

Metabolomics in Fundamental Plant Research



José Juan Ordaz-Ortiz, Anita Arroyo-Silva, and Moisés Guerrero-Esperanza

Abbreviations

[M + H] ⁺	Protonated ion
¹ H NMR	Proton nuclear magnetic resonance
2D-NMR	Bidimensional nuclear magnetic resonance
CE-MS	Capillary electrophoresis-mass spectrometry
cmQTL	Canalization metabolite quantitative trait loci
Da	Dalton
DART-MS	Direct analysis in real time-mass spectrometry
DEGs	Differentially expressed genes
DESI	Desorption electrospray ionization
FACS	Fluorescence-activated cell sorting
FT-ICR-MS	Fourier transform ion cyclotron resonance-mass spectrometry
GC	Gas chromatography
GC-MS	Gas chromatography-mass spectrometry
GC-QqQ-MS	Gas chromatography-triple quadrupole-mass spectrometry
GC-TOF-MS	Gas chromatography-time-of-flight mass spectrometry
GLC	Gas-liquid chromatography
HPLC	High-performance liquid chromatography
LC	Liquid chromatography
LC-ESI-MS	Liquid chromatography-electrospray-mass spectrometry
LC-HRMS	Liquid chromatography-high-resolution mass spectrometry
LCM	Laser-capture microdissection
LC-MS	Liquid chromatography-mass spectrometry
LMD	Laser microdissection
LMPC	Laser microdissection and pressure catapulting
m/z	Mass-to-charge ratio

J. J. Ordaz-Ortiz (✉) · A. Arroyo-Silva · M. Guerrero-Esperanza
Metabolomics and Mass Spectrometry Laboratory, Unidad de Genómica Avanzada,
Irapuato, Gto, Mexico
e-mail: jose.ordaz.ortiz@cinvestav.mx

MALDI	Matrix-assisted laser desorption/ionization
mGWAS	Metabolite genome-wide association studies
mQTL	Metabolite quantitative trait loci
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry, fragmentation
MSI	Mass spectrometry imaging
NMR	Nuclear magnetic resonance
PCA	Principal component analysis
PCR	Polymerase chain reaction
PLS-DA	Partial least squares discriminant analysis
SIMS	Secondary ion mass spectrometry
SNPs	Single-nucleotide polymorphisms
UHPLC-MS	Ultrahigh-pressure liquid chromatography-mass spectrometry
UPLC	Ultra-performance liquid chromatography
VOC	Volatile organic compounds
WGCNA	Correlation and weighted gene coexpression network analysis

1 Introduction

Plant metabolomics is an emerging research field focusing on the comprehensive analysis of metabolites, small molecules (<1500 Da) that play a critical role in plant growth, development, and response to environmental changes. By applying advanced analytical techniques, such as mass spectrometry and nuclear magnetic resonance spectroscopy, researchers can identify and quantify thousands of metabolites in a single sample, providing a holistic view of the metabolic pathways and processes that occur in plants. The insights gained from plant metabolomics research have broad implications for plant biology, agriculture, and biotechnology, as well as for human health and the environment.

Metabolomics is crucial to studying abiotic stress tolerance, pathogen resistance, robust ecotypes, and metabolic-assisted breeding of crops. The plant kingdom contains a considerable diversity of metabolites of approximately 200,000 compounds; the majority are still unknown. It is estimated that around 10,000 secondary metabolites have been discovered in different plant species [1], since metabolites have a wide range of physicochemical properties and functions. This makes it challenging for metabolomics techniques to study their diversity and gain insights into plant biology.

Since the word “metabolomics” was mentioned for the first time in the literature, it has evolved and been applied to many disciplines, such as plant biology. Plant metabolomics has become hugely modernized in the last decade. This chapter describes novel applications of mass spectrometry-based metabolomics approaches in recent fundamental plant research. We first commented on different metabolomics techniques, mass spectrometry, and nuclear magnetic resonance-based metabolomics. Mass spectrometry-based metabolomics is the widest in use; the main

approaches are targeted or quantitative metabolomics and untargeted, nontargeted, or discovery metabolomics, also known as global metabolomics. We then explored contemporary literature on gene identification and their functional characterization for crop improvement enabled by applying nontargeted and targeted metabolomic analysis in combination with genome-wide association studies, metabolite quantitative trait loci, and transcriptomics.

From there, we explored the current application of metabolomics methods for plant species identification, a trendy topic that is enabling researchers to support many aspects of food authentication, food quality control, and traceability; plant species identification is an essential factor in understanding biodiversity, the discovery of bioactives from herbal medicines and correlating chemical components from plants with chemical markers of patients who intakes herbal medication, in the same manner monitoring food intake in foodomics studies.

We exemplified by commenting on the use of molecular networking analysis and its application to classify plant species; by providing an example of the Malpighiaceae family, chemotaxonomic studies guided by metabolomics methods become hugely in use as it allows rapid classification of plant samples based on the endogenous chemical content. We also discussed pioneering work on classifying and discriminating cinnamon, vanilla, and coffee plant species using different metabolomics techniques.

We then explored the impact of climate change on the root metabolome and the differences in root abiotic and metabolic changes associated with other biotic factors interactions, highlighting key metabolites involved in root exudates when exposed to these types of stresses. In addition, we reviewed recent literature on plant biomarker discovery outlining different application areas, such as the food industry, where the identification of biomarkers has also worked in quality processes for food authenticity and food traceability matrices of plant origin. We outlined a list of key metabolites identified in various plant species using different analytical techniques including metabolites detected in transgenic plants.

Finally, we briefly explored the emerging field of single-cell metabolomics methods. We describe the latest development in mass spectrometry imaging, including different approaches for collecting single-cell from plant tissues and a revision of some essential techniques on mass spectrometry imaging. Mention has also been made on the challenging and future needs for plant metabolomics research.

2 Novel Gene Identifications and Their Functional Characterization for Crop Improvement

Metabolomics studies all small molecules – metabolites – content of a cell or whole organism. Plant metabolomics refers to comprehensive, nonbiased, high-throughput analyses of complex metabolite mixtures typical of plant extracts. The role of metabolomics in such studies is twofold: (1) to identify the spatial and temporal

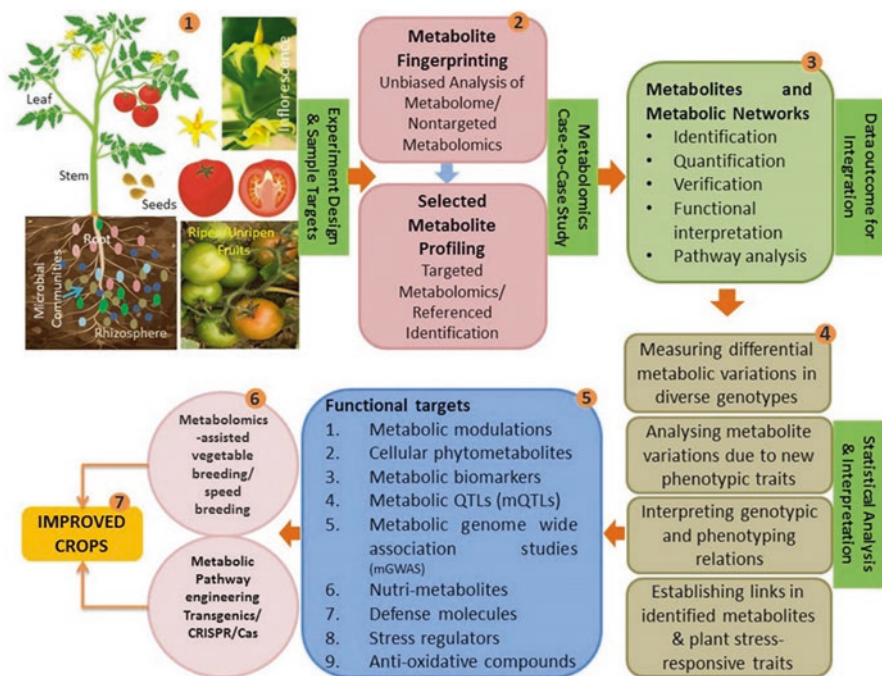


Fig. 1 Schematic presentation of metabolomics applications in crop improvement programs: (1) A representative sampling source of vegetable crop plants (tomato) as a biological source from which cellular metabolome can be extracted from almost all the plant parts and the rhizosphere under varying experimental environmental conditions; (2) data acquisition approach in metabolomics to be applied whether unbiased nontargeted fingerprinting is required or the analysis and quantification of a few selected target molecules is the need of the experiment; (3) study outcomes which needs biological interpretation for hypothesis questions; (4) possible answers to the hypothesis questions in the cellular chemistry and its entwining relations with the environmental impacts; (5) functional targets that could be achieved through metabolomics analysis of the vegetable plants; (6) result-oriented applications of the data outcomes in crop improvement practices; and (7) the “end product” of the experimental metabolomics exercise in vegetable crops. (Reproduced from Ref. [2])

distribution of the target compounds as influenced by plant development and environmental cues and (2) to identify related phytochemicals, which may be considered as either intermediate of biosynthesis or alternative or alternative products of promiscuous enzymes that support the biosynthesis of the target phytochemical [1].

This chapter describes the latest development and application of plant metabolomics in combination with metabolite genome-wide association studies (mGWAS), metabolite quantitative trait loci (mQTL), and transcriptomics for the discovery and characterization of genes and enzymes associated with the biosynthesis of specialized metabolites in significant crops such as maize, rice, and tomato (Fig. 1).

Different metabolomics techniques have been developed in the last two decades, including mass spectrometry-based metabolomics (MS); nuclear magnetic

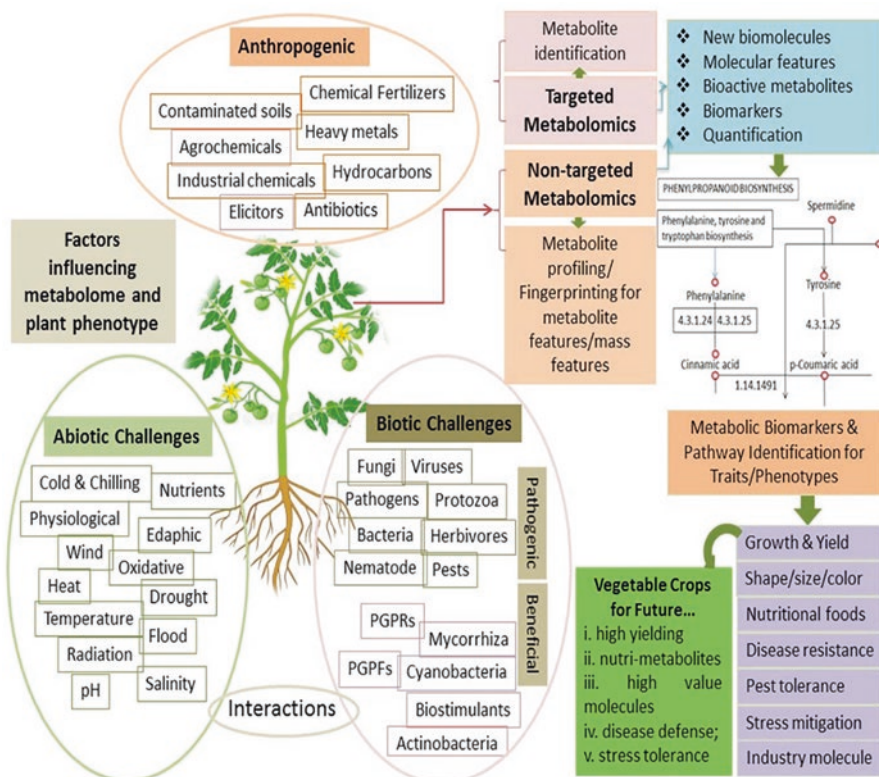


Fig. 2 Illustrative diagram of possible plant environmental interactions, which are supposed to influence the metabolic status of the crop plants. Analyzing the metabolome of plants exposed to such challenges using metabolomics approaches can yield competitive vegetable crops with better yield, high level of defense and stress-mitigating capabilities. (Reproduced from Ref. [2])

resonance spectroscopy (NMR); gas-chromatography-mass spectrometry (GC-MS); capillary electrophoresis-mass spectrometry (CE-MS); liquid chromatography-mass spectrometry (LC-MS); and more recently the implementation of high-resolution metabolomics with the aid of Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS). Mass spectrometry-based metabolomics is one of the most used in plant metabolomics, where researchers can undertake two main approaches: targeted and untargeted metabolomics (Fig. 2). Targeted metabolomics is a hypothesis-driven approach focusing on a specific set of metabolites. This approach is often used when the researcher has prior knowledge about the metabolites of interest or when a specific metabolic pathway is under investigation. In targeted metabolomics, a set of known metabolites is selected, and the mass spectrometer is set up to detect these specific metabolites. Targeted metabolomics can provide more accurate and quantitative information about specific metabolites, but it requires prior knowledge of the metabolites under investigation; on the other

hand, untargeted metabolomics is an exploratory approach that aims to identify all metabolites in every biological sample. This approach does not require prior knowledge of the metabolites under investigation, and it allows for detecting unexpected or unknown metabolites. In untargeted or nontargeted metabolomics, the mass spectrometer is set up to detect all metabolites within a certain mass range. The resulting data is analyzed using statistical methods and bioinformatics tools to identify metabolites. Untargeted metabolomics can provide a more comprehensive view of the metabolic profile of a biological sample. Still, it may miss some metabolites outside the mass spectrometer's detection range.

The combination of several omics has been recently implemented on gene discovery and their functional characterization to investigate gene relationship to metabolites supporting and accelerating crop improvements. For example, Wu et al. established a high-quality chromosome-level genome assembly of *Melilotus albus*, by resequencing 94 *Melilotus* accessions to characterize their phylogenetic relationships, the genetic exchange between *M. albus* and *M. officinalis*, and the differentiation of flower color and coumarin content. In addition, transcriptomics, metabolomics, and bulked segregant analysis (BSA) have been used to investigate *M. albus* near-isogenic lines segregating at the coumarin level to identify the key metabolites and enzymes in the coumarin biosynthesis pathway [3]. Similarly, Li et al., based on the integrative analysis of the transcriptomics and targeted carotenoid research, found that differentially expressed genes (DEGs) related to carotenoid metabolism had a stronger correlation with the critical carotenoid metabolite content in the panicle of foxtail millet. Correlation and weighted gene coexpression network analysis (WGCNA) identified and predicted the gene regulation network related to carotenoid metabolism [4].

Zheng et al. used a combination of metabolome and transcriptome of 11 tea cultivars and then a WGCNA-based biological system strategy to interpret metabolomic flux, predicted gene functions, and mined critical regulators involved in the flavonoid biosynthesis pathway; in this manner, they revealed new uncharacterized transcription factors (TFs) such as MADS, WRKYs, and SBP; and microRNAs (including 17 conserved and 15 novel microRNAs) that are potentially implicated in different steps of the catechin biosynthesis. In addition, they applied the metabolic-signature-based association method to capture additional critical regulators involved in the catechin pathway. This provides important clues for the functional characterization of five SCPL1A acyltransferase family members, which might be implicated in the production balance of anthocyanins, galloylated catechins, and proanthocyanins in tea cultivars [5].

Another approach has been implementing metabolite genome-wide association studies (mGWAS) and metabolite quantitative trait loci (mQTL), especially in cereal grains such as wheat. Chen et al. developed an approach that has also been applied in other major crops such as maize, barley, tomato, and blueberry, while the mQTL approach has been used in crops such as soybean, rice, and carrot [6]. Traditionally the output of mQTL/mGWAS was merely linkages/associations between chromosomal locations and metabolite contents, and this was basically due to the lack of genomic information for some crops until very recently, a wheat

mGWAS successfully identified 26 candidate genes with high confidence, among them two were validated to be involved in flavonoid metabolism pathway of wheat [6, 7].

This way, by combining interval mapping and genome-wide association studies (GWAS), the genetic determinants of tocochromanol accumulation in tomato (*Solanum lycopersicum*) fruits have been unveiled. Specifically, the content of vitamin E has been enhanced in tomato plants by identifying the genes involved in the chorismate-tyrosine pathway [8]. With this approach, Alseekh et al. reported a large-scale metabolic quantitative trait loci (mQTL) analysis on the well-characterized *Solanum pennellii* introgression lines to investigate the genomic regions associated with secondary metabolism in tomato fruit pericarp [9]. In total, 679 mQTLs were detected across the 76 introgression lines. Heritability analyses revealed that the environment affected mQTLs of secondary metabolism less than mQTLs of primary metabolism. Network analysis allowed to assess the interconnectivity of primary and secondary metabolism and compare their respective associations with morphological traits. Additionally, a real-time quantitative PCR platform demonstrated a transcriptional control mechanism of a subset of the mQTLs, including those for hydroxycinnamates, acyl-sugar, naringenin chalcone, and a range of glycoalkaloids. Intriguingly, many of these compounds displayed a dominant-negative mode of inheritance, contrary to the conventional wisdom that secondary metabolite contents decreased on domestication. Additionally, two candidate genes for glycoalkaloid mQTLs via virus-induced gene silencing were also validated [9].

More recently, Alseekh et al. identified several metabolite quantitative trait loci that reduce variability for both primary and secondary metabolites (phenylalanine, glucose-6-P, fructose-6-P, and maltose), which they named canalization metabolite quantitative trait loci (cmQTL); on their study nine cmQTL were validated using an independent population of backcross inbred lines, derived from the same parents, which allows increased resolution in mapping the QTL previously identified in the introgression lines. These cmQTL showed little overlap with QTL for the metabolite levels themselves. Moreover, the intervals they mapped harbor few metabolism-associated genes, suggesting that regulatory genes largely control the canalization of metabolism [10].

Maize has also been favored by the combination of GWAS and metabolomics profiling to highlight genes involved in the biosynthesis of several metabolites. Among these compounds, Liang et al. reported the identification of metabolites biomarkers for the tolerance to salt-induced osmotic stress; a citrate synthase, a glucosyltransferase, and a cytochrome P450 were found to be responsible for controlling the associations between the genotype and metabolites that induced the tolerance [11]. Owens et al. reported essential genes controlling maize grain carotenoid composition by using GWAS of quantified seed carotenoids across a panel of maize inbreds ranging from light yellow to dark orange in grain color; significant associations at the genome-wide level were detected within the coding region of *zep1* and *lut1*, carotenoid biosynthetic genes not previously shown to impact grain carotenoid composition in association studies, as well as within previously associated *lcyE* and

crtRB1 genes [12]. Similarly, Li et al. examined the genetic architecture of maize oil biosynthesis in a genome-wide association study using 1.03 million SNPs characterized in 368 inbred maize lines, including “high-oil” lines. Seventy-four loci were significantly associated with kernel oil concentration and fatty acid composition [4]. Another combination of maize metabolic profiling with GWAS has been reported by Riedelsheimer et al. who outlined an association of genetic variants and concentration of 118 metabolites in leaves of 289 diverse maize inbred lines from worldwide sources; genome-wide association mapping with correction for population structure and cryptic relatedness identified for 26 distinct metabolites with strong associations with SNPs explains up to 32.0% of the observed genetic variance [13]. Similarly, Chen et al. described a comprehensive profiling of 840 metabolites and a further metabolic genome-wide association study based on ~6.4 million SNPs obtained from 529 diverse accessions of *Oryza sativa*. They identified hundreds of common variants influencing numerous secondary metabolites with significant effects at high resolution. They observed substantial heterogeneity in the natural variation of metabolites and their underlying genetic architectures among different rice subspecies [13, 14]. Data mining identified 36 candidate genes modulating metabolite levels that are potentially physiological and nutritionally important. As a proof of concept, they functionally identified (annotated) five candidate genes influencing metabolic traits; the study provides first-time insights into the genetic and biochemical bases of rice metabolome variation and can be used as a powerful complementary tool to classical phenotypic trait mapping for rice improvement. Besides, Dong et al. reported a comprehensive metabolic profiling and natural variation analysis of phenolamides in rice using a liquid chromatography (LC)-mass spectrometry (MS)-based targeted metabolomics method; spatiotemporal controlled accumulations were observed for most phenolamides, together with their differential accumulations between the two major subspecies of rice.

Further GWAS on rice leaves identified *Os12g27220* and *Os12g27254* as spermidine hydroxycinnamoyl transferases that might underlie the natural variation of spermidine conjugate levels in rice [15]. Likewise, Chen et al. identified 32 candidate genes underlying metabolic traits in rice grains; 8 candidate genes were involved in the biosynthesis and transportation of amino acids and their derivatives. Three candidates were assigned to the choline levels and its lysophosphatidyl derivatives. Precise signals for trigonelline, a bioactive compound implicated in cell cycle control, resulted in the assignment of seven candidate genes for this metabolite. Furthermore, mGWAS in rice grains revealed 40 candidates (both regulatory and structural genes) involved in the biosynthesis, modification, and transportation of phenylpropanoids, including the C-glycosyl flavones, the primary class of flavonoids in cereals [7]. More recently, Yang et al. reported the identification of a gene *Os07g32020* (UGT707A3) that encodes a glucosyltransferase that converts naringenin and uridine diphosphate-glucose to naringenin-7-O- β -D-glucoside; the function of *Os07g32020* was verified with CRISPR/Cas9 mutant lines, which accumulated more naringenin and less naringenin-7-O- β -D-glucoside and apigenin-7-O- β -D-glucoside than wild-type Nipponbare [16].

It is evident that when working with a crop with a reference genome sequence, such as maize, rice, and wheat, to mention some of the global agricultural importance, enormous advances have been made in elucidating genes and their functional annotation for their improvements; sadly this is not the case for minor crops, medicinal plants, and other regional staple foods. Therefore, considerable effort must be made to genome sequence minor crops and develop a collection of segregating crop populations. On the other hand, plant metabolome lacks a free online accessible metabolite database for support in the annotation of unknown metabolites, and most of the studies carried out under a nontargeted metabolomic approach remain with several novel metabolites, making association studies such as transcriptomics, mQTL, and GWAS difficult when searching for gene annotation and their functional characterization. A summary of mQTL and mGWAS is presented in Table 1.

Table 1 Summary of mQTL and mGWAS studies in plants [17]

Species	Tissue	Population type	Method	Metabolic traits	Ref.
mQTL study					
<i>Arabidopsis</i>	Harvested seed	Recombinant inbred lines	HPLC	Tocopherol	[18]
<i>Arabidopsis</i>	Leaf	Recombinant inbred lines	GC-TOF-MS	Metabolome	[19]
<i>Arabidopsis</i>	Seed	Recombinant inbred lines	LC-MS	Flavonoids	[20]
<i>Arabidopsis</i>	Seedling	RILs and introgression lines	GC-TOF-MS	Metabolome	[21]
<i>Brassica napus</i>	Leaf/seed	Double haploid lines	HPLC	Glucosinolates	[22]
Maize	Leaf	RILs and natural accessions	GC-TOF-MS	Primary metabolites	[23]
Rice	Seed	Chromosomal segment substitution lines	LC-Q-TOF-MS	Metabolome	[24]
Rice	Seed	F2, F2-derived lines	GC-MS	Lipids	[25]
Rice	Flag leaf/germinating seed	RILs	LC-EI-MS	Metabolome	[26]
Tomato	Fruit	Introgression lines	GC-MS	Metabolome	[27]
Tomato	Fruit	Introgression lines	GC-MS	Metabolome	[28]
Tomato	Fruit	Introgression lines	LC-MS	Metabolome	[29]
Tomato	Fruit	Introgression lines	GC-MS	Primary metabolites	[30]
Tomato	Fruit	Introgression lines	UPLC	Secondary metabolites	[9]
Wheat	Flag leaf	Doubled haploid lines	LC-ESI-MS	Metabolome	[31]
Wheat	Flag leaf	Doubled haploid lines	GC-MS	Metabolome	[32]

(continued)

Table 1 (continued)

Species	Tissue	Population type	Method	Metabolic traits	Ref.
mGWAS study					
<i>Arabidopsis</i>	Seed	Natural accessions	LC-MS	Branched-chain amino acids	[33]
<i>Arabidopsis</i>	Leaf/seedling	Natural accessions	LC-MS	Glucosinolates	[34]
<i>Arabidopsis</i>	Leaf	Natural accessions	GC-TOF-MS	Metabolome	[35]
Maize	Kernel	Natural accessions	UPLC-MS	Metabolome	[36]
Maize	Grain	Natural accessions	HPLC	Carotenoid	[12]
Maize	Grain	Natural accessions	HPLC	Tocochromanol	[37]
Maize	Leaf	Natural accessions	GC-MS	Metabolome	[38]
Maize	Leaf	Natural accessions	GC-MS	Metabolome	[13]
Maize	Kernel	Natural accessions	LC-MS	Metabolome	[39]
Potato	Tuber	Natural accessions	GC-MS	Primary metabolites	[40]
Rice	Leaf	Natural accessions	LC-QTOF-MS	Secondary metabolites	[41]
Rice	Leaf	Natural accessions	LC-MS	Metabolome	[14]
Rice	Leaf	Natural accessions	LC-MS	Phenolamides	[15]
Tomato	Fruit	Natural accessions	GC-MS	Metabolome	[42]

3 Plant Species Identification

The plant metabolome constitutes an enormous reservoir of bioactive compounds; many of these are products of secondary or specialized metabolism. Their taxonomic distribution is in relatively narrow phylogenetic clades within *Plantae* [1]. While the identification of plant species based on morphological characteristics is a well-established practice in botany, being able to identify plant species is an essential factor in understanding biodiversity, the discovery of bioactives from herbal medicines, and correlating chemical components from plants with chemical markers of patients who intake herbal medication, as well as monitoring food intake in foodomics studies, food authentication, and fraud detection, among other applications. Since the small-molecule profile of an organism ultimately reflects the genes that distinguish it, the information content of the metabolome might be just as well suited to genomic fingerprinting and assessment of genetic relatedness between species as the genomes themselves [43].

Despite its practical importance, the establishment of phylogenetic diversification and distribution patterns of secondary plant metabolites is still in its early steps, and several plant families have not been deeply explored to date in this context; Mannocho-Russo et al. described a strategy for chemotaxonomic investigations using the Malpighiaceae botanical family as a model; their workflow (Fig. 3) was based on MS/MS untargeted metabolomics, spectral searches, and recently described in silico classification tools, which were mapped into the latest molecular phylogeny accepted for this family, the workflow combines several approaches to

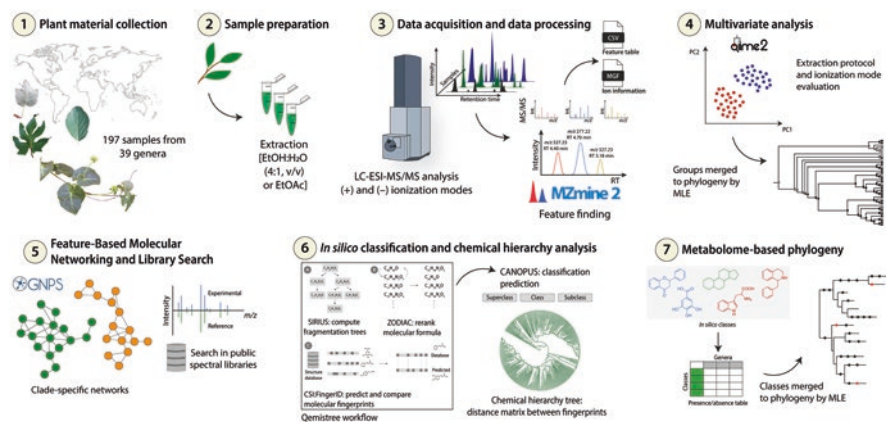


Fig. 3 Experimental workflow followed for the metabolomics and chemosystematics analyses of Malpighiaceae samples. (1) The samples were initially collected. (2) The extracts were prepared with different solvents (EtOH:H₂O (4:1, v/v) or EtOAc) and then (3) subjected to LC-ESI-MS/MS analysis in positive and negative ionization modes in an untargeted method. (4) The data acquired were processed for feature finding, and the exported data were used for multivariate analysis. The clustering groups observed were merged to the phylogeny using the maximum likelihood estimation (MLE) for preliminary chemotaxonomic investigations. (5) The data were also used for feature-based molecular networking and library searches workflows to observe clade-specific molecular families. (6) A chemical hierarchy analysis and *in silico* classifications were obtained and finally (7) merged to the currently accepted Malpighiaceae phylogeny to determine the ubiquitous and the taxa-specific *in silico* classes. (Reproduced from Ref. [44])

perform a comprehensive evolutionary chemical study. It is expected to be used in further chemotaxonomic investigations [44].

The metabolomic analysis revealed that different ionization modes and extraction protocols significantly impacted the chemical profiles, influencing the chemotaxonomic results. In addition to the library searches for metabolite annotation, the MS/MS data generated were visualized by molecular networking analysis (Fig. 4). Molecular families constructed by such analysis represent the similarity of fragmentation patterns obtained by tandem mass spectrometry (MS/MS) analysis. These molecular families consisted of nodes (representing MS/MS spectra) and edges of connecting these nodes (representing the cosine similarity between two nodes, which measure the relatedness in MS/MS spectra). The library matches retrieved from the analysis obtained in the positive ionization mode showed the presence of a high diversity of compound classes, including C-glycosylated and O-glycosylated flavonoids, lipids, alkaloids, quinic acid derivatives, amides, triterpenes, iridoids, and lignans [44].

Cinnamon is one of the oldest spices used in the world. A growing number of studies have illustrated varied phytochemical compositions among cinnamon species. Primary cinnamon metabolites, such as coumarin, cinnamaldehyde, cinnamic acid, cinnamyl alcohols, and proanthocyanidins, are shown to be differentially produced among various species; in this context, Zhang et al. developed a metabolomic



Fig. 4 Molecular families obtained from the feature-based molecular networking workflow and annotated based on spectral matches within the GNPS platform: (a) phenolic compounds, (b) alkaloids, and (c) lipids and terpenoids. Each node represents a tandem mass spectrometry spectra (MS/MS), while the edges that connect them represent the MS/MS fragmentation similarity (cosine >0.7). Pie charts indicate the relative abundance of ion features in each Malpighiaceae phylogenetic clade (A–J). Node sizes are relative to the summed peak areas of the precursor ion in MS1 scans. These are level 2 or 3 annotations according to the 2007 metabolomics standards initiative [45]. (Reproduced from Ref. [44])

ratio rule-based classification method for the automated metabolite profiling and differentiation of four cinnamon species using ultra-performance liquid chromatography-high-resolution mass spectrometry. Among the species studied were *Cinnamomum cassia* (Chinese cinnamon), *C. loureiroi* (Vietnamese cinnamon), *C. verum* (Ceylon cinnamon), and *C. burmannii* (Korintje cinnamon); proanthocyanidins, coumarin, and cinnamaldehyde were the preselected metabolites allowing the classification [46].

The genus *Vanilla*, a source of the most appreciated flavor worldwide, comprise over 110 species. Currently, only three species have commercial relevance, *Vanilla planifolia* Andrews, *V. tahitensis* J. W. Moore, and *V. pompona* Schiede. *V. planifolia* are preferred by industry due to its higher content of vanillin, the main flavor component; more attention needs to be made to other *Vanilla* species. Leyva et al. developed a nuclear magnetic resonance (NMR) metabolomic platform to profile for the first-time leaves that are known to accumulate putative vanillin precursors of *V. planifolia* and those of Peruvian *V. pompona*, *V. palmarum*, and *V. ribeiro* to determine metabolite difference among them. Their NMR analysis identified 36 metabolites, and multivariate analysis identified malic and homocitric acids, together with 2 vanillin precursors (glucoside A and B), as relevant markers for species identification [47].

Coffee is appreciated worldwide for its aroma, flavor, and stimulant properties. Souard et al. examined leaves of nine *Coffea* species grown in the same environmental conditions by an untargeted liquid chromatography high-resolution mass spectrometry (LC-HRMS) approach, with the primary objective of identifying metabolites that significantly contribute to the classification between *Coffea* species. Based on their results of multivariate analyses, 1637 variables (metabolites) were analyzed, from which 92% (1505 metabolites) were significantly different overall taxa. Among the species studied, when two well-known *C. arabica* and *C. canephora* were compared, a feature with an $m/z = 195.0870$ corresponding to $[M + H]^+$ of caffeine came out as the main discriminant compound. Caffeine concentration was approximately 800 times higher in *C. arabica* leaves than in *C. canephora*. Another feature observed at an m/z value of 247.0598 had much higher intensities in *C. arabica* than in *C. canephora*, but unfortunately, this feature was not identified. This metabolic fingerprinting study aimed to determine the specific differences between the metabolomes. All nine clusters of each species studied were observed on both PCA and PLS-DA score plots, with good discrimination between the eight *Coffea* species [48].

Several studies have described the use of metabolomics to distinguish herbal medicinal plants. For example, Lesiak et al., using seeds of the genus *Datura* plants, outlined direct analysis real-time mass spectrometry (DART-MS), which could provide diagnostic fingerprint profiles of nine *Datura* species seeds and whether chemometric processing of the observed profiles could enable species-level identification and differentiation. They confirmed that the seeds could be analyzed by DART-MS directly in a high-throughput manner without using a solvent extract. Each species exhibits a distinct chemical signature, and the processing of this data by multivariate statistical methods enables species-level differentiation. In addition, they observed that while intraspecies chemical signatures are similar, interspecies fingerprints are distinct enough to be discriminated against using multivariate statistical analysis tools [49]. Another example using seed samples of *Polygonatum* species was reported by Qi et al.; *Polygonatum* plant species have properties that make them sound like medicine and food in China. There were almost no differences in the contents of the metabolites in the amino acids and derivatives, nucleotides and products, and others (e.g., saccharides, alcohols, and vitamins) classes among the seed

samples. In contrast, the seed samples had some diversity in the contents of lipids, phenolic acids, lignans and coumarins, tannins, and quinones. The flavonoid, steroid, and terpenoid classes and contents varied among the *Polygonatum* seed samples; these compounds have relatively strong pharmacological effects; their findings indicate that different *Polygonatum* seeds differ in terms of their medicinal and nutritional value [50].

Other plant species that have been classified and identified using metabolomic approaches are *Mentha* species [51]; *Acorus* [52] – plants in *Acorus* have been used as herbal medicine by various linguistic groups for thousands of years; and *Phyllanthus* species [53]; among other medicinal plant species.

4 Plant Root Metabolome and Climate Change

Climate change is a relevant issue due to its adverse and high-impact consequences it can cause directly and indirectly at social, ecological, biological, and health levels [54, 55]. The main factors promoting climate change comprise natural and anthropogenic activity; agriculture can produce 30–40% of the total greenhouse gas emissions [56]. In addition, due to these issues, an increase in the use of pesticides is expected, which will significantly affect global crop production, as well as pathogen diseases, abiotic stress, and the decrease in the production of the major crops worldwide [57]. Quality and crop productivity are negatively affected by global warming. It is expected that dramatic increase in the following years due to the increasing annual temperature, solar radiation, changes in precipitation, and high CO₂ levels [54, 56, 57]. Some other factors which reduce crop production quality are floods and droughts [58, 59] and are affected by alterations in rainy seasons, pest invasions, crop disease, water supplies, price of products for agricultural processes, and premature consumption of fertilizers [54].

Humanity's well-being and economics depend strongly on the agricultural sector, which simultaneously depends on the ability to adapt crops to environmental conditions and is therefore considered a climate-dependent industry [58]. To improve productivity, nutrient quality, and crop resilience, it is necessary to adapt strategies and design technologies to help mitigate climate change's effects [55, 60]. The leading crop breeding technologies for adaptation comprise biotechnology techniques, such as next-generation sequencing and RNA-mediated silencing [58, 61, 62]. Another strategy to improve crops and their resistance to biotic and abiotic stress is the engineering of the root microbiota, which represents a promising technology in the future to face climate change [63]. This strategy arises from analyzing the plant microbiome in the rhizosphere and endosphere, in which several interactions occur between the plant, microorganisms, their metabolites, and the metabolites of the surrounding plants. Once the relevance of these interactions on plant development is understood, the rhizosphere microbiome should be manipulated and thus increase production and their well-being; reduce the need for farmland,

pesticides, and fertilizers; and thus reduce the intrinsic carbon footprint associated with these activities [64–66].

Plant evolution gave way to the adaptation through metabolite excretion to the soil (exudates) to interact with rhizosphere composition. The exudates may alter the composition and activity of the microbiome around, changing the pH, soil structure, and availability of oxygen or supplying organic compounds as a source of energy [67]. Some of the compounds exudated can also act as chemotrophic signals, which can attract pathogenic microorganisms to the plant, nematodes, or herbivorous arthropods [68]; however, these compounds can also work in the recruitment of beneficial microorganisms which can aid the plant in defense against pathogens, diseases, biotic stress conditions or enhance nutrient absorption [63, 64]. Under stress conditions, plants trigger the production of many secondary metabolites, with defense signals that promote the cope against pathogen organisms. Synthesis of secondary metabolites can relieve stress by modifying root microbiota to further degradation of different types of pollutants to carry out bioremediation [63, 69].

Among the strategies proposed for manipulating the rhizosphere microbiome is the direct inoculation of microorganisms in the soil. However, one of the biggest challenges to achieve is the determination of the species that act on the mechanisms involved, the competitive behavior with the native microbiota, and the effects it has on agricultural conditions. Alternatively, the metabolites observed in rhizosphere exudates should be used and applied in specific areas to stimulate native microorganisms [67]. Therefore, it is necessary to determine qualitatively and quantitatively the composition of the exudates under particular conditions, as well as the metabolomic analysis of the prominent participants in the interactions described and reasonably link the production of metabolites and their primary function in the face of possible types of stress [65, 70].

A metabolomic analysis is an excellent choice to study plant-rhizosphere interactions due to the complexity and quantity of compounds involved in the metabolic relation between roots and microorganisms. Instrumentation such as LC-MS, GC-MS, and NMR are the most widespread techniques and powerful tools used for identifying the compounds present in the rhizosphere [68, 71]. On the other hand, the exometabolomics concept consists of analyzing the metabolic traces of microorganisms present in certain conditions to understand the underlayer mechanisms that exist in the rhizosphere and the determination of substrate consumption by microorganisms [64]. In this context, we reviewed recent literature in metabolomics and exometabolomics, the findings in primary metabolites, and their function on specific rhizosphere interactions.

4.1 Metabolome Changes in Roots by Abiotic Stress

Metabolome alterations in the rhizosphere due to interactions result in significant interest in setting the defense mechanisms that play plants against external agents and expanding the landscape about key metabolites to degrade specific

contaminants. Therefore, Wang et al. explored changes in chemical composition in root exudates of the urban greening trees exposed to phenanthrene. In *Loropetalum chinense*, *Gardenia ellis*, and *Rhododendron simsii*, carbohydrate levels increased in the presence of phenanthrene, indicating a regulatory function mediated by the biopolymer's degradation, whereas in *Osmanthus fragrans* levels decreased, suggesting different behavior between species. Phenolic compounds increased in the presence of phenanthrene in *Ligustrum japonicum*, *R. simsii*, *O. fragrans*, *Gardenia jasminoides*, and *Camellia sasanqua*, implying an adaptation to attract rhizobia bacteria, with the aim of cope with exposure to phenanthrene [72]. Regarding heavy metals stress, Lu et al. analyzed two wheat genotypes with different tolerance to Cd: Aikang 58 with low accumulation and Zhenmai 10 with high accumulation. Both phenotypes showed an increase in phenylalanine and tyrosine in the presence of Cd, relating these changes in the shikimate-phenylpropanoid pathway. The rise in acetylglycine and histidine indicates a chelating activity to chelate Cd in vacuoles. At the same time, glutamate, glutamine, aspartic acid, asparagine, and lysine perform an osmotic balance to detoxify heavy metals. In the presence of Cd, an increase in maltose, isomaltose, sorbose, tagatose, and polyols assists the cell wall's structure. In contrast, the addition of glyceric acid, cis-aconitate, malic acid, salicylic acid, and citrate indicates a deterioration in the activity of the tricarboxylic acid cycle activity to assimilate carbon under stress conditions. These alterations in the metabolism promote a high ability to take out and defend the plant against reactive oxygen species, inducing molecular signaling and antioxidant enzymes [73].

A report studying the effect of acid drainage contamination on *Phragmites australis* by Kalu et al. analyzed root and rhizosphere metabolome. The main compounds found in roots at contaminated sites were adenosine, inosine, methionine, carnitine, and dimethylglycine. On the other hand, uridine, dopa, asymmetric dimethylglycine, adenosine, and phenylalanine had a lower abundance in contaminated sites. This alteration has the purpose of recruiting microorganisms that promote the growth of the plant while at the same time attracting microorganisms specialized in heavy metal detoxification. As for bacterial communities, the main phylum in samples grown at contaminated sites was proteobacteria, β -proteobacteria, and the *Methylocystis*, *Rhizobium*, and *Delftia* genera [74]. In salinity stress, Wang et al. analyzed the canola roots metabolome in the presence of NaCl. The abundance of proline and soluble sugars increased in canola roots under saline stress. However, the metabolites with the most significant difference between groups were lipids, primarily fatty acids, which increased compared to controls. In saline stress, lipids affect membrane permeability, fluidity, integrity, and protein transport activity; therefore, the reconstitution of lipids in cells becomes transcendental. This stress triggers the production of polyunsaturated fatty acids, which can help the activation of membrane ATP-loop activity, which is responsible for maintaining homeostasis, facilitating the pumping of Na^+ from the cytosol to the external medium, and blocking K^+ channels [70].

Finally, it is essential also to consider UV radiation stress. Mannucci et al. analyzed the effect of UV exposure on tomato plants and their metabolomic changes in

roots and leaves. In the roots of plants exposed to UV radiation, terpenoids and phenylpropanoid derivatives pathway synthesis increased compared to controls. For carbohydrates, degradation processes increased in radiation treatments, suggesting a wide degree of mobilization of reserve compounds to produce necessary precursors for secondary metabolites synthesis, such as flavonoids. As an effect of a rearrangement of lipid membrane composition, monogalactosylglycerol was found in the plants exposed, while 4 α -carboxy-5 α -cholesta-7,24-dien-3 β -ol levels decreased. Finally, phenolic compounds and p-coumaroyl glycolic acid decreased in the UV treatments, a compound with anti-inflammatory properties [75].

4.2 *Metabolic Changes Associated with Root and Other Biotic Factor Interactions*

The most relevant biotic factors that transcend the metabolome analysis include plant pathogenic microorganisms; it is essential to decipher the compounds produced by plants to counteract the conditions and even the mutual organisms that help defend against these agents. Another biotic stress to consider is the neighborhood of other plant species competing for nutrients. In addition, the importance of the interaction of mycorrhizal fungi and the benefits they have in association with root plants is known. Both metabolisms are relevant to develop strategies that can help to improve crops.

To find out how rye competition affects *Vicia villosa* Roth, Hazrati et al. analyzed the metabolic effect on the roots of these plants. Kaempferol-Rha-Xyl-Gal was the main compound found. This compound decreased in *V. villosa* when it was grown in the presence of rye. Thus, it is estimated that competition produces a deficit of several nutrients in hairy vetch, decreasing the production of flavonoids [76]. *Phytophthora sojae* is known to cause phytophthora root rot disease in soybean; Zhang et al. analyzed the rhizosphere of *Glycine max* inoculated with *P. sojae*. The post-inoculated rhizosphere of a resistant species had a greater abundance of metabolites related to cutin biosynthesis, suberin, wax, arginine, ansamycins, pyrimidine, galactose, linoleic acid, ABC transporters metabolism, and lysine degradation. Most of the metabolites in the post-inoculated rhizosphere include antibiotics, which are responsible for conferring plant resistance to pathogens. On the other hand, some compounds in the control rhizosphere contained compounds to attract possible pathogenic microorganisms to the plant, such as daidzein and genistein. Although some flavones and isoflavones repel zoospores, others have the opposite effect, each specific to the conditions the plants were exposed to. Besides, cutin, suberin, and wax biosynthesis are inferred to provide drought tolerance by preventing water loss and insect tolerance [77].

Finally, interaction with symbiotic microorganisms focuses on knowing the benefits they bring and how they achieve them. Therefore, a view with a metabolomic approach is interesting to know the main compounds in the rhizosphere used to

improve crop capacity. In this context, Zhang et al. analyzed the interaction between *Medicago truncatula* and rhizobium bacteria to get the metabolites in the formed nodules. Oxylipin-9-HODE decreased during the application of rhizobia bacteria, indicating decreased jasmonic acid precursors and inhibition due to the interaction [78]. Oxylipins are critical signaling molecules in the defensive response of plants to protect their tissues from attacks by herbivores or pathogens, and some contain antimicrobial properties.

Similarly, Sebastiana et al. analyzed cork oak roots colonized with *Pisolithus tinctorius*, an ectomycorrhizal fungus. In the study, the inoculated roots showed higher levels of γ -aminobutyric acid, alanine, β -glucose, and citrate. It was also demonstrated that the inoculated samples decreased quercitol, glycine-betaine, α -glucose, fructose, malate, and lactate levels. Inoculated samples influenced alkaloids, terpenoids, oxylipins, lipids, carbohydrates, amino acids, nucleic acids, and vitamins. In addition, a decrease in isomers of glucose, sucrose, sorbitol, and mannosyl glycerate was observed, as well as a reduction of isomers of fatty acids and compounds involved in the metabolism of tyrosine and histidine. The decrease of organic acids and glycine-betaine is related to apoplastic protective barriers, indicating the transfer of these metabolites to fungi, in the case of lipids, including monoacylglycerols, which are the main components of suberin and bark. This layer accumulates on the most exposed face of tree stems and roots and protects against drying and pathogen attack [79].

Another example was shown by Hernández et al. who analyzed the beneficial activity of *Rhizobium tropici* in *Phaseolus vulgaris* growth under phosphorous deficiency. It was observed that some organic acids, polyhydroxy acids, sugars, and polyols increased significantly in nodules with phosphorus (P) deficiency. In contrast, some amino acids and nitrogenous compounds decreased, reducing N fixation in P-deficient plants. In addition, they presented sugar accumulation, indicating demand for root photosynthate due to the decrease in the photosynthesis network. On the other hand, changes in carbohydrate content mean glycolysis/C binding pathways are induced in nodules under P deficiency stress [80]. P-deficient roots showed a decrease in the organic acid concentration, suggesting their exudation toward the rhizosphere; this has also been recently demonstrated in a study by Gomez-Zepeda et al. when using mass spectrometry imaging to locate organic acid exudates in P-deficient *Arabidopsis* plants. Organic acid exudation by roots is considered a core response to different types of abiotic stress and the interaction of roots with soil microbes. For decades it has been a target trait to produce plant varieties with increased capacities of inorganic orthophosphate uptake and aluminum tolerance [81].

Microbiota may vary according to the plant growing zone. In this way, Li et al. analyzed metabolome and microbiota in the roots of *Aconitum vilmorinianum* grown in two different sites in China (Luquan and Weixi). The difference observed in the metabolites was an increase in yunnaconitin and vilmorranine A in Weixi and a decrease in amino acids and some derivatives in Luquan. A correlation was found between 137 bacteria and 17 fungi with 75 differential metabolites in the 2 regions,

among which the fungus on *Cladosporium* stands out, with a high probability conditional on aconitine, demonstrating the appearance of this metabolite in Weixi samples. Regarding Luquan, three bacterial and six fungal biomarkers were found, while Weixi showed the presence of five bacterial and five fungal biomarkers. This finding in the microbiome may be due to the environmental temperature, while in Weixi usually snows, and Luquan rarely occurs [82].

Knowledge about metabolome interaction between roots and their environment is crucial to identify the relevant metabolites produced in this medium. With this, it is intended to know which metabolic pathways are altered in plants to produce detoxifying compounds, antibiotics, or those that recruit beneficial microorganisms for plants. On the other hand, the compounds produced by microorganisms and their identification are also convincing to analyze the possibilities of the inoculation of different bacterial species in scenarios of various types of stress. However, it is important to specifically identify these conditions since not all species will behave similarly under the same state or stress. Herewith, it is essential to trace how crops can be treated to face biotic and abiotic stress caused mainly by climate change and improve their production and development.

5 Plant Biomarker Identification

Metabolomics is one of the most recent powerful tools for studying plants and other organisms and is becoming a complementary technique to genomics, transcriptomics, and proteomics [83]. Metabolomics addresses the activity of small molecules (<1500 Da) produced by cells during their life cycles, that is, products of primary or secondary – specialized – metabolism, found in various biological systems, studying how metabolic profiles change within an organism in response to some situations, such as disease or stress [84].

Therefore, unlike other “omics,” metabolomics best describes phenotypes, can give instantaneous information on the physiological state of cells and thus provide a broader view of the biochemical state of an organism, and can track the metabolic network of a biological system and its perturbation in response to stimuli. Metabolomics aims not to identify every metabolite observed but to compare patterns of metabolites that change in each biological system. When these analyses are performed on enough biological replicate samples, it allows researchers to discriminate and classify samples and gain insight into changes in metabolome composition related to a particular physiological state, influence of stress or stimulus, genetic modification, or interaction with other organisms.

Metabolites provide a “fingerprint” of the complex interaction between the genome and the environment. They can generally be divided into two groups: primary metabolites, essential for maintaining processes directly involved in plant survival, growth, and reproduction, and secondary metabolites, which contribute to specialized processes in each organism synthesized to fulfill a nonessential function in the plant.

Due to the structural heterogeneity of metabolites and their different ranges of magnitude and concentration, their identification and measurement present a considerable challenge. For the plant kingdom alone, researchers have reported more than 400,000 plant species worldwide [85]. As for structurally distinct secondary metabolites, there are approximately 200,000 to 1,000,000 [86], which is why the field of plant metabolomics is the most advanced with a wide range of applications [1, 85].

All this information and understanding of the metabolome as it is affected by factors including environmental changes, physical changes, biotic stresses, abiotic stresses, and even internal changes in the plant as a function of its developmental stage can be used to monitor significant variations in metabolites and in the search for metabolites that can act as biomarkers.

The study of metabolomic biomarkers is one of the least explored areas in metabolomics. By 2022 only 16% of the publications examined the discovery or discrimination of biomarkers, while 46% of the publications refer to metabolic mechanisms and 33% examined metabolic profiles. However, many of these publications on metabolic profiles include preliminary and descriptive findings for more detailed analysis of the machine tool and discovery of biomarkers [87].

A metabolomic biomarker differs greatly from a protein biomarker and transcriptomic biomarkers because of the close relationship between individual metabolites. Factors measured in other “omics” technologies are independent, although there may be patterns of abundance that reflect a disease state. A metabolomic biomarker is not just a chain of changes in individual metabolites. Instead, it is composed of co-related metabolites that change together [88].

For the discovery and characterization of a metabolomic biomarker, validation based on the environment and research design is necessary to determine whether the proposed biomarker can distinguish between the changes to which plants are subjected [88], that is, for a biomarker to be classified as such, it must meet specific characteristics: be measurable, reflect the qualitative or quantitative interaction of the plant with the chemical of interest, be precise and sensitive, and be commonly shared among individuals in a population and plant species.

In this way, and through preliminary findings of metabolic profiling, some biomarkers have been identified in plants in response to exposure to stress factors: For example, a study on biomarker discovery [89] demonstrated that volatile organic compound (VOC) profiles could be used as diagnostic markers of stress in grapevine; this study shows that VOC emission can be considered as a universal response of grapevine to plant defense elicitors, given that the elicitors evaluated induced the emission of a standard set of VOCs encompassing chemically different compounds, including the sesquiterpenes α -farnesene and β -caryophyllene and that such a response is analogous with the induction of stilbene phytoalexins.

Similarly, plant metabolomics can help to identify resistant metabolites in plants that are subjected to stress conditions [17]; the selected biomarker can be used as a diagnostic metabolite for plant stress, as in the case of the study of wheat and barley resistance against *F. graminearum* infection where they point to various plant

hormones that respond to this infection [90]. Such is also the case of phenylpropanoids and organic acids, metabolites identified as biomarkers of nitrogen deficiency in leaves and roots of tea plants (*Camellia sinensis*) that are elevated when there is a nitrogen deficit [91]. For example, hexadecanoic acid and dotriacontane, highly expressed metabolites, were identified as potential biomarkers in rice seeds infected with *Rhizoctonia solani* toxin, metabolites involved in several important rice biosynthetic pathways, such as the biosynthesis of saturated fatty acids and the unsaturated fatty acids cutin, suberin, and wax [92].

However, the identification of metabolites not only corresponds to stress responses, but the detection of metabolic changes at different developmental stages also contributes to finding metabolites characteristic of each stage (biomarkers), as in the case of metabolomic analysis of rice where developmentally controlled phenolamide accumulation patterns are observed [15] or in *Arabidopsis* where patterns of glucosinolate, raffinose, and galactinol accumulation are present at the base of leaves during the senescence stage [90]. Analysis of the spatiotemporal metabolic profile of plant development also allows the identification of potential biomarkers to capture intrinsic genetic features of plant development, as in the study of rice tillering (branching), in which 21 metabolites captured nearly 83% of the metabolic variation [93], and the developmental phase of soybean during the transition from vegetative to reproductive stage, in which eight flavonoid kaempferol glycosides were identified as potential growth markers [94].

In the food industry, the identification of biomarkers has also worked in quality processes for food authenticity and food traceability matrices of plant origin, especially in the field of aromatic herbs and spices, which are very susceptible to food fraud, as in the case of thyme, an aromatic herb traditionally used for food purposes due to its organoleptic characteristics and medicinal properties. In this particular case, it was possible to determine the geographical traceability of thyme based on different origins (Spain, Poland, and Morocco), as well as to evaluate its processing by comparing sterilized thyme with non-sterilized thyme, where 24 differential markers belonging to different classes were identified: among monoterpenoids, diterpenoids, sesquiterpenoids, alkylbenzenes, and other diverse compounds for its authenticity [95].

Another example of this application that helps to detect adulterants in plants that are used commercially is observed in the study by Ivanovic et al. using wild garlic (*Allium ursinum*) and poisonous adulterants *Convallaria majalis* and *Arum maculatum* as a model for the detection of adulterants in edible plants; the metabolites isovitexin, vicenin II, azetidine-2-carboxylic acid, and trigonelline were elucidated as biomarkers of adulteration [96].

On the other hand, metabolomics approaches have also been used to characterize and diagnose plant diseases and thus crop improvement, for example, during the study of the interaction maize-*Fusarium graminearum*-*Bacillus amyloliquefaciens* or soybean-*Rhizoctonia-B. amyloliquefaciens*, a better understanding of the metabolic regulation of all interacting systems has been achieved, providing valuable insights potentially useful in plant breeding and metabolic bioengineering [97].

Metabolite markers against drought stress (malonate, leucine, 5-oxo-L-proline, saccharic acid, trans-cinnamate, succinate, and glyceric acid) have been reported by Khan et al. who identified biomarkers in the metabolome of chickpea (*Cicer arietinum* L.) when treated with plant growth regulators (salicylic acid and putrescine) and the PGPR growth consortium (*B. thuringiensis*, *Bacillus subtilis*, and *Bacillus megaterium*). Deliberative metabolic reprogramming of chickpea targeting biomarker synthesis pathways resulted in drought-tolerant chickpea varieties [98].

Biomarker identification can also be applied to predict phenotypic traits and provide early detection tools to identify and use them in plant breeding development [99]. In China, for example, hybrid rice combinations have been created using sterile lines and restorer lines to reduce seed deterioration during storage and establish galactose and gluconic acid as metabolic biomarkers that reflect the degree of seed aging [100]. Also, in understanding the functioning of plants growing under extreme conditions, the identification of biomarkers in these plants could provide information that would benefit crop improvement; for example, it was possible to identify associated metabolic biomarkers in an alpine medicinal plant (*Rhodiola crenulata*) that can survive in extreme altitude conditions where the shikimic acid-phenylalanine-phenylpropanoid flavonoid pathway was enhanced with phenylpropanoids upregulating much more than flavonoids [101].

Surveillance for potential pathogens is critical to plant innate immunity, so plants depend on the perception of pathogen-derived molecules to activate defense-related signaling cascades and specialized metabolites in response; in studies of the tomato plant (*Solanum lycopersicum*), by monitoring metabolic profiles of signaling cascades in response to pathogens, significant biomarkers were noted for several classes of metabolites including amino acid derivatives, lipid species, steroidal glycoalkaloids, hydroxybenzoic acids, hydroxycinnamic acids, and products, as well as flavonoids [102].

Other metabolites identified as biomarkers in the plant defense response to pathogens are hydroxycinnamic acids; the conjugation of these acids with amide groups contributes to the regulation of the dynamic metabolic pool of hydroxycinnamates; a wide range of biogenic amine compounds found in most plant cells and these conjugates can scavenge radicals, confer antimicrobial activity, and can be deposited in the cell wall; so finding the activity of these metabolites is indicative of the plant-pathogen response [103]. A summary of primary and secondary or specialized metabolites identified in various plant species is presented in Table 2.

The identification of biomarkers in plants can have diverse applications, as described above; however, to reach the validation of these metabolites, metabolic profiling studies are necessary; metabolomics has been widely applied in the study of plants showing a breakthrough in understanding how the phenotype is related to the metabolome and therefore the function of metabolites under normal conditions, stress, and during their development.

Table 2 Identifying key metabolites in various plant species using different analytical methods [104]

Plant species	Class	Analytical tools	Key metabolites	Ref.
Primary metabolites				
<i>Plantago ovata</i>	Fatty acids	GC-MS	α -Linolenic acid, linoleic acid, and palmitic acid	[105]
<i>P. ovata</i>	Fatty acids	GC-MS	Pentadecanoic acid, palmitic acid, heptadecanoic acid, stearic acid, oleic acid, linoleic acid, γ -linoleic acid, and arachidic acid	[106]
<i>Jatropha curcas</i>	Fatty acids	GC	Oleic acid, palmitic acid, and linoleic acid	[107]
<i>Paeonia rockii</i> , <i>P. potaninii</i> , and <i>P. lutea</i>	Fatty acids	GC-MS	α -Linolenic acid, oleic acid, and linoleic acid	[108]
<i>Cicer arietinum</i>	Fatty acids	GC-MS	Pentadecanoic acid, palmitic acid, palmitoleic acid, stearic acid, oleic acid, linoleic acid, α -linolenic acid, and arachidic acid	[109]
<i>P. ovata</i>	Amino acids	HPLC	Isoleucine, threonine, leucine, histidine, and lysine	[105]
<i>P. ovata</i>	Amino acids	HPLC	Aspartate, glutamine, glycine, alanine, arginine, serine, proline, isoleucine, and methionine	[106]
<i>Fritillaria thunbergii</i>	Amino acids	GC-MS	Tryptophan, phenylalanine, and histidine	[110]
<i>C. arietinum</i>	Amino acids	GC-MS	L-glutamic acid, L-tryptophan, phenylalanine, glycine, serine, L-threonine, L-valine, L-ornithine, and L-proline	[109]
<i>C. arietinum</i>	Sugars and sugar alcohols	GC-MS	Sucrose, cellobiose, galactose, methyl galactoside, and <i>myo</i> -inositol	[109]
<i>C. arietinum</i>	Sugar alcohols	GC-QqQ-MS	Galactitol, erythritol, arabitol, xylitol, mannitol, and inositol	[111]
Secondary metabolites				
<i>Beta vulgaris</i>	Terpenes	HPLC-MS	Oleanolic acid, hederagenin, akebonoic acid, and gypsogenin	[112]
<i>Ocimum gratissimum</i>	Terpenes	GC-MS	m-Chavicol, t-anethole, germacrene-D, naphthalene, ledene, eucalyptol, azulene, and camphor	[113]
<i>Mentha piperita</i>	Terpenes	GC-MS	Menthone, menthol, pulegone, and menthofuran	[114]
<i>M. arvensis</i>	Terpenes	GLC	Menthol, isomenthone, L-methone, and menthyl acetate	[115]
<i>Achyranthes bidentata</i>	Terpenes	HPLC	Oleanolic acid and ecdysterone	[116]

(continued)

Table 2 (continued)

Plant species	Class	Analytical tools	Key metabolites	Ref.
<i>Arabidopsis thaliana</i>	Phenolics	UHPLC-MS	Scopoletin, unbelliferone and esculetin, scopolin, skimming, and esculin	[117]
<i>P. ovata</i>	Phenolics	LC-MS	Luteolin, quercetagenin, syringetin, kaempferol, limocitrin, helilupolone, and catechin	[105]
<i>P. ovata</i>	Phenolics	LC-MS	Kaempferol, 3-(2'',3''-diacetylramnoside)-7-rhamnoside and apigenin 7-rhamnoside	[106]
<i>P. ovata</i>	Alkaloids	LC-MS	Lunamarine, hordatine B, and pinidine	[105]
<i>Dendrobium</i> snowflake "red star"	Alkaloids	¹ H and 2D NMR	Dendrobine and nobilone	[118]

Understanding the adaptative physiology and biochemistry of plants, as well as the underlying metabolic events, is relevant to have a global perception of the metabolomic status of plants, with the identification of biomarkers providing helpful information on metabolites involved in resistance responses, stress, growth, a better understanding of intra- and interspecific microbial interactions occurring at different heterogeneous levels within the plant habitat, identification of systemic responses of various crops to pathogen stress, and pathogens and their biological control, would allow crop scientists to identify unique metabolic markers that can be applied to early detection of a plant pathogen as well as to the development of bio fungicides, for example, for use during pre-harvest, post-harvest, and harvest storage and large-scale storage of crops. Identifying and applying metabolic biomarkers could favor controlled and semi-controlled planting systems shortly, and if properly integrated into crop protection strategies, food security could be mitigated. However, the applications of these biomarkers could be helpful in various areas, such as the food and pharmaceutical industry in food quality and safety processes, diagnosis and treatment of plant diseases, crop improvement, and analyzing genetic-modified crops. Still, the work done so far is relatively new, and efforts should continue to cover the tremendous potential presented by identifying metabolic biomarkers. Table 3 summarizes the identification of some metabolites in transgenic plants using different analytical techniques.

Table 3 Identification of important metabolites in transgenic plants using different analytical tools [104]

Transgenic plant	Analytical techniques	Key metabolites	Ref.
<i>Artemisia annua</i>	GC-TOF-MS	Borneol, phytol, (3-farnesene, germacrene-D, artemisinic acid, dihydroartemisinic acid, and artemisinin	[119]
<i>Lactuca sativa</i>	NMR	Asparagine, glutamine, valine, isoleucine, cx-ketoglutarate, succinate, fumarate, malate, sucrose, and fructose	[120]
<i>Lycopersicon esculentum</i>	GC-MS	γ -Aminobutyric acid, histidine, proline, pyrrol-2-carboxylate, galactitol/sorbitol, glycerol, maltitol, 3-phosphoglyceric acid, allantoin, homocysteine, caffeate, gluconate, ribonate, lysine, threonine, homoserine, tyrosine, tryptophan, leucine, arginine, and valine	[121]
<i>Nicotiana tabacum</i>	NMR	Chlorogenic acid, 4-O-caffeoylquinic acid, malic acid, threonine, alanine, glycine, fructose, (3-glucose, cx-glucose, sucrose, fumaric acid, and salicylic acid	[121]
<i>N. tabacum</i>	GC-MS	4-Aminobutanoic acid, asparagine, glutamine, glycine, leucine, phenylalanine, proline, serine, threonine, tryptophan, chlorogenic acid, quininic acid, threonic acid, citric acid, malic acid, and ethanolamine	[122]
<i>Oryza sativa</i>	GC-MS	Glycerol-3-phosphate, citric acid, linoleic acid, oleic acid, hexadecanoic acid, 2,3-dihydroxypropyl ester, sucrose, 9-octadecenoic acid, 2,3-dihydroxypropyl ester, sucrose, mannitol, and glutamic acid	[123]
<i>O. sativa</i>	LC-MS	Tryptophan, phytosphingosine, palmitic acid, 5-hydroxy-2-octadenoic acid 9,10,13-trihydroxyoctadec-11-enoic acid, and ethanolamine	[124]
<i>Populus</i>	GC-MS, HPLC	Caffeoyl and feruloyl conjugates, syringyl-to-guaiacyl ratio, asparagine, glutamine, aspartic acid, γ -amino-butyric acid, 5-oxo-proline, salicylic acid-2-O-glucoside, 2, 5-dihydroxybenzoic acid-5-O-glucoside, 2-methoxyhydroquinone-1-O-glucoside, 2-methoxyhydroquinone-4-O-glucoside, salicin, gallic acid, and dihydroxybenzoic acid	[125]
<i>Solanum tuberosum</i>	LC-TOF-MS	Glutathione, γ -aminobutyric acid, 3-cyanoalanine, 5-oxoproline, sucrose, glucose-1-phosphate, glucose-6-phosphate, fructose-6-phosphate, ethanolamine, adenosine, and guanosine	[126]
<i>Triticum aestivum</i>	GC-MS	Guanine and 4-hydroxycinnamic acid	[127]

(continued)

Table 3 (continued)

Transgenic plant	Analytical techniques	Key metabolites	Ref.
<i>T. aestivum</i>	LC-MS	Aminoacyl-tRNA biosynthesis, phenylalanine, tyrosine, tryptophan glyoxylic, tartaric acid, oxalic acids, sucrose, galactose, mannitol, leucine, valine, glutamate, proline, pyridoxamine, glutathione, arginine, citrulline, adenosine, hypoxanthine, allantoin, and adenosine monophosphate	[128]
<i>Zea mays</i>	¹ H NMR	Lactic acid, citric acid, lysine, arginine, glycine-betaine, raffinose, trehalose, galactose, and adenine	[129]

6 Plant Single-Cell Metabolomics

In the past, genomics, transcriptomics, and metabolomics techniques have been applied in bulk plant samples consisting of many cells; in such experiments, the biological process leading to cell heterogeneity is often considered not to be biologically relevant. However, cell heterogeneity has been shown to play important biological roles in many situations for which averaging would mask relevant metabolic processes [130]. Plants contain several cell types and exhibit complex regulatory mechanisms. Studies at the single-cell level have gradually become more common in plant science. Single-cell transcriptomics, spatial transcriptomics, and spatial metabolomics techniques have been combined to analyze plant development. These techniques have been used to study the transcriptomes and metabolomes of plant tissues at the single-cell level, enabling the systematic investigation of gene expression and metabolism in specific tissues and cell types during defined developmental stages [131]. However, single-cell technologies require labor-intensive protocols for plant cell isolation. On that respect several attempts have been developed; these strategies can be classified into three main groups: those that attempt to isolate material of specific cell type to perform the analysis on platforms used for regular metabolomics, which we will refer to as single-cell type metabolomics [132]; those based on micromanipulation of single cells; and those based on mass spectrometry imaging [130].

Several methods for harvesting cells have been developed for single-cell and single-cell type metabolomics, whereby cells can be obtained or extracted in situ. The in situ techniques include micropipetting for isolating the contents of specific cells, laser microdissection (LMD), laser microdissection and pressure catapulting (LMPC), laser capture microdissection (LCM), and fluorescence-activated cell sorting (FACS). Laser microdissection and pressure catapulting and laser capture microdissection use laser to excise single cells or microareas from fixed or frozen intact tissues and are becoming very popular for plant cell and tissue sampling. FACS, on the other hand, is used to obtain specific cell types; for example, those identified from root developmental zones by transgene-labelled nuclei or by immunolabelled-based collection and microfluidic sorting-based methods that exploit intrinsic cell properties [129]. Figure 5 summarizes different approaches for cell-specific metabolomics.

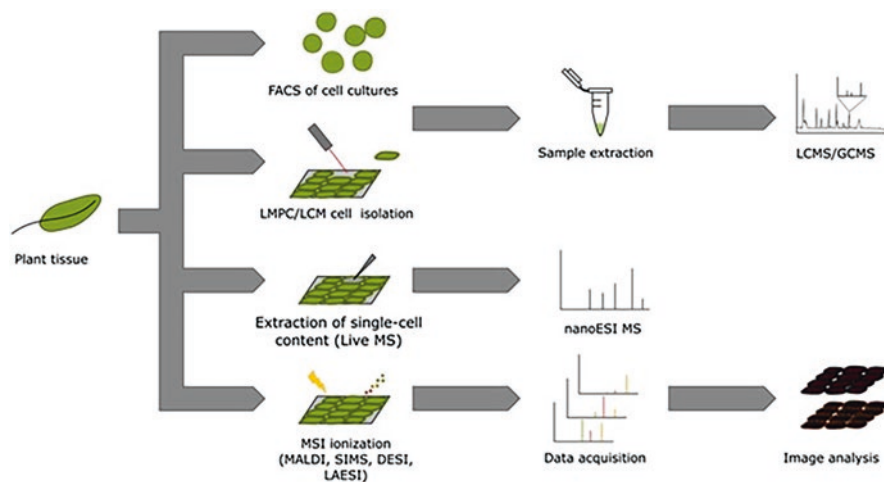


Fig. 5 Overview of experimental steps and data structure from the different approaches for cell-specific metabolomics. (Reproduced from Ref. [130])

However, to obtain single-cell suspension, it is a very challenging activity and deeply laborious. In addition, plant cells are rigid cells when compared to animal cells; rigid cell walls remain the main obstacle for single-cell technologies in plants. Since protoplasts must remain alive and be subjected to a minimal level of disturbance during isolation, for example, for protoplast isolation, the cell wall digestion procedure requires optimization for suitability for the tissue under study [133].

Mass spectrometry imaging (MSI) technique can provide spatially resolved information on the structure and content of metabolites including known and unknown endogenous metabolites, and it thus produces tissue molecular imaging maps. Three MSI techniques have been developed based on different ion sources: secondary ions MS (SIMS), desorption electrospray ionization (DESI), and matrix-assisted laser desorption/ionization (MALDI) [131]. Figure 6 summarizes different ionization techniques used for MSI. Among them, MALDI is the most popular ionization technique for MSI experiments. In MALDI mode a matrix applied to the sample is excited by a laser; this energy is further transferred to the sample resulting in the desorption/ionization event. Preparation for MALDI usually comprises cryo-sectioning and lyophilizing a frozen sample before applying the matrix by either a sprayer or a special device, as well solvent free sublimation [130]. However, MALDI remains a technique that still lacks significant improvement, for example, matrix selection and the choice of matrix method application, tissue sectioning technique, embedding protocols, sample preparation, and mounting. In other words, MALDI imaging technique requires optimization for every tissue and metabolites of interest; for example, Pérez-López et al. developed a protocol of MALDI imaging by sample imprinting in nylon membranes to locate fructans in stem and rhizome tissues of *Agave tequilana* plants [134]; in addition, the combination of ion mobility

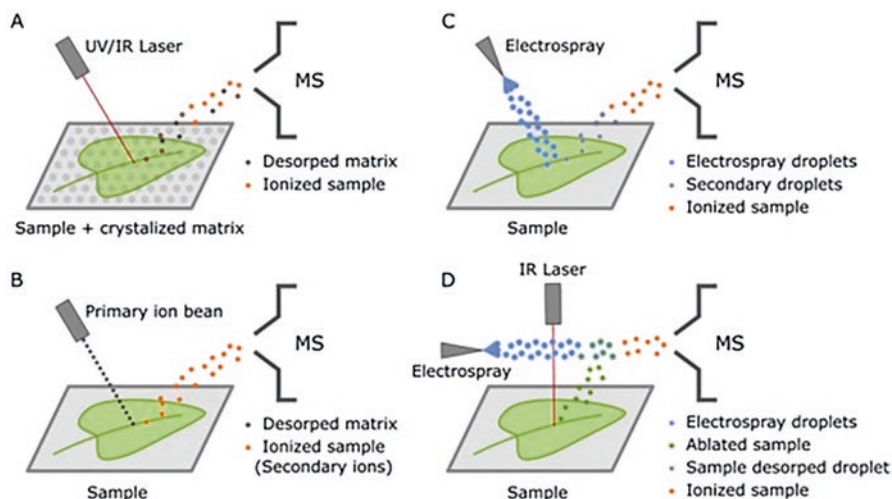


Fig. 6 Schematic representation of the different ionization strategies used for mass spectrometry imaging (MSI). (a) MALDI, (b) secondary ion mass spectrometry (SIMS), (c) desorption electrospray ionization (DESI), (d) laser-ablation electrospray ionization (LAESI). (Reproduced from Ref. [130])

spectrometry allowed the detection of fructan isomers even if these have not been mapped on their images obtained; Fig. 7 outlines the protocol developed.

More recently, DESI have become the newest development for mass spectrometry imaging to visualize plant metabolites; DESI offers a great advantage being matrix-free ionization alternative to MALDI. In DESI, a solvent stream originating from an electrospray probe is directed at an angle (most important parameter) toward the sample at ambient pressure, propelling secondary ions to the mass analyzer, enable direct analysis of unprocessed frozen samples sections which simplify samples preparation [135, 136]. Very recently, some metabolites detected with DESI source ranged from monoterpenoid alkaloids, which were localized in several of the major parts of the *Rauvolfia tetraphylla* plant when analyzed by MALDI and DESI-MSI [137], alkaloids in the leaves of *Gelsemium elegans* [138] through cannabinoids and flavonoids in the leaves of *Cannabis sativa* [139], among other recent applications of DESI.

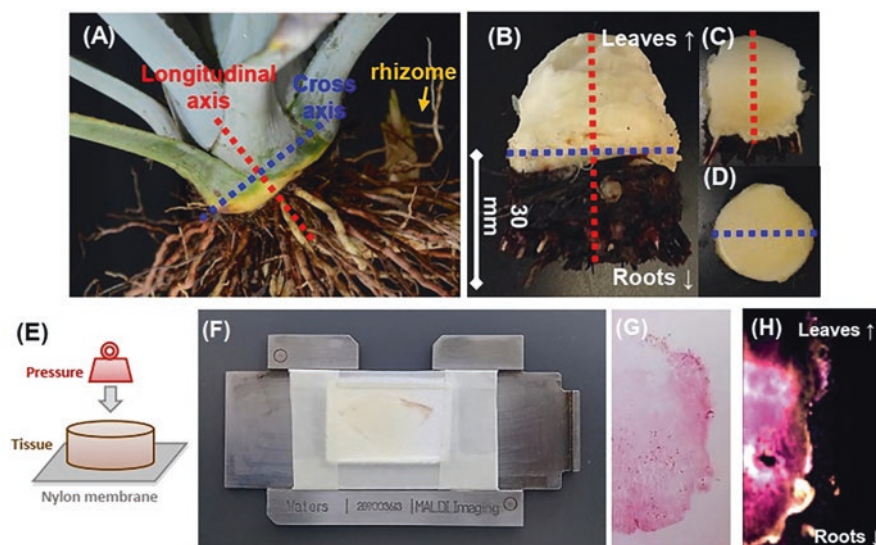


Fig. 7 Outline of tissue printing technique. (a) *Agave tequilana* plant showing the crown region (leaf/stem to root transition). Dotted red line indicates the longitudinal axis; dotted blue line indicates the transversal axis. (b) *A. tequilana* plant as in (a) with leaves and most roots removed and dissected longitudinally. (c) Longitudinal stem section. (d) Transverse stem section. (e) Representation of tissue printing process. (f) Tissue-printed transverse section mounted on MALDI Imaging plate. (g) PAS staining of a tissue-printed longitudinal section. (h) MALDI-ToF-MSI of a tissue-printed longitudinal section obtained, using a sprayer for matrix application and a QTOF SYNAPT G1 spectrometer with a spatial resolution of 100 pm per pixel. (Reproduced from Ref. [134])

7 Concluding Remarks

The utilization of a model plant like *Arabidopsis* or a crop with a reference genome sequence, such as - maize, rice, tomato, and wheat, to mention some of the global agricultural importance - offers a unique opportunity, where approaches such as mQTL, GWAS, mGWAS, and transcriptomics can effectively provide a vast potential to reveal gene annotation and their functional characterization. On the contrary, enormous efforts must be made for other minor crops, medicinal plants, and regional staple foods. However, plant metabolomics research needs to be focused on developing strategies to develop confident metabolite annotation through implementing a free online accessible database for metabolite identification. Plant metabolomics showed, as reviewed here, great potential to assist crop improvements, supporting exploring species identification for diversity and botanical purposes, food authentication, fraud, and traceability.

Without forgetting the knowledge of metabolome interaction between roots and their environment to identify the relevant metabolites produced, on the other hand, metabolomics approaches can play a crucial role in studying the interaction between

plants with biotic and abiotic stresses. Knowing the compounds produced by microorganisms and their identification in climate change conditions that enhance crop production and development is crucial.

Identifying and applying metabolic biomarkers could favor controlled and semi-controlled planting systems shortly, and if properly integrated into crop protection strategies, food security could be mitigated. However, the applications of these biomarkers could be helpful in various areas, such as the food and pharmaceutical industry in food quality and safety processes, diagnosis and treatment of plant diseases, and crop improvement. Still, the work done so far is relatively new, and efforts should continue to cover the tremendous potential presented by identifying metabolic biomarkers.

Plant metabolomics could benefit in developing new strategies to face challenges and demands in crop improvement. Spatiotemporal metabolomics can effectively support plant-soil interactions studies; the implementation of mass spectrometry imaging in combination with ion mobility spectroscopy could potentially reveal metabolites' location in plant tissue without the need for extraction, in addition, to providing isomer identification without forgetting the need to develop metabolite databases that can support full plant metabolome coverage.

References

1. Sumner, L.W., et al., *Modern plant metabolomics: advanced natural product gene discoveries, improved technologies, and future prospects*. Natural product reports, 2015. **32**(2): p. 212–229.
2. Singh, D.P., et al., *Metabolomics-Driven Mining of Metabolite Resources: Applications and Prospects for Improving Vegetable Crops*. International Journal of Molecular Sciences, 2022. **23**(20): p. 12062.
3. Wu, F., et al., *Genome and systems biology of *Melilotus albus* provides insights into coumarins biosynthesis*. Plant Biotechnology Journal, 2022a. **20**(3): p. 592–609.
4. Li, H., et al., *Genome-wide association study dissects the genetic architecture of oil biosynthesis in maize kernels*. Nature genetics, 2013. **45**(1): p. 43–50.
5. Zheng, C., et al., *Gene coexpression networks reveal key drivers of flavonoid variation in eleven tea cultivars (*Camellia sinensis*)*. Journal of agricultural and food chemistry, 2019. **67**(35): p. 9967–9978.
6. Chen, J., et al., *Exploring the genic resources underlying metabolites through mGWAS and mQTL in wheat: From large-scale gene identification and pathway elucidation to crop improvement*. Plant Communications, 2021a. **2**(4): p. 100216.
7. Chen, W., et al., *Comparative and parallel genome-wide association studies for metabolic and agronomic traits in cereals*. Nature communications, 2016. **7**(1): p. 1–10.
8. Burgos, E., et al., *Validated MAGIC and GWAS population mapping reveals the link between vitamin E content and natural variation in chorismate metabolism in tomato*. The Plant Journal, 2021. **105**(4): p. 907–923.
9. Alseikh, S., et al., *Identification and mode of inheritance of quantitative trait loci for secondary metabolite abundance in tomato*. The Plant Cell, 2015. **27**(3): p. 485–512.
10. Alseikh, S., et al., *Canalization of tomato fruit metabolism*. The Plant Cell, 2017. **29**(11): p. 2753–2765.

11. Liang, X., et al., *Metabolomics-driven gene mining and genetic improvement of tolerance to salt-induced osmotic stress in maize*. *New Phytologist*, 2021. **230**(6): p. 2355–2370.
12. Owens, B.F., et al., *A foundation for provitamin A biofortification of maize: genome-wide association and genomic prediction models of carotenoid levels*. *Genetics*, 2014. **198**(4): p. 1699–1716.
13. Riedelsheimer, C., et al., *Genome-wide association mapping of leaf metabolic profiles for dissecting complex traits in maize*. *Proceedings of the National Academy of Sciences*, 2012a. **109**(23): p. 8872–8877.
14. Chen, W., et al., *Genome-wide association analyses provide genetic and biochemical insights into natural variation in rice metabolism*. *Nature genetics*, 2014. **46**(7): p. 714–721.
15. Dong, X., et al., *Spatiotemporal distribution of phenolamides and the genetics of natural variation of hydroxycinnamoyl spermidine in rice*. *Molecular Plant*, 2015. **8**(1): p. 111–121.
16. Yang, Z., et al., *Genetic mapping identifies a rice naringenin O-glucosyltransferase that influences insect resistance*. *The Plant Journal*, 2021. **106**(5): p. 1401–1413.
17. Hong, J., et al., *Plant metabolomics: an indispensable system biology tool for plant science*. *International journal of molecular sciences*, 2016. **17**(6): p. 767.
18. Gilliland, L.U., et al., *Genetic basis for natural variation in seed vitamin E levels in Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences*, 2006. **103**(49): p. 18834–18841.
19. Rowe, H.C., et al., *Biochemical networks and epistasis shape the Arabidopsis thaliana metabolome*. *The Plant Cell*, 2008. **20**(5): p. 1199–1216.
20. Routaboul, J.-M., et al., *Metabolite profiling and quantitative genetics of natural variation for flavonoids in Arabidopsis*. *Journal of Experimental Botany*, 2012. **63**(10): p. 3749–3764.
21. Liseic, J., et al., *Identification of heterotic metabolite QTL in Arabidopsis thaliana RIL and IL populations*. *The Plant Journal*, 2009. **59**(5): p. 777–788.
22. Feng, J., et al., *Characterization of metabolite quantitative trait loci and metabolic networks that control glucosinolate concentration in the seeds and leaves of Brassica napus*. *New phytologist*, 2012. **193**(1): p. 96–108.
23. Wen, W., et al., *Genetic determinants of the network of primary metabolism and their relationships to plant performance in a maize recombinant inbred line population*. *The Plant Cell*, 2015. **27**(7): p. 1839–1856.
24. Matsuda, F., et al., *Dissection of genotype–phenotype associations in rice grains using metabolome quantitative trait loci analysis*. *The Plant Journal*, 2012. **70**(4): p. 624–636.
25. Ying, J.-Z., et al., *Identification of quantitative trait loci for lipid metabolism in rice seeds*. *Molecular plant*, 2012. **5**(4): p. 865–875.
26. Gong, L., et al., *Genetic analysis of the metabolome exemplified using a rice population*. *Proceedings of the National Academy of Sciences*, 2013. **110**(50): p. 20320–20325.
27. Schauer, N., et al., *Comprehensive metabolic profiling and phenotyping of interspecific introgression lines for tomato improvement*. *Nature biotechnology*, 2006. **24**(4): p. 447–454.
28. Toubiana, D., et al., *Metabolic profiling of a mapping population exposes new insights in the regulation of seed metabolism and seed, fruit, and plant relations*. *PLoS genetics*, 2012. **8**(3): p. e1002612.
29. Perez-Fons, L., et al., *A genome-wide metabolomic resource for tomato fruit from Solanum pennellii*. *Scientific Reports*, 2014. **4**(1): p. 3859.
30. Schauer, N., et al., *Mode of inheritance of primary metabolic traits in tomato*. *The Plant Cell*, 2008. **20**(3): p. 509–523.
31. Hill, C.B., et al., *Detection of QTL for metabolic and agronomic traits in wheat with adjustments for variation at genetic loci that affect plant phenology*. *Plant Science*, 2015. **233**: p. 143–154.
32. Hill, C.B., et al., *Whole-genome mapping of agronomic and metabolic traits to identify novel quantitative trait loci in bread wheat grown in a water-limited environment*. *Plant Physiology*, 2013. **162**(3): p. 1266–1281.

33. Angelovici, R., et al., *Genome-wide analysis of branched-chain amino acid levels in Arabidopsis seeds*. *The Plant Cell*, 2013. **25**(12): p. 4827–4843.
34. Chan, E.K.F., et al., *Combining genome-wide association mapping and transcriptional networks to identify novel genes controlling glucosinolates in Arabidopsis thaliana*. *PLoS biology*, 2011. **9**(8): p. e1001125.
35. Chan, E.K.F., et al., *The complex genetic architecture of the metabolome*. *PLoS genetics*, 2010. **6**(11): p. e1001198.
36. Shen, M., et al., *Leveraging non-targeted metabolite profiling via statistical genomics*. *PLoS One*, 2013. **8**(2): p. e57667.
37. Lipka, A.E., et al., *Genome-wide association study and pathway-level analysis of tocopherol levels in maize grain*. *G3: Genes, Genomes, Genetics*, 2013. **3**(8): p. 1287–1299.
38. Riedelsheimer, C., et al., *Genomic and metabolic prediction of complex heterotic traits in hybrid maize*. *Nature genetics*, 2012b. **44**(2): p. 217–220.
39. Wen, W., et al., *Metabolome-based genome-wide association study of maize kernel leads to novel biochemical insights*. *Nature communications*, 2014. **5**(1): p. 3438.
40. Carreno-Quintero, N., et al., *Untargeted metabolic quantitative trait loci analyses reveal a relationship between primary metabolism and potato tuber quality*. *Plant physiology*, 2012. **158**(3): p. 1306–1318.
41. Matsuda, F., et al., *Metabolome-genome-wide association study dissects genetic architecture for generating natural variation in rice secondary metabolism*. *The Plant Journal*, 2015. **81**(1): p. 13–23.
42. Sauvage, C., et al., *Genome-wide association in tomato reveals 44 candidate loci for fruit metabolic traits*. *Plant physiology*, 2014. **165**(3): p. 1120–1132.
43. Musah, R.A., et al., *A high throughput ambient mass spectrometric approach to species identification and classification from chemical fingerprint signatures*. *Scientific Reports*, 2015. **5**(1): p. 1–16.
44. Mannocho-Russo, H., et al., *Untargeted metabolomics sheds light on the diversity of major classes of secondary metabolites in the Malpighiaceae botanical family*. *Frontiers in Plant Science*, 2022. **13**.
45. Sumner, L.W., et al., *Proposed minimum reporting standards for chemical analysis: chemical analysis working group (CAWG) metabolomics standards initiative (MSI)*. *Metabolomics*, 2007. **3**: p. 211–221.
46. Zhang, M., et al., *Development of a Metabolite Ratio Rule-Based Method for Automated Metabolite Profiling and Species Differentiation of Four Major Cinnamon Species*. *Journal of Agricultural and Food Chemistry*, 2022a. **70**(17): p. 5450–5457.
47. Leyva, V.E., et al., *NMR-based leaf metabolic profiling of *V. planifolia* and three endemic *Vanilla* species from the Peruvian Amazon*. *Food Chemistry*, 2021. **358**: p. 129365.
48. Souard, F., et al., *Metabolomics fingerprint of coffee species determined by untargeted-profiling study using LC-HRMS*. *Food chemistry*, 2018. **245**: p. 603–612.
49. Lesiak, A.D., et al., *Plant seed species identification from chemical fingerprints: a high-throughput application of direct analysis in real time mass spectrometry*. *Analytical chemistry*, 2015. **87**(17): p. 8748–8757.
50. Qi, J., et al., *Untargeted Metabolomics Analysis Revealed the Major Metabolites in the Seeds of four Polygonatum Species*. *Molecules*, 2022. **27**(4): p. 1445.
51. Brahmi, F., et al., *Discrimination of *Mentha* species grown in different geographical areas of Algeria using ¹H-NMR-based metabolomics*. *Journal of Pharmaceutical and Biomedical Analysis*, 2020. **189**: p. 113430.
52. Cheng, Z., et al., *From folk taxonomy to species confirmation of *Acorus* (*Acoraceae*): evidences based on phylogenetic and metabolomic analyses*. *Frontiers in plant science*, 2020: p. 965.
53. Kiran, K.R., et al., *Untargeted metabolomics and DNA barcoding for discrimination of *Phyllanthus* species*. *Journal of Ethnopharmacology*, 2021. **273**: p. 113928.

54. Abbass, K., et al., *A review of the global climate change impacts, adaptation, and sustainable mitigation measures*. Environmental Science and Pollution Research, 2022: p. 1–21.
55. García, G.A., et al., *High night temperatures during grain number determination reduce wheat and barley grain yield: a field study*. Global Change Biology, 2015. **21**(11): p. 4153–4164.
56. Reidsma, P., et al., *Regional crop modelling in Europe: the impact of climatic conditions and farm characteristics on maize yields*. 2009. **100**(1-3): p. 51–60.
57. Santoro, A., et al., *A review of the role of forests and agroforestry systems in the FAO Globally Important Agricultural Heritage Systems (GIAHS) programme*. Forests, 2020. **11**(8): p. 860.
58. Munaweera, T., et al., *Modern plant biotechnology as a strategy in addressing climate change and attaining food security*. 2022. **11**(1): p. 1–28.
59. Ortiz, A.M.D., et al., *A review of the interactions between biodiversity, agriculture, climate change, and international trade: Research and policy priorities*. 2021. **4**(1): p. 88–101.
60. Chandio, A.A., et al., *Short and long-run impacts of climate change on agriculture: an empirical evidence from China*. 2020. **12**(2): p. 201–221.
61. Kausch, A.P., et al., *Edit at will: genotype independent plant transformation in the era of advanced genomics and genome editing*. Plant science, 2019. **281**: p. 186–205.
62. Wang, M.-B., et al., *RNA silencing and plant viral diseases*. Molecular Plant-Microbe Interactions, 2012. **25**(10): p. 1275–1285.
63. Hong, Y., et al., *Crafting the plant root metabolome for improved microbe-assisted stress resilience*. New Phytologist, 2022.
64. Chen, L., et al., *Metabolomics in plant-microbe interactions in the roots*. Advances in Botanical Research, 2021b. **98**: p. 133–161.
65. Hacquard, S., et al., *Impact of global change on the plant microbiome*. New Phytologist, 2022. **234**(6): p. 1907–1909.
66. Xun, W., et al., *Rhizosphere microbiome: Functional compensatory assembly for plant fitness*. Computational and Structural Biotechnology Journal, 2021. **19**: p. 5487–5493.
67. Pantigoso, H.A., D. Newberger, and J.M. Vivanco, *The rhizosphere microbiome: Plant-microbial interactions for resource acquisition*. Journal of Applied Microbiology, 2022.
68. van Dam, N.M. and H.J. Bouwmeester, *Metabolomics in the rhizosphere: tapping into belowground chemical communication*. Trends in plant science, 2016. **21**(3): p. 256–265.
69. Saeed, Q., et al., *Rhizosphere bacteria in plant growth promotion, biocontrol, and bioremediation of contaminated sites: a comprehensive review of effects and mechanisms*. International Journal of Molecular Sciences, 2021. **22**(19): p. 10529.
70. Wang, W., et al., *Transcriptomic and metabolomics-based analysis of key biological pathways reveals the role of lipid metabolism in response to salt stress in the root system of Brassica napus*. Plant Growth Regulation, 2022. **97**(1): p. 127–141.
71. Baldwin, M.A., *Mass spectrometers for the analysis of biomolecules*. Methods in enzymology, 2005. **402**: p. 3–48.
72. Wang, J., et al., *Non-targeted metabolomics reveal the impact of phenanthrene stress on root exudates of ten urban greening tree species*. Environmental Research, 2021. **196**: p. 110370.
73. Lu, M., et al., *Physiological and metabolomics responses of two wheat (Triticum aestivum L.) genotypes differing in grain cadmium accumulation*. Science of The Total Environment, 2021. **769**: p. 145345.
74. Kalu, C.M., et al., *Correlations between root metabolomics and bacterial community structures in the phragmites australis under acid mine drainage-Polluted wetland ecosystem*. Current Microbiology, 2022. **79**(1): p. 1–15.
75. Mannucci, A., et al., *Foliar and Root Comparative Metabolomics and Phenolic Profiling of Micro-Tom Tomato (Solanum lycopersicum L.) Plants Associated with a Gene Expression Analysis in Response to Short Daily UV Treatments*. Plants, 2022. **11**(14): p. 1829.
76. Hazrati, H., I.S. Fomsgaard, and P. Kudsk, *Targeted metabolomics unveil alteration in accumulation and root exudation of flavonoids as a response to interspecific competition*. Journal of Plant Interactions, 2021. **16**(1): p. 53–63.

77. Zhang, Z., et al., *Comparative metabolomics reveal the participation of soybean unique rhizosphere metabolites in susceptibility and resistance of host soybean to Phytophthora sojae*. Plant and Soil, 2022b: p. 1–15.
78. Zhang, N., et al., *Metabolomic profiling reveals suppression of oxylipin biosynthesis during the early stages of legume–rhizobia symbiosis*. FEBS letters, 2012. **586**(19): p. 3150–3158.
79. Sebastiana, M., et al., *Metabolomics and transcriptomics to decipher molecular mechanisms underlying ectomycorrhizal root colonization of an oak tree*. Scientific reports, 2021. **11**(1): p. 1–16.
80. Hernández, G., et al., *Global changes in the transcript and metabolic profiles during symbiotic nitrogen fixation in phosphorus-stressed common bean plants*. Plant Physiology, 2009. **151**(3): p. 1221–1238.
81. Gomez-Zepeda, D., et al., *Mass spectrometry-based quantification and spatial localization of small organic acid exudates in plant roots under phosphorus deficiency and aluminum toxicity*. The Plant Journal, 2021. **106**(6): p. 1791–1806.
82. Li, H., et al., *Metabolomics and microbiome reveal potential root microbiota affecting the alkaloidal metabolome in Aconitum vilmorinianum Kom*. BMC microbiology, 2022. **22**(1): p. 1–15.
83. Castro-Santos, P., C.M. Laborde, and R. Díaz-Peña, *Genomics, proteomics and metabolomics: their emerging roles in the discovery and validation of rheumatoid arthritis biomarkers*. Clin Exp Rheumatol, 2015. **33**(2): p. 279–286.
84. Dunn, W.B., N.J. Bailey, and H.E. Johnson, *Measuring the metabolome: current analytical technologies*. Analyst, 2005. **130**(5): p. 606–625.
85. Maroli, A.S., et al., *Omics in weed science: a perspective from genomics, transcriptomics, and metabolomics approaches*. Weed Science, 2018. **66**(6): p. 681–695.
86. Hartmann, T., *From waste products to ecochemicals: fifty years research of plant secondary metabolism*. Phytochemistry, 2007. **68**(22–24): p. 2831–2846.
87. Sakurai, N., *Recent applications of metabolomics in plant breeding*. Breeding Science, 2022: p. 21065.
88. Marchand, C.R., et al., *A framework for development of useful metabolomic biomarkers and their effective knowledge translation*. Metabolites, 2018. **8**(4): p. 59.
89. Lemaitre-Guillier, C., et al., *VOCs are relevant biomarkers of elicitor-induced defences in grapevine*. Molecules, 2021. **26**(14): p. 4258.
90. Watanabe, M., et al., *Comprehensive dissection of spatiotemporal metabolic shifts in primary, secondary, and lipid metabolism during developmental senescence in Arabidopsis*. Plant physiology, 2013. **162**(3): p. 1290–1310.
91. Lin, Z.-H., et al., *The GC-TOF/MS-based Metabolomic analysis reveals altered metabolic profiles in nitrogen-deficient leaves and roots of tea plants (Camellia sinensis)*. BMC Plant Biology, 2021. **21**(1): p. 1–13.
92. Dauda, W.P., et al., *Metabolomic analysis of sheath blight disease of rice (Oryza sativa L.) induced by Rhizoctonia solani phytotoxin*. Journal of Applied Microbiology, 2022.
93. Tarpley, L., et al., *Biomarker metabolites capturing the metabolite variance present in a rice plant developmental period*. BMC plant biology, 2005. **5**(1): p. 1–12.
94. Song, H.-H., et al., *Metabolomics investigation of flavonoid synthesis in soybean leaves depending on the growth stage*. Metabolomics, 2014. **10**(5): p. 833–841.
95. Rivera-Pérez, A., R. Romero-González, and A.G. Frenich, *Fingerprinting based on gas chromatography-Orbitrap high-resolution mass spectrometry and chemometrics to reveal geographical origin, processing, and volatile markers for thyme authentication*. Food Chemistry, 2022: p. 133377.
96. Ivanović, S., et al., *Plant Metabolomics as a Tool for Detecting Adulterants in Edible Plant: A Case Study of Allium ursinum*. Metabolites, 2022. **12**(9): p. 849.
97. Adeniji, A.A. and O.O. Babalola, *Metabolomic applications for understanding complex tripartite plant-microbes interactions: Strategies and perspectives*. Biotechnology reports, 2020. **25**: p. e00425.

98. Khan, N., A. Bano, and M.A. Babar, *Metabolic and physiological changes induced by plant growth regulators and plant growth promoting rhizobacteria and their impact on drought tolerance in Cicer arietinum L.* PloS one, 2019. **14**(3): p. e0213040.
99. Razzaq, A., et al., *Metabolomics: A way forward for crop improvement.* Metabolites, 2019. **9**(12): p. 303.
100. Chen, B.-X., et al., *Identification of metabolomic biomarkers of seed vigor and aging in hybrid rice.* Rice, 2022. **15**(1): p. 1–12.
101. Dong, T., et al., *Altitudinal Variation of Metabolites, Mineral Elements and Antioxidant Activities of Rhodiola crenulata (Hook. f. & Thomson) H. Ohba.* Molecules, 2021. **26**(23): p. 7383.
102. Zeiss, D.R., et al., *Altered metabolomic states elicited by Flg22 and FlgII-28 in Solanum lycopersicum: intracellular perturbations and metabolite defenses.* BMC plant biology, 2021a. **21**(1): p. 1–17.
103. Zeiss, D.R., L.A. Piater, and I.A. Dubery, *Hydroxycinnamate amides: Intriguing conjugates of plant protective metabolites.* Trends in plant science, 2021b. **26**(2): p. 184–195.
104. Patel, M.K., et al., *Plants metabolome study: Emerging tools and techniques.* Plants, 2021. **10**(11): p. 2409.
105. Patel, M.K., A. Mishra, and B. Jha, *Non-targeted metabolite profiling and scavenging activity unveil the nutraceutical potential of psyllium (Plantago ovata Forsk).* Frontiers in plant science, 2016. **7**: p. 431.
106. Patel, M.K., et al., *Metabolic profiling and scavenging activities of developing circumscissile fruit of psyllium (Plantago ovata Forsk) reveal variation in primary and secondary metabolites.* BMC plant biology, 2020. **20**(1): p. 1–15.
107. Sinha, P., et al., *Changes in oil content and fatty acid composition in Jatropha curcas during seed development.* Industrial Crops and Products, 2015. **77**: p. 508–510.
108. Zhang, Q.-Y., et al., *Fatty acid and associated gene expression analyses of three tree peony species reveal key genes for α -linolenic acid synthesis in seeds.* Frontiers in plant science, 2018. **9**: p. 106.
109. Pandey, S., et al., *Nitric oxide accelerates germination via the regulation of respiration in chickpea.* Journal of experimental botany, 2019. **70**(17): p. 4539–4555.
110. Cui, M.-C., et al., *Metabolic profiling investigation of Fritillaria thunbergii Miq. by gas chromatography–mass spectrometry.* Journal of food and drug analysis, 2018. **26**(1): p. 337–347.
111. Dias, D.A., et al., *Quantitative profiling of polar primary metabolites of two chickpea cultivars with contrasting responses to salinity.* Journal of Chromatography B, 2015. **1000**: p. 1–13.
112. Mikołajczyk-Bator, K., et al., *Characterisation and identification of triterpene saponins in the roots of red beets (Beta vulgaris L.) using two HPLC–MS systems.* Food Chemistry, 2016. **192**: p. 979–990.
113. Hazzoumi, Z., Y. Moustakime, and K. Amrani Joutei, *Effect of gibberellic acid (GA), indole acetic acid (IAA) and benzylaminopurine (BAP) on the synthesis of essential oils and the isomerization of methyl chavicol and trans-anethole in Ocimum gratissimum L.* SpringerPlus, 2014. **3**(1): p. 1–7.
114. Santoro, V.M., et al., *Effects of growth regulators on biomass and the production of secondary metabolites in peppermint (Mentha piperita) micropropagated in vitro.* 2013.
115. Naem, M., et al., *Triacontanol-mediated regulation of growth and other physiological attributes, active constituents and yield of Mentha arvensis L.* Plant Growth Regulation, 2011. **65**: p. 195–206.
116. Li, J. and Z. Hu, *Accumulation and dynamic trends of triterpenoid saponin in vegetative organs of Achyranthes bidentata.* Journal of integrative plant biology, 2009. **51**(2): p. 122–129.
117. Perkowska, I., et al., *Identification and quantification of coumarins by UHPLC-MS in Arabidopsis thaliana natural populations.* Molecules, 2021. **26**(6): p. 1804.

118. Morita, H., et al., *New picrotoxinin-type and dendrobine-type sesquiterpenoids from Dendrobium snowflake 'Red Star'*. *Tetrahedron*, 2000. **56**(32): p. 5801–5805.
119. Ma, C., et al., *Terpenoid metabolic profiling analysis of transgenic Artemisia annua L. by comprehensive two-dimensional gas chromatography time-of-flight mass spectrometry*. *Metabolomics*, 2009. **5**: p. 497–506.
120. Sobolev, A.P., et al., *Quality traits of conventional and transgenic lettuce (Lactuca sativa L.) at harvesting by NMR metabolic profiling*. *Journal of agricultural and food chemistry*, 2010. **58**(11): p. 6928–6936.
121. Roessner-Tunali, U., et al., *Metabolic profiling of transgenic tomato plants overexpressing hexokinase reveals that the influence of hexose phosphorylation diminishes during fruit development*. *Plant Physiology*, 2003. **133**(1): p. 84–99.
122. Jha, R.K., et al., *Introgression of a novel cold and drought regulatory-protein encoding CORA-like gene, SbCDR, induced osmotic tolerance in transgenic tobacco*. *Physiologia Plantarum*, 2021. **172**(2): p. 1170–1188.
123. Zhou, J., et al., *Metabolic profiling of transgenic rice with cryIAc and sck genes: An evaluation of unintended effects at metabolic level by using GC-FID and GC-MS*. *Journal of Chromatography B*, 2009. **877**(8-9): p. 725–732.
124. Chang, Y., et al., *Metabolic profiling based on LC/MS to evaluate unintended effects of transgenic rice with cryIAc and sck genes*. *Plant Molecular Biology*, 2012. **78**: p. 477–487.
125. Payyavula, R.S., et al., *Metabolic profiling reveals altered sugar and secondary metabolism in response to UGPase overexpression in Populus*. *BMC Plant Biology*, 2014. **14**: p. 1–14.
126. Iwaki, T., et al., *Metabolic profiling of transgenic potato tubers expressing Arabidopsis dehydration response element-binding protein 1A (DREB1A)*. *Journal of agricultural and food chemistry*, 2013. **61**(4): p. 893–900.
127. Stamova, B.S., et al., *Metabolic profiling of transgenic wheat over-expressing the high-molecular-weight Dx5 glutenin subunit*. *Metabolomics*, 2009. **5**: p. 239–252.
128. Niu, F., et al., *Metabolic profiling of DREB-overexpressing transgenic wheat seeds by liquid chromatography–mass spectrometry*. *The Crop Journal*, 2020. **8**(6): p. 1025–1036.
129. Piccioni, F., et al., *NMR metabolic profiling of transgenic maize with the CryIA (b) gene*. *Journal of agricultural and food chemistry*, 2009. **57**(14): p. 6041–6049.
130. de Souza, L.P., M. Borghi, and A. Fernie, *Plant single-cell metabolomics—challenges and perspectives*. *International journal of molecular sciences*, 2020. **21**(23): p. 8987.
131. Yu, X., Z. Liu, and X. Sun, *Single-cell and spatial multi-omics in the plant sciences: Technical advances, applications, and perspectives*. *Plant Communications*, 2022: p. 100508.
132. Misra, B.B., S.M. Assmann, and S. Chen, *Plant single-cell and single-cell-type metabolomics*. *Trends in Plant Science*, 2014. **19**(10): p. 637–646.
133. Mo, Y. and Y. Jiao, *Advances and applications of single-cell omics technologies in plant research*. *The Plant Journal*, 2022. **110**(6): p. 1551–1563.
134. Pérez-López, A.V., et al., *Localization and Composition of Fructans in Stem and Rhizome of Agave tequilana Weber var. azul*. *Frontiers in plant science*, 2021. **11**: p. 2309.
135. Takats, Z., J.M. Wiseman, and R.G. Cooks, *Ambient mass spectrometry using desorption electrospray ionization (DESI): instrumentation, mechanisms and applications in forensics, chemistry, and biology*. *Journal of mass spectrometry*, 2005. **40**(10): p. 1261–1275.
136. Takats, Z., et al., *Mass spectrometry sampling under ambient conditions with desorption electrospray ionization*. *Science*, 2004. **306**(5695): p. 471–473.
137. Lorensen, M.D.B.B., et al., *Spatial localization of monoterpenoid indole alkaloids in Rauvolfia tetraphylla by high resolution mass spectrometry imaging*. *Phytochemistry*, 2023a: p. 113620.
138. Wu, Z., et al., *A mass spectrometry imaging approach on spatiotemporal distribution of multiple alkaloids in Gelsemium elegans*. *Frontiers in Plant Science*, 2022b. **13**.
139. Lorensen, M.D.B.B., et al., *Leaves of Cannabis sativa and their trichomes studied by DESI and MALDI mass spectrometry imaging for their contents of cannabinoids and flavonoids*. *Phytochemical Analysis*, 2023b.