Review

Guide to Metabolomics Analysis: A Bioinformatics Workflow

Yang Chen [1,2], **En-Min Li** [1,2,*] **and Li-Yan Xu** [1,3,*]

1   The Key Laboratory of Molecular Biology for High Cancer Incidence Coastal Chaoshan Area, Shantou University Medical College, Shantou 515041, China; cy_koasde@163.com

2   Department of Biochemistry and Molecular Biology, Shantou University Medical College, Shantou 515041, China

3   Guangdong Provincial Key Laboratory of Infectious Diseases and Molecular Immunopathology, Institute of Oncologic Pathology, Shantou University Medical College, Shantou 515041, China

*   Correspondence: nmli@stu.edu.cn (E.-M.L.); lyxu@stu.edu.cn (L.-Y.X.); Tel.: +86-754-88900413 (E.-M.L.); Fax: +86-754-88900847 (E.-M.L.)

Abstract: Metabolomics is an emerging field that quantifies numerous metabolites systematically. The key purpose of metabolomics is to identify the metabolites corresponding to each biological phenotype, and then provide an analysis of the mechanisms involved. Although metabolomics is important to understand the involved biological phenomena, the approach's ability to obtain an exhaustive description of the processes is limited. Thus, an analysis-integrated metabolomics, transcriptomics, proteomics, and other omics approach is recommended. Such integration of different omics data requires specialized statistical and bioinformatics software. This review focuses on the steps involved in metabolomics research and summarizes several main tools for metabolomics analyses. We also outline the most abnormal metabolic pathways in several cancers and diseases, and discuss the importance of multi-omics integration algorithms. Overall, our goal is to summarize the current metabolomics analysis workflow and its main analysis software to provide useful insights for researchers to establish a preferable pipeline of metabolomics or multi-omics analysis.

Keywords: metabolomics; metabolomics analysis tools; metabolic pathways summary; multi-omics integration algorithms

1. Introduction

Metabolomics is a rapidly evolving field that deals with the high-throughput characterization of metabolites, and is the study of the metabolite composition of cell types, tissues, organs, or organisms [1,2]. Metabolomics is the collection of endogenous small molecules that mark specific fingerprints of cellular biochemistry [3]. It measures numerous low-molecular weight metabolites, such as amino acids, sugars, fatty acids, lipids, and steroids [4]. Small modifications in the chemical structure and some external stimuli (e.g., infections and allergens) can dramatically change the function of a metabolite [5–7]. Metabolites, in addition to being produced directly by the host organism, can be derived by host microbiota or transformed from dietary, xenobiotic or other exogenous sources [8]. It is worth emphasizing that lipids are important metabolites in the organism, and have a wide variety of properties, such as insolubility in water and solubility in non-polar organic solvents [9]. Lipids are involved in the regulation of many physiological reactions. An important step in energy metabolism is the hydrolysis of triglycerides (TG) in lipid droplets to release fatty acids that can be used or stored [10]. Abnormal lipid metabolism can lead to many diseases, such as obesity, atherosclerosis and diabetes [11–13]. The concept of "lipidomics" was first introduced by Richard et al. in 2003, and was followed by the introduction of more efficient and accurate lipidomics research [14]. In 2005, Markus Wenk also suggested that more and more studies show a direct link between lipids and many diseases, and that metabolomics and proteomics studies are no longer sufficient to provide

a clear picture of the causes of these diseases [15]. Therefore, it has become important to link lipid studies to diseases, with lipidomics no longer being just a single study of lipids, but rather a comprehensive discipline linking the proteome, metabolome, and various diseases. Lipidomics has been used to elucidate the metabolism of lipids by studying their composition, structure, and quantification in biological samples, and can be used to search for biomarkers and to study the mechanisms of lipids in various phenomena [10,16]. With the continuous innovation of mass spectrometry technology and the increasing awareness of the importance of the biological functions of lipids, lipidomics, as an important branch of metabolomics, has gradually attracted the attention of researchers [17].

Metabolomics is wildly used in cancers and metabolism-related diseases. For example, researchers found that, in bladder cancer, the metabolites in the tricarboxylic acid (TCA) cycle were significantly changed [18–22], along with changes in fatty acid metabolism [19–21,23]. In colorectal cancer, disordered methionine metabolism and abnormal TCA cycle function have been reported [24]. Amino acid metabolism [25–30], bile acid metabolism [25,31], choline metabolism [25–28], fatty acid metabolism [27–30], and glycolysis [32,33] have also been found to be abnormal in liver cancer through metabolomics analysis. In diabetes, many metabolic pathways have been found to be disordered, such as acetoacetate metabolism [34], acylcarnitine metabolism [35–37], palmitic acid metabolism [38–40], linolenic acid metabolism [38], cholesterol metabolism [41], carbohydrate metabolism [34,42–44], glycine and serine metabolism [45,46], and fatty acid metabolism [40,44,47–49]. Glycolysis [50,51], the TCA cycle [47,52–54], the urea cycle [55], and glutathione metabolism [56] have also been found to be abnormal in obesity. In Alzheimer's disease, abnormal amino acid metabolism [57–61], fatty acid metabolism [57], linoleic acid metabolism [62], cholesterol metabolism [57], glycine and serine metabolism [63], aspartate metabolism [64], glycerophospholipid metabolism [65] and polyamine metabolism [57] have been reported. In this context, metabolomics is useful in the identification of biomarkers associated with the diagnosis/prognosis of different oncological processes and the response to treatment [66,67].

Previous reviews have summarized metabolomic analysis platforms [68], and some have summarized metabolomic data preprocessing methods [69]. However, a detailed guide for the process of metabolomics analysis would be an essential addition to these previous works. This review aims to describe the overall metabolomics analysis process and summarize the currently available software and databases for analyzing metabolomics data, thus providing a standard protocol for analyzing metabolomics data to identify clinically or disease-relevant biomarkers for researchers.

2. The Analysis Workflow of Metabolomics

The specific characteristics of metabolomics data require the application of different bioinformatics tools following a specific workflow (Figure 1). The first step in the metabolomics workflow involves using different techniques to isolate and characterize different groups of metabolites. There are two main platforms of metabolomics analysis: mass spectrometry (MS) and nuclear magnetic resonance (NMR) spectroscopy. Each has its own advantages and disadvantages [70]. MS-based metabolomics is generally preceded by a separation step, which reduces the complexity of the biological sample and allows MS analysis of different sets of molecules at different time [71]. MS acquires spectral data in the form of a mass-to-charge ratio (m/z) and a relative intensity of the ionized compound [72]. The most common separation techniques in MS technology are liquid chromatography (LC) and gas chromatography (GC) columns (LC-MS and GC-MS, respectively) [73]. MS allows for the reliable identification of metabolites. Especially when MS is used in tandem with chromatographic separation methods, its resolving ability is improved. At the same time, MS has a short analysis cycle (with an analysis time ranging from 5 to 140 min) and allows for selective qualitative and quantitative analyses. Therefore, this technique is the most widely used. The main disadvantages of MS are the high cost of the instruments and the requirement for sample separation or purification prior to putting the sample into the

mass spectrometer. NMR is a spectroscopic technique based on the principle of energy absorption and re-emission by atomic nuclei due to variations in an external magnetic field [74]. The spectral data generated by NMR can be used to quantify the concentration and characterize the chemical structure of metabolites. The advantages of NMR are that it is a nondestructive and highly reproducible technique, and does not require extensive sample preparation [75]. In particular, NMR has the ability to provide a high degree of structural information in a short time. Nevertheless, NMR has a lower sensitivity, which means that lower concentrations of potentially important compounds can be masked by larger peaks, and thus cannot be identified [76]. The identification and quantification of metabolites is a complex task that cannot be easily automated. Therefore, careful data processing and statistical analysis are required to derive useful and reliable information from these profiles [77]. The application of NMR spectroscopy is not limited to liquid and solid samples but extends to intact tissue samples with high-resolution magic angle rotation (HRMAS) NMR spectroscopy [78,79]. LC-MS is more suitable for the detection of moderately polar compounds and substances with high polarity, in terms of specific types of substances, such as fatty acids, alcohols, phenols, vitamins, organic acids, polyamines, nucleotides, polyphenols, terpenes, flavonoids, lipids, and other compounds [73,80]. The inherent limitation of GC-MS is that it only detects volatile compounds or compounds that can be derivatized into volatiles. GC-MS can detect amino acids, organic acids, fatty acids, sugars, polyols, amines, sugar phosphates, and other substances [81,82]. It is worth noting that GC-MS measurement of water-soluble substances requires derivatization because GC-MS analysis can only be performed directly if the sample is volatile and stable to heat. Each separation method has its own resolution and sensitivity in the identification of metabolites, and its selection is based on the chemical and physical properties of each sample or the chemical and physical characteristics of the hypothetical target compounds, as well as the type of analysis to be performed (untargeted or targeted) [83,84].

The second step is the preprocessing of raw signals (chromatograms, spectra, or NMR data) by specific software for quantitative analysis of compounds. (e.g., XCMS [85], MAVEN [86] or MZmine3 [87]). In general, this step includes noise reduction, retention time correction, peak detection and integration, and chromatographic alignment. Several main platforms of the steps are discussed in Section 4.

In the third step, data processing is performed, and quality control (QC) is necessary. Data from QC samples are used to separate different-quality (high or low) data, balance the analytical platform's bias, and correct for noise in the signal. QC samples are used to determine the variance of metabolite features. If the variance of a feature is too high, it will be removed from the analysis [69,73]. Then, data normalization is used to reduce systematic bias or technical variation and to avoid misidentification due to disparate input of large amounts of metabolomics data. Subsequently, mass spectrometry peak data are used for compound identification by comparing it to authentic standard data (typically through an in-house library). In the absence of an in-house library, researchers can also apply public databases for compound identification. It is worth mentioning that researchers should be aware of the criteria for reporting metabolite annotation and identification. The Metabolomics Standards Initiative (MSI) was conceived in 2005 with the aim of enabling the effective application, sharing and reuse of data [88]. The standard proposes four different levels of metabolite identification observed in the scientific literature, including identified metabolites (level 1), presumptively annotated compounds (level 2), presumptively characterized compound classes (level 3), and unknown compounds (level 4). It is recommended that researchers define identification levels, common names, and structure codes (e.g., InChI or SMILES) in their publications and when submitting data to the repository [88]. For untargeted metabolomics studies, different databases, such as the Human Metabolome Database (HMDB [89]) or the Metabolite and Tandem MS Database (METLIN [90]), are used to identify metabolites from spectra.
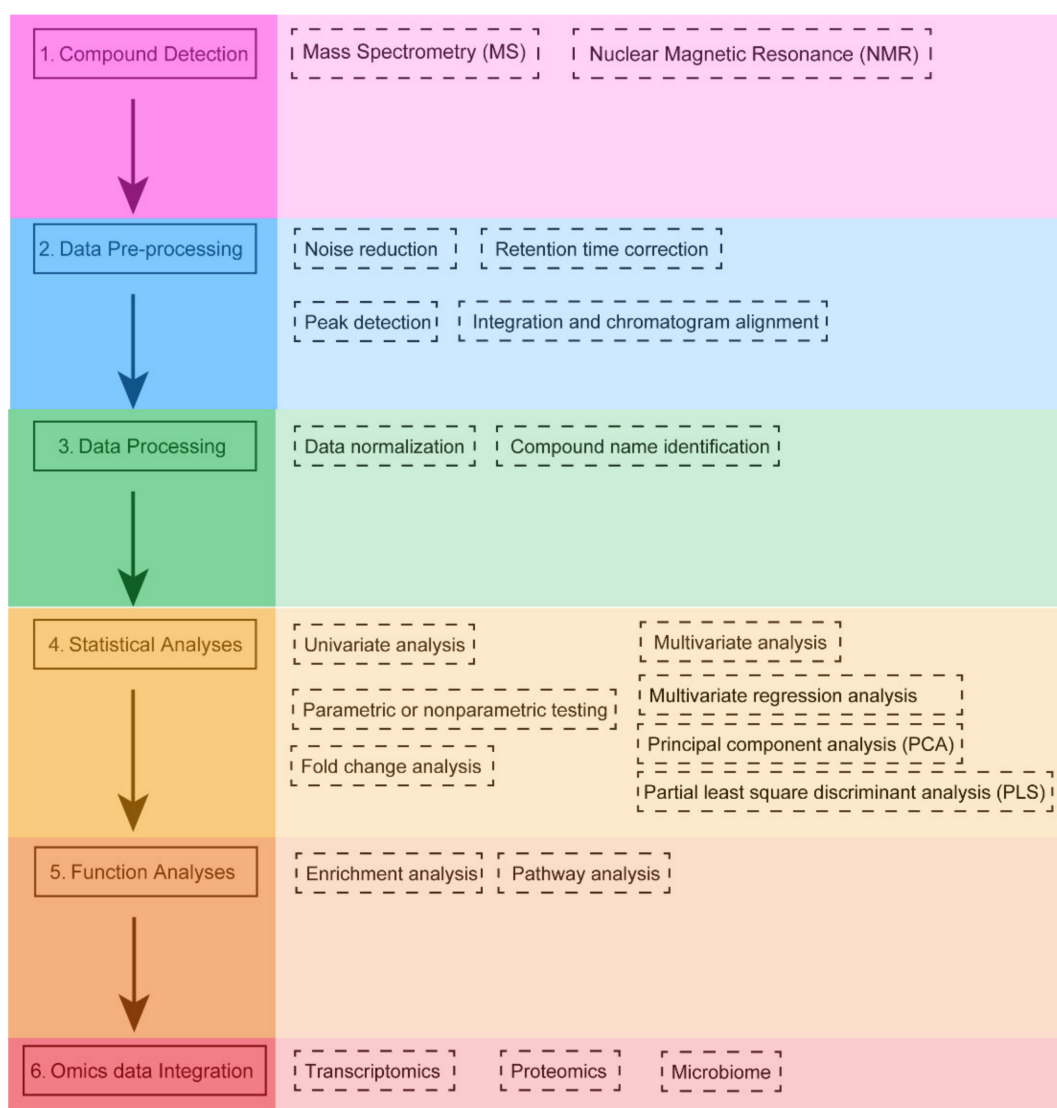
Figure 1. Typical workflow of metabolomics analysis. Metabolites are detected by using specific detection techniques (compound detection). Raw signals are then pre-processed to produce data in a suitable format for subsequent statistical analysis (data pre-processing). Then, data normalization is used to reduce the system and technical bias. For untargeted studies, metabolites are identified from spectral information in some given database (data processing). Univariate and multivariate statistical analyses are used to identify significantly expressed metabolites (statistical analyses). Next, the significantly expressed metabolites are subsequently linked to the biological context by using enrichment and pathway analysis (function analyses). Finally, metabolomics data may be integrated with other omics data (transcriptomics, proteomics, or the microbiome) to gain a comprehensive understanding of the molecular mechanisms of pathophysiological processes (Omics data Integration).

Typical statistical analyses of metabolomics data include univariate and multivariate approaches. They enable the evaluation of the input metabolomic dataset and identification of metabolites that undergo abnormal changes. Subsequently, functionally relevant metabolites can be distinguished by further data mining methods [91]. Traditional statistical methods determine the relationships between variables based solely on mathematical criteria, without fully taking into account biological correlations, which is one of their limitations [92]. So, the combined use of multiple statistical techniques is recommended when performing metabolomics analysis. In such a context, an appropriate *p*-value is used to rank the significantly expressed metabolites and determine a reliable threshold to select

the most significant one. The choice of this threshold may influence the final biological interpretation, and is therefore particularly critical [93].

In the next step, the screened metabolites are linked to their biological context by pathway and enrichment analysis. The aim of enrichment analysis is to explore the profile of functionally relevant metabolites to determine the link between changes in metabolite expression and biological context. This allows the use of a list of altered metabolites to suggest biological pathways or disease conditions that would indicate the subsequent steps in the study. The goal of pathway analysis is to identify pathways that have a significant impact on a specific biological process. Enrichment and pathway analyses are performed using specialized software tools [94] that map metabolites to known biochemical pathways based on information in public databases such as KEGG [95]. Subsequently, investigators typically use network visualization tools to present and understand their results.

Multi-omics data integration and analysis pipelines for studying the pathogenesis of disease and the influence of environmental risk factors are scarce. In this last step, an integrated multi-omics platform provides a reliable and understandable overview of metabolic changes [94]. The identified metabolites and metabolic pathways can be integrated with other omics data, which may help us to obtain more comprehensive information about the biological phenomena.

3. Statistical Analysis in Metabolomics

Depending on the experimental context, various types of data mining and statistical methods can be applied to metabolomics data. In the following overview, we summarize in detail the univariate and multivariate statistical analysis methods applied in metabolomics.

3.1. Univariate Analysis

Univariate analysis usually provides a preliminary overview of data characteristics that may be important in identifying the conditions under study. For two-group data (both unpaired and paired analyses), we can perform fold change analysis, *t*-tests, and volcano plots. For multi-group data, we can perform one-way analysis of variance (ANOVA), as well as related post hoc analysis and correlation analysis. Since each patient or biological sample usually has a large number of metabolites, and each metabolite needs to undergo a separate statistical test, a large number of false positive results can be obtained through multiple tests. To reduce this, multiple testing methods (e.g., Bonferroni, Bonferroni–Holm and Benjamini–Hochberg corrections) must be used to correct for *p*-values [96]. The Benjamin–Hochberg correction, also known as the false discovery rate (FDR), is one of the recommended methods because it allows for controlling the proportion of false positives in all significant results.

3.2. Multivariate Analysis

Since multi-omics data usually contain some characteristics that vary with phenotype or experimental conditions, the use of multivariate analyses that allow simultaneous observation and analysis of more than two statistical variables is recommended. Multivariate analysis includes multiple variance analysis (ANOVA), multiple regression analysis, factor analysis, principal component analysis (PCA), partial least squares discriminant analysis (PLS), cluster analysis, and machine learning (e.g., random forest and SVM). Because multivariate analysis uses the weighted averages to summarize the original variables in fewer variables, they are useful for exploratory data analysis. PCA analysis starts from the interrelationship between the original variables, linearly transforms them to several independent composite indicators (i.e., principal components) according to the principle of variance maximization, takes two to three principal components for graphing, visually describes the differences in metabolic patterns and clustering results between different groups, and searches for the original variables that contribute to intergroup classification as biomarkers through loadings plots. PCA is commonly used as a pre-analysis and quality control step for metabolomics data to observe whether there are intergroup classification

trends and data outlier points. PCA can also be used to analyze whether quality control samples are clustered together, or if they are scattered or have some variability, which would indicate problems with the quality of the assay. For example, Pasikanti et al. used PCA to analyze urine bladder cancer metabolomics data and observed that the QC samples were tightly clustered on the PCA score plot, thus validating the stability of the instrument's assay and the reliability of the metabolomics data [19]. PLS-DA is another commonly used classification method in metabolomics data analysis, which combines regression models with dimensionality reduction and discriminant analysis of regression results using certain discriminant thresholds. The difference between this dimensionality reduction method and PCA is that PLS-DA decomposes both the independent variable X matrix and the response variable Y matrix, and uses its covariance information in the decomposition so that the dimensionality reduction effect can extract the inter-group variation information more efficiently than PCA. In practice, a PLS-DA score plot is often used to visualize the classification effect of the model, and the greater the separation of the two groups in the plot, the more significant the classification effect [97]. It is important to note that cross-validation should be performed when using PLS-DA, and the PLS-DA results should be interpreted together with PCA to avoid overfitting problems in the metabolomic data. The goal of the metabolomic analysis is to screen potential biologically relevant markers to explore the metabolic mechanisms involved, and therefore requires variable screening with the help of certain feature screening methods. Random forest (RF) and SVM provide very flexible models for handling data with many covariates [97]. For example, RF is a nonparametric ensemble approach that prioritizes predictions by trying to find nonlinear patterns in metabolites that can explain the variation in each outcome [97]. RFs are very powerful tools if the relationship between metabolites and outcomes is complex and nonlinear, and have been used for missing data interpolation and outcome analysis in metabolomics [98,99]. A disadvantage of this approach is that as with PCA, it does not provide a measure of statistical significance or provide any *p*-values or equivalence measures. Nevertheless, RFs can provide analysts with a ranked list of the most important metabolites. All the mentioned machine learning methods have many of the same limitations as RF, as they can provide variable importance measures, but not a set of variables that can be considered statistically significant [97]. To evaluate the importance of each variable more objectively and comprehensively, a combination of the above methods is generally adopted in metabolomics studies for variable screening. A more common strategy is to perform univariate analysis, then combine the variable importance scores from multivariate models as screening criteria, and finally integrate them with the variables screened by machine learning models, such as selecting variables with FDR ≤ 0.05 and VIP > 1.5, and that are ranked high in the RF as potential biomarkers.

4. Software Tools for Metabolomics Data Analysis and Integration

Metabolomics analyses need powerful software tools to address the vast amount and variety of data. Excellent metabolomics software should include one or more of the following functions: (1) the ability to process of raw spectral data, (2) statistical analysis to find significantly expressed metabolites, (3) the ability to connect to metabolite databases for metabolite identification, (4) bioinformatics analysis and visualization of molecular interaction networks, and (5) the ability to integrate and analyze multi-omics data. We searched the PubMed literature database with the following keywords: "metabolomics software", "metabolomics & bioinformatics", "metabolomics analysis tool", "lipidomic software", "metabolomics protocol", "multi-omics software", and "multi-omics algorithm". Only original research articles regarding software and databases were included. In this section, we introduce several data analysis tools (Table 1) and show their characteristics and compare their advantages and disadvantages (Figures 2 and 3).

Table 1. Features of several most used metabolomics data analysis tools.

| Name	Year	Description	Functions							
			Data Pre-Processing	Data Processing		Statistical Analyses	Pathway Enrichment Analysis	Omics Data Integration		
				Normalization	Compound Name Identification			Transcriptomics	Proteomics	Microbiome
Mzmine3	2022	MZmine3 builds on the success of MZmine 2 with many features focused on improving the user-friendly graphical	Y	Y	Y	Y	-	-	-	-
MetaboAnalyst 5.0	2021	Comprehensive web-based tool for comprehensive metabolomics data analysis, interpretation, and integration with other omics data.	Y	Y	Y	Y	Y	Y	-	-
LipidSig	2021	Web-based tool for lipidomic data analysis	Y	Y	Y	Y	-	-	-	-
MS-DIAL 4.0	2020	Lipidome atlas in MS-DIAL 4.0	Y	Y	Y	Y	-	-	-	-
El-MAVEN	2019	Fast, Robust, and User-Friendly Mass Spectrometry Data Processing Engine for Metabolomics	Y	Y	Y	-	-	-	-	-
MetFlow	2019	Interactive and integrated web server for metabolomics data cleaning and differential metabolite discovery.	Y	Y	Y	Y	Y	-	-	-
LION	2019	Web-based ontology enrichment tool for lipidomic data analysis.	-	Y	Y	Y	Y	-	-	-
Omicsnet	2018	Web-based tool for creation and visual analysis of biological networks in 3D space	-	-	-	Y	Y	Y	Y	Y
METLIN	2018	Technology platform for the identification of known and unknown metabolites and other chemical entities.	-	-	Y	-	-	-	-	-
PaintOmics 3	2018	Web-based resource for the integrated visualization of multiple omics data types onto KEGG pathway diagrams.	-	-	-	-	Y	Y	Y	-
LipiDex	2018	Integrated Software Package for High-Confidence Lipid Identification	Y	-	Y	-	-	-	-	-
LipidMatch	2017	Automated workflow for rule-based lipid identification using untargeted high-resolution tandem mass spectrometry data	Y	-	Y	-	-	-	-	-
3Omics	2013	One-click web tool for fast analysis and visualization of multi-omics data.	Y	Y	-	Y	Y	Y	Y	-
IMPaLa	2011	Pathway analysis of transcriptomics or proteomics and metabolomics data.	-	-	-	-	Y	Y	Y	-
MetPA	2010	Pathway analysis for metabolomics data.	Y	-	-	-	Y	-	-	-
MassTRIX	2008	Tool for high precision MS data annotation.	Y	-	Y	-	Y	-	-	-
MetaCore™	2004	Commercial tool for functional analysis and integrated analysis of multi-omics data.	Y	-	-	-	Y	Y	Y	-
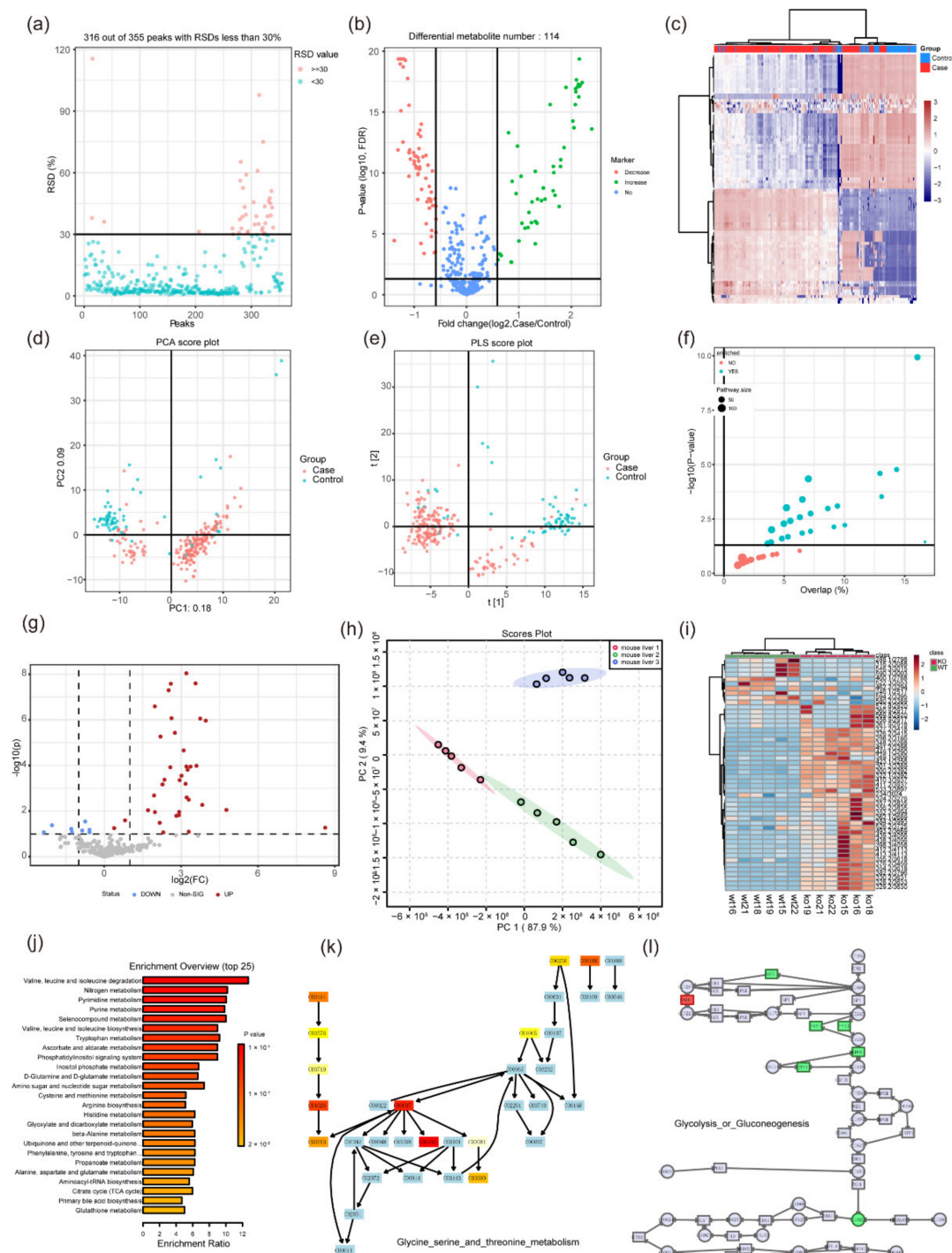
Figure 2. Some graphical visualization features of MetFlow and MetaboAnalyst 5.0. (**a**) RSD (relative standard deviation) plot in the data processing function of MetFlow. Features with a high percent RSD should be removed from the subsequent analysis (the suggested threshold is 20% for LC-MS and 30% for GC-MS). (**b**) Volcano plot and (**c**) heatmap of the differential metabolites in the statistical analysis function of MetFlow, the thresholds can be set autonomously by the submitter. (**d**) PCA analysis and (**e**) PLS analysis in MetFlow. (**f**) Pathway enrichment overview in MetFlow, each circle represents a different pathway. Circle size and color are based on the pathway size and *p*-value. (**g**) Volcano plot of the differential analysis in MetaboAnalyst 5.0. (**h**) PCA analysis plot in MetaboAnalyst 5.0. (**i**) Heatmap shows the differential metabolites in the statistical analysis function of MetaboAnalyst 5.0. (**j**) Pathway enrichment overview in MetaboAnalyst 5.0. Color shade is based on the *p*-value. (**k**) The demo-enriched metabolism pathway in MetaboAnalyst 5.0. Light blue

indicates that it is not an uploaded metabolite, but instead was used as background for enrichment analysis. Red indicates the metabolite is in the uploaded data and represents the different level. (**l**) An example of joint pathway analysis in MetaboAnalyst 5.0. By uploading candidate genes and metabolites, the corresponding pathway view is generated. Squares represent genes and circles represent metabolites. Red and green indicate the different levels. All images were obtained using the example data provided by the software.
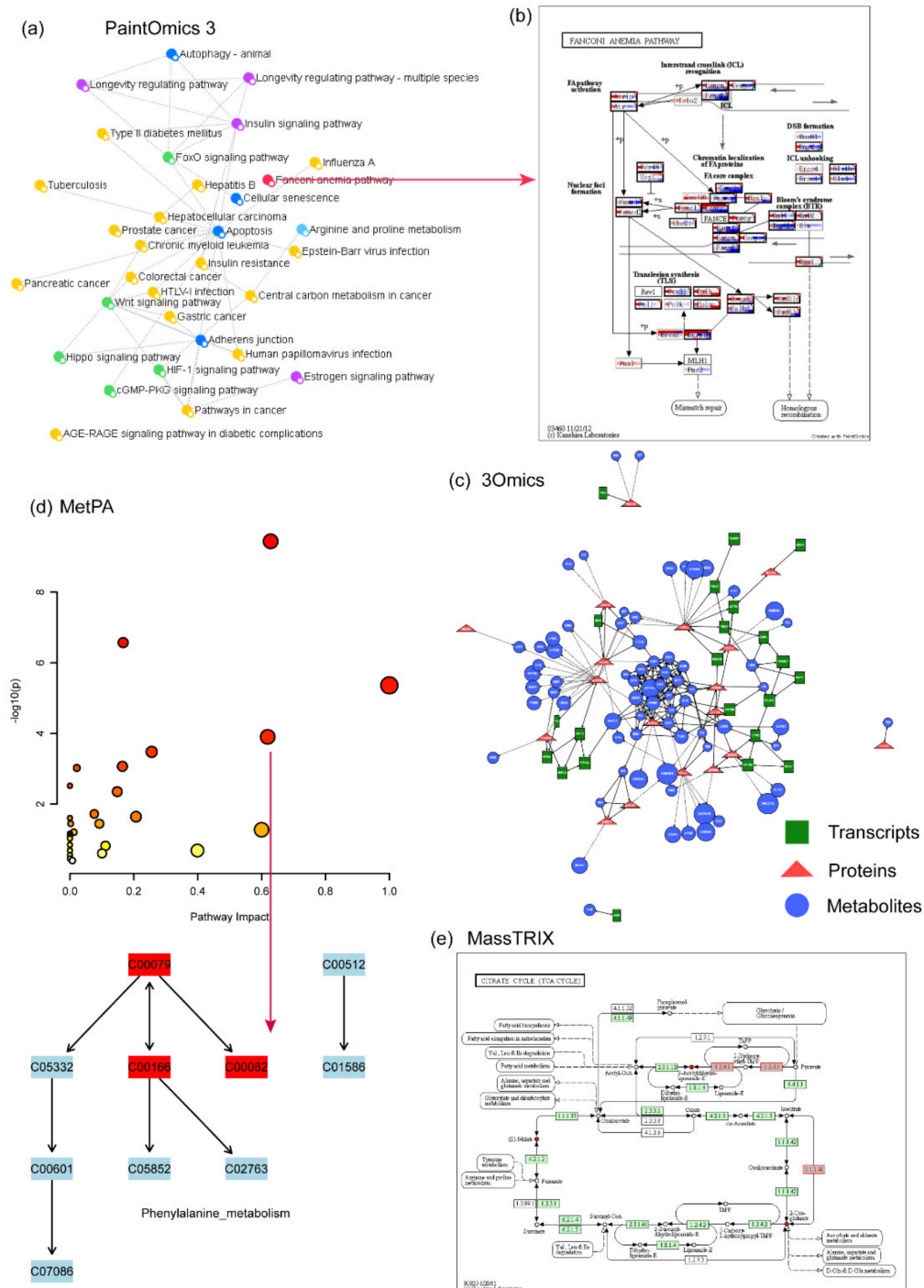

Figure 3. Example of other available metabolomics data analysis tools. (**a**) Pathway overview created by PaintOmics 3. By clicking on a circle, (**b**) the corresponding pathway view is generated, showing all genes involved in that pathway and their interactions. (**c**) A correlation network created

by 3Omics. (**d**) Pathway analysis of MetPA. MetPA is now integrated into the MetaboAnalyst 5.0 platform. (**e**) Pathway analysis of MassTRIX.

4.1. MS-DIAL

MS-DIAL was previously developed as free data pre-processing software for LC-MS data processing, but now the MS-DIAL 4.0 tool can also process LC-MS, GC-MS, and NMR data, in particular to obtain deconvoluted spectra from high-resolution GC-MS data as a prerequisite for compound identification (MS-DIAL 4.0, Hiroshi Tsugawa, Kanagawa, Japan) [100,101]. MS-DIAL offers multiple data-acquisition processing and includes the spectra for compound ID. It also includes normalization and statistical analysis options (http://prime.psc.riken.jp/, accessed on 31 March 2022) [100]. MS-DIAL has an internal GC/MS database, as well as silica retention time and MS/MS database for LC-MS/MS-based lipidomics [101].

4.2. MZmine 3

MZmine 3 (Tomáš Pluskal, Prague, Czech) is an open source software for mass spectrometry data processing that focuses on LC-MS data but can still handle GC-MS and NMR data (http://mzmine.github.io/, accessed on 13 January 2022). This software includes a complete workflow for LC-MS data analysis, including raw data processing, data filtering and peak identification, isotope detection, statistical analysis, and visualization [87].

4.3. El-MAVEN

El-MAVEN (Shubhra Agrawal, Cambridge, USA) is an open source desktop software for processing LC-MS, GC-MS and NMR data labeled in open formats (mzXML, mzML, CDF) [102]. This software has a graphical and command line interface, integrates with a cloud-based platform for storage, and conducts further analyses, such as relative fluxes and quantification [102]. El-MAVEN features a multi-file chromatography comparator, a peak feature detector, and an isotope calculator. El-MAVEN is more powerful, faster, and more user-friendly than Maven, and includes an additive calculator, fragment spectra matching, and peak editor. The El-MAVEN installer is available for Windows and Mac OS (www.elucidata.io/el-maven, accessed on 31 March 2022). Users can download the latest versions of these platforms-. Additionally, developers can follow the instructions to build El-MAVEN on Windows, Ubuntu, or Mac OS to set up the development environment (64-bit platforms only).

4.4. LipidMatch

LipidMatch can be used to annotate lipids detected by LC-MS (http://secim.ufl.edu/secim-tools/lipidmatch/, accessed on 31 March 2022). The LipidMatch fragment library contains over 250,000 lipid species spanning over 50 lipid types [103]. Users can annotate lipids in feature tables generated by its optimized peak picking and filtering strategy. LipidMatch is also used for the annotation of direct infusion and imaging experiments. The software is modular, which makes it suitable for a variety of workflows, and researchers can use it with a variety of peak picking software (e.g., MZmine 3, XCMS (Gary Siuzdak, California, USA), and MS-DIAL 4.0). LipidMatch also provides its lipid libraries in csv format and the R scripts for LipidMatch.

4.5. LipiDex

LipiDex (Joshua J Coon, Madison, WI, USA) is a unified software that can be used for lipid identification by LC-MS/MS. It has the ability to greatly reduce manual processing bias and improve the confidence of identification [104]. When using LipiDex, researchers first create a library of lipid spectra, then use fragment templates to build composite lipid spectra and mass spectrometry fragment models, and subsequently correlate spectral identifications with chromatographic peaks to generate LC-MS/MS lipidomic datasets with high confidence. LipiDex can automatically filter peak lists for additive peaks, en-

dogenous fragments, and dimers (https://github.com/coongroup/LipiDex, accessed on 31 March 2022).

4.6. MetFlow

MetFlow is a web-based tool developed in 2019 (http://metflow.zhulab.cn/, accessed on 13 January 2022) [105]. It offers a standardized workflow for metabolomics data processing and is an interactive web server. Researchers can also use it to perform data cleaning and differential analysis. Its functions include: (1) batch alignment, (2) data quality check and visualization, (3) missing value processing and outlier removal, (4) data normalization and integration, (5) statistical analysis, (6) performance validation, and (7) pathway enrichment analysis. The software enables users with little knowledge in programming and statistics to perform metabolomics data analysis. MetFlow is simple to operate. It has excellent graphic visualization ability (Figure 2a–f) and it can verify the results by uploading test data. However, its disadvantages are that the uploaded file format is fixed, and its pathway enrichment analysis cannot provide the visualization of specific pathways. Therefore, we cannot intuitively find the role of metabolites in the pathway.

4.7. MetaboAnalyst 5.0

MetaboAnalyst 5.0 is a comprehensive, freely accessible web-based metabolomics analysis platform (https://www.metaboanalyst.ca/, accessed on 13 January 2022). It was first developed in 2009 [106], then updated in 2012 (MetaboAnalyst 2.0 [107]), in 2015 (MetaboAnalyst 3.0 [108]), in 2019 (MetaboAnalyst 4.0 [94]) and in 2021 (MetaboAnalyst 5.0 [109]). It can be locally installed at the same time. MetaboAnalyst provides comprehensive online tools for metabolomics data analysis, statistical analysis, functional annotation, and visualization of data. MetaboAnalyst 5.0 improves its analytical performance and user interactivity. The platform provides four major functional modules that can be classified into 12 categories: (1) statistical analysis (statistics, biomarker analysis, multifactor/time series analysis, power analysis); (2) functional analysis (metabolome enrichment analysis, metabolic pathway analysis, mass spectrometry peak prediction of pathway activity); (3) data integration and systems biology (biomarker meta-analysis, joint-pathway analysis, and network explorer) and (4) data processing and utility functions (compound ID conversion, batch effect correction, lipidomics, and links to several spectra analysis tools). The advantages of MetaboAnalyst 5.0 are that it supports several formats of uploaded data, and the statistical methods are more selective (Figure 2g–i). The wide variety of pathway analysis methods can also meet most needs (Figure 2j–l). MetaboAnalyst 5.0 has a corresponding R package, which greatly improves the autonomy of metabolomics analysis. In addition, multiple databases are linked for multi-omics analysis. Nevertheless, MetaboAnalyst 5.0 did not have the analysis module for integration of the metabolome and microbiome, which is a disadvantage of most metabolomics analysis software.

4.8. LipidSig

LipidSig is a web-based platform for the comprehensive analysis of lipidomic data [110]. It contains five main functions: (1) profiling (for pre-processing data), (2) differential expression, (3) machine learning, (4) correlation analysis, and (5) network. LipidSig can also create interactive plots and generate downloadable images and corresponding tables (http://chenglab.cmu.edu.tw/lipidsig/, accessed on 31 March 2022).

4.9. LION

LION/web enables statistical analysis of lipids. Additionally, the most powerful feature of the software is the integration of more than 50,000 lipids with biophysical, biochemical and cell biological features, allowing a comprehensive enrichment of lipids [111]. Additionally, the authors developed a web-based interface based on LION for easy operation by researchers (www.lipidontology.com, accessed on 31 March 2022).

4.10. METLIN

The METLIN tandem mass spectrometry (MS/MS) database was created in 2003 and made publicly available in 2005 [112] to help identify metabolites. At that time, no such database existed for identifying metabolites. In 2018, to improve the coverage of metabolites and help annotate them, in silico MS/MS spectra were generated on additional molecules in METLIN. These data were based on machine learning algorithms, the METLIN database, and the unique fragmentation information (provided by stable isotopes) [90]. METLIN is a free cloud-based platform and metabolite database. It has since grown from a small collection of MS/MS spectra on 100 metabolites in its first iteration to more than 10,000 metabolites in 2012 [113], with an additional 12,000 metabolites and compounds having been analyzed in the last 5 years. METLIN data are widely used in a variety of tandem mass spectrometry instrument types (https://metlin.scripps.edu/, accessed on 13 January 2022).

4.11. PaintOmics 3

PaintOmics 3 is a web-based resource for the integrated visualization of multi-omics data types on KEGG pathway diagrams (www.paintomics.org, accessed on 13 January 2022) [114]. PaintOmics 3 combines data analysis with data visualization, providing researchers with an efficient framework for their multi-omics data. Unlike other visualization tools, PaintOmics 3 covers a comprehensive pathway analysis workflow (Figure 3a,b), including automatic feature name conversion, multi-layered feature matching, pathway enrichment, network analysis, heatmaps, trend charts, and more. It accepts a wide variety of omics types, including transcriptomics, proteomics, and metabolomics, as well as region-based approaches such as ATAC-seq or ChIP-seq data. However, the input data need to be pre-processed.

4.12. 3Omics

3Omics is a web-based tool that was developed in 2013 (http://3omics.cmdm.tw, accessed on 13 January 2022). It is used to analyze, integrate, and visualize transcriptome, proteome, and metabolome human data [115]. 3Omics supports correlation analysis, phenotype mapping, pathway enrichment analysis, and co-expression analysis (Figure 3c). In fact, depending on the input data, the software offers four parts of integrated analyses: (1) transcriptomics, proteomics, and metabolomics (T-P-M), (2) transcriptomics and proteomics (T-P), (3) proteomics and metabolomics (P-M) and (4) transcriptomics and metabolomics (T-M). A single omics analysis mode is also available in the tool. 3Omics can also carry out text mining of the biomedical literature through information Hyperlinked Over Protein (iHOP [116]) to supplement missing information. The drawback is that pathway enrichment analysis cannot provide the visualization of specific pathways.

4.13. IMPaLa

IMPaLA is a web tool for transcriptomics, proteomics, and metabolomics pathway analysis (http://impala.molgen.mpg.de, accessed on 13 January 2022) [117]. It was developed in 2011. The web tool uses over 3000 pre-annotated approaches from 11 databases to perform over-expression or enrichment analysis on uploaded metabolites and gene lists. Therefore, it is possible to identify pathways that may be regulated at the transcriptional level, metabolic level, or both. The output results of the tool include a ranked list of pathways, the size of each pathway and the p-value and q-values from the joint analysis of genes and metabolites. By clicking on the pathway name, users will be guided to a summary web page at the source database. Results can also be downloaded as a tab-delimited file.

4.14. MetPA

MetPA is a user friendly, web-based tool for the analysis and visualization of metabolomics data (http://metpa.metabolomics.ca, accessed on 13 January 2022) [118]. It combines pathway enrichment analysis programs and pathway topology feature analysis to help identify the most relevant metabolic pathways (Figure 3d). The results are displayed

in an interactive network visualization system that can be selected, dragged, and zoomed in and out. In addition, this tool offers a comprehensive compound library for metabolite name conversion, and it can also implement various univariate analyses. MetPA currently supports the analysis and visualization of 874 metabolic pathways in 11 common model organisms and it has been integrated into the MetaboAnalyst 5.0 platform.

4.15. MassTRIX

MassTRIX is a web-based software for metabolomics pathway enrichment analysis [119]. The input data of this tool require a mass peak list from high-precision MS experiments. MassTRIX marks the identified chemical compounds as differentially colored objects on the KEGG pathway maps (Figure 3e). Therefore, users can interpret the metabolic state of the organism based on the original organism and the true enzymatic capabilities in the case of submitted transcriptomics data. The tools' output page summarizes the number of identified metabolites on all available pathways and gives a list of all metabolites that are annotated on any given pathway of the organism. Here, users should note that in some cases multiple alternative annotations may be found. The MassTRIX web server is freely accessible at http://masstrix.org (accessed on 13 January 2022).

4.16. MetaCore™

MetaCore™ (http://thomsonreuters.com/metacore/, accessed on 13 January 2022) is a commercial tool used as a web-based application. The software can analyze different kinds of high-throughput molecular data. MetaCore™ is also a high-quality database of mammalian biology, with collections including metabolites and other molecular classes, bioactive molecules and their interactions, signal transduction and metabolic pathways. It also enables genomic analysis, identifies potentially important variants, and provides data visualization, analysis, and data mining. Unfortunately, no detailed information is available on how MetaCore™ works. Therefore, our review of this tool is limited.

4.17. OmicsNet

OmicsNet (www.omicsnet.ca, accessed on 31 March 2022) can integrate different omics data based on molecular interaction knowledge and visualization using network analysis. It also can annotate SNPs, microbial taxa, or LC-MS peaks for network analysis [120]. The network analysis can contain genes, proteins, transcription factors (TF), miRNAs and metabolites, and the creation of different types of biological networks is derived from multiple molecular interaction databases (PPI, TF-gene, miRNA-gene, and metabolic protein interactions).

In general, MS-DIAL, Mzmine3, and EI-MAVEN can perform data preprocessing, normalization, identification, and statistical analysis of metabolomic data. MetFlow and MetaboAnalyst 5.0 can perform most of the metabolomics analyses, including data processing, statistical analysis, and pathway analysis. Other software platforms (OmicsNet, PaintOmics3, 3Omics, IMPaLA, MetPA, MassTRIX, and others) have performed well in the subsequent analysis, including multi-omics integration and pathway analysis. In lipidomics, LipidMatch and LipiDex can be used for lipid identification, and LipidSig can perform most of the lipidomic analyses and may be the better choice. After that, the LION/web could be used for enrichment analysis. Researchers can choose different analysis software according to their needs.

5. The Integration Algorithm of Multi-Omics Data

Due to the complex and multi-factorial context of metabolic diseases, the results of metabolomic analysis should be followed by novel techniques to link the overall effectiveness between organisms, metabolites, microbiota, and individuals. Currently, the different molecular levels can be systematically divided into genomics, transcriptomics, proteomics, metabolomics, and microbiology. Genomics allows the evaluation of the whole genome of an organism and the analysis of the localization and function of genes. Transcriptomics

measures the expression of genes at a given time point. Subsequently, gene translation enables protein expression, giving rise to proteomics. Proteins can translate biologically active compounds (metabolites) into other metabolic molecules. Thus, metabolomics is an assessment of the level of metabolism in an organism, a process complicated by the fact that metabolites are in dynamic equilibrium and respond to external and internal factors. Finally, microbiomes characterize the gut microbial community [121]. Combining all of these analyses in the same biological context allows us to outline the interactions between multiple metabolite networks and gene expression markers in multiple tissues or locations, as well as determine the possible impact of microbial members on biosynthesis. Therefore, the development of an efficient and practical multi-omics algorithm is important to interpret the results of metabolomics.

In 2018, Pedersen et al. [122] proposed a calculation protocol, detailing and discussing dimensionality reduction technology and the subsequent method of integrating and interpreting multi-omics data. Dimensionality reduction of the different omics data was achieved through data normalization, the combination of co-abundant genes and metabolites, and the integration of existing biological knowledge. Using prior knowledge to overcome the functional redundancy among microbiome species is a major advancement of the method compared with existing alternative methods. Researchers can integrate multi-omics data with host physiology variables or any other phenotypes of interest to perform a three-pronged analysis to identify potential mechanistic connections through this framework and then test it through experimentation. Although it is a framework for a human metabolome-microbiome study, it is generalizable to other organisms and environmental metagenomes, and it could also be used for studies including other omics data (e.g., transcriptomics and proteomics). The R code of the protocol is available at https://bitbucket.org/hellekp/clinical-micro-meta-integration (accessed on 13 January 2022).

There are many multi-omics integration studies based on correlation analysis. In 2016, a multi-omics study by Kieffer et al. investigated the effect of a high-fat diet supplemented with resistant starch and found that the liver levels of the TCA metabolites fumarate and malate were decreased when mice were fed diets supplemented with resistant starch [123]. In 2019, a multi-omics integration revealed Parkinson's disease-specific patterns in microbial-host sulfur co-metabolism that may contribute to PD severity [124]. Multi-omics integration analysis based on an order statistic algorithm was also applied to Alzheimer's disease in 2020 [125]. In this context, a multi-omics integration analysis of metabolomics with those from other omics will help to understand the disease mechanisms and further screen key molecular markers, and then help to indicate subsequent validation experiments.

6. Conclusions and Prospects

A large amount of informative data are generated by the rapidly evolving field of metabolomics. At the same time, these data need to be integrated and analyzed with other omic data to be fully interpreted. The most common approach today is to simultaneously monitor transcript, protein and metabolite levels and obtain structural and dynamic changes in the underlying biological network of interest through integration analysis. This invokes the need for suitable statistical and computational methods to analyze and integrate these diverse and large amounts of data, and to visualize and map the data metabolites. Such multi-omics integration analysis can greatly contribute to the rapid identification of relevant metabolites and the biological processes when they are involved under specific research conditions.

To date, there are many tools available for processing and analyzing metabolomics data, and we reviewed and compared several of the commonly used tools. Overall, despite the undeniable validity of the tools reviewed, there are still several challenges in the field of metabolomics that need to be addressed. One of the biggest challenges is the reliable identification of known compounds, as well as the identification of unknown compounds.

The fine structure of isotopes enables the determination of molecular formulae of unknowns to discover substances present in the "dark metabolome". To achieve reliable identification of compounds, database-based search methods are often used, where retention times, accurate masses, isotopic properties, and fragment mass spectra must be provided to reliably resolve compounds in complex samples. Other challenges are mainly in the field of data integration, to support a thorough comprehensive evaluation of the experimental data and a deeper understanding of the biological processes.

With the increasing availability of multiple types of histological data, how to effectively use them to understand the observed abnormal biological mechanisms in metabolomics is still an open issue in the analysis. To achieve this goal, further development and improvement of computational techniques to identify and accurately quantify metabolites by integrating a priori knowledge and to integrate and finely visualize multi-omics data pathways is essential and will be the focus of the future field of bioinformatics.

References

1. Patti, G.J.; Yanes, O.; Siuzdak, G. Innovation: Metabolomics: The apogee of the omics trilogy. *Nat. Rev. Mol. Cell Biol.* **2012**, *13*, 263–269. [CrossRef] [PubMed]
2. Zhang, A.; Sun, H.; Wang, X. Serum metabolomics as a novel diagnostic approach for disease: A systematic review. *Anal. Bioanal. Chem.* **2012**, *404*, 1239–1245. [CrossRef] [PubMed]
3. Gowda, G.A.; Zhang, S.; Gu, H.; Asiago, V.; Shanaiah, N.; Raftery, D. Metabolomics-based methods for early disease diagnostics. *Expert Rev. Mol. Diagn.* **2008**, *8*, 617–633. [CrossRef] [PubMed]
4. Turi, K.N.; Romick-Rosendale, L.; Ryckman, K.K.; Hartert, T.V. A review of metabolomics approaches and their application in identifying causal pathways of childhood asthma. *J. Allergy Clin. Immunol.* **2018**, *141*, 1191–1201. [CrossRef]
5. Idle, J.R.; Gonzalez, F.J. Metabolomics. *Cell Metab.* **2007**, *6*, 348–351. [CrossRef]
6. Pacchiarotta, T.; Deelder, A.M.; Mayboroda, O.A. Metabolomic investigations of human infections. *Bioanalysis* **2012**, *4*, 919–925. [CrossRef]
7. Scrivo, R.; Casadei, L.; Valerio, M.; Priori, R.; Valesini, G.; Manetti, C. Metabolomics approach in allergic and rheumatic diseases. *Curr. Allergy Asthma Rep.* **2014**, *14*, 445. [CrossRef]
8. Johnson, C.H.; Patterson, A.D.; Idle, J.R.; Gonzalez, F.J. Xenobiotic metabolomics: Major impact on the metabolome. *Annu. Rev. Pharmacol. Toxicol.* **2012**, *52*, 37–56. [CrossRef]
9. Fahy, E.; Subramaniam, S.; Brown, H.A.; Glass, C.K.; Merrill, A.H., Jr.; Murphy, R.C.; Raetz, C.R.; Russell, D.W.; Seyama, Y.; Shaw, W.; et al. A comprehensive classification system for lipids. *J. Lipid. Res.* **2005**, *46*, 839–861. [CrossRef]
10. Walther, T.C.; Farese, R.V., Jr. Lipid droplets and cellular lipid metabolism. *Annu. Rev. Biochem.* **2012**, *81*, 687–714. [CrossRef]
11. Beloribi-Djefaflia, S.; Vasseur, S.; Guillaumond, F. Lipid metabolic reprogramming in cancer cells. *Oncogenesis* **2016**, *5*, e189. [CrossRef] [PubMed]
12. Musunuru, K.; Kathiresan, S. Surprises From Genetic Analyses of Lipid Risk Factors for Atherosclerosis. *Circ. Res.* **2016**, *118*, 579–585. [CrossRef] [PubMed]
13. Musunuru, K.; Kathiresan, S. Genetics of Common, Complex Coronary Artery Disease. *Cell* **2019**, *177*, 132–145. [CrossRef] [PubMed]

14. Han, X.; Gross, R.W. Global analyses of cellular lipidomes directly from crude extracts of biological samples by ESI mass spectrometry: A bridge to lipidomics. *J. Lipid. Res.* **2003**, *44*, 1071–1079. [CrossRef] [PubMed]

15. Wenk, M.R. The emerging field of lipidomics. *Nat. Rev. Drug Discov.* **2005**, *4*, 594–610. [CrossRef]

16. van Meer, G.; Voelker, D.R.; Feigenson, G.W. Membrane lipids: Where they are and how they behave. *Nat. Rev. Mol. Cell Biol.* **2008**, *9*, 112–124. [CrossRef]

17. Zullig, T.; Kofeler, H.C. High Resolution Mass Spectrometry in Lipidomics. *Mass Spectrom. Rev.* **2021**, *40*, 162–176. [CrossRef]

18. Pasikanti, K.K.; Esuvaranathan, K.; Hong, Y.; Ho, P.C.; Mahendran, R.; Raman Nee Mani, L.; Chiong, E.; Chan, E.C. Urinary metabotyping of bladder cancer using two-dimensional gas chromatography time-of-flight mass spectrometry. *J. Proteome Res.* **2013**, *12*, 3865–3873. [CrossRef]

19. Pasikanti, K.K.; Esuvaranathan, K.; Ho, P.C.; Mahendran, R.; Kamaraj, R.; Wu, Q.H.; Chiong, E.; Chan, E.C. Noninvasive urinary metabonomic diagnosis of human bladder cancer. *J. Proteome Res.* **2010**, *9*, 2988–2995. [CrossRef]

20. Huang, Z.; Lin, L.; Gao, Y.; Chen, Y.; Yan, X.; Xing, J.; Hang, W. Bladder cancer determination via two urinary metabolites: A biomarker pattern approach. *Mol. Cell. Proteom. MCP* **2011**, *10*, M111.007922. [CrossRef]

21. Wittmann, B.M.; Stirdivant, S.M.; Mitchell, M.W.; Wulff, J.E.; McDunn, J.E.; Li, Z.; Dennis-Barrie, A.; Neri, B.P.; Milburn, M.V.; Lotan, Y.; et al. Bladder cancer biomarker discovery using global metabolomic profiling of urine. *PLoS ONE* **2014**, *9*, e115870. [CrossRef] [PubMed]

22. Srivastava, S.; Roy, R.; Singh, S.; Kumar, P.; Dalela, D.; Sankhwar, S.N.; Goel, A.; Sonkar, A.A. Taurine—A possible fingerprint biomarker in non-muscle invasive bladder cancer: A pilot study by 1H NMR spectroscopy. *Cancer Biomark. Sect. A Dis. Markers* **2010**, *6*, 11–20. [CrossRef] [PubMed]

23. Cheng, X.; Liu, X.; Liu, X.; Guo, Z.; Sun, H.; Zhang, M.; Ji, Z.; Sun, W. Metabolomics of Non-muscle Invasive Bladder Cancer: Biomarkers for Early Detection of Bladder Cancer. *Front. Oncol.* **2018**, *8*, 494. [CrossRef] [PubMed]

24. Cheng, Y.; Xie, G.; Chen, T.; Qiu, Y.; Zou, X.; Zheng, M.; Tan, B.; Feng, B.; Dong, T.; He, P.; et al. Distinct urinary metabolic profile of human colorectal cancer. *J. Proteome Res.* **2012**, *11*, 1354–1363. [CrossRef]

25. Chen, T.; Xie, G.; Wang, X.; Fan, J.; Qiu, Y.; Zheng, X.; Qi, X.; Cao, Y.; Su, M.; Wang, X.; et al. Serum and urine metabolite profiling reveals potential biomarkers of human hepatocellular carcinoma. *Mol. Cell. Proteom. MCP* **2011**, *10*, M110.004945. [CrossRef]

26. Shariff, M.I.; Gomaa, A.I.; Cox, I.J.; Patel, M.; Williams, H.R.; Crossey, M.M.; Thillainayagam, A.V.; Thomas, H.C.; Waked, I.; Khan, S.A.; et al. Urinary metabolic biomarkers of hepatocellular carcinoma in an Egyptian population: A validation study. *J. Proteome Res.* **2011**, *10*, 1828–1836. [CrossRef]

27. Ladep, N.G.; Dona, A.C.; Lewis, M.R.; Crossey, M.M.; Lemoine, M.; Okeke, E.; Shimakawa, Y.; Duguru, M.; Njai, H.F.; Fye, H.K.; et al. Discovery and validation of urinary metabotypes for the diagnosis of hepatocellular carcinoma in West Africans. *Hepatology* **2014**, *60*, 1291–1301. [CrossRef]

28. Cox, I.J.; Aliev, A.E.; Crossey, M.M.; Dawood, M.; Al-Mahtab, M.; Akbar, S.M.; Rahman, S.; Riva, A.; Williams, R.; Taylor-Robinson, S.D. Urinary nuclear magnetic resonance spectroscopy of a Bangladeshi cohort with hepatitis-B hepatocellular carcinoma: A biomarker corroboration study. *World J. Gastroenterol.* **2016**, *22*, 4191–4200. [CrossRef]

29. Chen, J.; Wang, W.; Lv, S.; Yin, P.; Zhao, X.; Lu, X.; Zhang, F.; Xu, G. Metabonomics study of liver cancer based on ultra performance liquid chromatography coupled to mass spectrometry with HILIC and RPLC separations. *Anal. Chim. Acta* **2009**, *650*, 3–9. [CrossRef]

30. Shariff, M.I.; Ladep, N.G.; Cox, I.J.; Williams, H.R.; Okeke, E.; Malu, A.; Thillainayagam, A.V.; Crossey, M.M.; Khan, S.A.; Thomas, H.C.; et al. Characterization of urinary biomarkers of hepatocellular carcinoma using magnetic resonance spectroscopy in a Nigerian population. *J. Proteome Res.* **2010**, *9*, 1096–1103. [CrossRef]

31. Liang, Q.; Liu, H.; Wang, C.; Li, B. Phenotypic Characterization Analysis of Human Hepatocarcinoma by Urine Metabolomics Approach. *Sci. Rep.* **2016**, *6*, 19763. [CrossRef] [PubMed]

32. Osman, D.; Ali, O.; Obada, M.; El-Mezayen, H.; El-Said, H. Chromatographic determination of some biomarkers of liver cirrhosis and hepatocellular carcinoma in Egyptian patients. *Biomed. Chromatogr. BMC* **2017**, *31*, e3893. [CrossRef] [PubMed]

33. Wu, H.; Xue, R.; Dong, L.; Liu, T.; Deng, C.; Zeng, H.; Shen, X. Metabolomic profiling of human urine in hepatocellular carcinoma patients using gas chromatography/mass spectrometry. *Anal. Chim. Acta* **2009**, *648*, 98–104. [CrossRef] [PubMed]

34. Salek, R.M.; Maguire, M.L.; Bentley, E.; Rubtsov, D.V.; Hough, T.; Cheeseman, M.; Nunez, D.; Sweatman, B.C.; Haselden, J.N.; Cox, R.D.; et al. A metabolomic comparison of urinary changes in type 2 diabetes in mouse, rat, and human. *Physiol. Genom.* **2007**, *29*, 99–108. [CrossRef]

35. Adams, S.H.; Hoppel, C.L.; Lok, K.H.; Zhao, L.; Wong, S.W.; Minkler, P.E.; Hwang, D.H.; Newman, J.W.; Garvey, W.T. Plasma acylcarnitine profiles suggest incomplete long-chain fatty acid beta-oxidation and altered tricarboxylic acid cycle activity in type 2 diabetic African-American women. *J. Nutr.* **2009**, *139*, 1073–1081. [CrossRef]

36. Mihalik, S.J.; Goodpaster, B.H.; Kelley, D.E.; Chace, D.H.; Vockley, J.; Toledo, F.G.; DeLany, J.P. Increased levels of plasma acylcarnitines in obesity and type 2 diabetes and identification of a marker of glucolipotoxicity. *Obesity* **2010**, *18*, 1695–1700. [CrossRef]

37. Ha, C.Y.; Kim, J.Y.; Paik, J.K.; Kim, O.Y.; Paik, Y.H.; Lee, E.J.; Lee, J.H. The association of specific metabolites of lipid metabolism with markers of oxidative stress, inflammation and arterial stiffness in men with newly diagnosed type 2 diabetes. *Clin. Endocrinol.* **2012**, *76*, 674–682. [CrossRef]

38. Li, X.; Xu, Z.; Lu, X.; Yang, X.; Yin, P.; Kong, H.; Yu, Y.; Xu, G. Comprehensive two-dimensional gas chromatography/time-of-flight mass spectrometry for metabonomics: Biomarker discovery for diabetes mellitus. *Anal. Chim. Acta* **2009**, *633*, 257–262. [CrossRef]

39. Liu, L.; Li, Y.; Guan, C.; Li, K.; Wang, C.; Feng, R.; Sun, C. Free fatty acid metabolic profile and biomarkers of isolated post-challenge diabetes and type 2 diabetes mellitus based on GC-MS and multivariate statistical analysis. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **2010**, *878*, 2817–2825. [CrossRef]

40. Mozaffarian, D.; Cao, H.; King, I.B.; Lemaitre, R.N.; Song, X.; Siscovick, D.S.; Hotamisligil, G.S. Circulating palmitoleic acid and risk of metabolic abnormalities and new-onset diabetes. *Am. J. Clin. Nutr.* **2010**, *92*, 1350–1358. [CrossRef]

41. Lee, Y.; Pamungkas, A.D.; Medriano, C.A.D.; Park, J.; Hong, S.; Jee, S.H.; Park, Y.H. High-resolution metabolomics determines the mode of onset of type 2 diabetes in a 3-year prospective cohort study. *Int. J. Mol. Med.* **2018**, *41*, 1069–1077. [CrossRef] [PubMed]

42. Messana, I.; Forni, F.; Ferrari, F.; Rossi, C.; Giardina, B.; Zuppi, C. Proton nuclear magnetic resonance spectral profiles of urine in type II diabetic patients. *Clin. Chem.* **1998**, *44*, 1529–1534. [CrossRef] [PubMed]

43. Suhre, K.; Meisinger, C.; Doring, A.; Altmaier, E.; Belcredi, P.; Gieger, C.; Chang, D.; Milburn, M.V.; Gall, W.E.; Weinberger, K.M.; et al. Metabolic footprint of diabetes: A multiplatform metabolomics study in an epidemiological setting. *PLoS ONE* **2010**, *5*, e13953. [CrossRef] [PubMed]

44. Lu, J.; Zhou, J.; Bao, Y.; Chen, T.; Zhang, Y.; Zhao, A.; Qiu, Y.; Xie, G.; Wang, C.; Jia, W.; et al. Serum metabolic signatures of fulminant type 1 diabetes. *J. Proteome Res.* **2012**, *11*, 4705–4711. [CrossRef] [PubMed]

45. Ferrannini, E.; Natali, A.; Camastra, S.; Nannipieri, M.; Mari, A.; Adam, K.P.; Milburn, M.V.; Kastenmuller, G.; Adamski, J.; Tuomi, T.; et al. Early metabolic markers of the development of dysglycemia and type 2 diabetes and their physiological significance. *Diabetes* **2013**, *62*, 1730–1737. [CrossRef]

46. Floegel, A.; Stefan, N.; Yu, Z.; Muhlenbruch, K.; Drogan, D.; Joost, H.G.; Fritsche, A.; Haring, H.U.; Hrabe de Angelis, M.; Peters, A.; et al. Identification of serum metabolites associated with risk of type 2 diabetes using a targeted metabolomic approach. *Diabetes* **2013**, *62*, 639–648. [CrossRef]

47. Fiehn, O.; Garvey, W.T.; Newman, J.W.; Lok, K.H.; Hoppel, C.L.; Adams, S.H. Plasma metabolomic profiles reflective of glucose homeostasis in non-diabetic and type 2 diabetic obese African-American women. *PLoS ONE* **2010**, *5*, e15234. [CrossRef]

48. Hodge, A.M.; English, D.R.; O'Dea, K.; Sinclair, A.J.; Makrides, M.; Gibson, R.A.; Giles, G.G. Plasma phospholipid and dietary fatty acids as predictors of type 2 diabetes: Interpreting the role of linoleic acid. *Am. J. Clin. Nutr.* **2007**, *86*, 189–197. [CrossRef]

49. Chow, L.S.; Li, S.; Eberly, L.E.; Seaquist, E.R.; Eckfeldt, J.H.; Hoogeveen, R.C.; Couper, D.J.; Steffen, L.M.; Pankow, J.S. Estimated plasma stearoyl co-A desaturase-1 activity and risk of incident diabetes: The Atherosclerosis Risk in Communities (ARIC) study. *Metab. Clin. Exp.* **2013**, *62*, 100–108. [CrossRef]

50. Nakashima, K. Glycolytic and gluconeogenic metabolites and enzymes in the liver of obese-hyperglycemic mice (KK) and alloxan diabetic mice. *Nagoya J. Med. Sci.* **1969**, *32*, 143–158.

51. Wood, I.S.; Stezhka, T.; Trayhurn, P. Modulation of adipokine production, glucose uptake and lactate release in human adipocytes by small changes in oxygen tension. *Pflug. Arch. Eur. J. Physiol.* **2011**, *462*, 469–477. [CrossRef] [PubMed]

52. Bohm, A.; Halama, A.; Meile, T.; Zdichavsky, M.; Lehmann, R.; Weigert, C.; Fritsche, A.; Stefan, N.; Konigsrainer, A.; Haring, H.U.; et al. Metabolic signatures of cultured human adipocytes from metabolically healthy versus unhealthy obese individuals. *PLoS ONE* **2014**, *9*, e93148. [CrossRef] [PubMed]

53. Lillefosse, H.H.; Clausen, M.R.; Yde, C.C.; Ditlev, D.B.; Zhang, X.; Du, Z.Y.; Bertram, H.C.; Madsen, L.; Kristiansen, K.; Liaset, B. Urinary loss of tricarboxylic acid cycle intermediates as revealed by metabolomics studies: An underlying mechanism to reduce lipid accretion by whey protein ingestion? *J. Proteome Res.* **2014**, *13*, 2560–2570. [CrossRef] [PubMed]

54. Ho, J.E.; Larson, M.G.; Ghorbani, A.; Cheng, S.; Chen, M.H.; Keyes, M.; Rhee, E.P.; Clish, C.B.; Vasan, R.S.; Gerszten, R.E.; et al. Metabolomic Profiles of Body Mass Index in the Framingham Heart Study Reveal Distinct Cardiometabolic Phenotypes. *PLoS ONE* **2016**, *11*, e0148361. [CrossRef] [PubMed]

55. Cho, K.; Moon, J.S.; Kang, J.H.; Jang, H.B.; Lee, H.J.; Park, S.I.; Yu, K.S.; Cho, J.Y. Combined untargeted and targeted metabolomic profiling reveals urinary biomarkers for discriminating obese from normal-weight adolescents. *Pediatric Obes.* **2017**, *12*, 93–101. [CrossRef]

56. Newgard, C.B.; An, J.; Bain, J.R.; Muehlbauer, M.J.; Stevens, R.D.; Lien, L.F.; Haqq, A.M.; Shah, S.H.; Arlotto, M.; Slentz, C.A.; et al. A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance. *Cell Metab.* **2009**, *9*, 311–326. [CrossRef]

57. Graham, S.F.; Chevallier, O.P.; Elliott, C.T.; Holscher, C.; Johnston, J.; McGuinness, B.; Kehoe, P.G.; Passmore, A.P.; Green, B.D. Untargeted metabolomic analysis of human plasma indicates differentially affected polyamine and L-arginine metabolism in mild cognitive impairment subjects converting to Alzheimer's disease. *PLoS ONE* **2015**, *10*, e0119452. [CrossRef]

58. Toledo, J.B.; Arnold, M.; Kastenmuller, G.; Chang, R.; Baillie, R.A.; Han, X.; Thambisetty, M.; Tenenbaum, J.D.; Suhre, K.; Thompson, J.W.; et al. Metabolic network failures in Alzheimer's disease: A biochemical road map. *Alzheimer's Dement. J. Alzheimer's Assoc.* **2017**, *13*, 965–984. [CrossRef]

59. Proitsi, P.; Kim, M.; Whiley, L.; Pritchard, M.; Leung, R.; Soininen, H.; Kloszewska, I.; Mecocci, P.; Tsolaki, M.; Vellas, B.; et al. Plasma lipidomics analysis finds long chain cholesteryl esters to be associated with Alzheimer's disease. *Transl. Psychiatry* **2015**, *5*, e494. [CrossRef]

60. Kim, M.; Nevado-Holgado, A.; Whiley, L.; Snowden, S.G.; Soininen, H.; Kloszewska, I.; Mecocci, P.; Tsolaki, M.; Vellas, B.; Thambisetty, M.; et al. Association between Plasma Ceramides and Phosphatidylcholines and Hippocampal Brain Volume in Late Onset Alzheimer's Disease. *J. Alzheimer's Dis. JAD* **2017**, *60*, 809–817. [CrossRef]

61. Proitsi, P.; Kim, M.; Whiley, L.; Simmons, A.; Sattlecker, M.; Velayudhan, L.; Lupton, M.K.; Soininen, H.; Kloszewska, I.; Mecocci, P.; et al. Association of blood lipids with Alzheimer's disease: A comprehensive lipidomics analysis. *Alzheimer's Dement. J. Alzheimer's Assoc.* **2017**, *13*, 140–151. [CrossRef] [PubMed]

62. Snowden, S.G.; Ebshiana, A.A.; Hye, A.; An, Y.; Pletnikova, O.; O'Brien, R.; Troncoso, J.; Legido-Quigley, C.; Thambisetty, M. Association between fatty acid metabolism in the brain and Alzheimer disease neuropathology and cognitive performance: A nontargeted metabolomic study. *PLoS Med.* **2017**, *14*, e1002266. [CrossRef] [PubMed]

63. Guiraud, S.P.; Montoliu, I.; Da Silva, L.; Dayon, L.; Galindo, A.N.; Corthesy, J.; Kussmann, M.; Martin, F.P. High-throughput and simultaneous quantitative analysis of homocysteine-methionine cycle metabolites and co-factors in blood plasma and cerebrospinal fluid by isotope dilution LC-MS/MS. *Anal. Bioanal. Chem.* **2017**, *409*, 295–305. [CrossRef] [PubMed]

64. Paglia, G.; Stocchero, M.; Cacciatore, S.; Lai, S.; Angel, P.; Alam, M.T.; Keller, M.; Ralser, M.; Astarita, G. Unbiased Metabolomic Investigation of Alzheimer's Disease Brain Points to Dysregulation of Mitochondrial Aspartate Metabolism. *J. Proteome Res.* **2016**, *15*, 608–618. [CrossRef] [PubMed]

65. Koal, T.; Klavins, K.; Seppi, D.; Kemmler, G.; Humpel, C. Sphingomyelin SM(d18:1/18:0) is significantly enhanced in cerebrospinal fluid samples dichotomized by pathological amyloid-beta42, tau, and phospho-tau-181 levels. *J. Alzheimer's Dis. JAD* **2015**, *44*, 1193–1201. [CrossRef] [PubMed]

66. Mamas, M.; Dunn, W.B.; Neyses, L.; Goodacre, R. The role of metabolites and metabolomics in clinically applicable biomarkers of disease. *Arch. Toxicol.* **2011**, *85*, 5–17. [CrossRef] [PubMed]

67. Puchades-Carrasco, L.; Lecumberri, R.; Martinez-Lopez, J.; Lahuerta, J.J.; Mateos, M.V.; Prosper, F.; San-Miguel, J.F.; Pineda-Lucena, A. Multiple myeloma patients have a specific serum metabolomic profile that changes after achieving complete remission. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **2013**, *19*, 4770–4779. [CrossRef]

68. O'Shea, K.; Misra, B.B. Software tools, databases and resources in metabolomics: Updates from 2018 to 2019. *Metabolomics* **2020**, *16*, 36. [CrossRef]

69. Mattoli, L.; Gianni, M.; Burico, M. Mass Spectrometry Based Metabolomic Analysis as a Tool for Quality Control of Natural Complex Products. *Mass Spectrom. Rev.* **2022**. [CrossRef]

70. Beisken, S.; Eiden, M.; Salek, R.M. Getting the right answers: Understanding metabolomics challenges. *Expert Rev. Mol. Diagn.* **2015**, *15*, 97–109. [CrossRef]

71. Burnap, R.L. Systems and photosystems: Cellular limits of autotrophic productivity in cyanobacteria. *Front. Bioeng. Biotechnol.* **2015**, *3*, 1. [CrossRef]

72. Ma, S.; Huang, J. Regularized gene selection in cancer microarray meta-analysis. *BMC Bioinform.* **2009**, *10*, 1. [CrossRef] [PubMed]

73. Theodoridis, G.; Gika, H.G.; Wilson, I.D. Mass spectrometry-based holistic analytical approaches for metabolite profiling in systems biology studies. *Mass Spectrom. Rev.* **2011**, *30*, 884–906. [CrossRef] [PubMed]

74. Ho, C.S.; Lam, C.W.; Chan, M.H.; Cheung, R.C.; Law, L.K.; Lit, L.C.; Ng, K.F.; Suen, M.W.; Tai, H.L. Electrospray ionisation mass spectrometry: Principles and clinical applications. *Clin. Biochem. Rev.* **2003**, *24*, 3–12. [PubMed]

75. Pan, Z.; Raftery, D. Comparing and combining NMR spectroscopy and mass spectrometry in metabolomics. *Anal. Bioanal. Chem.* **2007**, *387*, 525–527. [CrossRef]

76. Veenstra, T.D. Metabolomics: The final frontier? *Genome Med.* **2012**, *4*, 40. [CrossRef]

77. Ebbels, T.M.; Lindon, J.C.; Coen, M. Processing and modeling of nuclear magnetic resonance (NMR) metabolic profiles. *Methods Mol. Biol.* **2011**, *708*, 365–388. [CrossRef]

78. DeFeo, E.M.; Wu, C.L.; McDougal, W.S.; Cheng, L.L. A decade in prostate cancer: From NMR to metabolomics. *Nat. Rev. Urol.* **2011**, *8*, 301–311. [CrossRef]

79. Vignoli, A.; Ghini, V.; Meoni, G.; Licari, C.; Takis, P.G.; Tenori, L.; Turano, P.; Luchinat, C. High-Throughput Metabolomics by 1D NMR. *Angew. Chem. Int. Ed. Engl.* **2019**, *58*, 968–994. [CrossRef]

80. Chaleckis, R.; Meister, I.; Zhang, P.; Wheelock, C.E. Challenges, progress and promises of metabolite annotation for LC-MS-based metabolomics. *Curr. Opin. Biotechnol.* **2019**, *55*, 44–50. [CrossRef]

81. Lai, Z.; Fiehn, O. Mass spectral fragmentation of trimethylsilylated small molecules. *Mass Spectrom. Rev.* **2018**, *37*, 245–257. [CrossRef] [PubMed]

82. Wei, J.; Xiang, L.; Cai, Z. Emerging environmental pollutants hydroxylated polybrominated diphenyl ethers: From analytical methods to toxicology research. *Mass Spectrom. Rev.* **2021**, *40*, 255–279. [CrossRef] [PubMed]

83. Castle, A.L.; Fiehn, O.; Kaddurah-Daouk, R.; Lindon, J.C. Metabolomics Standards Workshop and the development of international standards for reporting metabolomics experimental results. *Brief. Bioinform.* **2006**, *7*, 159–165. [CrossRef] [PubMed]

84. Zhang, A.; Sun, H.; Wang, P.; Han, Y.; Wang, X. Modern analytical techniques in metabolomics analysis. *Analyst* **2012**, *137*, 293–300. [CrossRef] [PubMed]

85. Smith, C.A.; Want, E.J.; 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. [CrossRef]

86. Clasquin, M.F.; Melamud, E.; Rabinowitz, J.D. LC-MS data processing with MAVEN: A metabolomic analysis and visualization engine. *Curr. Protoc. Bioinform.* **2012**, *37*, 14.11.1–14.11.23. [CrossRef]

87. Pluskal, T.; Castillo, S.; Villar-Briones, A.; Oresic, M. MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinform.* **2010**, *11*, 395. [CrossRef]

88. Salek, R.M.; Steinbeck, C.; Viant, M.R.; Goodacre, R.; Dunn, W.B. The role of reporting standards for metabolite annotation and identification in metabolomic studies. *Gigascience* **2013**, *2*, 13. [CrossRef]

89. Wishart, D.S.; Feunang, Y.D.; Marcu, A.; Guo, A.C.; Liang, K.; Vazquez-Fresno, R.; Sajed, T.; Johnson, D.; Li, C.; Karu, N.; et al. HMDB 4.0: The human metabolome database for 2018. *Nucleic Acids Res.* **2018**, *46*, D608–D617. [CrossRef]

90. Guijas, C.; Montenegro-Burke, J.R.; Domingo-Almenara, X.; Palermo, A.; Warth, B.; Hermann, G.; Koellensperger, G.; Huan, T.; Uritboonthai, W.; Aisporna, A.E.; et al. METLIN: A Technology Platform for Identifying Knowns and Unknowns. *Anal. Chem.* **2018**, *90*, 3156–3164. [CrossRef]

91. Sugimoto, M.; Kawakami, M.; Robert, M.; Soga, T.; Tomita, M. Bioinformatics Tools for Mass Spectroscopy-Based Metabolomic Data Processing and Analysis. *Curr. Bioinform.* **2012**, *7*, 96–108. [CrossRef] [PubMed]

92. Reshetova, P.; Smilde, A.K.; van Kampen, A.H.; Westerhuis, J.A. Use of prior knowledge for the analysis of high-throughput transcriptomics and metabolomics data. *BMC Syst. Biol.* **2014**, *8* (Suppl. S2), S2. [CrossRef] [PubMed]

93. Xia, J.; Wishart, D.S. MSEA: A web-based tool to identify biologically meaningful patterns in quantitative metabolomic data. *Nucleic Acids Res.* **2010**, *38*, W71–W77. [CrossRef] [PubMed]

94. Chong, J.; Soufan, O.; Li, C.; Caraus, I.; Li, S.; Bourque, G.; Wishart, D.S.; Xia, J. MetaboAnalyst 4.0: Towards more transparent and integrative metabolomics analysis. *Nucleic Acids Res.* **2018**, *46*, W486–W494. [CrossRef]

95. Ogata, H.; Goto, S.; Sato, K.; Fujibuchi, W.; Bono, H.; Kanehisa, M. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **1999**, *27*, 29–34. [CrossRef]

96. Noble, W.S. How does multiple testing correction work? *Nat. Biotechnol.* **2009**, *27*, 1135–1137. [CrossRef]

97. Antonelli, J.; Claggett, B.L.; Henglin, M.; Kim, A.; Ovsak, G.; Kim, N.; Deng, K.; Rao, K.; Tyagi, O.; Watrous, J.D.; et al. Statistical Workflow for Feature Selection in Human Metabolomics Data. *Metabolites* **2019**, *9*, 143. [CrossRef]

98. Chen, T.; Cao, Y.; Zhang, Y.; Liu, J.; Bao, Y.; Wang, C.; Jia, W.; Zhao, A. Random forest in clinical metabolomics for phenotypic discrimination and biomarker selection. *Evid.-Based Complementary Altern. Med.* **2013**, *2013*, 298183. [CrossRef]

99. Gromski, P.S.; Xu, Y.; Kotze, H.L.; Correa, E.; Ellis, D.I.; Armitage, E.G.; Turner, M.L.; Goodacre, R. Influence of missing values substitutes on multivariate analysis of metabolomics data. *Metabolites* **2014**, *4*, 433–452. [CrossRef]

100. Tsugawa, H.; Cajka, T.; Kind, T.; Ma, Y.; Higgins, B.; Ikeda, K.; Kanazawa, M.; VanderGheynst, J.; Fiehn, O.; Arita, M. MS-DIAL: Data-independent MS/MS deconvolution for comprehensive metabolome analysis. *Nat. Methods* **2015**, *12*, 523–526. [CrossRef]

101. Tsugawa, H.; Ikeda, K.; Takahashi, M.; Satoh, A.; Mori, Y.; Uchino, H.; Okahashi, N.; Yamada, Y.; Tada, I.; Bonini, P.; et al. A lipidome atlas in MS-DIAL 4. *Nat. Biotechnol.* **2020**, *38*, 1159–1163. [CrossRef] [PubMed]

102. Agrawal, S.; Kumar, S.; Sehgal, R.; George, S.; Gupta, R.; Poddar, S.; Jha, A.; Pathak, S. El-MAVEN: A Fast, Robust, and User-Friendly Mass Spectrometry Data Processing Engine for Metabolomics. *Methods Mol. Biol.* **2019**, *1978*, 301–321. [CrossRef] [PubMed]

103. Koelmel, J.P.; Kroeger, N.M.; Ulmer, C.Z.; Bowden, J.A.; Patterson, R.E.; Cochran, J.A.; Beecher, C.W.W.; Garrett, T.J.; Yost, R.A. LipidMatch: An automated workflow for rule-based lipid identification using untargeted high-resolution tandem mass spectrometry data. *BMC Bioinform.* **2017**, *18*, 331. [CrossRef] [PubMed]

104. Hutchins, P.D.; Russell, J.D.; Coon, J.J. LipiDex: An Integrated Software Package for High-Confidence Lipid Identification. *Cell Syst.* **2018**, *6*, 621–625.e625. [CrossRef] [PubMed]

105. Shen, X.; Zhu, Z.J. MetFlow: An interactive and integrated workflow for metabolomics data cleaning and differential metabolite discovery. *Bioinformatics* **2019**, *35*, 2870–2872. [CrossRef]

106. Xia, J.; Psychogios, N.; Young, N.; Wishart, D.S. MetaboAnalyst: A web server for metabolomic data analysis and interpretation. *Nucleic Acids Res.* **2009**, *37*, W652–W660. [CrossRef]

107. Xia, J.; Mandal, R.; Sinelnikov, I.V.; Broadhurst, D.; Wishart, D.S. MetaboAnalyst 2.0—A comprehensive server for metabolomic data analysis. *Nucleic Acids Res.* **2012**, *40*, W127–W133. [CrossRef]

108. Xia, J.; Sinelnikov, I.V.; Han, B.; Wishart, D.S. MetaboAnalyst 3.0—Making metabolomics more meaningful. *Nucleic Acids Res.* **2015**, *43*, W251–W257. [CrossRef]

109. Pang, Z.; Chong, J.; Zhou, G.; de Lima Morais, D.A.; Chang, L.; Barrette, M.; Gauthier, C.; Jacques, P.E.; Li, S.; Xia, J. MetaboAnalyst 5.0: Narrowing the gap between raw spectra and functional insights. *Nucleic Acids Res.* **2021**, *49*, W388–W396. [CrossRef]

110. Lin, W.J.; Shen, P.C.; Liu, H.C.; Cho, Y.C.; Hsu, M.K.; Lin, I.C.; Chen, F.H.; Yang, J.C.; Ma, W.L.; Cheng, W.C. LipidSig: A web-based tool for lipidomic data analysis. *Nucleic Acids Res.* **2021**, *49*, W336–W345. [CrossRef]

111. Molenaar, M.R.; Jeucken, A.; Wassenaar, T.A.; van de Lest, C.H.A.; Brouwers, J.F.; Helms, J.B. LION/web: A web-based ontology enrichment tool for lipidomic data analysis. *Gigascience* **2019**, *8*, giz061. [CrossRef] [PubMed]

112. Smith, C.A.; O'Maille, G.; Want, E.J.; Qin, C.; Trauger, S.A.; Brandon, T.R.; Custodio, D.E.; Abagyan, R.; Siuzdak, G. METLIN: A metabolite mass spectral database. *Ther. Drug Monit.* **2005**, *27*, 747–751. [CrossRef] [PubMed]

113. Tautenhahn, R.; Cho, K.; Uritboonthai, W.; Zhu, Z.; Patti, G.J.; Siuzdak, G. An accelerated workflow for untargeted metabolomics using the METLIN database. *Nat. Biotechnol.* **2012**, *30*, 826–828. [CrossRef] [PubMed]

114. Hernandez-de-Diego, R.; Tarazona, S.; Martinez-Mira, C.; Balzano-Nogueira, L.; Furio-Tari, P.; Pappas, G.J., Jr.; Conesa, A. PaintOmics 3: A web resource for the pathway analysis and visualization of multi-omics data. *Nucleic Acids Res.* **2018**, *46*, W503–W509. [CrossRef]

115. Kuo, T.C.; Tian, T.F.; Tseng, Y.J. 3Omics: A web-based systems biology tool for analysis, integration and visualization of human transcriptomic, proteomic and metabolomic data. *BMC Syst. Biol.* **2013**, *7*, 64. [CrossRef]

116. Fernandez, J.M.; Hoffmann, R.; Valencia, A. iHOP web services. *Nucleic Acids Res.* **2007**, *35*, W21–W26. [CrossRef]

117. Kamburov, A.; Cavill, R.; Ebbels, T.M.; Herwig, R.; Keun, H.C. Integrated pathway-level analysis of transcriptomics and metabolomics data with IMPaLA. *Bioinformatics* **2011**, *27*, 2917–2918. [CrossRef]

118. Xia, J.; Wishart, D.S. MetPA: A web-based metabolomics tool for pathway analysis and visualization. *Bioinformatics* **2010**, *26*, 2342–2344. [CrossRef]

119. Suhre, K.; Schmitt-Kopplin, P. MassTRIX: Mass translator into pathways. *Nucleic Acids Res.* **2008**, *36*, W481–W484. [CrossRef]

120. Zhou, G.; Xia, J. OmicsNet: A web-based tool for creation and visual analysis of biological networks in 3D space. *Nucleic Acids Res.* **2018**, *46*, W514–W522. [CrossRef]

121. Altmae, S.; Esteban, F.J.; Stavreus-Evers, A.; Simon, C.; Giudice, L.; Lessey, B.A.; Horcajadas, J.A.; Macklon, N.S.; D'Hooghe, T.; Campoy, C.; et al. Guidelines for the design, analysis and interpretation of 'omics' data: Focus on human endometrium. *Hum. Reprod. Update* **2014**, *20*, 12–28. [CrossRef] [PubMed]

122. Pedersen, H.K.; Forslund, S.K.; Gudmundsdottir, V.; Petersen, A.O.; Hildebrand, F.; Hyotylainen, T.; Nielsen, T.; Hansen, T.; Bork, P.; Ehrlich, S.D.; et al. A computational framework to integrate high-throughput '-omics' datasets for the identification of potential mechanistic links. *Nat. Protoc.* **2018**, *13*, 2781–2800. [CrossRef] [PubMed]

123. Kieffer, D.A.; Piccolo, B.D.; Marco, M.L.; Kim, E.B.; Goodson, M.L.; Keenan, M.J.; Dunn, T.N.; Knudsen, K.E.; Martin, R.J.; Adams, S.H. Mice Fed a High-Fat Diet Supplemented with Resistant Starch Display Marked Shifts in the Liver Metabolome Concurrent with Altered Gut Bacteria. *J. Nutr.* **2016**, *146*, 2476–2490. [CrossRef] [PubMed]

124. Hertel, J.; Harms, A.C.; Heinken, A.; Baldini, F.; Thinnes, C.C.; Glaab, E.; Vasco, D.A.; Pietzner, M.; Stewart, I.D.; Wareham, N.J.; et al. Integrated Analyses of Microbiome and Longitudinal Metabolome Data Reveal Microbial-Host Interactions on Sulfur Metabolism in Parkinson's Disease. *Cell Rep.* **2019**, *29*, 1767–1777.e1768. [CrossRef] [PubMed]

125. Bai, B.; Wang, X.; Li, Y.; Chen, P.C.; Yu, K.; Dey, K.K.; Yarbro, J.M.; Han, X.; Lutz, B.M.; Rao, S.; et al. Deep Multilayer Brain Proteomics Identifies Molecular Networks in Alzheimer's Disease Progression. *Neuron* **2020**, *105*, 975–991.e977. [CrossRef]