Natural Resources for Human Health
eISSN: 2583-1194

Original Research

**Received** 12 March 2026
**Revised** 18 April 2026
**Accepted** 10 May 2026

Natural Resources for Human Health 2026, Vol. 6 (3): 787–795

**KEYWORDS**:
plant metabolomics; high-resolution mass spectrometry; nuclear magnetic resonance; metabolite annotation; molecular networking; machine learning; multi-omics; spatial metabolomics

# Beyond Isolation: Modern Analytical and Computational Approaches to Unraveling Plant Metabolomes

Yangjingyu Li[1], V. Shanmugapriya[2], Hirvita Bhatt[3], Abhinav Rathour[4], Surayyo Khaydarova[5], Hari Hara Subramanyan P.V.[6], Roshini B[7]

[1] Department of Biocomputing and Engineering, Lincoln University College, Petaling Jaya, Selangor, Malaysia, lisagz@lincoln.edu.my
[2] Department of Biochemistry, Vinayaka Missions Medical College, Vinayaka Mission's Research Foundation (DU), Karaikal, Tamil Nadu, India, priyasri@live.com
[3] Department of Pharmacognosy, Parul Institute of Pharmacy, Parul University, Vadodara, Gujarat, India, hirvita.bhatt40020@paruluniversity.ac.in
[4] Centre for Research Impact & Outcome, Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura, Punjab 140401, India, abhinav.rathour.orp@chitkara.edu.in
[5] Jizzakh State Pedagogical University, Uzbekistan, surayyo.uz@mail.ru
[6] Meenakshi College of Physiotherapy, Meenakshi Academy of Higher Education and Research, hari@maher.ac.in
[7] Meenakshi College of Allied Health Sciences, Meenakshi Medical College Hospital & Research Institute, Meenakshi Academy of Higher Education and Research, roshiniahs@maher.ac.in

**ABSTRACT:** Plant metabolomes consist of a complex and dynamic mixture of small molecules, which are chemically diverse and spatially organized, and that link the genotype, environment, development, and phenotype. While conventional phytochemical isolation is important for structural confirmation, its compound-by-compound workflow doesn't comprehensively capture this complexity. This review discusses the analytical and computational shift from targeted isolation to integrated plant metabolomics. The robust profiling of primary metabolites by gas chromatography–mass spectrometry is complemented by the ability of liquid chromatography–high-resolution mass spectrometry to profile lipids and specialized metabolites. Nuclear magnetic resonance spectroscopy provides quantitative reproducibility and structural information, and capillary electrophoresis, ion mobility, and mass-spectrometry imaging provide chemical variation information for ions, isomers, and spatial information. Reliable discovery requires good experimental design, rapid quenching, representative extraction, pooled quality-control samples, internal standards, and transparent metadata. Computational workflows allow raw signals to be transformed into aligned features, correct analytical drift, prioritize molecular formulae, match spectral library, simulate fragmentation, construct molecular network, and integrate statistical learning and pathway and multi-omics evidence.

There are many challenges that persist, such as limited metabolome coverage, ion suppression, batch effects, unrecognized features, ambiguous feature annotation, poor cross-study comparability, and lack of reference standards. Curated training data,

Natural Resources for Human Health
eISSN: 2583-1194

reporting of uncertainties, chemical validation, and the FAIR deposition of data are needed for emerging spatial and single-cell methods, field-deployable instrumentation, machine learning, foundation models and automated laboratories to facilitate discovery. Complementary platforms, coupled with evidence-aware computation and biological-specific confirmation, can enable the translation of multidimensional signals to reproducible metabolite identities, biomarkers, pathways and causal hypotheses. This integrated approach enables crop improvement, stress biology, food quality, chemotaxonomy, nutrition and natural product discovery.

## 1. INTRODUCTION

The plant metabolome refers to all primary metabolites involved in plant growth and all specialized metabolites involved in defense, signalling, reproduction, environmental adaptation, nutrition and ecological interactions. (2) It is quite complex, mirroring vast chemical diversity, wide concentration changes, tissue and organelle compartmentation, developmental timing, genotype, microbiota and rapidly changing environmental conditions. (2) Conventional phytochemistry involves bioactivity-guided fractionation, repeated chromatography and spectroscopic structure elucidation, which are still essential for obtaining pure compounds and confirmation of their structures but are preferred for abundant, stable, extractable molecules and lack of systems-level context. (3) Modern metabolomics, on the other hand, simultaneously processes hundreds, even thousands, of signals and links the patterns of these chemicals to experimental conditions, to phenotypes, to pathways, and to molecular mechanisms. There is no single analytical platform that can cover the entire metabolome, and integrated workflows involve the use of gas or liquid chromatography–mass spectrometry, nuclear magnetic resonance, capillary electrophoresis, ion mobility and imaging methods that are all combined with computational preprocessing and annotation. The iterative workflow that is the subject of this review is described in Figure 1: A defined biological question leads to random sampling, rapid quenching, extraction, blank samples, pooled quality controls, and internal standards are applied to complementary platforms to generate multidimensional data, and algorithms were used to detect and align features, make suggestions of identities, organize molecular families and infer pathways, which are then subject to targeted assays and biological experiments for validation of the result. The diverse application of plant metabolomics includes identification of stress-response pathways, identification of climate-resilient plants, improvement of flavor and nutritional content, authentication of botanical materials, chemotaxonomic relationships and prioritization of bioactive natural products. (6) It is most useful when chemical observations are combined with genomic variability, gene expression, proteins, enzymes, spatial context and measurable characteristics of the plant or plant part, and not just listed as chemicals. This review, therefore, assesses the latest analytical platforms, computational processing, metabolite annotation, structural elucidation, statistical interpretation and multi-omics integration. It also covers issues of limited coverage, uncertainty in annotation, reproducibility, reference standards, reporting and FAIR data practices. The central idea is that isolation is no longer the first step in all discovery, but rather a targeted validation step in a context-preserving, uncertainty-quantified, and testable biological knowledge-producing pipeline. (7)

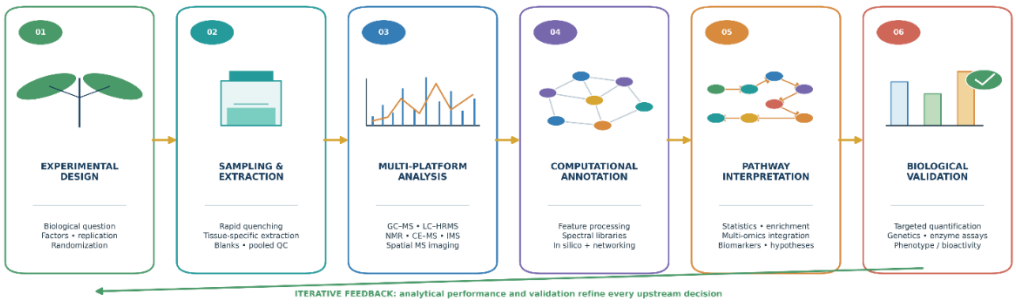


**Figure 1.** Integrated plant-metabolomics workflow from experimental design, sampling, and extraction to multi-platform analysis, computational annotation, pathway interpretation, and biological validation.

## 2. MODERN ANALYTICAL PLATFORMS FOR PLANT-METABOLOME PROFILING

Plant metabolome profiling is reliable; not only is it started before acquiring the instrument, but it is also carried out at the level of the analytical method. The samples must be representative of the biological material to be analysed, collected at specific developmental and circadian time points, randomized, quenched immediately, stored in a consistent manner and extracted with solvents appropriate to the chemical space to be analysed. Process blanks, extraction replicates, pooled QC samples, internal standards, dilution series and randomized injection order indicate contamination, drift, carryover, saturation and unstable features. Gas chromatography–mass spectrometry offers reproducible retention behavior and has a large number of electron-ionization libraries for volatile compounds and derivatized sugars, amino acids, organic acids and other primary metabolites, but it is limited by the requirement for derivatization that can generate artifacts and cannot be used for thermolabile or high-mass molecules. Liquid chromatography coupled to high-resolution tandem mass spectrometry provides a good coverage of semi-polar specialized metabolites, lipids, glycosides, alkaloids, phenolics and pigments with accurate mass and fragmentation data, but quantification and annotation are complicated by ion suppression, adduct formation, source-dependent spectra and retention-time variation. Nuclear magnetic resonance spectroscopy, which is inherently quantitative, intrinsically reproducible, intrinsically structural and non-destructive, is less sensitive than other methods, and will detect low concentrations of metabolites in complex extracts. (11) Capillary electrophoresis–mass spectrometry is well suited for polar and ionic metabolites, and ion-mobility spectrometry can provide collision-cross-section and gas-phase separation, which can aid in isomer separation and the suppression of spectral congestion. However, while bulk extraction of metabolites can provide the chemical gradient, mass spectrometry imaging can directly map metabolites across tissues, uncovering chemical gradients and cellular neighborhoods that bulk extraction does not capture, despite the challenges of matrix effect, spatial resolution and confident metabolite identification (12). A comparison of metabolite coverage, sensitivity, resolution, strengths and limitations of these platforms is shown in Table 1. An authentic standard, with validated calibration, is used to quantify defined analytes, whereas widely targeted workflow optimizes discovery, and untargeted workflow optimizes monitoring of large scheduled panels on a more consistent basis, but with a narrower novelty. The orthogonal platforms can be combined to enhance the coverage and confidence if they are harmonized using common samples, data metadata, data quality, and data retention or data spectral evidence. The question being addressed in the experiment should be the one that came from the biological question, and a deeper analysis will not make up for sampling not being confounded, not being replicated, not reproducible or not having quality controls. (14)

**Table 1.** Comparative characteristics, metabolite coverage, sensitivity, resolution, advantages, and limitations of modern plant-metabolomics platforms.

| Platform | Principal coverage | Sensitivity / resolution | Major advantages | Main limitations |
|---|---|---|---|---|
| GC–MS | Volatiles; derivatized sugars, amino and organic acids | High sensitivity; reproducible EI spectra | Mature libraries; robust retention indices; strong primary-metabolite profiling | Derivatization; thermolabile/high-mass compounds poorly represented |
| LC–HRMS/MS | Semi-polar metabolites, lipids, alkaloids, phenolics, glycosides | Very high mass resolution and sensitivity | Broad coverage; accurate mass; rich MS/MS; flexible chromatography | Ion suppression, adducts, source dependence, incomplete libraries |
| NMR | Abundant soluble metabolites and structural motifs | Lower sensitivity; high spectral reproducibility | Non-destructive; quantitative; strong structural evidence | Overlap in mixtures; costly instrumentation; low-abundance features missed |

| Platform | Principal coverage | Sensitivity / resolution | Major advantages | Main limitations |
|---|---|---|---|---|
| CE–MS | Polar, ionic and charged metabolites | High separation efficiency; small sample volume | Excellent for organic acids, nucleotides and charged species | Migration-time variability; limited robustness and interfaces |
| Ion mobility–MS | Isomers and congested MS features | Adds millisecond separation and collision cross section | Orthogonal structural descriptor; improved peak capacity | CCS libraries and method harmonization still developing |
| MS imaging | Spatial metabolite distributions in tissues | Micrometre-scale spatial maps; method dependent | Preserves anatomical context; label-free localization | Matrix effects, identification confidence, quantification and data volume |

## 3. COMPUTATIONAL PROCESSING, METABOLITE ANNOTATION, AND STRUCTURAL ELUCIDATION

The chromatograms, spectra, images, and metadata are converted to interpretable chemical evidence using computational analysis. Algorithms can include centroiding, peak detection, deconvolution, grouping of isotopes and adducts, retention alignment, gap filling, and construction of feature tables, among others, and are usually preceded by the transformation of vendor files to open formats. The approaches implemented in XCMS, MZmine, and MS-DIAL are complementary, and not just setting default parameters will suffice for parameter optimization. Unsuitable features are removed from the data before normalization, transformation, scaling, batch correction, and careful missing value treatment in (15). Quality assessment tests blank ratios, pooled-QC precision, signal drift, injection order, peak shape, carryover, saturation, and dilution behavior. The basic information for Annotation is exact mass, isotope pattern, retention behavior, collision cross section, and tandem spectra. While strong hypotheses can be made based on a spectral-library match, definitive identification generally requires an authentic spectral standard which has been analyzed under similar conditions; reporting should therefore provide confidence levels, not simply report that every feature has been identified as a metabolite. (17) When reference spectra are not available, the candidate structures or chemical classes are ranked by the formula prediction and in silico fragmentation systems (SIRIUS, CSI: FingerID, CANOPUS, MetFrag). Molecular networking groups fragmentation spectra into families to enable propagation of annotations and the detection of analogues, while feature-based networking maintains quantitative and chromatographic context. Recurrent fragments or neutral losses are discovered, across unknown molecular families, by (19) substructure discovery and (20) topic-model approaches. When using machine learning, both curated training sets and balanced evaluation are necessary to classify samples, predict retention or fragmentation, prioritize structures, and recognize chemical classes, and applicability domains, and protection against leakage are important. This evidence-based computational intelligence framework, presented here in Figure 2, spans from raw data and metadata through to quality-controlled features, identities, molecular families, pathways, biomarkers, and biological hypotheses. Table 2 tabulates some of the major tools and databases used for pre-processing, statistical analysis, annotation, networking, interpretation of pathways, and sharing. Reproducible workflows should maintain software versions, parameters, code, provenance, raw data, processed matrices, and metadata in community repositories. In addition to human inspection, the FAIR principles enhance findability, accessibility, interoperability and reuse of the artifacts, with chemical isolation and orthogonal confirmation of the computational predictions to be considered at the discretion of the humans who must evaluate artifacts and audit their model behavior. (20)

Natural Resources for Human Health
eISSN: 2583-1194

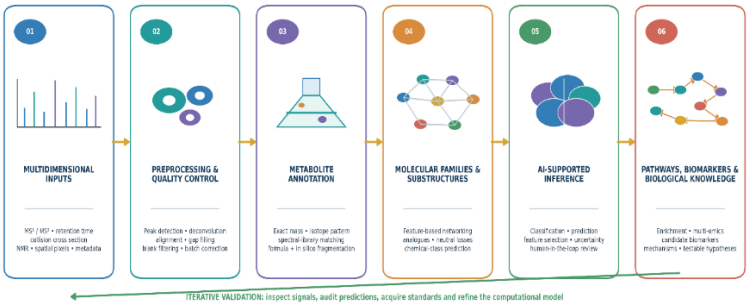


**Figure 2.** Computational intelligence framework for transforming multidimensional plant-metabolomics data into metabolite identities, molecular families, pathway networks, biomarkers, and biological knowledge.

**Table 2.** Computational tools and databases for preprocessing, statistical analysis, spectral annotation, molecular networking, pathway analysis, and data sharing.

| Tool/resource | Primary role | Core input or method | Representative output | Key consideration |
|---|---|---|---|---|
| XCMS | LC/GC–MS preprocessing | Peak detection, alignment, grouping | Feature table and corrected retention times | Parameters require dataset-specific optimization |
| MZmine | Modular MS data processing | Feature finding, deconvolution, annotation | Curated feature lists and exports | Manual inspection remains valuable |
| MS-DIAL | DIA/DDA processing and libraries | Deconvolution, alignment, spectral matching | Features, spectra and annotations | Library and acquisition compatibility matter |
| GNPS / FBMN | Spectral libraries and molecular networking | MS/MS similarity with feature metadata | Molecular families and propagated hypotheses | Network links indicate similarity, not identity |
| SIRIUS / CSI:FingerID / CANOPUS | Formula, structure and class prediction | Isotope patterns and fragmentation trees | Ranked formulae, candidates and classes | Candidate space and model domain constrain results |
| MetFrag | In silico fragmentation | Candidate structures and MS/MS peaks | Ranked structural candidates | Ranking is hypothesis generation |
| MetaboAnalyst | Statistics and pathway analysis | Feature or concentration matrices | Models, enrichment and visual summaries | Design, scaling and validation choices affect inference |

| Tool/resource | Primary role | Core input or method | Representative output | Key consideration |
|---|---|---|---|---|
| MetaboLights / FAIR repositories | Data and metadata sharing | Raw/processed data and ISA metadata | Reusable studies with persistent identifiers | Complete metadata and licensing enable reuse |

## 4. BIOLOGICAL INTERPRETATION AND MULTI-OMICS INTEGRATION

Biological interpretation starts with a pre-specified analysis plan which aims to distinguish exploratory discovery from confirmatory testing. The effects of the metabolites of interest and uncertainty are estimated in univariate models, adjusted for multiple tests, and adjusted for covariates; multivariate models (principal-component analysis, multivariate classification, partial least-squares) show global structure and potential outliers. Small, high dimensional data sets are prone to overfitting, which is why supervised classifiers need to be accompanied by nested cross validation, external test sets where possible, permutation testing and transparent feature selection. To measure the suitability of candidates as biomarkers, they should be stable within batches, seasons, genotypes, environments, and independent populations, and measured using targeted methods. Coordinate metabolite changes can be related to coordinated biosynthesis, degradation, transport, redox balance and signaling processes and used to draw conclusions about pathway enrichment; however, incomplete plant-specific databases, existing ambiguity in the assignment of isomers, common intermediates, and pathway topology, can lead to distorted conclusions. Network-based methods use the concept of similarity of spectra, biochemical reactions, correlations, and prior knowledge to cluster all known and unknown features. Genomics enables integration of metabolite quantitative-trait loci and biosynthetic gene clusters; genomics, transcriptomics, proteomics enable identification of candidate enzymes and regulatory modules; phenomics enables linking chemistry to growth, resilience, yield, color, flavor or disease resistance. Correlation is not enough; test for causation with genetics, isotope tracing, enzyme assays, perturbation experiments or targeted complementation. Spatial, temporal and single-cell metabolomics provides new information that is not provided by bulk averaging, enabling probing of tissue-specific biosynthesis, transport, storage and responses at infection or developmental interfaces. Major applications and their translation value are summarized in Table 3. In stress biology, integrated profiles allow for the identification of early response signatures and adaptive pathways, in breeding for predicting traits, and in chemotaxonomy and food science to authenticate species, origin, processing and quality. Natural-product discovery is based on activity data, molecular networks, and taxonomical information to prioritize novel scaffolds and to prevent re-discovery while nutrition studies correlate the crop chemistry and exposure and bioavailability. (26) Strongest studies provide triangulation of complementary analytical platforms, computational studies, and biological experiments, uncertainty, and validation of priority metabolites in new material. This integration makes metabolomics a mechanism bridge between molecular variation and plant phenotype from a descriptive fingerprinting to a mechanistic one. (27)

**Table 3.** Major applications of integrated plant metabolomics, analytical strategies, computational methods, biological outputs, and translational value.

| Application | Analytical strategy | Computational / multi-omics method | Biological output | Translational value |
|---|---|---|---|---|
| Stress biology | Time-course LC–MS, GC–MS and imaging | Network analysis with transcriptomics/proteomics | Early signatures and adaptive pathways | Mechanistic targets for resilience |
| Crop improvement | Population-scale targeted and untargeted profiling | mQTL, GWAS and genomic prediction | Metabolite–trait associations | Selection for yield, quality and climate adaptation |

| Application | Analytical strategy | Computational / multi-omics method | Biological output | Translational value |
|---|---|---|---|---|
| Chemotaxonomy/authentication | Reproducible fingerprints across accessions | Classification, feature selection, spectral libraries | Species, origin and adulteration markers | Botanical quality control and traceability |
| Nutrition and food quality | Quantitative multi-platform profiling | Pathways, processing models and exposure integration | Nutrients, antinutrients, flavor and stability markers | Healthier crops and optimized processing |
| Natural-product discovery | Bioactivity-linked LC–MS/MS and isolation | Molecular networking, in silico annotation, taxonomic priors | Prioritized scaffolds and analogues | Faster dereplication and lead discovery |
| Spatial/single-cell biology | MS imaging, microsampling and cell-resolved assays | Image segmentation and spatial multi-omics | Cell-type and tissue chemical maps | Mechanistic localization of biosynthesis and transport |

## 5. CHALLENGES, FUTURE PERSPECTIVES, AND CONCLUSION

The extraction and ionization coverage, extreme dynamic range, matrix effects, compound instability, isomerism and spatial heterogeneity are still limiting plant metabolomics. Most features detected are not identified or only putatively annotated as authentic standards and curated spectra represent a small portion of plant chemical diversity. To enable cross-study comparison, it is essential to have standardized metadata, reference materials, pooled quality controls, harmonized confidence terminology, interoperable formats, and public deposition. Portable and field-deployable instruments can record responses to the environment, quickly; imaging and single-cell methods can identify metabolite exchange within tissues; both require careful calibration and validation. The prediction of spectra, retention, structures, pathways, and bioactivity, and learning of reusable chemical representations across repositories could be possible with the help of artificial intelligence and foundation models. Generative predictions need to be accompanied by provenance and uncertainty and should be used in conjunction with authentic standards, isolation, synthesis and biological testing. Autonomous laboratories and digital twins might automatically choose experiments and optimize data acquisition, but require machine readable protocols, variability in training data, and human supervision. Priorities are to grow and link a library for plants, handle unknown features across studies, build uncertainty-aware models, incorporate spatial and isotope-resolved information, and promote reusability of datasets. Finally, modern plant metabolomics transcends isolation, with complementary analytical platforms and transparent computation, alongside multi-omics context and targeted validation. This strategy focuses classical isolation on informative molecules, and places validated structures into pathways, tissues, phenotypes, and applications. Evidence-aware integration has the potential to translate intricate chemical signals into repeatable knowledge that can be applied to resilient crops, nutritious foods, ecological understanding, and innovation of natural products. (35)

**Declarations**

Acknowledgment: The authors gratefully acknowledge the support and guidance received during the preparation of this manuscript.

**Conflict of Interest:** The authors declare that they have no conflict of interest.

**Funding:** This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

**REFERENCES**

[1] Fiehn O. Metabolomics—the link between genotypes and phenotypes. Plant Mol Biol. 2002;48:155-171.
[2] Fernie AR, Trethewey RN, Krotzky AJ, Willmitzer L. Metabolite profiling: from diagnostics to systems biology. Nat Rev Mol Cell Biol. 2004;5:763-769.
[3] Sumner LW, Mendes P, Dixon RA. Plant metabolomics: large-scale phytochemistry in the functional genomics era. Phytochemistry. 2003;62:817-836.
[4] Saito K, Matsuda F. Metabolomics for functional genomics, systems biology, and biotechnology. Annu Rev Plant Biol. 2010;61:463-489.
[5] Jorge TF, Rodrigues JA, Caldana C, et al. Mass spectrometry-based plant metabolomics: metabolite responses to abiotic stress. Mass Spectrom Rev. 2016;35:620-649.
[6] Hall RD. Plant metabolomics: from holistic hope, to hype, to hot topic. New Phytol. 2006;169:453-468.
[7] Alseekh S, Aharoni A, Brotman Y, et al. Mass spectrometry-based metabolomics: a guide for annotation, quantification and best reporting practices. Nat Methods. 2021;18:747-756.
[8] Dunn WB, Broadhurst D, Begley P, et al. Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. Nat Protoc. 2011;6:1060-1083.
[9] Lisec J, Schauer N, Kopka J, Willmitzer L, Fernie AR. Gas chromatography mass spectrometry-based metabolite profiling in plants. Nat Protoc. 2006;1:387-396.
[10] De Vos RCH, Moco S, Lommen A, Keurentjes JJB, Bino RJ, Hall RD. Untargeted large-scale plant metabolomics using liquid chromatography coupled to mass spectrometry. Nat Protoc. 2007;2:778-791.
[11] Emwas AH, Roy R, McKay RT, et al. NMR spectroscopy for metabolomics research. Metabolites. 2019;9:123.
[12] Ramautar R, Somsen GW, de Jong GJ. CE-MS for metabolomics: developments and applications in the period 2010-2012. Electrophoresis. 2013;34:86-98.
[13] Boughton BA, Thinagaran D, Sarabia D, Bacic A, Roessner U. Mass spectrometry imaging for plant biology: a review. Phytochem Rev. 2016;15:445-488.
[14] Ernst M, Silva DB, Silva RR, Vêncio RZN, Lopes NP. Mass spectrometry in plant metabolomics strategies: from analytical platforms to data acquisition and processing. Nat Prod Rep. 2014;31:784-806.
[15] Smith CA, Want EJ, O'Maille G, Abagyan R, Siuzdak G. XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. Anal Chem. 2006;78:779-787.
[16] Pluskal T, Castillo S, Villar-Briones A, Orešič M. MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. BMC Bioinformatics. 2010;11:395.
[17] Sumner LW, Amberg A, Barrett D, et al. Proposed minimum reporting standards for chemical analysis. Metabolomics. 2007;3:211-221.
[18] Dührkop K, Fleischauer M, Ludwig M, et al. SIRIUS 4: a rapid tool for turning tandem mass spectra into metabolite structure information. Nat Methods. 2019;16:299-302.
[19] Nothias LF, Petras D, Schmid R, et al. Feature-based molecular networking in the GNPS analysis environment. Nat Methods. 2020;17:905-908.
[20] Wilkinson MD, Dumontier M, Aalbersberg IJ, et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data. 2016;3:160018.
[21] Xia J, Psychogios N, Young N, Wishart DS. MetaboAnalyst: a web server for metabolomic data analysis and interpretation. Nucleic Acids Res. 2009;37:W652-W660.
[22] Broadhurst D, Goodacre R, Reinke SN, et al. Guidelines and considerations for the use of system suitability and quality control samples in mass spectrometry assays applied in untargeted clinical metabolomic studies. Metabolomics. 2018;14:72.
[23] Fukushima A, Kusano M. Recent progress in the development of metabolome databases for plant systems biology. Front Plant Sci. 2013;4:73.
[24] Alexandrov T. Spatial metabolomics and imaging mass spectrometry in the age of artificial intelligence. Annu Rev Biomed Data Sci. 2020;3:61-87.
[25] Kusano M, Fukushima A, Redestig H, Saito K. Metabolomic approaches toward understanding nitrogen metabolism in plants. J Exp Bot. 2011;62:1439-1453.

[26] Wolfender JL, Marti G, Thomas A, Bertrand S. Current approaches and challenges for the metabolite profiling of complex natural extracts. J Chromatogr A. 2015;1382:136-164.
[27] Perez de Souza L, Alseekh S, Scossa F, Fernie AR. Ultra-high-performance liquid chromatography high-resolution mass spectrometry variants for metabolomics research. Nat Methods. 2021;18:733-746.
[28] Blaženović I, Kind T, Ji J, Fiehn O. Software tools and approaches for compound identification of LC-MS/MS data in metabolomics. Metabolites. 2018;8:31.
[29] Dührkop K, Nothias LF, Fleischauer M, et al. Systematic classification of unknown metabolites using high-resolution fragmentation mass spectra. Nat Biotechnol. 2021;39:462-471.
[30] Haug K, Cochrane K, Nainala VC, et al. MetaboLights: a resource evolving in response to the needs of its scientific community. Nucleic Acids Res. 2020;48:D440-D444.
[31] Rappez L, Stadler M, Triana S, et al. SpaceM reveals metabolic states of single cells. Nat Methods. 2021;18:799-805.
[32] Domingo-Almenara X, Guijas C, Billings E, et al. The METLIN small molecule dataset for machine learning-based retention time prediction. Nat Commun. 2019;10:5811.
[33] van der Hooft JJJ, Wandy J, Barrett MP, Burgess KEV, Rogers S. Topic modeling for untargeted substructure exploration in metabolomics. Proc Natl Acad Sci USA. 2016;113:13738-13743.
[34] Wang M, Carver JJ, Phelan VV, et al. Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. Nat Biotechnol. 2016;34:828-837.
[35] Misra BB. New tools and resources in metabolomics: 2016-2017. Electrophoresis. 2018;39:909-923.