Chapter 4
Fundamentals of Mass Spectrometry-Based Metabolomics

Emilio S. Rivera, Marissa A. Jones, Emma R. Guiberson, and Jeremy L. Norris

Abstract Metabolomics involves the study of a complex and diverse array of compounds that can be thought of as the ultimate end products of the complex systems that are characteristic of molecular biology. The compounds that constitute the metabolome are small in size relative to the genome and proteome and include amino acids, carbohydrates, organic acids, lipids, and nucleotides with a mass less than 1800 Da. Understanding the role of these metabolites and the way in which changes in these important molecules impact biological processes has great potential to improve public health through better understanding of disease mechanisms. A comprehensive understanding of the metabolome will ultimately lead to better candidate biomarkers and drug targets enabling improvements in patient care. Metabolomics experiments can be divided primarily into two experimental strategies: targeted and untargeted. This monograph details these two approaches and the specific considerations for sample preparation, analytical separations, instrumental considerations, and data analysis that are required in the practice of these important technologies. Furthermore, selected applications of targeted and untargeted experiments are showcased to demonstrate the role of metabolomics as part of multi-omics studies and how metabolites can be spatially mapped in biological systems using imaging mass spectrometry.

E. S. Rivera · J. L. Norris (✉)
Mass Spectrometry Research Center, Vanderbilt University, Nashville, TN, USA

Department of Biochemistry, Vanderbilt University, Nashville, TN, USA
e-mail: emilio.s.rivera@vanderbilt.edu

M. A. Jones · E. R. Guiberson
Mass Spectrometry Research Center, Vanderbilt University, Nashville, TN, USA

Department of Chemistry, Vanderbilt University, Nashville, TN, USA
e-mail: marissa.a.jones@vanderbilt.edu; emma.r.guiberson@vanderbilt.edu

G. Sindona et al. (eds.), *Toxic Chemical and Biological Agents*, NATO Science for Peace and Security Series A: Chemistry and Biology,
https://doi.org/10.1007/978-94-024-2041-8_4

4.1 Introduction

Metabolomics is a field which studies the chemically diverse set of biological molecules that are essential components of living systems. These include molecules such as amino acids, carbohydrates, organic acids, lipids, and nucleotides. The compounds that make up the metabolome typically have molecular weights less than 1800 Da. The metabolites in this mass range are of special interest to the scientific community due to their dynamic nature and their close relation to phenotype. Metabolomics differs from its -omic counterparts (genomics, transcriptomics, and proteomics) in that metabolites are typically the final expression of a complex series of molecular events that make up system's biology [1]. Through the study of metabolomics, there is great potential to not only expand our fundamental understanding of cellular processes, but these discoveries hold the promise to change human health. The understanding of the end point of disease mechanisms could lead to improved treatment and detection through the discovery of biomarkers and drug targets which would improve patient outcomes [2]. As a result, the field of metabolomics research has expanded greatly. Publications have correspondingly increased exponentially, for example citations related to metabolomics research as increased 665 times from 1992 to 2017 [3].

With this rise in interest, a variety of technologies have been developed and are currently being applied to the study of metabolites. The principal methods include NMR (nuclear magnetic resonance) [4], GC-MS/MS (gas chromatography tandem mass spectrometry) [5], LC-MS/MS (liquid chromatography tandem mass spectrometry) [6], and IMS (imaging mass spectrometry) [7]. NMR-based methods, which have steadily been increasing for the past 15 years, have a few distinct differences from MS-based platforms: they are largely non-destructive, quantitative, and require minimal sample preparation (no derivitization, sample treatment or chromatographic separation). However, NMR is 10 to 100 times less sensitive than LC-MS and GC-MS [8]. Although each method has applications for which it is best suited, LC-MS and GC-MS methods account for ~80% of all published metabolomics studies. Due to the popularity of these methods, this review will focus primarily on mass spectrometry based methods.

4.1.1 Targeted Vs. Untargeted Assays

In general, metabolic experiments can be characterized by one of two possible experimental approaches: targeted and untargeted assays (Fig. 4.1). Targeted approaches probe a specific hypothesis, monitoring a limited number of known metabolites. Alternatively, untargeted approaches are often used for hypothesis generation and focus on broad coverage of diverse metabolites to identify both known and unknown metabolic changes.

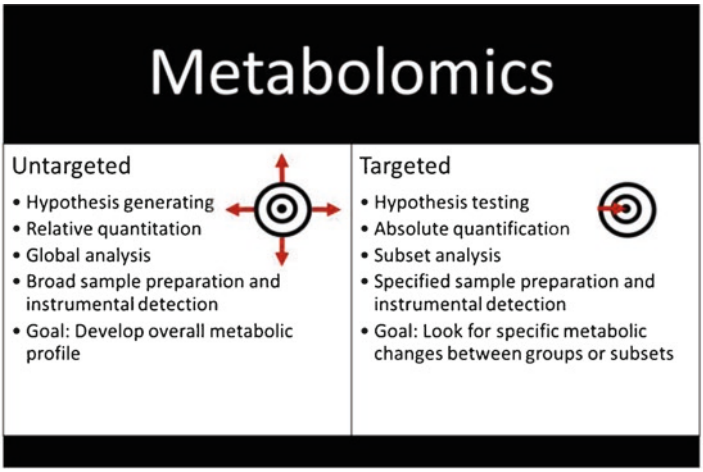

Fig. 4.1 Untargeted Vs. Targeted Metabolomics. A list of attributes of untargeted metabolomics in contrast to targeted metabolomics [9]

The targeted approach to metabolomics requires many considerations in order to produce an assay that is suitable for the biological question under consideration. The type of quantitation (*i.e.*, absolute or relative), sample preparation, separation strategy, mass spectrometer parameters, and data analysis approach all must be implemented to satisfy the ultimate experimental goal. One major advantage of targeted metabolomics is that it allows for absolute quantitation. With known analytes of interest being investigated, the necessary steps to obtain absolute quantitation are limited, only requiring the creation of a calibration curve for each analyte to be run in parallel with samples [9]. Relative quantitation is even easier to obtain experimentally as it entails comparing intensities among samples.

In targeted metabolomics, characteristics of analytes such as solubility, polarity, and pH can be used to determine the optimal assay parameters. Analytes that can be grouped together such as amino acids, or short chain fatty acids, can be processed together for all downstream steps such as extraction, derivatization, and chromatography. For example, amino acids can be extracted in aqueous solvent, as they all contain a primary amine, which allows for derivatization by dansyl chloride to aid in chromatography [10]. In addition, these derivatized products can all be separated by C18 reverse phase chromatography and require similar instrument parameters.

When taking the untargeted approach to a metabolomics experiment, the objective is to broaden the effectiveness of the workflow for simultaneous analysis of a wide variety of metabolites [11]. As the goal of an untargeted experiment is to develop an overall metabolic profile of a biological system, it is critical that methods of extraction and separation chosen be broadly inclusive of a diverse range of metabolites. To achieve these goals, metabolite experiments are comprised of an experimental design, sample preparation, separation method, mass spectrometric analysis, data processing and data interpretation (Fig. 4.2).

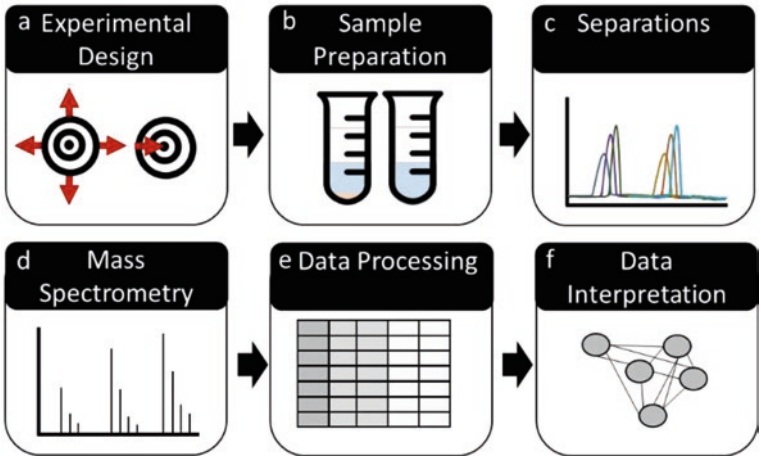

Fig. 4.2 All metabolomics experiments consist of the following: (**a**) Experimental design (targeted or untargeted experiments), (**b**) sample preparations (metabolite extraction and reconstitution), (**c**) separations (liquid chromatography), (**d**) mass spectrometry, (**e**) data processing (preprocessing and metabolite identification), and (**f**) data interpretation (metabolic network mapping)

4.2 Sample Preparation

The method by which an analyte is obtained and processed for analysis is crucial. Several considerations must be made for both targeted and untargeted extraction approaches.

For targeted metabolomics, the extraction method is dependent upon the sample matrix and class of analytes to be measured. The sample preparation process could be as simple as dilution of the sample with solvent prior to analysis [12], or as complicated as a multistep extraction involving sample preparation columns and buffer exchanges followed by a multistep derivatization. The goal of sample preparation is to mitigate any interference in the measurement of the analytes of interest that may arise from the complex biological matrix with minimal sample manipulation. Before choosing an extraction method, it is important to consider necessary down-stream manipulations. To continue with the above example, amino acids may need to be extracted from a complex matrix such as cell culture and may require derivatization for effective reverse phase chromatography. In this case, the extraction can be accomplished effectively with a mix of methanol, water and formic acid [13], which are ideal for downstream reverse phase chromatography. Regardless of the extraction method, one major consideration is loss of analyte during sample processing [14]. To address this problem, a known internal standard is introduced at or near the beginning of sample preparation to account for loss of analyte as well as any inconsistency such as pipetting error, retention time drifts in chromatography, or instrumental drift [15]. Internal standards can be utilized for normalization of analytes, where abundances are often reported as ‘response ratios’ to their respective internal

standard [16]. Resuspension is also a consideration, as the solvent chosen can have effects on downstream chromatography such as reproducibility of retention times or peak shape [17].

The optimal extraction method in untargeted approaches depends on the complexity of the sample matrix as well as the class of analytes. Having little to no information about the metabolites of interest, however, it can be difficult to discern the optimal extraction protocol. For this reason, straightforward and versatile techniques such as protein precipitation [18], Folch extraction [19], and Bligh-Dyer extraction [20] involving multiple immiscible solvents are often employed in untargeted metabolomics. Here, different classes of biomolecules are separated into isolated liquid fractions; polar metabolites suspend in the aqueous layer while lipids separate into a hydrophobic fraction such as chloroform. This phenomenon allows for simple, broad extractions of metabolites and even lends itself to multiomics workflows as each different class of biomolecules from a single sample can be easily taken for class-specific sample preparation [21].

The addition of internal standards in untargeted workflows is also common practice [22]. In this context, an internal standard could be used for normalization purposes where each analyte is reported relative to the internal standard, or it could provide a retention time reference point to provide insight on chromatographic drift over the course of an experiment [22]. More common, however, is the practice of sample pooling [23]. This involves pooling equal volumes from each sample for downstream quality control. This approach operates on the premise that a pooled sample contains every possible analyte from an entire untargeted experiment in a single injection and can thus be used to gauge both chromatography and instrumental efficiency. While analytes may in some cases be diluted in the pooled sample, this methodology works to provide qualitative insight to an experiment. Quality control of chromatographic and instrumental drift can be determined by periodically injecting the pooled sample mix between samples (after every 10 injections), over the course of the experiment.

Again, as with targeted metabolomics, the last major consideration for sample preparation is the composition of the final resuspension solvent. Trying to use generic solvents which are broadly compatible with any unknown analytes present in the sample will help to avoid analyte precipitation or having sample conditions incompatible with chromatography. Some biases can be made however, tailoring the resuspension solvent to the analytes being measured. For example, if measuring lipids, a solvent that will minimize lipid precipitation is necessary. Most lipids have been found to be soluble in chloroform making it an attractive choice for resuspension; however, chloroform would not be compatible with most reversed-phase or HILIC methods, and therefore cannot be used for resuspension in most applications. Instead, methanol, which solubilizes most lipids could serve as a substitute resuspension solvent.

4.3 Separations

Analytical separation of metabolites prior to mass analysis provide a means for more comprehensive analysis metabolites, enabling greater depth of coverage. There are analytical tasks that do not require separations for the analysis of metabolites; however, these approaches sacrifice broad metabolite coverage in favor of other important performance characteristics of the assay. For example, direct-infusion high-resolution MS (DI-HRMS) allows for the analysis of metabolites without the need for chromatographic alignment and extensive sample preparation [24]. In addition, direct-infusion methods also allow for maximum sample throughput [2]. Many imaging mass spectrometry techniques also do not use any chromatographic separations. However, these technologies allow for the unique ability to spatially localize specific m/z to regions of a sample, which can be of unique importance in clinical applications [25]. Although separation-free techniques can be used for metabolomics analysis, isomeric compounds cannot be separated and ion suppression effects must be mitigated [2]. To address these challenges, typically liquid chromatography, gas chromatography, capillary electrophoresis, and ion mobility are used.

4.3.1 Liquid Chromatography

One of the first widely accepted types of liquid chromatography in a column format was normal-phase chromatography which was derived from thin layer chromatography (TLC) [26]. Normal phase separations employ a polar stationary phase, often consisting of silica [26, 27]. This polar stationary phase is ideal for retaining and separating polar molecules in highly nonpolar solvents such as hexanes, which can be incompatible with downstream components and not provide the necessary polarity for efficient electrospray ionization [28]. While normal phase has lost much of its popularity due to its major limitations, it is still used in limited capacities due to its effective class separations of analytes such as lipids, as well as its compatibility with organic solvents which are necessary for the stability of some molecules [29].

In contrast to normal-phased chromatography, reversed-phase chromatography is defined by a nonpolar stationary phase which retains and separates nonpolar, hydrophobic analytes very effectively [30]. Historically, reversed-phase chromatography has been the gold standard in LC-MS, which has percolated into LC-MS based metabolomics [31]. Reversed-phase chromatography offers versatility in mobile phase/sample composition and can be used in flow regimes from nanoflow (< 1 μL/min) to analytical flow (>100 μL/min < 1 mL/min). Furthermore, reversed-phase chromatography produces highly reproducible retention times and peak shapes [32, 33]. One large hurdle associated with the use of reversed-phase chromatography for metabolites, however, is the inherently polar properties of the majority of endogenous small molecules. As discussed above, this problem has led to the

development of many derivatizing strategies for small molecules to make them less polar and aid in reversed-phase retention.

Hydrophilic interaction chromatography (HILIC) is a relatively new chromatography technique which is a variation of normal phase chromatography [34]. Briefly, HILIC relies on a thin layer of water which surrounds the polar stationary phase, allowing for analytes to interact with the water layer rather than the stationary phase directly [34]. This interaction with water lends to the retention of polar, hydrophilic molecules without the need for mobile phases which are incompatible with mass spectrometry. This normal-phase variant has increased opportunities for performing metabolomics without concern for analyte hydrophobicity. In addition, HILIC provides a method of separation capable of retaining and effectively resolving polar metabolites without the need for derivatization as with reversed-phase chromatography or incompatible solvents like normal-phase. Despite its clear advantages over reversed-phase and normal-phase in the context of metabolites, it does have limitations. Retention time and peak shape have been observed to be less robust than reversed-phase requiring a great deal of care in buffering of mobile phases as well as long re-equilibration periods between injections [35]. All of these factors and others have led to hesitance in the field towards adopting HILIC, with some asserting that a new method of separating polar molecules is still needed [36].

Because reversed-phase and HILIC techniques offer complementary coverage of the metabolome, they are often used together to provide a more comprehensive analysis of sample analytes [21]. Many common extraction methods such as the Folch extraction or the Bligh-Dyer extraction afford separations of metabolite classes into distinct sample fractions [19, 20]. This fractionation allows for nonpolar to be analyzed by downstream reversed-phase, and polar metabolites from the same sample to be analyzed by HILIC [21]. While this approach can significantly increase analysis time, it provides a much more comprehensive view of the metabolites in a given sample set.

4.3.2 Gas Chromatography

Gas chromatography (GC) has also been shown to provide a high degree of sensitivity and reproducibility for volatile analytes. Rather than using changes in solvent composition to separate analytes as in LC, GC takes advantage of analytes having different boiling points by ramping temperature [37]. When coupled to a mass spectrometer, GC offers reliable platform for metabolomics [38]. One consideration when integrating these techniques is an ionization source. In most GC experiments, electron-impact (EI) or chemical ionization (CI) are used for ionization before mass analysis [39]. Much like LC-MS, GC-MS can be used to effectively separate and analyze complex mixtures and is effective in both targeted and untargeted experiments. However, there are certain limitations associated with GC-based metabolomics. As GC relies on analyte volatility, it is vital that analytes be volatile enough to transition into the gas phase easily in order for GC-based methods to be effective

[40]. Historically, in the event that analytes of interest are not sufficiently volatile, derivatizations such as alkylation have been necessary to increase volatility for effective analysis by GC [40]. While effective, these derivatization techniques can be laborious and complicate data [41]. For these reasons, GC is not as widely used in metabolomics workflows as LC.

4.3.3 Capillary Electrophoresis

Another form of separation which has been gaining popularity in the field of metabolomics is capillary electrophoresis (CE). CE is not a form of chromatography because it lacks a stationary phase, a defining component of all chromatography [42]. Instead, CE separation is achieved by applying high voltage to a capillary, inducing an electrophoretic migration of ions. The electrophoretic mobility of the analytes is dependent upon the ions charge-to-size ratio [42], wherein separation of ions with differing electrophoretic mobilities is achieved. One strength of CE is its high resolution, which is directly correlated to the potential applied to the column as well as narrow peak widths provided in part by the inherent electroosmotic flow, rather than laminar flow as in traditional chromatography [43, 44]. This resolution coupled to mass spectrometry is conducive to both targeted and untargeted metabolomics. In the past, the integration of these two technologies was a limiting factor [45]. In recent years however, advancements have been made which allow for easy coupling of CE to mass spectrometry [46]. Current limitations of CE include a lack of robustness, especially related to clogging [47].

4.3.4 Ion Mobility

Another separation technology which has been demonstrated to be effective for the analysis of metabolites is ion mobility (IM) [48]. By applying a high voltage gradient opposing a gas flow, charged analytes are driven by the voltage gradient in one direction, and by the gas flow the opposite direction. These competing forces allow gas-phase separation of ions based on differing size-to-charge [49]. Because IM operates in the gas phase, it is frequently coupled with mass spectrometry, often being integrated within the mass analyzers of an instrument [50]. Ion mobility provides a degree of separation which can be comparable to that of LC-MS or GC-MS, on a much shorter timescale. Where chromatographic methods separate metabolites in a matter of minutes to hours, [51, 52] IM operates on the order of milliseconds [53]. IM is usually measured in drift time, and can be used to calculate an ion's collision cross section with proper calibration [54]. When coupled to mass spectrometry, IM provides a high degree of separation, having been shown to separate isobaric species, as well as offering this orthogonal drift time information for each analyte. Moreover, IM can be utilized in conjunction with chromatography

up-stream of a mass spectrometer, providing a second level of separation as well as affording higher peak capacity [9]. Ion mobility is not without shortcomings however, especially that the addition of ion mobility in a metabolomics experiment has been shown to reduce overall sensitivity [53].

4.4 Mass Spectrometry

In metabolomics, the instrument of choice is dependent on the experiment being conducted, where different mass spectrometer platforms are ideal for different types of assays. For targeted experiments, an instrument capable of interrogating many known molecules on a time-scale that is compatible with the chromatographic timescale is critical. Targeted assays are commonly quantitative; therefore, it is important that the instrument selected has good quantitative capabilities and sensitivity. Usually this type of work is done by a triple-quadrupole or QTRAP system [15]. Untargeted experiments have different needs, as unknown molecules must be selected for fragmentation in a manner that permits broad coverage of the analytes. Orbitrap-based and quadrupole time-of-flight (QTOF) systems have proven themselves optimal for these types of workflows [6, 55]. There are many other types of mass spectrometers which can be used for LC-MS-based metabolomics, but this chapter will focus specifically on these as they are the platforms that play a central role in the field.

Triple quadrupole and QTRAP platforms are the dominant mass spectrometers used in the field of targeted metabolomics [55]. These instruments are very similar in design, sharing an electrospray source followed by optics for the transmission of ions to an initial quadrupole capable of isolating specific *m/z* windows. A second quadrupole is then used as a collision cell for collision-induced dissociation (CID) to fragment precursor ions for MS/MS analysis [56]. These two instruments differ in their final stage where a triple quadrupole is equipped with a third quadrupole used to isolate a particular *m/z* of the fragments created in the collision cell for transmission to the detector. The QTRAP is equipped with a ion trap rather than a conventional quadrupole, which is capable of not only of performing subsequent fragmentation events on product ions but also accumulating ions for increased sensitivity [57]. This instrument can also be operated as a conventional triple quadrupole instrument. Both systems are capable of isolating a precursor ion, fragmenting, and monitoring the presence of specific fragments on the order of milliseconds [58]. The monitoring of a specific fragment of a specific parent ion is referred to as a selected reaction monitoring (SRM), and this approach can be multiplexed so that multiple analytes can be monitored in the same assay with high specificity. This approach is referred to as multiple reaction monitoring (MRM), and is paramount for targeted LC-MS techniques [59]. Practically, these mass spectrometers are capable of quantitatively monitoring upwards of 50 unique transitions within milliseconds [58]. This level of speed allows for dozens of metabolites to be measured with

sufficient coverage over a chromatographic peak, producing quantitative, reproducible data for each metabolite.

Untargeted metabolomics comes with a different set of requirements due to the chemical diversity of analytes that are characterized in a single analysis. Here, an instrument's ability to quickly and efficiently identify a molecule which may be of interest for further investigation by MS/MS is critical. There exist two main methods of addressing this problem: data-dependent-acquisition (DDA) and data-independent-acquisition (DIA). Various vendors have different names for these processes, but in general these two techniques prevail in untargeted metabolomics workflows [9]. DDA has proven to be effective in the field of proteomics for years [60]. For DDA, tandem mass spectra are collected for *m/z* values selected from a previously acquired MS spectrum; where the instrument is set to perform an initial MS scan, determine the N most abundant ions, and isolate those for MS/MS analysis [61]. This entire process must take place in milliseconds [62], and is repeated several times over the course of a peak for effective and reliable MS/MS data.

DIA has grown more prominent recently, where independent of the MS data, the instrument indiscriminately isolates mass windows (which in some cases can be chosen by the operator) across the entire MS mass range [63]. Each of these windows is sequentially isolated for fragmentation, meaning that ions within each window will proceed to the collision cell, producing fragments that are subsequently analyzed [64]. One exception to this generalization is in Waters platforms, which refer to their version of DIA as '$MS^{E'}$ and fragments all precursor ions simultaneously following an initial MS scan. Regardless of the how the DIA is carried out, as with DDA, this entire process must be repeated several times within a peak width for reliable fragmentation and quantitation [64]. The difficulty in this approach is pairing fragment ions with their parent ion counterparts. This requires sophisticated software programs for data annotation [65]. Thus, the specific data analysis approach must be carefully considered before acquiring data to ensure success.

It should be noted that fragmentation, while integral to both workflows, serves a slightly different purpose for targeted and untargeted metabolomics. In a targeted experiment, fragmentation is used as validation of the previously known identity of a given molecule. The resulting fragment is measured, meaning that quantitation is based solely on the abundance of fragment ions [66]. In untargeted metabolomics, MS/MS is used to extract more information about the possible identity of a given molecule [11]. In this case, any quantitative measurements are most frequently made using MS data [11]. Lastly, the mass resolving power of the instrument tends to differ between targeted and untargeted workflows. Untargeted metabolomics benefits from higher mass resolving power as exact mass measurements in the MS data can aid in identifying a metabolite [67]. For this reason, Orbitrap-based and QTOF systems have prevailed as the ideal platforms for an untargeted experiment. Orbitraps offer varying levels of mass resolving power, ranging from 15,000–240,000 [68], and can be modulated by the operator to suit an experiment. While not as high performance, QTOF instruments offer mass resolving powers of up to 60,000 on current platforms [69]. In contrast, a targeted experiment relies on specific fragmentation and retention time for identification of a metabolite, rather than mass

resolving power [9]. A targeted metabolomics experiment is also greatly affected by the speed of the mass spectrometer being used. For this reason, most instrument platforms used in targeted workflows are low resolution mass spectrometers capable of rapid scan rates that enable high-speed MRM analyses. Both targeted and untargeted workflows require high levels of data management and processing, which can often prove to be the most cumbersome and time-intensive component of a metabolomics experiment.

4.5 Data Analysis

The validity of identifications in metabolomics studies are crucial for making biological conclusions. A useful framework for considering the confidence in the identification of a specific metabolite has been established [9]. In this framework, there are five levels of validation ranging from a unique feature such as an accurate mass measurement at level five (the lowest level of validation), to a validated identification which is measured against a reference standard with a confirmed structure at level one (the highest level of validation) (Fig.4.3).

One universal challenge between targeted and untargeted experiments is determining which spectra in these large data sets are representative of a real metabolite. In order to make confident identifications, informative features relevant to

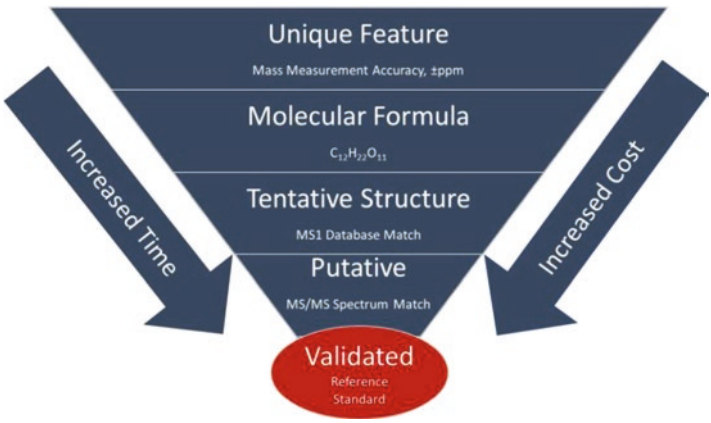

Fig. 4.3 Levels of Confidence for Metabolite Validation. From top to bottom the least validated level 5 to validated level 1 is shown. With increased level of validation comes increased time and cost. A level five validated feature is a unique feature with accurate mass measurement. A level four validation is a unique molecular formula. A level three validated feature includes tentative structure and matches a precursor to an MS1 data based. A level two validated feature, a putative identification matches MS/MS spectra to a database. Level two and three validations utilize orthogonal measurements and can be techniques other than mass spectrometry such as NMR, collisional cross section, spectroscopy, or retention time. A completely validated metabolite, a level one validation matches a metabolite to a reference standard [9]

biological processes must be differentiated from extraneous ones. For example, Mahieu et al. showed that although a dataset had over 25,000 features, with the subtraction of isotopes, adducts, artifacts, and contaminants, less than 1000 were metabolites [70]. Preprocessing methods facilitate recognition of data as either meaningful or irrelevant.

There are a variety of tools for data preprocessing such as noise filtering, spectral deconvolution, chromatogram alignment, and retention time correction. Data processing such as peak detection, peak alignment, metabolite identification, quality control, normalization, statistical analysis, metabolite quantification, and *in silico* fragmentation are also used [71]. Both targeted and untargeted metabolomics methods share similar data preprocessing. With targeted methods such as MRM, chromatographic features linked to specific MS/MS transitions are often used. A variety of commercial software, such as LCQuan (ThermoFisher Scientific), allow for the identification of internal standards for relative quantitation or the import of calibration curves for absolute quantitation [72].

To reduce the dimensionality of the data, classification and clustering tools are used [73]. Once metabolites are identified, relative or absolute quantitation can be performed to determine the overall role of observed metabolic changes in a global framework. There are a variety of commercially available tools for targeted and untargeted LC-MS data analysis including LCQuan, Agilent Masshunter, Bruker's Profile Analysis, Thermo SIEVE, Waters' Progensis QI and more [72, 74–77]. In addition there are a number of open source, vendor-independent tools including XCMS/XCMS Online, Mzmine 2, and MS-DIAL [78–80].

After data preprocessing, metabolite identification remains challenging owing to incomplete spectral libraries and incompatibility between databases and data types. For example, some databases are compatible with MS^n data while others are only designed to search compounds. Table 4.1 provides a list of relevant spectral libraries and databases to assist in the identification of metabolites. When determining which database best fits an experiment, it is important not only to consider the total number of compounds and the data type, but also the original data used to build the data base. For example, it is possible to limit false identifications for a human-based experiment by selecting HMDB rather than an *in silco* prediction-based database.

Once metabolites are identified, pathways become integral in identifying the collective role metabolites play in relation to a scientific question. Table 4.2 lists a few metabolic pathway analysis tools, their number of reference pathways and the number of organisms on which they are based. Using these tools, data sets can be mapped into known biological networks to aid in the interpretation of the results and provide context that will help generate future hypotheses. Although pathway analysis can often bring about answers to a variety of biological questions, it is important to note that many experiments are temporal and looking at the accumulation of metabolites or the change between experimental groups. In order to directly follow a metabolic pathway, heavy labeling experiments are needed [87].

Table 4.1 A variety of spectral libraries and databases are available for metabolite identification. From left to right the database or spectral library, the number of total compounds, target data types, organism base/ Focus and a brief description are shown [71, 81–86]

Spectral Library/ database	Total Compounds	Targets	Organism Base/ focus	Description
MoNA	>200,000	EI, MS/MS, MSn	Multiple species, Curated	Curated Spectra
Metlin	>500,000	CID-MS/MS	Multiple species	Commonly used, Original use QTOF
NIST	>574,000	EI-MS, CID-MS/MS	Multiple species	Curated database
m/z cloud	8904	MSn	Multiple species	Multiple stage MSn
KEGG	18,612	Metabolites	Multiple species	Pathway database
HMDB	114,100	Metabolites	Human	Spectra, physical and biological properties
ChemSpider	67,000,000	All small molecules	Curated data, compounds	Curated data
Mass Bank	>38,000	EI, MS/MS, MSn	Multiple species	Long standing community database
MINE	>571,000	Metabolites	In silico predicted metabolites	Predicted database

Table 4.2 Pathway analysis databases provide the biological context for individual metabolite measurements within a system. From left to right the database, number of reference pathways, and organisms included are shown

Database	Reference Pathways	Organisms
KEGG[88]	372	>700
MetaCyc[89]	1100	1500
WikiPathways[90]	100	20

4.6 Other Approaches to Metabolomics

Pathway analysis introduces a unique view of identified metabolites and their biological relevance. Metabolomics offers a plethora of biological significance through the metabolites identified, but many times metabolomics through routine LC-MS lack dimensions of information. There are highly complementary approaches that can be used in the study of metabolism. Three specific examples are described below: multiomic sample preparation methodologies, imaging MS for the addition of spatial information, and NMR for high reproducibility [88–90].

4.6.1 Combining Metabolomics with Other Omics Technolgies

LC-MS has been successfully utilized for metabolomics, but previous sample preparation methods made metabolomics incompatible with proteomic and lipidomic analysis. In order to maximize the data extracted from a sample, a new method known as sample preparation for multi-omics technologies (SPOT) has been developed for high-throughput multi-omics analysis by various collaborations at Vanderbilt University [25, 91]. This technology allows for proteomic, transcriptomic, and metabolomic analysis from the same sample, with common sample preparation methodology. This common preparation allows for high-throughput sample analysis, which would be optimal for applications such as rapid threat assessment. This LC-MS based method allows for temporally resolved data sets in addition to multi-omics analyses, optimal for addressing complex bio-logical questions.

This novel multi-omics sample preparation method utilizes cells but can be applied to tissue samples as well. Cells are lysed, undergo a freeze-thaw cycle, and then are sonicated in an ice bath. Aliquots are then lysed and precipitated with 75:25 Acetone: Ethanol for 2 h, then spun down. The resulting supernatant is then collected for metabolomic analysis while the precipitate is used for proteomics analysis. SPOT applied to metabolomics is best utilized for untargeted analysis. 50 microliters of supernatant extracted from SPOT sample preparation were analyzed through either reverse phase LC or HILIC in a global untargeted analysis with simultaneous analysis of molecular fragmentation. This approach showed reproducible results comparable to traditional metabolomic methods and is efficient with the ability to take cells from pellets to desalted samples ready for MS analysis within 9 h. Additionally, this method led to the extraction of changing metabolites key for biological information. SPOT was applied to human acute promyelocytic leukemia (HL-60) cells that were exposed to zinc intoxication. Additionally, data was collected at various time points throughout the analysis from 6 h to 24 h.

This investigation highlighted three pathways that appeared significantly modified with zinc treatment: tryptophan metabolism, purine metabolism, and eicosanoid signaling. Metabolomics allowed for the discovery of cellular responses not found with proteomics and transcriptomics on the same sample. These pathways were previously identified with genomic technologies and are supported by these metabolomic data sets extracted using SPOT. The continued use of the SPOT protocol will answer many biological questions, through the incorporation of high-throughput, time-resolved, large-scale data sets for untargeted multi-omics analysis.

4.6.2 Metabolomic Analysis with Imaging Mass Spectrometry

IMS has been used with high success for metabolomic analysis. Although LC-MS is more suitable for absolute quantitation, IMS maintains the spatial information from a tissue section. A paper from 2018 utilized IMS, in coordination with immunohistochemistry, qPCR, western blotting and enzyme assays, to elucidate the

regional differences in glucose metabolism in the brain [91]. IMS is an untargeted, label-free technology, but can also be used to visualize the localization of targeted metabolites such as those related to glucose metabolism. In this paper ATP, ADP, HP and HBP were all identified to determine regional differences in metabolism within the brain. This approach allows for direct measurement of metabolites generated through specific pathways and in specific brain regions. Then, using immunohistochemistry and Nissl staining, these regional differences can be visualized a high spatial resolution.

MALDI IMS sample preparation differs from LC-MS methods, due to its retention of spatial information. Tissue can be sectioned and mounted onto glass slides, then sprayed with matrix to improved ionization efficiency. This maintains the relationship of the robust chemical information provided by MS with spatial location in the tissue, offering new correlations between the molecular makeup of the tissue and the various regions and substructures. Overall this approach allowed for the identification of key metabolites and gives insight into the relationship between brain regions and pathways. Metabolites are tentatively identified by exact mass, and then confirmed by MS^3 fragmentation experiments.

IMS showed the regional variations between areas that use glucose for glycolysis versus areas that use glucose primarily for the pentose phosphate pathway (PPP). For example, more of the glucose in the thalamus is entering the PPP over other regions such as the amygdala where more glucose is utilized in other pathways such as glycolysis. However, in white matter tracts and regions with low glycolysis and PPP, ATP production is high. Additionally, this investigation showed an increase in lactate during fasting that shows regional localization to specific brain substructures. Overall IMS allows for spatially resolved metabolomics, also showing the ability to conduct high resolution metabolomics with the addition of spatial information in regions of interest to study specific pathways.

4.6.3 Other Metabolomics Methods: Nuclear Magnetic Resonance

Previous methods described for metabolomics utilize mass spectrometry for analysis. While MS is higher in sensitivity by orders of magnitude, other technologies such as nuclear magnetic resonance (NMR) have been growing in their applicability to metabolomic research. NMR has a variety of advantages over MS [8]. The sample preparation for NMR is relatively easy compared to LC-MS, high experimental reproducibility, and NMR is nondestructive for samples. One of the major benefits of NMR, however, is the ability to quantify the metabolite levels explicitly. Due to these advantages and the high automatability, NMR-based metabolomics has been increasing over the last 15 years. With NMR technologies such as MRI and ssNMR, living cells and entire organs can be analyzed due to the nondestructive nature of NMR, applications that are currently inaccessible for MS. Ultimately the choice between NMR and MS relies on the priorities of the experiment: high sensitivity and more identifications, or nondestructive analysis.

NMR suffers from limited spectral bandwidth when analyzing complex metabolomics samples, however, which can make untargeted complex mixtures difficult to analyze. 2D NMR spectra allows for more information at overlapping resonances, helping with further separation of peaks. 2D NMR involves the plotting of two frequency axes against each other, allowing for visualization of correlation be-tween different peaks using either homonuclear or heteronuclear correlations [92]. While 2D NMR applied to metabolomics can be cumbersome, the Gi-raudeau group has recently described a fast quantitative 2D NMR workflow for metabolomics and lipidomics [93]. This approach specifically mentions UF COSY (ultrafast correlation spectroscopy), ^{1}H-^{13}C HSQC (heteronuclear single-quantum correlation spectroscopy), and ZF-TOCSY (Z-filter total correlation spectroscopy) as their approaches, but their workflow can be applied to any 2D NMR approach. Previously 2D NMR experiments required long acquisition, up to several hours per spectrum, as well as difficulties in quantitation. This new fast 2D NMR workflow reduces acquisition time and allows for quantitation for both targeted and untargeted approaches. For targeted approaches, standard additions or calibration are incorporated into the sample design, which untargeted approaches utilize involved data processing and statistics.

4.7 Conclusion

Metabolomics is a growing field with a variety of analytical and computational tools for analyzing a broad, dynamic, and diverse chemical and biological spaces. Strategies exist for analyzing the metabolome for both hypothesis generation and hypothesis testing. Specifically, mass spectrometry enables the interrogation of this chemical space to answer a biological question. However, the experimental design including design (targeted/untargeted), sample preparation, separations, data acquisition, and data analysis tailored towards the ultimate question is integral to a successful experiment. As the endpoint of biochemical processes, the metabolome is uniquely suited to provide a broad, yet specific view biologically processes that closely relate to phenotype, especially for biological and medicinal applications.

References

1. Ryan D, Robards K (2006) Metabolomics: the greatest omics of them all?. https://doi.org/10.1021/AC0614341
2. Ren J-L, Zhang A-H, Kong L, Wang X-J (2018) Advances in mass spectrometry-based metabolomics for investigation of metabolites. RSC Adv 8(40):22335–22350. https://doi.org/10.1039/C8RA01574K
3. Guo S, Tian J, Zhu B, Yang S, Yu K, Zhao Z (2018) Trends in metabolomics research: A Scientometric analysis (1992–2017). Curr Sci 114(11)
4. Markley JL, Brüschweiler R, Edison AS, Eghbalnia HR, Powers R, Raftery D, Wishart DS (2017) The future of NMR-based metabolomics. Curr Opin Biotechnol 43:34–40. https://doi.org/10.1016/J.COPBIO.2016.08.001

5. Papadimitropoulos M-EP, Vasilopoulou CG, Maga-Nteve C, Klapa MI (2018) Untargeted GC-MS metabolomics. Humana Press, New York, pp 133–147. https://doi.org/10.1007/978-1-4939-7643-0_9
6. Zhou B, Xiao JF, Tuli L, Ressom HW (2012) LC-MS-based metabolomics. Mol BioSyst 8(2):470. https://doi.org/10.1039/C1MB05350G
7. Dueñas ME, Larson EA, Lee YJ (2019) Toward mass spectrometry imaging in the metabolomics scale: increasing metabolic coverage through multiple on-tissue chemical modifications. Front Plant Sci 10(860). https://doi.org/10.3389/fpls.2019.00860
8. Emwas A-H, Roy R, McKay RT, Tenori L, Saccenti E, Gowda GAN, Raftery D, Alahmari F, Jaremko L, Jaremko M et al (2019) NMR spectroscopy for metabolomics research. Meta 9(7):123. https://doi.org/10.3390/metabo9070123
9. Schrimpe-Rutledge AC, Codreanu SG, Sherrod SD, McLean JA (2016) Untargeted metabolomics strategies—challenges and emerging directions. J Am Soc Mass Spectrom 27(12):1897–1905. https://doi.org/10.1007/s13361-016-1469-y
10. Gray WR (1967) Dansyl Chloride Procedure. *Methods Enzymol* 11(C):139–151. https://doi.org/10.1016/S0076-6879(67)11014-8
11. Vinayavekhin N, Saghatelian A Untargeted metabolomics. In: Current protocols in molecular biology. Wiley, Hoboken, p 2010. https://doi.org/10.1002/0471142727.mb3001s90
12. Enders JR, McIntire GL (2015) A dilute-and-shoot LC-MS method for quantitating opioids in Oral fluid. J Anal Toxicol 39(8):662–667. https://doi.org/10.1093/jat/bkv087
13. Ser Z, Liu X, Tang NN, Locasale JW (2015) Extraction parameters for metabolomics from cultured cells. Anal Biochem 475:22–28. https://doi.org/10.1016/j.ab.2015.01.003
14. Boyd RK (1993) Quantitative trace analysis by combined chromatography and mass spectrometry using external and internal standards. Rapid Commun Mass Spectrom:257–271. https://doi.org/10.1002/rcm.1290070402
15. Roberts LD, Souza AL, Gerszten RE, Clish CB Targeted Metabolomics. *Curr Protoc Mol Biol* 2012:1. (SUPPL.98. https://doi.org/10.1002/0471142727.mb3002s98
16. Dunn WB, Ellis DI (2005) Metabolomics: current analytical platforms and methodologies. TrAC Trends Anal Chem 24(4):285–294. https://doi.org/10.1016/j.trac.2004.11.021
17. Cajka T, Fiehn O (2014) Comprehensive analysis of lipids in biological systems by liquid chromatography-mass spectrometry. TrAC Trends Anal Chem:192–206. https://doi.org/10.1016/j.trac.2014.04.017
18. Jiang L, He L, Fountoulakis M (2004) Comparison of protein precipitation methods for sample preparation prior to proteomic analysis. J Chromatogr A 1023(2):317–320. https://doi.org/10.1016/j.chroma.2003.10.029
19. Folch J, Lees M, Sloane GH (2019) A simple method for the isolation and purification of total Lipides from animal tissues* Downloaded From; 2019
20. Bligh EG, Dyer WJ A rapid method of total lipid extraction and purification
21. Gutierrez DB, Gant-Branum RL, Romer CE, Farrow MA, Allen JL, Dahal N, Nei YW, Codreanu SG, Jordan AT, Palmer LD et al (2018) An integrated, high-throughput strategy for Multiomic systems level analysis. J Proteome Res 17(10):3396–3408. https://doi.org/10.1021/acs.jproteome.8b00302
22. Gertsman I, Barshop BA (2018) Promises and pitfalls of untargeted metabolomics. J Inherit Metab Dis:355–366. https://doi.org/10.1007/s10545-017-0130-7
23. Broadhurst D, Goodacre R, Reinke SN, Kuligowski J, Wilson ID, Lewis MR, Dunn WB (2018) Guidelines and considerations for the use of system suitability and quality control samples in mass spectrometry assays applied in untargeted clinical Metabolomic studies. *Metabolomics*. https://doi.org/10.1007/s11306-018-1367-3
24. Haijes HA, van der Ham M, Gerrits J, van Hasselt PM, Prinsen HCMT, de Sain-van der Velden MGM, Verhoeven-Duif NM, Jans JJM (2019) Direct-infusion based metabolomics unveils biochemical profiles of inborn errors of metabolism in cerebrospinal fluid. Mol Genet Metab 127(1):51–57. https://doi.org/10.1016/J.YMGME.2019.03.005
25. Norris JL, Caprioli RM Analysis of tissue specimens by matrix-assisted laser desorption/ ionization imaging mass spectrometry in biological and clinical research. https://doi.org/10.1021/cr3004295

26. Jandera P (2005) Liquid chromatography | normal phase. In: Encyclopedia of analytical science. Elsevier, pp 142–152. https://doi.org/10.1016/b0-12-369397-7/00324-1
27. Scott RPW (2000) Chromatography: Liquid | mechanisms: normal phase. In: Encyclopedia of separation science. Elsevier, pp 706–711. https://doi.org/10.1016/b0-12-226770-2/00301-x
28. Jiang P, Lucy CA (2016) Coupling normal phase liquid chromatography with electrospray ionization mass spectrometry: strategies and applications. Anal Methods:6478–6488. https://doi.org/10.1039/c6ay01419d
29. Abbott SR (1980) Practical aspects of normal-phase chromatography. J Chromatogr Sci 18(10):540–550. https://doi.org/10.1093/chromsci/18.10.540
30. Patti GJ, Yanes O, Siuzdak GI (2012) Metabolomics: the apogee of the omics trilogy. Nat Rev Mol Cell Biol:263–269. https://doi.org/10.1038/nrm3314
31. Dettmer K, Aronov PA, Hammock BD Mass spectrometry-based metabolomics. Mass Spectrom Rev 26(1):51–78. https://doi.org/10.1002/mas.20108
32. Patterson RE, Ducrocq AJ, McDougall DJ, Garrett TJ, Yost RA (2015) Comparison of blood plasma sample preparation methods for combined LC-MS lipidomics and metabolomics. J Chromatogr B Anal Technol Biomed Life Sci 1002:260–266. https://doi.org/10.1016/j.jchromb.2015.08.018
33. Cutillas P (2005) Principles of nanoflow liquid chromatography and applications to proteomics. Curr Nanosci 1(1):65–71. https://doi.org/10.2174/1573413052953093
34. Alpert AJ (1990) Hydrophilic-interaction chromatography for the separation of peptides, nucleic acids and other polar compounds. J Chromatogr A 499(C):177–196. https://doi.org/10.1016/S0021-9673(00)96972-3
35. Simon R, Enjalbert Q, Biarc J, Lemoine J, Salvador A (2012) Evaluation of Hydrophilic Interaction Chromatography (HILIC) versus C18 reversed-phase chromatography for targeted quantification of peptides by mass spectrometry. J Chromatogr A 1264:31–39. https://doi.org/10.1016/j.chroma.2012.09.059
36. Naser FJ, Mahieu NG, Wang L, Spalding JL, Johnson SL, Patti GJ (2018) Two complementary reversed-phase separations for comprehensive coverage of the semipolar and nonpolar metabolome. Anal Bioanal Chem 410(4):1287–1297. https://doi.org/10.1007/s00216-017-0768-x
37. Martin AJP, Synge RLM (1941) A new form of chromatogram employing two liquid phases. Biochem J 35(12):1358–1368
38. Roessner U, Wagner C, Kopka J, Tretheway RN, Willmitzer L (2001) Simultaneous analysis of metabolites in tuber by gas chromatography-mass spectrometry. Plant J 23:1
39. Castrillo JI, Hayes A, Mohammed S, Gaskell SJ, Oliver SG (2003) An optimized protocol for metabolome analysis in yeast using direct infusion electrospray mass spectrometry. Phytochemistry 62:929–937
40. Halket JM, Waterman D, Przyborowska AM, Patel RKP, Fraser PD, Bramley PM (2005) Chemical derivatization and mass spectral libraries in metabolic profiling by GC/MS and LC/MS/MS. J Exp Bot 56(410):219–243. https://doi.org/10.1093/jxb/eri069
41. Jonsson P, Johansson AI, Gullberg J, Trygg J, A J, Grung B, Marklund S, Sjöström M, Antti H, Moritz T (2005) High-throughput data analysis for detecting and identifying differences between samples in GC/MS-based metabolomic analyses. Anal Chem 77(17):5635–5642. https://doi.org/10.1021/ac050601e
42. Monnig CA, Kennedy RT (2014) Capillary electrophoresis. Food Toxic Anal Tech Strateg Dev 1997(12):561–597. https://doi.org/10.1021/acs.analchem.5b04125
43. Ewing AG, Wallingford RA, Olefirowicz TM (1989) Capillary electrophoresis. Anal Chem 61(4):292A–303A. https://doi.org/10.1021/ac00179a002
44. VanOrman BB, Liversidge GG, McIntire GL, Olefirowicz TM, Ewing AG (1990) Effects of buffer composition on electroosmotic flow in capillary electrophoresis. J Microcolumn Sep 2(4):176–180. https://doi.org/10.1002/mcs.1220020404
45. Olivares JA, Nguyen NT, Yonker CR, Smith RD (1987) On-line mass spectrometric detection for capillary zone electrophoresis. Anal Chem 59(8):1230–1232. https://doi.org/10.1021/ac00135a034

46. Soga T, Ohashi Y, Ueno Y, Naraoka H, Tomita M, Nishioka T (2003) Quantitative metabolome analysis using capillary electrophoresis mass spectrometry. J Proteome Res 2(5):488–494. https://doi.org/10.1021/pr034020m
47. Nowak PM, Woźniakiewicz M, Gładysz M, Janus M, Kościelniak P (2017) Improving repeatability of capillary electrophoresis—a critical comparison of ten different capillary inner surfaces and three criteria of peak identification. Anal Bioanal Chem 409(18):4383–4393. https://doi.org/10.1007/s00216-017-0382-y
48. Dwivedi P, Wu P, Klopsch SJ, Puzon GJ, Xun L, Hill HH (2008) Metabolic profiling by Ion Mobility Mass Spectrometry (IMMS). Metabolomics 4(1):63–80. https://doi.org/10.1007/s11306-007-0093-z
49. Miller RA, Eiceman GA, Nazarov EG, King AT (2000) Novel micromachined high-field asymmetric waveform-ion mobility spectrometer. Sensors Actuators B Chem 67(3):300–306. https://doi.org/10.1016/S0925-4005(00)00535-9
50. Kanu AB, Dwivedi P, Tam M, Matz L, Hill HH (2008) Ion mobility-mass spectrometry. J Mass Spectrom 43(1):1–22. https://doi.org/10.1002/jms.1383
51. Berm EJJ, Paardekooper J, Brummel-Mulder E, Hak E, Wilffert B, Maring JG (2015) A simple dried blood Spot method for therapeutic drug monitoring of the tricyclic antidepressants amitriptyline, nortriptyline, imipramine, clomipramine, and their active metabolites using LC-MS/MS. Talanta 134:165–172. https://doi.org/10.1016/j.talanta.2014.10.041
52. Tautenhahn R, Bottcher C, Neumann S (2008) Highly sensitive feature detection for high resolution LC/MS. BMC Bioinf 9. https://doi.org/10.1186/1471-2105-9-504
53. Zhang X, Quinn K, Cruickshank-Quinn C, Reisdorph R, Reisdorph N (2018) The application of ion mobility mass spectrometry to metabolomics. *Curr Opin Chem Biol*:60–66. https://doi.org/10.1016/j.cbpa.2017.11.001
54. Picache JA, Rose BS, Balinski A, Leaptrot KL, Sherrod SD, May JC, McLean JA (2019) Collision cross section compendium to annotate and predict multi-Omic compound identities. Chem Sci 10(4):983–993. https://doi.org/10.1039/c8sc04396e
55. Xiao JF, Zhou B, Ressom HW (2012) Metabolite identification and quantitation in LC-MS/MS-based metabolomics. TrAC Trends Anal Chem:1–14. https://doi.org/10.1016/j.trac.2011.08.009
56. Yost RA, Enke CG (1978) Selected ion fragmentation with a tandem quadrupole mass spectrometer. J Am Chem Soc 100(7):2274–2275. https://doi.org/10.1021/ja00475a072
57. Matraszek-Zuchowska I, Wozniak B, Posyniak A (2016) Comparison of the multiple reaction monitoring and enhanced product ion scan modes for confirmation of stilbenes in bovine urine samples using LC–MS/MS QTRAP® system. Chromatographia 79(15–16):1003–1012. https://doi.org/10.1007/s10337-016-3121-1
58. Hopfgartner G, Varesio E, Tschäppät V, Grivet C, Bourgogne E, Leuthold LA (2004) Triple quadrupole linear ion trap mass spectrometer for the analysis of small molecules and macromolecules. J Mass Spectrom 39(8):845–855. https://doi.org/10.1002/jms.659
59. Zhou J, Yin Y (2016) Strategies for large-scale targeted metabolomics quantification by liquid chromatography-mass spectrometry. Analyst 141(23):6362–6373. https://doi.org/10.1039/c6an01753c
60. Bateman NW, Goulding SP, Shulman NJ, Gadok AK, Szumlinski KK, MacCoss MJ, Wu CC (2014) Maximizing peptide identification events in proteomic workflows using Data-Dependent Acquisition (DDA). Mol Cell Proteomics 13(1):329–338. https://doi.org/10.1074/mcp.M112.026500
61. Mullard G, Allwood JW, Weber R, Brown M, Begley P, Hollywood KA, Jones M, Unwin RD, Bishop PN, Cooper GJS et al (2015) A new strategy for MS/MS data acquisition applying multiple data dependent experiments on orbitrap mass spectrometers in non-targeted metabolomic applications. Metabolomics 11(5):1068–1080. https://doi.org/10.1007/s11306-014-0763-6
62. Schwudke D, Liebisch G, Herzog R, Schmitz G, Shevchenko A (2007) Shotgun lipidomics by tandem mass spectrometry under data-dependent acquisition control. Methods Enzymol:175–191. https://doi.org/10.1016/S0076-6879(07)33010-3

63. Doerr A (2014) DIA mass spectrometry. Nat Methods 2014:121
64. Venable JD, Dong MQ, Wohlschlegel J, Dillin A, Yates JR (2004) Automated approach for quantitative analysis of complex peptide mixtures from tandem mass spectra. Nat Methods 1(1):39–45. https://doi.org/10.1038/nmeth705
65. Bilbao A, Varesio E, Luban J, Strambio-De-Castillia C, Hopfgartner G, Müller M, Lisacek F (2015) Processing strategies and software solutions for data-independent acquisition in mass spectrometry. Proteomics 15(5–6):964–980. https://doi.org/10.1002/pmic.201400323
66. Kondrat RW, McClusky GA, Cooks RG (1978) Multiple reaction monitoring in mass spectrometry/mass spectrometry for direct analysis of complex mixtures. Anal Chem 50(14):2017–2021. https://doi.org/10.1021/ac50036a020
67. Rathahao-Paris E, Alves S, Junot C, Tabet JC (2016) High resolution mass spectrometry for structural identification of metabolites in metabolomics. Metabolomics:1–15. https://doi.org/10.1007/s11306-015-0882-8
68. Michalski A, Damoc E, Lange O, Denisov E, Nolting D, Müller M, Viner R, Schwartz J, Remes P, Belford M et al (2012) Ultra high resolution linear ion trap orbitrap mass spectrometer (Orbitrap elite) facilitates top down LC MS/MS and versatile peptide fragmentation modes. Mol Cell Proteomics 11(3). https://doi.org/10.1074/mcp.O111.013698
69. Ghaste M, Mistrik R, Shulaev V (2016) Applications of Fourier transform ion cyclotron resonance (FT-ICR) and Orbitrap based high resolution mass spectrometry in metabolomics and Lipidomics. Int J Mol Sci. https://doi.org/10.3390/ijms17060816
70. Mahieu NG, Patti GJ (2017) Systems-level annotation of a metabolomics data set reduces 25 000 features to fewer than 1000 unique metabolites. Anal Chem 89(19):10397–10406. https://doi.org/10.1021/acs.analchem.7b02380
71. Blaženović I, Kind T, Ji J, Fiehn O (2018) Software tools and approaches for compound identification of LC-MS/MS data in metabolomics. Meta 8(2):31. https://doi.org/10.3390/metabo8020031
72. LCQuan. Thermo Fisher Scientific: Hemel Hempstead, Hetfordshire, UK
73. Depke T, Franke R, Brönstrup M (2017) Clustering of MS2 spectra using unsupervised methods to aid the identification of secondary metabolites from Pseudomonas Aeruginosa. J Chromatogr B 1071:19–28. https://doi.org/10.1016/j.jchromb.2017.06.002
74. MassHunter. Agilent: Santa Clara, CA, USA
75. Profile Analysis. Bruker
76. SIEVE. Thermo Fisher Scientific
77. Progenesis. Waters, Milford, MA, USA
78. Tautenhahn R, Patti GJ, Rinehart D, Siuzdak G (2012) XCMS online: A web-based platform to process untargeted metabolomic data. Anal Chem 84(11):5035–5039. https://doi.org/10.1021/ac300698c
79. Tsugawa H, Cajka T, Kind T, Ma Y, Higgins B, Ikeda K, Kanazawa M, VanderGheynst J, Fiehn O, Arita M (2015) MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis. Nat Methods 12(6):523–526. https://doi.org/10.1038/nmeth.3393
80. Katajamaa M, Miettinen J, Oresic M (2006) MZmine: toolbox for processing and visualization of mass spectrometry based molecular profile data. Bioinformatics 22(5):634–636. https://doi.org/10.1093/bioinformatics/btk039
81. MassBank of North America(2019). https://mona.fiehnlab.ucdavis.edu/. Accessed Sep 12, 2019
82. Wishart DS, Feunang YD, Marcu A, Guo AC, Liang K, Ázquez-Fresno RV, Sajed T, Johnson D, Li C, Karu N et al (2018) HMDB 4.0: the human metabolome database for 2018. Nucleic Acids Res 46. https://doi.org/10.1093/nar/gkx1089
83. NIST Standard Reference Database 1A v17|NIST (2019). https://www.nist.gov/srd/nist-standard-reference-database-1a-v17. Accessed Sep 12, 2019
84. Horai H, Arita M, Kanaya S, Nihei Y, Ikeda T, Suwa K, Ojima Y, Tanaka K, Tanaka S, Aoshima K et al (2010) MassBank: A public repository for sharing mass spectral data for life sciences. J Mass Spectrom 45(7):703–714. https://doi.org/10.1002/jms.1777

85. Jeffryes JG, Colastani RL, Elbadawi-Sidhu M, Kind T, Niehaus TD, Broadbelt LJ, Hanson AD, Fiehn O, Tyo KEJ, Henry CS (2015) MINEs: open access databases of computationally predicted enzyme promiscuity products for untargeted metabolomics. J Cheminform 7(1):44. https://doi.org/10.1186/s13321-015-0087-1
86. mzCloud – Statistics (2019). https://www.mzcloud.org/Stats. Accessed Sep 12, 2019
87. Huang X, Chen Y-J, Cho K, Nikolskiy I, Crawford PA, Patti GJX (2014) ^{13}CMS: global tracking of isotopic labels in untargeted metabolomics. Anal Chem 86(3):1632–1639. https://doi.org/10.1021/ac403384n
88. Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, Katayama T, Kawashima S, Okuda S, Tokimatsu T et al (2007) KEGG for linking genomes to life and the environment. Nucleic Acids Res 36(Database):D480–D484. https://doi.org/10.1093/nar/gkm882
89. Karp PD, Riley M, Paley SM, Pellegrini-Toole A (2002) The MetaCyc database. Nucleic Acids Res 30(1):59–61. https://doi.org/10.1093/nar/30.1.59
90. Pico AR, Kelder T, van Iersel MP, Hanspers K, Conklin BR, Evelo C (2008) WikiPathways: pathway editing for the people. PLoS Biol 6(7):e184. https://doi.org/10.1371/journal.pbio.0060184
91. Kleinridders A, Ferris HA, Reyzer ML, Rath M, Soto M, Manier ML, Spraggins J, Yang Z, Stanton RC, Caprioli RM et al (2018) Regional differences in brain glucose metabolism determined by imaging mass spectrometry. Mol Metab 12:113–121. https://doi.org/10.1016/j.molmet.2018.03.013
92. Aue WP, Bartholdi E, Ernst RR (1976) Two-dimensional spectroscopy. Application to nuclear magnetic resonance. J Chem Phys 64(5):2229–2246. https://doi.org/10.1063/1.432450
93. Martineau E, Dumez JN, Giraudeau P (2019., No. February, 1–14) Fast quantitative 2D NMR for metabolomics and lipidomics: a tutorial. Magn Reson Chem. https://doi.org/10.1002/mrc.4899