S$	$[M-H]^-$	469.1198	3.0	389.1566 $[M-H-SO_3]^-$	Arctigenic acid + SO_3	P,U,F
M31	9.64	$C_{36}H_{44}O_{15}$	$[M-H]^-$	715.2645	4.3	**517.1920 $[M-H-Glc-2H_2O]^-$**, 505.1926 $[M-H-Glc-2H_2O-C]^-$, 411.1450 $[M-H-Glc-C_7H_{10}O_3]^-$	Isolappaol C + Glc	U,F
M32	9.66	$C_{36}H_{42}O_{16}$	$[M-H]^-$	729.2407	1.2	**517.1861$[M-H-GluA-2H_2O]^-$**, 505.1888 $[M-H-GluA-2H_2O-C]^-$, 411.1449 $[M-H-GluA-C_7H_{10}O_3]^-$, 357.1333 $[M-H-GluA-C_{10}H_{12}O_4]^-$	Lappaol E + GluA	P,U,F
M33	9.70	$C_{30}H_{34}O_{13}S$	$[M-H]^-$	633.1683	4.1	**517.1880 $[M-H-SO_3-2H_2O]^-$**, 505.1971 $[M-H-SO_3-2H_2O-C]^-$, 411.1461 $[M-H-SO_3-C_7H_{10}O_3]^-$	Isolappaol C + SO_3 **2**	F
M34	9.71	$C_{25}H_{26}O_{12}$	$[M-H]^-$	517.1370	2.4	341.1063 $[M-H-GluA]^-$	Arctigenin-2CH_2-2H + GluA	U
M35	9.84	$C_{20}H_{24}O_7$	$[M-H]^-$	375.1461	1.7	331.1571 $[M-H-CO_2]^-$, **316.1313 $[M-H-CO_2-CH_3]^-$**	Arctigenic acid-CH_2 **1**	P,U,F
M36	9.84	$C_{20}H_{22}O_{10}S$	$[M-H]^-$	453.0871	1.6	**373.1327 $[M-H-SO_3]^-$**, 358.1473 $[M-SO_3-O]^-$, 235.0982 $[M-H-SO_3-C_7H_6O_3]^-$	Arctigenin-CH_2 + O + SO_3	P,U,F,Li,K

Table 1 (*continued*)

NO.	t_R (min)	Molecular formula	Selected ion	Measured ion	Error (mDa)	(ESI−/ESI+) MS fragmentations	Identification or characterization	Source
M37	9.85	$C_{36}H_{44}O_{15}$	$[M-H]^-$	715.2646	4.4	**517.1917 $[M-H-Glc-2H_2O]^-$**, 411.1520 $[M-H-Glc-C_7H_{10}O_3]^-$	Lappaol C + Glc	U,F
M38	9.88	$C_{29}H_{32}O_{10}$	$[M-H]^-$	539.1960	4.3	503.1805 $[M-H-2H_2O]^-$, **491.1753 $[M-H-2H_2O-C]^-$**	Isolappaol C-CH_2	F
37	9.96	$C_{50}H_{58}O_{18}$	$[M-H]^-$	945.3574	2.9	927.3487 $[M-H-H_2O]^-$, 897.3442 $[M-H-2H_2O-C]^-$, 825.2891 $[M-H-6H_2O-C]^-$, **813.2893 $[M-H-6H_2O-2C]^-$**, 719.2555 $[M-H-C_8H_{18}O_7]^-$, 707.2465 $[M-H-C_8H_{18}O_7-C]^-$	Unknown	F
M39	9.99	$C_{26}H_{30}O_{12}$	$[M-H]^-$	533.1703	4.4	381.1289 $[M + Na-GluA]^+$, 341.1423 $[M + H-GluA-H_2O]^+$, **137.0615 $C_8H_9O_2$**	Arctigenin-CH_2 + GluA **1**	P,U,F,H,Li,Lu,K
M40	10.01	$C_{20}H_{22}O_9S$	$[M-H]^-$	437.0918	1.2	357.1371 $[M-H-SO_3]^-$	Arctigenin-CH_2 + SO_3 **1**	P,F,Li
M41	10.07	$C_{30}H_{34}O_{13}S$	$[M-H]^-$	633.1691	4.9	505.1909 $[M-H-SO_3-2H_2O-C]^-$, **411.1497 $[M-H-SO_3-C_7H_{10}O_3]^-$**	Lappaol C + SO_3 **2**	F
39	10.09	$C_{50}H_{58}O_{18}$	$[M-H]^-$	945.3562	1.7	927.3511 $[M-H-H_2O]^-$, 825.2864 $[M-H-6H_2O-C]^-$, 813.2885 $[M-H-6H_2O-2C]^-$, 607.2216 $[M-H-C_{17}H_{22}O_7]^-$, 589.2120 $[M-H-C_{17}H_{22}O_7-H_2O]^-$, 571.2025 $[M-H-C_{17}H_{22}O_7-2H_2O]^-$, **559.2009 $[M-H-C_{17}H_{22}O_7-2H_2O-C]^-$**	Unknown	F
40	10.12	$C_{40}H_{46}O_{14}$	$[M-H]^-$	749.2849	4.0	665.2424 $[M-H-4H_2O-C]^-$, **653.2432 $[M-H-4H_2O-2C]^-$**, **559.1986 $[M-H-C_7H_{10}O_3-2H_2O-C]^-$**	Lappaol H	P,U,F
M42	10.13	$C_{29}H_{32}O_{10}$	$[M-H]^-$	539.1954	3.7	503.1799 $[M-H-2H_2O]^-$, **491.1736 $[M-H-2H_2O-C]^-$**	Lappaol C-CH_2	F
M43	10.16	$C_{26}H_{30}O_{12}$	$[M-H]^-$	533.1688	2.9	381.1288 $[M + Na-GluA]^+$, 341.1447 $[M + H-GluA-H_2O]^+$, **137.0609 $C_8H_9O_2$**	Arctigenin-CH_2 + GluA **2**	P,U,F,H,Li,Lu,K
M44	10.32	$C_{20}H_{22}O_9S$	$[M-H]^-$	437.0930	2.4	357.1345 $[M-H-SO_3]^-$	Arctigenin-CH_2 + SO_3 **2**	P,U,F,H,Li,Lu,K
M45	10.33	$C_{25}H_{28}O_{11}$	$[M-H]^-$	503.1577	2.4	**351.1280 $[M + Na-GluA]^+$**, 329.1326 $[M + H-GluA]^+$	Arctigenin-CH_2-OCH_2 + GluA	P,U
M46	10.33	$C_{25}H_{28}O_{12}$	$[M-H]^-$	519.1533	3.0	343.1195 $[M-H-GluA]^-$	Arctigenin-$2CH_2$ + GluA	U
M47	10.37	$C_{19}H_{20}O_9S$	$[M-H]^-$	423.0779	2.9	343.1224 $[M-H-SO_3]^-$	Arctigenin-$2CH_2$ + SO_3	P,U,F,Li,Lu,K
M48	10.38	$C_{20}H_{24}O_7$	$[M-H]^-$	375.1457	1.3	**298.1194 $[M-H-C_2H_5O_3]^-$**, 253.1060 $[M-H-C_7H_6O_2]^-$	Arctigenic acid-CH_2 **2**	U,F
M49	10.49	$C_{21}H_{24}O_{10}S$	$[M-H]^-$	467.1016	0.4	**387.1481 $[M-H-SO_3]^-$**,298.1220 $[M-H-SO_3-C_3H_5O_3]^-$	Arctigenin + O + SO_3 **1**	U,F,Li
M50	10.51	$C_{20}H_{22}O_9S$	$[M-H]^-$	437.0921	1.5	357.1384 $[M-H-SO_3]^-$	Arctigenin-CH_2 + SO_3 **3**	P,U,F,H,Li,Lu,K
M51	10.56	$C_{26}H_{30}O_{12}$	$[M-H]^-$	533.1669	1.0	381.1253 $[M + Na-GluA]^+$, 341.1396 $[M + H-GluA-H_2O]^+$, **137.0606 $C_8H_9O_2$**	Arctigenin-CH_2 + GluA **3**	P,U,Li,K
44	10.57	$C_{30}H_{34}O_{10}$	$[M-H]^-$	553.2109	3.5	**517.1858 $[M-H-2H_2O]^-$**, 487.1645 $[M-H-2H_2O-CH_2O]^-$	Arctignan A	F
M52	10.61	$C_{20}H_{22}O_9S$	$[M-H]^-$	437.0921	1.0	357.1355 $[M-H-SO_3]^-$	Arctigenin-CH_2 + SO_3 **4**	P,U,F,H,Li,S,Lu,K,B
M53	10.67	$C_{20}H_{20}O_9S$	$[M-H]^-$	435.0798	4.8	355.1147 $[M-H-SO_3]^-$	Arctigenin-CH_2-2H + SO_3 **1**	U,F
M54	10.87	$C_{20}H_{22}O_9S$	$[M-H]^-$	437.0947	4.1	357.1318 $[M-H-SO_3]^-$	Arctigenin-CH_2 + SO_3 **5**	P,U,F,Li,K
M55	11.05	$C_{20}H_{20}O_9S$	$[M-H]^-$	435.0786	3.6	355.1225 $[M-H-SO_3]^-$	Arctigenin-CH_2-2H + SO_3 **2**	U,F,Li
46	11.06	$C_{50}H_{58}O_{18}$	$[M-H]^-$	945.3588	4.3	927.3484 $[M-H-H_2O]^-$, 747.2638 $[M-H-C_{10}H_{14}O_4]^-$, 663.2255 $[M-H-C_{11}H_{22}O_8]^-$, **517.1858 $[M-H-C_{20}H_{28}O_{10}]^-$**, 505.1848 $[M-H-C_{20}H_{28}O_{10}-C]^-$	Unknown	F
M56	11.22	$C_{18}H_{18}O_7S$	$[M-H]^-$	377.0739	4.4	**297.1160 $[M-H-SO_3]^-$**, 253.1198 $[M-H-SO_3-CO_2]^-$	Enterlactone sulfate 1 (arctigenin-$3CH_2$-2O + SO_3)	U,F,K
M57	11.28	$C_{30}H_{34}O_{13}S$	$[M-H]^-$	633.1671	2.9	517.1912 $[M-H-SO_3-2H_2O]^-$, **505.1921 $[M-H-SO_3-2H_2O-C]^-$**, 411.1509 $[M-H-SO_3-C_7H_{10}O_3]^-$, 357.1391 $[M-H-SO_3-C_{10}H_{12}O_4]^-$	Lappaol E + SO_3	F

Table 1 (*continued*)

NO.	t_R (min)	Molecular formula	Selected ion	Measured ion	Error (mDa)	(ESI−/ESI+) MS fragmentations	Identification or characterization	Source
49	11.28	$C_{50}H_{58}O_{18}$	$[M-H]^-$	945.3565	2.0	927.3463 $[M-H-H_2O]^-$, 897.3375 $[M-H-2H_2O-C]^-$, 665.2451 $[M-H-C_{11}H_{20}O_8]^-$, 517.1912 $[M-H-C_{20}H_{28}O_{10}]^-$, **505.1929 $[M-H-C_{20}H_{28}O_{10}-C]^-$**	Unknown	F
M58	11.33	$C_{27}H_{30}O_{12}$	$[M-H]^-$	545.1704	4.5	**369.1299** $[M-H-GluA]^-$, 354.1186 $[M-H-GluA-CH_3]^-$, 203.1186 $C_{11}H_7O_4$, 174.0690 $C_{11}H_{10}O_2$, 159.0475 $C_{10}H_7O_2$	7,8-Didehydroarctigenin isomer + GluA	P,U
50	11.38	$C_{50}H_{58}O_{18}$	$[M-H]^-$	945.3561	1.6	897.3373 $[M-H-2H_2O-C]^-$, **665.2415 $[M-H-C_{11}H_{20}O_8]^-$**, 653.2426 $[M-H-C_{12}H_{20}O_8]^-$, 505.1906 $[M-H-C_{21}H_{28}O_{10}]^-$	Unknown	F
M59	11.39	$C_{18}H_{18}O_7S$	$[M-H]^-$	377.0741	4.6	**297.1163 $[M-HSO_3]^-$**, 253.1237 $[M-HSO_3-CO_2]^-$	Enterlactone sulfate 2 (arctigenin-3CH_2-2O + SO_3)	U,F
51	11.42	$C_{40}H_{46}O_{14}$	$[M-H]^-$	749.2842	3.3	731.2740 $[M-H-H_2O]^-$, 713.2694 $[M-H-2H_2O]^-$, 701.2621 $[M-H-2H_2O-C]^-$, **665.2447 $[M-H-4H_2O-C]^-$**, 653.2400 $[M-H-4H_2O-2C]^-$, 571.2007 $[M-H-C_7H_{10}O_3-2H_2O]^-$	Unknown	F
52	11.47	$C_{30}H_{34}O_{10}$	$[M-H]^-$	553.2106	3.2	517.1921 $[M-H-2H_2O]^-$, **505.1910 $[M-H-2H_2O-C]^-$**, 411.1481 $[M-H-C_7H_{10}O_3]^-$	Isolappaol C	P,U,F,Li
M60	11.54	$C_{21}H_{24}O_{10}S$	$[M-H]^-$	467.0995	−1.7	387.1483 $[M-HSO_3]^-$	Arctigenin + O + SO_3 **2**	F,Li
M61	11.59	$C_{20}H_{20}O_9S$	$[M-H]^-$	435.0799	4.9	355.1301 $[M-HSO_3]^-$	Arctigenin-CH_2-2H + SO_3 **3**	U,F,Li
54*	11.65	$C_{27}H_{34}O_{11}$	$[M-H+HCOOH]^-$	579.2104	2.6	**371.1502 $[M-H-Glc]^-$**, 357.1101 $[M-H-Glc-H_2O]^-$, 121.0287 $C_7H_5O_2$	Arctiin	P,F,Li,K
M62	11.69	$C_{30}H_{32}O_{12}S$	$[M-H]^-$	615.1511	−2.5	535.2059 $[M-HSO_3]^-$, 517.1884 $[M-HSO_3-H_2O]^-$, **505.1880 $[M-HSO_3-H_2O-C]^-$**, 475.1444 $[M-HSO_3-H_2O-3CH_2]^-$, 381.1364 $[M-HSO_3-C_8H_{10}O_3]^-$, 282.0869 $C_{17}H_{14}O_4$	Lappaol A isomer + SO_3 **1**	F
55	11.71	$C_{30}H_{34}O_{10}$	$[M-H]^-$	553.2114	4.0	517.1755 $[M-HSO_3-2H_2O]^-$, **505.1920 $[M-H-2H_2O-C]^-$**	Lappaol C	P,U,F,Li
M63	11.76	$C_{27}H_{32}O_{12}$	$[M-H]^-$	547.1818	0.2	**371.1509 $[M-H-GluA]^-$**, 357.1331 $[M-H-GluA-CH_2]^-$	Arctigenin + GluA	P,U,F,Li,Lu,K,B
57	11.85	$C_{21}H_{26}O_7$	$[M-H]^-$	389.1620	2.0	371.1501 $[M-H-H_2O]^-$, 345.1702 $[M-CO_2]^-$, **330.1486 $[M-H-COOH-CH_2]^-$**	Arctigenic acid	P,U,F,H,Li,Lu
M64	11.86	$C_{30}H_{32}O_{12}S$	$[M-H]^-$	615.1570	3.4	517.1903 $[M-HSO_3-H_2O]^-$, **505.1894 $[M-HSO_3-H_2O-C]^-$**, 282.0868 $C_{17}H_{14}O_4$	Isolappaol A + SO_3 **1**	F
58	11.88	$C_{50}H_{56}O_{17}$	$[M-H]^-$	927.3447	0.8	879.3217 $[M-H-2H_2O-C]^-$, 825.3047 $[M-H-5H_2O-C]^-$, **813.2912 $[M-H-5H_2O-2C]^-$**, 801.3015 $[M-H-5H_2O-3C]^-$	Unknown	F
M65	11.91	$C_{36}H_{40}O_{15}$	$[M-H]^-$	711.2339	5.0	505.1936 $[M-H-GluA-H_2O-C]^-$	Isolappaol A + GluA	P,U,Li
61	11.95	$C_{50}H_{56}O_{17}$	$[M-H]^-$	927.3455	1.6	879.3209 $[M-H-2H_2O-C]^-$, 843.2980 $[M-H-4H_2O-C]^-$, **813.2946 $[M-H-5H_2O-2C]^-$**, 801.2982 $[M-H-5H_2O-3C]^-$	Unknown	F
M66	12.01	$C_{30}H_{32}O_{12}S$	$[M-H]^-$	615.1555	1.9	**505.1817 $[M-HSO_3-H_2O-C]^-$**, 381.1395 $[M-HSO_3-C_8H_{10}O_3]^-$, 282.0761 $C_{17}H_{14}O_4$	Lappaol A + SO_3 **1**	F
M67	12.08	$C_{27}H_{32}O_{12}$	$[M-H]^-$	547.1843	2.7	**371.151535 $[M-H-GluA]^-$**, 357.1322 $[M-H-GluA-CH_2]^-$	Arctigenin isomer + GluA	P
M68	12.11	$C_{50}H_{58}O_{17}$	$[M-H]^-$	929.3621	2.5	**911.3586 $[M-H-H_2O]^-$**, 893.3672 $[M-H-2H_2O]^-$, 881.3408 $[M-H-2H_2O-C]^-$	Unknown	F
M69	12.15	$C_{36}H_{40}O_{15}$	$[M-H]^-$	711.2325	3.6	505.1938 $[M-H-GluA]^-$	Lappaol A + GluA	P,U,Li
M70	12.18	$C_{21}H_{24}O_9S$	$[M-H]^-$	451.1110	4.7	371.1494 $[M-HSO_3]^-$	Arctigenin + SO_3	P,U,F,Li,Lu,K
M71	12.22	$C_{20}H_{22}O_9S$	$[M-H]^-$	437.0929	2.3	357.1378 $[M-HSO_3]^-$	Arctigenin-CH_2 + SO_3 **6**	P,U,Li,Lu

Table 1 (*continued*)

NO.	t_R (min)	Molecular formula	Selected ion	Measured ion	Error (mDa)	(ESI−/ESI+) MS fragmentations	Identification or characterization	Source
66	12.25	$C_{50}H_{56}O_{17}$	$[M-H]^-$	927.3450	1.1	897.3360 $[M-H-H_2O-C]^-$, 825.2929 $[M-H-5H_2O-C]^-$, **813.2922 $[M-H-5H_2O-2C]^-$**	Unknown	F
67	12.36	$C_{21}H_{22}O_6$	$[M-H]^-$	369.1360	2.2	354.1140 $[M-H-CH_3]^-$, 203.0398 $C_{11}H_7O_4$, **174.0670 $C_{11}H_{10}O_2$**, 159.0449 $C_{10}H_7O_2$	7,8-Didehydroarctigenin isomer	F
68	12.36	$C_{50}H_{56}O_{17}$	$[M-H]^-$	927.3428	−1.1	909.3333 $[M-H-H_2O]^-$, 879.3340 $[M-H-2H_2O-C]^-$, 825.2908 $[M-H-5H_2O-C]^-$, **813.2943 $[M-H-5H_2O-2C]^-$**	Unknown	F
M72	12.37	$C_{30}H_{32}O_{12}S$	$[M-H]^-$	615.1479	4.3	517.1956 $[M-H-SO_3-H_2O]^-$, **505.1948 $[M-H-SO_3-H_2O-C]^-$**	Lappaol A isomer + SO_3 **2**	F
M73	12.57	$C_{30}H_{32}O_{12}S$	$[M-H]^-$	615.1580	4.4	517.1956 $[M-H-SO_3-H_2O]^-$, **505.1932 $[M-H-SO_3-H_2O-C]^-$**, 282.0826 $C_{17}H_{14}O_4$	Isolappaol A + SO_3 **1**	F
M74	12.59	$C_{21}H_{24}O_9S$	$[M-H]^-$	451.1106	4.3	371.1583 $[M-H-SO_3]^-$	Arctigenin isomer +SO_3	U,F
72	12.66	$C_{40}H_{46}O_{14}$	$[M-H]^-$	749.2836	2.7	597.2433 $[M-H-C_8H_8O_3]^-$, 165.0543 $C_9H_9O_3$, **151.0429 $C_8H_7O_3$**	Unknown	F
M75	12.77	$C_{27}H_{30}O_{12}$	$[M-H]^-$	545.1677	1.8	**369.1388 [M-H-GluA]$^-$**, 174.0641 $C_{11}H_{10}O_2$, 159.0413 $C_{10}H_7O_2$	7,8-Didehydroarctigenin + GluA	P,U
77	12.80	$C_{30}H_{32}O_{10}$	$[M-H]^-$	551.1954	3.7	**505.1921 $[M-H-H_2O-CO]^-$**, 397.1359 $[M-H-C_7H_{10}O_3-C]^-$	Arctignan C	F
M76	12.80	$C_{30}H_{32}O_{12}S$	$[M-H]^-$	615.1563	2.7	517.1924 $[M-H-SO_3-H_2O]^-$, **505.1918 $[M-H-SO_3-H_2O-C]^-$**, 282.0875 $C_{17}H_{14}O_4$	Lappaol A + SO_3 **2**	F
79	13.05	$C_{50}H_{56}O_{17}$	$[M-H]^-$	927.3483	4.4	**909.3328 $[M-H-H_2O]^-$**, 879.3239 $[M-H-2H_2O-C]^-$, 825.3026 $[M-H-5H_2O-C]^-$, 813.2978 $[M-H-5H_2O-2C]^-$	Unknown	F
80	13.07	$C_{40}H_{44}O_{13}$	$[M-H]^-$	731.2745	4.1	677.2421 $[M-H-3H_2O]^-$, **665.2379 $[M-H-3H_2O-C]^-$**, 653.2529 $[M-H-3H_2O-2C]^-$	Unknown	F
82	13.12	$C_{31}H_{36}O_{10}$	$[M-H]^-$	567.2280	5.0	**531.2053 $[M-H-2H_2O]^-$**, 519.2139 $[M-H-2H_2O-C]^-$	Lappaol D	P,U,F
M77	13.14	$C_{21}H_{24}O_7$	$[M-H]^-$	387.1466	2.2	**298.1193 $[M-H-C_3H_5O_3]^-$**, 283.1055 $[M-H-C_3H_5O_3-CH_3]^-$	Arctigenin + O	P,U
83	13.24	$C_{40}H_{44}O_{13}$	$[M-H]^-$	731.2754	5.0	713.2668 $[M-H-H_2O]^-$, 683.2544 $[M-H-2H_2O-C]^-$, 677.2445 $[M-H-3H_2O]^-$, **665.2429 $[M-H-3H_2O-C]^-$**, 653.2457 $[M-H-3H_2O-2C]^-$, 571.2076 $[M-H-C_7H_{10}O_3-H_2O]^-$, 559.2043 $[M-H-C_7H_{10}O_3-H_2O-C]^-$	Arctignan D	F
84	13.28	$C_{40}H_{44}O_{13}$	$[M-H]^-$	731.2736	3.2	713.2700 $[M-H-H_2O]^-$, 683.2442 $[M-H-2H_2O-C]^-$, 677.2417 $[M-H-3H_2O]^-$, **665.2424 $[M-H-3H_2O-C]^-$**, 653.2433 $[M-H-3H_2O-2C]^-$, 571.2093 $[M-H-C_7H_{10}O_3-H_2O]^-$, 559.2032 $[M-H-C_7H_{10}O_3-H_2O-C]^-$	Arctignan E	F
M78	13.34	$C_{21}H_{22}O_9S$	$[M-H]^-$	449.0942	3.6	369.1390 $[M-H-SO_3]^-$	7,8-Didehydroarctigenin + SO_3	F
86*	13.40	$C_{20}H_{22}O_6$	$[M-H]^-$	357.1361	−1.1	**342.1114 $[M-H-CH_3]^-$**, 313.1420 $[M-H-CO_2]^-$, 209.0842 $[M-H-C_8H_8O_2-C]^-$	Matairesinol	P,U,F,H,Li,S,Lu,K,B
M79	13.48	$C_{40}H_{46}O_{13}$	$[M-H]^-$	733.2889	2.9	715.2755 $[M-H-H_2O]^-$, 685.2692 $[M-H-2H_2O-C]^-$, 591.2262 $[M-H-C_7H_{10}O_3]^-$, **517.1920** $[M-H-C_{10}H_{16}O_5]^-$, 505.1934 $[M-H-C_{10}H_{16}O_5-C]^-$, 383.1521 $C_{22}H_{23}O_6$	Arctignan D + 2H	F
M80	13.51	$C_{20}H_{22}O_6$	$[M-H]^-$	357.1356	1.8	342.1165 $[M-H-CH_3]^-$	Arctigenin-CH_2 1 (isomer of Matairesinol)	U,F

Table 1 (*continued*)

NO.	t_R (min)	Molecular formula	Selected ion	Measured ion	Error (mDa)	(ESI−/ESI+) MS fragmentations	Identification or characterization	Source
M81	13.60	$C_{40}H_{46}O_{13}$	$[M-H]^-$	733.2892	3.2	715.2733 $[M-H-H_2O]^-$, 697.2676 $[M-H-2H_2O]^-$, 685.2706 $[M-H-2H_2O-C]^-$, 591.2296 $[M-H-C_7H_{10}O_3]^-$, 517.1870 $[M-H-C_{10}H_{16}O_5]^-$, **505.1880 $[M-H-C_{10}H_{16}O_5-C]^-$**, 383.1438 $C_{22}H_{23}O_6$	Arctignan E + 2H	F
M82	13.64	$C_{19}H_{20}O_5$	$[M-H]^-$	327.1261	2.9	191.0697 $[M-H-C_8H_8O_2]^-$, **121.0332 $C_7H_5O_2$**	(2*R*,3*R*)-2-(3-Hydroxybenzyl)-3-(3-hydroxy-4-methoxybenzyl)butyrolactone (arctigenin-$2CH_2$-O)	P,U,F,Li
M83	13.80	$C_{20}H_{22}O_6$	$[M-H]^-$	357.1347	0.9	342.1168 $[M-H-CH_3]^-$, **205.0453 $C_{11}H_9O_4$**	Arctigenin-CH_2 2 (isomer of matairesinol)	P,U,F,Li,K
M84	13.92	$C_{18}H_{18}O_4$	$[M-H]^-$	297.1114	−1.3	253.1200 $[M-H-CO_2]^-$	Enterlactone (arctigenin-$2CH_2$-2O)	F
M85	14.16	$C_{19}H_{18}O_6$	$[M-H]^-$	341.1036	1.1	267.0750 $[M-H-C_3H_6O_2]^-$	Arctigenin-$2CH_2$-2H	U,F
97	14.75	$C_{30}H_{32}O_9$	$[M-H]^-$	535.1989	2.1	517.1890 $[M-H-H_2O]^-$, **505.1926 $[M-H-H_2O-C]^-$**, 283.1099 $C_{17}H_{15}O_4$, 269.0774 $C_{16}H_{13}O_4$	Isolappaol A	P,U,F
98	14.84	$C_{30}H_{32}O_9$	$[M-H]^-$	535.1988	2.0	517.1889 $[M-H-H_2O]^-$, **505.1900 $[M-H-H_2O-C]^-$**, 381.1389 $[M-H-C_8H_{10}O_3]^-$, 282.0898 $C_{17}H_{14}O_4$	Lappaol A	P,U,F,Li
100	15.12	$C_{30}H_{32}O_9$	$[M-H]^-$	535.1957	−1.1	505.1964 $[M-H-H_2O-C]^-$	Lappaol A isomer	F
101	15.14	$C_{21}H_{22}O_6$	$[M-H]^-$	369.1374	3.6	174.0660 $C_{11}H_{10}O_2$	7,8-Didehydroarctigenin	P,U,F,Li
102[a]	15.17	$C_{21}H_{24}O_6$	$[M-H]^-$	371.1522	2.7	357.1301 $[M-H-H_2O]^-$, 174.0640 $C_{11}H_{10}O_2$, **121.0294 $C_7H_5O_2$**	Arctigenin	P,U,F,H,Li,S,Lu,K,B
105	15.62	$C_{40}H_{42}O_{12}$	$[M-H]^-$	713.2594	−0.4	683.2465 $[M-H-H_2O-C]^-$, **665.2446 $[M-H-2H_2O-C]^-$**, 653.2501 $[M-H-2H_2O-2C]^-$, 295.2203 $C_{18}H_{15}O_4$	Lappaol F	F
M86	16.14	$C_{48}H_{54}O_{18}$	$[M-H]^-$	917.3226	−0.6	743.3159 $[M + H-GluA]^+$, **725.2975 $[M + H-GluA-H_2O]^+$**, 507.2039 $[M + H-GluA-C_{13}H_{16}O_4]^+$, 489.2100 $[M + H-GluA-C_{13}H_{16}O_4-H_2O]^+$	Diarctigenin + GluA	F
114	19.19	$C_{42}H_{46}O_{12}$	$[M-H]^-$	741.2908	−0.3	**725.3047 $[M + H-H_2O]^+$**, 707.2900 $[M + H-2H_2O]^+$, 507.2056 $[M + H-C_{13}H_{16}O_4]^+$, 489.2108 $[M + H-C_{13}H_{16}O_4-H_2O]^+$, 285.1182 $C_{17}H_{17}O_4$, 271.1081 $C_{16}H_{15}O_4$	Diarctigenin	F

Notes. [a] Unambitiously identified by comparison with reference standards; P: plasma; U: urine; F: feces; H: heart; Li: liver; S: spleen; Lu: lung; K: kidneys; B: brain.

M15. Metabolite **M4**, with the formula of $C_{20}H_{24}O_{11}S$, had oxygen more than **M11**/**M13**/**M19**/**M21** metabolites. It was then characterized as demethyl-dehydroxy-arctigenic acid sulfate. Matairesinol underwent hydroxylation and glucuronidation or sulfation could result in phase I and II products, including metabolites **M8**, **M12**, **M14**, **M18**, and **M36**. Notably, arctigenin underwent demethylation coupled with glucuronidation and sulfation, resulting in **M39**/**M43**/**M51** and **M40**/**M44**/**M50**/**M52**/**M54**/**M71**, respectively. Metabolite **M46** was characterized as a demethylation production of metabolites **M39**/**M43**/**M51**. Furthermore, **M34** was a dehydrogenation product of **M46**. Metabolites **M49**/**M60** presented $[M-H]^-$ ion at *m*/*z* 467.10 ($C_{21}H_{23}O_{10}S$), which had 79.957 Da (SO_3) more than hydroxyl-arctigenin (**M77**); and then, they were characterized as the sulfation products of **M77**.

3.3.2.2. Sesquilignans-related metabolites. In the extract of FA, arctignan A/isolappaol C/lappaol C/lappaol E and lappaol A isomer/isolappaol A/lappaol A were the main sesquilignans. Thus, based on their structures, correlated metabolites were tentatively characterized in mice biosamples.

3.3.2.2.1 Phase I reactions-related. Metabolite **M10** presented $[M-H]^-$ ion at *m*/*z* 571.2222 (C30H35O11), and as compared to compound **29**, it had two additional hydrogens. Meanwhile, they had similar fragmentation behaviors, such as fragment ions at *m*/*z* 505.1823 ($C_{29}H_{29}O_8$) which were similar to fragment ions at *m*/*z* 503.1741 ($C_{29}H_{27}O_8$) produced by compound **29**. Thus, metabolite **M10** was characterized as a hydrogenation product of compound **29**. Metabolites **M31**/**M37** and

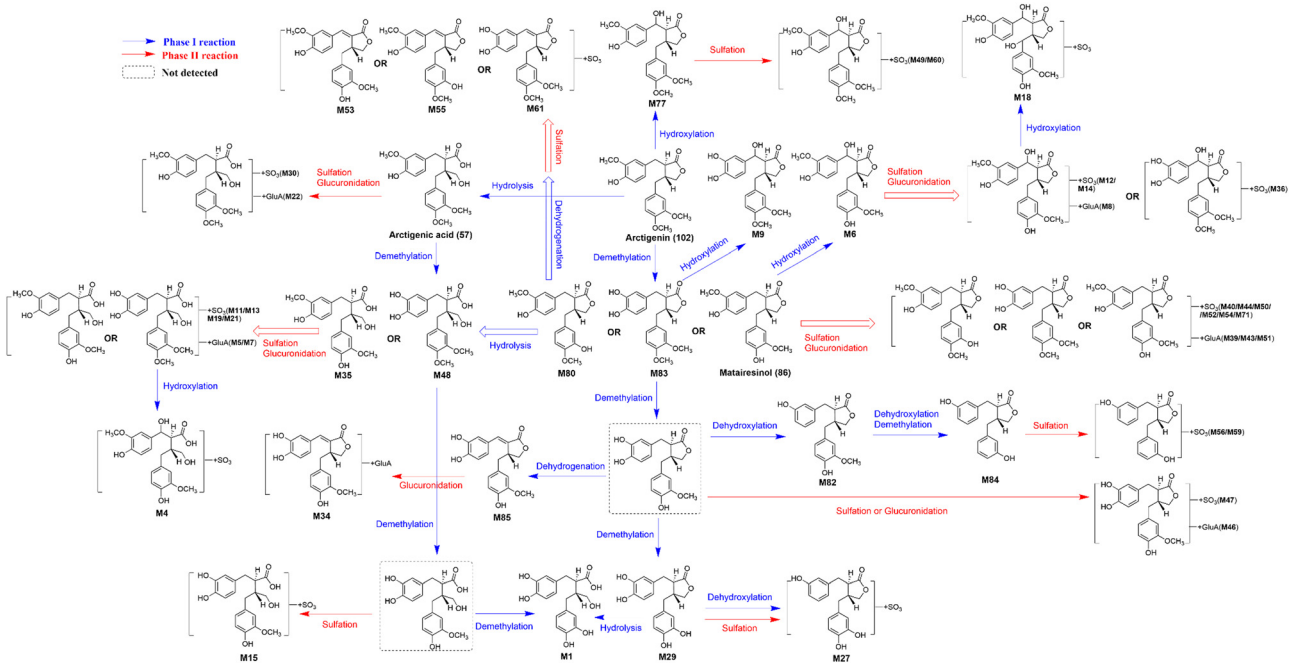

Fig. 7 Proposed metabolic pathways of arctigenin-related in H1N1 infected mice.

M38/**M42** were the glycosylation products and demethylation products of isolappaol C/lappaol C, respectively.

3.3.2.2.2 Phase II reactions-related. The metabolic reactions of sulfation and glucuronidation of arctignan A/isolappaol C/lappaol C/lappaol E could result in the metabolites **M17**/**M23**/**M26**/**M33**/**M41**/**M57** and **M24**/**M25**/**M28**/**M32**, respectively. Similarly, the sulfation or glucuronidation products of lappaol A isomer/isolappaol A/lappaolA were also detected, including **M62**/**M64**/**M66**/**M72**/**M73**/**M76** and **M65**/**M69**.

3.3.2.3 Dilignans-related metabolites. Five dilignans-related metabolites were characterized in mice samples, including four phase I metabolites and one phase II product. Metabolite **M16** exerted a deprotonated ion at *m*/*z* 765.2726 ($C_{40}H_{45}O_{15}$), which gained one oxygen more than lappaol H. Metabolite **M16** could produce fragment ions at *m*/*z* 747.2606 ($C_{40}H_{43}O_{14}$) $[M\text{-}H\text{-}H_2O]^-$, 681.2353 ($C_{39}H_{37}O_{11}$) $[M\text{-}H\text{-}4H_2O\text{-}C]^-$, and 575.1934 ($C_{32}H_{31}O_{10}$) $[M\text{-}H\text{-}C_7H_{10}O_4\text{-}H_2O\text{-}OCH_2]^-$, and this fragmentation behavior was similar to that of lappaol H (**40**) with the fragment ions at *m*/*z* 731.2716 ($C_{40}H_{43}O_{13}$) $[M\text{-}H\text{-}H_2O]^-$, 665.2383 ($C_{39}H_{37}O_{10}$) $[M\text{-}H\text{-}4H_2O\text{-}C]^-$, and 559.1967 ($C_{32}H_{31}O_9$) $[M\text{-}H\text{-}C_7H_{10}O_4\text{-}H_2O\text{-}OCH_2]^-$. Metabolite **M20** presented $[M\text{-}H]^-$ ion at *m*/*z* 911.3374 with the formula $C_{46}H_{56}O_{19}$. As compared to lappaol H, they had one more $C_6H_{10}O_5$ (162.053 Da). Meanwhile, the fragment ions of metabolite **M20** were the same as lappaol H. Thus, metabolites **M16** and **M20** were characterized as the hydroxylation product and the glycosylation product of lappaol H, respectively. Metabolites **M79**/**M81** exerted $[M\text{-}H]^-$ ion at *m*/*z* 733.28 ($C_{40}H_{45}O_{13}$) and fragment ions at *m*/*z* 715.27 ($C_{40}H_{43}O_{12}$) and 685.27 ($C_{39}H_{41}O_{11}$). Compared to arctignan D/arctignan E, metabolites **M79**/**M81** had similar fragmentation behavior, and the difference was that **M79**/**81** gained two additional hydrogens. Then, metabolites **M79**/**M81** were characterized as the hydrogenation products of arctignan D/arctignanE.

3.3.2.4 New lignans-related metabolites. Only one new-lignans-related metabolite was characterized. **M68** presented $[M\text{-}H]^-$ ion at *m*/*z* 929.3621 ($C_{50}H_{57}O_{17}$), and compared to compounds **58**/**61**/**66**/**68**/**79**, **M68** gained two additional hydrogens. Then, **M68** was characterized as the hydrogenation product of compounds **58**/**61**/**66**/**68**/**79**.

3.3.3 The metabolic character of FA in H1N1-induced mice model. Gennearly, absorbed constituents were the primary substances to achieve therapeutic effects. As result, 124 xenobiotics (38 prototypes and 86 metabolites) in H1N1 infected mice were detected and characterized, including 47 in plasma (11 prototypes and 36 metabolites), 101 in feces (38 prototypes and 63 metabolites), 68 in urine (10 prototypes and 58 metabolites), 14 in the lung (3 prototypes and 11 metabolites), 8 in the heart (3 prototypes and 5 metabolites), 32 in the liver (8 prototypes and 24 metabolites), 3 in the spleen (2 prototypes and 1 metabolite), 18 in kidneys (3 prototypes and 15 metabolites), and 4 in the brain (2 prototypes and 2 metabolites). Meanwhile, the metabolic reactions of FA *in vivo* were also summarized, including phase I and phase II reactions. It was found that the main phase I metabolic reactions of monolignans and their glycosides were demethylation, dehydroxylation, demethoxylation, and hydroxylation. Demethylation and hydroxylation are major phase imetabolic reactions of sesquilignans. Meanwhile, phase II reactions of these two types included sulfation and glucuronidation. The phase I metabolic reactions of dilignans included hydroxylation and hydrogenation reactions, and the phase II reaction was glucuronidation.

H1N1 virus could cause lung edema, inflammation, *etc.*, which significantly impacting the host's survival condition. Thus, the absorbed compounds (prototypes or metabolites) in the lung might be major therapeutic substances. In the lung, 14 metabolites were characterized including matairesinol + O (**M9**), arctigenic acid-CH_2 + SO_3 (**M19**), arctigenin-CH_2 + GluA (**M39**/**M43**), arctigenin-2CH_2 + SO_3 (**M47**), arctigenin-CH_2 + SO_3 (**M44**/**M50**/**M52**/**M71**), arctigenin + GluA (**M63**), arctigenin + SO_3 (**M70**), arctigenic acid (**57**), matairesinol (**86**), and arctigenin (**102**). Based on these metabolites, the compounds, which could present anti-influenza effects in FA might bearctigenic acid, matairesinol, and arctigenin.

3.4 Network pharmacology analysis

Absorbed constituents were the primary substances for further pharmacological mechanisms. Thus, the prototypes and metabolites were subjected to the target prediction.

A total of 592 targets from 52 constituents (8 prototypes and 44 metabolites), with a probability of more than 0.1, were obtained through the Swiss Target Prediction. Then, the

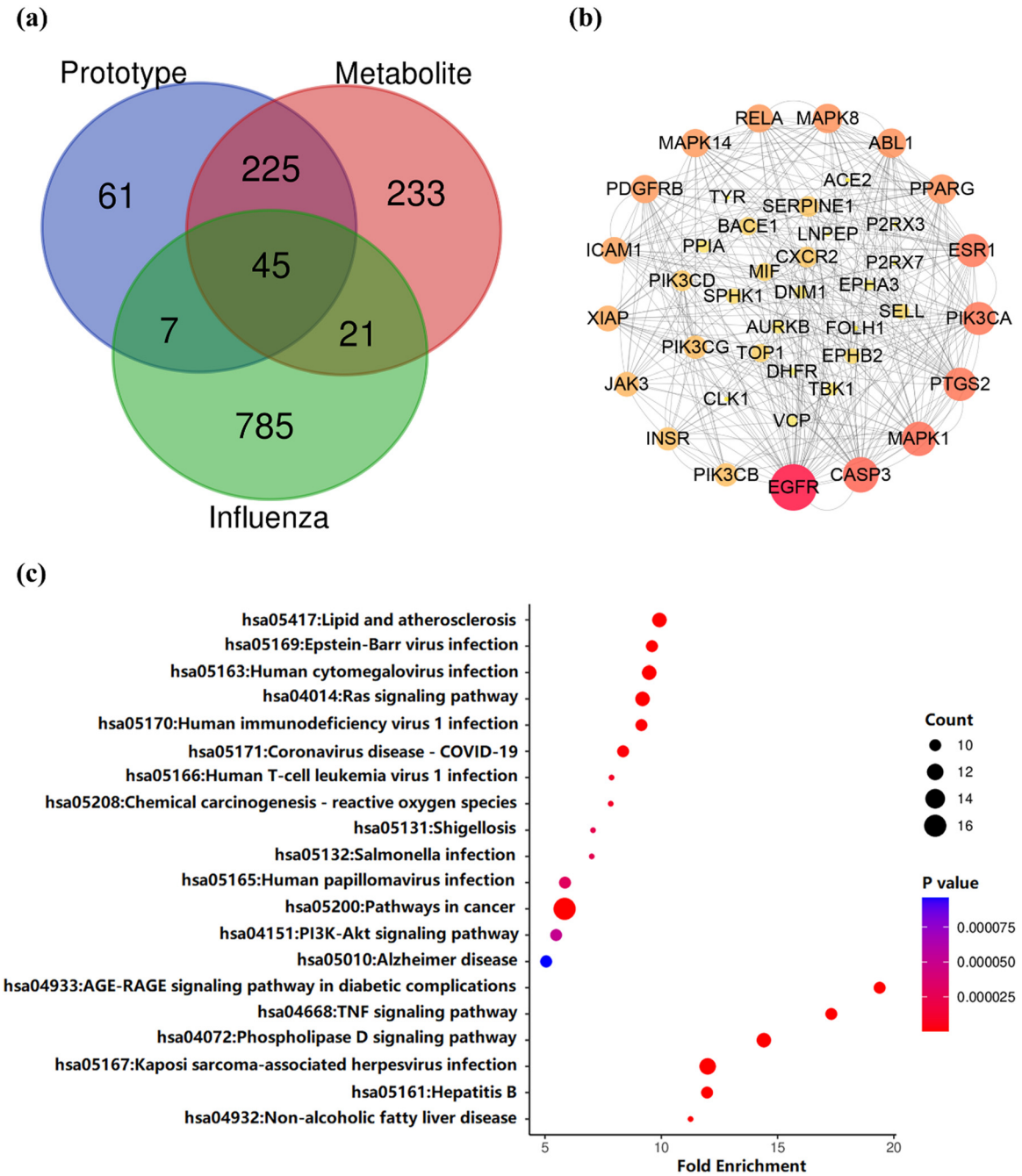

Fig. 8 The predicted targets and pharmacological information of FA. (a) The overlapped targets between influenza-related targets and targets of FA; (b) protein–protein interactions among targets of overlapped targets; (c) top 20 KEGG pathways.

compound-target network was constructed using Cytoscape software. The network consisted of 645 nodes and 1040794 interactions. Meanwhile, the influenza-related targets, with 'Score_gda' more than 0.1, were obtained in DisGeNet, and 858 targets were obtained. As shown in Fig. 8a, 45 overlapped targets between FA-related candidates (prototypes and metabolites) and influenza-related targets were found, and these could be treated as the core targets of FA in the treatment of influenza, including EGFR, CASP3, MAPK1, PTGS2, PIK3CA, ESR1, *etc.* Further, these targets, with a degree value of more than 3 in Fig. 8b, were used for constructing protein–protein interactions (PPI) and KEGG analysis. Meanwhile, the protein–protein interaction network was also constructed and the main targets from the PPI network were screened out by analyzing their degree value in the network. In the PPI network, PI3K complex (PIK3CA, PIK3CB), MAPK complex, EGFR, and CASP3 were the main targets for FA to achieve the anti-influenza effects. As reported in the literature, the phosphatidylinositol-3-kinase (PI3K) was identified to be activated upon influenza A virus (IAV) infection, and the early and transient induction of PI3K signaling is caused by viral attachment to cells and promotes virus entry.[42] This gave a direction that FA might depress the entry of the influenza virus to achieve therapeutic effects.

3.4.1 KEGG classification of target proteins. The relationship between target proteins and the pathways was analyzed using the data extracted from the DAVID database, and the pathways were screened according to the KEGG analysis with BH-corrected *P*-values less than 0.05. The results showed that the main target proteins were mainly involved in pathways in cancer, kaposi sarcoma-associated herpesvirus infection, ras signaling pathway, human cytomegalovirus infection, TNF signaling pathway, PI3K-Akt signaling pathway, *etc.* (Fig. 8c). Among them, the PI3K-Akt signaling pathway (correlated to PI3K complex) and TNF signaling pathway are regarded as the core pathways for FA for treating influenza.

Finally, the reported components with anti-influenza ability in FA had lignans and caffeoylquinic acids. Lignans were one of the major anti-influenza materials in FA. For example, arctiin and arctigenin, two lignans in FA, were reported to present obvious anti-influenza *in vitro* and *in vivo*. Hayashi *et al.* reported that arctiin and arctigenin exerted protective effects in immunocompetent and immunocompromised mice infected with influenza A virus.[43] Meanwhile, caffeoylquinic acids in FA also presented anti-influenza effects, such as chlorogenic acid and 3,4-dicaffeoylquinic acid. It has been reported that chlorogenic acid could inhibit neuraminidase activity and block the release of newly formed virus particles from infected cells.[44] 3,4-Dicaffeoylquinic acid could increase TRAIL expression and extend the lifetimes of mice infected with the influenza A virus.[45] These data indicate that lignans and caffeoylquinic acids are the core materials for FA to achieve its anti-influenza effects. The predicted mechanism, such as PI3K signaling and TNF signaling, might be regulated by lignans and caffeoylquinic acids in FA. Notably, this hypothesis should be validated by more experiments in the future.

4. Conclusion

Revealing the functional mechanism of complex systems (such as TCM) is quite difficult since there are hundreds and thousands of chemical components in TCM, and exploring their functional mechanism is of great importance to practical applications. Oral administration was the primary way for TCM access to the public, and notably, the absorbed compounds in normal and infected hosts were quite different. Thus, the characterization of *in vivo* substances under the disease model was an effective way to reveal their functional mechanism. In our work, a strategy integrating pharmacological evaluation, chemical profiling, and network pharmacology was proposed to reveal the potential pharmacological mechanism of TCM, and Fructus Arctii was used as an example. As a result, FA could significantly improve the survival rate of H1N1-infected mice with amelioration of lung inflammation. Meanwhile, a total of 123 compounds (62 targeted components and 61 non-targeted components) were identified or tentatively characterized in FA, providing the material basis for further quality evaluation or exploring the pharmacological mechanism. Furthermore, 124 xenobiotics were tentatively characterized in nine biosamples (38 prototypes and 86 metabolites). Considering the metabolites in mice's lungs, the anti-influenza effects of FA might mainly be arctigenic acid, matairesinol, and arctigenin. All xenobiotics were screened out in mice's plasma to construct the pharmacological network. The network pharmacology results showed that FA could target EGFR, CASP3, MAPK1, PTGS2, PIK3CA, ESR1, *etc.*, which are mainly involved in the pathways in cancer, kaposi sarcoma-associated herpesvirus infection, ras signaling pathway, PI3K-Akt signaling pathway, *etc.* Among them, the PI3K-Akt signaling pathway (correlated to PI3K complex) and TNF signaling pathway are regarded as the core pathways for FA for treating influenza. It was the first time to evaluate anti-influenza effects and potential mechanisms of FA and obtain the chemical information of FA *in vitro* and *in vivo*. In the meantime, new insight for exploring the potential pharmacological mechanism of TCM has been provided.

Author contributions

Min Li: Conceptualization, formal analysis, methodology, writing-original draft preparation, funding acquisition. Zhong-Yu Huang: conceptualization, formal analysis, methodology. Yu-lin-lan Yuan and Shuang-Shuang Cui: visualization, investigation. Hai-Jun Li: writing-review andediting. Feng-Xiang Zhang: writing-review and editing, supervision.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This work was supported by the Hainan Provincial Natural Science Foundation of China (Grant No. 822QN315), the Youth

Cultivation Fund of the First Affiliated Hospital of Hainan Medical University (Grant No. HYYFYPY202002), and the Natural Science Foundation of China (No. 82004256).

References

1 F. X. Zhang, Z. T. Li, M. Li, Y. L. L. Yuan, S. S. Cui, J. X. Chen and R. M. Li, *J. Pharm. Biomed. Anal.*, 2021, 193.
2 Q. H. Ma, M. Y. Ren and J. B. Luo, *J. Ethnopharmacol.*, 2021, **264**, 112800.
3 J. Li, M. Li, K. Ye, Q. Jiang, M. Wang, X. Wen and J. Yang, *J. Ethnopharmacol.*, 2021, **267**, 113517.
4 F. X. Zhang, Y. L. L. Yuan, S. S. Cui, M. Li, X. Tan, Z. C. Qiu and R. M. Li, *Food Funct.*, 2021, **12**, 4325–4336.
5 G. B. Zhang, Q. Y. Li, Q. L. Chen and S. B. Su, *J. Evidence-Based Complementary Altern. Med.*, 2013, 621423.
6 H. Hao, X. Zheng and G. Wang, *Trends Pharmacol. Sci.*, 2014, **35**, 168–177.
7 S. Zheng, J. P. Baak, S. Li, W. Xiao, H. Ren, H. Yang, Y. Gan and C. Wen, *Phytomedicine*, 2020, **79**, 153336.
8 H. T. Zhang, Y. Y. Gao, J. Y. Zhang, K. Wang, T. Jin, H. Y. Wang, K. F. Ruan, F. Wu and Z. H. Xu, *J. Ethnopharmacol.*, 2020, 255.
9 R. Jaiswal and N. Kuhnert, *Food Funct.*, 2011, **2**, 63–71.
10 J. Xia, Z. Guo, S. Fang, J. Gu and X. Liang, *Foods*, 2021, 10.
11 H. Zhang, Y. Wang, K. Ruan and Z. Xu, *Traditional Chinese Drug Research and Clinical Plarmacology*, 2020, **31**, 163–168.
12 R. Guo, M. Zhao, H. Liu, R. Su, Q. Mao, L. Gong, X. Cao and Y. Hao, *Biomedicine*, 2021, **141**, 111676.
13 J. Wang, Y. Zhang, J. Fan and H. Li, *Chin. Tradit. Pat. Med.*, 2019, **41**, 1505–1510.
14 K. Hayashi, K. Narutaki, Y. Nagaoka, T. Hayashi and S. Uesato, *Biol. Pharm. Bull.*, 2010, **33**, 1199–1205.
15 X. Xu, H. Chen, Y. Zheng and Z. Zhou, *J. Southwest Univ. Natl., Nat. Sci. Ed.*, 2017, **39**, 172–176.
16 D. S. Ming, E. S. Guns, A. Eberding and G. H. N. Towers, *Pharm. Biol.*, 2004, **42**, 44–48.
17 J. He, X.-Y. Huang, Y.-N. Yang, Z.-M. Feng, J.-S. Jiang and P.-C. Zhang, *J. Asian Nat. Prod. Res.*, 2016, **18**, 423–428.
18 B. Yang, H. Xin, F. Wang, J. Cai, Y. Liu, Q. Fu, Y. Jin and X. Liang, *J. Sep. Sci.*, 2017, **40**, 3231–3238.
19 Y. Yang, X. Huang, W. Wang, N. Du, J. Zhang, Z. Feng, J. Jiang and P. Zhang, *Acta Pharm. Sin. B*, 2017, **52**, 779–784.
20 X.-Y. Huang, Z.-M. Feng, Y.-N. Yang, J.-S. Jiang and P.-C. Zhang, *J. Asian Nat. Prod. Res.*, 2015, **17**, 504–511.
21 Y.-N. Yang, F. Zhang, Z.-M. Feng, J.-S. Jiang and P.-C. Zhang, *J. Asian Nat. Prod. Res.*, 2012, **14**, 981–985.
22 J. Hu, L. Ma, T. Yao, C. Ran, H. Liu, H. Li and H. Lei, *J. Fruit Sci.*, 2012, **29**, 905–910.
23 K. Qin, Q. Liu, H. Cai, G. Cao, T. Lu, B. Shen, Y. Shu and B. Cai, *Pharmacogn. Mag.*, 2014, **10**, 541–546.
24 J. Liu, Y.-Z. Cai, R. N. S. Wong, C. K.-F. Lee, S. C. W. Tang, S. C. W. Sze, Y. Tong and Y. Zhang, *J. Agric. Food Chem.*, 2012, **60**, 4067–4075.
25 K. Qin, B. Wang, W. Li, H. Cai, D. Chen, X. Liu, F. Yin and B. Cai, *J. Sep. Sci.*, 2015, **38**, 1491–1498.
26 F.-X. Zhang, Z.-T. Li, M. Li, Y.-L.-L. Yuan, S.-S. Cui, G.-H. Wang and R.-M. Li, *J. Chromatogr. B: Anal. Technol. Biomed. Life Sci.*, 2020, 1157.
27 Q. Gao, Y. Zhang, S. Wo and Z. Zuo, *Planta Med.*, 2013, **79**, 471–479.
28 F. He, D.-Q. Dou, Y. Sun, L. Zhu, H.-B. Xiao and T.-G. Kang, *Planta Med.*, 2012, **78**, 800–806.
29 F. X. Zhang, Z. T. Li, X. Yang, Z. N. Xie, M. H. Chen, Z. H. Yao, J. X. Chen, X. S. Yao and Y. Dai., *J. Ethnopharmacol.*, 2021, **268**, 113660.
30 A. Daina, O. Michielin and V. Zoete, *Nucleic Acids Res.*, 2019, **47**, W357–W364.
31 J. Pinero, A. Bravo, N. Queralt-Rosinach, A. Gutierrez-Sacristan, J. Deu-Pons, E. Centeno, J. Garcia-Garcia, F. Sanz and L. I. Furlong, *Nucleic Acids Res.*, 2017, **45**, D833–D839.
32 D. Szklarczyk, A. L. Gable, D. Lyon, A. Junge, S. Wyder, J. Huerta-Cepas, M. Simonovic, N. T. Doncheva, J. H. Morris, P. Bork, L. J. Jensen and C. Mering, *Nucleic Acids Res.*, 2019, **47**, D607–D613.
33 P. Shannon, A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski and T. Ideker, *Genome Res.*, 2003, **13**, 2498–2504.
34 D. W. Huang, B. T. Sherman and R. A. Lempicki, *Nucleic Acids Res.*, 2009, **37**, 1–13.
35 P. B. Ampomah and L. H. K. Lim, *Apoptosis*, 2020, **25**, 1–11.
36 F.-x Zhang, Z.-t Li, X. Yang, Z.-n Xie, M.-h Chen, Z.-h Yao, J.-x Chen, X.-s Yao and Y. Dai, *J. Ethnopharmacol.*, 2021, 268.
37 F.-X. Zhang, M. Li, Z.-H. Yao, C. Li, L.-R. Qiao, X.-Y. Shen, K. Yu, Y. Dai and X.-S. Yao, *Biomed. Chromatogr.*, 2018, 32.
38 F.-x Zhang, Z.-t Li, C. Li, M. Li, Z.-h Yao, X.-s Yao and Y. Dai, *J. Pharm. Pharmacol.*, 2020, **72**, 1879–1892.
39 A. Ichihara, Y. Numata, S. Kanai and S. Sakamura, *J. Chin. Agric. Chem. Soc.*, 1977, **41**, 1813–1814.
40 D. Wang, A. S. Badarau, M. K. Swamy, S. Shaw, F. Maggi, L. E. da Silva, V. Lopez, A. W. K. Yeung, A. Mocan and A. G. Atanasov, *Front. Plant Sci.*, 2019, 10.
41 F.-x Zhang, M. Li, L.-r Qiao, Z.-h Yao, C. Li, X.-y Shen, Y. Wang, K. Yu, X.-s Yao and Y. Dai, *J. Pharm. Biomed. Anal.*, 2016, **122**, 59–80.
42 E. R. Hrincius, R. Dierkes, D. Anhlan, V. Wixler, S. Ludwig and C. Ehrhardt, *Cell. Microbiol.*, 2011, **13**, 1907–1919.
43 K. Hayashi, K. Narutaki, Y. Nagaoka, T. Hayashi and S. Uesato, *Biol. Pharm. Bull.*, 2010, **33**, 1199–1205.
44 Y. Ding, Z. Cao, L. Cao, G. Ding, Z.-Z. Wang and W. Xiao, *Sci. Rep.*, 2017, **7**, 45723.
45 T. Takemura, T. Urushisaki, M. Fukuoka, J. Hosokawa-Muto, T. Hata, Y. Okuda, S. Hori, S. Tazawa, Y. Araki and K. Kuwata, *J. Evidence-Based Complementary Altern. Med.*, 2012, 946867.

