

Review

Metabolomics in the Context of Plant Natural Products Research: From Sample Preparation to Metabolite Analysis

Mohamed A. Salem ^{1,*}, Leonardo Perez de Souza ², Ahmed Serag ³, Alisdair R. Fernie ^{2,4},
Mohamed A. Farag ^{5,6}, Shahira M. Ezzat ^{5,7} and Saleh Alseekh ^{2,4,*}

¹ Department of Pharmacognosy, Faculty of Pharmacy, Menoufia University, Gamal Abd El Nasr st., Shibin Elkom, Menoufia 32511, Egypt

² Max Planck Institute of Molecular Plant Physiology, Am Mühlenberg 1, 14476 Potsdam-Golm, Germany; LPerez@mpimp-golm.mpg.de (L.P.d.S.); Fernie@mpimp-golm.mpg.de (A.R.F.)

³ Pharmaceutical Analytical Chemistry Department, Faculty of Pharmacy, Al-Azhar University, Cairo 11751, Egypt; Ahmedserag777@azhar.edu.eg

⁴ Center of Plant Systems Biology and Biotechnology (CPSBB), 4000 Plovdiv, Bulgaria

⁵ Pharmacognosy Department, Faculty of Pharmacy, Cairo University, Cairo 11562, Egypt; Mohamed.farag@pharma.cu.edu.eg (M.A.F.); shahira.ezzat@pharma.cu.edu.eg (S.M.E.)

⁶ Chemistry Department, School of Sciences & Engineering, The American University in Cairo, New Cairo 11835, Egypt

⁷ Department of Pharmacognosy, Faculty of Pharmacy, October University for Modern Sciences and Arts (MSA), Giza 11787, Egypt

* Correspondence: Correspondence: mohamed.salem@phrm.menofia.edu.eg (M.A.S.); Alseekh@mpimp-golm.mpg.de (S.A.); Tel.: +49-(0)331-567-8211 (S.A.)

Received: 25 November 2019; Accepted: 11 January 2020; Published: 15 January 2020

Abstract: Plant-derived natural products have long been considered a valuable source of lead compounds for drug development. Natural extracts are usually composed of hundreds to thousands of metabolites, whereby the bioactivity of natural extracts can be represented by synergism between several metabolites. However, isolating every single compound from a natural extract is not always possible due to the complex chemistry and presence of most secondary metabolites at very low levels. Metabolomics has emerged in recent years as an indispensable tool for the analysis of thousands of metabolites from crude natural extracts, leading to a paradigm shift in natural products drug research. Analytical methods such as mass spectrometry (MS) and nuclear magnetic resonance (NMR) are used to comprehensively annotate the constituents of plant natural products for screening, drug discovery as well as for quality control purposes such as those required for phytomedicine. In this review, the current advancements in plant sample preparation, sample measurements, and data analysis are presented alongside a few case studies of the successful applications of these processes in plant natural product drug discovery.

Keywords: metabolomics; plant natural products; drug discovery; metabolite extraction; liquid chromatography; gas chromatography; mass spectrometry; NMR

1. Introduction

Nature provides a rich source of numerous bioactive compounds that have been extensively employed in traditional medicine since time immemorial [1]. In recent years, the Food and Drug Administration (FDA) has approved an impressive number of modern drugs that are also natural products or directly derived therefrom [2]. These mostly constitute compounds belonging to secondary metabolic pathways with prominent examples including taxol from *Taxus brevifolia*, vinblastine from *Catharanthus roseus*, doxorubicin from *Streptomyces peuceitius*, and cyclosporine from

Tolypocladium inflatum [3,4]. The first step in the discovery of lead compounds from natural sources is the release of bioactive metabolites from their biomass through various extraction techniques viz., supercritical fluid extraction [5,6], microwave-assisted and ultrasonic-assisted extraction [7], molecular distillation methods [8], and membrane separation technology [9]. Moreover, bioassay-guided fractionation using chromatographic methods such as preparative high performance liquid chromatography (HPLC) is applied for the isolation and purification of active metabolites from their crude extracts [10]. Further technologies such as nuclear magnetic resonance spectroscopy (NMR), mass spectrometry (MS), and ultraviolet-visible spectroscopy (UV-Vis) have allowed the detailed characterization and ultimately, the structural elucidation of these agents [11]. Finally, the bioactivity (effects in cell lines, animal models, and human volunteers) is investigated for assessing the pharmacological potential of the candidate compounds. Nevertheless, some pitfalls were observed when employing this classical approach for lead compound discovery, where degradation or chemical modification of the bioactive compounds during the process of isolation and purification often occurs. Furthermore, important biological information that was present in the original extract might be lost during activity-guided fractionation as the samples are not fully analyzed [12]. Moreover, better therapeutic effects were reported when using the whole extracts, as practiced in traditional medicine, rather than a single-compound based remedy. This effect could be attributed to the synergy between bioactive components (e.g., studies showed the synergistic effects of different plant extracts and doxorubicin in cancer treatment [13], Apocynaceae plants and antibiotics against *Acinetobacter baumannii* [14], and catechin and resveratrol as antioxidants [15]).

Unlike the classical approach in natural products research, metabolomics experiments offer an improved expedited route for drug discovery [16]. The basic goal of metabolomics is to provide a comprehensive qualitative and/or quantitative analysis of all metabolites present in a living system [17]. Interestingly, this concept could be extended in natural product drug discovery via studying the relationship between the whole metabolome of natural-derived remedies and their biological effects [18]. Implementing such an approach not only overcomes the aforementioned pitfalls of the classical techniques used in natural product research, but also provides a broader insight of the biochemical status and gene functions of the studied organisms. Furthermore, the signature between specific compounds in the metabolome and their bioactivity could also be revealed, aided by advanced bioinformatics tools, a process that can be highly useful in pharmacological standardization and biological fingerprinting of natural extracts [19]. The workflow of metabolomics experiments involves an efficient extraction of these endogenous metabolites to be subjected for qualitative and quantitative analysis. However, unlike other omics technologies, no single analytical platform is capable of analyzing all metabolites simultaneously due to their extreme complexity and huge chemical diversity. Recent developments in analytical chemistry platforms such as hyphenating mass spectrometry with gas chromatography (GC), liquid chromatography (LC) or capillary electrophoresis (CE), and nuclear magnetic resonance (NMR) spectroscopy have led to a highly efficient set up for metabolome analysis [20], however, these do not yet reach comprehensibility [21]. Although huge datasets are generated from these instruments, the evolution of chemometrics and multivariate data analysis algorithms provides a powerful tool for extracting useful information from such high dimensionality results [22]. They enable, for example, the detection of compounds that correlate to the medicinal efficacy in test animal or human systems based on their analytical spectral fingerprints [23]. Moreover, pattern recognition and classification algorithms have also allowed the implementation of metabolomics as an effective tool for the quality control of herbal medicinal products [24,25]. Nevertheless, metabolite identification remains the most challenging aspect of metabolomics experiments [26]. Large mass spectral and NMR spectral databases have been created to untangle such problems [27]. Furthermore, bioinformatics tools based on molecular networking such as GNPS [28] and MetGem [29] have been implemented not only to assign known metabolites from their complex mixtures, but also to elucidate the chemical structures of novel compounds of interest.

This review will discuss the recent developments of metabolomics in the context of plant natural product drug discovery including current advances in sample preparation techniques and analytical

profiling platforms. In addition, computational tools employed for metabolomics data processing and metabolites identification are reviewed. Finally, some successful applications of metabolomics in the identification of bioactive agents and in the quality control of natural products are outlined, and an outlook for the increasing use of metabolomics in these fields is provided.

2. Sample Preparation for Metabolomic Studies

In metabolomics studies, biological samples are collected, extracted, measured, and finally, the resulting data are analyzed (Figure 1). Sample preparation is a crucial step in metabolomics as it greatly affects the reliability of the metabolomics results. Minor changes in the sample collection, extraction, or storage greatly affect metabolite stability and hence can lead to major changes in the observed metabolome. Metabolomics samples have to be collected uniformly and rapidly to avoid changes due to the fast enzymatic turnover rate [30]. The ultimate aim is to minimize the biologically-irrelevant changes resulting from sample processing. Improper handling of biological samples is the most likely source of bias in metabolomic studies [30]. In order to validate plant metabolomics studies, the minimum parameters related to experimental design, sample extraction to data analysis should ideally follow the Metabolomics Standards Initiative (MSI) [31].

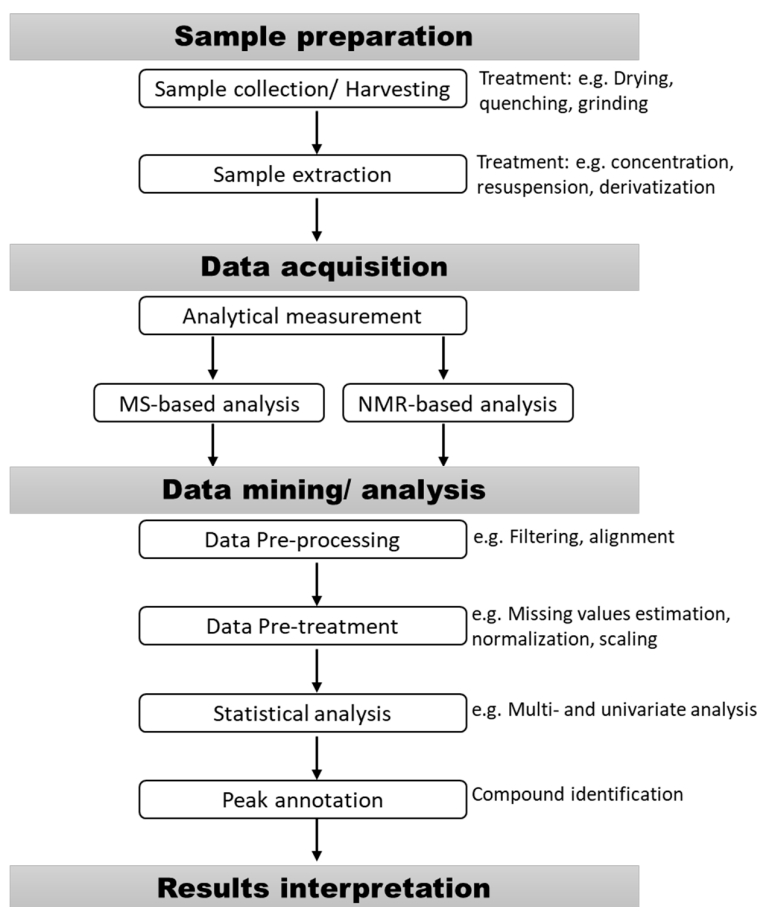


Figure 1. Key steps of a plant metabolomics study. Additional steps can be performed or modified in certain approaches based on the analytical methods or point of interest.

2.1. Sample Collection

A wide range of biological samples have been extensively reported in metabolomics studies. This includes tissues (animal-and plant-derived tissues), fluids (such as urine, whole blood, serum, sweat, cerebrospinal fluid, breast milk, amniotic fluid, saliva, etc.), and cell cultures (human, animal,

plant, algae, etc.) [32,33]. Such complex samples have different matrices, and therefore, require different sample preparation protocols. Here, we focus on natural products that are mainly derived from plant origin, while also giving a general overview of other tissues. Depending on the natural abundance of some metabolites and the level of detection, various amounts of biological material per sample have to be considered. Usually, 1–100 mg of tissue, 10–250 μ L of fluids, or 10^5 – 10^7 cells per each biological replicate is required [33]. A minimum number of 3–5 biological replicates is recommended for each condition in metabolomic studies [31]. Biological replicates (parallel measurements of samples from different individuals) rather than technical replicates (repeated measurements of the same sample) are to be considered [34].

2.2. Harvesting Methods

Sample freezing methods such as using dry ice or liquid nitrogen are highly recommended during the harvesting of fresh samples to avoid enzyme-induced metabolic changes [31]. Removing unwanted components such as soil particles is also recommended before collection. Long-term storage of samples prior to extraction should, however, be avoided. For short term storage (e.g., for few days up to two weeks), samples can be kept in liquid nitrogen, dry ice, or a -80 °C freezer [31]. Prior to extraction, the harvested samples can be exposed to processing methods such as lyophilization, cell lysis, and/or grinding, depending on the biological material. The conditions related to cultivation parameters, collected tissue type, seasonality, developmental stage, harvesting time, and sample processing should be reported for each condition since metabolites are greatly affected by such parameters [35], with environmental aspects being reported to induce both qualitative and quantitative variations in the metabolite composition of both plant primary and secondary metabolites [36].

2.3. Sample Extraction

Unlike the genome and proteome, which can be captured using a single extraction protocol, the metabolome is difficult to capture within a single solvent due to the diverse chemistry of metabolites [37]. Therefore, complex biological samples, which as stated above, have different matrices, and therefore, require different extraction protocols. Generally, metabolites are preferentially extracted with the solvent following the rule of thumb “like dissolves like”. Polar and semi-polar metabolites can be extracted with hydrophilic solvents such as hydro-alcoholic solutions, while, lipids can be extracted with more hydrophobic solvents. Several protocols for metabolome extraction have been developed and extensively reviewed [32,33,35,38–40]. The selected method must be rapid and efficient, whilst at the same time, cover a wide range of target metabolites, thus maintaining a high level of precision and accuracy.

Liquid–liquid fractionation provides significant simplifying steps compared to single extraction methods [41]. Partial purification of each fraction is achieved via removing interfering compounds such as hydrophobic molecules that are enriched in lipophilic solvents and more polar metabolites in hydrophilic solvents [42]. While different extraction protocols are varied with regard to the selected solvent(s) and the amount that is to be used, most protocols include a deproteinization step. The presence of proteins in the analyzed samples can severely affect the instrument’s accuracy, precision, and lifetime [43]. Extraction efficiency can usually be enhanced by selecting an extraction solvent mixture consisting of one or two steps compared to single extraction solvent [33]. Liquid–liquid extraction methods were based on the so called gold standard extraction protocols of the ‘Folch’ and ‘Bligh and Dyer’ methods utilizing chloroform/methanol mixture in different proportions [44,45]. Liquid–liquid extraction using chloroform/methanol/water have been used for the analysis of lipids and polar metabolites [46–49]. As a cleaner and safer solvent, methyl *tert*-butyl ether (MTBE) has also been used as an alternative to chloroform for liquid–liquid extraction [50], particularly for the recovery of metabolites and lipids from bacterial, plant, algal, flies, and diverse mammalian samples [41,42,51–60]. The use of isotope-labeled internal standards during sample preparation is recommended to assess matrix effects [39,61]. For the selection of suitable solvent(s) and an extraction

method that is suitable for different metabolite classes, readers are referred to the available extraction protocol reviews [32,35,39,40,62,63].

2.4. Extract Concentration, Dilution, Enrichment, and Re-Suspension

In addition to liquid–liquid extraction, other methods such as organic solvent precipitation as well as single step or multiple step solid-phase extraction (SPE) have also been used to partially purify samples prior to analytical measurement [64]. After sample extraction, solvents are evaporated to concentrate the metabolites. Samples are typically concentrated in a vacuum concentrator at room temperature without heating. The use of a nitrogen flow evaporator is recommended in cases where metabolites are sensitive to oxidative modifications as, for example, some lipids. Finally, the dry extracts are typically resuspended in an analysis-compatible solvent [43]. Dried samples can be subjected to further steps such as derivatization before GC/MS analysis [49]. Short term storage of extracted liquid samples, even at low temperature (−20 °C), is not recommended, however, samples can be stored, if necessary, in a dry state for a short time prior to analysis [65].

3. Analytical Methods for Metabolites Analysis

The diversity of metabolites in the plant kingdom is staggering; a commonly quoted estimate is that plants produce somewhere between 100,000 and one million metabolites [66,67]. Metabolite measurements have been carried out for decades because of the fundamental regulatory importance of metabolites as components of biochemical pathways, the importance of certain metabolites in the human diet, and their use as diagnostic markers for a wide range of biological conditions including disease and response to chemical treatment [37]. There are wide ranges of metabolomics approaches, which ultimately aim to measure the entire small molecule complement of the cell. Current metabolomics strategies are mainly reliant on four major approaches: gas chromatography–mass spectrometry (GC-MS), liquid chromatography–mass spectrometry (LC-MS), capillary electrophoresis–mass spectrometry (CE-MS), and nuclear magnetic resonance (NMR) spectroscopy. A number of detailed protocols [49,63,68,69], in addition to several excellent technical reviews [70,71] regarding the utilization of these analytical tools in metabolomics experiments, have been published, however, a brief technical overview of these four major methods is provided herein.

3.1. Gas Chromatography–Mass Spectrometry (GC-MS)

GC-MS has been one of the most popular metabolomics techniques to measure the levels of volatile and semi-volatile organic compounds in a wide variety of samples. In GC-MS, polar metabolites are derivitized to render them volatile and then separated by GC. Various derivatization methods such as alkylation, acylation, methoximation, trimethylsilylation, and silylation can be used [72,73]. Electron impact ionization results in highly reproducible fragmentation patterns that are essential for large-scale experiments [37]. A great advantage of GC-MS is that it is both relatively sensitive and highly robust, and can routinely and reproducibly measure hundreds of analytes across thousands of samples [72,74–76]. That said, a technical challenge in GC-MS profiling is to separate each metabolite signal from overlapping peaks in the raw GC-MS chromatogram [77]. Several peak-picking software packages equipped with sophisticated peak deconvolution functions are available to deal with a relatively high-throughput data of thousands of samples [77]. In addition, various databases have been developed to aid in assigning structures to spectral peaks observed in metabolomics experiments [78–80]. Furthermore, limitations of GC-MS have recently been improved by the development of two-dimensional GC × GC coupled with high resolution mass spectrometry [81–84]. In this approach, two columns with different properties (non-polar vs. polar) are connected through a modulator, allowing further separation of compounds that co-elute from the first column, thereby giving rise to enhanced resolution and peak capacity [85,86]. The GC-MS method has had myriad applications in plant, pharmacological, and medical metabolomics studies, and is considered as one of most suitable techniques of the accurate determination of primary metabolites, however, it is severely compromised in measuring some secondary metabolites. Several classes of volatile and

non-volatile metabolites such as phenolics, alkaloids, and terpenoids have been analyzed and identified by GC-MS [79,87].

3.2. Liquid Chromatography-Mass Spectrometry (LC-MS)

Liquid chromatography-mass spectrometry (LC-MS) has become the most comprehensive technique to measure a wide range of diverse metabolites. Unlike GC-MS, it does not require prior sample treatment, and crude extracts obtained by simple extraction can be introduced directly to the LC-MS. The choice of columns including reversed phase, ion exchange, and hydrophobic interaction provides metabolite separation on the basis of differential chemical properties. Nowadays, reversed-phase columns such as C18 or C8 are the most commonly used for LC gradient separation [63]. The development of ultra-performance LC rendered the technique even more powerful with regard to resolution, sensitivity, and throughput. LC-MS is a unique method for measuring plant secondary metabolites such as flavonoids and alkaloids [88–90], membrane lipids (lipidomics) [60,91,92], and primary metabolites such as amino acids [93]. Despite the fact that LC-MS is the most comprehensive technique in hand, LC-MS by no means approaches the metabolic complement of a typical plant cell [21]. This is mainly due to the poor availability of standard compounds for secondary metabolites [94]. In the past years, a considerable improvement in the number of metabolites that can be measured and annotated has been realized, which is mainly due to the improved machine performance afforded by the development of Ultra-Performance Liquid Chromatography (UPLC) coupled with high-resolution mass analysis methods such as time-of-flight (TOF) MS, Fourier transform (FT) MS, and Orbitrap-based MS [88,90,95,96]. In addition, this improvement has relied on increased efforts in the collection of standard compounds and sharing of reference extracts [97–99] for use in peak annotation authentication and by the increased sophistication of computational approaches for compound annotation (see the web resources listed in [27,100]).

3.3. Capillary Electrophoresis-Mass Spectrometry (CE-MS)

Capillary electrophoresis (CE) separates polar and charged compounds on the basis of their charge-to-mass ratio [74]. CE offers fast and high-resolution separation of charged analytes from small injection volumes. Coupled to mass spectrometry (MS), it represents a powerful analytical technique providing (exact) mass information and enables molecular characterization based on fragmentation. Although exquisitely sensitive and able to readily capture metabolite classes that the other platforms do not, including nucleotides and highly charged metabolites [101], CE is time consuming and is further hampered by the fact that it covers a range of metabolites that have highly diverse extraction requirements. Therefore, the use of CE-MS in plant metabolism remains relatively rare [74,102]. In addition, CE has poor migration time reproducibility and a lack of reference libraries, which may only be partially overcome by the prediction of migration time [103,104]. However, CE has some distinct advantages over other instruments employed for metabolomics: mainly the facts that it utilizes low separation volume, which is particularly suitable for the study of biological fluids in small experimental animals, and operates under a homogeneous separation environment [105].

3.4. Nuclear Magnetic Resonance (NMR)

NMR spectroscopy is a powerful analytical tool that has traditionally occupied a vital position for structure elucidation of natural products [106–110]. Unlike MS, NMR is not discriminatory and is less biased as the results does not rely on the type of ionization condition or the preferences of the used instruments [111]. Hence, this technique allows for the simultaneous detection of the abundant primary metabolites (organic acids, amino acids, and sugars) alongside secondary metabolites (alkaloids, terpenoids, and flavonoids) as typically found in plant natural extracts [112]. In addition, NMR is a very useful technique for the structure elucidation of novel and/or unexpected compounds including those with identical masses and/or different isotopomer distributions [12,113]. Moreover, NMR-based methods are highly reproducible, noninvasive, nondestructive, and require minimal sample preparation as the sample does not get in physical contact with the device as in MS [114].

Another major strength in NMR spectroscopy lies in the direct proportionality between the NMR spectrum signals and the corresponding real molar levels of the detected metabolites, making absolute quantification of all detected metabolites possible without the need for calibration curves of individual analytes, posing it as a powerful tool in quality control purposes of drug extracts [115]. However, a five times delay of the spin–lattice relaxation times (T_1) of the slowest relaxing nuclei in the extract is essential in such quantitative experiments to avoid baseline distortions and increase the accuracy of signal integration. This will significantly increase the duration of these experiments compared to other routine NMR metabolomics methods.

Nevertheless, NMR also has a number of disadvantages in metabolomics analysis [116]. The technique's major disadvantage is the low sensitivity as the polarization rate of the active NMR nuclei obeys the Boltzmann distribution law. However, recent developments in NMR hardware have at least partially improved this drawback to some extent. The use of superconducting magnets that are commonly operated in metabolomics experiments at field resonance of 600 MHz can increase the sensitivity and resolution of the detected signals [117]. Moreover, the development of cryogenic probes in which the electronics are cooled down to a very low temperature (e.g., 20 K) has resulted in a better sensitivity than standard probes via keeping the noise generated due to random thermal motion of electrons at a minimum [118]. There are currently two types of cryoprobe available, the first is cooled by a closed cycle helium cooler system, which yields a signal-to-noise enhancement up to a factor of five. The second type is a liquid nitrogen-cooled system, commonly used in metabolomics experiments, that is able to provide a sensitivity enhancement of up to a factor of three, but with the capability to be cooled down and warmed up relatively quickly [119]. However, such probes may be of limited relevance for samples with high ionic and/or dielectric conductivity such as those with salts or polar solvents, commonly encountered in plant-derived samples [120]. That said, the low concentrations of natural products remains a limitation of NMR analysis. To partially compensate for this, special probes known as microprobes have been also developed in which the analysis of a few microliters of samples could be possible, thus overcoming the limited availability of the samples [121] as is typical in natural product studies. Dynamic nuclear polarization (DNP) is another interesting advancement that has also been employed to enhance the sensitivity of NMR signals. The fundamental principle behind DNP involves transferring the polarization of the electron spins to the nuclei of interest via microwave irradiation close to the electron Larmor frequency, resulting in a temporary hyperpolarization in this spin-active nuclei (usually naturally abundant ^{13}C , ^{15}N , and ^1H). After the polarization transfer step is completed, the sample is transferred to the NMR spectrometer to collect the enhanced (>1000-fold) NMR signals [122].

3.4.1. 2D NMR Based Metabolomics Strategies

Although NMR methods have largely been employed primarily to identify compounds from pure samples as in classical natural product research, they have recently played a larger role in the analysis of mixtures of unfractionated natural extracts [123]. The bottleneck in such complicated analysis is the congregation of the ^1H NMR spectra, which could be untangled using two main strategies. The first strategy aims to simplify the abundant ^1H NMR spectra through the use of diffusion and/or relaxation filters or through creating a projection of the ^1H decoupled spectrum like in 2D ^1H J-resolved NMR experiments. Relaxation filters are pulse sequences that selectively attenuate, or even remove, signals of components with shorter T_1 or spin–spin relaxation times, T_2 [124]. Hence, they could be utilized to simplify the ^1H NMR spectra through selective elimination of signals belonging to high molecular weight or rigid structure (quickly relaxing) compounds. They have also been implemented to attenuate water signals using the water attenuation T_2 relaxation Carr–Purcell–Meiboom–Gill (WATR-CPMG) pulse sequence [125]. On the other hand, diffusion filter based experiments such as diffusion ordered spectroscopy (DOSY) commonly simplify ^1H NMR spectra based on the difference of the translational diffusion coefficients between molecules of different molecular size, achieving what has been described as “NMR chromatography” [126]. The DOSY experiments can be performed in 1D mode, where two or more ^1H NMR experiments are acquired for compounds possessing different diffusion coefficients or in 2D mode (2D DOSY), where

the proton spectrum is depicted in one dimension and the diffusion coefficient is spread in an orthogonal dimension [127]. Interestingly, the extension of 2D DOSY to 3D DOSY is gaining much popularity in metabolomics analysis via incorporating another proton dimension as in 3D COSY-DOSY and 3D TOCSY-DOSY, or another carbon dimension as in 3D DOSY-HSQC [128]. Such DOSY methods have been recently applied in quality control and metabolic fingerprinting of commercial natural extracts [128,129]. Furthermore, they have also been used to identify novel metabolites from natural resources, for example, novel bromopyrrole alkaloids, and agesamides A and B were identified in the crude bromopyrrole fractions from Okinawan marine sponges using the 2D DOSY method [130]. 2D ^1H J-resolved NMR experiments represent another approach that simplify ^1H NMR spectra where the overlapping resonance of the ^1H dimension is resolved through representing the coupling constant of each signal (J value) in a second dimension [131]. 2D ^1H J-resolved NMR has been widely used in metabolomics classifications studies [24,132], however, prior knowledge of the chemical composition of the studied samples is required, which make this experiment of limited utility in the analysis of novel compound discovery [108]. On the other hand, a newer approach called pure-shift ^1H NMR spectroscopy has emerged in which better removal of homonuclear J couplings has been achieved [133]. In this experiment, J coupling multiplicities are collapsed into single lines via implementing specific pulse sequences that result in a broadband homo-decoupled spectrum. This method has been applied recently to investigate the metabolome of *Physalis peruviana* fruit aqueous extracts [134]. The simplified data allowed for the identification of glutamic acid, a metabolite not observed in previous studies of the same extract due to the heavy overlap of its NMR signals.

The second strategy that can reliably handle ^1H NMR overlapping peaks is the spreading of the crowded ^1H NMR peaks in a second frequency dimension via magnetization transfer [135]. This transfer could be between nuclei of the same type as in H–H correlated spectroscopy (COSY) and H–H total correlated spectroscopy (TOCSY), nuclei of different types such as heteronuclear single quantum coherence spectroscopy (HSQC) and heteronuclear multiple bond correlation (HMBC) experiments, or through space as in nuclear overhauser effect spectroscopy (NOESY) and rotating frame overhauser enhancement spectroscopy (ROESY) experiments [135]. Full details of these common 2D NMR methods that have been used extensively in metabolomics analysis of natural products, and are thus beyond the scope of this review, but can be found elsewhere [12,108,135,136]. It is noteworthy to mention that spreading the crowded signals in indirect proton dimensions such as in COSY and TOCSY results in a relatively short acquisition time due to the high natural abundance of protons, however, some overlap might still exist as a region of only 10 ppm is employed. In contrast, using carbon as a second dimension such as in HSQC and HMBC will require a much longer acquisition time, but allow the resolution of the overlapping signals and the indirect dimension to be increased to more than 200 ppm units. Nevertheless, it should be noted that acquiring data using HSQC or HMBC can aid in the detection of less abundant compounds compared to the direct acquisition from the carbon channel as it is typical in ^{13}C -NMR. It is worth mentioning that different pulse sequences have been developed for quantitative HSQC experiments. For example, a new phase modulated pulse sequence called the quantitative, offset-compensated, CPMG-adjusted HSQC (Q-OCCAHSQC) experiment has been proposed to enable the use of HSQC spectra for the quantitative determination of many primary metabolites [137]. However, a combination of different 2D NMR experiments is favored for better analytical performance and enhanced identification strategies (e.g., COSY, (1H,13C)-HMQC, (1H,13C)-HMBC, and NOESY) were used to analyze the unfractionated defensive secretion of the walking stick *Parectatosoma mocquersyi* identifying the novel monoterpene parectadiol [138]. COSY, HMQC, and HMBC have also aided in the identification of six terpenoids/sterol secondary metabolites in the crude extract of *Sarcophyton* marine coral [139]. Similarly, a combination of ^1H , ^1H COSY, HSQC, and HMBC have identified and quantified structurally related α - and β -bitter acids and their degradation products in hop resin [140] using run times comparable to those used in HPLC. Interestingly, in the same study, HMBC spectra coupled with multivariate techniques were utilized as a novel approach for the classification of 13 commercial hop cultivars. Application of 2D-NMR metabolomics for drug discovery is considerably

easier in cases where the secondary metabolites are naturally abundant as typical in the case of hop resin, which contains much lower levels of primary metabolites [54]. In an interesting application of 2D NMR methods, differential analyses of NMR spectra (DANS) has been oriented for natural products drug discovery and comparative metabolomics. DANS uses a simple algorithm that graphically compares the 2D NMR spectra of different biological states to spot the peaks that are discriminatory among this set of spectra. For example, 2D double quantum filtered (DQF)-COSY was employed for DANS analysis of seven different culturing protocols of the filamentous fungus, *Tolypocladium cylindrosporum*, leading to the detection and identification of two novel indole alkaloids, TC-705A and TC-705B, in the unfractionated extracts [141]. Furthermore, DQF-COSY experiments have been applied for the identification of signaling molecules in the model organism *Caenorhabditis elegans* [142] and polyene antibiotic “bacillaene” from *Bacillus subtilis* [143]. On the other hand, multidimensional NMR experiments have already been used to increase the resolution of the newly developed benchtop NMR instruments [144]. Benchtop NMR provides inexpensive and direct NMR access within any metabolomics laboratory, however, they suffer from decreased chemical shift dispersion and peaks overlapping as they are typically operated in the range of 40–80 MHz proton resonance frequency [145]. Benchtop NMR maintain the same J-coupling frequency as the high field instruments since it is independent of the Larmor frequency, making it suitable for some metabolomics and quality control applications, particularly in 2D modes [146].

3.4.2. Solid State NMR Based Metabolomics

Another exciting application of NMR spectroscopy is the analysis of semi-solid samples such as fresh plant leaf or intact tissues via the implementation of high resolution solid state magic angle spinning (HR-MAS) NMR [147]. In this technique, the sample is mixed with a minimal volume of solvent. Next, an NMR spectrum is acquired using a special HR-MAS probe rotated at the angle of 54.74° (the ‘magic angle’) and a spinning rate of 4 kHz. HR-MAS NMR provides spectra with a similar resolution to that of classical liquid-state NMR techniques due to the elimination of chemical anisotropies and dipolar coupling that hamper analysis of the solid states samples. This technique requires few sample preparation steps, thus it is gaining much attention in the quality control of herbal medicine and food analysis [148,149]. HR-MAS NMR has been also reported for the metabolic profiling of red algae [150], carrot [151], and Arabidopsis [152] tissues. However, the method is not fully quantitative as the local microenvironment in tissues or cells (e.g., pH) could affect the metabolite chemical shifts, thus hindering its potential, as does the difficulty in incorporating a reference compound [153].

3.4.3. Hyphenated NMR-Based Metabolomics

Whereas standalone NMR spectroscopy has unquestionable merits, on-line hyphenation of separation techniques and NMR spectroscopy (HPLC-NMR) can provide another powerful workflow for de novo identification of novel natural products in crude extracts [154,155]. In hyphenated HPLC-NMR, a flow cell matching the HPLC module is used instead of conventional NMR probes [156]. The separated analytes are eluted from the HPLC column to the crossed flow cell, and their on-line— ^1H NMR spectra are recorded. Some drawbacks were observed in this setup, particularly in the case of applying a gradient mobile phase [157]. Hence, most instrumental setups have changed from on-flow cell to stopped flow HPLC-NMR, and further to the loop/cartridge storage devices that showed wider application in natural product chemistry [158,159] and secondary metabolite identification [160]. Another major development in hyphenated NMR techniques is high-performance liquid chromatography-solid phase extraction-NMR spectroscopy (HPLC-SPE-NMR) where a post-column SPE cartridge is used as an analyte enrichment device prior to NMR acquisition [161]. In this setup, the HPLC eluted component is allowed to be captured onto a solid adsorbent, allowing the HPLC mobile phase to be removed. The adsorbed analyte is then transferred to the NMR probe using the minimal volume of solvents with sufficient elution power (e.g., CD_3OD , CD_3CN). The decrease in analyte chromatographic peak volume via the use of minimal solvent confers increased sensitivity of this technique over conventional HPLC-NMR, in addition to

significant cost reduction since employing a deuterated HPLC mobile phase is no longer mandatory [162]. However, careful optimization of SPE trapping is necessary from an analytical point of view, as some metabolites may not be trapped onto the SPE stationary phase. In addition, the use of a small volume of deuterated NMR solvents for flushing the required analyte from the SPE stationary phase is a critical parameter that needs to be carefully optimized. HPLC-SPE-NMR has been applied in the structural characterization of secondary metabolites belonging to different classes including flavonoids [163], terpenoids [164], steroids [165], aromatic alkaloids [166], diarylheptanoids [167], iridoids [168], and saponins [169] and has great potential in fields such as chemical ecology in which targeted metabolites are often signaling molecules present at low levels [170]. Furthermore, multiple complementary hyphenated approaches including PDA and MS have been integrated with HPLC-SPE-NMR as versatile platforms for authentication and the structural identification of secondary metabolites directly from extracts [171,172]. The sample is split after the initial HPLC separation with a small portion being routed for PDA and MS analysis and the remainder to the SPE cartridges for collection [173]. HPLC-PDA-MS-SPE-NMR has been used for the dereplication of six novel quinolinone alkaloids, named haplacutine A–F from the *Haplophyllum acutifolium* crude extract [174] and 23 coumarins including six new compounds from the crude ethyl acetate extract of *Coleonema album* leaves [175] and even minor metabolites could also be detected and fully characterized using this advanced platform [176]. These developments thus suggest that NMR, despite its lower sensitivity will retain a role in drug discovery in the decades to come.

4. Metabolomic Data Processing and Interpretation

GC-MS, LC-MS, and NMR are arguably the most relevant techniques within the context of natural product discovery. GC-MS is a robust platform with great peak resolution in the chromatographic dimension, capable of providing both stable separation and mass spectra fragmentation via its most common ionization source, namely electron ionization (EI). This renders data processing as well as metabolite annotation relatively facile for GC-MS. Well established methods for obtaining robust retention indexes based on series of standards [49], together with the great stability of EI spectra, result in an outstanding machine independent reproducibility of the results obtained and facilitate the use of standard databases for the identification of analytes such as the Golm Metabolome Database (GMD) and MassBank of North America (MoNA), among others [27]. The high resolution of the chromatographic separation and stability of ionization also facilitates the deconvolution of mass features into compound mass spectra and several relatively user-friendly tools are available that are capable of performing all steps from the raw data to peak area and assessment and metabolite annotation matching against databases [177–179]. However, the portion of the metabolome that can be covered by GC is significantly limited by the necessary properties of its analytes, namely, being volatile and stable at the high temperatures of analysis. Several compounds can be derivatized to fit such properties, this is unfortunately not the case for most natural products. Moreover, compound annotation relies almost exclusively in matching against known compounds in databases. De novo annotation of compounds based solely on an analyte mass spectra and chromatographic behavior is uncommon. This is because the interpretation of EI spectra is very complex due to the high degree of fragmentation, which can be partially mitigated by the adoption of a milder ionization source such as chemical ionization. Finally, GC has a disadvantage in relation to LC as it cannot be easily translated into a purification technique for further characterization of unknowns by more powerful structural techniques such as NMR. Still, a few interesting attempts have been made in expanding GC-MS structural elucidation such as the work by Matsuo et al. [180]. Here, the authors achieved considerable improvements over traditional spectra and retention index matching approaches by integrating multiple cheminformatics procedures including the use of quality controls to reduce mass spectra background noise and remove artifact peaks, principal component analysis for the selection of biologically relevant variables, EI-MS spectral search in databases, and both retention index filtering and predictions [180]. A further peculiarity of GC-MS is the lack of linearity with respect to quantitation across peaks, requiring quantitative analysis to be made on an analyte by analyte basis [181,182].

LC-MS is the most versatile technique in terms of metabolome coverage and certainly the most relevant for rapid dereplication of natural products in complex mixtures. Its high sensitivity combined with the multitude of stationary phases with different chemistries and ionization sources covering a broad spectra of different compounds allows it to be optimized for nearly every class of natural product. The greatest challenges faced by LC-MS are related to the susceptibility of the results to the many factors influencing separation and ionization efficiency such as changes in solvent composition, column stability, electrospray formation, ion suppression, and other so-called matrix effects. These factors negatively affect spectra and retention reproducibility. The resulting shifts of compound retention times across multiple runs are hard to predict or account for with procedures such as the calculation of retention indexes for GC. The lack of reproducibility in both retention and spectra make the identification of known compounds based on database matches significantly less efficient than for GC. Data processing is also a more challenging task since the lack of ionization stability combined with the lower resolution of LC separation complicates combining mass features into compound spectra, resulting in much more complex datasets with an overwhelming amount of mass features. Despite all of these challenges, the great flexibility of LC-MS can lead to intensive developments for this technique within the context of metabolomics and several pipelines are available to process, analyze, and interpret LC-MS based metabolomics data [32].

Considering that the focus of this review is on the discovery of plant natural products, we mainly describe the processing of untargeted metabolomics experiments, defined as the unbiased processing of all signals within a dataset. Several excellent freeware are currently available for the processing of mass chromatograms with some popular options including XCMS [183], MZmine 2 [184], OpenMS [185], and MS-DIAL [179], all of which have been extensively used for diverse sets of metabolomics data. All of these software work on similar principles, identifying m/z signals above a certain threshold as well as the boundaries confining the chromatographic peak representing these signals over time (often referred as features) and returning the respective peak areas/heights. After peak detection, they are usually aligned across multiple samples prior to data analysis. One of the first challenges in processing the resulting data matrixes is the sheer number of features resulting from this process, usually in the order of tens of thousands, much of which are uninformative or redundant, being attributed, for example, to background noise, contaminations and in source fragmentation. Several practices are commonly included in experimental design to assist in the elimination of these signals in the initial steps of data analysis. The inclusion of extraction blanks and technical replicates of quality controls are good examples of such practices and have been shown to be useful for the identification of features that do not originate from the samples or that exhibit low reproducibility [186]. This data cleaning process can significantly facilitate further steps of data analysis, particularly for the identification of features that are most relevant for the separation of experimental groups, for instance, in activity guided fractionation of complex extracts [187]. Another challenging and essential aspect of processing untargeted metabolomics experiments is the identification of features representing fragments, adducts, and isotopes of a same compound. As mentioned before, this process is much more challenging for LC-MS in relation to GC-MS, however, there are a few tools such as CAMERA [188], which can assign putative identity for features based on their retention time and correlation in relation to one another as well as mass differences matching common adducts and in source fragmentation patterns in electrospray ionization.

Particularly relevant for the identification of compounds by mass spectrometry are the experiments including tandem MS fragmentation of either specific ions in a data-dependent acquisition mode (DDA), or of all ions in a data-independent acquisition mode (DIA). Both types of experiment provide second order MS spectra (MS/MS), which are more informative and suitable for compound identification based on matching spectra signals against databases. In DDA experiments, specific ions are isolated and fragmented, therefore all signals in the MS/MS spectra correspond to fragments of the isolated ions. In DIA experiments, larger isolation windows are combined to fragment all ions in first order spectra, therefore there is no link between parental and daughter ions but with the advantage of a comprehensive fragmentation. Several specialized tools as well as extensions to the tools previously mentioned were developed to process such MS/MS data, a few

examples including MetDIA [189], which is based on algorithms from XCMS, and MS-DIAL [179], which was developed with a particular focus on processing DIA data despite also being able to process DDA and GC-MS data.

MS/MS spectra are currently the starting point for most mass spectrometry based structural elucidation approaches. The initial procedure usually consists of using the experimental data to search matches through different databases. Several relevant databases for natural product discovery were recently reviewed by Wolfender et al. [190]. We highlight here some of the open source alternatives that are particularly interesting for including curated experimental MS/MS data. Some extensive metabolome databases such as Metlin [191], MassBank [192], MoNA, European MassBank, Global Natural Products Social Molecular Networking resource (GNPS) [28], and Human Metabolome Database (HMDB) [193] include experimental MS/MS data from a large number of natural products from different sources. Additionally, other databases oriented toward specific organisms such as ECMDB for *E. coli*, YMDB for yeast, and ResPect for plants, also provide access to extensive collections of experimental MS/MS data [194]. Most of these alternatives have their own inner structures, hence the functionalities for querying and extracting data vary. An interesting feature that is worth noting for large scale metabolomics is a database integration with other data processing and analysis pipelines. Metlin provides the largest individual collection of MS/MS experimental data acquired in multiple collision energies and despite being the only one of the larger databases mentioned to impose restrictions on downloading data, it is directly integrated with the XCMS online processing platform [191]. MZmine 2 incorporates functions to directly query Kyoto Encyclopedia of Genes and Genomes (KEGG), HMDB, Yeast Metabolome Database (YMDB), LipidMaps, the European MassBank, ChempSpider, and Metacyc [184]. MS-DIAL allows for the uploading and automatic matching of spectra against user-defined libraries in the msp. format, the commonly used text-based format for metabolomics spectra [179]. It also provides multiple libraries for download in this format including MassBank, ReSpect, and GNPS, among others.

One of the major obstacles for mass spectrometry based annotation in metabolomics is the lack of characterized standards for the majority of the different compounds found in nature. Therefore, several alternatives have been developed to cope with this limitation. Most large metabolomics databases such as Metlin and HMDB have significantly expanded their coverage by including in silico generated MS/MS spectra of known compounds for which no experimental data are available, usually based on machine learning or quantum mechanics calculations [195]. Similar approaches have also been used to predict putative compounds based on the MS/MS spectra. These functionalities are provided in tools such as CSI:fingerID [196] and MS-FINDER [197]. Another interesting and popular approach provided via the GNPS database is to construct networks based on spectra similarity and extend annotations based on the assumption that related metabolites exhibit similar second order spectra [28].

Additionally to the mass spectra, LC-MS provides another measure with intrinsic structural information, which is the retention time. However, as previously mentioned, there are multiple factors affecting retention time in liquid chromatography that are hard to predict and control to a similar extent of what is done in gas chromatography. Therefore, few approaches for metabolite annotation and cataloguing include retention time information. It is worth mentioning a few recent efforts in integrating this extra information in spectral libraries such as the WEIZMASS [97] as well as strategies for in silico prediction of retention indexes as that provided by PredRet [198].

Finally, there are also approaches for compound annotation that are independent of the acquisition of tandem mass spectra fragmentation. Most untargeted metabolomics platforms nowadays rely on high resolution and high accuracy m/z measurements provided mainly by QTOF and Orbitrap-based systems. Directly matching putative molecular ion masses against databases is not as reliable as MS/MS spectral matching, it can however, still bring valuable insights in putative structures and the higher accuracy measurements provided by those systems can provide a significantly smaller search space. This is particularly useful when genomics and taxonomic data can be integrated into the search to even further restrict the space of reasonable matches. Unfortunately, this information is still often scarce and not well structured, with few databases providing extensive

collections of curated data. One exception worth mentioning is the KNApSack database, which provides the distribution of over 50,000 compounds across over 20,000 different species [199].

5. Successful Applications of Metabolomics in Natural Products Discovery

5.1. Metabolomics for Secondary Metabolites Identification

In the science of natural product discovery using metabolomics, the separation of metabolites is usually performed using GC or HPLC, which is generally coupled with on-line MS detection. MS fragmentation spectra provide precise information on the structures, and it is thus common to record multiple-stage MS data with the spectra of two or more products (MS_2 , MS_n , where n is the number of production stages) [200]. As above-mentioned, several MS-based databases and software tools are now applied for natural products identification; these databases include the Golm Metabolome Database (GMD) that uses GC retention indices and electron impact (EI) mass spectra, all acquired under defined conditions as mass spectral tags (MSTs) for both neat authentic standards as well as plant extracts in GC-MS data [78]. When GC-MS data of a plant extract are obtained, the MST of its components can be matched to the GMD to record either known metabolites or unique identifiers for unidentified peaks in complex plant extracts. There is also the decision tree search, which predicts substructures of unknown metabolites, even if the full structure cannot be predicted [78]. METLIN, a MS/MS database, includes about 62,000 spectra representing more than 12,000 metabolites. These spectra were acquired under standardized conditions using electrospray ionization in positive and negative ionization modes on a quadrupole time-of-flight (QTOF) mass spectrometer [201]. The limitations of this database are the lack of chromatographic data and closed design. Meanwhile, MassBank is a database of high resolution MS-based platforms (GC-MS and HPLC-MS). All MassBank has spectral information, and some have chromatographic information, but there are no retention times (or indices) data [192]. In contrast, ReSpect is an MS_2 database specific to plant metabolites. The main advantage of this database is that spectral records are annotated with taxonomic information about the species from which a particular metabolite has been extracted and to which structural class the metabolite belongs [202]. Similarly, the Global Natural Products Social Molecular Networking resource (GNPS) is an MS_2 database for natural products [28]. Several databases are considered as an important tool to facilitate the dereplication process [80,203].

NMR is the most powerful technique with respect to structure elucidation, as it is highly reproducible, quantitative, requires simple sample preparation, and is able to measure analytes over a wide range of solvent conditions [204]. It is, however, limited by being slower and less sensitive than MS, and as such, it cannot easily be used as a routine technology in metabolomics. Instead, its application is mainly limited to measuring the most abundant metabolites in the sample [205,206]. In natural product discovery, fractions with the metabolites of interest are separated through chromatographic runs and then subjected to off-line NMR analysis. It is more common in applications of metabolomics to make NMR analysis of plant extracts or biological fluids without chromatographic separation [80]. The applications and limitations of NMR databases that are particularly useful for natural products discovery are discussed in Table 1. To search in one of the mentioned NMR databases, the obtained spectra should be processed and interpreted using external software and entered into a web interface as a peak table. The rNMR software interfaces convert spectral peaks to a searchable file format, so can be used in the MMCD database [207]. On the other hand, SpinAssign is used for the identification of compounds using the ^{13}C -HSQC spectra of complex cell extracts [208], while the MetaboHunter and COLMAR tools are designed to search the NMR spectra of mixtures against those of a subset of metabolites in HMDB and BMRB [209,210].

Table 1. Nuclear magnetic resonance (NMR) databases and computational methods.

Database	Applications	Notes	References
NAPROC-13	Is useful in cases of dereplication as it has more than 20,000 of natural product spectra and inclusion of metabolite classification	Its closed design, there is no user access to spectral data. ¹³ C-NMR spectra, molecular formula or predicted molecular weight can only be searched separately.	[211]
NMRShiftDB	Not limited to NP and has 51,000 collected spectra. It accepts submissions	It lists NMR chemical shifts, but not peak size.	[212,213]
SDBS	Not limited to NP and has 29,000 collected spectra. It uses multiple kinds of spectra in a single search.	Does not accept submissions.	[214]
BML-NMR	The spectral depth of 16 different one- and two-dimensional experiments for each compound provides high quality references	Covers only 208 compounds, but each compound is measured with 16 different NMR parameter sets which provides high quality references. These metabolites were selected based on their importance within metabolic pathways and their detection potential by NMR. They were analyzed at pH 6.6, 7.0, and 7.4.	[215]
BMRB database	Contains NMR data for various biomolecules with a focus on protein, peptide, and nucleic acid spectra. Spectra are available for downloading in raw and processed data formats	The database mainly covers the compounds involved in the lignin biosynthesis which is a plant cell wall constituent and the compounds obtained by plant cell wall deconstruction.	[216]

NAPROC-13: Carbon-13 Database of Natural Products and Related Substances; DB: Database; SDBS: Spectral Database for Organic Compounds; BML-NMR: Birmingham Metabolite Library; BMRB: Biological Magnetic Resonance Data Bank.

There is an increasing tendency in metabolomics to analyze the same sample by both NMR and MS. Taking advantage of both methods enables a better coverage of the metabolome [217,218]. NMR allows the identification of trends in metabolic alteration across core metabolic pathways, while MS provides the identification of the minor metabolites. The integration of NMR and MS can increase the size of the obtained dataset with the added complexity of the simultaneous processing, analysis, and interpretation of two dissimilar data types [218,219]. Indeed, as described below, LC-online-NMR has been proposed as a route toward improving coverage of the metabolome. However, despite the advocacy of this approach, relatively few examples of its utility have been published to date [106,114].

5.2. Metabolomics in Defining or Refining the Pathway Structure of Plant Natural Products

Given that this subject has been extensively reviewed [3], we will keep this section short, suffice to say that the chemistry of many medicinal plants has remained largely unexplored. Despite the advances of synthetic biology for the production of medicinal compounds in heterologous hosts (described in the section below), the native plant species often remain the most reliable and economic source for their production. It is thus of fundamental importance to investigate the metabolic composition of medicinal plants to characterize their natural metabolic diversity and to define their in planta biosynthetic routes in order to develop strategies to further increase their content. In this vein, the cases of benzoisoquinoline and monoterpene indole alkaloids, cannabinoids, caffeine, ginsenosides, withanolides, artemisinin, and taxol represent interesting case studies in which the

contribution of recent integrative metabolomics and genomic approaches in augmenting earlier biochemical strategies have helped to elucidate their metabolic pathways and cellular compartmentation [3].

5.3. Metabolomics to Aid Metabolic Engineering of Secondary Metabolites Production

A recent approach in the field of metabolomics is its use as a generic debugging tool for engineered microbial production systems that arise as a result of advanced synthetic biology. The synthetic biology of secondary metabolite production is not restricted to the awakening of the cryptic metabolite production encoded in newly sequenced genomes, but it also serves in modifying and recombining the modular biosynthesis apparatus found in nature, which leads to the generation of novel chemistry, in addition to the more ambitious re-engineering strategies such as refactoring, which means replacing the native regulatory machinery by fine-tuned designer systems or the improvement of production levels through the redeployment of primary metabolism [220].

In a metabolomics study of the consequences of an overexpression of ncRNA-based regulatory element in *Streptomyces coelicolor* [221], where ncRNA affected not only the nearby metabolites, but also the metabolite levels that showed rapid changes throughout the metabolic network of the organism. The reproducibility of the metabolite dynamics was ensured by repeated metabolite profiling. Another application is the detection of the accumulated toxic side products and intermediates, or the depletion of required precursors, which is considered as metabolic bottlenecks. In these cases, metabolomics provides an unbiased overview of the metabolic status of a system and its changes due to the overproduction of compounds of interest, which in combination with systems biological modeling can drive cycles of refined engineering [222]. The *Streptomyces lydicus* metabolome was characterized in various cultures where the critical precursors of streptolydigin were identified through the depletion in overproducing cultures. Here, the added precursors (glutamic acid and proline) led to an increase in streptolydigin production, with no effect on the strain growth. Therefore, metabolomics approaches enabled the identification of the major bottleneck in the system with no need for understanding the underlying metabolic and regulatory network [223].

Given the massive diversity of secondary metabolism in plants, a recent approach to metabolic engineering has been combinatorial biosynthesis, an engineering strategy that allows for the generation of novel, structurally similar, but distinct compounds, which is achieved by combining biosynthetic genes of related metabolic pathways from different organisms in a single host [224–226]. This requires the use of enzymes that have relaxed specificity and thus are able to incorporate derivatives of their natural substrates. Other than pioneering studies on the synthesis of novel carotenoids in nonphotosynthetic *Escherichia coli* [227], a few other reports of successful targeted combinatorial biosynthesis of (novel) terpenoids in plants exist. For example, the biosynthesis of monoterpenoid indole alkaloids (MIAs) in the Madagascar periwinkle (*Catharanthus roseus*) was carried out following the expression of two different bacterial halogenases, yielding new-to-nature halogenated MIA derivatives [228]. Similarly, biosynthetic genes of five different plants were combined in yeast to produce a monoglycosylated triterpene saponin that does not occur in any of the parent plant species [229]. Moreover, artemisinin was recently transferred from the medicinal plant of origin into the crop plant tobacco [230]. In all of the examples above, metabolomics was instrumental in confirming the novel natural products. It is important to note that this subject area is currently in bloom and it seems reasonable to anticipate that many more examples will follow shortly. Furthermore, the labelling approach in metabolomics such as ^{13}C -based metabolomics studies, can also be applied for discovering novel secondary metabolism [231–233]. Although this technique needs a single source of carbon or nitrogen, this can be achieved by tracing the distribution of a class-specific precursor to side branches of the metabolic network [234].

5.4. Metabolomics and Dereplication

Another application of metabolomics is the dereplication of the natural product biosynthesis at different development stages through the use of various analytical methods, the bioassay guided isolation, thus rapid dereplication of known activities is efficiently delivered [235]. High resolution

Fourier transform mass spectrometry (HRFTMS) is usually used for dereplication of secondary metabolites together with LTQ-Orbitrap and high resolution NMR. Identification of high resolution MS and NMR of the metabolomes is undertaken using the online and in-house databases. Multivariate analysis allows the Fourier transformation of the FID (free induction decay) of multiple samples data, in order to statistically validate the parameters to produce biologically novel secondary metabolites [236,237]. Combining PCA with dereplication and molecular networking was applied on microbial extracts for the discovery of novel compounds where the molecular networking provided a chance to compare the metabolic profiles of complex crude fermentation extracts, leading to efficient chemical dereplication [238].

The dereplication process was summarized by Tawfike et al. (2013, Figure 2) where the extracts, fractions, and purified secondary metabolites were first tested for possible biological activity. After that, the chemical profiling of the active extract or fraction was analyzed using high resolution mass spectrometry and NMR spectroscopy to identify the constituents of interest at an early stage to minimize the efforts by focusing on the active principles. Following the application of various metabolomic tools and multivariate statistics, the obtained huge datasets could be utilized for the identification of the biomarkers of the biologically-active extracts. The identified biomarkers can then be employed in the production of the novel secondary metabolites to provide sustainability of the interesting metabolites in genetic and metabolic engineering [113].

In this section, we will discuss some examples for the application of metabolomics in dereplication. Dereplication was applied for the identification and quantification of tocopherol using HPLC analysis of Brazil oil nut from different geographic locations in Brazil. This study proved the presence of the different amounts of α - and β -tocopherol in some samples with the absence of tocopherols in other samples [239]. Dereplication was also used to select superior banana genotypes for breeding based on the phenolic contents and carotenoid profiles, where the HPLC analysis showed an appreciable amount of pro-vitamin A carotenoids in the active germoplasm samples compared with the main cultivars that are currently marketed [240]. Coffee genotypes of three different regions in Brazil were distinguished using GC-MS coupled with statistical analysis where 44 metabolites were characterized as the chemomarkers for differentiation of the origin and genotype [241]. Design of experiments (DoE) and partial least squares (PLS) were used to prove that the physicochemical properties of the solvent have a significant effect on its extraction capacity [242]. Moreover, dereplication was used to detect bromotyrosine-derived metabolites in 14 species of sponge belonging to genus *Aplysina* through measuring their UV absorption using LC-photodiode array detector (PDA)-MS analysis [243]. Additionally, dereplication could differentiate between green and brown Brazilian propolis bead on their terpenoid content using GC-MS [244] as well as the discovery of biomarkers and novel active constituents that may be active against different neglected diseases [245]. UHPLC-(TOF)-MS, together with hierarchical clustering analyses (HCA), was applied to differentiate between six *Lippia* species based on their chemical constituents [246]. In addition, dereplication was successfully applied to determine 11 anti-inflammatory biomarkers in 57 extracts of Asteraceae leaves using HPLC-MS among 1241 peaks [247]. Similarly, 44 metabolites were identified as chemomarkers to determine the geographical origin and genotype of coffee in Brazil using GC-Q-MS [241].

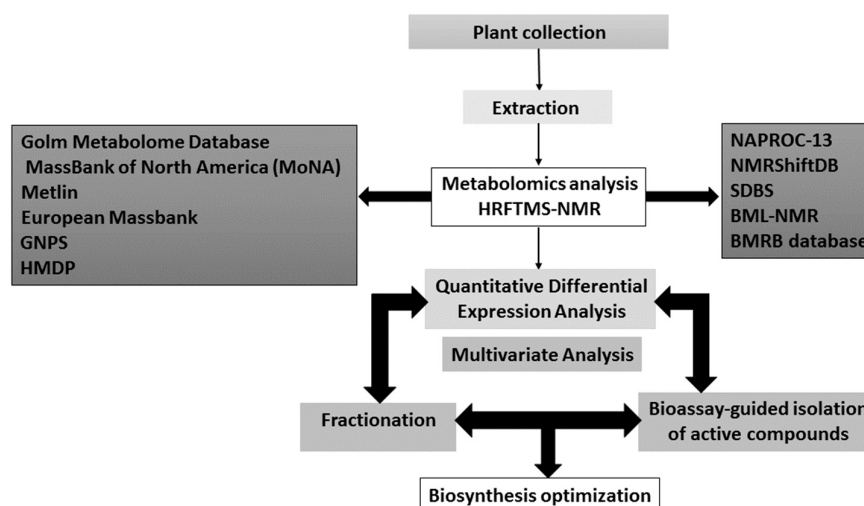


Figure 2. A flowchart showing the application of metabolomics in dereplication.

5.5. Metabolomics for Quality Control of Natural Products

Metabolomics is also finding utility in the quality control of natural products, being used to monitor the variation of metabolic profiles among individuals, environmental alterations during growth and harvesting, post harvesting treatment, extraction, and method of isolation. The change in the chemical profiles have a significant effect on the efficacy of phytomedicines prepared from these herbs [18]. Unsupervised principal component analysis (PCA) and supervised partial least square analysis/partial least square analysis with discriminant analysis (PLS/PLS-DA), combined with ^1H -NMR, are the most common techniques used in quality control. These methods were used to study the effect of location on the percentage of various constituents of chamomile (*Matricaria recutita* L.) [248]. Determination of Δ^9 -tetrahydrocannabinolic acid (THCA) and cannabidiolic acid (CBDA) was carried out to discriminate between different *Cannabis sativa* [249]. Similarly, discrimination of different samples of St John's Wort (*Hypericum perforatum*) from different batches of the same supplier, according to the variation of the content of the antidepressant flavonoids, has also been reported [250]. Metabolomics has also been applied to detect adulterations of herbal preparations with similar species but with low levels of constituents responsible for the activity. A company produced antimalarial capsules containing *Artemisia afra*, which is free from artemisinin instead of *Artemisia annua*, and this could be easily detected by ^1H -NMR and PCA [251]. Metabolomics can additionally be used to profile the biofluids of individuals who have been medicated with natural products and a vast literature describes such studies [30,252,253].

6. Conclusions and Future Perspectives

Metabolomics experiments offer an improved expedited route for plant natural product drug discovery and is increasingly applied in a bioactivity-guided approach from natural extracts. This review provides an overview for the most widely employed analytical tools in metabolomics experiments for natural products' drug ability viz. GC-MS, LC-MS, CE-MS, and NMR with recent advances in their applications. Furthermore, different sample preparation protocols, data processing tools, and some successful applications of metabolomics in natural products discovery are outlined, highlighting its advantages and limitations. Although metabolomics has clearly transcended the classical approach of lead compound discovery as previously mentioned, the main challenge in such experiments appears to simultaneously explore the extreme complexity and the huge chemical diversity of metabolites present in any natural organism. In addition, the annotation of metabolites

in such complicated environments remains another obstacle that needs to be tackled. The establishment of organism databases and sharing of spectral data in public repositories will aid in that regard, and avoid dereplication of already existing metabolites in bioactivity guided studies for active agents. Bioinformatic tools that can link the metabolomics dataset with bioactivity results have yet to be implemented for all spectral datasets. Finally, we believe that there is urgent need to apply the integrative approaches of genomics and metabolomics to understanding the metabolism of natural products in plants and elucidate their metabolic pathways.

Funding: M.A.F. thanks the Alexander von Humboldt Foundation for financial support and research grant number 30 from the Academy of Scientific Research & Technology (ASRT), Egypt. S.A. and ARF acknowledge the PlantaSYST project by the European Union Horizon 2020 Research and Innovation Programme (SGA-CSA no. 664621 and no. 739582 under FPA no. 664620).

Conflicts of Interest: The authors declare no conflict of interest.

Reference

1. Cragg, G.M.; Newman, D.J. Natural Product Drug Discovery in the Next Millennium. *Pharm. Biol.* **2001**, *39*, 8–17, doi:10.1076/phbi.39.s1.8.0009.
2. Newman, D.J.; Cragg, G.M. Natural Products as Sources of New Drugs from 1981 to 2014. *J. Nat. Prod.* **2016**, *79*, 629–661, doi:10.1021/acs.jnatprod.5b01055.
3. Scossa, F.; Benina, M.; Alseekh, S.; Zhang, Y.; Fernie, A.R. The Integration of Metabolomics and Next-Generation Sequencing Data to Elucidate the Pathways of Natural Product Metabolism in Medicinal Plants. *Planta Med.* **2018**, *84*, 855–873, doi:10.1055/a-0630-1899.
4. Skirycz, A.; Kierszniowska, S.; Méret, M.; Willmitzer, L.; Tzotzos, G. Medicinal Bioprospecting of the Amazon Rainforest: A Modern Eldorado? *Trends Biotechnol.* **2016**, *34*, 781–790, doi:10.1016/j.tibtech.2016.03.006.
5. Yoshioka, T.; Nagatomi, Y.; Harayama, K.; Bamba, T. Development of an analytical method for polycyclic aromatic hydrocarbons in coffee beverages and dark beer using novel high-sensitivity technique of supercritical fluid chromatography/mass spectrometry. *J. Biosci. Bioeng.* **2018**, *126*, 126–130, doi:10.1016/j.jbiosc.2018.01.014.
6. Morales, D.; Piris, A.J.; Ruiz-Rodriguez, A.; Prodanov, M.; Soler-Rivas, C. Extraction of bioactive compounds against cardiovascular diseases from *Lentinula edodes* using a sequential extraction method. *Biotechnol. Prog.* **2018**, *34*, 746–755, doi:10.1002/btpr.2616.
7. Joana Gil-Chávez, G.; Villa, J.A.; Fernando Ayala-Zavala, J.; Basilio Heredia, J.; Sepulveda, D.; Yahia, E.M.; González-Aguilar, G.A. Technologies for Extraction and Production of Bioactive Compounds to be Used as Nutraceuticals and Food Ingredients: An Overview. *Compr. Rev. Food Sci. Food Saf.* **2013**, *12*, 5–23, doi:10.1111/1541-4337.12005.
8. Gan, Z.; Liang, Z.; Chen, X.; Wen, X.; Wang, Y.; Li, M.; Ni, Y. Separation and preparation of 6-gingerol from molecular distillation residue of Yunnan ginger rhizomes by high-speed counter-current chromatography and the antioxidant activity of ginger oils in vitro. *J. Chromatogr. B* **2016**, *1011*, 99–107, doi:10.1016/j.jchromb.2015.12.051.
9. Williams, S.R.; Oatley, D.L.; Abdrahman, A.; Butt, T.; Nash, R. Membrane Technology for the Improved Separation of Bioactive Compounds. *Procedia Eng.* **2012**, *44*, 2112–2114, doi:10.1016/j.proeng.2012.09.064.
10. Salam, A.M.; Lyles, J.T.; Quave, C.L. Methods in the Extraction and Chemical Analysis of Medicinal Plants. In *Methods and Techniques in Ethnobiology and Ethnobotany*; Albuquerque, U.P., de Lucena, R.F.P., Cruz da Cunha, L.V.F., Alves, R.R.N., Eds.; Springer: New York, NY, USA, 2019; pp. 257–283, doi:10.1007/978-1-4939-8919-5_17.
11. Thomford, N.E.; Senthane, D.A.; Rowe, A.; Munro, D.; Seele, P.; Maroyi, A.; Dzobo, K. Natural Products for Drug Discovery in the 21st Century: Innovations for Novel Drug Discovery. *Int. J. Mol. Sci.* **2018**, *19*, 1578, doi:10.3390/ijms19061578.
12. Robinette, S.L.; Brunschweiler, R.; Schroeder, F.C.; Edison, A.S. NMR in metabolomics and natural products research: Two sides of the same coin. *Acc. Chem. Res.* **2012**, *45*, 288–297, doi:10.1021/ar2001606.
13. Sharma, G.; Tyagi, A.K.; Singh, R.P.; Chan, D.C.F.; Agarwal, R. Synergistic Anti-Cancer Effects of Grape Seed Extract and Conventional Cytotoxic Agent Doxorubicin Against Human Breast Carcinoma Cells. *Breast Cancer Res. Treat.* **2004**, *85*, 1–12, doi:10.1023/b:brea.0000020991.55659.59.

14. Chusri, S.; Siriyong, T.; Na-Phatthalung, P.; Voravuthikunchai, S.P. Synergistic effects of ethnomedicinal plants of Apocynaceae family and antibiotics against clinical isolates of *Acinetobacter baumannii*. *Asian Pac. J. Trop. Med.* **2014**, *7*, 456–461, doi:10.1016/S1995-7645(14)60074-2.
15. Skroza, D.; Generalić Mekinić, I.; Svilović, S.; Šimat, V.; Katalinić, V. Investigation of the potential synergistic effect of resveratrol with other phenolic compounds: A case of binary phenolic mixtures. *J. Food Compos. Anal.* **2015**, *38*, 13–18, doi:10.1016/j.jfca.2014.06.013.
16. Shyur, L.-F.; Yang, N.-S. Metabolomics for phytomedicine research and drug development. *Curr. Opin. Chem. Biol.* **2008**, *12*, 66–71, doi:10.1016/j.cbpa.2008.01.032.
17. Goodacre, R.; Vaidyanathan, S.; Dunn, W.B.; Harrigan, G.G.; Kell, D.B. Metabolomics by numbers: Acquiring and understanding global metabolite data. *Trends Biotechnol.* **2004**, *22*, 245–252, doi:10.1016/j.tibtech.2004.03.007.
18. Wang, M.; Lamers, R.J.; Korthout, H.A.; van Nesselrooij, J.H.; Witkamp, R.F.; van der Heijden, R.; Voshol, P.J.; Havekes, L.M.; Verpoorte, R.; van der Greef, J. Metabolomics in the context of systems biology: Bridging traditional Chinese medicine and molecular pharmacology. *Phytother. Res.* **2005**, *19*, 173–182, doi:10.1002/ptr.1624.
19. Ulrich-Merzenich, G.; Zeitler, H.; Jobst, D.; Panek, D.; Vetter, H.; Wagner, H. Application of the “-Omic-” technologies in phytomedicine. *Phytomedicine* **2007**, *14*, 70–82, doi:10.1016/j.phymed.2006.11.011.
20. Tebani, A.; Afonso, C.; Bekri, S. Advances in metabolome information retrieval: Turning chemistry into biology. Part I: Analytical chemistry of the metabolome. *J. Inherit. Metab. Dis.* **2018**, *41*, 379–391, doi:10.1007/s10545-017-0074-y.
21. Alseekh, S.; Fernie, A.R. Metabolomics 20 years on: What have we learned and what hurdles remain? *Plant J.* **2018**, *94*, 933–942, doi:10.1111/tpj.13950.
22. Pinto, R.C. Chemometrics Methods and Strategies in Metabolomics. In *Metabolomics: From Fundamentals to Clinical Applications*; Sussulini, A., Ed.; Springer International Publishing: Cham, Switzerland, 2017; pp. 163–190, doi:10.1007/978-3-319-47656-8_7.
23. Farrag, A.R.H.; Abdallah, H.M.I.; Khatatb, A.R.; Elshamy, A.I.; Gendy, A.E.-N.G.E.; Mohamed, T.A.; Farag, M.A.; Efferth, T.; Hegazy, M.-E.F. Antiulcer activity of *Cyperus alternifolius* in relation to its UPLC-MS metabolite fingerprint: A mechanistic study. *Phytomedicine* **2019**, *62*, 152970, doi:10.1016/j.phymed.2019.152970.
24. Yang, S.Y.; Kim, H.K.; Lefeber, A.W.M.; Erkelens, C.; Angelova, N.; Choi, Y.H.; Verpoorte, R. Application of Two-Dimensional Nuclear Magnetic Resonance Spectroscopy to Quality Control of Ginseng Commercial Products. *Planta Med.* **2006**, *72*, 364–369, doi:10.1055/s-2005-916240.
25. Zeng, Z.-D.; Liang, Y.-Z.; Chau, F.-T.; Chen, S.; Daniel, M.K.-W.; Chan, C.-O. Mass spectral profiling: An effective tool for quality control of herbal medicines. *Anal. Chim. Acta* **2007**, *604*, 89–98, doi:10.1016/j.aca.2007.09.057.
26. Dunn, W.B.; Erban, A.; Weber, R.J.M.; Creek, D.J.; Brown, M.; Breitling, R.; Hankemeier, T.; Goodacre, R.; Neumann, S.; Kopka, J.; et al. Mass appeal: Metabolite identification in mass spectrometry-focused untargeted metabolomics. *Metabolomics* **2013**, *9*, 44–66, doi:10.1007/s11306-012-0434-4.
27. Perez de Souza, L.; Naake, T.; Tohge, T.; Fernie, A.R. From chromatogram to analyte to metabolite. How to pick horses for courses from the massive web resources for mass spectral plant metabolomics. *GigaScience* **2017**, *6*, 1–20, doi:10.1093/gigascience/gix037.
28. Wang, M.; Carver, J.J.; Phelan, V.V.; Sanchez, L.M.; Garg, N.; Peng, Y.; Nguyen, D.D.; Watrous, J.; Kapon, C.A.; Luzzatto-Knaan, T.; et al. Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nat. Biotechnol.* **2016**, *34*, 828, doi:10.1038/nbt.3597.
29. Olivon, F.; Elie, N.; Grellet, G.; Roussi, F.; Litaudon, M.; Touboul, D. MetGem Software for the Generation of Molecular Networks Based on the t-SNE Algorithm. *Anal. Chem.* **2018**, *90*, 13900–13908, doi:10.1021/acs.analchem.8b03099.
30. Riekeberg, E.; Powers, R. New frontiers in metabolomics: From measurement to insight. *F1000Research* **2017**, *6*, 1148, doi:10.12688/f1000research.11495.1.
31. Sumner, L.W.; Amberg, A.; Barrett, D.; Beale, M.H.; Beger, R.; Daykin, C.A.; Fan, T.W.; Fiehn, O.; Goodacre, R.; Griffin, J.L.; et al. Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics* **2007**, *3*, 211–221, doi:10.1007/s11306-007-0082-2.

32. Vuckovic, D. Current trends and challenges in sample preparation for global metabolomics using liquid chromatography-mass spectrometry. *Anal. Bioanal. Chem.* **2012**, *403*, 1523–1548, doi:10.1007/s00216-012-6039-y.
33. Cajka, T.; Fiehn, O. Toward Merging Untargeted and Targeted Methods in Mass Spectrometry-Based Metabolomics and Lipidomics. *Anal. Chem.* **2016**, *88*, 524–545, doi:10.1021/acs.analchem.5b04491.
34. Blainey, P.; Krzywinski, M.; Altman, N. Replication. *Nat. Methods* **2014**, *11*, 879–880, doi:10.1038/nmeth.3091.
35. Kim, H.K.; Verpoorte, R. Sample preparation for plant metabolomics. *Phytochem. Anal.* **2010**, *21*, 4–13, doi:10.1002/pca.1188.
36. Ernst, M.; Silva, D.B.; Silva, R.R.; Vencio, R.Z.; Lopes, N.P. Mass spectrometry in plant metabolomics strategies: From analytical platforms to data acquisition and processing. *Nat. Prod. Rep.* **2014**, *31*, 784–806, doi:10.1039/c3np70086k.
37. Fernie, A.R.; Trethewey, R.N.; Krotzky, A.J.; Willmitzer, L. Innovation—Metabolite profiling: From diagnostics to systems biology. *Nat. Rev. Mol. Cell Biol.* **2004**, *5*, 763–769, doi:10.1038/nrm1451.
38. Dunn, W.B.; Winder, C.L. Sample Preparation Related to the Intracellular Metabolome of Yeast: Methods for Quenching, Extraction, and Metabolite Quantitation. *Method Enzym.* **2011**, *500*, 277–297, doi:10.1016/B978-0-12-385118-5.00015-3.
39. Bylda, C.; Thiele, R.; Kobold, U.; Volmer, D.A. Recent advances in sample preparation techniques to overcome difficulties encountered during quantitative analysis of small molecules from biofluids using LC-MS/MS. *Analyst* **2014**, *139*, 2265–2276, doi:10.1039/c4an00094c.
40. Mushtaq, M.Y.; Choi, Y.H.; Verpoorte, R.; Wilson, E.G. Extraction for Metabolomics: Access to the Metabolome. *Phytochem. Anal.* **2014**, *25*, 291–306, doi:10.1002/pca.2505.
41. Salem, M.A.; Giavalisco, P. Semi-targeted Lipidomics of Plant Acyl Lipids Using UPLC-HR-MS in Combination with a Data-Independent Acquisition Mode. In *Plant Metabolomics: Methods and Protocols, Methods in Molecular Biology*; António, C., Ed.; Springer: New York, NY, USA; Humana Press: Totowa, NJ, USA, 2018; Volume 1778, pp. 137–155.
42. Salem, M.; Bernach, M.; Bajdzienko, K.; Giavalisco, P. A Simple Fractionated Extraction Method for the Comprehensive Analysis of Metabolites, Lipids, and Proteins from a Single Sample. *J. Vis. Exp.* **2017**, *124*, e55802, doi:10.3791/55802.
43. Raterink, R.-J.; Lindenburg, P.W.; Vreeken, R.J.; Ramautar, R.; Hankemeier, T. Recent developments in sample-pretreatment techniques for mass spectrometry-based metabolomics. *TrAC Trends Anal. Chem.* **2014**, *61*, 157–167, doi:10.1016/j.trac.2014.06.003.
44. Bligh, E.G.; Dyer, W.J. A rapid method of total lipid extraction and purification. *Can. J. Biochem. Physiol.* **1959**, *37*, 911–917, doi:10.1139/o59-099.
45. Folch, J.; Lees, M.; Stanley, G.H.S. A Simple Method for the Isolation and Purification of Total Lipides from Animal Tissues. *J. Biol. Chem.* **1957**, *226*, 497–509.
46. Weckwerth, W.; Wenzel, K.; Fiehn, O. Process for the integrated extraction, identification and quantification of metabolites, proteins and RNA to reveal their co-regulation in biochemical networks. *Proteomics* **2004**, *4*, 78–83, doi:10.1002/pmic.200200500.
47. Fiehn, O.; Kopka, J.; Trethewey, R.N.; Willmitzer, L. Identification of uncommon plant metabolites based on calculation of elemental compositions using gas chromatography and quadrupole mass spectrometry. *Anal. Chem.* **2000**, *72*, 3573–3580, doi:10.1021/ac991142i.
48. Fiehn, O.; Kopka, J.; Dormann, P.; Altmann, T.; Trethewey, R.N.; Willmitzer, L. Metabolite profiling for plant functional genomics. *Nat. Biotechnol.* **2000**, *18*, 1157–1161, doi:10.1038/81137.
49. Lisec, J.; Schauer, N.; Kopka, J.; Willmitzer, L.; Fernie, A.R. Gas chromatography mass spectrometry-based metabolite profiling in plants. *Nat. Protoc.* **2006**, *1*, 387–396, doi:10.1038/nprot.2006.59.
50. Giavalisco, P.; Li, Y.; Matthes, A.; Eckhardt, A.; Hubberten, H.M.; Hesse, H.; Segu, S.; Hummel, J.; Kohl, K.; Willmitzer, L. Elemental formula annotation of polar and lipophilic metabolites using ^{13}C , ^{15}N and ^{34}S isotope labelling, in combination with high-resolution mass spectrometry. *Plant J. Cell Mol. Biol.* **2011**, *68*, 364–376, doi:10.1111/j.1365-313X.2011.04682.x.
51. Bozek, K.; Wei, Y.; Yan, Z.; Liu, X.; Xiong, J.; Sugimoto, M.; Tomita, M.; Paabo, S.; Sherwood, C.C.; Hof, P.R.; et al. Organization and evolution of brain lipidome revealed by large-scale analysis of human, chimpanzee, macaque, and mouse tissues. *Neuron* **2015**, *85*, 695–702, doi:10.1016/j.neuron.2015.01.003.

52. Khrameeva, E.E.; Bozek, K.; He, L.; Yan, Z.; Jiang, X.; Wei, Y.; Tang, K.; Gelfand, M.S.; Prufer, K.; Kelso, J.; et al. Neanderthal ancestry drives evolution of lipid catabolism in contemporary Europeans. *Nat. Commun.* **2014**, *5*, 3584, doi:10.1038/ncomms4584.
53. Bozek, K.; Wei, Y.; Yan, Z.; Liu, X.; Xiong, J.; Sugimoto, M.; Tomita, M.; Paabo, S.; Pieszek, R.; Sherwood, C.C.; et al. Exceptional evolutionary divergence of human muscle and brain metabolomes parallels human cognitive and physical uniqueness. *PLoS Biol.* **2014**, *12*, e1001871, doi:10.1371/journal.pbio.1001871.
54. Kuyukina, M.S.; Ivshina, I.B.; Philp, J.C.; Christofi, N.; Dunbar, S.A.; Ritchkova, M.I. Recovery of Rhodococcus biosurfactants using methyl tertiary-butyl ether extraction. *J. Microbiol. Methods* **2001**, *46*, 149–156.
55. Matyash, V.; Liebisch, G.; Kurzchalia, T.V.; Shevchenko, A.; Schwudke, D. Lipid extraction by methyl-tert-butyl ether for high-throughput lipidomics. *J. Lipid Res.* **2008**, *49*, 1137–1146, doi:10.1194/jlr.D700041-JLR200.
56. Salem, M.A.; Li, Y.; Bajdzienko, K.; Fisahn, J.; Watanabe, M.; Hoefgen, R.; Schottler, M.A.; Giavalisco, P. RAPTOR Controls Developmental Growth Transitions by Altering the Hormonal and Metabolic Balance. *Plant Physiol.* **2018**, *177*, 565–593, doi:10.1104/pp.17.01711.
57. Farag, M.A.; Meyer, A.; Ali, S.E.; Salem, M.A.; Giavalisco, P.; Westphal, H.; Wessjohann, L.A. Comparative Metabolomics Approach Detects Stress-Specific Responses during Coral Bleaching in Soft Corals. *J. Proteome Res.* **2018**, *17*, 2060–2071, doi:10.1021/acs.jproteome.7b00929.
58. Alhajturki, D.; Muralidharan, S.; Nurmi, M.; Rowan, B.A.; Lunn, J.E.; Boldt, H.; Salem, M.A.; Alseekh, S.; Jorzig, C.; Feil, R.; et al. Dose-dependent interactions between two loci trigger altered shoot growth in BG-5 × Krotzenburg-0 (Kro-0) hybrids of Arabidopsis thaliana. *New Phytol.* **2018**, *217*, 392–406, doi:10.1111/nph.14781.
59. Salem, M.A.; Li, Y.; Wiszniewski, A.; Giavalisco, P. Regulatory-associated protein of TOR (RAPTOR) alters the hormonal and metabolic composition of Arabidopsis seeds, controlling seed morphology, viability and germination potential. *Plant J. Cell Mol. Biol.* **2017**, *92*, 525–545, doi:10.1111/tjp.13667.
60. Salem, M.A.; Juppner, J.; Bajdzienko, K.; Giavalisco, P. Protocol: A fast, comprehensive and reproducible one-step extraction method for the rapid preparation of polar and semi-polar metabolites, lipids, proteins, starch and cell wall polymers from a single sample. *Plant Methods* **2016**, *12*, 45, doi:10.1186/s13007-016-0146-2.
61. Doppler, M.; Kluger, B.; Bueschl, C.; Schneider, C.; Krska, R.; Delcambre, S.; Hiller, K.; Lemmens, M.; Schuhmacher, R. Stable Isotope-Assisted Evaluation of Different Extraction Solvents for Untargeted Metabolomics of Plants. *Int. J. Mol. Sci.* **2016**, *17*, 1017, doi:10.3390/ijms17071017.
62. Roessner, U.; Dias, D.A. Plant Tissue Extraction for Metabolomics. *Metab. Tools Nat. Prod. Discov. Methods Protoc.* **2013**, *1055*, 21–28, doi:10.1007/978-1-62703-577-4_2.
63. De Vos, R.C.; Moco, S.; Lommen, A.; Keurentjes, J.J.; Bino, R.J.; Hall, R.D. Untargeted large-scale plant metabolomics using liquid chromatography coupled to mass spectrometry. *Nat. Protoc.* **2007**, *2*, 778–791, doi:10.1038/nprot.2007.95.
64. Cajka, T.; Fiehn, O. Comprehensive analysis of lipids in biological systems by liquid chromatography-mass spectrometry. *Trends Anal. Chem. TrAC* **2014**, *61*, 192–206, doi:10.1016/j.trac.2014.04.017.
65. Fernie, A.R.; Aharoni, A.; Willmitzer, L.; Stitt, M.; Tohge, T.; Kopka, J.; Carroll, A.J.; Saito, K.; Fraser, P.D.; DeLuca, V. Recommendations for reporting metabolite data. *Plant Cell* **2011**, *23*, 2477–2482, doi:10.1105/tpc.111.086272.
66. Dixon, R.A.; Strack, D. Phytochemistry meets genome analysis, and beyond. *Phytochemistry* **2003**, *62*, 815–816, doi:10.1016/s0031-9422(02)00712-4.
67. Rai, A.; Saito, K.; Yamazaki, M. Integrated omics analysis of specialized metabolism in medicinal plants. *Plant J. Cell Mol. Biol.* **2017**, *90*, 764–787, doi:10.1111/tjp.13485.
68. Kruger, N.J.; Troncoso-Ponce, M.A.; Ratcliffe, R.G. ¹H NMR metabolite fingerprinting and metabolomic analysis of perchloric acid extracts from plant tissues. *Nat. Protoc.* **2008**, *3*, 1001–1012, doi:10.1038/nprot.2008.64.
69. Tohge, T.; Fernie, A.R. Combining genetic diversity, informatics and metabolomics to facilitate annotation of plant gene function. *Nat. Protoc.* **2010**, *5*, 1210–1227, doi:10.1038/nprot.2010.82.
70. Sumner, L.W.; Mendes, P.; Dixon, R.A. Plant metabolomics: Large-scale phytochemistry in the functional genomics era. *Phytochemistry* **2003**, *62*, 817–836, doi:10.1016/s0031-9422(02)00708-2.

71. Lu, W.; Su, X.; Klein, M.S.; Lewis, I.A.; Fiehn, O.; Rabinowitz, J.D. Metabolite Measurement: Pitfalls to Avoid and Practices to Follow. *Annu. Rev. Biochem.* **2017**, *86*, 277–304, doi:10.1146/annurev-biochem-061516-044952.
72. Fiehn, O. Metabolomic—The link between genotypes and phenotypes. *Plant Mol. Biol.* **2002**, *48*, 155–171.
73. Kopka, J.; Fernie, A.; Weckwerth, W.; Gibon, Y.; Stitt, M. Metabolite profiling in plant biology: Platforms and destinations. *Genome Biol.* **2004**, *5*, 109, doi:10.1186/gb-2004-5-6-109.
74. Obata, T.; Fernie, A.R. The use of metabolomics to dissect plant responses to abiotic stresses. *Cell. Mol. Life Sci.* **2012**, *69*, 3225–3243, doi:10.1007/s00018-012-1091-5.
75. Roessner, U.; Luedemann, A.; Brust, D.; Fiehn, O.; Linke, T.; Willmitzer, L.; Fernie, A. Metabolic profiling allows comprehensive phenotyping of genetically or environmentally modified plant systems. *Plant Cell* **2001**, *13*, 11–29, doi:10.1105/tpc.13.1.11.
76. Halket, J.M.; Przyborowska, A.; Stein, S.E.; Mallard, W.G.; Down, S.; Chalmers, R.A. Deconvolution gas chromatography/mass spectrometry of urinary organic acids-potential for pattern recognition and automated identification of metabolic disorders. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 279–284, doi:10.1002/(sici)1097-0231(19990228)13:4<279::aid-rcm478>3.0.co;2-i.
77. Saito, K.; Matsuda, F. Metabolomics for Functional Genomics, Systems Biology, and Biotechnology. *Annu. Rev. Plant Biol.* **2010**, *61*, 463–489, doi:10.1146/annurev-arplant.043008.092035.
78. Kopka, J.; Schauer, N.; Krueger, S.; Birkemeyer, C.; Usadel, B.; Bergmüller, E.; Dormann, P.; Weckwerth, W.; Gibon, Y.; Stitt, M.; et al. GMD@CSB.DB: The Golm Metabolome Database. *Bioinformatics* **2005**, *21*, 1635–1638, doi:10.1093/bioinformatics/bti236.
79. Schauer, N.; Steinhauser, D.; Strelkov, S.; Schomburg, D.; Allison, G.; Moritz, T.; Lundgren, K.; Roessner-Tunali, U.; Forbes, M.G.; Willmitzer, L.; et al. GC-MS libraries for the rapid identification of metabolites in complex biological samples. *FEBS Lett.* **2005**, *579*, 1332–1337, doi:10.1016/j.febslet.2005.01.029.
80. Johnson, S.R.; Lange, B.M. Open-access metabolomics databases for natural product research: Present capabilities and future potential. *Front. Bioeng. Biotechnol.* **2015**, *3*, 22, doi:10.3389/fbioe.2015.00022.
81. Liu, Z.; Sirimanne, S.R.; Patterson, D.G., Jr.; Needham, L.L. Comprehensive two-dimensional gas chromatography for the fast separation and determination of pesticides extracted from human serum. *Anal. Chem.* **1994**, *66*, 3086–3092, doi:10.1021/ac00091a016.
82. Mondello, L.; Casillia, A.; Tranchida, P.Q.; Dugo, G.; Dugo, P. Comprehensive two-dimensional gas chromatography in combination with rapid scanning quadrupole mass spectrometry in perfume analysis. *J. Chromatogr. A* **2005**, *1067*, 235–243, doi:10.1016/j.chroma.2004.09.040.
83. Yu, Z.; Huang, H.; Reim, A.; Charles, P.D.; Northage, A.; Jackson, D.; Parry, I.; Kessler, B.M. Optimizing 2D gas chromatography mass spectrometry for robust tissue, serum and urine metabolite profiling. *Talanta* **2017**, *165*, 685–691, doi:10.1016/j.talanta.2017.01.003.
84. Egert, B.; Weinert, C.H.; Kulling, S.E. A peaklet-based generic strategy for the untargeted analysis of comprehensive two-dimensional gas chromatography mass spectrometry data sets. *J. Chromatogr. A* **2015**, *1405*, 168–177, doi:10.1016/j.chroma.2015.05.056.
85. Wong, Y.F.; Hartmann, C.; J Marriott, P. Multidimensional gas chromatography methods for bioanalytical research. *Bioanalysis* **2014**, *6*, 2461–2479, doi:10.4155/bio.14.186.
86. Mostafa, A.; Edwards, M.; Gorecki, T. Optimization aspects of comprehensive two-dimensional gas chromatography. *J. Chromatogr. A* **2012**, *1255*, 38–55, doi:10.1016/j.chroma.2012.02.064.
87. Hill, C.B.; Roessner, U. Metabolic Profiling of Plants by GC–MS. In *The Handbook of Plant Metabolomics*; Weckwerth, W., Kahl, G., Eds.; Wiley-VCH: Weinheim, Germany, 2013; pp. 1–23, doi:10.1002/9783527669882.ch1.
88. Iijima, Y.; Nakamura, Y.; Ogata, Y.; Tanaka, K.; Sakurai, N.; Suda, K.; Suzuki, T.; Suzuki, H.; Okazaki, K.; Kitayama, M.; et al. Metabolite annotations based on the integration of mass spectral information. *Plant J. Cell Mol. Biol.* **2008**, *54*, 949–962, doi:10.1111/j.1365-313X.2008.03434.x.
89. Matsuda, F.; Hirai, M.Y.; Sasaki, E.; Akiyama, K.; Yonekura-Sakakibara, K.; Provart, N.J.; Sakurai, T.; Shimada, Y.; Saito, K. AtMetExpress development: A phytochemical atlas of Arabidopsis development. *Plant Physiol.* **2010**, *152*, 566–578, doi:10.1104/pp.109.148031.
90. Matsuda, F.; Yonekura-Sakakibara, K.; Niida, R.; Kuromori, T.; Shinozaki, K.; Saito, K. MS/MS spectral tag-based annotation of non-targeted profile of plant secondary metabolites. *Plant J. Cell Mol. Biol.* **2009**, *57*, 555–577, doi:10.1111/j.1365-313X.2008.03705.x.

91. Okazaki, Y.; Shimojima, M.; Sawada, Y.; Toyooka, K.; Narisawa, T.; Mochida, K.; Tanaka, H.; Matsuda, F.; Hirai, A.; Hirai, M.Y.; et al. A chloroplastic UDP-glucose pyrophosphorylase from *Arabidopsis* is the committed enzyme for the first step of sulfolipid biosynthesis. *Plant Cell* **2009**, *21*, 892–909, doi:10.1105/tpc.108.063925.
92. Salem, M.A.; Giavalisco, P. Semi-targeted Lipidomics of Plant Acyl Lipids Using UPLC-HR-MS in Combination with a Data-Independent Acquisition Mode. *Methods Mol. Biol.* **2018**, *1778*, 137–155, doi:10.1007/978-1-4939-7819-9_10.
93. Gu, L.; Jones, A.D.; Last, R.L. LC-MS/MS assay for protein amino acids and metabolically related compounds for large-scale screening of metabolic phenotypes. *Anal. Chem.* **2007**, *79*, 8067–8075, doi:10.1021/ac070938b.
94. Bocker, S.; Rasche, F. Towards de novo identification of metabolites by analyzing tandem mass spectra. *Bioinformatics* **2008**, *24*, i49–i55, doi:10.1093/bioinformatics/btn270.
95. Keurentjes, J.J.; Fu, J.; de Vos, C.H.; Lommen, A.; Hall, R.D.; Bino, R.J.; van der Plas, L.H.; Jansen, R.C.; Vreugdenhil, D.; Koornneef, M. The genetics of plant metabolism. *Nat. Genet.* **2006**, *38*, 842–849, doi:10.1038/ng1815.
96. Giavalisco, P.; Kohl, K.; Hummel, J.; Seiwert, B.; Willmitzer, L. ¹³C isotope-labeled metabolomes allowing for improved compound annotation and relative quantification in liquid chromatography-mass spectrometry-based metabolomic research. *Anal. Chem.* **2009**, *81*, 6546–6551, doi:10.1021/ac900979e.
97. Shahaf, N.; Rogachev, I.; Heinig, U.; Meir, S.; Malitsky, S.; Battat, M.; Wyner, H.; Zheng, S.; Wehrens, R.; Aharoni, A. The WEIZMASS spectral library for high-confidence metabolite identification. *Nat. Commun.* **2016**, *7*, 12423, doi:10.1038/ncomms12423.
98. Giavalisco, P.; Hummel, J.; Lisec, J.; Inostroza, A.C.; Catchpole, G.; Willmitzer, L. High-resolution direct infusion-based mass spectrometry in combination with whole ¹³C metabolome isotope labeling allows unambiguous assignment of chemical sum formulas. *Anal. Chem.* **2008**, *80*, 9417–9425, doi:10.1021/ac8014627.
99. Nakabayashi, R.; Kusano, M.; Kobayashi, M.; Tohge, T.; Yonekura-Sakakibara, K.; Kogure, N.; Yamazaki, M.; Kitajima, M.; Saito, K.; Takayama, H. Metabolomics-oriented isolation and structure elucidation of 37 compounds including two anthocyanins from *Arabidopsis thaliana*. *Phytochemistry* **2009**, *70*, 1017–1029, doi:10.1016/j.phytochem.2009.03.021.
100. Tohge, T.; Fernie, A.R. Web-based resources for mass-spectrometry-based metabolomics: A user's guide. *Phytochemistry* **2009**, *70*, 450–456, doi:10.1016/j.phytochem.2009.02.004.
101. Ishii, N.; Nakahigashi, K.; Baba, T.; Robert, M.; Soga, T.; Kanai, A.; Hirasawa, T.; Naba, M.; Hirai, K.; Hoque, A.; et al. Multiple high-throughput analyses monitor the response of *E. coli* to perturbations. *Science* **2007**, *316*, 593–597, doi:10.1126/science.1132067.
102. Fernie, A.R.; Tohge, T. The Genetics of Plant Metabolism. *Annu. Rev. Genet.* **2017**, *51*, 287–310, doi:10.1146/annurev-genet-120116-024640.
103. Williams, B.J.; Cameron, C.J.; Workman, R.; Broeckling, C.D.; Sumner, L.W.; Smith, J.T. Amino acid profiling in plant cell cultures: An inter-laboratory comparison of CE-MS and GC-MS. *Electrophoresis* **2007**, *28*, 1371–1379, doi:10.1002/elps.200600609.
104. Soga, T.; Imaizumi, M. Capillary electrophoresis method for the analysis of inorganic anions, organic acids, amino acids, nucleotides, carbohydrates and other anionic compounds. *Electrophoresis* **2001**, *22*, 3418–3425, doi:10.1002/1522-2683(200109)22:16<3418::aid-elps3418>3.0.co;2-8.
105. Ren, J.-L.; Zhang, A.-H.; Kong, L.; Wang, X.-J. Advances in mass spectrometry-based metabolomics for investigation of metabolites. *RSC Adv.* **2018**, *8*, 22335–22350, doi:10.1039/C8RA01574K.
106. Kim, H.K.; Choi, Y.H.; Verpoorte, R. NMR-based plant metabolomics: Where do we stand, where do we go? *Trends Biotechnol.* **2011**, *29*, 267–275, doi:10.1016/j.tibtech.2011.02.001.
107. Kim, H.K.; Saifullah, Khan, S.; Wilson, E.G.; Kricun, S.D.P.; Meissner, A.; Goraler, S.; Deelder, A.M.; Choi, Y.H.; Verpoorte, R. Metabolic classification of South American *Ilex* species by NMR-based metabolomics. *Phytochemistry* **2010**, *71*, 773–784, doi:10.1016/j.phytochem.2010.02.001.
108. Mahrous, E.A.; Farag, M.A. Two dimensional NMR spectroscopic approaches for exploring plant metabolome: A review. *J. Adv. Res.* **2015**, *6*, 3–15, doi:10.1016/j.jare.2014.10.003.
109. Grkovic, T.; Pouwer, R.H.; Vial, M.-L.; Gambini, L.; Noël, A.; Hooper, J.N.A.; Wood, S.A.; Mellick, G.D.; Quinn, R.J. NMR Fingerprints of the Drug-like Natural-Product Space Identify Iotrochotazine A: A

- Chemical Probe to Study Parkinson's Disease. *Angew. Chem. Int. Ed.* **2014**, *53*, 6070–6074, doi:10.1002/anie.201402239.
110. Starks, C.M.; Williams, R.B.; Norman, V.L.; Lawrence, J.A.; O'Neil-Johnson, M.; Eldridge, G.R. Phenylpropanoids from *Phragmipedium calurum* and their antiproliferative activity. *Phytochemistry* **2012**, *82*, 172–175, doi:10.1016/j.phytochem.2012.06.017.
 111. Cai, S.-S.; Short, L.C.; Syage, J.A.; Potvin, M.; Curtis, J.M. Liquid chromatography–atmospheric pressure photoionization-mass spectrometry analysis of triacylglycerol lipids—Effects of mobile phases on sensitivity. *J. Chromatogr. A* **2007**, *1173*, 88–97, doi:10.1016/j.chroma.2007.10.008.
 112. Kim, H.K.; Choi, Y.H.; Verpoorte, R. NMR-based metabolomic analysis of plants. *Nat. Protoc.* **2010**, *5*, 536, doi:10.1038/nprot.2009.237.
 113. Tawfik, A.F.; Viegmann, C.; Edrada-Ebel, R. Metabolomics and Dereplication Strategies in Natural Products. In *Metabolomics Tools for Natural Product Discovery: Methods and Protocols*; Roessner, U., Dias, D.A., Eds.; Humana Press: Totowa, NJ, USA, 2013; pp. 227–244, doi:10.1007/978-1-62703-577-4_17.
 114. Zhang, L.; Hatzakis, E.; Patterson, A.D. NMR-Based Metabolomics and Its Application in Drug Metabolism and Cancer Research. *Curr. Pharmacol. Rep.* **2016**, *2*, 231–240, doi:10.1007/s40495-016-0067-9.
 115. Simmler, C.; Napolitano, J.G.; McAlpine, J.B.; Chen, S.-N.; Pauli, G.F. Universal quantitative NMR analysis of complex natural samples. *Curr. Opin. Biotechnol.* **2014**, *25*, 51–59, doi:10.1016/j.copbio.2013.08.004.
 116. Nagana Gowda, G.A.; Raftery, D. Can NMR solve some significant challenges in metabolomics? *J. Magn. Reson.* **2015**, *260*, 144–160, doi:10.1016/j.jmr.2015.07.014.
 117. Halabalaki, M.; Vougiannopoulou, K.; Mikros, E.; Skaltsounis, A.L. Recent advances and new strategies in the NMR-based identification of natural products. *Curr. Opin. Biotechnol.* **2014**, *25*, 1–7, doi:10.1016/j.copbio.2013.08.005.
 118. Kovacs, H.; Moskau, D.; Spraul, M. Cryogenically cooled probes—A leap in NMR technology. *Prog. Nucl. Magn. Reson. Spectrosc.* **2005**, *46*, 131–155, doi:10.1016/j.pnmrs.2005.03.001.
 119. Dona, A.C. CHAPTER 1 Instrumental Platforms for NMR-based Metabolomics. In *NMR-Based Metabolomics*; The Royal Society of Chemistry: London, UK, 2018; pp. 1–21, 10.1039/9781782627937-00001.
 120. Deborde, C.; Moing, A.; Roch, L.; Jacob, D.; Rolin, D.; Giraudeau, P. Plant metabolism as studied by NMR spectroscopy. *Prog. Nucl. Magn. Reson. Spectrosc.* **2017**, *102–103*, 61–97, doi:10.1016/j.pnmrs.2017.05.001.
 121. Anklin, C. Chapter 3 Small-volume NMR: Microprobes and Cryoprobes. In *Modern NMR Approaches to the Structure Elucidation of Natural Products: Volume 1: Instrumentation and Software*; The Royal Society of Chemistry: London, UK, 2016; Volume 1, pp. 38–57.
 122. Ardenkjær-Larsen, J.H.; Fridlund, B.; Gram, A.; Hansson, G.; Hansson, L.; Lerche, M.H.; Servin, R.; Thaning, M.; Golman, K. Increase in signal-to-noise ratio of >10,000 times in liquid-state NMR. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 10158–10163, doi:10.1073/pnas.1733835100.
 123. Forst, R.R.; Schroeder, F.C. NMR-spectroscopic analysis of mixtures: From structure to function. *Curr. Opin. Chem. Biol.* **2011**, *15*, 38–47, doi:10.1016/j.cbpa.2010.10.010.
 124. Novoa-Carballal, R.; Fernandez-Megia, E.; Jimenez, C.; Riguera, R. NMR methods for unravelling the spectra of complex mixtures. *Nat. Prod. Rep.* **2011**, *28*, 78–98, doi:10.1039/C005320C.
 125. Rabenstein, D.L.; Fan, S. Proton nuclear magnetic resonance spectroscopy of aqueous solutions: Complete elimination of the water resonance by spin-spin relaxation. *Anal. Chem.* **1986**, *58*, 3178–3184, doi:10.1021/ac00127a057.
 126. Johnson, C.S. Diffusion ordered nuclear magnetic resonance spectroscopy: Principles and applications. *Prog. Nucl. Magn. Reson. Spectrosc.* **1999**, *34*, 203–256, doi:10.1016/S0079-6565(99)00003-5.
 127. Morris, K.F.; Johnson, C.S. Diffusion-ordered two-dimensional nuclear magnetic resonance spectroscopy. *J. Am. Chem. Soc.* **1992**, *114*, 3139–3141, doi:10.1021/ja00034a071.
 128. Balayssac, S.; Trefi, S.; Gilard, V.; Malet-Martino, M.; Martino, R.; Delsuc, M.-A. 2D and 3D DOSY 1H NMR, a useful tool for analysis of complex mixtures: Application to herbal drugs or dietary supplements for erectile dysfunction. *J. Pharm. Biomed. Anal.* **2009**, *50*, 602–612, doi:10.1016/j.jpba.2008.10.034.
 129. Politi, M.; Zloh, M.; Pintado, M.E.; Castro, P.M.L.; Heinrich, M.; Prieto, J.M. Direct metabolic fingerprinting of commercial herbal tinctures by nuclear magnetic resonance spectroscopy and mass spectrometry. *Phytochem. Anal.* **2009**, *20*, 328–334, doi:10.1002/pca.1131.
 130. Tsuda, M.; Yasuda, T.; Fukushi, E.; Kawabata, J.; Sekiguchi, M.; Fromont, J.; Kobayashi, J.i. Agesamides A and B, Bromopyrrole Alkaloids from Sponge *Agelas* Species: Application of DOSY for Chemical Screening of New Metabolites. *Org. Lett.* **2006**, *8*, 4235–4238, doi:10.1021/ol061464q.

131. Ludwig, C.; Viant, M.R. Two-dimensional J-resolved NMR spectroscopy: Review of a key methodology in the metabolomics toolbox. *Phytochem. Anal.* **2010**, *21*, 22–32, doi:10.1002/pca.1186.
132. Georgiev, M.I.; Ali, K.; Alipieva, K.; Verpoorte, R.; Choi, Y.H. Metabolic differentiations and classification of Verbascum species by NMR-based metabolomics. *Phytochemistry* **2011**, *72*, 2045–2051, doi:10.1016/j.phytochem.2011.07.005.
133. Aguilar, J.A.; Faulkner, S.; Nilsson, M.; Morris, G.A. Pure Shift ¹H NMR: A Resolution of the Resolution Problem? *Angew. Chem.* **2010**, *122*, 3993–3995, doi:10.1002/ange.201001107.
134. Lopez, J.M.; Cabrera, R.; Maruenda, H. Ultra-Clean Pure Shift ¹H-NMR applied to metabolomics profiling. *Sci. Rep.* **2019**, *9*, 6900, doi:10.1038/s41598-019-43374-5.
135. Breton, R.C.; Reynolds, W.F. Using NMR to identify and characterize natural products. *Nat. Prod. Rep.* **2013**, *30*, 501–524, doi:10.1039/C2NP20104F.
136. Parkinson, J.A. Chapter 2—NMR Spectroscopy Methods in Metabolic Phenotyping. In *The Handbook of Metabolic Phenotyping*; Lindon, J.C., Nicholson, J.K., Holmes, E., Eds.; Elsevier: Amsterdam, The Netherlands, 2019; pp. 53–96, doi:10.1016/B978-0-12-812293-8.00002-5.
137. Koskela, H.; Heikkilä, O.; Kilpeläinen, I.; Heikkinen, S. Quantitative two-dimensional HSQC experiment for high magnetic field NMR spectrometers. *J. Magn. Reson.* **2010**, *202*, 24–33, doi:10.1016/j.jmr.2009.09.021.
138. Dossey, A.T.; Walse, S.S.; Conle, O.V.; Edison, A.S. Parectadial, a Monoterpenoid from the Defensive Spray of Parectatosoma mocquersyi. *J. Nat. Prod.* **2007**, *70*, 1335–1338, doi:10.1021/np070151g.
139. Farag, M.A.; Porzel, A.; Al-Hammady, M.A.; Hegazy, M.-E.F.; Meyer, A.; Mohamed, T.A.; Westphal, H.; Wessjohann, L.A. Soft Corals Biodiversity in the Egyptian Red Sea: A Comparative MS and NMR Metabolomics Approach of Wild and Aquarium Grown Species. *J. Proteome Res.* **2016**, *15*, 1274–1287, doi:10.1021/acs.jproteome.6b00002.
140. Farag, M.A.; Mahrous, E.A.; Lübken, T.; Porzel, A.; Wessjohann, L. Classification of commercial cultivars of Humulus lupulus L. (hop) by chemometric pixel analysis of two dimensional nuclear magnetic resonance spectra. *Metabolomics* **2014**, *10*, 21–32, doi:10.1007/s11306-013-0547-4.
141. Schroeder, F.C.; Gibson, D.M.; Churchill, A.C.L.; Sojikul, P.; Wursthorn, E.J.; Krasnoff, S.B.; Clardy, J. Differential Analysis of 2D NMR Spectra: New Natural Products from a Pilot-Scale Fungal Extract Library. *Angew. Chem. Int. Ed.* **2007**, *46*, 901–904, doi:10.1002/anie.200603821.
142. Pungalaya, C.; Srinivasan, J.; Fox, B.W.; Malik, R.U.; Ludewig, A.H.; Sternberg, P.W.; Schroeder, F.C. A shortcut to identifying small molecule signals that regulate behavior and development in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 7708–7713, doi:10.1073/pnas.0811918106.
143. Butcher, R.A.; Schroeder, F.C.; Fischbach, M.A.; Straight, P.D.; Kolter, R.; Walsh, C.T.; Clardy, J. The identification of bacillaene, the product of the PksX megacomplex in *Bacillus subtilis*. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 1506–1509, doi:10.1073/pnas.0610503104.
144. McCarney, E.R.; Dykstra, R.; Galvosas, P. Evaluation of benchtop NMR Diffusion Ordered Spectroscopy for small molecule mixture analysis. *Magn. Reson. Imaging* **2019**, *56*, 103–109, doi:10.1016/j.mri.2018.09.033.
145. Blümich, B. Low-field and benchtop NMR. *J. Magn. Reson.* **2019**, *306*, 27–35, doi:10.1016/j.jmr.2019.07.030.
146. Gouilleux, B.; Marchand, J.; Charrier, B.; Remaud, G.S.; Giraudeau, P. High-throughput authentication of edible oils with benchtop Ultrafast 2D NMR. *Food Chem.* **2018**, *244*, 153–158, doi:10.1016/j.foodchem.2017.10.016.
147. Kruk, J.; Dąbkowicz, M.; Jodłowska, E.; Zacharzewska, A.; Łakomiec, J.; Czaja, K.; Kujawski, J. NMR Techniques in Metabolomic Studies: A Quick Overview on Examples of Utilization. *Appl. Magn. Reson.* **2017**, *48*, 1–21, doi:10.1007/s00723-016-0846-9.
148. Pérez, E.M.S.; Iglesias, M.J.; Ortiz, F.L.; Pérez, I.S.; Galera, M.M. Study of the suitability of HRMAS NMR for metabolic profiling of tomatoes: Application to tissue differentiation and fruit ripening. *Food Chem.* **2010**, *122*, 877–887, doi:10.1016/j.foodchem.2010.03.003.
149. Daolio, C.; Beltrame, F.L.; Ferreira, A.G.; Cass, Q.B.; Cortez, D.A.G.; Ferreira, M.M.C. Classification of commercial Catuaba samples by NMR, HPLC and chemometrics. *Phytochem. Anal.* **2008**, *19*, 218–228, doi:10.1002/pca.1019.
150. Broberg, A.; Kenne, L. Use of High-Resolution Magic Angle Spinning Nuclear Magnetic Resonance Spectroscopy for in Situ Studies of Low-Molecular-Mass Compounds in Red Algae. *Anal. Biochem.* **2000**, *284*, 367–374, doi:10.1006/abio.2000.4722.
151. Miglietta, M.L.; Lamanna, R. ¹H HR-MAS NMR of carotenoids in aqueous samples and raw vegetables. *Magn. Reson. Chem.* **2006**, *44*, 675–685, doi:10.1002/mrc.1826.

152. Sekiyama, Y.; Chikayama, E.; Kikuchi, J. Profiling Polar and Semipolar Plant Metabolites throughout Extraction Processes Using a Combined Solution-State and High-Resolution Magic Angle Spinning NMR Approach. *Anal. Chem.* **2010**, *82*, 1643–1652, doi:10.1021/ac9019076.
153. Van der Graaf, M.; Heerschap, A. Effect of Cation Binding on the Proton Chemical Shifts and the Spin–Spin Coupling Constant of Citrate. *J. Magn. Reson. Ser. B* **1996**, *112*, 58–62, doi:10.1006/jmrb.1996.0109.
154. Jaroszewski, J.W. Hyphenated NMR Methods in Natural Products Research, Part 1: Direct Hyphenation. *Planta Med.* **2005**, *71*, 691–700, doi:10.1055/s-2005-871298.
155. Seger, C.; Sturm, S.; Stuppner, H. Mass spectrometry and NMR spectroscopy: Modern high-end detectors for high resolution separation techniques—State of the art in natural product HPLC-MS, HPLC-NMR, and CE-MS hyphenations. *Nat. Prod. Rep.* **2013**, *30*, 970–987, doi:10.1039/C3NP70015A.
156. Gebretsadik, T.; Linert, W.; Thomas, M.; Berhanu, T.; Frew, R. LC-NMR for Natural Products Analysis: A Journey from an Academic Curiosity to a Robust Analytical Tool. *Sci* **2019**, *1*, 31.
157. Keifer, P.A. Chemical-shift referencing and resolution stability in methanol:water gradient LC–NMR. *J. Magn. Reson.* **2010**, *205*, 130–140, doi:10.1016/j.jmr.2010.04.009.
158. Bringmann, G.; Wohlfarth, M.; Rischer, H.; Heubes, M.; Saeb, W.; Diem, S.; Herderich, M.; Schlauer, J. A Photometric Screening Method for Dimeric Naphthylisoquinoline Alkaloids and Complete On-Line Structural Elucidation of a Dimer in Crude Plant Extracts, by the LC–MS/LC–NMR/LC–CD Triad. *Anal. Chem.* **2001**, *73*, 2571–2577, doi:10.1021/ac001503q.
159. Waridel, P.; Wolfender, J.-L.; Lachavanne, J.-B.; Hostettmann, K. ent-Labdane glycosides from the aquatic plant *Potamogeton lucens* and analytical evaluation of the lipophilic extract constituents of various *Potamogeton* species. *Phytochemistry* **2004**, *65*, 945–954, doi:10.1016/j.phytochem.2004.01.018.
160. Zehl, M.; Braunberger, C.; Conrad, J.; Crnogorac, M.; Krasteva, S.; Vogler, B.; Beifuss, U.; Krenn, L. Identification and quantification of flavonoids and ellagic acid derivatives in therapeutically important *Drosera* species by LC–DAD, LC–NMR, NMR, and LC–MS. *Anal. Bioanal. Chem.* **2011**, *400*, 2565–2576, doi:10.1007/s00216-011-4690-3.
161. Clarkson, C.; Sibum, M.; Mensen, R.; Jaroszewski, J.W. Evaluation of on-line solid-phase extraction parameters for hyphenated, high-performance liquid chromatography–solid-phase extraction–nuclear magnetic resonance applications. *J. Chromatogr. A* **2007**, *1165*, 1–9, doi:10.1016/j.chroma.2007.07.016.
162. Griffiths, L.; Horton, R. Optimization of LC–NMR. III—Increased signal-to-noise ratio through column trapping. *Magn. Reson. Chem.* **1998**, *36*, 104–109, doi:10.1002/(sici)1097-458x(199802)36:2<104::aid-omr230>3.0.co;2-f.
163. Xu, Y.-J.; Capistrano, R.; Dhooghe, L.; Foubert, K.; Lemièr, F.; Maregesi, S.; Baldé, A.; Apers, S.; Pieters, L. Herbal Medicines and Infectious Diseases: Characterization by LC-SPE-NMR of Some Medicinal Plant Extracts Used against Malaria. *Planta Med.* **2011**, *77*, 1139–1148, doi:10.1055/s-0030-1270719.
164. Castro, A.; Moco, S.; Coll, J.; Vervoort, J. LC-MS-SPE-NMR for the Isolation and Characterization of neo-Clerodane Diterpenoids from *Teucrium luteum* subsp. *flavovirens*. *J. Nat. Prod.* **2010**, *73*, 962–965, doi:10.1021/np9005025.
165. Gao, H.; Zehl, M.; Kaehlig, H.; Schneider, P.; Stuppner, H.; Moreno, Y.; Banuls, L.; Kiss, R.; Kopp, B. Rapid Structural Identification of Cytotoxic Bufadienolide Sulfates in Toad Venom from *Bufo melanostictus* by LC-DAD-MSn and LC-SPE-NMR. *J. Nat. Prod.* **2010**, *73*, 603–608, doi:10.1021/np900746k.
166. Chen, C.-K.; Lin, F.-H.; Tseng, L.-H.; Jiang, C.-L.; Lee, S.-S. Comprehensive Study of Alkaloids from *Crinum asiaticum* var. *sinicum* Assisted by HPLC-DAD-SPE-NMR. *J. Nat. Prod.* **2011**, *74*, 411–419, doi:10.1021/np100819n.
167. Lai, Y.-C.; Chen, C.-K.; Lin, W.-W.; Lee, S.-S. A comprehensive investigation of anti-inflammatory diarylheptanoids from the leaves of *Alnus formosana*. *Phytochemistry* **2012**, *73*, 84–94, doi:10.1016/j.phytochem.2011.02.008.
168. Clarkson, C.; Hansen, S.H.; Smith, P.J.; Jaroszewski, J.W. Identification of Major and Minor Constituents of *Harpagophytum procumbens* (Devil’s Claw) Using HPLC-SPE-NMR and HPLC-ESIMS/APCIMS. *J. Nat. Prod.* **2006**, *69*, 1280–1288, doi:10.1021/np0601612.
169. Nyberg, N.T.; Baumann, H.; Kenne, L. Solid-Phase Extraction NMR Studies of Chromatographic Fractions of Saponins from *Quillaja saponaria*. *Anal. Chem.* **2003**, *75*, 268–274, doi:10.1021/ac025725v.
170. Kuhlisch, C.; Pohnert, G. Metabolomics in chemical ecology. *Nat. Prod. Rep.* **2015**, *32*, doi:10.1039/c5np00003c.

171. Kongstad, K.T.; Özdemir, C.; Barzak, A.; Wubshet, S.G.; Staerk, D. Combined Use of High-Resolution α -Glucosidase Inhibition Profiling and High-Performance Liquid Chromatography–High-Resolution Mass Spectrometry–Solid-Phase Extraction–Nuclear Magnetic Resonance Spectroscopy for Investigation of Antidiabetic Principles in Crude Plant Extracts. *J. Agric. Food Chem.* **2015**, *63*, 2257–2263, doi:10.1021/jf506297k.
172. Agnolet, S.; Jaroszewski, J.W.; Verpoorte, R.; Staerk, D. 1H NMR-based metabolomics combined with HPLC-PDA-MS-SPE-NMR for investigation of standardized Ginkgo biloba preparations. *Metabolomics* **2010**, *6*, 292–302, doi:10.1007/s11306-009-0195-x.
173. Tang, H.; Xiao, C.; Wang, Y. Important roles of the hyphenated HPLC-DAD-MS-SPE-NMR technique in metabonomics. *Magn. Reson. Chem.* **2009**, *47*, S157–S162, doi:10.1002/mrc.2513.
174. Staerk, D.; Kesting, J.R.; Sairafianpour, M.; Witt, M.; Asili, J.; Emami, S.A.; Jaroszewski, J.W. Accelerated dereplication of crude extracts using HPLC–PDA–MS–SPE–NMR: Quinolinone alkaloids of Haplophyllum acutifolium. *Phytochemistry* **2009**, *70*, 1055–1061, doi:10.1016/j.phytochem.2009.05.004.
175. Lima, R.d.C.L.; Gramsbergen, S.M.; Van Staden, J.; Jäger, A.K.; Kongstad, K.T.; Staerk, D. Advancing HPLC-PDA-HRMS-SPE-NMR Analysis of Coumarins in Coleonema album by Use of Orthogonal Reversed-Phase C18 and Pentafluorophenyl Separations. *J. Nat. Prod.* **2017**, *80*, 1020–1027, doi:10.1021/acs.jnatprod.6b01020.
176. Liu, B.; Kongstad, K.T.; Qinglei, S.; Nyberg, N.T.; Jäger, A.K.; Staerk, D. Dual High-Resolution α -Glucosidase and Radical Scavenging Profiling Combined with HPLC-HRMS-SPE-NMR for Identification of Minor and Major Constituents Directly from the Crude Extract of Pueraria lobata. *J. Nat. Prod.* **2015**, *78*, 294–300, doi:10.1021/np5009416.
177. Luedemann, A.; Strassburg, K.; Erban, A.; Kopka, J. TagFinder for the quantitative analysis of gas chromatography–Mass spectrometry (GC-MS)-based metabolite profiling experiments. *Bioinformatics* **2008**, *24*, 732–737, doi:10.1093/bioinformatics/btn023.
178. Cuadros-Inostroza, Á.; Caldana, C.; Redestig, H.; Kusano, M.; Lisec, J.; Peña-Cortés, H.; Willmitzer, L.; Hannah, M.A. TargetSearch—A Bioconductor package for the efficient preprocessing of GC-MS metabolite profiling data. *BMC Bioinform.* **2009**, *10*, 428, doi:10.1186/1471-2105-10-428.
179. 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, doi:10.1038/nmeth.3393.
180. Matsuo, T.; Tsugawa, H.; Miyagawa, H.; Fukusaki, E. Integrated Strategy for Unknown EI-MS Identification Using Quality Control Calibration Curve, Multivariate Analysis, EI-MS Spectral Database, and Retention Index Prediction. *Anal. Chem.* **2017**, *89*, 6766–6773, doi:10.1021/acs.analchem.7b01010.
181. Beale, D.J.; Pinu, F.R.; Kouremenos, K.A.; Poojary, M.M.; Narayana, V.K.; Boughton, B.A.; Kanojia, K.; Dayalan, S.; Jones, O.A.H.; Dias, D.A. Review of recent developments in GC-MS approaches to metabolomics-based research. *Metabolomics* **2018**, *14*, 152, doi:10.1007/s11306-018-1449-2.
182. Fiehn, O. Metabolomics by Gas Chromatography–Mass Spectrometry: Combined Targeted and Untargeted Profiling. *Curr. Protoc. Mol. Biol.* **2016**, *114*, 30.4.1–30.4.32, doi:10.1002/0471142727.mb3004s114.
183. 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, doi:10.1021/ac051437y.
184. Pluskal, T.; Castillo, S.; Villar-Briones, A.; Orešič, M. MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinform.* **2010**, *11*, 395, doi:10.1186/1471-2105-11-395.
185. Röst, H.L.; Sachsenberg, T.; Aiche, S.; Bielow, C.; Weissner, H.; Aicheler, F.; Andreotti, S.; Ehrlich, H.-C.; Gutenbrunner, P.; Kenar, E.; et al. OpenMS: A flexible open-source software platform for mass spectrometry data analysis. *Nat. Methods* **2016**, *13*, 741–748, doi:10.1038/nmeth.3959.
186. Schiffman, C.; Petrick, L.; Perttula, K.; Yano, Y.; Carlsson, H.; Whitehead, T.; Metayer, C.; Hayes, J.; Rappaport, S.; Dudoit, S. Filtering procedures for untargeted LC-MS metabolomics data. *BMC Bioinform.* **2019**, *20*, 334, doi:10.1186/s12859-019-2871-9.
187. Nothias, L.-F.; Nothias-Esposito, M.; da Silva, R.; Wang, M.; Protsyuk, I.; Zhang, Z.; Sarvepalli, A.; Leyssen, P.; Touboul, D.; Costa, J.; et al. Bioactivity-Based Molecular Networking for the Discovery of Drug Leads in Natural Product Bioassay-Guided Fractionation. *J. Nat. Prod.* **2018**, *81*, 758–767, doi:10.1021/acs.jnatprod.7b00737.

188. Kuhl, C.; Tautenhahn, R.; Böttcher, C.; Larson, T.R.; Neumann, S. CAMERA: An Integrated Strategy for Compound Spectra Extraction and Annotation of Liquid Chromatography/Mass Spectrometry Data Sets. *Anal. Chem.* **2012**, *84*, 283–289, doi:10.1021/ac202450g.
189. Li, H.; Cai, Y.; Guo, Y.; Chen, F.; Zhu, Z.-J. MetDIA: Targeted Metabolite Extraction of Multiplexed MS/MS Spectra Generated by Data-Independent Acquisition. *Anal. Chem.* **2016**, *88*, 8757–8764, doi:10.1021/acs.analchem.6b02122.
190. Wolfender, J.-L.; Nuzillard, J.-M.; van der Hooft, J.J.J.; Renault, J.-H.; Bertrand, S. Accelerating Metabolite Identification in Natural Product Research: Toward an Ideal Combination of Liquid Chromatography–High-Resolution Tandem Mass Spectrometry and NMR Profiling, in Silico Databases, and Chemometrics. *Anal. Chem.* **2019**, *91*, 704–742, doi:10.1021/acs.analchem.8b05112.
191. 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, doi:10.1021/acs.analchem.7b04424.
192. Horai, H.; Arita, M.; Kanaya, S.; Nihei, Y.; Ikeda, T.; Suwa, K.; Ojima, Y.; Tanaka, K.; Tanaka, S.; Aoshima, K.; et al. MassBank: A public repository for sharing mass spectral data for life sciences. *J. Mass Spectrom.* **2010**, *45*, 703–714, doi:10.1002/jms.1777.
193. Wishart, D.S.; Feunang, Y.D.; Marcu, A.; Guo, A.C.; Liang, K.; Vázquez-Fresno, R.; Sajed, T.; Johnson, D.; Li, C.; Karu, N.; et al. HMDB 4.0: The human metabolome database for 2018. *Nucleic Acids Res.* **2017**, *46*, D608–D617, doi:10.1093/nar/gkx1089.
194. Scalbert, A.; Brennan, L.; Manach, C.; Andres-Lacueva, C.; Dragsted, L.O.; Draper, J.; Rappaport, S.M.; van der Hooft, J.J.; Wishart, D.S. The food metabolome: A window over dietary exposure. *Am. J. Clin. Nutr.* **2014**, *99*, 1286–1308, doi:10.3945/ajcn.113.076133.
195. Blaženović, I.; Kind, T.; Torbašinović, H.; Obrenović, S.; Mehta, S.S.; Tsugawa, H.; Wermuth, T.; Schauer, N.; Jahn, M.; Biedendieck, R.; et al. Comprehensive comparison of in silico MS/MS fragmentation tools of the CASMI contest: Database boosting is needed to achieve 93% accuracy. *J. Cheminformatics* **2017**, *9*, 32, doi:10.1186/s13321-017-0219-x.
196. Dührkop, K.; Shen, H.; Meusel, M.; Rousu, J.; Böcker, S. Searching molecular structure databases with tandem mass spectra using CSI:FingerID. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 12580–12585, doi:10.1073/pnas.1509788112.
197. Tsugawa, H.; Kind, T.; Nakabayashi, R.; Yukihiro, D.; Tanaka, W.; Cajka, T.; Saito, K.; Fiehn, O.; Arita, M. Hydrogen Rearrangement Rules: Computational MS/MS Fragmentation and Structure Elucidation Using MS-FINDER Software. *Anal. Chem.* **2016**, *88*, 7946–7958, doi:10.1021/acs.analchem.6b00770.
198. Stanstrup, J.; Neumann, S.; Vrhovšek, U. PredRet: Prediction of Retention Time by Direct Mapping between Multiple Chromatographic Systems. *Anal. Chem.* **2015**, *87*, 9421–9428, doi:10.1021/acs.analchem.5b02287.
199. Afendi, F.M.; Okada, T.; Yamazaki, M.; Hirai-Morita, A.; Nakamura, Y.; Nakamura, K.; Ikeda, S.; Takahashi, H.; Altaf-Ul-Amin, M.; Darusman, L.K.; et al. KnapSack Family Databases: Integrated Metabolite–Plant Species Databases for Multifaceted Plant Research. *Plant Cell Physiol.* **2011**, *53*, e1, doi:10.1093/pcp/pcr165.
200. Murray, K.K.; Boyd, R.K.; Eberlin, M.N.; Langley, G.J.; Li, L.; Naito, Y. Definitions of terms relating to mass spectrometry (IUPAC Recommendations 2013). *Pure Appl. Chem.* **2013**, *85*, 1515–1609.
201. Smith, C.; 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.
202. Sawada, Y.; Nakabayashi, R.; Yamada, Y.; Suzuki, M.; Sato, M.; Sakata, A.; Akiyama, K.; Sakurai, T.; Matsuda, F.; Aoki, T. RIKEN tandem mass spectral database (ReSpect) for phytochemicals: A plant-specific MS/MS-based data resource and database. *Phytochemistry* **2012**, *82*, 38–45.
203. Mohimani, H.; Gurevich, A.; Mikheenko, A.; Garg, N.; Nothias, L.F.; Ninomiya, A.; Takada, K.; Dorrestein, P.C.; Pevzner, P.A. Dereplication of peptidic natural products through database search of mass spectra. *Nat. Chem. Biol.* **2017**, *13*, 30–37, doi:10.1038/nchembio.2219.
204. Pan, Z.; Raftery, D. Comparing and combining NMR spectroscopy and mass spectrometry in metabolomics. *Anal. Bioanal. Chem.* **2007**, *387*, 525–527.
205. Gowda, G.N.; 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.
206. Marshall, D.D.; Powers, R. Beyond the paradigm: Combining mass spectrometry and nuclear magnetic resonance for metabolomics. *Prog. Nucl. Magn. Reson. Spectrosc.* **2017**, *100*, 1–16.

207. Lewis, I.A.; Schommer, S.C.; Markley, J.L. rNMR: Open source software for identifying and quantifying metabolites in NMR spectra. *Magn. Reson. Chem.* **2009**, *47*, S123–S126.
208. Chikayama, E.; Sekiyama, Y.; Okamoto, M.; Nakanishi, Y.; Tsuboi, Y.; Akiyama, K.; Saito, K.; Shinozaki, K.; Kikuchi, J. Statistical indices for simultaneous large-scale metabolite detections for a single NMR spectrum. *Anal. Chem.* **2010**, *82*, 1653–1658.
209. Bingol, K.; Bruschweiler-Li, L.; Li, D.-W.; Bruschweiler, R. Customized metabolomics database for the analysis of NMR 1H–1H TOCSY and 13C–1H HSQC-TOCSY spectra of complex mixtures. *Anal. Chem.* **2014**, *86*, 5494–5501.
210. Bingol, K.; Zhang, F.; Bruschweiler-Li, L.; Bruschweiler, R. TOCCATA: A customized carbon total correlation spectroscopy NMR metabolomics database. *Anal. Chem.* **2012**, *84*, 9395–9401.
211. López-Pérez, J.L.; Therón, R.; del Olmo, E.; Díaz, D. NAPROC-13: A database for the dereplication of natural product mixtures in bioassay-guided protocols. *Bioinformatics* **2007**, *23*, 3256–3257.
212. Steinbeck, C.; Krause, S.; Kuhn, S. NMRShiftDB constructing a free chemical information system with open-source components. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 1733–1739.
213. Steinbeck, C.; Kuhn, S. NMRShiftDB—compound identification and structure elucidation support through a free community-built web database. *Phytochemistry* **2004**, *65*, 2711–2717.
214. Yamamoto, O.; Someno, K.; Wasada, N.; Hiraishi, J.; Hayamizu, K.; Tanabe, K.; Tamura, T.; Yanagisawa, M. An integrated spectral data base system including IR, MS, 1H-NMR, 13C-NMR, ESR and Raman spectra. *Anal. Sci.* **1988**, *4*, 233–239.
215. Ludwig, C.; Easton, J.M.; Lodi, A.; Tiziani, S.; Manzoor, S.E.; Southam, A.D.; Byrne, J.J.; Bishop, L.M.; He, S.; Arvanitis, T.N. Birmingham metabolite library: A publicly accessible database of 1-D ¹H and 2-D ¹H J-resolved NMR spectra of authentic metabolite standards (BML-NMR). *Metabolomics* **2012**, *8*, 8–18.
216. Ulrich, E.; Akutsu, H.; Doreleijers, J.; Harano, Y.; Ioannidis, Y.; Lin, J.; Livny, M.; Mading, S.; Maziuk, D.; Miller, Z. BioMagResBank. *Nucleic Acids Res.* **2008**, *36*, D402–D408.
217. Yanshole, V.V.; Snytnikova, O.A.; Kiryutin, A.S.; Yanshole, L.V.; Sagdeev, R.Z.; Tsentalovich, Y.P. Metabolomics of the rat lens: A combined LC-MS and NMR study. *Exp. Eye Res.* **2014**, *125*, 71–78.
218. Bingol, K.; Bruschweiler-Li, L.; Yu, C.; Somogyi, A.; Zhang, F.; Bruschweiler, R. Metabolomics beyond spectroscopic databases: A combined MS/NMR strategy for the rapid identification of new metabolites in complex mixtures. *Anal. Chem.* **2015**, *87*, 3864–3870.
219. Bingol, K.; Bruschweiler, R. Two elephants in the room: New hybrid nuclear magnetic resonance and mass spectrometry approaches for metabolomics. *Curr. Opin. Clin. Nutr. Metab. Care* **2015**, *18*, 471.
220. Wohlleben, W.; Mast, Y.; Muth, G.; Röttgen, M.; Stegmann, E.; Weber, T. Synthetic biology of secondary metabolite biosynthesis in actinomycetes: Engineering precursor supply as a way to optimize antibiotic production. *FEBS Lett.* **2012**, *586*, 2171–2176.
221. Jankevics, A.; Merlo, M.E.; de Vries, M.; Vonk, R.J.; Takano, E.; Breitling, R. Metabolomic analysis of a synthetic metabolic switch in *Streptomyces coelicolor* A3 (2). *Proteomics* **2011**, *11*, 4622–4631.
222. Breitling, R.; Achcar, F.; Takano, E. Modeling challenges in the synthetic biology of secondary metabolism. *ACS Synth. Biol.* **2013**, *2*, 373–378.
223. Cheng, J.-S.; Liang, Y.-Q.; Ding, M.-Z.; Cui, S.-F.; Lv, X.-M.; Yuan, Y.-J. Metabolic analysis reveals the amino acid responses of *Streptomyces lydicus* to pitching ratios during improving streptolydigin production. *Appl. Microbiol. Biotechnol.* **2013**, *97*, 5943–5954.
224. Arendt, P.; Pollier, J.; Callewaert, N.; Goossens, A. Synthetic biology for production of natural and new-to-nature terpenoids in photosynthetic organisms. *Plant J. Cell Mol. Biol.* **2016**, *87*, 16–37, doi:10.1111/tpj.13138.
225. Julsing, M.K.; Koulman, A.; Woerdenbag, H.J.; Quax, W.J.; Kayser, O. Combinatorial biosynthesis of medicinal plant secondary metabolites. *Biomol. Eng.* **2006**, *23*, 265–279, doi:10.1016/j.bioeng.2006.08.001.
226. Pollier, J.; Moses, T.; Goossens, A. Combinatorial biosynthesis in plants: A (p)review on its potential and future exploitation. *Nat. Prod. Rep.* **2011**, *28*, 1897–1916, doi:10.1039/c1np00049g.
227. Umeno, D.; Arnold, F.H. A C35 carotenoid biosynthetic pathway. *Appl. Environ. Microbiol.* **2003**, *69*, 3573–3579, doi:10.1128/aem.69.6.3573-3579.2003.
228. Rungtaphan, W.; Qu, X.; O'Connor, S.E. Integrating carbon-halogen bond formation into medicinal plant metabolism. *Nature* **2010**, *468*, 461–464, doi:10.1038/nature09524.
229. Moses, T.; Papadopoulou, K.K.; Osbourn, A. Metabolic and functional diversity of saponins, biosynthetic intermediates and semi-synthetic derivatives. *Crit. Rev. Biochem. Mol. Biol.* **2014**, *49*, 439–462, doi:10.3109/10409238.2014.953628.

230. Fuentes, P.; Zhou, F.; Erban, A.; Karcher, D.; Kopka, J.; Bock, R. A new synthetic biology approach allows transfer of an entire metabolic pathway from a medicinal plant to a biomass crop. *eLife* **2016**, *5*, e13664, doi:10.7554/eLife.13664.
231. Hiller, K.; Metallo, C.M.; Kelleher, J.K.; Stephanopoulos, G. Nontargeted elucidation of metabolic pathways using stable-isotope tracers and mass spectrometry. *Anal. Chem.* **2010**, *82*, 6621–6628.
232. Creek, D.J.; Chokkathukalam, A.; Jankevics, A.; Burgess, K.E.; Breitling, R.; Barrett, M.P. Stable isotope-assisted metabolomics for network-wide metabolic pathway elucidation. *Anal. Chem.* **2012**, *84*, 8442–8447.
233. Bueschl, C.; Krska, R.; Kluger, B.; Schuhmacher, R. Isotopic labeling-assisted metabolomics using LC–MS. *Anal. Bioanal. Chem.* **2013**, *405*, 27–33.
234. Ellis, D.I.; Goodacre, R. Metabolomics-assisted synthetic biology. *Curr. Opin. Biotechnol.* **2012**, *23*, 22–28.
235. Moldenhauer, J.; Chen, X.H.; Borriss, R.; Piel, J. Biosynthesis of the antibiotic bacillaene, the product of a giant polyketide synthase complex of the trans-AT family. *Angew. Chem. Int. Ed.* **2007**, *46*, 8195–8197.
236. Blunt, J.W.; Munro, M.H. Data, 1 H-NMR databases, data manipulation. *Phytochem. Rev.* **2013**, *12*, 435–447.
237. Buckingham, J. *Dictionary of Natural Products on DVD*; Version 21: 1; Chapman & Hall/CRC Press: Boca Raton, FL, USA, 2012.
238. Purves, K.; Macintyre, L.; Brennan, D.; Hreggviðsson, G.; Kuttner, E.; Ásgeirsdóttir, M.; Young, L.; Green, D.; Edrada-Ebel, R.; Duncan, K. Using molecular networking for microbial secondary metabolite bioprospecting. *Metabolites* **2016**, *6*, 2.
239. Funasaki, M.; Menezes, I.S.; dos Santos BARROSO, H.; Zanotto, S.P.; Carioca, C.R.F. Tocopherol profile of Brazil nut oil from different geographic areas of the Amazon region. *Acta Amaz.* **2012**, *43*.
240. Borges, C.V.; de Oliveira Amorim, V.B.; Ramlov, F.; da Silva Ledo, C.A.; Donato, M.; Maraschin, M.; Amorim, E.P. Characterisation of metabolic profile of banana genotypes, aiming at biofortified *Musa* spp. cultivars. *Food Chem.* **2014**, *145*, 496–504.
241. da Silva Taveira, J.H.; Borém, F.M.; Figueiredo, L.P.; Reis, N.; Franca, A.S.; Harding, S.A.; Tsai, C.-J. Potential markers of coffee genotypes grown in different Brazilian regions: A metabolomics approach. *Food Res. Int.* **2014**, *61*, 75–82.
242. Pilon, A.C.; Carnevale Neto, F.; Freire, R.T.; Cardoso, P.; Carneiro, R.L.; Da Silva Bolzani, V.; Castro-Gamboa, I. Partial least squares model and design of experiments toward the analysis of the metabolome of *Jatropha gossypifolia* leaves: Extraction and chromatographic fingerprint optimization. *J. Sep. Sci.* **2016**, *39*, 1023–1030.
243. Silva, M.M.; Bergamasco, J.; Lira, S.P.; Lopes, N.P.; Hajdu, E.; Peixinho, S.; Berlinck, R.G. Dereplication of bromotyrosine-derived metabolites by LC-PDA-MS and analysis of the chemical profile of 14 aplysina sponge specimens from the Brazilian coastline. *Aust. J. Chem.* **2010**, *63*, 886–894.
244. Bittencourt, M.L.; Ribeiro, P.R.; Franco, R.L.; Hilhorst, H.W.; de Castro, R.D.; Fernandez, L.G. Metabolite profiling, antioxidant and antibacterial activities of Brazilian propolis: Use of correlation and multivariate analyses to identify potential bioactive compounds. *Food Res. Int.* **2015**, *76*, 449–457.
245. Castro, C.B.; Pires, D.O. Brazilian coral reefs: What we already know and what is still missing. *Bull. Mar. Sci.* **2001**, *69*, 357–371.
246. Funari, C.S.; Eugster, P.J.; Martel, S.; Carrupt, P.-A.; Wolfender, J.-L.; Silva, D.H.S. High resolution ultra high pressure liquid chromatography–time-of-flight mass spectrometry dereplication strategy for the metabolite profiling of Brazilian *Lippia* species. *J. Chromatogr. A* **2012**, *1259*, 167–178.
247. Chagas-Paula, D.A.; Oliveira, T.B.; Zhang, T.; Edrada-Ebel, R.; Da Costa, F.B. Prediction of anti-inflammatory plants and discovery of their biomarkers by machine learning algorithms and metabolomic studies. *Planta Med.* **2015**, *81*, 450–458.
248. Wang, Y.; Tang, H.; Nicholson, J.K.; Hylands, P.J.; Sampson, J.; Whitcombe, I.; Stewart, C.G.; Caiger, S.; Oru, I.; Holmes, E. Metabolomic strategy for the classification and quality control of phytomedicine: A case study of chamomile flower (*Matricaria recutita* L.). *Planta Med.* **2004**, *70*, 250–255.
249. Choi, Y.H.; Kim, H.K.; Hazekamp, A.; Erkelens, C.; Lefeber, A.W.; Verpoorte, R. Metabolomic Differentiation of Cannabis sativa Cultivars Using 1H NMR Spectroscopy and Principal Component Analysis. *J. Nat. Prod.* **2004**, *67*, 953–957.
250. Rasmussen, B.; Cloarec, O.; Tang, H.; Stærk, D.; Jaroszewski, J.W. Multivariate analysis of integrated and full-resolution 1H-NMR spectral data from complex pharmaceutical preparations: St. John's wort. *Planta Med.* **2006**, *72*, 556–563.

251. Van der Kooy, F.; Verpoorte, R.; Meyer, J.M. Metabolomic quality control of claimed anti-malarial *Artemisia afra* herbal remedy and *A. afra* and *A. annua* plant extracts. *S. Afr. J. Bot.* **2008**, *74*, 186–189.
252. Johnson, C.H.; Ivanisevic, J.; Siuzdak, G. Metabolomics: Beyond biomarkers and towards mechanisms. *Nat. Rev. Mol. Cell Biol.* **2016**, *17*, 451–459, doi:10.1038/nrm.2016.25.
253. Zhao, Q.; Zhang, J.-L.; Li, F. Application of Metabolomics in the Study of Natural Products. *Nat. Prod. Bioprospecting* **2018**, *8*, 1–14, doi:10.1007/s13659-018-0175-9.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).