Emerging Applications of Mass Spectrometry-Based Metabolic Fingerprinting in Clinics

Jing Cao, Yuning Wang, Yong Zhang,* and Kun Qian*

A number of recently developed mass spectrometry (MS)-based metabolic fingerprinting analyses are promoting metabolic analysis systems into practical clinical applications, including but not limited to disease diagnosis. Herein, recent cutting-edge research on mass spectrometry-based metabolic fingerprinting analytical methodology and clinical applications are described. These developments resolve challenges regarding the practical clinical applications of MS-based metabolic fingerprinting analysis systems, such as rapid signal readout, high throughput and direct MS analysis, and intelligent data mining of complex biological samples.

1. Introduction

Metabolomics has been considered an irreplaceable platform of biological sciences, which is the closest to the phenotype and most predictive of phenotype compared with other -omics such as genomics, transcriptomics, and proteomics. Metabolic fingerprinting is a commonly used approach towards untargeted metabolomics, aiming to compare the fingerprints of metabolites that change in response to alterations, such as disease, environment, toxin exposure, etc. The state of the system can be evaluated by revealing the composition, interaction, or change of the fingerprints of metabolites. For instance, metabolic fingerprinting is adapted as a diagnostic tool by comparing the metabolic fingerprinting between disease subjects and healthy controls.

Mass spectrometry (MS), nuclear magnetic resonance (NMR), and Raman spectroscopy are the most used analytical avenues for metabolic analysis in complex biological samples. Raman spectroscopy can provide molecular and structural information of molecules and be considered a potential tool for metabolic analysis in biological and biomedical research, possessing advantages of non-destructive, label free, and rapid. However, Raman spectroscopy suffers limited chemical resolution compared to NMR and MS. NMR spectroscopy is capable of identifying and quantitatively analyzing numerous biological samples owing to the chemical shift of nuclear spins. MS analysis can provide not only molecular weights information of the biomarkers but also their chemical or biological structures, which plays an irreplaceable role in the metabolomics field. NMR spectroscopy is highly reproducible but has relatively low sensitivity compared to MS platforms because of the signal overlapping at certain spectral regions, leading to larger sample size required. MS-based techniques achieved high sensitivity and thence low sample consumption, which is particularly meaningful for valuable clinical samples in clinics. MS-based metabolic fingerprinting has motivated numerous studies aiming at clinical applications, including cancer diagnosis, biomarker discovery, and tissue-type identification. The processing of MS big data is a key challenge of MS-based metabolic fingerprinting for clinical application with the necessary accuracy. Data mining, such as machine learning of omics information, has achieved a lot of success for clinical application. In this minireview, we provide a summary of state-of-the-art development in the field of medical application by MS-based metabolic fingerprinting in three years (Scheme 1).

2. Methodology

To maximize the clinical application performance, we first need to understand the workflow of metabolic fingerprinting. The methodology of metabolic fingerprinting should be planned as an integrated unit, including sample preparation, MS detection, and data analysis (Scheme 1).

A variety of biological samples have been applied in metabolic fingerprinting, including tissues and biological fluids, such as serum, plasma, saliva, urine, sputum, cerebrospinal fluid...
(CSF), and feces. Notably, sample preparation is highly dependent on the choice of MS techniques. Numerous MS techniques have been used for MS-based metabolomics applications, such as liquid/gas chromatography (LC/GC) MS, laser desorption ionization mass spectrometry (LDI) MS, and rapid evaporative ionization (REI) MS. Chromatography-MS (LC/GC MS) requires many sample pretreatment procedures (e.g., desalination, derivatization, and chromatography for metabolomics). LDI is a soft ionization process, which generates mainly singly charged gaseous ions from solid or liquid samples under the laser. LDI MS is tolerant of salts, proteins, and contaminants, which makes it free of complex and time-consuming sample pretreatments when dealing with complex biological samples. REI produces gaseous ions by thermal evaporation, which is a by-product of the electrosurgical process (i.e., radiofrequency wave). Furthermore, REI is an ambient ionization technique, thus can achieve direct real-time analysis of biological fluids with almost no sample pretreatment.

After MS data acquisition, multivariate statistical methods are required to interpret the information obtained from MS datasets, including unsupervised clustering and supervised machine learning (ML) models. Unsupervised methods, such as principal component analysis (PCA), cluster samples based on the variation of metabolic fingerprints without the prior knowledge of the sample label. Supervised ML (e.g., partial least squares discriminant analysis [PLS-DA], random forest [RF], elastic net [EN]) supports the construction of predictive models based on the learning of the data matrix (X) toward a matrix (Y) that contains label information, such as disease or healthy control. Due to the ability to rapidly process complex and heterogeneous data, ML has been immensely used for statistical analysis of MS data toward disease diagnosis and biomarker discovery. Notably, a power analysis is required before data analysis to determine the minimum sample size required of statistically validated data.

3. Clinical Application of Mass Spectrometry-Based Metabolic Fingerprinting

MS-based metabolic fingerprinting have been conducted on the following diseases: cancers (including brain, lung, pancreatic, colorectal, gastric, and gynecological cancer), inflammatory or infectious disease, kidney disease, and cardiovascular/cerebrovascular diseases. Compared with conventional diagnostic imaging and biopsy methods, MS-based metabolic fingerprinting affords desirable accuracy with minimum invasiveness and low costs.

3.1. LC/GC MS-Based Metabolic Fingerprinting

Coupled with chromatographic separation (e.g., GC, LC), MS has emerged as an essential instrument in the clinical laboratory for the qualitative or quantitative metabolic analysis of biological samples with high sensitivity and specificity. A combination of untargeted and targeted LC coupled tandem mass spectrometry was utilized by Song et al., interpreting the plasma metabolome (404 polar metabolites) and lipidome (998 lipids) quantitated by 71 internal standards in coronavirus disease 2019 (COVID-19) group and healthy group (Figure 1a). A plasma biomarker panel of 10 metabolites (with the smallest p-value) presents the meaningful biological functions, which were suitable for COVID-19 diagnosis, achieving an AUC of 0.975 by a logistic regression model. To identify potential early-stage chronic kidney disease (CKD) metabolite biomarkers, Chen et al. reported the high-throughput metabolomics of 2155 subjects by ultra-performance liquid chromatography coupled with mass spectrometry (UPLC MS) (Figure 1b). Five biomarkers were identified using an advanced bioinformatics approach (including machine learning and statistic feature reduction) and multistep validations. Notably, Chen et al. first identify 5-MTO as a critical biomarker for early-stage CKD detection. Yu et al. conducted untargeted serum metabolic analysis to discover potential biomarkers for gastric cancer diagnosis via UPLC MS. According to the filter criteria (p < 0.05 of t-test, fold change >1.5 or <2/3, and variable importance in the projection (VIP) >1 in PLS-DA), a panel of three metabolites was selected as diagnostic markers for gastric cancer, showing high potential diagnostic value with AUC of 0.986 by PLS-DA. Similarly, Ghosh et al. performed GC MS-based serum metabolic fingerprinting and systemic inflammatory profiling of asthma COPD overlap (ACO), finding eleven metabolites significantly regulated in ACO compared with COPD or asthma only by OPLS-DA, including fatty acids, energy metabolites, and cholesterol. To investigate the metabolic diversity of different cancer, Li et al. applied LC MS for metabolic profiling of 928 cell lines from more than 20 cancer types, confirming the existence of lineage distinctions in model cancer cell lines.

Besides disease diagnosis and mechanism investigation, chromatography-MS has also been applied in the research of...
drug metabolism study. To investigate the attenuation mechanism of Caowu (traditional Chinese medicine) compatibility with Yunnan Baiyao, Ren et al. [32] adopted UPLC MS-based serum metabolomics of male rats (control group and treatments groups), identifying 63 endogenous metabolites related to the potential toxicity of Caowu according to \( p < 0.05 \) in t-test and \( \text{VIP} > 1 \) in PLS-DA. Further, four core attenuated metabolic pathways were identified, revealing the possible mechanism of inflammatory inhibition. Zimmermann et al. [33] applied LC MS for identification of bacteria-produced drug metabolites by evaluating the capability of 76 human gut bacteria to metabolize 271 oral medicines for mapping drug metabolism of the human microbiome.

Chromatography-MS has been widely utilized as a conventional MS tool for disease diagnosis, disease mechanism investigation, and drug metabolism study toward clinical application. However, sample cleanup (e.g., desalination) and chromatographic separation (e.g., LC/GC) procedures lead to labor- and time-consuming sample preparation and large sample consumption, limiting the application of LC/GC MS for the point-of-care test.

### 3.2. LDI MS-Based Metabolic Fingerprinting

LDI MS has exhibited great potential for metabolic analysis in terms of the advantages of high throughput, fast analysis, and minimal sample preparation owing to the tolerance of salts, proteins, and contaminants toward large-scale clinical use. Noteworthy, the performance of LDI MS is highly dependent on the selection of matrix. The traditional organic matrices limit the metabolic analysis in low mass range (\( m/z < 700 \)) by LDI MS because of the strong background noise. Recently, LDI MS-based metabolic fingerprinting has been applied for various diseases diagnosis, owing to the development of well-designed nanocomposites as matrices and adaptive application of the machine learning algorithms for data analysis.

In recent years, cancer studies using LDI MS-based metabolic fingerprinting have been performed by many research groups...
(including our group). Cao et al.\textsuperscript{[8b]} reported a high-performance brain tumor diagnosis by machine learning of LDI MS-based metabolic fingerprinting with Pd-Au synthetic alloys as matrix, achieving diagnostic AUC of 0.917 with accuracy of 89.9\%, sensitivity of 94.0\%, and specificity of 85.7\% (Figure 2a). The optimized synthetic alloy enhanced the LDI efficacy and achieved direct detection of 100 nL of serum in seconds with high salt tolerance and protein endurance. Further, four biomarkers were selected with gradual changes to the healthy state during the radiotherapy process, indicating the potential of the platform for treatment response study. Huang et al.\textsuperscript{[20]} conducted machine learning of serum metabolic fingerprinting for early-stage lung adenocarcinoma (LA) diagnosis. Ferric particle-assisted LDI MS achieved direct metabolic patterns extraction using only 50 nL of serum with diagnostic AUC of 0.921, sensitivity of 90\%, and specificity of 93\% by sparse regression machine learning, which is superior to traditional imaging and biopsy methods. A panel of 7 biomarkers (i.e., uracil, uric acid, 3-hydroxyypicolinic acid) were confirmed beneficial for discriminating early-stage LA from healthy controls, achieving AUC of 0.894. Su et al.\textsuperscript{[8a]} designed mesoporous PdPtAu alloys as matrix with enhanced optoelectric/thermal effect and mesoporous surface morphology, yielding sensitivity of 92.0\% and specificity of 92.0\% for early diagnosis of gastric cancer. Pei et al.\textsuperscript{[25]} constructed a FeOOH@Metal-Organic Framework-assisted LDI MS platform to diagnose gynecological cancers with AUC of 0.921–0.997, sensitivity of 97.2–98.7\%, and specificity of 97.3–98.7\% using orthogonal partial least squares discriminant analysis (OPLS-DA) (Figure 2b). The FeOOH@Metal-Organic Framework composites showed selective metabolic analysis and enhanced ionization efficiency.

Figure 2. LDI MS-based metabolic fingerprinting for clinical application. a) Metabolic fingerprinting on synthetic alloys for medulloblastoma diagnosis and radiotherapy evaluation. Reproduced with permission.\textsuperscript{[8b]} Copyright 2020, John Wiley and Sons. b) FeOOH@MOF core-satellite nanocomposites for the serum metabolic fingerprinting of gynecological cancers. Reproduced with permission.\textsuperscript{[25]} Copyright 2020, John Wiley and Sons. c) COF@Au reveals specific serum metabolic fingerprints as the point of Crohn’s disease diagnosis. Reproduced with permission.\textsuperscript{[26]} Copyright 2021, John Wiley and Sons. d) Rapid deep learning-aided diagnosis of stroke by serum metabolic fingerprint based multi-modal recognition. Reproduced with permission.\textsuperscript{[27a]} Copyright 2020, John Wiley and Sons.
owing to size-exclusion effect and good light response, thereby achieving direct metabolic fingerprinting extraction using 1 µL of serum by LDI MS without any pretreatment.

Besides cancer diagnosis, LDI MS-based metabolic fingerprinting has also been widely used in the research of other diseases, such as Crohn’s disease and cardiovascular/cerebrovascular diseases. Yang et al.\cite{26} designed a covalent organic framework@Au (COF-V@Au) for serum metabolic fingerprinting of Crohn’s disease patients and healthy controls, achieving of AUC of 0.984 by OPLS-DA (Figure 2c). The COF-V@Au matrix possesses a large specific surface area and nano-porous structure to provide abundant sites for metabolites adsorption, achieving great ionization efficiency. Xu et al.\cite{27a} conducted deep learning to integrate serum metabolic fingerprinting extracted by nano-assisted LDI MS and clinical index of stroke patients, achieving enhanced AUC of 0.845 for stroke screening compared with only serum metabolic fingerprinting or clinical index (Figure 2d). This work highlighted the importance of multi-modal data integration. A novel deep learning model-based feature selection approach was developed, addressing the limitation of deep learning for feature selection. Zhang et al.\cite{36} established a laser-assisted LDI-REIMS-based salivary metabolic fingerprinting approach to provide accurate diagnostic results. This approach achieved discriminative metabolic fingerprinting (>90%) for type 2 diabetes and controls using OPLS-DA (Q² of 0.734) and successfully cross-validated by the time-consuming and redundant LC-MS, confirming the reliability of the LA-REIMS for fecal metabolic fingerprinting.

In summary, REIMS-based metabolic fingerprinting allowed the construction of real-time predictive models reflecting the metabolic perturbations of various diseases without sample pretreatment, resulting in a powerful candidate for intraoperative detection and large-scale screening and diagnosis.

4. Conclusion

Mass spectrometry is a powerful tool for metabolic fingerprinting of biological fluids or tissues for clinical application owing to its high sensitivity and resolution. This minireview intends to provide an overview of the research progress on MS-based metabolic fingerprinting in recent three years, focusing on three typical MS techniques (LC/GC MS, LDI MS, and REIMS) and their clinical applications.

LC/GC MS is a dominant tool for metabolic analysis toward disease diagnosis, disease mechanism investigation, and drug metabolism study. However, sample cleanup (e.g., desalination) and chromatographic separation (e.g., LC/GC) procedures lead to labor- and time-consuming sample preparation and large sample consumption, limiting the application of LC/GC MS for
point-of-care test. By contrast, LDI MS requires minimal sample preparation due to the selective and sensitive ionization with the assistance of the recently developed nanocomposites as matrix. LDI MS-based metabolic fingerprinting shows high potential for large-scale clinical application of various diseases, owing to the high throughput, rapid analysis, and high diagnostic performance. REIMS requires almost no sample pretreatment owing to the ambient ionization process. REIMS-based metabolic fingerprinting achieves real-time prediction of the metabolic perturbations of various diseases, making it a powerful candidate for large-scale screening and diagnosis, especially intraoperative tissue-type identification.

Although numerous progresses have been made by MS-based metabolic fingerprinting toward clinical application, there are still some challenges to be addressed. First, most of the recently developed applications by MS-based metabolic fingerprinting are focusing on a satisfactory diagnostic performance toward disease diagnosis. Expanding the application for therapeutic response assessment and prognosis is a promising direction for precision medicine, especially for LDI MS and REIMS. Second, it remains a great challenge for the identification and biological interpretation of metabolic biomarkers. With the selected signals as differential features, it is necessary to understand the biological mechanism of how the biomarkers interact with the biological systems. Furthermore, rather than the single metabolic fingerprinting analysis, the integration of multi-omics datasets (such as metabolomics, proteomics, and genomics) would be beneficial to interpret the complex biological systems for clinical application. Developing efficient multi-omics data acquisition platforms and intelligent data interpretation algorithms (such as machine learning) are efficient methods to address this issue, which will be a breakthrough point for MS-based metabolic fingerprinting in clinical use.

Acknowledgements

The authors are grateful for the financial support from Projects 2017YFE0124400, 2017YFC0909000, 2021YFF0703500, and 2021YFA0910104 by MOST, Projects 81971771 and 81771983 by NSFC, and Project 2021-01-00-02-E0053 by Shanghai Institutions of Higher Learning. This work was also sponsored by the Shanghai Rising-Star Programme (19QA1404800), Innovation Group Project of Shanghai Municipal Health Commission (2019CJ(Q03), Innovation...
Conflict of Interest
The authors declare no conflict of interest.

Keywords
clinics, fingerprinting, mass spectrometry, metabolites

Received: September 29, 2021
Revised: November 29, 2021
Published online: January 23, 2022
Jing Cao received her B.S. degree from the School of Engineering, Sun Yat-sen University. She is currently a joint Ph.D. candidate under the supervision of Prof. Kun Qian at the School of Biomedical Engineering, Shanghai Jiao Tong University and Prof. Yong Zhang at the Department of Biomedical Engineering, National University of Singapore. Her current research is focused on the development of novel nanodevice for in vitro diagnostics.

Yuning Wang received her Ph.D. in analytical chemistry from Fudan University. Currently, she is an assistant professor at the School of Biomedical Engineering, Shanghai Jiao Tong University. Her research interests focus on the development of novel signal amplification methods based on mass spectrometry for biomolecules detection.

Yong Zhang is a Provost's chair professor in the Department of Biomedical Engineering, National University of Singapore (NUS). His current research interests include nanomaterials, nanobiophotonics, nanomedicine, and microfluidic devices. Professor Zhang has authored over 300 peer-reviewed research articles in international journals such as Nature Medicine, Nature Biomedical Engineering and PNAS, and has received numerous awards such as IES Prestigious Engineering Achievement Award and Humboldt Research Award. He is a fellow of Royal Society of Chemistry (FRSC) and a highly cited researcher named by Clarivate Analytics.

Kun Qian received his B.S./M.S. from Fudan University, his Ph.D. from University of Queensland, and held the position of research fellow at Stanford University. He is now a principal investigator (PI) and distinguished professor at the School of Biomedical Engineering and School of Medicine, Shanghai Jiao Tong University. His research interests and efforts are on design, synthesis, engineering, and advanced biomedical applications of materials and devices. His group aims at the development and commercialization of bioanalytical methods and mass spectrometry techniques for diverse biomolecules in biomedical fields.