Metabolomics (2019) 15:96
https://doi.org/10.1007/s11306-019-1557-7

ORIGINAL ARTICLE

Comprehensive metabolomic and proteomic analyses reveal candidate biomarkers and related metabolic networks in atrial fibrillation

Juntuo Zhou[2] · Lijie Sun[1] · Liwen Chen[1] · Shuwang Liu[1] · Lijun Zhong[3] · Ming Cui[1]

Received: 23 January 2019 / Accepted: 15 June 2019 / Published online: 21 June 2019

Abstract

Introduction Atrial fibrillation (AF) is an abnormal heart rhythm characterized by an irregular beating of the atria and is associated with an increased risk of heart failure, dementia, and stroke. Currently, the perturbation of plasma content due to AF disease onset is not well known.

Objectives To investigate dysregulated molecules in blood plasma of untreated AF patients, with the goal of identifying biomarkers for disease screening and pathological studies.

Methods LC-MS based untargeted metabolomics, lipidomics and proteomics analyses were performed to find candidate biomarkers. A targeted quantification assay and an ELISA were performed to validate the results of the omics analyses.

Results We found that 24 metabolites, 16 lipids and 16 proteins were significantly dysregulated in AF patients. Pathway enrichment analysis showed that the purine metabolic pathway and fatty acid metabolism were perturbed by AF onset. FA 20:2 and FA 22:4 show great linear correlational relationship with the left atrial area and could be considered for AF disease stage monitoring or prognosis evaluation.

Conclusion we used a comprehensive multiple-omics strategy to systematically investigate the dysregulated molecules in the plasma of AF patients, thereby revealing potential biomarkers for diagnosis and providing information for pathological studies.

Keywords Atrial fibrillation · Metabolomics · Proteomics · Biomarkers

Abbreviations

AF	Atrial fibrillation
ECG	Electrocardiogram
LC–MS	Liquid chromatography–mass spectrometry
DDA	Data-dependent acquisition
PRM	Parallel reaction monitoring
MS/MS	Tandem mass spectrometry
QC	Quality control
PCA	Principle component analysis
FA	Fatty acid
ACar	Acylcarnitine
OxFA	Oxidized fatty acid
GSS	Glutathione synthetase

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s11306-019-1557-7) contains supplementary material, which is available to authorized users.

Juntuo Zhou and Lijie Sun have contributed equally to this study.

✉ Lijun Zhong
zhonglijun@bjmu.edu.cn

✉ Ming Cui
mingcui@bjmu.edu.cn

[1] Department of Cardiology and Institute of Vascular Medicine, Peking University Third Hospital; Key Laboratory of Cardiovascular Molecular Biology and Regulatory Peptides, Ministry of Health; Key Laboratory of Molecular Cardiovascular Science, Ministry of Education; Beijing Key Laboratory of Cardiovascular Receptors Research, No. 49, Hua Yuan North Rd, Hai Dian District, Beijing 100191, China

[2] Beijing Advanced Innovation Center for Big Data-Based Precision Medicine, Beihang University, Beijing 100083, China

[3] Center of Medical and Health Analysis, Peking University Health Science Center, Beijing 100191, China

1 Introduction

Atrial fibrillation (also called AFib or AF) is a quivering or irregular heartbeat (arrhythmia) that can lead to many heart-related complications (Iwasaki et al. 2011), including blood clots (Vink et al. 2005), stroke (Kaarisalo et al. 1997), and heart failure (Chamberlain et al. 2011), etc. The estimated number of individuals with AF globally in 2010 was 33.5 million (Chugh et al. 2014), and in the United States, it is estimated that the number of adults with AF will more than double by the year 2050 (Go et al. 2001). Additionally, increases in the frequency of AF cases have been predicted (Chugh et al. 2014; Miyasaka et al. 2006). In the clinic, prior to a final confirmation of AF by a 12-lead (current golden standard) or single-lead ECG documenting P-waves (January et al. 2014), several methods are used in early screening for the disease, such as pulse palpation (Cooke et al. 2006), handheld single-lead ECGs (Lau et al. 2013) and modified blood pressure monitors (Marazzi et al. 2012). These methods have their own advantages, as well as shortcomings (Freedman et al. 2017). More importantly, the existing screening methods mentioned above mostly depend on the onset of AF, thus may leave episodes of paroxysmal AF undetected. Therefore, the identification of more reliable plasma biomarkers for the early diagnosis or prognosis evaluation of AF is both meaningful and urgent and can additionally aid in our understanding of the disease's pathological mechanism and provide new therapeutic targets.

Proteomics (Pandey and Mann 2000) and metabolomics (Dettmer et al. 2007) aim to study the entire proteome and metabolome expressed in a given biological sample. As complementary techniques to genomics (Gstaiger and Aebersold 2009) and transcriptomics, proteomics and metabolomics are more sensitive in response to external factors and can better reflect the true physiological status of a biological system. The rapid development of LC–MS-based omics techniques (Ong and Mann 2005; Tyanova et al. 2016) has allowed the reliable analysis of the expression patterns of hundreds to thousands of metabolites or proteins from biological samples (Steger et al. 2016; Wishart 2016), thereby providing useful information for biomarker screening and pathological and biological studies (Aebersold and Mann 2003; Cravatt et al. 2007; Saito and Matsuda 2010).

Omics techniques have been used to facilitate the study of AF (Woods and Olgin 2014; Hyman and Deo 2017); however, most researchers have focused on the influence of drug therapy (Christersson et al. 2018), left atrial appendage closure (Fastner et al. 2018; Rusnak et al. 2018; Sattler et al. 2017) or other surgical operations (Jung et al. 2018) on AF patients. Few studies have used samples from untreated AF patients in omics analyses, with the exception of Alonso and colleagues, who found two bile acids to be correlated with an increased risk of AF (Zeller et al. 2015; Ko et al. 2016). In the present study, through the integration of proteomic and metabolomic profiling, we characterize dysregulated metabolites, lipids and proteins from untreated AF patients relative to control patients who have angiocardiopathy without AF. The results will provide candidate plasma biomarkers for AF diagnosis or prognosis, as well as provide clues for pathological study and the identification of new therapeutic targets.

2 Experimental

2.1 Chemical materials

Ammonium acetate was purchased from Sigma-Aldrich (St. Louis, MO, USA). Formic acid, HPLC grade isopropanol, acetonitrile and methanol were purchased from Fisher Scientific. Deionized water was produced by a Milli-Q system. The chemical standards were of analytical grade with typical purity of > 99%.

2.2 Ethical statement

The study was approved by the Clinical Ethics Committee of Peking University Third Hospital and conforms to the principles in the Declaration of Helsinki. The samples were obtained only from patients who agreed to undergo the exam for the purpose of laboratory research, and informed consent was obtained from all patients who were asked to donate blood. All methods were performed in accordance with the relevant guidelines and regulations.

2.3 Study population and design

Analyzed plasma samples were obtained from two cohorts (Fifty AF patients and fifty angiocardiopathy patients without AF). The angiocardiopathy patients included as control group were patients with cardiovascular disease such as atherosclerosis, hypertension or coronary heart disease, but no AF. The two cohorts were matched with age and gender. The fifty AF patients were diagnosed by 12-lead ECG, and blood samples were collected before treatment. All participants were recruited from Peking University Third Hospital (Beijing, China) from December 2016 to December 2017. Detailed information of the patients including medical history and data on medication were listed in Table S1 and Table S5. Obsolete cerebral infarction was defined as cerebral infarction diagnosed by CT or MRI and the age of onset was more than 30 years, and was known by collecting medical history. Left atrial area (LAA) was measured by two-dimensional ultrasound on the four-chamber view of the apex of the heart, and the LAA measurement was performed

by qualified cardiologists according to the American Society of Echocardiography standard.

2.4 Samples preparation for metabolomics

After an overnight fast, 4 mL forearms systemic circulation vein blood were collected. Blood samples were transferred into 4 mL EDTA-K2 vacuum blood collection tubes (BD Vacutainer) and placed on ice. Plasma was separated by centrifugation at 3000×*g* for 20 min, and then stored at − 80 °C until analysis. Metabolites and lipids were extracted from plasma samples using liquid–liquid extraction as follows: 100 µL plasma was extracted by fourfold volume of cold chloroform: methanol (2:1). The mixture was vortexed and centrifuged at 13,000×*g* for 15 min. Then the upper aqueous phase (hydrophilic metabolites) and the lower organic phase (hydrophobic metabolites) were separately collected and evaporated at room temperature under vacuum. The evaporated samples were stored at − 80 °C until LC–MS analysis.

2.5 High-performance liquid chromatography for metabolomics

Metabolomics and lipidomics were performed on an Ultimate 3000 UHPLC system coupled with Q-Exactive HF MS (Thermo Scientific). For the aqueous phase (metabolomics), an Xbridge amide column (100 × 2.1 mm i.d., 3.5 µm; Waters) was employed for compound separation at 30 °C. The samples were suspended with 100 µL of acetonitrile:water (1:1, v/v) solution and the injection volume was 10 µL. For the lipid, chromatographic separation was performed on a reversed phase X-select CSH C18 column (2.1 mm × 100 mm, 2.5 µm, Waters, USA) at 40 °C. The samples were suspended with 100 µL of chloroform:methanol (1:1, v/v) solution and then diluted threefold with isopropanol:acetonitrile:H2O (2:1:1, v/v/v) solution. The injection volume was 10 µL. Detailed methods can be found in Supplementary Materials.

2.6 Mass spectrometry for metabolomics

Data-dependent acquisition (DDA) and parallel reaction monitoring (PRM) based targeted quantification assay were performed using the Q-Exactive HF MS (Thermo Scientific). For DDA–MS, acquisition was performed in positive ion mode and negative ion mode separately in TOP 10 mode. For PRM–MS, the m/z of the 14 targeted lipids (13 target lipids and 1 internal standard) were set in the inclusion list and each acquisition cycle consists of 1 full MS1 scan at 60,000 resolution from 200 to 1200 m/z and 14 targeted MS2 scans at 30,000 resolution targeting the designated lipids. Samples (n = 100 in total) were analyzed in random order. Quality control (QC) samples were prepared by pooling equally volumes of all study samples, and were analyzed between every 10 samples during the entire LC–MS analytical sequence.

Detailed methods can be found in Supplementary Materials.

2.7 DDA–MS data analysis for metabolomics

Raw data collected from the DDA–MS were processed on MS–DIAL software (Tsugawa et al. 2015) according to user guide. Briefly, the raw MS data were converted from the vendor file format (.wiff) into the common file format of Reifycs Inc. (.abf) using the Reifycs ABF converter (http://www.reifycs.com/AbfConverter/index.html). After conversion, the MS–DIAL software was used for feature detection, spectra deconvolution, metabolite identification and peak alignment between samples. The MS/MS spectra based metabolite identification was performed in MS–DIAL by searching the acquired MS/MS spectra against the MassBank database provided by MS–DIAL software, containing MS1 and MS/MS information of metabolites (8068 records in positive ion mode and 4782 records in negative ion mode). The MS/MS spectra based lipid identification was performed in MS–DIAL by searching the acquired MS/MS spectra against the software's internal in silico MS/MS spectra database (version: LipidDBs-VS23-FiehnO), which includes MS1 and MS/MS information of common lipid species. The tolerance for MS1 and MS/MS search were set to 0.01 Da and 0.05 Da, separately. Other parameters used in MS–DIAL were set as default. Raw mass spectrometric data have been deposited in MassIVE with ID: MSV000083774.

2.8 PRM–MS data analysis

Raw data were processed on Skyline software according to the protocols (https://skyline.gs.washington.edu/labkey/project/home/software/Skyline/begin.view). Raw MS data files were imported to the software for peak extraction. One precursor-product ion pair (transition) was pre-selected for each target lipid, and the peak area corresponding to each transition was calculated by the software. Finally, result containing lipid identifications and quantifications were exported in table format for further statistical analysis.

2.9 Samples preparation for proteomics

Total of 48 AF patients and 48 healthy controls were used in the proteomics study. We created three plasma pools for each group, each pool contained equal amounts of plasma from 16 subjects. Each pooled sample (4 µL crude plasma)

was subjected to Multiple Affinity Removal System (MARS) Human-14 column (Agilent) for high abundance protein depletion in accordance with the manufacturer's recommendations. The flow-through fractions (low abundance proteins) were collected and then digested according to the manufacturer's protocol for filter-aided sample preparation (FASP). Proteins were digested by trypsin at a protein to enzyme ratio of 50:1 in 50 mM NH_4HCO_3 buffer overnight at 37 °C, and released peptides were collected by centrifugation and evaporated under vacuum.

2.10 LC–MS analysis for proteomics

The samples (1 μg) were analyzed on a house-made C18 column (75 μm × 10 cm, 3 μm) by using a U3000 UHPLC connected to a Q-Exactive HF mass spectrometer (Thermo Scientific). The peptides were separated by a linear gradient from 5 to 35% ACN with 0.1% formic acid at 300 nL/min for 60 min and linearly increased to 80% ACN in 1 min and maintained for 3 min. The column was re-equilibrated at 5% ACN for 5 min. The MS acquisition was performed in TOP-20 DDA mode. Raw MS data file analysis was performed with MaxQuant software version 1.6.2.0 (http://www.maxquant.org/). For protein identification, MS/MS data were submitted to the UniProt human protein database (release 3.43, 72,340 sequences) using the Andromeda search engine with the following settings: trypsin cleavage; fixed modification of carbamidomethylation of cysteine; variable modifications of methionine oxidation; a maximum of two missed cleavages; false discovery rate calculated by searching the decoy database. Other parameters were set as default. Label-free quantitation (LFQ) was also performed in MaxQuant. The Minimum ratio count for LFQ was set to 2, and the match-between-runs option was set to 1 min. Detailed method can be found in Supplementary Materials. Raw mass spectrometric data have been deposited in MassIVE with ID: MSV000083774.

2.11 Statistical analysis

Statistical analysis, principle component analysis (PCA), hierarchical clustering analysis and pathway enrichment analysis for DDA data were performed using Metaboanalyst 4.0 web service, an online tool for analyzing omics data. For the metabolomics analyses between group AF and control, grouped *t* test were performed, and significant metabolites were identified based on the following criteria: (i) a p-value < 0.05 and (ii) a fold change ≥ 2. Enriched metabolic pathways were retrieved from KEGG or SMPDB database and redrawn using Adobe Photoshop. For the proteomics analyses, grouped t-test were performed and a p-value < 0.05 was used as cutoff. χ^2-test was used for the analysis of categorical variable. Functional protein association network construction and GO analysis were performed using STRING web service. Column graph, ROC analysis and correlation analysis for the PRM results were performed using GraphPad Prism 6. Quantitative results are reported in the form of the mean ± SEM. Associations between the quantitation of lipids and left atrial area were evaluated by Pearson correlation coefficient.

3 Results

3.1 Overview of the study workflow

In this study, 50 plasma samples were collected from patients with clinically confirmed AF, while 50 control plasma samples were collected from patients with common angiocardiopathy without AF. The patients' age, sex, history of disease and AF related diagnostic records are presented by patient set in Table S1, and more detailed information of the patients are presented in Table S5. Clinical characteristics, including age, sex, smoking status, and history of diseases were comparable between the discovery and validation sets (Table S1). The comprehensive multi-omics workflow is shown in Fig. 1. LC–MS-based untargeted metabolomics, lipidomics and proteomics were performed to investigate possible dysregulated biological molecules in the plasma at the onset of AF. After statistical analysis and data mining, 24 metabolites, 16 lipids and 16 proteins were found significantly dysregulated in AF patients. Further bioinformatic analyses were employed to reveal perturbed pathways and provide clues to the pathological mechanism. We then validated part of the omics results with the use of a LC–MS-based targeted quantification assay (PRM) and an ELISA. Lastly, ROC curves for the dysregulated molecules and correlations between these molecules, and AF clinical factors were calculated to evaluate their potential use as biomarkers for AF diagnosis or prognosis.

3.2 Metabolomic profiling and dysregulated metabolites in AF

After processing the raw MS data, a principal component analysis (PCA) was used to create an overview of the metabolomic expression profiles of all the samples in positive (Fig. 2a) and negative (Fig. 2b) ion modes. Pooled quality control (QC) samples were closely clustered in the PCA score plots, while the AF and control groups showed a trend of partial separation. After peak alignment and the removal of missing values, over two thousand features with acquired MS/MS spectra were obtained. 394 features in positive ion mode and 224 in negative ion mode were reliably identified by MS/MS spectra comparison. Of these features, further statistical analysis revealed a

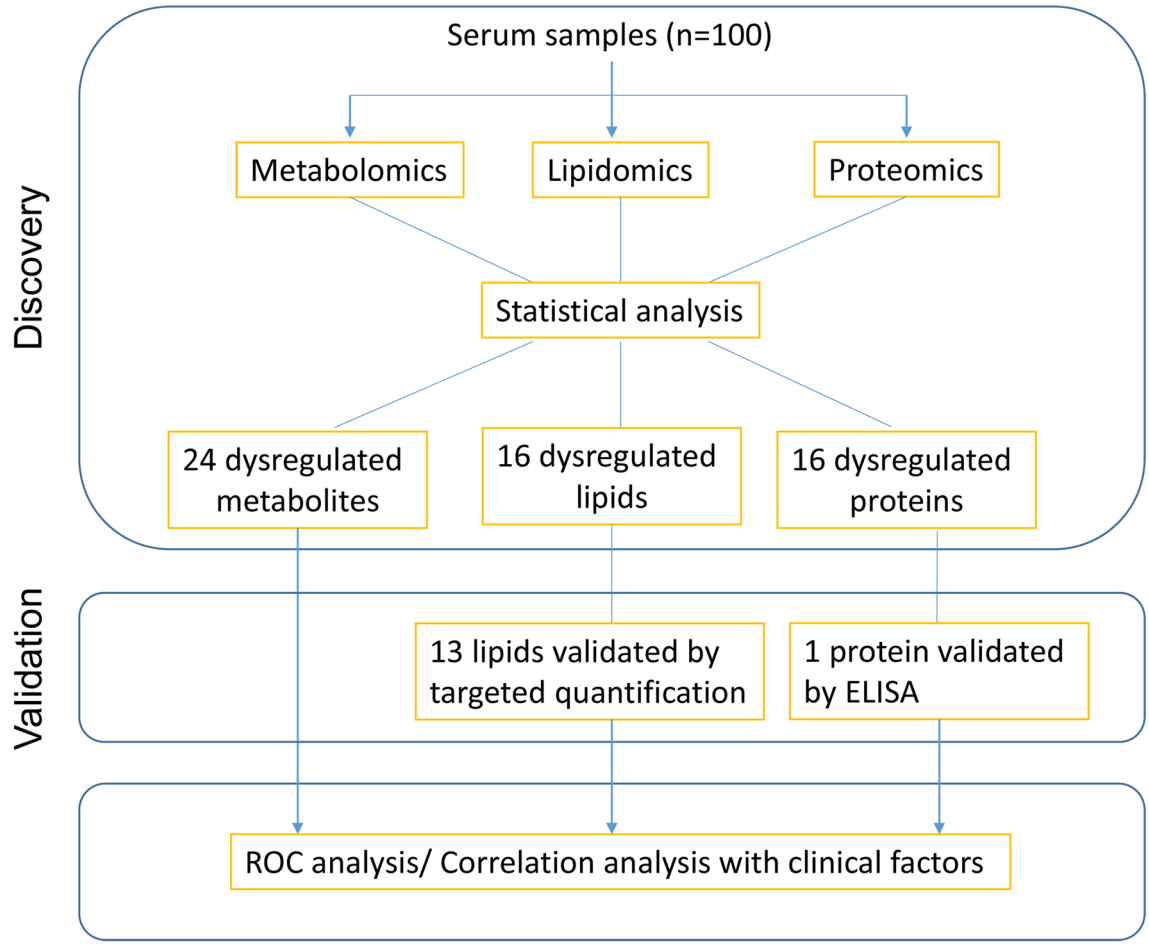


Fig. 1 Flowchart of the study workflow. One hundred patients were included in the study, of which 50 were AF patients without therapeutic treatment. 50 patients with angiocardiopathy but no AF onset were included as a control group. In the discovery stage, untargeted metabolomics, lipidomics and proteomics were performed to investigate dysregulated molecules in the patients' plasma. In the validation stage, the omics results were validated by a targeted quantitative assay and an ELISA. Pathway enrichment, correlation analysis and ROC analysis were also performed to further gain insights from the omics data

total of 24 dysregulated metabolites (16 in positive ion mode, Fig. 2c; 10 in negative ion mode, Fig. 2d; Guanosine 5′-monophosphate and inosine were detected in both ion modes) between the AF and control groups (FC > 2 and t-test p-value < 0.05, Table S2). Hierarchical clustering analysis was performed on these metabolites, and a heat map was used to visualize their expression profiles (Figure S1a). Correlation analysis showed that eight of the metabolites (uridine 5′-diphosphate, AMP, uridine 5′-monophosphate, hypoxanthine, para-aminobenzoic acid, adenine, adenosine 3′-monophosphate and guanosine 5′-monophosphate) had expression patterns that were positively correlated, while others showed no correlation (Figure S1b).

3.3 Lipidomic profiling and dysregulated lipids in AF

As with the data processing procedures for the metabolomic features, PCA was used for an overview of the expression patterns of the lipids in all the samples. Similar to results found for the metabolomic profiles, the lipid QC samples were closely clustered in the PCA score plots both in positive (Fig. 3a) and negative (Fig. 3b) ion modes. However, the AF and control groups did not show obvious trends of separation. In total, 821 features in positive ion mode and 259 in negative were reliably identified by MS/MS spectra comparison. Statistical analysis revealed two lipids in positive ion mode (Fig. 3c) and 14 lipids in negative ion mode (Fig. 3d) that were dysregulated between the AF and control groups (FC > 2 and t-test p-value < 0.05, Table S3). The majority of the dysregulated lipids (12/16) was fatty acids which were upregulated in the AF group,

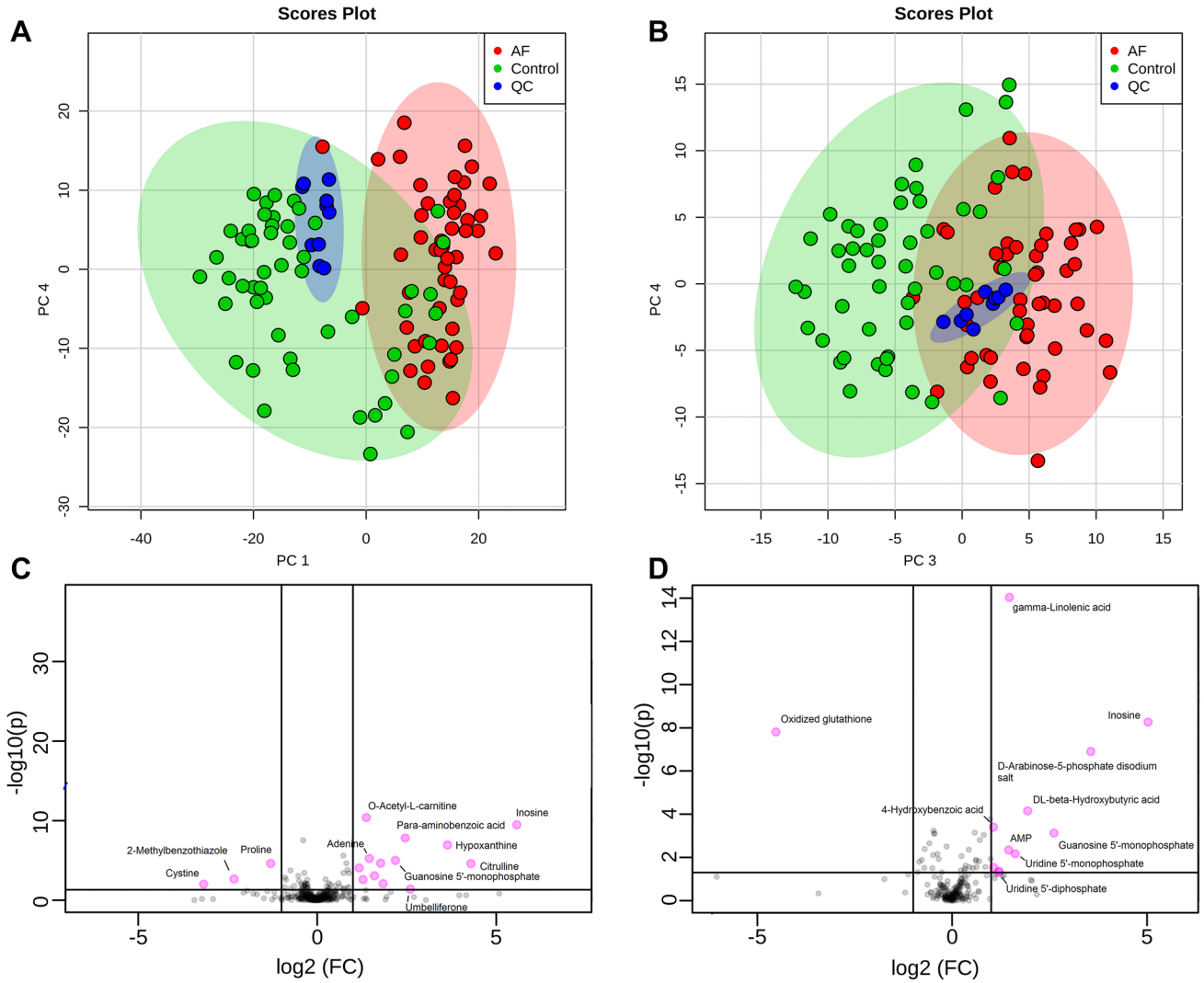


Fig. 2 Overview of the untargeted metabolomic results. **a** PCA score plot for the metabolites detected in positive ion mode. The samples in different groups are presented by different colors: red, AF (n = 50); green, control (n = 50); blue, QC. Circles represent the 95% confident interval. **b** PCA score plot for the metabolites detected in negative ion mode. The samples in different groups are presented by different colors: red, AF (n = 50); green, control (n = 50); blue, QC. Circles represent the 95% confident interval. **c** Scatter plots presenting fold change (FC) and t-test p-value of the identified metabolites in positive ion mode. The X-axis represents the log2-transformed FC, and the Y-axis represents the log10-transformed p-value. **d** Scatter plots presenting FC and t-test p-value of the identified metabolites in negative ion mode. The X-axis represents the log2-transformed FC, and the Y-axis represents the log10-transformed p-value

as shown in Figure S1c. The other two upregulated lipids were acyl-carnitines, which is in agreement with the metabolomic results, which indicated an upregulation of acetyl-carnitines in AF patients (Fig. 2a, Table S2). A heat map was used to show the expression profiles of the 16 dysregulated lipids (Figure S1c). Correlation analysis demonstrated that the 12 fatty acids had positively correlated expression patterns, while the other lipids exhibited little to no correlation (Figure S1d).

3.4 Bioinformatic analysis reveals perturbed metabolic pathways

The 24 dysregulated metabolites and 16 dysregulated lipids were used to search against KEGG and SMPDB pathway databases. Two metabolic pathways were found to be significantly perturbed by AF disease onset with FDR adjusted p-value of < 0.05 (Fig. 4a, b). One is the pathway of purine metabolism, of which seven metabolites (adenine, 3′-AMP, adenosine 3′ 5′-diphosphate, adenosine monophosphate, guanosine monophosphate, hypoxanthine and inosine) were observed to be upregulated in

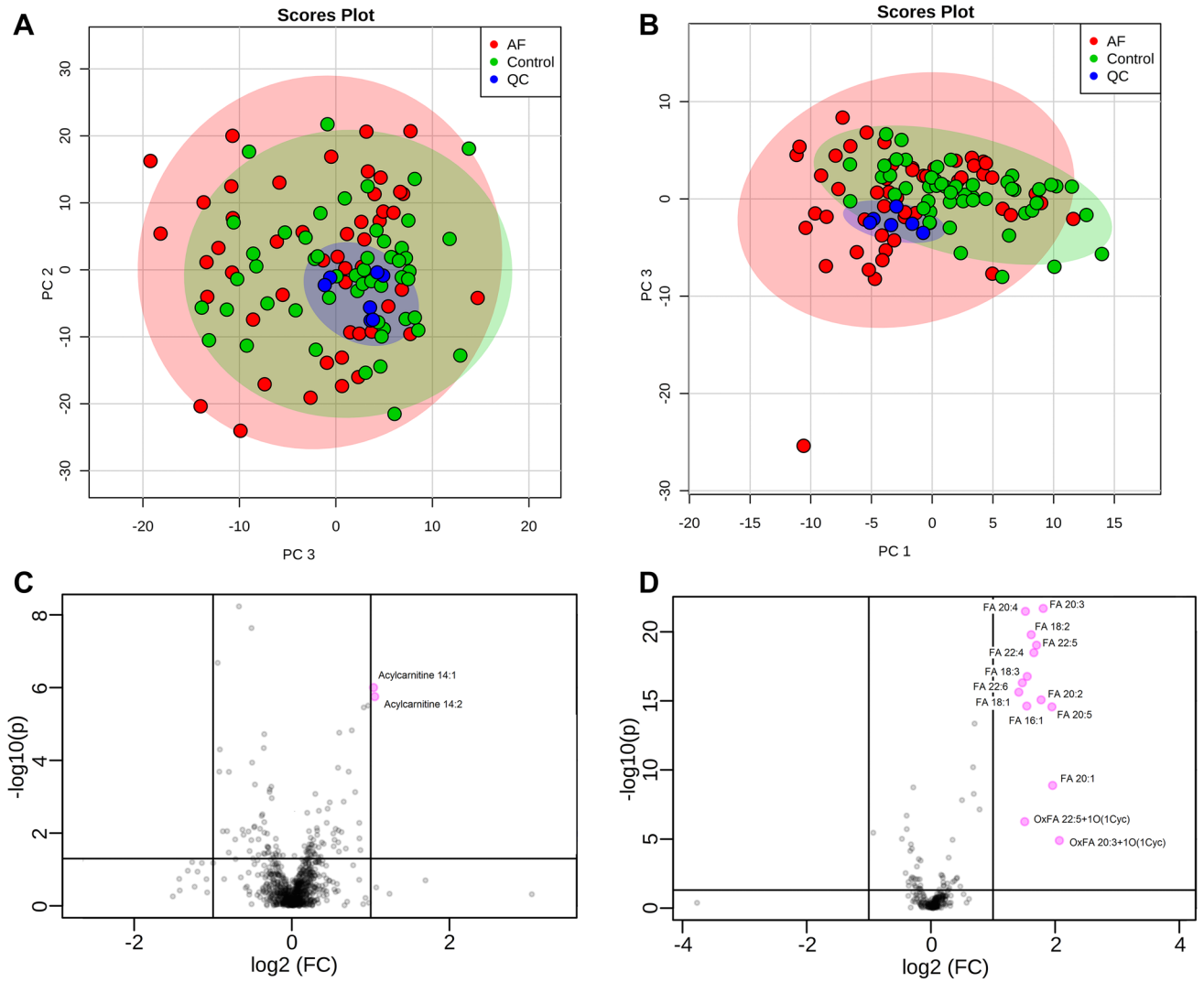


Fig. 3 Overview of the untargeted lipidomic results. **a** PCA score plot for lipids detected in positive ion mode. Samples in different groups are presented by different colors: red, AF (n = 50); green, control (n = 50); blue, QC. Circles represent 95% confident interval. **b** PCA score plot for lipids detected in negative ion mode. Samples in different groups are presented by different colors: red, AF (n = 50); green, control (n = 50); blue, QC. Circles represent 95% confident interval. **c** Scatter plots presenting FC and t-test p-value of the identified lipids in positive ion mode. X axis represents log2 transformed FC, and Y axis represents log10 transformed p-value. **d** Scatter plots presenting FC and t-test p-value of the identified lipids in negative ion mode. X axis represents log2 transformed FC, and Y axis represents log10 transformed p-value

AF patients in the metabolomic analysis. The other is the pathway of alpha linolenic acid and linoleic acid metabolism, of which nine fatty acids (FA 18:1, 18:2, 20:3, 20:4, 20:5, 22:4, 22:5, 22:6) were seen to be upregulated in AF. Functional protein association network of the 16 dysregulated proteins was constructed and proteins enriched in different molecular functions or pathways were shown by different node colors (Fig. 4c).

3.5 Overview of proteomic results and dysregulated proteins

Plasma protein profiling was analyzed by label-free proteomics. 747 proteins were identified. After statistical analysis, 5 upregulated and 11 downregulated proteins were found, as shown in Fig. 5a and Table S4. Further GO analysis (Table S6) revealed potential links between proteomics

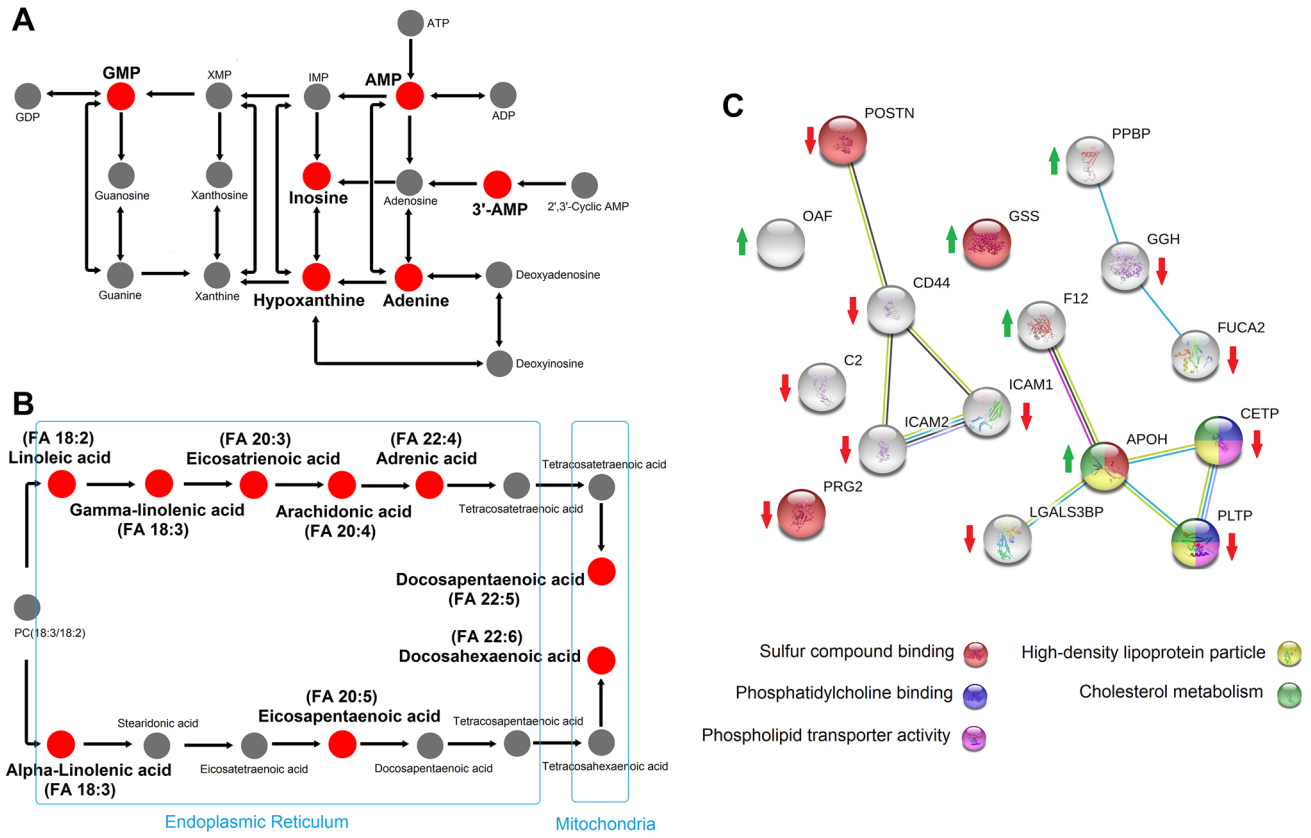


Fig. 4 Metabolic pathways enriched by dysregulated metabolites and lipids. **a** The pathway of purine metabolism. Seven metabolites (red-colored) were detected up-regulated in AF by our metabolomics analysis. The pathway diagram was derived from KEGG database. **b** The pathway of Alpha linolenic acid and linoleic acid metabolism. Nine fatty acids (red-colored) were detected up-regulated in AF. The pathway diagram was derived from SMPDB pathway database. **c** Functional protein association network of the 16 dysregulated proteins constructed using STRING web service. Proteins enriched in different molecular functions or pathways are shown by different node colors and interaction types are presented by different colors: yellow, text-mining; black, co-expression; blue, curated database; purple, experimentally determined

and metabolomics results, as some proteins were enriched in the biological process of triglyceride transport, having the molecular function of phospholipid binding and sulfur-compound binding, and are localized to the high-density lipoprotein particle. To validate the proteomic results, we randomly chose one dysregulated protein, glutathione synthetase (GSS), to be quantified using an ELISA. As shown in Fig. 5b, GSS showed statistically significant upregulation in the AF group, which is in accordance with the proteomic result.

3.6 Candidate biomarker validation and correlation analysis with left atrial area

A PRM-based targeted lipid quantification assay was performed to quantify 13 lipids, including 11 dysregulated lipids found in the lipidomic analysis and 2 others (FA 16:0 and FA 18:0). C^{13} FA 16:0 was used as an internal standard to exclude measurement errors. As shown in Fig. 6a, these FAs showed significant upregulation in the AF group except for FA 18:0, which is consist with the lipidomic results. Further ROC analysis was performed using the expression levels of the FAs. Nine FAs showed an AUC greater than 0.8 (Table S7), and the ROCs of the FAs with the top 5 AUCs (FA 18:2, 18:3, 20:3, 20:4, and 22:4) are presented in Fig. 6b–f. The correlation between the FA expression patterns and the left atrial area (a clinical parameter for AF prognosis evaluation) was also performed. Two FAs (20:2 and 22:4) were found to have significant correlations with the left atrial area in AF patients (Fig. 6g, i) while no significant correlations were found in the control group (Fig. 6h, j).

4 Discussion

Atrial fibrillation is one of the most common and serious abnormal heart arrhythmias, yet ideal biomarkers for early diagnosis or prognosis prediction is still needed (Woods and

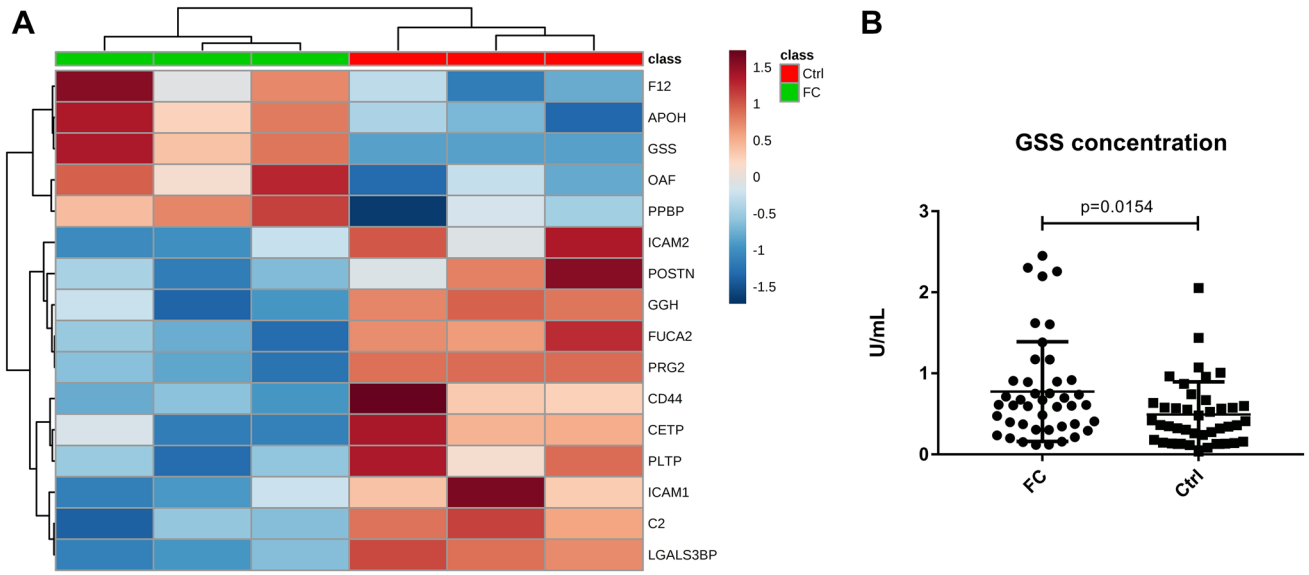
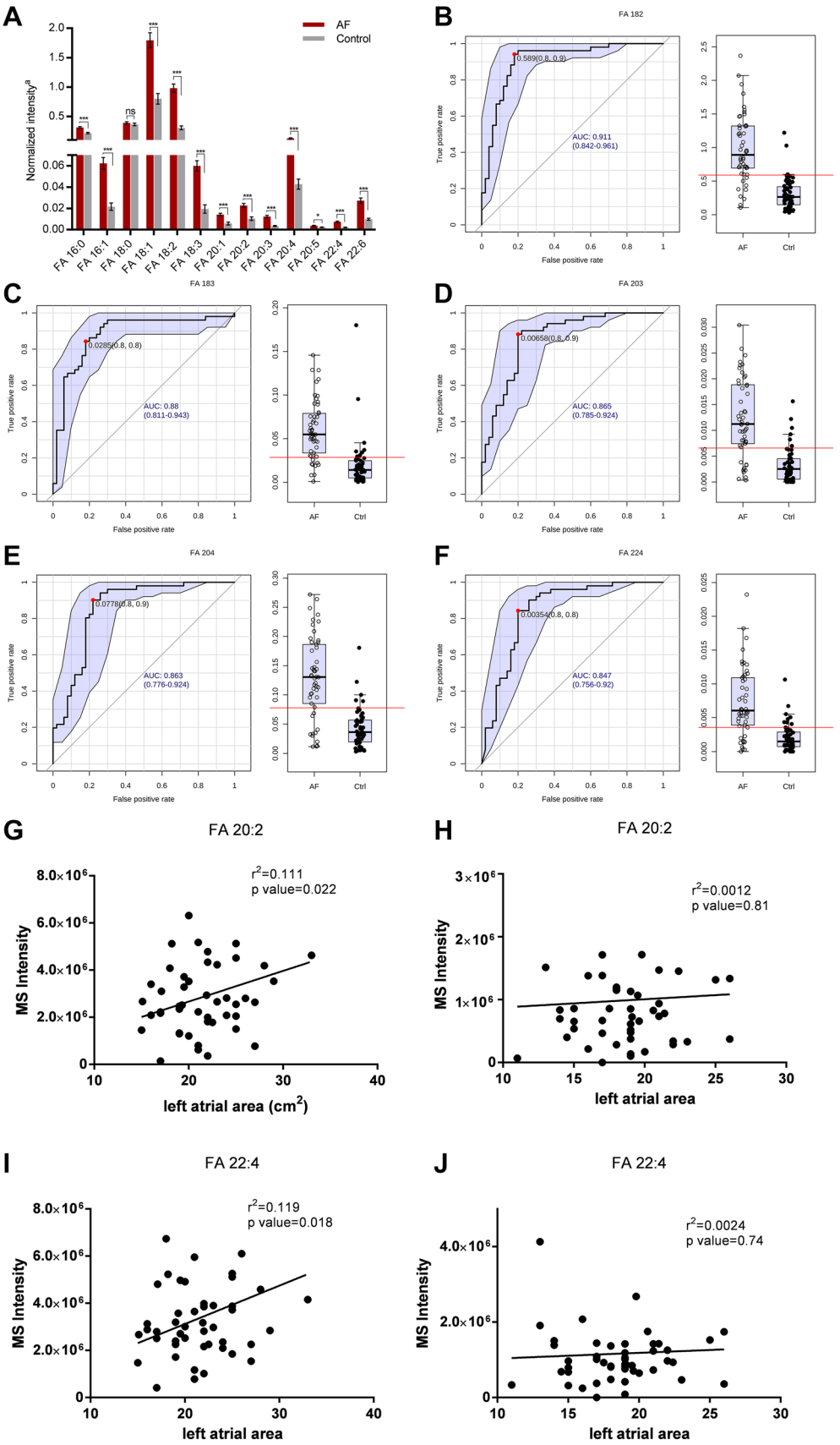


Fig. 5 Dysregulated proteins revealed by proteomics. **a** Heat map presenting the expressive patterns of 16 dysregulated proteins. Sample category is presented in red (AF, n = 3 pooled biological replicates) and green (control, n = 3 pooled biological replicates), and the intensity of protein is presented from blue (low intensity) to red (high intensity). **b** Grouped scatter plot presenting the concentration (U/mL) of GSS analyzed by ELISA (AF group, n = 45; control group, n = 45)

Olgin 2014). The early detection of asymptomatic AF would provide an opportunity to prevent or decrease unfavorable disease outcomes by the early implementation of appropriate therapeutic strategies (Chugh et al. 2014). In the present study (as illustrated in the flowchart in Fig. 1), we employed three untargeted omics analytical strategies to comprehensively analyze the dysregulated molecules in plasma from AF patients, and further PRM and ELISA assays validated the reliability of our omics data. Finally, bioinformatics and multi-omics correlative analysis revealed AF related metabolic pathways. Thus, to the best of our knowledge, the present study is the first to reliably provide a comprehensive view of plasma content profiles of untreated AF patients. Our results will provide not only potential biomarkers for AF diagnosis or prognosis but also clues to further understand the AF pathological mechanism.

LC–MS-based omics strategies such as proteomics and metabolomics have been proven to be powerful tools for the identification of biomarkers and in pathological studies. In comparison to genetics, proteomics and metabolomics are more closely related to phenotypic expression and influential environmental factors, and thus can better reflect the state of the human body. Among the 24 dysregulated metabolites identified, eight with highly positively correlated expression levels (Figure S1) were found to belong to the purine metabolic pathway (Fig. 4a). This result indicates that the purine pathway may be involved in AF disease progress and that these metabolites could potentially be developed into an analytical panel of biomarkers for AF diagnosis. Previous studies (Suzuki et al. 2012; Letsas et al. 2010; Liu et al. 2011) have reported that plasma uric acid is upregulated in AF patients and is correlated with AF burden (Letsas et al. 2010). Interestingly, uric acid is the end product of purine metabolism. Another metabolite related to purine metabolism, adenosine, has been reported to induce AF (Li et al. 2016). Taken together with the findings of the current study, the purine metabolic pathway is strongly implicated in AF disease progression. Another interesting molecular group from our results is the fatty acids, which were found in the lipidomic analysis. Twelve unsaturated fatty acids with carbon numbers ranging from 18 to 22 were found to be upregulated in AF patients, and their expression levels were highly correlated with each other (Figure S1). Nine of these fatty acids belong to the alpha linolenic acid and linoleic acid metabolic pathway, which occurs mainly in the endoplasmic reticulum (Fig. 4b). Fatty acids have been reported to be downregulated in DCCV treated AF patients (Jung et al. 2018). However, our results showed that in untreated AF patients, FAs were upregulated. This suggests that the levels of FAs may be used as a potential marker for AF disease diagnosis and in evaluation of therapeutic effect. Also, FAs and glucose are the principal energy substrates in cardiomyocytes to produce ATP (Harada et al. 2017), thus the upregulation of free FAs in plasma may be related to the FA intake or metabolism dysfunction of cardiomyocyte. Additionally, another lipid species, acylcarnitines, was found to be upregulated in AF in both the metabolomic (acetyl-carnitine) and lipidomic (acylcarnitine 14:1 and 14:2) results, validating the reliability of the omics analyses.


A
AF
Control
Normalized intensity[a]
FA 16:0
FA 16:1
FA 18:0
FA 18:1
FA 18:2
FA 18:3
FA 20:1
FA 20:2
FA 20:3
FA 20:4
FA 20:5
FA 22:4
FA 22:6
B
FA 182
0.589(0.8, 0.9)
AUC: 0.911
(0.842-0.961)
True positive rate
False positive rate
AF
Ctrl
C
FA 183
0.0285(0.8, 0.8)
AUC: 0.88
(0.811-0.943)
D
FA 203
0.00658(0.8, 0.9)
AUC: 0.865
(0.785-0.924)
E
FA 204
0.0778(0.8, 0.9)
AUC: 0.863
(0.776-0.924)
F
FA 224
0.00354(0.8, 0.8)
AUC: 0.847
(0.756-0.92)
G
FA 20:2
r2=0.111
p value=0.022
MS Intensity
left atrial area (cm2)
H
FA 20:2
r2=0.0012
p value=0.81
left atrial area
I
FA 22:4
r2=0.119
p value=0.018
J
FA 22:4
r2=0.0024
p value=0.74

◂**Fig. 6** Validation and clinical correlation analysis of the 13 dysregulated lipids. **a** Relative abundance of 13 dysregulated fatty acids analyzed by targeted quantitative assay (AF group, n = 50; Healthy control group, n = 50). Data are presented as the mean ± SEM of each independent sample. Significance was determined with a t-test. *p < 0.05, **p < 0.01, ***p < 0.001, ns: nonsignificant, p > 0.05. **b–f** ROC curves of FAs (18:2, 18:3, 20:3, 20:4, 22:4, respectively. AF group, n = 50; Healthy control group, n = 50). AUC, 95% confidence intervals, scatter plots and cutoff values are also shown. **g**, **h** Scatter plot presenting the relationship between the left atrial area and the intensity of FA 20:2 in the AF group (**g**, n = 50) and the control group (**h**, n = 50). Linear regression r^2 values and p-values are presented. **i**, **j** Scatter plot presenting the relationship between the left atrial area and the intensity of FA 22:4 in AF group (**i**, n = 50) and control group (**j**, n = 50). Linear regression r^2 values and p-values are presented

LC–MS-based proteomic analysis is a widely used tool for plasma biomarker screening. Due to the limited throughput of the proteomic workflow, pooled samples were used for the biomarker screening procedure. One of the identified dysregulated proteins from the screening, GSS, was randomly chosen for validation using an ELISA with individual samples; the ELISA showed consistent results with the proteomic outcome. Taking together the multi-omics results we acquired, potential links are found between dysregulated proteins and metabolites/lipids. GSS is the second enzyme in the glutathione (GSH) biosynthesis pathway (Liang et al. 1999), catalyzing the formation of glutathione, which is an important antioxidant (Foyer et al. 1995). In the metabolomics results, cystine and oxidized glutathione were detected down-regulated in AF patients (Table S2). Considering the increased oxidative stress from AF (Watanabe 2012), we may infer that the glutathione related metabolic pathway is perturbed. Bioinformatics analysis also reveals that dysregulated proteins were enriched in lipid metabolism related functions or pathways such as lipid binding, transfer and cholesterol metabolism (Fig. 4c). Considering the up-regulation of FAs in lipidomic results, FA related lipid metabolic pathways may be involved in AF pathological progress. But considering the plasma is not the place where biochemical reaction is performed, the expressive correlation between proteins and metabolites/lipids in plasma may not reflect the true reaction status in cardiomyocytes, so the exact regulative mechanism need to be further validated and studied in tissues or cells.

The LC–MS-based targeted quantification assay (MRM assay) is seen as the gold standard for accurate metabolite quantification (Garcíacañaveras et al. 2012). PRM is a newly applied strategy for targeted quantitative analyses of proteins (Peterson et al. 2012), metabolites (Zhou et al. 2016) and lipids (Zhou et al. 2017) with the advantages of good reliability and feasibility (Zhou and Yin 2016). In the present study, we used a PRM-based quantification assay to validate our lipidomic results, reliably quantifying 13 fatty acids in plasma samples (Fig. 6a). The results showed great consistency between the lipidomic analysis and the targeted quantification assay. Except the 11 dysregulated lipids found in the lipidomic analysis, we also included FA 16:0 and 18:0 in the targeted quantification assay, which were not found to be changed in the lipidomic results. FA 18:0 was found unchanged in the targeted quantification, however, FA 16:0 was changed. Due to this discrepancy, we rechecked the results of the lipidomic analysis, and found that FA 16:0 indeed showed trends of upregulation (FC of 1.6) but did not reach the cutoff set for the screening (FC of 2). Therefore, we verified the upregulation of FAs in AF patients by both lipidomics and targeted quantification. The nine FAs which have an AUC greater than 0.8 can be potential biomarkers for AF diagnosis. What's more, FA 20:2 and FA 22:4 show great linear correlational relationship with the left atrial area, which is a widely employed index in the clinic for AF prognosis evaluation (Sievers et al. 2004). These two lipids could be considered for AF disease stage monitoring or prognosis evaluation.

5 Conclusions

In conclusion, we have demonstrated there to be a metabolic disturbance in the plasma of AF patients, and several dysregulated molecules were identified including metabolites, lipids and proteins. Our results can help in the identification of novel biomarkers or effective therapeutic targets and provide clues for a better understanding of AF pathology. However, further validation of these candidate biomarkers in larger sample cohorts using a double-blind test is needed before they can be applied in clinical diagnosis. Additionally, functional studies with biological experiments are also needed to illustrate the exact molecular mechanism of AF onset.

Authors' contributions JZ, LS, LZ and MC designed the study. LS, LC and SL collected samples and performed clinical related analyses. JZ and LZ performed metabolomics and proteomics experiments. JZ analyzed the data. LC and SL reviewed statistical analyses. JZ and MC wrote the manuscript. All authors read and approved the final manuscript.

Funding This study was supported by the Fund for Fostering Young Scholars of Peking University Health Science Center (Grant No. BMU2018PY006) and the National Natural Science Foundation of China (Grant No. 81570235).

Compliance with ethical standards

Conflict of interests The authors declare that they have no conflict of interests.

Ethics approval The study was approved by the Clinical Ethics Committee of Peking University Third Hospital and conforms to the prin-

ciples in the Declaration of Helsinki. The samples were obtained only from patients who agreed to undergo the exam for the purpose of laboratory research, and informed consent was obtained from all patients who were asked to donate blood. All methods were performed in accordance with the relevant guidelines and regulations.

Informed consent Awritten informed consent was obtained from all the included patients.

References

Aebersold, R., & Mann, M. (2003). Mass spectrometry-based proteomics. *Nature, 422,* 198–207.

Chamberlain, A. M., Redfield, M. M., Alonso, A., Weston, S. A., & Roger, V. L. (2011). Atrial fibrillation and mortality in heart failure: A community study. *Circulation Heart Failure, 4,* 740–746.

Christersson, C., et al. (2018). Effect of apixaban compared with warfarin on coagulation markers in atrial fibrillation. *Heart, 105*(3), 235–242.

Chugh, S. S., et al. (2014). Worldwide epidemiology of atrial fibrillation: A global burden of disease 2010 study. *Circulation, 129,* 837–847.

Cooke, G., Doust, J., & Sanders, S. (2006). Is pulse palpation helpful in detecting atrial fibrillation? A systematic review. *Journal of Family Practice, 55,* 130–134.

Cravatt, B. F., Simon, G. M., & Iii, J. R. Y. (2007). The biological impact of mass-spectrometry-based proteomics. *Nature, 450,* 991–1000.

Dettmer, K., Aronov, P. A., & Hammock, B. D. (2007). Mass spectrometry-based metabolomics. *Mass Spectrometry Reviews, 26,* 51–78.

Fastner, C., et al. (2018). Interventional left atrial appendage closure affects the metabolism of acylcarnitines. *International Journal of Molecular Sciences, 19*(2), 500.

Foyer, C. H., et al. (1995). Overexpression of glutathione reductase but not glutathione synthetase leads to increases in antioxidant capacity and resistance to photoinhibition in poplar trees. *Plant Physiology, 109,* 1047–1057.

Freedman, B., et al. (2017). Screening for atrial fibrillation: A report of the AF-SCREEN international collaboration. *Circulation, 135,* 1851–1867.

Garcíacañaveras, J. C., Donato, M. T., Castell, J. V., & Lahoz, A. (2012). Targeted profiling of circulating and hepatic bile acids in human, mouse, and rat using a UPLC-MRM-MS-validated method. *Journal of Lipid Research, 53,* 2231–2241.

Go, A. S., et al. (2001). Prevalence of diagnosed atrial fibrillation in adults: National implications for rhythm management and stroke prevention: The AnTicoagulation and Risk Factors in Atrial Fibrillation (ATRIA) Study. *JAMA, 285,* 2370–2375.

Gstaiger, M., & Aebersold, R. (2009). Applying mass spectrometry-based proteomics to genetics, genomics and network biology. *Nature Reviews Genetics, 10,* 617–627.

Harada, M., Melka, J., Sobue, Y., & Nattel, S. (2017). Metabolic considerations in atrial fibrillation—mechanistic insights and therapeutic opportunities. *Circulation Journal, 81,* 1749–1757.

Hyman, M. C., & Deo, R. (2017). Proteomics of atrial fibrillation: Evolving From a coarse understanding to a fine phenotype. *JAMA Cardiology, 2,* 474.

Iwasaki, Y. K., Nishida, K., Kato, T., & Nattel, S. (2011). Atrial fibrillation pathophysiology: Implications for management. *Circulation, 124,* 2264–2274.

January, C. T., et al. (2014). 2014 AHA/ACC/HRS guideline for the management of patients with atrial fibrillation: A report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and the Heart Rhythm Society. *Journal of the American College of Cardiology, 64,* e1–76.

Jung, Y., et al. (2018). Lipidomic profiling reveals free fatty acid alterations in plasma from patients with atrial fibrillation. *PLoS ONE, 13,* e0196709.

Kaarisalo, M. M., et al. (1997). Atrial fibrillation and stroke. Mortality and causes of death after the first acute ischemic stroke. *Stroke, 28,* 311–315.

Ko, D., et al. (2016). Metabolomic profiling in relation to new-onset atrial fibrillation (from the framingham heart study). *American Journal of Cardiology, 118,* 1493–1496.

Lau, J., et al. (2013). Performance of an automated iPhone ECG algorithm to diagnose atrial fibrillation in a community AF screening program (SEARCH-AF). *Heart Lung and Circulation, 22,* S205–S205.

Letsas, K. P., et al. (2010). Uric acid elevation in atrial fibrillation. *Hellenic Journal of Cardiology, 51,* 209–213.

Li, N., et al. (2016). Adenosine-induced atrial fibrillation. *Circulation, 134,* 486–498.

Liang, Z. Y., Pilonsmits, E. A., Jouanin, L., & Terry, N. (1999). Overexpression of glutathione synthetase in Indian mustard enhances cadmium accumulation and tolerance. *Plant Physiology, 119,* 73–79.

Liu, T., Zhang, X., Korantzopoulos, P., Wang, S., & Li, G. (2011). Uric acid levels and atrial fibrillation in hypertensive patients. *Internal Medicine, 50,* 799–803.

Marazzi, G., et al. (2012). Comparison of microlife BP A200 plus and omron M6 blood pressure monitors to detect atrial fibrillation in hypertensive patients. *Advances in Therapy, 29,* 64–70.

Miyasaka, Y., et al. (2006). Secular trends in incidence of atrial fibrillation in Olmsted County, Minnesota, 1980 to 2000, and implications on the projections for future prevalence. *Circulation, 114,* 119–125.

Ong, S. E., & Mann, M. (2005). Mass spectrometry-based proteomics turns quantitative. *Nature Chemical Biology, 1,* 252–262.

Pandey, A., & Mann, M. (2000). Proteomics to study genes and genomes. *Nature, 405,* 837–846.

Peterson, A. C., Russell, J. D., Bailey, D. J., Westphall, M. S., & Coon, J. J. (2012). Parallel reaction monitoring for high resolution and high mass accuracy quantitative, targeted proteomics. *Molecular and Cellular Proteomics Mcp, 11,* 1475.

Rusnak, J., et al. (2018). Interventional left atrial appendage closure may affect metabolism of essential amino acids and bioenergetic efficacy. *International Journal of Cardiology, 268,* 125–131.

Saito, K., & Matsuda, F. (2010). Metabolomics for functional genomics, systems biology, and biotechnology. *Annual Review of Plant Biology, 61,* 463–489.

Sattler, K., et al. (2017). Occlusion of left atrial appendage affects metabolomic profile: Focus on glycolysis, tricarboxylic acid and urea metabolism. *Metabolomics, 13,* 127.

Sievers, B., Kirchberg, S., Addo, M., Bakan, A., Brandts, B., & Trappe, H. J. (2004). Assessment of left atrial volumes in sinus rhythm and atrial fibrillation using the biplane area-length method and cardiovascular magnetic resonance imaging with TrueFISP. *Journal of Cardiovascular Magnetic Resonance, 6,* 855–863.

Steger, M., et al. (2016). Phosphoproteomics reveals that Parkinson's disease kinase LRRK2 regulates a subset of Rab GTPases. *eLife, 5,* e12813.

Suzuki, S., et al. (2012). Gender-specific relationship between plasma uric acid level and atrial fibrillation prevalence. *Circulation Journal, 76,* 607–611.

Tsugawa, H., et al. (2015). MS-DIAL: Data-independent MS/MS deconvolution for comprehensive metabolome analysis. *Nature Methods, 12,* 523–526.

Tyanova, S., Temu, T., & Cox, J. (2016). The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nature Protocols, 11,* 2301.

Vink, R., et al. (2005). Risk of thromboembolism and bleeding after general surgery in patients with atrial fibrillation. *American Journal of Cardiology, 96,* 822–824.

Watanabe, E. (2012). Uric acid and atrial fibrillation. *Circulation Journal, 76,* 584–585.

Wishart, D. S. (2016). Emerging applications of metabolomics in drug discovery and precision medicine. *Nature Reviews Drug Discovery, 15,* 473–484.

Woods, C. E., & Olgin, J. (2014). Atrial fibrillation therapy now and in the future: Drugs, biologicals, and ablation. *Circulation Research, 114,* 1532–1546.

Zeller, T., et al. (2015). Metabolomics and incidence of atrial fibrillation in African Americans: The atherosclerosis risk in communities (ARIC) study. *PLoS ONE, 10,* e0142610.

Zhou, J., Liu, H., Liu, Y., Liu, J., Zhao, X., & Yin, Y. (2016). Development and evaluation of a parallel reaction monitoring strategy for large-scale targeted metabolomics quantification. *Analytical Chemistry, 88,* 4478–4486.

Zhou, J., Liu, C., Si, D., Jia, B., Zhong, L., & Yin, Y. (2017). Workflow development for targeted lipidomic quantification using parallel reaction monitoring on a quadrupole-time of flight mass spectrometry. *Analytica Chimica Acta, 972,* 62.

Zhou, J., & Yin, Y. (2016). Strategies for large-scale targeted metabolomics quantification by liquid chromatography-mass spectrometry. *Analyst, 141,* 6362.