



# Evaluation of Untargeted Metabolomic Strategy for the Discovery of Biomarker of Breast Cancer

Xujun Ruan<sup>1†</sup>, Yan Wang<sup>2†</sup>, Lirong Zhou<sup>2</sup>, Qiuling Zheng<sup>2\*</sup>, Haiping Hao<sup>1\*</sup> and Dandan He<sup>3\*</sup>

<sup>1</sup>Key Laboratory of Drug Metabolism and Pharmacokinetics, State Key Laboratory of Natural Medicines, China Pharmaceutical University, Nanjing, China, <sup>2</sup>Department of Pharmaceutical Analysis, College of Pharmacy, China Pharmaceutical University, Nanjing, China, <sup>3</sup>Experimental Center of Molecular and Cellular Biology, The Public Laboratory Platform, China Pharmaceutical University, Nanjing, China

## OPEN ACCESS

### Edited by:

Guo Ma,  
Fudan University, China

### Reviewed by:

Lei Zhou,  
Guangdong Pharmaceutical  
University, China  
Qinglin Li,  
University of Chinese Academy of  
Sciences, China  
Yan-Fang Xian,  
The Chinese University of Hong Kong,  
China  
Changhui Liu,  
Guangzhou University of Chinese  
Medicine, China

### \*Correspondence:

Qiuling Zheng  
qiuling\_zheng@cpu.edu.cn  
Haiping Hao  
haipinghao@cpu.edu.cn  
Dandan He  
cpudandanhe@163.com

<sup>†</sup>These authors have contributed  
equally to this work

### Specialty section:

This article was submitted to  
Drug Metabolism and Transport,  
a section of the journal  
Frontiers in Pharmacology

**Received:** 11 March 2022

**Accepted:** 25 April 2022

**Published:** 30 May 2022

### Citation:

Ruan X, Wang Y, Zhou L, Zheng Q,  
Hao H and He D (2022) Evaluation of  
Untargeted Metabolomic Strategy for  
the Discovery of Biomarker of  
Breast Cancer.  
Front. Pharmacol. 13:894099.  
doi: 10.3389/fphar.2022.894099

Discovery of disease biomarker based on untargeted metabolomics is informative for pathological mechanism studies and facilitates disease early diagnosis. Numerous of metabolomic strategies emerge due to different sample properties or experimental purposes, thus, methodological evaluation before sample analysis is essential and necessary. In this study, sample preparation, data processing procedure and metabolite identification strategy were assessed aiming at the discovery of biomarker of breast cancer. First, metabolite extraction by different solvents, as well as the necessity of vacuum-dried and re-dissolution, was investigated. The extraction efficiency was assessed based on the number of eligible components (components with MS/MS data acquired), which was more reasonable for metabolite identification. In addition, a simplified data processing procedure was proposed involving the OPLS-DA, primary screening for eligible components, and secondary screening with constraints including VIP, fold change and *p* value. Such procedure ensured that only differential candidates were subjected to data interpretation, which greatly reduced the data volume for database search and improved analysis efficiency. Furthermore, metabolite identification and annotation confidence were enhanced by comprehensive consideration of mass and MS/MS errors, isotope similarity, fragmentation match, and biological source confirmation. On this basis, the optimized strategy was applied for the analysis of serum samples of breast cancer, according to which the discovery of differential metabolites highly encouraged the independent biomarkers/indicators used for disease diagnosis and chemotherapy evaluation clinically. Therefore, the optimized strategy simplified the process of differential metabolite exploration, which laid a foundation for biomarker discovery and studies of disease mechanism.

**Keywords:** untargeted metabolomics, strategy evaluation, biomarker discovery, breast cancer, UPLC-MS

## INTRODUCTION

Metabolites have been realized to play an important role in the onset of diseases, and are of great significance for disease diagnosis and prevention. Metabolomics is attracting increasing attentions in various areas, such as pathological mechanism studies, pathway analysis, and the exploration of novel biomarkers for diseases, including cancers (Armitage and Southam, 2016; Kumar and Misra,

2019; Liu et al., 2021; Long et al., 2021). According to different research goals, there comes up with untargeted metabolomics and targeted metabolomics (Jasbi et al., 2019; Casari et al., 2021; Baek et al., 2022; Harrieder et al., 2022). Targeted metabolomics concentrates more on the analysis of defined metabolites, which has limited scope but achieves enhanced detection sensitivity and enables the absolute quantification with the application of standards (Roberts et al., 2012; Cai et al., 2015; Zhou and Yin, 2016). Untargeted metabolomics, on the other hand, has superiority in high-throughput detection, which offers a comprehensive and in-depth insight of metabolome profiling (Yuan et al., 2012; Heiles, 2021) and is regarded as the basis for biomarker discovery.

Mass spectrometry (MS) has been regarded as a powerful analytical technique owing to its high detection speed and sensitivity. Its coupling with chromatographic separation, such as gas chromatography and liquid chromatography (LC), has been widely used for metabolomic analysis by providing both molecular weight and structural information (Liu et al., 2013; Alonso et al., 2015; Nash and Dunn, 2019; Hou et al., 2020; Harrieder et al., 2022). Efforts have been done in method development for sample preparation, chromatographic separation and derivatization-based detection to improve the metabolome coverage (Yuan et al., 2018; An et al., 2021; Meng et al., 2021). Besides, computational approaches are also dramatically developed to assist data interpretation and metabolite global annotation (Bonini et al., 2020; Chen et al., 2021; Duehrkop et al., 2021).

Breast cancer has become one of the leading causes threatening health in women and its incidence is increasing within recent years. With the development of new therapeutic strategies, the mortality of breast cancer has gradually reduced (DeSantis et al., 2019). Metabolomics has been widely applied for metabolic pathway analysis and biomarker discovery for breast cancer based on the analysis of different biological samples, including cell lines, plasma, serum, tissues, urine and saliva (Tsutsui et al., 2013; Tenori et al., 2015; Zhong et al., 2016; Porto-Figueira et al., 2018), which facilitates the early diagnosis, treatment target exploration and mechanism studies of the disease (Günther, 2015; McCartney et al., 2018; Park et al., 2019; Silva et al., 2019; Long et al., 2021). Methodologies with various sample preparation steps, detection methods, data processing and metabolite annotation procedures were developed, upon which the evaluation is essential before sample analysis due to the difference of sample property or experimental purpose. Herein, sample preparation, data processing procedure, and metabolite identification strategy of untargeted metabolomics were evaluated and subsequently applied for the analysis of serum samples of breast cancer. First, metabolite extraction by different solvents was assessed and evaluation based on the number of components with MS/MS data acquired (defined as eligible components) was regarded to be more reasonable in consideration of metabolite identification. In addition, a simplified data processing procedure was proposed involving orthogonal projections to latent structures discriminant analysis (OPLS-DA) for all detected components, followed up with a primary screening based on the availability of MS/MS data

and underwent a secondary screening with criteria of VIP, fold change (FC) and *p* value. Thus, only differential candidates were subjected to database search, identification and annotation, which greatly reduced the data volume and improved the analysis efficiency. Furthermore, the confidence and accuracy of metabolite identification were enhanced by comprehensively considering mass and MS/MS errors, isotope similarity, fragmentation match and biological source confirmation. On this basis, the evaluated strategy was applied for the analysis of serum samples of breast cancer, upon which the discovery of potential biomarkers would be informative for early diagnosis and chemotherapeutic evaluation of the disease.

## MATERIALS AND METHODS

### Chemicals and Reagents

Formic acid (FA) was purchased from Sigma-Aldrich (Saint Louis, MO, United States). Ultra-pure H<sub>2</sub>O was prepared by a Milli-Q Pure Water System (Bedford, MA, United States). LC-MS grade methanol (MeOH) and acetonitrile (ACN) were purchased from Merck (Darmstadt, Germany).

### Clinical Sample Collection and Sample Preparation

Serum samples were collected from Huashan Hospital, Fudan University and stored at  $-80^{\circ}\text{C}$  before analysis. The research protocol was approved by the Ethical Committee of Huashan Hospital, Fudan University (KY2021-034), and written informed consents were provided by all participants.

For sample preparation, 400  $\mu\text{l}$  of designed solvent (pre-cooled on ice) was mixed with 100  $\mu\text{l}$  of serum sample and followed up with a 2 min vortex for sufficient extraction and protein precipitation. The obtained mixture was centrifuged at 15,000 g (5 min at  $4^{\circ}\text{C}$ ) and directly stored at  $4^{\circ}\text{C}$  (1 h) to ensure the complete protein precipitation. Supernatant (200  $\mu\text{l}$ ) was transferred for an additional centrifugation (18,000 g for 5 min at  $4^{\circ}\text{C}$ ). Vacuum-dried and re-dissolution by extraction solvent was performed prior to MS analysis if necessary. Quality control (QC) sample was prepared by pooling aliquots of each serum sample investigated in this study, including 58 samples from breast cancer vs. healthy control and 12 samples from breast cancer patients before and after chemotherapy. The prepared QC sample was applied for extraction evaluation and instrumental performance monitoring.

### UPLC-MS

Chromatographic separation was performed by a Waters ACQUITY I-Class UPLC system equipped with an ACQUITY UPLC HSS T3 column (2.1  $\times$  100 mm, 1.8  $\mu\text{m}$ , Waters). Parameters were set as follows: column temperature  $40^{\circ}\text{C}$ ; flow rate of 0.4 ml/min; injection volume of 4  $\mu\text{l}$ ; mobile phase A was H<sub>2</sub>O containing 0.1% FA and B was ACN. The gradient elution condition was referenced by previous study (He et al., 2021): 0–1.00 min, 0% B; 1.01–4.00 min, 0–35% B; 4.01–15.50 min, 35–95% B; 15.51–18.00 min, maintaining at 95% B;

18.01–23.00 min, back to 0% B. MS detection and data acquisition were performed by a Q-TOF MS (Waters, Xevo, G2-XS QToF) in both positive and negative ion modes. MS parameters were set as follows:  $m/z$  range at 50–700 Da; capillary voltage at + 3.0 kV or –2.5 kV for positive or negative ion mode respectively; sampling cone at 40 V; source temperature at 110°C; desolvation temperature at 450°C; cone gas at 50 L/h, desolvation gas at 600 L/h. MassLynx (version 4.1, Waters) was used for data acquisition.

## Data Processing

Data processing was performed by Progenesis QI (version 2.0, Waters; denoted as QI), which mainly contained steps of: creation of a new experiment; data import; review alignment; experiment design setup; peak picking; review deconvolution; compounds identification; review compounds; and compound statistics. Notably, adduct ion forms, such as  $[M + H]^+$ ,  $[M-H]^-$ ,  $[M + Na]^+$ ,  $[M + K]^+$ ,  $[2M-H]^-$ , and  $[M + FA-H]^-$ , were all included. Different adduct ion forms of a same metabolite were automatically combined to provide an accurate identity. For metabolites having different adduct ion forms, the one with the highest MS intensity was selected to represent the abundance. Databases for metabolite identification and annotation included Human Metabolome Database (HMDB) (<http://www.hmdb.ca/>), MoNA (<http://mona.fiehnlab.ucdavis.edu/>), and METLIN (<https://metlin.scripps.edu/>). Parameters were set as follows: 10 ppm for precursor ion match; 20 ppm for fragment ion match; > 80% for the isotope similarity;  $\geq 40$  score of identification. The MS/MS spectra for potential biomarkers were manually checked to confirm the assigned identities. Relative quantification of selected metabolites by QI was performed by integration of corresponding extracted ion chromatograms (EICs) and normalized against the total ion chromatogram (TIC). Manually check of EIC was required to ensure the accurate peak picking and correct integration.

## Statistical Analysis

Components detected were exported by QI and subjected to EZinfo software (version 3.0, Waters) for OPLS-DA for statistical difference confirmation. Components with variable importance calculated based on S-plots ( $VIP > 1$ ,  $p$  value  $< 0.05$  (student's t-test) and  $FC > 1.5$ ) were screened. Notably, for the same substance detected under different ion modes, the one with the smallest FC were adapted for further data analysis. Bioinformatics analysis of assigned differential metabolites, including pathway analysis and receiver operating characteristic (ROC) curve, were performed on MetaboAnalyst (version 5.0) (<http://www.metaboanalyst.ca/>).

# RESULTS AND DISCUSSION

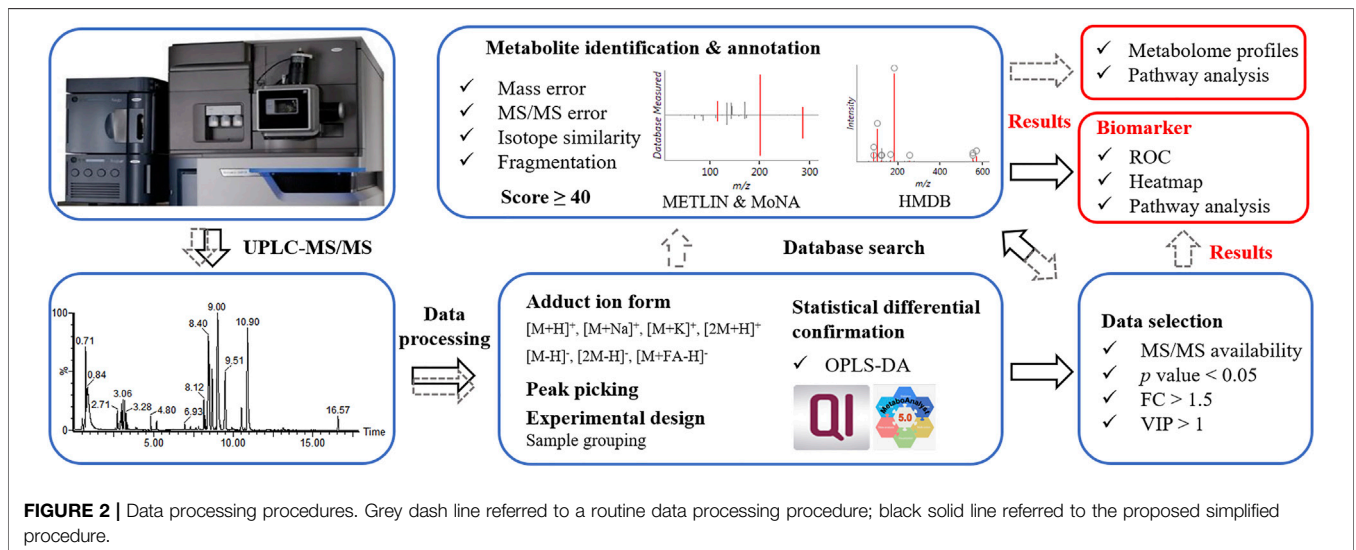
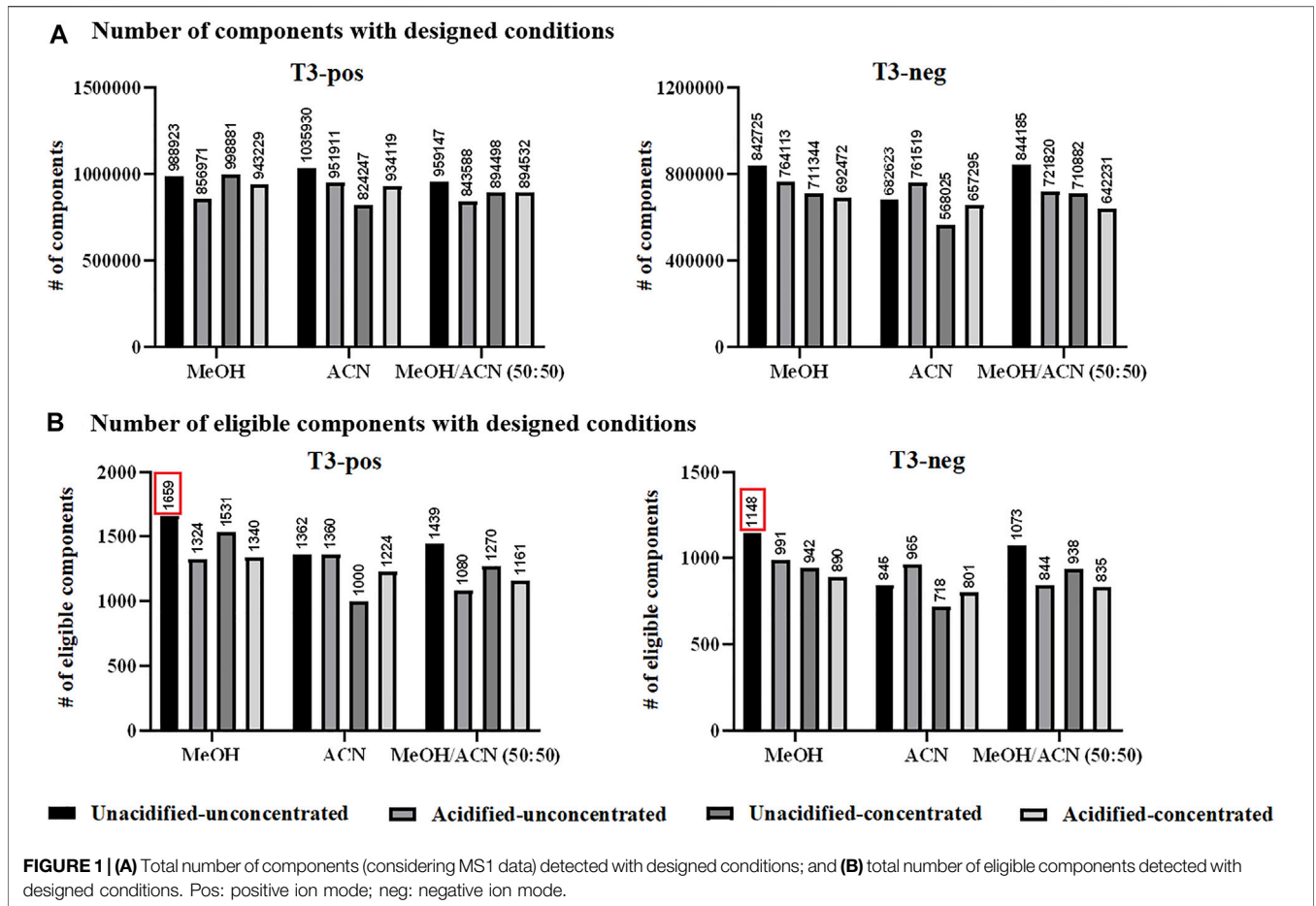
## Evaluation of Untargeted Metabolomics

**Sample preparation.** Metabolite extraction and protein precipitation by two commonly used solvents, MeOH and ACN, as well as their mixture (50:50 by volume) were

evaluated. The option of 0.1% FA adoption was also explored, as well as vacuum-dried concentration and re-dissolution considering metabolites with low abundance. The analysis of designed solvent-treated QC samples resulted to over 800,000 ions in the positive ion mode, among which ACN-extraction appeared to be the optimal one. In the negative ion mode, at least 500,000 ions were detected upon designed extraction conditions, and MeOH/ACN mixture had a higher efficiency (**Figure 1A**). Nevertheless, it was still hesitant to choose one appropriate solvent while such comparison was rough and inappropriate. Aiming at metabolite identification, evaluation based on eligible components was proposed to be more reasonable. Accordingly, eligible components obtained by designed conditions were in the range of 1,000–1,700 in the positive ion mode, among which MeOH-extraction was optimum with 1,659 components having MS/MS data. Similarly, extraction by designed solvents resulted to eligible components ranging from 700 to 1,200 in the negative ion mode, among which MeOH also emerged by having 1,148 components detected with MS/MS spectra (**Figure 1B**). Interestingly that no significant improvement was observed after concentration, which was probably due to the decomposition of unstable compounds during the tedious vacuum-dried step. Thus, MeOH extraction was applied for subsequent serum sample preparation and the obtained mixture after protein precipitation was directly subjected to MS analysis without further treatment. Moreover, the number of components varied with different extraction solvents further supported the importance of evaluation before sample analysis.

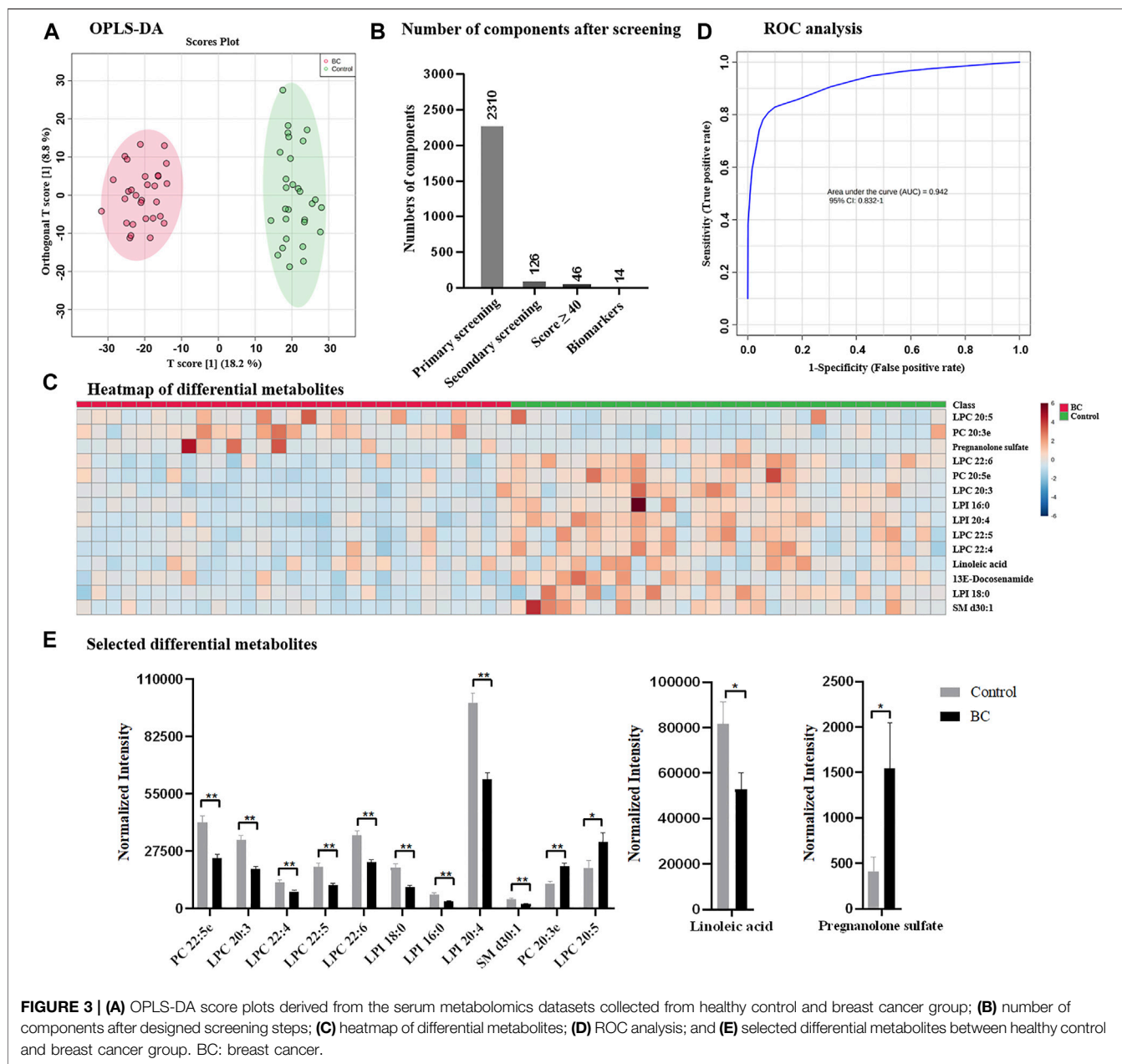
**Data processing procedure.** Upon UPLC-MS analysis, resulted components (considering MS1 data here) underwent OPLS-DA to ensure the statistical difference between tested groups. For routine procedure (**Figure 2**, in grey dash line), components with MS/MS data were subjected to database search, metabolite identification and differential analysis. Such procedure involved annotation and confirmation for every component, which was informative for revealing the difference in metabolome profiles and metabolomic pathways between tested samples. Aiming at the discovery of biomarker, we proposed a simplified data processing procedure with improved analysis efficiency. As shown in **Figure 2** (in black solid line), OPLS-DA was applied for multivariate analysis. Primary screening was performed to select eligible components, among which differential ones were further isolated by secondary screening with criteria of  $p$  value  $< 0.05$ ,  $VIP > 1$  and  $FC > 1.5$ . Thus, instead of massive identification, only differential candidates were allowed to subsequent database search and identification, which greatly narrowed down the data volume for interpretation.

**Metabolite identification strategy.** For metabolite identification and annotation, parameters including  $m/z$  error, fragment error, isotope similarity, and fragmentation match were mainly concerned. Fragmentation match was performed based on databases including METLIN, MoNA and HMDB. The former two contained MS/MS spectra that experimentally collected, while HMDB offered theoretical



fragmentation based on molecular structures. Metabolite identification was relied on the Score value from QI, which comprehensively evaluated the match degree of all above

mentioned parameters. Based on previous study (Hou et al., 2020), the annotation accuracy and confidence were enhanced by setting a threshold of Score  $\geq 40$  for acceptance, which was



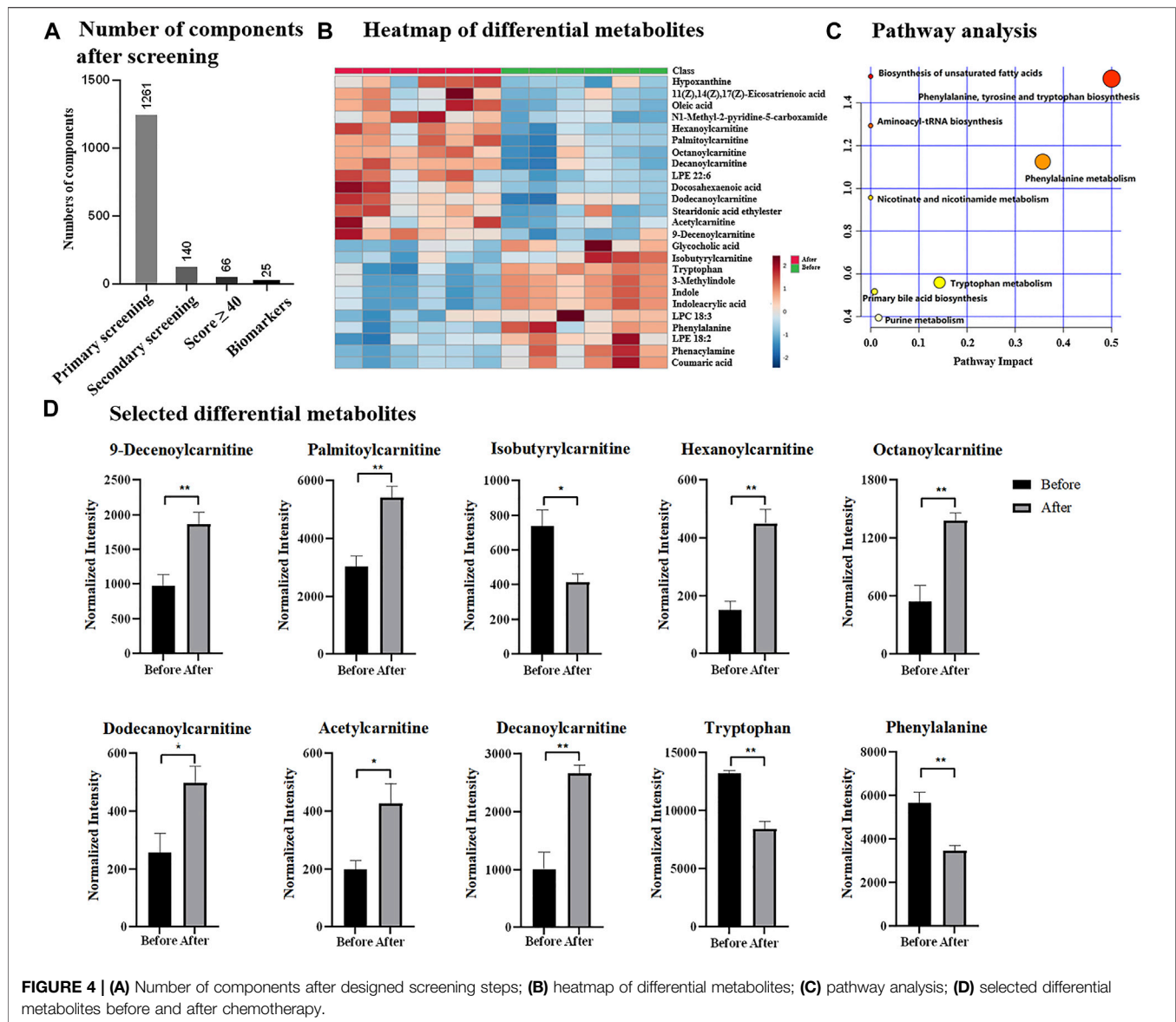
**FIGURE 3 | (A)** OPLS-DA score plots derived from the serum metabolomics datasets collected from healthy control and breast cancer group; **(B)** number of components after designed screening steps; **(C)** heatmap of differential metabolites; **(D)** ROC analysis; and **(E)** selected differential metabolites between healthy control and breast cancer group. BC: breast cancer.

higher than the one commonly used ( $\geq 35$ ). Moreover, biological source information provided complementary information for further confirmation, upon which the structural and biological information was integrated for identification and annotation. Notably that it was possible to exclude a small portion of potential biomarkers based on such data processing procedure and identification strategy, especially for those with MS1 data only or those were not included in the databases we used. However, it was known that the identification of such components was challenge due to limited information provided, which would decrease the annotation confidence and reliability. Thus, such potential

biomarkers would not be considered in this study as the identification and annotation with high confidence and accuracy were taken as the primary standard and purpose.

## Exploration of Potential Biomarker of Breast Cancer

Following the optimized strategy, 58 serum samples containing 29 from healthy control and 29 from breast cancer patients were investigated. The instrumental stability and reproducibility were assessed by QC sample based on five successive injections prior to sample analysis and interval injections during sample run



**FIGURE 4 | (A)** Number of components after designed screening steps; **(B)** heatmap of differential metabolites; **(C)** pathway analysis; **(D)** selected differential metabolites before and after chemotherapy.

(Supplementary Figure S1A). Upon UPLC-MS analysis, total 8,399 components (considering MS1 data only) were detected in the positive and negative ion modes. According to OPLS-DA, healthy control and breast cancer group exhibited distinguish metabolome profiles and were completely separated with  $R^2Y$  of 96.4% and  $Q^2$  of 83.3% (Figure 3A). Following the proposed simplified procedure, primary screening resulted to 2,310 eligible components with qualified MS/MS data and only 126 differential candidates were remained after secondary screening for subsequent database search (Figure 3B), which greatly reduced the data volume and saved efforts for data interpretation. The identification by QI further narrowed down the number of metabolites to 46 with Score value  $\geq 40$ . Upon biological source confirmation, a total of 14 endogenous metabolites were finally emerged (details found in Supplementary Table S1; Supplementary Figure S2).

Heatmap analysis in Figure 3C also confirmed that they had quantitative difference between healthy control and breast cancer group, according to which a considerable number of metabolites were observed down-regulated in breast cancer group. ROC curve analysis was applied to verify the representative of annotated metabolites and AUC of 0.942 in Figure 3D confirmed the reliability of identified metabolites, which implied them as potential biomarkers for clinical diagnosis of breast cancer. Differential metabolites mainly belonged to lipids, fatty acid, and fatty amide, some of which were also reported previously (Long et al., 2021) (Figure 3E). Lipids accounted for the majority of differential metabolites, including two phosphorylcholines (PCs), five lysophosphatidylcholines (LPCs) and three lysophosphatidylinositol (LPIs), which were also reported to be closely related to the occurrence of breast cancer (Cala et al., 2018;

Song et al., 2020; Long et al., 2021). Besides, linoleic acid, down-regulated in breast cancer group, was reported to be closely related to the regulation of breast cancer involved miRNA expression (Elieh Ali Komi et al., 2021). Pregnanolone sulfate, significantly up-regulated in breast cancer group based on our results, suggested a possible relationship between breast cancer and steroids. (Bicikova et al., 2001). It was worth to mention that the sample size enrolled in this study was limited, which might lead to insufficient discovery and miss other potential biomarkers. A larger sample population or detailed classification in subtypes, disease progression or treatment would be more expected.

Chemotherapy is one of major treatments for breast cancer, thus, the discovery of potential indicators is also essential for monitoring the process or effect of chemotherapy. Herein, total 12 serum samples collected from six breast cancer patients were investigated, including six samples before chemotherapeutic treatment and six samples after chemotherapeutic treatment. Upon UPLC-MS analysis, total 6,130 components (considering MS1 data only) were detected and based OPLS-DA (**Supplementary Figure S3A**), two groups were statistically different by receiving a  $R^2Y$  of 99.8% and  $Q^2$  of 77.8%. With the simplified data processing procedure, 1,261 eligible components with qualified MS/MS data were isolated, while 140 of them met the criteria of secondary screening and were remained as differential candidates for database search. Based on QI identification, 66 components received Score value  $\geq 40$ , and 25 of them were finally emerged as potential biomarkers according to biological source confirmation (**Figure 4A**, details found in **Supplementary Table S2** and **Supplementary Figure S4**). In addition, heatmap analysis illustrated significant quantitative difference among two groups, including 11 metabolites were down-regulated and 14 were up-regulated in chemotherapeutic group (**Figure 4B** and **Supplementary Table S2**). ROC analysis in **Supplementary Figure S3B** confirmed them as representative and potential indicators for monitoring the process or treatment of chemotherapy by achieving AUC of 1.00. Notably that the AUC value of 1.00 was not common, which was probably due to the indeed significant difference of potential metabolites before and after chemotherapeutic treatment. Besides, it was also because of limited sample size, which resulted to overfitting during ROC analysis. The identified potential indicators mainly distributed in biosynthesis of unsaturated fatty acids, aminoacyl-tRNA biosynthesis, nicotinate and nicotinamide metabolism, as well as phenylalanine, tyrosine and tryptophan biosynthesis (**Figure 4C**), some of which had also been reported previously as biomarkers for breast cancer diagnosis (Long et al., 2021). Among differential metabolites included eight carnitine species, suggesting the possible dysregulation of carnitine metabolism and fatty acid  $\beta$ -oxidation process. Phenylalanine and tryptophan were reported previously to be dysregulated and were highly expressed in breast cancer based on tissue or cell investigation (Du et al., 2019; Long et al., 2021). Similarly, indole was also considered as one of biomarkers for breast cancer and discovered to be up-regulated based on plasma analysis (Jasbi et al., 2019). They were observed down-regulated after chemotherapy

according to results in **Figure 4D** and **Supplementary Table S2**, suggesting them as potential indicators to track chemotherapeutic treatment and disease recovery. Notably that differential metabolites identified in this study had overlap but not identical with previous studies, which could be attributed to the difference in sample types (tissue, serum or cultured cells). In addition, the differential metabolites identified for chemotherapeutic samples were different from that of breast cancer vs. healthy group, suggesting the possibility that the potential biomarkers for breast cancer diagnosis and chemotherapeutic monitoring were independent. It also implied that for clinical applications, the same potential biomarkers for breast cancer diagnosis were not sufficient for evaluation of chemotherapy, as separated metabolic pathways were probably affected. Thus, the discovery of indicators specified for the evaluation of chemotherapy were highly encouraged and large sample size would be more expected, especially for samples collected from the same subtype, progression of the disease or treatment.

## CONCLUSION

In this study, sample preparation, data processing procedure and metabolite identification strategy of untargeted metabolomics were evaluated. Metabolite extraction by different solvents was investigated and reasonable assessment based on the number of eligible components was proposed. A simplified data processing procedure was proposed, involving OPLS-DA for statistical difference confirmation, primary screening based on MS/MS data availability and secondary screening according to criteria including FC, VIP and  $p$  value. Such procedure allowed database search for differential components only, which greatly narrowed down the data volume, improved the analysis efficiency and facilitated to reduce false identification results. For identification and annotation, mass error, MS/MS error, isotope similarity and fragmentation match were comprehensively considered and Score  $\geq 40$  was set to enhance the identification confidence and accuracy. The evaluated strategy was applied for the analysis of serum samples of breast cancer and the discovery of distinguished metabolites highly encouraged the exploration of independent biomarkers for disease diagnosis and treatment clinically. Therefore, the evaluated strategy was beneficial for the discovery of potential biomarker with the simplified procedure and unambiguous annotation.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding authors.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethical Committee of Huashan Hospital, Fudan

University. The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

XR: Data curation; Methodology; Software; Validation; Writing—original draft; YW: Data curation; Software; Validation LZ: Data curation; Software QZ: Supervision; Writing—review and editing HH: Supervision, Writing review and editing, Conceptualization DH: Supervision, Writing review and editing, Conceptualization, Visualization.

## FUNDING

This work was supported by the National Natural Science Foundation of China (81720108032, 81930109), the National

## REFERENCES

- Alonso, A., Marsal, S., and Julià, A. (2015). Analytical Methods in Untargeted Metabolomics: State of the Art in 2015. *Front. Bioeng. Biotechnol.* 3, 23. doi:10.3389/fbioe.2015.00023
- An, N., Zhu, Q. F., Wang, Y. Z., Xiong, C. F., Hu, Y. N., and Feng, Y. Q. (2021). Integration of Chemical Derivatization and In-Source Fragmentation Mass Spectrometry for High-Coverage Profiling of Submetabolomes. *Anal. Chem.* 93 (32), 11321–11328. doi:10.1021/acs.analchem.1c02673
- Armitage, E. G., and Southam, A. D. (2016). Monitoring Cancer Prognosis, Diagnosis and Treatment Efficacy Using Metabolomics and Lipidomics. *Metabolomics* 12 (10), 1–15. doi:10.1007/s11306-016-1093-7
- Baek, J., He, C., Afshinnia, F., Michailidis, G., and Pennathur, S. (2022). Lipidomic Approaches to Dissect Dysregulated Lipid Metabolism in Kidney Disease. *Nat. Rev. Nephrol.* 18 (1), 38–55. doi:10.1038/s41581-021-00488-2
- Biciková, M., Számel, I., Hill, M., Tallová, J., and Stárka, L. (2001). Allopregnanolone, Pregnenolone Sulfate, and Epitestosterone in Breast Cyst Fluid. *Steroids* 66 (1), 55–57. doi:10.1016/s0039-128x(00)00140-9
- Bonini, P., Kind, T., Tsugawa, H., Barupal, D. K., and Fiehn, O. (2020). Retip: Retention Time Prediction for Compound Annotation in Untargeted Metabolomics. *Anal. Chem.* 92 (11), 7515–7522. doi:10.1021/acs.analchem.9b05765
- Cai, Y., Weng, K., Guo, Y., Peng, J., and Zhu, Z.-J. (2015). An Integrated Targeted Metabolomic Platform for High-Throughput Metabolite Profiling and Automated Data Processing. *Metabolomics* 11 (6), 1575–1586. doi:10.1007/s11306-015-0809-4
- Cala, M. P., Aldana, J., Medina, J., Sánchez, J., Guio, J., Wist, J., et al. (2018). Multiplatform Plasma Metabolic and Lipid Fingerprinting of Breast Cancer: a Pilot Control-Case Study in Colombian Hispanic Women. *PLoS One* 13 (2), e0190958. doi:10.1371/journal.pone.0190958
- Casari, I., Manfredi, M., Metharom, P., and Falasca, M. (2021). Dissecting Lipid Metabolism Alterations in SARS-CoV-2. *Prog. Lipid Res.* 82, 101092. doi:10.1016/j.plipres.2021.101092
- Chen, L., Lu, W., Wang, L., Xing, X., Chen, Z., Teng, X., et al. (2021). Metabolite Discovery through Global Annotation of Untargeted Metabolomics Data. *Nat. Methods* 18, 1377–1385. doi:10.1038/s41592-021-01303-3
- DeSantis, C. E., Ma, J., Gaudet, M. M., Newman, L. A., Miller, K. D., Goding Sauer, A., et al. (2019). Breast Cancer Statistics, 2019. *CA Cancer J. Clin.* 69 (6), 438–451. doi:10.3322/caac.21583
- Du, S., Wang, Y., Alatrash, N., Weatherly, C. A., Roy, D., MacDonnell, F. M., et al. (2019). Altered Profiles and Metabolism of L- and D-Amino Acids in Cultured Human Breast Cancer Cells vs. Non-tumorigenic Human Breast Epithelial Cells. *J. Pharm. Biomed. Anal.* 164, 421–429. doi:10.1016/j.jpba.2018.10.047
- Key Research and Development Programme of China (2021YFA1301300), the Project for Major New Drug Innovation and Development (2018ZX09711001-002-003), and the Project of State Key Laboratory of Natural Medicines at China Pharmaceutical University (SKLNMZZ202020).

## ACKNOWLEDGMENTS

We thank Chengyu Chu from Huashan Hospital, Fudan University for clinical sample collection.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fphar.2022.894099/full#supplementary-material>

- Dührkop, K., Nothias, L. F., Fleischauer, M., Reher, R., Ludwig, M., Hoffmann, M. A., et al. (2021). Systematic Classification of Unknown Metabolites Using High-Resolution Fragmentation Mass Spectra. *Nat. Biotechnol.* 39 (4), 462–471. doi:10.1038/s41587-020-0740-8
- Elieh Ali Komi, D., Shekari, N., Soofian-Kordkandi, P., Javadian, M., Shanehbandi, D., Baradaran, B., et al. (2021). Docosahexaenoic Acid (DHA) and Linoleic Acid (LA) Modulate the Expression of Breast Cancer Involved miRNAs in MDA-MB-231 Cell Line. *Clin. Nutr. ESPEN* 46, 477–483. doi:10.1016/j.clnesp.2021.09.006
- Günther, U. L. (2015). Metabolomics Biomarkers for Breast Cancer. *Pathobiology* 82 (3-4), 153–165. doi:10.1159/000430844
- Harrieder, E.-M., Kretschmer, F., Böcker, S., and Witting, M. (2022). Current State-Of-The-Art of Separation Methods Used in LC-MS Based Metabolomics and Lipidomics. *J. Chromatogr. B* 1188, 123069. doi:10.1016/j.jchromb.2021.123069
- He, D., Su, Y., Meng, D., Wang, X., Wang, J., and Ye, H. (2021). A Pilot Study Optimizing Metabolomic and Lipidomic Acquisition in Serum for Biomarker Discovery in Nonalcoholic Fatty Liver Disease. *J. Mass Spectrom. Adv. Clin. Lab.* 22, 17–25. doi:10.1016/j.jmsacl.2021.10.001
- Heiles, S. (2021). Advanced Tandem Mass Spectrometry in Metabolomics and Lipidomics-Methods and Applications. *Anal. Bioanal. Chem.* 413 (24), 5927–5948. doi:10.1007/s00216-021-03425-1
- Hou, Y., He, D., Ye, L., Wang, G., Zheng, Q., and Hao, H. (2020). An Improved Detection and Identification Strategy for Untargeted Metabolomics Based on UPLC-MS. *J. Pharm. Biomed. Anal.* 191, 113531. doi:10.1016/j.jpba.2020.113531
- Jasbi, P., Wang, D., Cheng, S. L., Fei, Q., Cui, J. Y., Liu, L., et al. (2019). Breast Cancer Detection Using Targeted Plasma Metabolomics. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* 1105, 26–37. doi:10.1016/j.jchromb.2018.11.029
- Kumar, A., and Misra, B. B. (2019). Challenges and Opportunities in Cancer Metabolomics. *Proteomics* 19 (21-22), e1900042. doi:10.1002/pmic.201900042
- Liu, L., Wang, M., Yang, X., Bi, M., Na, L., Niu, Y., et al. (2013). Fasting Serum Lipid and Dehydroepiandrosterone Sulfate as Important Metabolites for Detecting Isolated Postchallenge Diabetes: Serum Metabolomics via Ultra-high-performance LC-MS. *Clin. Chem.* 59 (9), 1338–1348. doi:10.1373/clinchem.2012.200527
- Liu, L., Zhao, J., Zhang, R., Wang, X., Wang, Y., Chen, Y., et al. (2021). Serum Untargeted Metabolomics Delineates the Metabolic Status in Different Subtypes of Non-alcoholic Fatty Liver Disease. *J. Pharm. Biomed. Anal.* 200, 114058. doi:10.1016/j.jpba.2021.114058
- Long, N. P., Heo, D., Kim, H. Y., Kim, T. H., Shin, J. G., Lee, A., et al. (2021). Metabolomics-guided Global Pathway Analysis Reveals Better Insights into the Metabolic Alterations of Breast Cancer. *J. Pharm. Biomed. Anal.* 202, 114134. doi:10.1016/j.jpba.2021.114134



- McCartney, A., Vignoli, A., Biganzoli, L., Love, R., Tenori, L., Luchinat, C., et al. (2018). Metabolomics in Breast Cancer: A Decade in Review. *Cancer Treat. Rev.* 67, 88–96. doi:10.1016/j.ctrv.2018.04.012
- Meng, X., Pang, H., Sun, F., Jin, X., Wang, B., Yao, K., et al. (2021). Simultaneous 3-Nitrophenylhydrazine Derivatization Strategy of Carbonyl, Carboxyl and Phosphoryl Submetabolome for LC-MS/MS-Based Targeted Metabolomics with Improved Sensitivity and Coverage. *Anal. Chem.* 93 (29), 10075–10083. doi:10.1021/acs.analchem.1c00767
- Nash, W. J., and Dunn, W. B. (2019). From Mass to Metabolite in Human Untargeted Metabolomics: Recent Advances in Annotation of Metabolites Applying Liquid Chromatography-Mass Spectrometry Data. *TrAC Trends Anal. Chem.* 120, 115324. doi:10.1016/j.trac.2018.11.022
- Park, J., Shin, Y., Kim, T. H., Kim, D. H., and Lee, A. (2019). Plasma Metabolites as Possible Biomarkers for Diagnosis of Breast Cancer. *PLoS One* 14, e0225129. doi:10.1371/journal.pone.0225129
- Porto-Figueira, P., Pereira, J. A. M., and Câmara, J. S. (2018). Exploring the Potential of Needle Trap Microextraction Combined with Chromatographic and Statistical Data to Discriminate Different Types of Cancer Based on Urinary Volatome Biosignature. *Anal. Chim. Acta* 1023, 53–63. doi:10.1016/j.aca.2018.04.027
- Roberts, L. D., Souza, A. L., Gerszten, R. E., and Clish, C. B. (2012). Targeted Metabolomics. *Curr. Protoc. Mol. Biol.* Chapter 30, Unit 30.2.1–24. doi:10.1002/0471142727.mb3002s98
- Silva, C., Perestrelo, R., Silva, P., Tomás, H., and Câmara, J. S. (2019). Breast Cancer Metabolomics: from Analytical Platforms to Multivariate Data Analysis. A Review. *Metabolites* 9 (5), 102. doi:10.3390/metabo9050102
- Song, L., Liu, Z., Hu, H. H., Yang, Y., Li, T. Y., Lin, Z. Z., et al. (2020). Proto-oncogene Src Links Lipogenesis via Lipin-1 to Breast Cancer Malignancy. *Nat. Commun.* 11 (1), 5842. doi:10.1038/s41467-020-19694-w
- Tenori, L., Oakman, C., Morris, P. G., Gralka, E., Turner, N., Cappadona, S., et al. (2015). Serum Metabolomic Profiles Evaluated after Surgery May Identify Patients with Oestrogen Receptor Negative Early Breast Cancer at Increased Risk of Disease Recurrence. Results from a Retrospective Study. *Mol. Oncol.* 9 (1), 128–139. doi:10.1016/j.molonc.2014.07.012
- Tsutsui, H., Mochizuki, T., Inoue, K., Toyama, T., Yoshimoto, N., Endo, Y., et al. (2013). High-Throughput LC-MS/MS Based Simultaneous Determination of Polyamines Including N-Acetylated Forms in Human Saliva and the Diagnostic Approach to Breast Cancer Patients. *Anal. Chem.* 85 (24), 11835–11842. doi:10.1021/ac402526c
- Yuan, B. F., Zhu, Q. F., Guo, N., Zheng, S. J., Wang, Y. L., Wang, J., et al. (2018). Comprehensive Profiling of Fecal Metabolome of Mice by Integrated Chemical Isotope Labeling-Mass Spectrometry Analysis. *Anal. Chem.* 90 (5), 3512–3520. doi:10.1021/acs.analchem.7b05355
- Yuan, M., Breitkopf, S. B., Yang, X., and Asara, J. M. (2012). A Positive/negative Ion-Switching, Targeted Mass Spectrometry-Based Metabolomics Platform for Bodily Fluids, Cells, and Fresh and Fixed Tissue. *Nat. Protoc.* 7 (5), 872–881. doi:10.1038/nprot.2012.024
- Zhong, L., Cheng, F., Lu, X., Duan, Y., and Wang, X. (2016). Untargeted Saliva Metabonomics Study of Breast Cancer Based on Ultra Performance Liquid Chromatography Coupled to Mass Spectrometry with HILIC and RPLC Separations. *Talanta* 158, 351–360. doi:10.1016/j.talanta.2016.04.049
- Zhou, J., and Yin, Y. (2016). Strategies for Large-Scale Targeted Metabolomics Quantification by Liquid Chromatography-Mass Spectrometry. *Analyst* 141 (23), 6362–6373. doi:10.1039/c6an01753c

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Ruan, Wang, Zhou, Zheng, Hao and He. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.