Published in final edited form as:

J Chromatogr A. 2019 March 29; 1589: 105–115. doi:10.1016/j.chroma.2018.12.061.

A systemic workflow for profiling metabolome and lipidome in tissue

Dasheng Lu^{a,b,c}, Liming Xue^a, Chao Feng^a, Yu'e Jin^a, Chunhua Wu^b, Cen Xie^c, Frank J. Gonzalez^c, Guoquan Wang^{a,*}, and Zhijun Zhou^{b,*}

^aShanghai Municipal Center for Disease Control and Prevention, 1380 Zhongshan West Road, Shanghai, 200336, China

^bSchool of Public Health/MOE Key Lab for Public Health, Fudan University, Shanghai, 200032, China

^cLaboratory of Metabolism, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD, 20892, USA

Abstract

Simple metabolome and lipidome sample preparation procedures involving two successive extractions using small pieces of tissue, and a subsequent metabolite identification (MetID) strategy were developed. The sample preparation can significantly circumvent incomplete analysis due to insufficient amounts of tissue as a result of splitting into several aliquots for multiple measurements, with advantages over the similar previously reported methods in metabolite coverage, extraction efficiency, method robustness and friendly experimental operation. A MetID strategy, based on the integration of MS information mining (including adduct ions, in-source CID, MS information from both ESI (+) and ESI (–), characteristic fragmentation ions (CFIs), constant neutral losses (CNLs) and multimers) and *in silico* MS simulation, was demonstrated. A large number of adduct ions (83 features), in-source CID (123 features), ESI (+/–) ionization (20 features), CFIs& CNLs (more than 120 features) and multimers (17 features) were mined by manually or *in silico* recognition/filtering, which provide the most suspicious structures for subsequent *in silico* MS simulation. The unknown features presented the same score distribution as the known (83 features) features with scores > 25% (geomean score: 52%) and with satisfactory match for the main ions of interest. The MS/MS noise and fragment ions of coeluted quasi-molecular ions of interest are the main reason for the low score in the simulation. Manual check/evaluation is always suggested for the simulation with a score less than 50%. This strategy presents satisfactory performance with 2.5 times more metabolites structurally characterized compared with that of the traditional method based on accurate-mass-based MS and MS/MS library matching. This strategy would be useful for potentially identifying metabolites without available MS/MS information in the library.

*Corresponding authors. wangguoquan@scdc.sh.cn (G. Wang), zjzhou@fudan.edu.cn (Z. Zhou).

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.chroma.2018.12.061>.

Keywords

Metabolome; Lipidome; *In silico* MS simulation; MS information mining

1. Introduction

Metabolomics identifies and quantifies the complements of all small molecules (termed metabolome) within a biological system [1–3], and lipidomics is a subset of metabolomics devoted to lipids (also called lipidome) [2]. Recently, metabolomics has emerged as a cornerstone in the field of systems biology [3]. Previous studies revealed that metabolomics is a powerful technique that might be of value to gain pathophysiological knowledge of various diseases/injuries to help clinicians in the diagnostic process, especially to determine the etiology or severity of the disease and to assess therapeutic response or predict drug induced injuries and corresponding mechanism [1,4–6]. Metabolomics has integrated into or even become an important alternative to the traditional methods of diagnostics at the molecular and cellular level based modern methods in toxicology [1,7].

Apart from the commonly used biofluids such as urine and blood [8,9], tissues are also used as sample matrices to explore the systematic (whole body)/local (organs) metabolic perturbation associated with body/organ function and injury, screen biomarkers and explore the mechanism related to the perturbation [10,11]. More importantly, its metabolite coverage not only involves the hydrophilic but also the hydrophobic compounds such as lipids, which make possible more comprehensive metabolite coverage. For a comprehensive study, an integrated strategy inclusive of metabolomics and molecular/cellular, biochemical and histological analysis is needed to investigate disease or toxicity mechanism. However, the limited amount of tissue is sometimes not sufficient, especially for biopsy tissue from patients and tissues available from small animal models, such as mice, worms, and flies [6,12,13]. Therefore, in a metabolomics study, it is necessary to develop an efficient, tissue-saving experimental pipeline to cover metabolome and lipidome. Metabolome and lipidome components exhibit different physicochemical properties, resulting in generally using two different sample preparation techniques for their separate extraction and enrichment [14–16]. For the metabolome, a chilled mixture of methanol (MeOH) and water (normally 1/1, v/v) is generally used as an extraction solvent [17,18]. For the lipidome, a chilled mixture of methanol, water and chloroalkane (normally including dichloromethane (DCM) and trichlorometane (TCM)) or the chilled methyl tert-butyl ether (MTBE) are commonly used as the extraction solvent [4,14,15,17,18]. Thus, more tissue would be consumed if these two methods were used both for the extraction of metabolome and lipidome.

Recently, two methods, as the most representative studies, were reported to use a single small tissue sample to simultaneously prepare metabolome and lipidome extracts for subsequent analysis [12,16]. One method used MTBE-MeOH-H₂O as the extraction solvent for homogenization and subsequently liquid-liquid extraction (LLE) by adding water to induce a separation between metabolome and lipidome [12]. This method yielded results for the lipidome but less information involving the total metabolome. Additionally, the tissue used in this report was freeze-dried powder instead of the widely adopted wet sample.

Another method was based on two successive extractions [16]. Namely, MeOH-H₂O (1:1, v/v) and DCM-MeOH (3:1, v/v) were used for separate extraction by use of a bead beater. This method presented satisfactory results for both the metabolome and lipidome. However, DCM is highly toxic. It was widely reported with complex extraction operation and even with dripping losses because it needs to pass through the upper residue layer to transfer the lower solvent layer out after LLE phase separation [4,14,15,17,18]. This drawback could be potentially corrected using an alternative solvent-MTBE [4,5,14,15,17,18]. However, to date, no comprehensive comparison has been done involving these extraction solvents in lipidomics.

Metabolite identification (MetID) is a bottleneck and time-consuming work in non-target metabolomics [19–21]. It's expensive to purchase all authentic standards for metabolite confirmation and some compounds are not available commercially or have very limited availability [19,20]. According to a survey on the main endogenous metabolite suppliers (including IROA, Sigma-Aldrich and Cambridge Isotope Laboratory), there are less than 900 standards commercially available, which is much less than 2651 and 4229 of metabolites respectively reported in urine and blood [8,9]. Therefore, the most economic, feasible and practical method for identification of the unknown metabolite would primarily be based on mass spectrometric data. The canonical MetID is based on accurate-mass-based MS and MS/MS matching using in-house and online library, which is limited by the availability of experimental MS/MS spectra in public libraries. Metlin has the most comprehensive metabolite library [19], yet among over 1 million molecules, only a small percentage (16,000 of metabolites, drugs, xenobiotics, and toxicants) have available MS/MS data [19]. mzCloud is an Orbitrap MS based MSⁿ library and now it only involves 7859 compounds inclusive of 1203 endogenous metabolites. Therefore, the main MetID obstacle in metabolomics has now shifted from identifying molecules with known MS/MS spectra to identifying molecules that do not exist in the databases or that exist yet do not have experimental MS/MS information [19,20]. MS information mining and *in silico* MS simulation were developed as the mainstream techniques of biotransformation and MetID of exogenous chemicals (such as pharmaceuticals and drugs) [22,23]. These techniques have the potential to be explored for MetID in metabolomics studies [19,21]. Recently, a large number of *in silico* MS/MS fragmentation algorithms (including MetFrag, MIDAS, MAGMa, CSI:FingerID, CFM-ID, FingerID, MassFrontier (HighChem), MS-Fragmenter (ACD/Labs), Molecular Structure Correlator (Agilent) and so on) have been available [24] and metabolome databases based *in-silico* MS simulation [20] and Metlin based fragment similarity search [19] have also been proposed. However, publicly available studies are still very limited and, more importantly, lack a systematical combination of these techniques to optimize their MetID functions in metabolomics.

This study is aimed to develop a simultaneous extraction method to cover metabolome and lipidome using a single small piece of tissue and to propose a MetID strategy. The method was validated and evaluated by comparison with previously reported methods using rat liver as a representative tissue sample. The MetID strategy was focused on the integration of MS information mining (including adduct ions, in-source CID, MS information from both ESI (+) and ESI (–), characteristic fragmentation ions (CFIs), constant neutral losses (CNLs) and multimers) and *in silico* MS simulation, facilitating the identification of metabolites and other small molecules that have no library MS/MS data.

2. Experimental

2.1. Materials and preparation

Methanol (MeOH) and acetonitrile (ACN) were HPLC-grade and purchased from Merck (Darmstadt, Germany). Dichloromethane (DCM, pesticide residue grade), methyl-tert-butyl methyl ether (MTBE, HPLC grade) and isopropanol (IPA, HPLC grade) were from Tedia Company, Inc. (Fairfield, Ohio, USA). Ammonium formate (purity 99.0%) and formic acid (FA, purity 98.0%) were purchased from Sigma-Aldrich with LC-MS ultra-grade. Ultra-pure water was purchased from Wastons (A.S. Watson Group Ltd., Hong Kong). Millex-GV (PVDF, 0.22 μ m) was purchased from Minipore (Millipore Corp., USA). Besides the Mass Spectrometry Metabolite Library (IROA Technologies, Boston, USA), other authentic metabolite standards were mainly purchased from Sigma-Aldrich (St. Louis, USA), Cambridge Isotope Laboratory (MA, USA) and Avanti Polar (Alabaster, AL, USA). Authentic standard solutions were prepared by dissolving these standards into appropriate solvent or solvent mixture and consequently diluting to 1 μ g/mL and stored at -20°C .

2.2. Experimental animals and sample collection

Wistar rats (200–260 g, 6- to 8-weeks-old) used in this study were supplied by the Shanghai Institute of Pharmaceutical Industry and all animal studies were performed under the approval of the Institutional Authority for Laboratory Animal Care. The rats were housed in an environmentally controlled room at room temperature on a 12 h light/dark cycle, with food and water provided *ad libitum*. After fed with a standard diet for 10 days to adapt to the environment, the rats were killed and the liver tissues immediately excised and frozen, lyophilized and stored at -80°C until chemical analysis.

2.3. Sample preparation

For sample pretreatment, two successive extractions were employed. Liver samples, were thawed, and 208–315 mg of them sampled from the right liver lobe were transferred to a 1.5 mL tube with five steel beads for hydrophilic metabolite extraction. An ice-cold mixture of MeOH- H_2O (1/1, v/v) was added with a volume equivalent to 28 mg tissue/100 μ L solvent. After freezing in liquid nitrogen for 30 s, the samples were homogenized based on the vibration of the bead in the JXFSTPRP-192 Bead Beater (Shanghai Jingxin Technology Company, China) under the following conditions: 20 s at 60 Hz of vibration frequency. This homogenization was repeated twice. After homogenization, the centrifugation (4°C) was performed for 20 min at 13,000 rpm. The hydrophilic extracted solvent (50 μ L) was totally transferred to another tube and 50 μ L of it was evaporated to nearly dryness using a vacuum concentrator (Thermo Fisher Scientific, USA, with the operation parameters of 45°C , 0.1 psi and 3000 rpm) and subsequently reconstituted into 250 μ L of 95% ACN aqueous solution. After shaking for 30 s and sonication for 1 min, the extracts were centrifuged for 20 min at 13,000 rpm and 4°C . The supernatant was filtered using Millex-GV, transferred into injection bottle and finally stored at -20°C until instrumental analysis.

After the hydrophilic metabolite extraction, the residual pellet was further used for hydrophobic metabolite extraction. The ice-cold MTBE was used as solvent with the same volume and extraction process (including homogenization, centrifugation, evaporation,

reconstitution and storage) as the above-mentioned hydrophilic metabolite extraction except that H₂O-ACN-IPA (1/1/2, v/v/v) was used as the reconstitution solvent. For comparison, a similar extractions method reported based on DCM-MeOH (3/1, v/v) was also carried out in parallel to evaluate the performance of TBME-based extraction method.

2.4. UHPLC-HRMS analysis

Analysis was performed using an UltiMate 3000 Hyperbaric LC system coupled to a Q Exactive MS (Thermo Fisher Scientific, USA). For hydrophilic metabolites, HILIC chromatographic separation was carried out using an Acquity UHPLC BEH HILIC column (100 mm × 2.1 mm, 1.7 μm, Waters, Milford, MA, USA). Mobile phase A (ACN-water, 95/5, v/v) and mobile phase B (ACN-water, 50/50, v/v) both contained with 0.1% formic acid and 10 mM ammonium formate were utilized at a flow rate of 0.4 ml/min. Mobile phase gradient was as follows: 0–2 min, 95% A; 2–40 min, linear from 95% to 5% A; and 40–45 min, 95% A for equilibration. For hydrophobic metabolites, a reverse phase chromatographic separation was applied using Acquity CSH C18 (2.1 × 100 mm, 1.7 μm, Waters). Mobile phase A (ACN-water, 60/40, v/v) and mobile phase B (IPA-ACN, 90/10, v/v) were used at a flow rate of 0.4 ml/min. Mobile phase gradient was as follows: 0–2.0 min, 40%–43% B; 2.0–2.1 min, 43%–50% B; 2.1–12.0 min, 50%–54% B; 12.0–12.1 min, 54%–70% B; 12.1–18.0 min, 70%–99%B; 18.0–18.1 min 99%–40% B; and 18.1–20.0 min, 40% B. The temperatures of column oven were set at 35 and 55 °C for the analysis of hydrophilic and hydrophobic metabolites, respectively. Autosampler were set at 4 °C. Injection volume was 5 μl.

The ESI source was operated in the positive/negative mode with the following parameters: capillary temperature, 350 °C; source voltage and spray voltage, 3.8 KeV (positive)/–3.3 KeV (negative) and 3.3 KeV (positive)/–2.5 KeV (negative) for hydrophilic and hydrophobic metabolites, respectively; sheath gas (nitrogen) flow, 40 arb (arbitrary units); and aux gas flow, 10 arb. Data was acquired using full MS scan (resolution: 70,000; AGC target: 1×10^6 ; Maximum IT: 120 ms; scan range: m/z 55–825 (for metabolome)/120–1800 (for lipidome) and collision induced dissociation (CID)-based data dependent MS/MS (DDMS²) acquisition (resolution: 175,000; AGC target: 1×10^5 ; Maximum IT: 120 ms; Loop count: 5; TopN = 5; Isolation window: m/z 1.0; Scan range: m/z 55–825; NCE/stepped NCE: 30, 40, 50; Underfill ratio: 1.0%; intensity threshold: 8.3×10^3 ; Apex trigger: 2–6 s; Dynamic exclusion: 6 s). Except for MetID samples whose data were acquired using full MS scan plus DDMS² mode or single ion monitoring (SIM) mode when DDMS² acquisition isn't triggered in the mode of full MS scan plus DDMS², other samples were run using the full MS scan mode.

2.5. Quality control

The liver samples were immediately frozen at –80 °C after collected. The sample preparation was performed at 4 °C and sample extracts stored at –20 °C. UHPLC-HRMS maintenances was performed prior to running samples. To cover the low-molecular metabolites with high MS accuracy, in-house tune calibration (caffeine (2 μg/ml) for ESI (+), acetic acid (1 μg/ml) and 3-methylglutaric acid (1 μg/ml) for ESI (–)) was executed. TICs of each scan have less than 15% of deviation in peak area. The maintenance needs to

ensure the test sample (generally the QC samples) to provide less than 20% area-deviation when using three mass tolerance (MT) values (normally including 1, 4 and 10 ppm) to extract typical diagnostic ions. Before running the samples, the HPLC was conditioned by repeatedly running blanks or QC samples until less than 5% retention time (RT) drift was obtained.

The samples were run based on the following sequence: solvent blank, procedure blank, 5 QC samples, MetID samples, diluted QC samples, 12 test samples, QC sample, 12 test samples and finally QC sample and solvent blank. The test samples were randomized. QC samples were prepared by mixing the appropriate volume (generally 5 μ l) of each test sample. The diluted QC samples, applied to remove peaks with poor concentration-based linear-dynamic range, were prepared by diluting the QC sample with a diluted factor of 2, 5 and 10. Similar to the preparation of QC, the MetID samples were prepared by combining the samples from the same group (based on the experimental design) and run in full-scan 2 DDMS data acquisition mode for subsequent MetID.

2.6. Data pretreatment and metabolite identification strategy

For the hydrophilic metabolites, the raw data of all samples were uploaded into Compound Discoverer 2.1 (Thermo Fisher Scientific, USA) for spectrum properties filter ($S/N > 3$), peak alignment (RT ± 0.2 min and MT ± 4 ppm), feature extraction (Adduct ions: $[M+H]^+$, $[M+NH_4]^+$ and $[M-nH_2O+H]^+$ for ESI (+) and $[M-H]^-$, $[M+HCOO]^-$ and $[M-nH_2O-H]^-$ for ESI (-), peak with $S/N \geq 10$ and intensity $\geq 1 \times 10^6$) and feature grouping (RT ± 0.2 min and MT ± 4 ppm). The processed data of both ionization modes were combined and exported into a Microsoft Excel spreadsheet for data filtering. The filtering was performed to reduce noise influence, remove duplicates and avoid over-fitting in the applied multivariate models with the following parameters: peak area ($\geq 1 \times 10^6$), procedure blank interference (Ratio of peak area between test sample and procedure blank < 5), inter-group deviation (RSD $\leq 45\%$ for inter-group samples), repeatability (RSD $\leq 30\%$ for QC), linearity dynamics ($R \geq 0.7$ for diluted QC) and duplicated features in ESI ($\pm$) ionization modes (RT < 0.1 min and mass deviation of precursor ions < 0.001 Da in two ionization modes). To reduce the potential influence of statistical analysis due to the greatly different variance of individual metabolite concentrations, the filtered data was subjected to logarithm transformation and normalization (Pareto Scaling) followed by principal component analysis (PCA) and orthogonal partial least squares-discriminant analysis (OPLS-DA) using Simca 4.1 (Umetrics, Umea, Sweden).

After multivariable analysis, the specific features with value of Variable Importance in the Projection (VIP) equal to or higher than 2 were structurally characterized based on our MetID strategy (Fig. 1). Accurate mass MS matching (MT ± 5 ppm) and isotopic distribution recognition ($\geq 50\%$) based on online and offline databases (including HMDB, KEGG, ChemSpider, BioCyc, Endogenous Metabolite database 4400 compounds provide by Thermo Fisher Scientific and RefMet (<http://metabolomicsworkbench.org/databases/refmet/browse.php>)) were first used to screen for potential metabolites. Once the feature was identified as a potential metabolite, MS/MS matching (MT ± 8 ppm) based on the offline Metlin PCDL Manager (Agilent Technologies, USA), online Metlin, mzVault and mzCloud

and in-house database (containing 601 compounds) was executed. The mzCloud and mzVault match were performed base on Similarity Forward method and HighChem-HighRes search algorithm, respectively. A MS/MS match value (0–100%) between the MS/MS for an unknown feature and the matching library spectrum would be displayed.

After mass matching identification, the remaining unidentified feature/metabolites were further identified using a novel method integrating MS information mining (including MS information of two ionizations, adduct ions and in-source CID, CFIs, CNLs and multimers) and *in silico* MS simulation. This process was performed using Compound Discoverer 2.1 and Massfrontier 7.0. The recognition of CID products was performed semi-automatically by comparing the RT, TIC shape and MS² between the suspicious in-source CID peak and its corresponding parent peak. CFIs similarity matching was performed by comparing the unknown feature spectrum and their similar spectrum in mzCloud library based on Fragment Ion search (FISh) technology provided by Compound Discoverer and Massfrontier. The multimers were identified by matching with their corresponding identified-monomer based on CFIs similarity and RT matching. Besides these abovementioned techniques, the traditional CFIs/CNLs recognition was used manually for MS information mining, especially for phase II conjugated metabolites and typical lipids (Table S1, Supplementary materials). The possible structure that predicted in MS information mining was tried to assign to the unknown features based on *in silico* MS simulation. The *in silico* MS simulation is conducted based on fragmentation prediction and FISh algorithm of Massfrontier. A similarity matching value (0–100%) expressed in CFIs matching and *in silico* MS simulation was used to evaluate MetID confidence.

For lipid metabolites, the MetID samples were imported into Lipidsearch 4.1 (Thermo Fisher Scientific, USA) for MetID. The MetID were performed using *in silico* MS simulation with the following parameters: 5 and 8 ppm of MT for the precursor and product ions, respectively; product ions with response 5% higher than their corresponding precursor ions; 5.0 (suggested by the vendor) of m-Score (expressed the number of matches with predicted product ion peaks in the MS/MS spectrum) threshold; ion adducts (for ESI (+), [M + H]⁺, [M + NH₄]⁺ and [M + Na]⁺; for ESI (–), [M – H][–], [M + FA – H][–]) and default lipid sub-classes.

3. Results and discussion

3.1. Metabolite profiling in different liver segment

Different segments in one organ can have different biological functions [10]. For example, liver is a complex tissue with eight functional segments functionally and morphologically segregated according to vascular supply [25–27]. Different liver segments have their own vascular inflow, outflow and biliary/lymphatic drainage and different levels of enzymatic systems [10,26,27], resulting in different chemical compositions in different hepatic zones and consequently have potentially distinct effects on discrimination analysis and metabolic pathway studies in metabolomics [10]. However, the past literature seldom mentions the endogenous metabolite profiling differences due to the different organ segments sampled in a batch of samples.

A pilot study was done using rat liver tissue as the representative sample. The right lobe of liver was cut into 12 pieces along the direction of the right hepatic vein. These samples were performed instrumental analysis blindly. PCA scoring plot of the 12 samples and corresponding QC are shown in Fig. S1 (Supplementary materials). The plot shows differences between the superior (L-8-L12) and inferior subsegments (L1-L7). Further analysis was done by analyzing the peak area deviation of features extracted from these samples (Fig. S2, Supplementary materials). Totally 550 and 348 features were obtained in ESI (+) and ESI (–) ionization mode, respectively. Among these features, the majority (88% and 82% for ESI (+) and ESI (–), respectively) are satisfactory as revealed by relative standard deviation (RSD) $\leq 30\%$, with a minority questionable ($2 < \text{RSD} \leq 3$, 5% and 8% for ESI (+) and ESI (–), respectively) and unsatisfactory (RSD ≥ 3 , 7% and 10% for ESI (+) and ESI (–), respectively). Here, the liver used was normal liver. If the livers were from rats under treatment (such as exposure to toxicant), or a disease model, the RSD value would potentially be larger because different liver segments could have different focal effects. These questionable and unsatisfactory features may potentially result in, in worst cases, uninterpretable metabolic pathways which can't be rescued even by the most sophisticated multivariate analysis. Therefore, the tissue is suggested to be sampled from the same location within an organ, especially for metabolomics studies where only a small piece of tissue is destined for analysis.

3.2. One successive extraction plus phase separation vs. two successive extraction

The one step extraction (using MTBE-MeOH-H₂O) plus phase separation (by adding water) method previously reported presents excellent lipidome results but less total metabolome information from freeze-dried tissue [12]. To further investigate its application in metabolomics and its feasibility in the widely used sample state-wet tissue, this reported method was repeated here. The traditional metabolome extraction method (MeOH-H₂O (1/1, v/v) as extraction solvent) was also parallelly performed as a control to evaluate the reported method's performance in metabolome extraction. For the control, 1142 metabolome features were acquired in ESI ($\pm$); among them, 1026 features successfully passed the data filtering as validated features in data pretreatment. Compared with the control, only 71% of the metabolome features were detected in the reported method and among these detected features, the majority present at much lower intensity. The possible reason is that MTBE owns middle polarity which makes it possible for the middle-polar metabolites to preferentially be transferred from the water phase to the organic phase during phase separation. Because of this reason, the current study aimed to improve this method based on two successive extractions, namely, to firstly use MeOH-H₂O (1/1, v/v) to extract metabolome and then to use the left residues to extract lipidome using MTBE.

3.3. Comparison of hydrophobic extraction between the MTBE and DCM-MeOH methods

To evaluate the performance of the two successive extractions method, another two successive extractions using MeOH-H₂O and DCM-MeOH (3:1, v/v) to separately extract hydrophilic and hydrophobic metabolite [16], was performed along with two successive extractions. After data filtering the data from two methods, PCA multivariable analysis and comparisons were conducted, along with response intensity and response deviation of features. The scoring plots of PCA shows that the DCM-MeOH method presented less

scattered sample points than that of MTBE (Fig. S3, Supplementary materials). There are totally 370 and 1691 feature separately acquired in the ESI (–) and ESI (+) mode in our proposed method. The sum number of these features is 2.7 times than that obtained in DCM-MeOH methods (330 and 635 features in ESI (–) and ESI (+), respectively (Fig. 2A), which is also well demonstrated in total ion chromatography (TIC) (Fig. 3). Among these features, except 891 (ESI (–): 319 and ESI (+): 572) features obtained in both methods, there are 74 (ESI (–): 11 and ESI (+): 63) and 1170 (ESI (–): 51 and ESI (+): 1119) features obtained from DCM-MeOH method and our proposed method, respectively (Fig. 2C). 662 lipid compounds (ESI (–): 388 and ESI (+): 274) were identified in MTBE method, which is more than 2 times of that of DCM-MeOH method (ESI (–): 200 and ESI (+): 122) (Fig. 4). A major difference existed in lipid number focused on Cer, DG, PC, PE, CL, PG, DMePE, PS and SM (Fig. 4).

In addition, this proposed method also presents higher feature intensity as compared with the DCM-MeOH method. Among the 891 features coextracted in both methods, 69% have higher response in the MTBE method and 25% have higher response in the DCM-MeOH method (Fig. 2D). After structural characterization, just like the feature results, all lipid species except LPI and LPS, present higher response in the MTBE method as compared with the DCM-MeOH method, especially for Cer, DG, PC, PE, CL, PG, DMePE, PS and SM (Fig. 4). These data suggest that our proposed method has better performance than the DCM-MeOH in tissue dispersion and lipid solubility. Lipids readily form aggregates (such as micelles) once the extraction solvent becomes more polar [28,29]. Therefore, after the MeOH-H₂O extraction in this study, the lipids of the extracted residues potentially form micelles/bilayers with their hydrophilic tails inside and their polar head outside [28,29]. Meanwhile, a water film may be formed outside of micelles based on the hydrogen bond activation between water molecule and lipid polar heads. The micelle/bilayer and film may potentially result in lipid solubility issues in subsequent lipid extractions. MTBE (4.8g/100 g H₂O) has higher aqueous-solubility than DCM (1.32 g/100 g H₂O), and therefore MTBE is more readily disrupts the film carrier and dissolves the lipids. This may be the potential reason why the MTBE method presents better lipid extraction performance than the DCM method.

In extraction repeatability comparison, three ranks based on the response RSD among samples, are defined as follows, satisfactory: RSD ≤ 20%, questionable: 20 < RSD ≤ 30% and unsatisfactory: RSD > 30%. Comparing with the extraction repeatability of DCM-MeOH method, our proposed method presents slightly better results in ESI (–) mode with 91%, 4% and 5% of satisfactory, questionable and unsatisfactory peaks (corresponding values of DCM-MeOH method are 88%, 4% and 5%), respectively, but slightly worse results in ESI (+) mode with 75%, 11% and 14% of satisfactory, questionable and unsatisfactory peaks (corresponding values of DCM-MeOH method are 86%, 8% and 6%), respectively (Fig. 2B). In conclusion, the discussion above shows that our MTBE based two successive extraction presented better performance than DCM-MeOH method.

3.4. Metabolite identification

In this study, 743 ESI (+) and 283 ESI (–) metabolome features were well validated from data process, of which, 527 (ESI (+)) and 213 (ESI (–)) features were identified as the suspicious metabolites using accurate-MS matching and isotopic distribution recognition. These metabolites were further identified using accurate MS/MS matching of in-house and online databases. After the MS/MS matching, only 26% of the suspicious metabolites (130 and 63 metabolites for ESI (+) and ESI (–), respectively) were identified as the putative metabolites (Table S2, Supplementary materials). These used MetID techniques are all canonical and completely dependent on the metabolite MS database. Till now, the metabolite MS database is limited, especially for the extremely limited experimental-MS/MS data which always plays a critically important role in MetID in metabolomics [19]. That is the reason why such a low identification rate was obtained in the present study.

The unidentified features were further structurally characterized using the integrated technique based on MS information mining (including MS information of two ionizations, adduct ions and in-source CID and characteristic fragmentation ions (CFIs)) and *in silico* MS simulation. Totally, there were 349 and 125 metabolites tentatively identified in ESI (+) and ESI (–) in this study, which proves the advantages over the canonical MS matching in facilitating the identification of small molecules that have no library MS/MS data.

3.4.1. MS information mining

3.4.1.1. Adduct ions. The existence of adduct ions will increase the complexity of MetID and misinterpretation of metabolic pathways. The commercial and free *in silico* tools can discriminate these adduct ions by identifying CNLs existed between quasi-molecular ion and its coeluted adduct ions, facilitating MetID of structurally unassigned features [19,21]. For part of metabolites that aren't recognized by the tools due to only adduct ions presented in MS ionization, manual discrimination and analysis is always needed. In this study, formic acid (FC) and ammonium formate were used as buffer of LC mobile phase and therefore the adduct ions $[M + NH_3 + H]^+$ (48 ions) and $[M + FC - H]^-$ (22 ions) were widely detected in ESI (+) and ESI (–), respectively. $[M + Cl]^-$ (5 ions), $[M + ACN + H]$ (4 ions) and $[M + Na]^+$ (4 ions) were also formed. Among the identified features with adduct ions, feature L642 (Fig. 5A) and L641 (Fig. 5B&C) are the representative MetID example. Both features were observed with the same RT (1.879 min) but a molecule of HCl of difference in molecular formula (L642: $C_9H_{12}N_2O_6$ based on m/z 243.0618 ($[M - H]^-$), L641: $C_9H_{13}ClN_2O_6$ based on m/z 279.03903 ($[M - H]^-$)). Feature L641 was also observed with a specific Cl isotopic distribution (Fig. 5B) and base peak ion (m/z 243.0618, Fig. 5C) same to quasi-molecular ion of feature L642. These observations could be useful to predict that feature L641 is derived from the adduct ion $[M + Cl]^-$ of feature L642. Feature L642 was putatively predicted as uridine based on accurate mass MS matching. Uridine was inputted for *in silico* MS simulation and its fragment ions are well matched with the query spectrum of L641 (Fig. 5C). Furthermore, this prediction was further confirmed by CFIs similarity matching and the matching result shows that the query spectrum of L641 were well matched with the uridine reference spectrum of mzCloud (Fig. 5C). This identification was finally confirmed using uridine as a standard reference (Fig. S4, Supplementary materials).

3.4.1.2. In-source CID MS.: Just like the adduct ions, in-source CID MS data will provide rich and useful information for structural characterization of feature. Unlike adduct ions, the in-source CID MS information can't be recognized using *in silico* tools because it's complex and highly dependent on a metabolite-itself, consequently limiting corresponding algorithm to be written. In this study, there are totally 123 features identified with the in-source CID MS information. Among them, dehydration (62 features) is widely occurred and its MS information would be useful for MetID as demonstrated in feature L333 (m/z 274.20105, $[M+H]^+$), L335 (m/z 256.19055, $[M-H_2O+H]^+$) and L334 (m/z 238.18008, $[M-2H_2O+H]^+$) (Fig. S5A, Supplementary materials). Feature L335 and L334 were identified to be respectively formed by losing H_2O and $2H_2O$ from feature L333 during in-source CID. The ion $[M+H]^+$ doesn't produce an MS/MS in DDMS² but $[M-2H_2O+H]^+$ does. The m/z 58.06593 and 142.08614 are both typical CFIs of carnitine. Based on the carnitine parent structure and the predicted molecular formula ($C_{14}H_{27}NO_4$) of L333, three possible structures can be predicted based on the derivation of hydroxyl (ester or ether) and carboxyl (ether) group of the carnitine. This prediction along with the occurrence of $[M-2H_2O+H]^+$ makes it easy to conclude that L333 can be tentatively identified as heptanoylcarnitine (HMDB0013238). The MS *in silico* simulation confirmed the conclusion with all L333 MS peaks well assigned to fragmentation of heptanoylcarnitine (Fig. S5B, Supplementary materials). Additionally, the L333 MS peaks are matched with the predicted heptanoylcarnitine LC-MS/MS spectrum provided by HMDB.

Feature L204 (MW: 171.1008Da, Fig. 6B), L206 (MW: 157.1216 Da, Fig. 6C), L205 (MW: 115.0636, Fig. 6D) and L203 (MW: 132.0901, Fig. 6E) exhibited a typical instance for MS information mining and MetID and all were identified as in-source CID product of feature L201 (MW: 202.1430 Da, Fig. 6A) by comparing their RT (8.67 min) and MS/MS information (Fig. 6G-I). After accurate mass-based MS matching plus isotopic distribution recognition and adduct-ion filtering, the feature L201 was well matched with asymmetric dimethylarginine (ADMA) and its isomer symmetric dimethylarginine (SDMA). Based on this mined MS information, *in silico* MS simulation was successfully performed with each ion from feature L201 and its in-source CID products well matched the predicted MS fragmentation of ADMA or SDMA (Fig. 6G-I), which supports the putative identification that feature L201 and its in-source CID products all are originated from a same metabolite, ADMA or SDMA.

3.4.1.3. MS information from both positive and negative ionization modes.: Some metabolites possess structures allowing ionization in both ESI (+) and ESI (-) modes. In this study, there are totally 20 features with both positive and negative ionization acquisition data. The data would potentially provide informative MS data for mining, consequently facilitating MetID, as demonstrated here for feature L202 acquired in ESI (-). Feature L202 has the same molecular mass and RT as feature L201 which was tentatively identified as ADMA or SDMA in ESI(+) (Fig. 6F). The *in silico* MS simulation results of L202 further confirmed this identification (Fig. 6J). Unfortunately, no MS fragment difference was *in silico* simulated between the isomers of ADMA and SDMA. Finally, authentic ADMA standard confirmed the identification that the feature L201/L202 is ADMA (Fig. S6, Supplementary materials).

3.4.1.4. CFIs similarity matching and CNLs recognition. The differential metabolism network regulated by exogenous and endogenous factors, especially the slight changes in intermediary metabolism will result in the diversity of structural analogs/derivatives of metabolite [30,31]. Metabolite analogs/derivatives may present same/similar CFIs or CNLs due to similarities in structure [32], which makes it possible to use CFIs and CNLs of a metabolite in the MS library to identify its analogs/derivatives that are not in the MS library [21]. Conjugated metabolites belong to derivatives with the same conjugated function group. The typical conjugates with glucuronide, sulfate and taurine were identified in this study. Each conjugate presents CNLs and/or CFIs associated with its corresponding conjugated function group (Table S1), which makes for easy identification. For conjugates of glucuronides, they present the neutral loss of m/z 176.03209 Da in ESI (+) and CFIs m/z 175.02379, 113.0240 and 85.02792 in ESI (–) as demonstrated in feature L87 (MW: 336.1423 Da) (Fig. S8A, Supplementary materials). Based on the identification of a glucuronide conjugate, the MW of the aglycon of L87 is 160.11021 Da (336.1423–176.03209 (neutral loss)). This mined information in combination with MS/MS information facilitate to conclude that L87 can be tentatively identified as glucuronidation of 2-hydroxycaprylic acid (HMDB02264) using accurate mass MS and MS/MS matching coupled with *in silico* MS simulation (Fig. S8A). Similarly, another 2 glucuronide conjugates were tentatively identified.

For conjugates of sulfate, the constant neutral loss of SO_3 moiety (79.9563 Da) and CFIs of m/z 79.95570 and 96.95856 will occur in ESI (–), just like the feature L284 (MW: 262.0145) (Fig. S8B, Supplementary materials) as demonstrated here. Considering the neutral loss (79.9563 Da), L284 can be predicted as a metabolite produced by sulfation of an intermediate with a MW 182.0588. The intermediate was putatively identified as either homovanillic acid (HMDB0000118) or dihydrocaffeic acid (HMDB00423) in HMDB database based on accurate mass matching. For the former, its sulfate (2-hydroxy-4-methoxyacetophenone-5-sulfate) will potentially yield the fragmentation ion m/z 137.02442 with a mass deviation of 26 ppm compared with queried ion m/z 137.05951 (Fig. S8B). However, for the latter, its sulfate can be well *in silico* simulated with all predicted fragment ions well matched with the query ions (Fig. S8B). More importantly, the query ions m/z 137.05951 and 181.06168 are existed in mass spectra of dihydrocaffeic acid (MID: 3789, Metlin MS/MS library) using CFIs similarity matching. Like this, the rest 8 features with sulfate CFIs were identified as sulfate conjugates in this study.

For taurine conjugates, m/z 126.02194 and m/z 124.00590, 106.97923 and 79.9556 are CFIs in ESI (+) and ESI (–), respectively. This is demonstrated in feature 2208N45d (MW: 265.0621 Da). L581 is formed by amination between a reactive intermediate and taurine with a loss of H_2O and the corresponding MW of the intermediate is 158.058 Da (265.0621–125.01466 (taurine) + 18.01056 (H_2O)) (Fig. S8C, Supplementary materials). The intermediate was well matched with succinylacetone (HMDB00635) in HMDB based on accurate mass matching. L581 can be tentatively identified as tauro-succinylacetone. Similarly, Other remaining 38 features with tauro-CFIs were also tentatively identified as taurine conjugates.

Besides the conjugates, the other metabolites can also be identified based on CFIs similarity matching. L53 owns the quasi-molecular ion m/z 448.08078 and no matched metabolite was successfully obtained based on the accurate-mass-based MS and MS/MS matching using in-house and on-line metabolite database. L53 mainly possesses four CFIs: m/z 143.04488, 177.03284, m/z 226.04800 and 270.0L201 for matching (Fig. S8D, Supplementary materials). Based on CFIs similarity matching, the former two ions were identified to belong to CFIs of *N*-glycylcysteine (HMDB0028838) and the other two ions were CFIs of dianthramine (identified previously, C0847). *N*-glycylserine and dianthramine can form several structures with the equal MW as L53 when a molecule of H₂O is lost. These possible structures were performed *in silico* MS simulation and the results show that 2-((2-carboxy-5-((glycyl-L-cysteinyl) oxy) phenyl) amino)-4-hydroxybenzoic acid is the optimized structure with all its possible fragment ions well matched with the query ion of L53. Therefore, L53 was tentatively identified as 2-((2-carboxy-5-((glycyl-L-cysteinyl) oxy) phenyl) amino)-4-hydroxybenzoic acid.

Multimers were reported to preferably form as sample concentration is high and their MetID based on accurate MS and MS/MS matching often leads to false annotation [21]. CFIs similarity matching, as a technique performed by relying on the fragment ions to identify similar structures within the library instead of on the precursor mass, is very useful for MetID of multimers when being in combination with RT matching. This MetID was demonstrated in feature L253 (Fig. S7–C, Supplementary materials) which was identified as creatinine dimer because it has the same CFIs (CFI similarity score: 89) and RT as creatinine (feature L252, Fig. S7A, Supplementary materials) as well as 2 times of MW as creatinine (Fig. S7–B/D (Supplementary materials): standard MS spectra from mzCloud library). Just like MetID of L253, the other 16 dimers and trimers were also identified with CFI similarity scores ranged from 85% to 100% in this study (Table S2).

3.4.2. In silico MS simulation—Totally, 347 unknown features were performed *in silico* MS simulation by inputting the most possible structures predicted using MS information mining into Compound discoverer/Massfrontier for fragmentation prediction and FISH (Table S2). To assess the performance of this simulation, part of identified features (totally 83 features) were conducted the simulation along with the unknown features. The application and advantage of *in silico* MS simulation were well demonstrated in the above section-MS information mining, highlighting its omnipotence in MetID of features without experimental MS/MS information. For the unknown features, the matching scores are ranged from 25% to 100% (geomean: 52%) (Table S2). The distribution of these scores is generally same to that obtained from the identified features (matching scores ranged from 26% (2-Pyrimidine Acetic Acid) to 100%, geomean: 53%), suggesting that these scores lower than 60% obtained from the unknown feature don't mean their low confidence or even failure in MetID. *In silico* MS simulation is a complicated process and highly dependent on the accurate fragmentation prediction and the purity of MS/MS acquisition spectrum. Our data shows that for the most of simulations, the score 25% generally would represent a satisfactory annotation for the main interested ions. The low score value is mainly due to the unmatched noise ions and the fragmentation ions of the coeluted quasi-molecular ions of interest. The co-eluted ions will be indiscriminately selected along with the quasi-molecular

ions of interest in the unit-resolution quadrupole mass analyzer for subsequent MS/MS in Orbitrap mass analyzer, which brought extra MS/MS fragment ions into the spectral of corresponding features of interest, resulting in the MS/MS impurity [22]. Additionally, the simulation is failed to identify the features (e.g. isomers) owning two or more suspicious structures with the undifferentiated MS/MS spectrum. In conclusion, to avoid the false positive MetID in *in silico* MS simulation, manual check/evaluation is always suggested for the simulation with a score less than 50%.

Lipidsearch 4.1 SP1 was used for lipidome identification in this study. Among 1691 features acquired based on full scan plus DDMS² acquisition in ESI ($\pm$), totally 662 lipid compounds (ESI (-): 388 and ESI (+): 274, Fig. 2) were identified with a 39% of identification ratio (Fig. 4). The ratio is dependent on the number of features with acquired MS/MS data and therefore if SIM acquisition mode were involved in these features of our interest, the corresponding ratio would be increased. These identified lipids presented the satisfactory identification quality with 35%, 42% and 22% that are completely assigned with lipid class and fatty acid (FA), lipid class and some FA, and lipid class or FA and with 1% identified by other fragment ions (e.g., H₂O loss). Lipidsearch provides more than 1.7 million lipid ions and their predicted fragment ions for database query, which facilitates totally 34 subclasses of lipids identified in this study (Figs. 3,4).

4. Conclusions

To reduce tissue consumption for those samples with too little amount to afford large number of measurements and more importantly to enhance the method throughput in metabolomics and lipidomics, a successive two extractions method based on two extraction solvent system (MeOH-H₂O and MTBE for metabolome and lipidome, respectively) and low-temperature homogenization using bead beater was developed using a representative tissue sample-liver. To evaluate the performance of this method, two similar methods, including one step extraction plus phase separation method and two successive extractions (using DCM as extraction solvent of lipidome) method, were performed in parallel. The results showed that our extraction method presented better performance in metabolite coverage and extraction efficiency as compared with these two reported methods, and in experimental operation convenience in contrast to the DCM extraction based two successive extraction. At the same time, considering that different organ segments potentially own different bio-functions and pathological changes, the corresponding effect on omics-discrimination analysis were evaluated using the representative sample-liver. The results suggest that same organ segment is preferred to sample in a batch of samples.

A MetID strategy based on integration of MS information mining and *in silico* MS simulation was proposed with the identified metabolites about 2.5 times more than that of accurate-mass-based MS and MS/MS matching in this study. The integrated MS information mining based on *in silico* and manual recognition/filtering of adduct ions, in-source CID MS, MS information of all possible ionizations, CFIs, CNLs and multimers, plays a critical role in both providing the most suspicious candidates for subsequent *in silico* simulation and reference-standard-based MetID and confirming MetID results. In combination with the mining, *in silico* MS simulation compressively presents its facilitation in the metabolite

identification in metabolomics, especially for identifying molecules that are not present in the databases or that are present yet do not have experimental MS/MS data.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This study was supported by the Shanghai Municipal Science Foundation (18ZR1432200) and National Natural Science Foundation of China (Grant No. 81872643) and the 3-year Action Program of Shanghai Municipal Government.

References

- [1]. Ramirez T, Daneshian M, Kamp H, Bois FY, Clench MR, Coen M, Donley B, Fischer SM, Ekman DR, Fabian E, Guillou C, Heuer J, Hogberg HT, Jungnickel H, Keun HC, Krennrich G, Krupp E, Luch A, Noor F, Peter E, Riefke B, Seymour M, Skinner N, Smirnova L, Verheij E, Wagner S, Hartung T, van Ravenzwaay B, Leist M, Metabolomics in toxicology and preclinical research, *Altex30* (2013) 209–225. [PubMed: 23665807]
- [2]. Ulmer CZ, Yost RA, Chen J, Mathews CE, Garrett TJ, Liquid chromatography-mass spectrometry metabolic and lipidomic sample preparation workflow for suspension-cultured mammalian cells using Jurkat T lymphocyte cells, *J. Proteomics Bioinform* 8 (2015) 126–132. [PubMed: 26401069]
- [3]. Putri SP, Nakayama Y, Matsuda F, Uchikata T, Kobayashi S, Matsubara A, Fukusaki E, Current metabolomics: practical applications, *J. Biosci. Bioeng* 115 (2013) 579–589. [PubMed: 23369275]
- [4]. Gregory KE, Bird SS, Gross VS, Marur VR, Lazarev AV, Walker WA, Kristal BS, Method development for fecal lipidomics profiling, *Anal. Chem* 85 (2013) 1114–1123. [PubMed: 23210743]
- [5]. Abbott SK, Jenner AM, Mitchell TW, Brown SH, Halliday GM, Garner B, An improved high-throughput lipid extraction method for the analysis of human brain lipids, *Lipids* 48 (2013) 307–318. [PubMed: 23355308]
- [6]. Huan T, Troyer D, Li L, Metabolite analysis and histology on the exact same tissue: comprehensive metabolomic profiling and metabolic classification of prostate cancer, *Sci. Rep* 6 (2016) 32272. [PubMed: 27578275]
- [7]. Gonzalez FJ, Fang Z-Z, Ma X, Transgenic mice and metabolomics for study of hepatic xenobiotic metabolism and toxicity, *Expert Opin. Drug Metab. Toxicol* 11 (2015) 869–881. [PubMed: 25836352]
- [8]. Psychogios N, Hau DD, Peng J, Guo AC, Mandal R, Bouatra S, Sinelnikov I, Krishnamurthy R, Eisner R, Gautam B, Young N, Xia JG, Knox C, Dong E, Huang P, Hollander Z, Pedersen TL, Smith SR, Bamforth F, Greiner R, McManus B, Newman JW, Goodfriend T, Wishart DS, The Human Serum Metabolome, *PLoS One* 6 (2011), e16957. [PubMed: 21359215]
- [9]. Bouatra S, Aziat F, Mandal R, Guo A, Wilson MR, Knox C, Bjorn Dahl TC, Krishnamurthy R, Saleem F, Liu P, Dame ZT, Poelzer J, Huynh J, Yallou FS, Psychogios N, Dong E, Bogumil R, Roehring C, Wishart DS, The human urine metabolome, *PLoS One* 8 (2013), e73076. [PubMed: 24023812]
- [10]. Naz S, dos Santos DCM, Garcia A, Barbas C, Analytical protocols based on LC-MS, GC-MS and CE-MS for nontargeted metabolomics of biological tissues, *Bioanalysis* 6 (2014) 1657–1677. [PubMed: 25077626]
- [11]. Price KE, Tissue-targeted Metabonomics: Metabolic Profiling by Microdialysis and NMR Spectroscopy, University of Kansas, 2008, pp. 259, Ann Arbor.
- [12]. Chen SL, Hoene M, Li J, Li YJ, Zhao XJ, Haring HU, Schleicher ED, Weigert C, Xu GW, Lehmann R, Simultaneous extraction of metabolome and lipidome with methyl tert-butyl ether

from a single small tissue sample for ultra-high performance liquid chromatography/mass spectrometry, *J. Chromatogr. A* 1298 (2013) 9–16. [PubMed: 23743007]

- [13]. Zukunft S, Pohn C, Röhrling C, Möller G, Hrabě de Angelis M, Adamski J, Tokarz J, High-throughput extraction and quantification method for targeted metabolomics in murine tissues, *Metabolomics* 14 (2018) 18. [PubMed: 29354024]
- [14]. Cajka T, Fiehn O, Comprehensive analysis of lipids in biological systems by liquid chromatography-mass spectrometry, *Trends Analyt. Chem* 61 (2014) 192–206.
- [15]. Li M, Yang L, Bai Y, Liu HW, Analytical methods in lipidomics and their applications, *Anal. Chem* 86 (2014) 161–175. [PubMed: 24215393]
- [16]. Vorkas PA, Isaac G, Anwar MA, Davies AH, Want EJ, Nicholson JK, Holmes E, Untargeted UPLC-MS profiling pipeline to expand tissue metabolome coverage: application to cardiovascular disease, *Anal. Chem* 87 (2015) 4184–4193. [PubMed: 25664760]
- [17]. Zhou LL, Zhang WP, Xie WP, Chen HM, Yu WL, Li HS, Shen GL, Tributyl phosphate impairs the urea cycle and alters liver pathology and metabolism in mice after short-term exposure based on a metabolomics study, *Sci. Total Environ* 603(2017) 77–85. [PubMed: 28623793]
- [18]. Stella R, Dervilly-Pinel G, Bovo D, Mastroianni E, Royer AL, Angeletti R, Le Bizet B, Biancotto G, Metabolomics analysis of liver reveals profile disruption in bovines upon steroid treatment, *Metabolomics* 13 (2017) 80.
- [19]. Guijas C, Montenegro-Burke JR, Domingo-Almenara X, Palermo A, Warth B, Hermann G, Koellensperger G, Huan T, Uritboonthai W, Aisporna AE, Wolan DW, Spilker ME, Benton HP, Siuzdak G, METLIN: a technology platform for identifying knowns and unknowns, *Anal. Chem* (2018) 3156–3164. [PubMed: 29381867]
- [20]. Lai Z, Tsugawa H, Wohlgemuth G, Mehta S, Mueller M, Zheng Y, Ogiwara A, Meissen J, Showalter M, Takeuchi K, Kind T, Beal P, Arita M, Fiehn O, Identifying metabolites by integrating metabolome databases with mass spectrometry cheminformatics, *Nat. Methods* 15 (2018) 53–56. [PubMed: 29176591]
- [21]. Lynn KS, Cheng ML, Chen YR, Hsu C, Chen A, Lih TM, Chang HY, Huang CJ, Shiao MS, Pan WH, Sung TY, Hsu WL, Metabolite identification for mass spectrometry-based metabolomics using multiple types of correlated ion information, *Anal. Chem* 87 (2015) 2143–2151. [PubMed: 25543920]
- [22]. Lu D, Zhang S, Wang D, Feng C, Liu S, Jin Y, Xu Q, Lin Y, Wu C, Tang L, She J, Wang G, Zhou Z, Identification of flurochloridone metabolites in rat urine using liquid chromatography/high resolution mass spectrometry, *J. Chromatogr. A* 1445 (2016) 80–92. [PubMed: 27063369]
- [23]. Prasad B, Garg A, Takwani H, Singh S, Metabolite identification by liquid chromatography-mass spectrometry, *Trac-Trends Anal. Chem* 30 (2011) 360–387.
- [24]. Blazenovic I, Kind T, Torbasinovic H, Obrenovic S, Mehta SS, Tsugawa H, Wermuth T, Schauer N, Jahn M, Biedendieck R, Jahn D, Fiehn O, Comprehensive comparison of in silico MS/MS fragmentation tools of the CASMI contest: database boosting is needed to achieve 93% accuracy, *J. Cheminform* 9 (2017) 32. [PubMed: 29086039]
- [25]. Gotra A, Sivakumaran L, Chartrand G, Vu K-N, Vandenbroucke-Menu F, Kauffmann C, Kadoury S, Gallix B, de Guise JA, Tang A, Liver segmentation: indications, techniques and future directions, *Insights Imaging* 8 (2017) 377–392. [PubMed: 28616760]
- [26]. Abdalla EK, Denys A, Chevalier P, Nemr RA, Vauthey J-N, Total and segmental liver volume variations: implications for liver surgery, *Surgery* 135 (2004) 404–410. [PubMed: 15041964]
- [27]. Taniguchi H, Oguro A, Takeuchi K, Miyata K, Takahashi T, Inaba T, Nakahashi H, Difference in regional hepatic blood flow in liver segments–non-invasive measurement of regional hepatic arterial and portal blood flow in human by positron emission tomography with H₂ 15O–, *Ann. Nucl. Med* 7 (1993) 141–145. [PubMed: 8217488]
- [28]. Han X, Variables in mass spectrometry for lipidomics, in: *Lipidomics: Comprehensive Mass Spectrometry of Lipids*, John Wiley & Sons, Inc., Hoboken, New Jersey, 2016, pp. 89–120, 08 4.
- [29]. Ridgway N, McLeod R, *Biochemistry of Lipids, Lipoproteins and Membranes*, Elsevier, 2015.
- [30]. Linster CL, Van Schaftingen E, Hanson AD, Metabolite damage and its repair or pre-emption, *Nat. Chem. Biol* 9 (2013) 72–80. [PubMed: 23334546]

- [31]. Mekala LP, Mohammed M, Chintalapati S, Chintalapati VR, Stable isotope-assisted metabolic profiling reveals growth mode dependent differential metabolism and multiple catabolic pathways of l-phenylalanine in *rubrivivaxbenzoatilyticus*JA2,J. Proteome Res 17 (2018) 189–202. [PubMed: 29043820]
- [32]. Garg N, Kapono C, Lim YW, Koyama N, Vermeij MJA, Conrad D, Rohwer F, Dorrestein PC, Mass spectral similarity for untargeted metabolomics data analysis of complex mixtures, Int. J. Mass Spectrom 377 (2015), 719–717. [PubMed: 25844058]

Fig. 1.
A workflow for the structural elucidation of unknown metabolites.

Fig. 2.

Comparison of number, intensity and RSDs of features acquired from two extraction methods.

Fig. 3.

Comparison of TIC acquired from two extraction methods (A&B: ESI (-), C&D: ESI (+), A&C: MTBE, B&D: DCM/MeOH (3:1, v/v)). Cer: Ceramides, CerG: Simple Glc series, CL: Cardiolipin, DG: diglyceride, dMePE: dimethylphosphatidylethanolamine, GM3: Gangliosides, LdMePE: lysodimethylphosphatidylethanolamine, LPC: lysophosphatidylcholine, LPE: lysophosphatidylethanolamine, LPG:lysophosphatidylglycerol, LPI: lysophosphatidylinositol, LPMe: lysophosphatidylmethanol, LPS: lysophosphatidylserine, MGDG:

Monogalactosyldiacylglycerol, PA: phosphatidic acid, PC: phosphatidylcholine, PE: phosphatidylethanolamine, PG: phosphatidylglycerol, phSM: sphingomyelin(phytosphingosine), PI: phosphatidylinositol, PIP: phosphatidylinositol, Pme: phosphatidylmethanol, PS: phosphatidylserine, SM: sphingomyelin, TG: triglyceride.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Fig. 4.

Number and peak area of identified lipids extracted from two methods (Red numbers mean the number of lipid species). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

Fig. 5. Identification of feature L641 and L642 as uridine based on its adduct ions and *in silico* MS simulation. A: EICs of feature L642 (m/z 243.0618) and L641 (m/z 279.03903) at RT 1.879 min; B: MS¹ spectrum of feature L641 with specific chlorine isotopic distribution pattern; C: Feature L641 was tentatively identified as uridine using *in silico* MS fragmentation (the top) and mzCloud based CFIs similarity matching (the bottom) and finally its identification was confirmed using uridine authentic standard (Fig. S4).

Fig. 6. Identification of feature L201 and L202 as ADMA based on MS information of in-source CID and two MS ionizations and *in silico* MS simulation. Feature L204(B, m/z 172.10809), L206 (C, m/z 158.12885), L205 (D, m/z 116.07095) and L203 (E, m/z 133.09645) are in-source CID products of feature L201 (A, m/z 203.15022) and they were acquired in ESI (+); Feature L202 (F, m/z 201.13451) acquired in ESI (–) is identified as same metabolite as feature L201; MS/MS of feature L201 (G), L204 (H), L203 (I) and L202 (G) are well

matched with *in silico* MS fragments of ADMA; The identification of feature L201 and L202 was confirmed using ADMA authentic standard (Fig. S6).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript