 doi:10.20944/preprints202107.0193.v1

Review

Plants metabolome study: emerging tools and techniques

Manish Kumar Patel[1]*#, Sonika Pandey[2#], Manoj Kumar[3], Intesaful Haque[4], Sikander PAL[5] Narendra Singh Yadav[6]*

[1] Department of Postharvest Science of Fresh Produce, Agricultural Research Organization, Volcani Center, P.O. Box 15159, HaMaccabim Road 68, Rishon LeZion, 7505101, Israel; patelm1402@gmail.com (M.K.P)
[2] Independent scholar, India; sonikapandey14@gmail.com (S.P)
[3] Institute of Plant Sciences, Agricultural Research Organization, Volcani Center, Rishon LeZion, 7505101 Israel; manojbiochem16@gmail.com (M.K).
[4] Fruit tree science department, Newe Ya'ar Research Center, Agriculture Research Organization, Volcani Center. Ramat Yishay, 3009500, Israel; intesafulhaque@gmail.com (M.I.H)
[5] Plant Physiology Laboratory, Department of Botany, University of Jammu, Jammu, 180006, India, sikanderpal@jammuuniversity.ac.in
[6] Department of Biological Sciences, University of Lethbridge, Lethbridge T1K 3M4, Alberta, Canada; nsyadava2004@gmail.com

* Correspondence: patelm1402@gmail.com (M.K.P); nsyadava2004@gmail.com (N.S.Y)
# Equal contribution: Manish Kumar Patel (M.K.P); Sonika Pandey (S.P)

Abstract:

Metabolomics is now considered to be a wide-ranging, sensitive and practical approach to acquire useful information on the composition of a metabolite pool present in any organism, including plants. Investigating metabolomic regulation in plants is essential to understand their adaptation, acclimation and defense response to environmental stresses through the production of numerous metabolites. Moreover, metabolomics can be easily applied for the phenotyping of plants; and thus, it has great potential to be used in molecular breeding and genome editing programs to develop superior next generation crops. This review describes the recent analytical tools and techniques available to study plants metabolome, along with their significance of sample preparation using targeted and non-targeted method. Advanced analytical tools, like gas chromatography-mass spectrometry (GC-MS), liquid chromatography mass-spectroscopy (LC-MS), capillary electrophoresis-mass spectrometry (CE-MS), fourier transform ion cyclotron resonance-mass spectrometry (FTICR-MS) and matrix-assisted laser desorption/ionization (MALDI) have speed up metabolic profiling in plants. Further, we deliver a complete overview of bioinformatics tools and plant metabolome database that can be utilized to advance our knowledge to plant biology.

Keywords: metabolomics; plant biology; metabolomics databases; data analysis; metabolomics software tools; mass spectrometry; omics

1. Metabolomics: plant biology perspective

Metabolomics is one of the fastest developing and attractive discipline of omics field, which has huge potential to be used in crop improvement programs. It is vital to reviewing the abiotic/biotic stress tolerances and metabolomics assisted-breeding of crop plants [1]. Primary metabolites are essential for plant growth and development as they are involved in various physiological and biochemical processes [2]. Primary metabolites include different classes of metabolites such as sugars, polyols and amino acids, serving as vital functions such as osmolytes and osmoprotectants in plants under biotic and abiotic stresses [3,4]. Secondary metabolites (SMs) play a key role in protecting plants against various environmental stresses. It has been estimated that approximately 100,000 SMs have been reported within different plant species and are classified into various groups,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 July 2021 doi:10.20944/preprints202107.0193.v1

nitrogen-containing compounds, terpenes, thiols and phenolic compounds [5]. Recent metabolomics platforms play key role in exploration of unknown regulatory networks that control plant growth and development [1]. Additional crucial metabolomics application, called ecological metabolomics, deals with the study of the biochemical interactions among plants across different temporal and spatial networks [6]. It describes the biochemical nature of various vital ecological phenomena, such as the effects of parasite load, incidence of disease and infection. It also helps to decode the potential effect of AS and BS on any critical biochemical process through the detection of metabolites [1]. Modern metabolomics platforms are being exploited to explain complex biological pathways and explore hidden regulatory networks controlling crop growth and health.

The performance of metabolomics study relies on its methodologies and instruments to comprehensively identify and measure each metabolite [7]. In fact, it is very difficult because of the complexity of the various metabolic characteristics and molecular abundances. Metabolomics or metabolite profiling terms are alternatively used to define three types of approaches, such as untargeted metabolomics, targeted metabolomics and semi-targeted metabolomics [3,8]. A number of integrated technologies and methodologies such as mass spectrometry (MS) based methods including GC–MS, LC–MS, CE-MS, FTICR-MS and MALDI are currently being used for large-scale analysis of highly complex mixtures of plant extracts [9]. In fact, these analytical methods have shown their potential in many plant species, including halophytes, medicinal plants and food crops such as *Salicornia brachiata*, *Plantago ovata*, *Cuminum cyminum*, *Solanum lycopersicum*, *Oryza sativa*, *Triticum aestivum* and *Zea mays* [10,11,12,13,14,15,16,17]. However, integrated methods are increasingly used in metabolomics analysis, due to the inherent weakness of each analytical platform [7].

Last decades have witnessed major developments in various 'Omics' fields, such as genomics, transcriptomics, proteomics, metabolomics and phenomics. The various omics platforms have an endless potential to enhance the current understanding of complex biological pathways, allowing us to develop new approaches for crops improvement [2]. Metabolomics is one of the most complex approach among other omics approaches and has received attention in agriculture science, especially for plant selections in molecular breeding program. Therefore, metabolomics is used to acquire a vast amount of useful knowledge by accurate and high throughput peak annotation through the snapshot of the plant metabolome for the novel genes and pathways elucidation [18]. The combination of metabolomic integrated with transcriptomic analysis, was successfully used to find out several possible approaches such as breeding, and genome editing involved in the activation of metabolic pathways and gene expression [19]. Nevertheless, plant metabolomics has become an effective tool for exploring different aspects of system biology, which is greatly expanding our knowledge of the metabolic and signaling pathways in plant growth, development and response to stress for improving the quality and yield of crops [20]. This review describes the plant metabolomics including different analytical techniques used, bioinformatics tools and plant metabolome database.

2. Significance of sample preparation in plant metabolites

In plant metabolomics study, plant samples are harvested, stored, metabolites extraction and quantification, followed by data interpretation. Sample preparation is a key step in plant metabolomics as it significantly changes the quantity of the metabolites. Thus, considering all the factors, harvesting and storage of plant samples should be fast as to reduce the changes of biochemical reaction in the plant cells [21]. Inappropriate handling during the sample collection is the most likely source of bias in plant metabolomic studies [22]. Sample harvesting, storage and extract preparation should ideally follow the Metabolomics Standards Initiative (MSI) in order to justify plant metabolomics studies [23].

2.1. Sample harvesting and storage

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 July 2021 doi:10.20944/preprints202107.0193.v1

Commonly, four major steps are involved in plant metabolomics; harvesting, storage, extraction and sample analysis (Figure 1). Plant sample harvesting must be carried out with caution, as the metabolome of the plant is sensitive to enzymatic reactions that can degrade different metabolites. In addition, metabolites vary with the different development stages, plant age and time of sample harvesting [9]. Mostly, 10-100 mg of plant samples are required for each biological sample in metabolomics studies. Usually, immediately after harvesting, the plant samples are snap frozen in liquid nitrogen to prevent metabolic changes. Similarly, various storage techniques, such as freeze-drying, oven-drying and air-drying, are essential for the processing of metabolomics [22, 24].

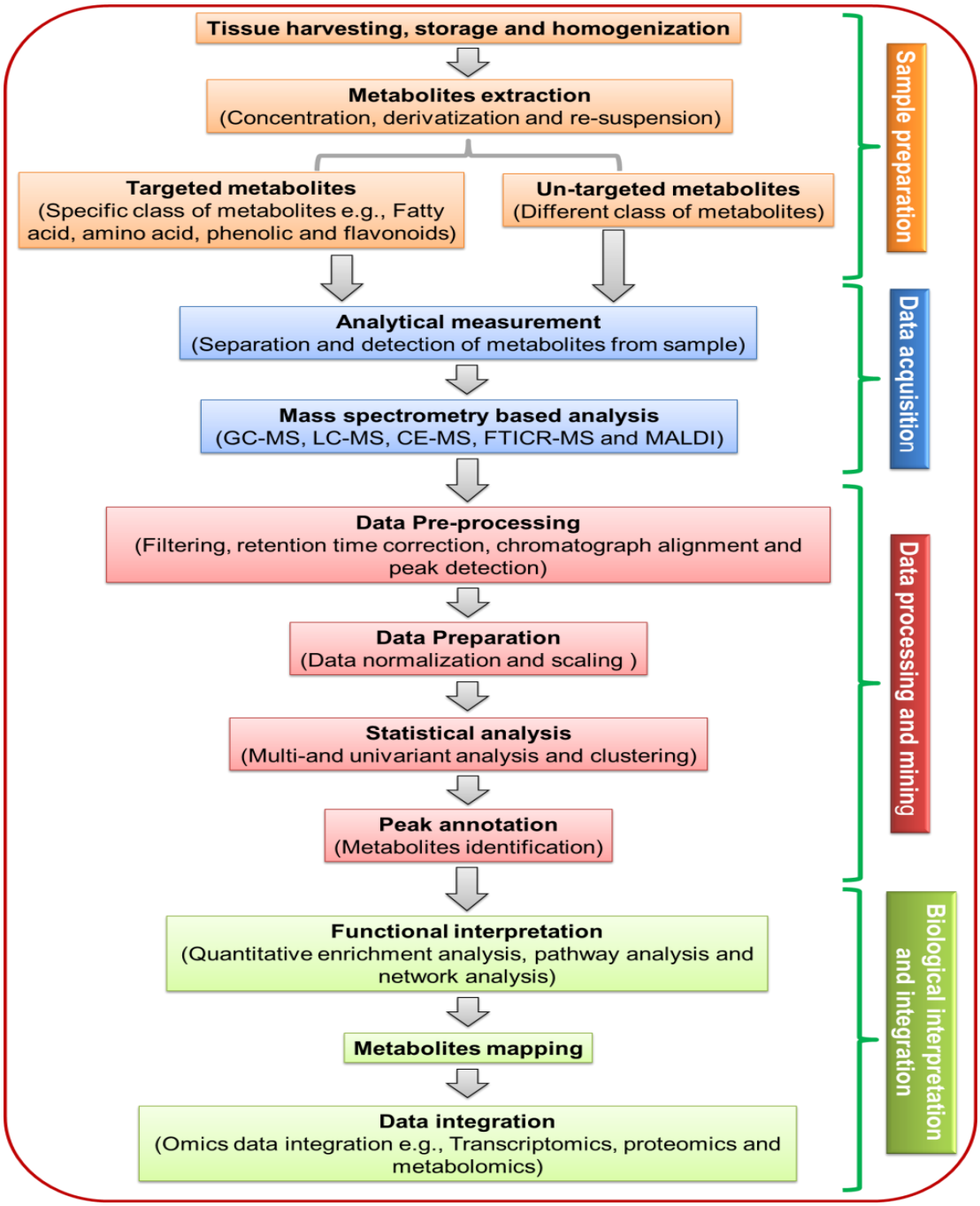

Figure 1. Schematic representation of the multi-step workflow of a plant metabolomics study. Sample preparation, data acquisition, data processing and biological interpretation are key steps in plant metabolomics. Now days, for data acquisition, different MS-based analytical tools (GC-MS, LC-MS CE-MS, FTICR-MS and MALDI) are available. The most important step data processing and mining includes correction of baseline shifts, background noise reduction, chromatograph

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 July 2021 doi:10.20944/preprints202107.0193.v1

alignment and peaks detection. Biological interpretation and integration includes enrichment analysis, networks and pathways analysis for a comprehensive scope of the metabolome. GC-MS, gas chromatography-mass spectrometry; LC-MS, liquid chromatography mass-spectroscopy; CE-MS, capillary electrophoresis-mass spectrometry; FTICR-MS, fourier transform ion cyclotron resonance-mass spectrometry; MALDI, matrix-assisted laser desorption/ionization.

2.2. Sample Preparation

Sample preparation plays a key role in metabolomic study, as it includes extraction of metabolites using different extraction methods (Figure 1). Among the extraction methods quenching, mechanical and ultrasound extraction methods are promising in metabolomic analysis [25]. In addition, high quality, yield and chemical versatility can be obtained by integrating ultrasound extraction method and mechanical grinding [26]. Apart from extraction methods, the choice of solvents is also crucial, as single solvent cannot extract a variety of metabolites (e.g. polar or non-polar). The isolation of a wide variety of metabolites can be achieved using a solvent system composed of chloroform: methanol: water [27,28]. This solvent system is widely used for wide variety of metabolites such as polar compounds, nonpolar compounds and hydrophilic metabolites. Diverse solvent systems were reported for the plant metabolomics such as extraction with pure methanol [29,30], the mixture of methanol: water [31] and methanol: methyl-tert-butyl-ether: water [32]. A specific solvent gradient extraction method was developed to recover almost all type metabolites in a single protocol [33]. In addition, hot methanol (70% v/v) was used for the extraction of phenolic compounds from *Brassica oleraceae* using ultra-high-performance liquid chromatography–diode array detector–tandem mass spectrometry [34]. Various methods are used for sample preparation, such as microwave-assisted extraction [35], ultrasound-assisted extraction [36], Swiss rolling technique [37] and enzyme-assisted extraction [38]. In plant metabolomics, new extraction methods are also developing day by day in line depending on the nature of the compounds and selection of analytical systems.

3. Analytical techniques used for plant metabolome

Along with sample preparation, different MS-based analytical systems are available for data acquisition. In plant metabolomics, single analytical tools cannot be used to identify all the metabolites present in a sample; instead, a set of various techniques are needed to provide the largest amount of metabolite coverage [1]. Various metabolomics tools include MS based technique namely GC-MS, LC-MS CE-MS, FTICR-MS and MALDI for sensitive and specific qualitative and quantitative analyses of metabolites (Figure 1) [9,39]. All five mentioned analytical methods identifying metabolites in plant tissue directly or indirectly have advantages and disadvantages (Table 1). Also, the combination of analytical methods can be used to ensure the efficacy of metabolite profiling.

Table 1: Advantages and disadvantages of common analytical techniques used in MS-based metabolomics.

Analytical method	Advantage	Disadvantage
GC-MS	• Suitable for the identification of volatile compounds • Large commercial and public libraries • Identification and quantification of small metabolites (~500 daltons)	• Sample pre-processing process and requires derivatization • Many metabolites thermally-unstable or unsuitable for non-volatile compounds
LC-MS	• Easy sample preparation • No derivatization • Several separation modes are available • Multiple MS detectors	• Few commercial libraries • Adduct ions are needed for metabolites detection

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 July 2021 doi:10.20944/preprints202107.0193.v1

	• Large number of detectable metabolites	
CE-MS	• Evaluating ionic metabolites based on the proportion of charge and size ratio • Fast and high-resolution of charged compounds • No derivatization	• Low sensitivity and reproducibility • Poor migration time and lack of reference libraries
FTICR-MS	• Mass resolving power • Detection sensitivity and mass accuracy	• Expensive • Lack of detection for non-ionizable compounds • Slow MS/MS
MALDI-MSI	• Quantification by peak intensities • Resolution up to 10 mm • Direct on tissue identification by tandem–MS fragmentation • Mass range up to 20 kDa	• Unsuitable for higher molecular mass compound • Expensive equipment to purchase • Time consuming • Limited by size of the metabolites

3.1. Gas chromatography-mass spectrometry (GC-MS)

GC-MS is an ideal technique for the identification and quantification of small metabolites (~500 daltons). These molecules include amino acids, fatty acids, hydroxyl acids, alcohols, sugars, sterols and amines, are identified mostly using chemical derivatization in order to make them volatile enough for gas chromatography [40]. Moreover, different methods of derivatization, such as alkylation, acylation, methoximation, trimethylsillation and silylation, can also be used. In GC-MS, two steps are involved in derivatization procedure. The first step requires the conversion of all the carbonyl groups using methoxyamine hydrochloride into corresponding oximes and the second step is followed by trimethylsilylation reaction to increasing the volatility of the derivative metabolites using derivatizing reagent such as N-Methyl-N-(trimethylsilyl) trifluoroacetamide (MSTFA) and N,O-bis- (trimethylsilyl)-trifluoroacetamide (BSTFA) [41-43]. In this procedure, the active hydrogen is replaced from the –NH, –SH, –OH and −COOH of specific metabolites with [−Si(CH3)3] and are converted into thermally stable, less polar and volatile trimethylsilyl (TMS)-ether, TMS-ester, TMS-amine, or TMS-sulphide groups, respectively [42].

There are two major forms of ionization used in GC–MS that comprises of electron ionization (EI) and chemical ionization (CI). Till now, the majority of GC–MS methods in metabolomics utilize EI. GC with EI detector equipped with single quadrupole (Q) mass analyzer is the oldest and most advanced analytical tool with robustness, high sensitivity, resolution and reproducibility, but suffers from sluggish scanning speeds and also poor mass accuracy (~50–200 ppm). Therefore, GC with a time-of-flight mass spectrometry (TOF-MS) analyzer is more preferred for metabolic profiling as it provides higher mass accuracy, faster acquisition times and improved deconvolution for complex mixtures [44]. Among all metabolomics technique, GC-MS is one of the most standardized, efficient, productive technique in plant metabolomics and it is considered as a most versatile platform for metabolites analysis [45]. In addition, GC-MS has availability of the huge number of well-established libraries of both commercial and in-house metabolite databases [46,47,48]. Metabolite profiling is utilized as an essential tool for screening of GM crops with regard to quality and health requirements and in categorization to investigation of potential changes in metabolic contents e.g., *T. aestivum* [49], *O. sativa* [50] and *Z. mays* [51].

3.2. Liquid chromatography-mass spectrometry (LC-MS)

LC-MS is one of the most comprehensive analytical techniques in plant metabolome research, which is used to measure a wide variety of complex metabolites. LC-MS approach is appropriate for plant metabolites with high molecular weight (>500 kDa), heat-labile functional groups, chemically

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 July 2021 doi:10.20944/preprints202107.0193.v1

unstable functional groups and high-vapor-point, as it does not require volatilization of the metabolites. LC-MS is also quite effective techniques in profiling of SMs (e.g., alkaloids, phenolics, flavonoids and terpenes), lipids (e.g., phospholipids, sphingolipids and glycerolipids) and sterols, and steroids [52,53,54].

LC–MS can also be used with many various ionization methods and depending on choice of specific separating columns based on the chemical characteristics of both mobile and stationary phases [53]. Currently, reverse-phase columns such as C18 or C8 are the most widely used columns for LC gradient separation. In reverse-phase separations, organic solvent/aqueous mixed mobile phases are often used, such as water: acetonitrile or water: methanol. Atmospheric pressure ionization (API) and electron spray ionization (ESI) are the most widely used ionization tools for LC–MS [55]. ESI and API have provided limited structural information of the compound because they introduce less internal energy and produce only a few fragments [56]. Structural information is typically obtained by number of fragments using collision induced dissociation (CID) on tandem MSn. Commonly, two tandem MSn analytical tool configurations are commonly available with the LC-MS-based metabolite analysis: tandem-in-time and tandem-in-space. The ion trap MS is used by tandem-in-time instruments, such as quadruple ion traps (QIT-MS), FTICR-MS and orbitrap. The tandem-in-space tool facilitates two sequential steps of mass spectrometric analysis (MS2); it includes two mass analyzers separated by a collision cell [57,58]. Although, LC-MS requires standard reference compounds to identify and quantify SMs, this restricts the analysis of metabolites that are not commercially available [59,60].

3.3. Capillary electrophoresis-mass spectrometry (CE-MS)

CE-MS is a strong analytical technique for evaluating a large variety of ionic metabolites based on the proportion of charge and size ratio [53]. It provides fast and high-resolution of charged compounds from small injection volumes and enables the metabolites characterization based on mass fragmentation [22]. The coverage of CE-MS metabolites majorly overlaps with GC-MS, but requires no derivatization, thus this technique save time and consumables. CE is performed in a fused silica capillary tube, the ends of which are dipped in buffer solutions and across which high voltages (20–30 kV) are employed [43]. Furthermore, CE has low sensitivity and reproducibility, poor migration time and lack of reference libraries; therefore, it is the least appropriate platform for studying metabolites from complex plant samples [61,62]. However, CE has some distinct rewards over other metabolomics tools; primarily the fact that it uses low volume of separation, which is especially appropriate for the study of plant metabolome study [22,63].

3.4. Fourier transform ion cyclotron resonance-mass spectrometry (FTICR-MS)

FTICR-MS provides the highest resolving power and mass accuracy among all kinds of mass spectrometry [64]. Its specific analytical features have made FTICR an important technique for proteomics and metabolomics. FTICR-MS with direct injection high-resolution mass spectrometry (DIHR-MS) is an efficient ion trapping system that has very low detection limits, very high resolution and high mass accuracy [43]. It's also well compatible with multi-stage mass spectrometry (MSn) analyzers. However, the instrument associated with a high magnetic field, complex ion-ion interactions and high cost are major barriers to its widespread application and use in plant metabolomics studies [21].

3.5. Matrix-assisted laser desorption/ionization (MALDI)

Recently, the applications of MALDI-Mass Spectrometry Imaging (MSI) and other MSI tools use a non-target approach for the qualitative or quantitative imaging of a broad variety of metabolites [65]. In plants, many studies have used MALDI-MSI to assess the spatial distribution of lipids, sugars and other classes of metabolite from plant parts such as flowers, leaves and roots [66,67]. In

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 July 2021 doi:10.20944/preprints202107.0193.v1

addition, MALDI-MSI has permitted the simultaneous analysis of the distribution of many peptides and proteins actively from a plant tissue section. This method involves coating a thin film of a matrix comprising either sinapinic acid, α-Cyano-4-hydroxycinnamic acid (CHCA) and 2,5-dihdroxybenzoic acid (2,5-DHBA) on the tissue surface. At each stage, a laser beam is inserted across the matrix-coated tissue to obtain a mass spectrum [42]. For protein/metabolites imaging MALDI is the most used method of ionization, combined with a wide variety of different mass analyzers, namely ToF, ToF-ToF, QqToF (quadrupole time of flight), Fourier ICR transform (FT-ICR) and ion-trap (both linear and spherical). All of these have their own merits and have previously been addressed and reviewed [68]. Other different ionization techniques such as secondary ion mass spectrometry (SIMS), desorption electrospray ionization (DESI) and laser ablation electrospray ionization (LAESI) have been also investigated [69].

4. Bioinformatics tools and databases for plants metabolome analysis

Metabolomics generate a huge amount of metabolic data using wide range of analytical instruments. During the last decades, bioinformatics tools (web-based programs) are designed for metabolomics raw data processing, data mining, data assessment, data interpretation, statistical analysis as well as mathematical modelling of metabolomic networks (Figure 1). In general, acquired data is processed for the correction of baseline shifts, background noise reduction, peak detection and alignment, and finally deconvolution of mass spectra (Figure 1, Table 1). Many bioinformatic tools are designed for the data pre-processing including XCMS (https://xcmsonline.scripps.edu), METLIN (http://metlin.scripps.edu) AMDIS (Automated Mass Spectral Deconvolution and Identification System), MeltDB, MetaboAnalys, MetAlign and AnalyzerPro for different analytical techniques (Table 1). XCMS is an online bioinformatics platform that facilitates direct uploading of raw data and assists the user in data processing and statistical analysis [70]. For LC-MS experiments, XCMS has been developed for programmed data transfer that has reduced data processing time and improved the effectiveness of an online system [71]. METLIN is another online database, which has been used in various studies related to plant metabolic profiling of stress response. It is useful for plant metabolic profiling of specific metabolites and it is not time-consuming for data processing, mining and annotation [72]. In addition, MeltDB (https://meltdb.cebitec.uni-bielefeld.de) is an important web-based platform used for data assessment, processing, and statistical analysis in plant metabolomics [73]. Finally, the MetaboAnalyst online platform also includes a flexible enrichment analysis tool including some topological and visualization possibilities [74]. Various computational web-based, statistical and online bioinformatics tools are commonly used for data analysis in plant metabolomics (Table 1).

Table 2. Available/accessible bioinformatics and statistical tools for metabolite identification.

Database name	Website (URL)	Data Input	Major Function	Reference
ADAP	http://www.du-lab.org/software.htm/	GC/TOF-MS	Data processing	[75]
AMDIS	http://www.amdis.net/	GC-MS	Data processing	[76]
BinBase	http://fiehnlab.ucdavis.edu/db or https://fiehnlab.ucdavis.edu/projects/binbase-setup	GC–MS	Metabolite annotation	[77]
FiehnLib	http://fiehnlab.ucdavis.edu/db or https://fiehnlab.ucdavis.edu/projects/fiehnlib	GC-qTOF-MS	Metabolic profiling	[78]
GMDB	https://jcggdb.jp/rcmg/glycodb	MALDI-TOF	Metabolite	[79]

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 July 2021 doi:10.20944/preprints202107.0193.v1

	/Ms_ResultSearch		annotation	
KEGG	http://www.genome.jp/kegg/	--	Metabolic models	[80]
KNApSAcK	http://kanaya.naist.jp/KNApSAcK/	FT/ICR-MS	Metabolite database	[81]
MarVis	http://marvis.gobics.de/	LC-MS	Metabolite annotation	[82]
MassBase	http://webs2.kazusa.or.jp/massbase/	MS	Metabolite annotation	[83]
MAVEN	https://maven.apache.org/	LC-MS	Data processing	[84]
MeltDB 2.0	https://meltdb.cebitec.uni-bielefeld.de	GC-MS & LC-MS	Data processing	[72]
MetaboAnalyst	www.metaboanalyst.ca/	GC-MS & LC-MS	Statistical analysis	[85]
Metabolome Express	https://www.metabolome-express.org	GC-MS	Data processing, visualization and statistical analysis	[86]
MetaboSearch	http://omics.georgetown.edu/metabosearch.html	MS	Data annotation	[87]
Metabox	https://github.com/kwanjeeraw/metabox	MS	Analysis workflow	[88]
MetAlign	www.metalign.nl	GC-MS & LC-MS	Data processing & Statistical analysis	[89]
metaP-server	http://metabolomics.helmholtz-muenchen.de/metap2/	LC-MS/MS	Data analysis	[90]
MetAssign	http://mzmatch.sourceforge.net/	LC-MS	Data annotation	[91]
MetFrag	https://ipb-halle.github.io/MetFrag/	MS	Metabolite annotation	[92]
MET-IDEA	http://bioinfo.noble.org/gateway/index.php?option=com_wrapper&Itemid=57	GC-MS & LC-MS	Data processing	[93]
MetiTree	http://www.metitree.nl/	MS	Data annotation	[94]
METLIN	https://metlin.scripps.edu/	LC-MS & MS/MS	Metabolite annotation	[95]
MMCD	http://mmcd.nmrfam.wisc.edu/ or https://www.g6g-softwaredirectory.com/bio/metabolomics/dbs-kbs/20670-Univ-Madison-WI-MMCD.php	MS	Metabolite annotation	[96]
Molfind	http://metabolomics.pharm.uconn.edu/Software.html	HPLC/MS	Metabolite annotation	[97]
Mzcloud	https://www.mzcloud.org/	MS/MS & MSn	Metabolite annotation	[98]

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 July 2021 doi:10.20944/preprints202107.0193.v1

MZedDB	http://maltese.dbs.aber.ac.uk:8888/hrmet/index.html	MS	Data annotation	[99]
MZmine2	http://mzmine.github.io/	LC-MS	Data processing	[100]
NIST	http://www.nist.gov/srd/nist1a.cfm or https://www.nist.gov/srd/nist-standard-reference-database-1a	GC-MS, LC–MS & MS/MS	Metabolite annotation	[101]
PRIMe	http://prime.psc.riken.jp/	GC-MS, LC-MS & CE–MS	Metabolite annotation	[102]
XCMS	https://xcmsonline.scripps.edu	GC-MS, LC-MS & MS2	Data processing	[69]

The basic goal of pathway analysis is to combine biochemical information with collected metabolomics data to recognize metabolite patterns that match with metabolic pathways [103]. It is possible to consider metabolic pathways as groups of metabolites that share a common biological process and are related by one or more enzymatic reactions. A broad set of metabolic pathways are covered by comprehensive metabolic pathway databases, such as the KEGG database [104], PlantCyc [105], MetaCyc [106], AraCyc [107] and the small molecule pathway database (SMPDB) [108] (Table 2). A number of software, such as, metabolite set enrichment analysis (MSEA), MPEA, IMPaLA, MBRole, VANTED, MetaboAnalyst, Paintomics, ProMeTra, Metscape2, and MetaMapRR can perform statistical and other metabolite enrichment analyses (Table 2). MSEA methods can be methodically distinguished into over-representation (ORA), single-sample profiling (SSP) and quantitative enrichment (QEA) analysis [109]. Metscape2 [110], which is an add-on to the common Cytoscape software [111] that allows data on metabolites, genes, and pathways to be displayed in the scope of metabolic networks. In addition, platform-independent online resources such as Paintomics [112], ProMeTra [113] and MetaMapRR [114] are also accessible.

Table 3. Database for metabolite enrichment analysis and pathway visualization.

Database	URL	References
AraCyc	https://www.plantcyc.org/typeofpublication/aracyc	[106]
Cytoscape	http://www.cytoscape.org/	[115]
IMPaLA	http://impala.molgen.mpg.de	[116]
iPath	http://pathways.embl.de/	[117]
KEGG	http://www.genome.jp/kegg/	[103]
MapMan	http://mapman.gabipd.org/web/guest/mapman	[118]
MBRole	http://csbg.cnb.csic.es/mbrole/	[119]
Metabolonote	http://metabolonote.kazusa.or.jp/	[120]
MetaCrop	http://metacrop.ipk-gatersleben.de	[121]
MetaCyc	http://www.metacyc.org	[105]
MetPA	http://metpa.metabolomics.ca/MetPA/	[122]
MPEA	http://ekhidna.biocenter.helsinki.fi/poxo/mpea/	[123]
MSEA	http://www.msea.ca. or http://www.metaboanalyst.ca	[108]
Pathcase	http://nashua.case.edu/PathwaysMAW/Web/	[124]
PathwayExplorer	http://genome.tugraz.at/pathwayexplorer/pathwayexplorer_description.shtml	[125]

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 July 2021 doi:10.20944/preprints202107.0193.v1

PMN/PlantCyc	http://www.plantcyc.org	[104]
SMPDB	http://www.smpdb.ca	[107]
VANTED	https://immersive-nalytics.infotech.monash.edu/vanted/	[126]
WikiPathways	http://wikipathways.org	[127]

5. Conclusions

Metabolomics has achieved a prominent role in plant science research. In plant sciences, it has wide applications ranging from investigating the stress-specific metabolites for different climatic stresses, evaluating candidate metabolic gene functions to analyzing the biological mechanism in plant cells and dissecting the genotype-phenotype relationship in response to the different BSs and ASs. In this review, we provide an overview of different sample collection, harvesting method, storage and sample preparation in the plant metabolomics experiments. Furthermore, the most widely used analytical tools in metabolomics for agriculture research viz. GC-MS, LC-MS, CE-MS, FTICR-MS and MALDI with new development in their applications. In addition, we discussed computational software and database employed for metabolomics data processing in plant science. The integration of comprehensive bioinformatics tools with omics strategies professionally dissects novel metabolic networks for crop improvement. Metabolomics has excelled classical approach for novel metabolites discovery and simultaneously explore the complexity and enormous chemical diversity of metabolites in any crop plants. The integration of metabolomics with other "omics" technologies, e.g., genomics, transcriptomics, proteomics is able to deliver novel insights into genetic regulations of crop plants in the context to their cellular function and metabolic network. The complete elucidation of physio-biochemical and molecular mechanisms underlying plant developmental and stress-responsive biology primarily depends on the comprehensive investigations using omics techniques that make metabolomics more applicable in agriculture sciences. Metabolomics has tremendous potential in the field of plant research, as metabolites are more appropriate to the plant phenotype than DNAs, RNAs or proteins [128]. Therefore, studies in this field will effort on both ways, one is the systematic study of the biochemical and genetic mechanisms of metabolic variations in crop plants using both targeted and non-targeted methods; other is metabolomic platform can be used for metabolic profiling of genome-edited plants using CRISPR/Cas9 system for risk evaluation and regulatory affairs related with genetically modified crops [129]. Thus, we can say metabolomics will be able to contribute a lot to agriculture science such as crop breeding and genome editing for crop improvement, better grain yield and elucidate their unknown and novel metabolic pathways.

Author Contributions: Conceptualization, M.K.P, NSY; writing—original draft preparation, M.K.P and S.P; writing—review and editing, M.K.P., S.P., M.K., I.H., S. Pal and N.S.Y; supervision, M.K.P. and N.S.Y.; project administration, M.K.P. and N.S.Y.; funding acquisition, N.S.Y. All authors have read and agreed to the published version of the manuscript.

Funding:

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

Abbreviations

CE-MS; Capillary electrophoresis-mass spectrometry

EI; Electron ionization

FTICR-MS; Fourier transform ion cyclotron resonance-mass spectrometry

GC-MS; Gas chromatography-mass spectrometry,

HPLC; High-performance liquid chromatography

 doi:10.20944/preprints202107.0193.v1

LC-MS; Liquid chromatography mass-spectroscopy

MALDI; matrix-assisted laser desorption/ionization,

MS; Mass spectrometry

MSI; Mass Spectrometry Imaging

TOF; Time-of-flight

SMs; Secondary metabolites

References

1. Razzaq, A.; Sadia, B.; Raza, A.; Khalid Hameed, M.; Saleem, F. Metabolomics: A way forward for crop improvement. *Metabolites* **2019**, *9*, 303.

2. Kumar, R.; Bohra, A.; Pandey, A.K.; Pandey, M.K.; Kumar, A. Metabolomics for plant improvement: Status and prospects. *Front. Plant. Sci.* **2017**, *8*, 1302.

3. Shulaev, V.; Cortes, D.; Miller, G.; Mittler, R. Metabolomics for plant stress response. *Physiol Plant.* **2008**, *132*, 199–208.

4. Patel, M.K.; Kumar, M.; Li, W.; Luo, Y.; Burritt, D.J.; Alkan, N.; Tran, L.-S.P. Enhancing salt tolerance of plants: From metabolic reprogramming to exogenous chemical treatments and molecular approaches. *Cells* **2020**, *9*, 2492.

5. Zagorchev, L.; Seal, C.E.; Kranner, I.; Odjakova, M. A central role for thiols in plant tolerance to abiotic stress. *Int. J. Mol. Sci.* **2013**, *14*, 7405–7432.

6. Peters, K.; Worrich, A.; Weinhold, A.; Alka, O.; Balcke, G.; Birkemeyer, C.; Bruelheide, H.; Calf, O.W.; Dietz, S.; Dührkop, K.; Gaquerel, E.; Heinig, U.; Kücklich, M.; Macel, M.; Müller, C.; Poeschl, Y.; Pohnert, G.; Ristok, C.; Rodríguez, V.M.; Ruttkies, C.; Schuman, M.; Schweiger, R.; Shahaf, N.; Steinbeck, C.; Tortosa, M.; Treutler, H.; Ueberschaar, N.; Velasco, P.; Weiß, B.M.; Widdig, A.; Neumann, S.; Dam, N.M.v. Current Challenges in Plant Eco-Metabolomics. *Int. J. Mol. Sci.* **2018**, *19*, 1385.

7. Hong, J.; Yang, L.; Zhang, D.; Shi, J. Plant Metabolomics: An indispensable system biology tool for plant science. *Int. J. Mol. Sci.* **2016**, *17*, 767.

8. Piasecka, A.; Kachlicki, P.; Stobiecki, M. Analytical Methods for Detection of plant metabolomes changes in response to biotic and abiotic stresses. *Int. J. Mol. Sci.* **2019**, *20*, 379.

9. Patel, M.K.; Mishra, A.; Jha, B. Untargeted metabolomics of halophytes. In Marine Omics: Principles and Applications; Kim, S. Eds.; CRC Press: Boca Raton, FL, USA, **2016**, pp. 309–325.

10. Mishra, A.; Patel, M.K.; Jha, B. Non–targeted metabolomics and scavenging activity of reactive oxygen species reveal the potential of *Salicornia brachiata* as a functional food. *J. Funct. Foods.* **2015**, *13*, 21–31.

11. Pandey, S.; Patel, M.K.; Mishra, A.; Jha, B. Physio-biochemical composition and untargeted metabolomics of cumin (*Cuminum cyminum* L.) make it promising functional food and help in mitigating salinity stress. *PLoS ONE* **2015**, *10*, e0144469.

12. Patel, M.K.; Mishra, A.; Jaiswar, S.; Jha, B. Metabolic profiling and scavenging activities of developing circumscissile fruit of psyllium (*Plantago ovata* Forssk.) reveal variation in primary and secondary metabolites. *BMC Plant Biol.* **2020**, *20*, 116.

13. Patel, M.K.; Mishra, A.; Jha, B. Non-targeted Metabolite Profiling and Scavenging Activity Unveil the Nutraceutical Potential of Psyllium (*Plantago ovata* Forsk). *Front. Plant Sci.* **2016**, *7*, 431.

14. Bénard, C.; Bernillon, S.; Biais, B.; Osorio, S.; Maucourt, M.; Ballias, P.; Deborde, C.; Colombié, S.; Cabasson, C.; Jacob, D. Metabolomic profiling in tomato reveals diel compositional changes in fruit affected by source–sink relationships. *J. Exp. Bot.* **2015**, *66*, 3391–3404.

15. Xie, Z.; Wang, C.; Zhu, S.; Wang, W.; Xu, J.; Zhao, X. Characterizing the metabolites related to rice salt tolerance with introgression lines exhibiting contrasting performances in response to saline conditions. *Plant Growth Regul.* **2020**, *92*, 157–167.

16. Francki, M.G.; Hayton, S.; Gummer, J.; Rawlinson, C.; Trengove, R.D. Metabolomic profiling and genomic analysis of wheat aneuploid lines to identify genes controlling biochemical pathways in mature grain. *Plant Biotechnol. J.* **2016**, *14*, 649–660.

17. Rao, J.; Cheng, F.; Hu, C.; Quan, S.; Lin, H.; Wang, J.; Chen, G.; Zhao, X.; Alexander, D.; Guo, L. Metabolic map of mature maize kernels. *Metabolomics* **2014**, *10*, 775–787.

18. Tohge, T.; de Souza, L.P.; Fernie, A.R. Genome-enabled plant metabolomics. *J. Chromatogr. B* **2014**, *966*, 7–20.

19. Xu, J.; Chen, Z.; Wang, F.; Jia, W.; Xu, Z. Combined transcriptomic and metabolomic analyses uncover rearranged gene expression and metabolite metabolism in tobacco during cold acclimation. *Sci. rep.* **2020**, *10*, 1–13.

20. Hamany Djande, C.Y.; Pretorius, C.; Tugizimana, F.; Piater, L. A.; Dubery, I.A. Metabolomics: A Tool for Cultivar Phenotyping and Investigation of Grain Crops. *Agronomy* **2020**, *10*, 831.

21. Tanna, B.; Mishra, A. Metabolomics of seaweeds: Tools and techniques. In Plant Metabolites and Regulation Under Environmental Stress; Ahmad, P., Ahanger, M.A., Singh, V.P., Tripathi, D.K., Alam, P., Alyemeni, M.N., Eds.; Academic Press: Cambridge, MA, USA, **2018**; pp. 37–52.

22. Salem, M.A.; Perez de Souza, L.; Serag, A.; Fernie, A.R.; Farag, M.A.; Ezzat, S.M.; Alseekh, S. Metabolomics in the Context of Plant Natural Products Research: From Sample Preparation to Metabolite Analysis. *Metabolites* **2020**, *10*, 37.

23. Sumner, L.W.; Amberg, A.; Barrett, D.; Beale, M.H.; Beger, R.; Daykin, C.A.; Fan, T.W.M.; Fiehn, O.; Goodacre, R.; Griffin, J.L.; Hankemeier, T. Proposed minimum reporting standards for chemical analysis. *Metabolomics* **2007,** 3, 211–221.

24. Harbourne, N.; Marete, E.; Jacquier, J.C.; O'Riordan, D. Effect of drying methods on the phenolic constituents of meadowsweet (*Filipendula ulmaria*) and willow (*Salix alba*). *Lwt-Food Sci. Technol.* **2009**, *42*, 1468–1473.

25. Parida, A.K.; Panda, A.; Rangani, J. Metabolomics-guided elucidation of abiotic stress tolerance mechanisms in plants. In Plant Metabolites and Regulation Under Environmental Stress; Ahmad, P., Ahanger, M.A., Singh, V.P., Tripathi, D.K., Alam, P., Alyemeni, M.N., Eds.; Academic Press: Cambridge, MA, USA, **2018**; pp. 89–131.

26. Gong, Z.G.; Hu, J.; Wu, X.; Xu, Y.J. The recent developments in sample preparation for mass spectrometry-based metabolomics. *Crit. Rev. Anal. Chem*. **2017**, *8347*, 1–7.

27. Silva-Navas, J.; Moreno-Risueno, M.A.; Manzano, C. Flavonols mediate root phototropism and growth through regulation of proliferation-to differentiation transition. *Plant Cell* **2016**, *28*, 1372–1387.

28. Corrales, A.R.; Carrillo, L.; Lasierra, P. Multifaceted role of cycling DOF factor 3 (CDF3) in the regulation of flowering time and abiotic stress responses in Arabidopsis. *Plant Cell Environ* **2017**, *40*, 748–764.

29. Sánchez-Parra, B.; Frerigmann, H.; Pérez Alonso,M.-M. Characterization of four bifunctional plant IAM/PAM-amidohydrolases capable of contributing to auxin biosynthesis. *Plants* **2014**, *3*, 324–347.

30. Lehmann, T.; Janowitz, T.; Sánchez-Parra, B. Arabidopsis NITRILASE 1 contributes to the regulation of root growth and development through modulation of auxin biosynthesis in seedlings. *Front. Plant Sci.* **2017**, *8*, 36.

31. t'Kindt, R.; Morreel, K.; Deforce, D. Joint GC–MS and LC–MS platforms for comprehensive plant metabolomics: repeatability and sample pre-treatment. *J. Chromatogr. B* **2009**, *877*, 3572–3580.

32. Giavalisco, P.; Li, Y.; Matthes, A. Elemental formula annotation of polar and lipophilic metabolites using (13) C, (15) N and (34) S isotope labelling, in combination with high-resolution mass spectrometry. *Plant J.* **2011**, *68*, 364–376.

33. Yuliana, N.D.; Khatib, A.; Verpoorte, R.; Choi, Y.H. Comprehensive extraction method integrated with NMR metabolomics: a new bioactivity screening method for plants, adenosine a1 receptor binding compounds in *Orthosiphon stamineus*, Benth. *Anal. Chem.* **2011**, *83*, 6902–6906

34. Gratacós-Cubarsí, M.; Ribas-Agustí, A.; García-Regueiro, J.A.; Castellari, M. Simultaneous evaluation of intact glucosinolates and phenolic compounds by UPLC-DAD-MS/MS in *Brassica oleracea* L. var. botrytis. *Food Chem.* **2010**, *121*, 257–263.

35. Wu, X.; Li, N.; Li, H.; Tang, H. An optimized method for NMR-based plant seed metabolomic analysis with maximized polar metabolite extraction efficiency, signal-to-noise ratio, and chemical shift consistency. *Analyst* **2014**, *139*, 1769–1778.

36. Altemimi, A.; Watson, D.G.; Choudhary, R.; Dasari, M.R.; Lightfoot, D.A. Ultrasound assisted extraction of phenolic compounds from peaches and pumpkins. *PLoS ONE* **2016**, *11*, e0148758.

37. Velickovic, D.; Chu, R.K.; Myers, G.L.; Ahkami, A.H.; Anderton, C.R. An approach for visualizing the spatial metabolome of an entire plant root system inspired by the Swiss-rolling technique. *J. Mass Spectrom.* **2020**, *55*, 4363.

38. Zuorro, A.; Lavecchia, R.; Medici, F.; Piga, L. Enzyme-assisted production of tomato seed oil enriched with lycopene from tomato pomace. *Food Bioprocess Tech.* **2013**, *6*, 3499–3509.

39. Bianchi, F.; Ilag, L.; Termopoli, V.; Mendez, L. Advances in MS-Based Analytical Methods: Innovations and Future Trends. *J. Anal. Methods Chem*. **2018**.

40. Fiehn, O. Metabolomics by gas chromatography–mass spectrometry: combined targeted and untargeted profiling. *Curr. Protoc. Mol. Biol.* **2016**, *114*, 1–32.

41. Kopka, J. Current challenges and developments in GC–MS based metabolite profiling technology. J. Biotechnol. **2006**, *124*, 312–322.

42. Kopka, J.; Schauer, N.; Krueger, S.; Birkemeyer, C.; Usadel, B.; Bergmüller, E.; Dörmann, P.; Weckwerth, W.; Gibon, Y.; Stitt, M.; Willmitzer, L. GMD@ CSB. DB: the Golm metabolome database. *Bioinformatics* **2005**, *21*, 1635–1638.

43. Jorge, TF.; Rodrigues, J.A.; Caldana, C.; Schmidt, R.; van Dongen, J.T.; Thomas-Oates, J.; António, C. Mass spectrometry-based plant metabolomics: Metabolite responses to abiotic stress. *Mass Spectrom. Rev.* **2016**, *35*, 620–649.

44. Kumar, M.; Kuzhiumparambil, U.; Pernice, M.; Jiang, Z.; Ralph, P.J. Metabolomics: an emerging frontier of systems biology in marine macrophytes. *Algal Res.* **2016**, *16*, 76–92.

45. Tsugawa, H.; Tsujimoto, Y.; Arita, M.; Bamba, T.; Fukusaki, E.; Fiehn, O. GC/MS based metabolomics: Development of a data mining system for metabolite identification by using soft independent modeling of class analogy (simca). *BMC Bioinformatics* **2011**, *12*, 131–131.

46. Koek, M.M.; Jellema, R.H.; van der Greef, J.; Tas, A.C.; Hankemeier, T. Quantitative metabolomics based on gas chromatography mass spectrometry: Status and perspectives. *Metabolomics* **2011**, 7, 307–328.

47. Mastrangelo, A.; Ferrarini, A.; Rey-Stolle, F.; García, A.; Barbas, C. From sample treatment to biomarker discovery: A tutorial for untargeted metabolomics based on GC-(EI)-Q-MS. *Anal. Chim. Acta* **2015**, *900*, 21–35.

48. Beale, D.J.; Pinu, F.R.; Kouremenos, K.A.; Poojary, M.M.; Narayana, V.K.; Boughton, B.A.; Dias, D.A. Review of recent developments in GC–MS approaches to metabolomics-based research. *Metabolomics* **2018**, *14*, 1–31.

49. Stamova, B.S.; Roessner, U.; Suren, S.; Laudencia-Chingcuanco, D.; Bacic, A.; Beckles, D.M. Metabolic profiling of transgenic wheat over-expressing the high-molecular-weight Dx5 glutenin subunit. *Metabolomics* **2009**, *5*, 239–252.

50. Zhou, J.; Ma, C.F.; Xu, H.L.; Yuan, K.L.; Lu, X.; Zhu, Z.; Wu, Y.N.; Xu, G.W. 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. *J. Chromatogr. B* **2009**, *877*, 725–732.

51. Sissener, N.H.; Hemre, G.-I.; Lall, S.P.; Sagstad, A.; Petersen, K.; Williams, J.; Rohloff, J.; Sanden, M. Are apparent negative effects of feeding genetically modified MON810 maize to Atlantic salmon, Salmo salar caused by confounding factors? *Br. J. Nutr.* **2011**, *106*, 42–56.

52. Matsuda, F.; Hirai, M.Y.; Sasaki, E.; Akiyama, K.; Yonekura-Sakakibara, K.; Provart, N.J.; Sakurai, T.; Shimada, Y.; Saito, K. AtMetExpress development: a phytochemical atlas of Arabidopsis development. *Plant Physiol.* **2010**, *152*, 566–578.

53. Okazaki, Y.; Kamide, Y.; Hirai, M.Y.; Saito, K.; 2011. Plant lipidomics based on hydrophilic interaction chromatography coupled to ion trap time-of-flight mass spectrometry. *Metabolomics* **2011**, *9*, 121–131.

54. Obata, T.; Fernie, A.R. The use of metabolomics to dissect plant responses to abiotic stresses. *Cell Mol. Life Sci.* **2012**, *69*, 3225–3243.

55. Okazaki, Y.; and Saito, K. Recent advances of metabolomics in plant biotechnology. *Plant biotechnol. Rep.* **2012**, *6*, 1–15.

56. Holčapek, M.; Jirásko, R.; Lísa, M. Recent developments in liquid chromatography–mass spectrometry and related techniques. *J. Chromatogr.* **2012**, *1259*, 3–15.

57. Allwood, J.W.; Goodacre, R. An introduction to liquid chromatography–mass spectrometry instrumentation applied in plant metabolomic analyses. *Phytochem. Anal.* **2010**, *21*, 33–47.

58. Schiffmann, C.; Hansen, R.; Baumann, S.; Kublik, A.; Nielsen, P.H.; Adrian, L.; Von Bergen, M.; Jehmlich, N.; Seifert, J. Comparison of targeted peptide quantification assays for reductive dehalogenases by selective reaction monitoring (SRM) and precursor reaction monitoring (PRM). *Anal. Bioanal. Chem.* **2014**, *406*, 283–291.

59. Last, R.L.; Jones, A.D.; Shachar-Hill, Y. Towards the plant metabolome and beyond. *Nat .Rev. Mol. Cell Biol.* **2007**, *8*, 167–174.

60. Alseekh, S.; Fernie, A.R. Metabolomics 20 years on: What have we learned and what hurdles remain?. *Plant J.* **2018**, *94*, 933–942.

61. Soga, T.; Imaizumi, M. Capillary electrophoresis method for the analysis of inorganic anions, organic acids, amino acids, nucleotides, carbohydrates and other anionic compounds. *Electrophoresis* **2001**, *22*, 3418–3425.

62. Williams, B.J.; Cameron, C.J.; Workman, R.; Broeckling, C.D.; Sumner, L.W.; Smith, J.T. Amino acid profiling in plant cell cultures: An inter-laboratory comparison of CE-MS and GC-MS. *Electrophoresis* **2007**, *28*, 1371–1379.

63. Ren, J.L.; Zhang, A.H.; Kong, L.; Wang, X.J. Advances in mass spectrometry-based metabolomics for investigation of metabolites. *RSC Adv.* **2018**, *8*, 22335–22350.

64. Nikolaev, E.N.; Kostyukevich, Y.I.; Vladimirov, G.N. (2016). Fourier transform ion cyclotron resonance (FT ICR) mass spectrometry: Theory and simulations. *Mass Spectrom. Rev.* **2016**, *35*, 219–258.

65. Baker, T.C.; Han, J.; Borchers, C.H. Recent advancements in matrix-assisted laser desorption/ionization mass spectrometry imaging. *Curr. Opin. Biotechnol.* **2017**, *43*, 62–69.

66. Cha, S.; Zhang, H.; Ilarslan, HI.; Wurtele, ES.; Brachova, L.; Nikolau, B.J.; Yeung, E.S. Direct profiling and imaging of plant metabolites in intact tissues by using colloidal graphite-assisted laser desorption ionization mass spectrometry. *Plant J.* **2008**, *55*, 348–360.

67. Jun, J.H.; Song, Z.; Liu, Z.; Nikolau, B.J.; Yeung, E.S.; Lee, Y.J. High-spatial and high-mass resolution imaging of surface metabolites of *Arabidopsis thaliana* by laser desorption-ionization mass spectrometry using colloidal silver. *Anal. Chem.* **2010**, *82*, 3255–3265.

68. Goodwin, R.J.; Pennington, S.R.; Pitt, A.R. Protein and peptides in pictures: imaging with MALDI mass spectrometry. *Proteomics* **2008**, *8*, 3785–3800.

69. Lee, Y.J.; Perdian, D.C.; Song, Z.; Yeung, E.S.; Nikolau, B.J. Use of mass spectrometry for imaging metabolites in plants. *Plant J.* 2012, *70*, 81–95.

70. Tautenhahn, R.; Patti, G.J.; Rinehart, D.; Siuzdak, G. XCMS Online: A web-based platform to process untargeted metabolomic data. *Anal. Chem.* **2012**, *84*, 5035–5039.

71. Montenegro-Burke, J.R.; Aisporna, A.E.; Benton, H.P.; Rinehart, D.; Fang, M.; Huan, T.;Warth, B.; Forsberg, E.; Abe, B.T.; Ivanisevic, J. Data streaming for metabolomics: Accelerating data processing and analysis from days to minutes. *Anal. Chem.* **2017**, *89*, 1254–1259.

72. Siuzdak, G. METLIN: A metabolite mass spectral database. *Ther. Drug Monit.* **2005**, *27*, 747–751.

73. Kessler, N.; Neuweger, H.; Bonte, A.; Langenkämper, G.; Niehaus, K.; Nattkemper, T.W.; Goesmann, A. MeltDB 2.0–advances of the metabolomics software system. *Bioinformatics* **2013**, *29*, 2452–2459.

74. Xia, J.; Wishart, D.S. Using MetaboAnalyst 3.0 for comprehensive metabolomics data analysis. *Curr. Protoc. Bioinformatics* **2016**, *55*, 14–10.

75. Jiang, W.; Qiu, Y.; Ni, Y.; Su, M.; Jia, W.; Du, X. An Automated Data Analysis Pipeline for GC− TOF− MS Metabonomics Studies. *J. Proteome Res.* **2010**, *9*, 5974–5981.

76. Behrends, V.; Tredwell, G.D.; Bundy, J.G. A software complement to AMDIS for processing GC-MS metabolomic data. *Anal. Biochem.* **2011**, *415*, 206–208.

77. Skogerson, K.; Wohlgemuth, G.; Barupal, D.K.; Fiehn, O. The volatile compound BinBase mass spectral database. *BMC bioinformatics* **2011**, *12*, 1–15.

78. Kind, T.; Wohlgemuth, G.; Lee, D.Y.; Lu, Y.; Palazoglu, M.; Shahbaz, S.; Fiehn, O. FiehnLib: mass spectral and retention index libraries for metabolomics based on quadrupole and time-of-flight gas chromatography/mass spectrometry. *Anal. Chem.* **2009**, *81*, 10038–10048.

79. Kameyama, A.; Kikuchi, N.; Nakaya, S.; Ito, H.; Sato, T.; Shikanai, T.; Takahashi, Y.; Takahashi, K.; Narimatsu, H. A strategy for identification of oligosaccharide structures using observational multistage mass spectral library. *Anal. Chem*. **2005**, *77*, 4719–4725.

80. Aoki, K.F.; Kanehisa, M. Using the KEGG database resource. *Curr. Protoc. Bioinform.* **2005**, *11*, 1–12.

81. Shinbo, Y.; Nakamura, Y.; Altaf-Ul-Amin, M.; Asahi, H.; Kurokawa, K.; Arita, M.; Saito, K.; Ohta, D.; Shibata, D; Kanaya, S. KNApSAcK: a comprehensive species-metabolite relationship database. In Plant metabolomics; Saito, K., Richard, A.D., Willmitzer**,** L., Eds.; Springer, Berlin, Heidelberg **2006**; pp. 165-181.

82. Kaever, A.; Landesfeind, M.; Feussner, K.; Mosblech, A.; Heilmann, I.; Morgenstern, B.; Feussner, I.; Meinicke, P. MarVis-Pathway: Integrative and exploratory pathway analysis of non-targeted metabolomics data. *Metabolomics* **2015**, *11*, 764–777.

83. Carroll, A.J. Online Metabolomics Databases and Pipelines. *In Tech Rijeka* **2012**, *10*, 47–73.

84. Clasquin, M.F.; Melamud, E.; Rabinowitz, J.D. LC-MS data processing with MAVEN: A metabolomic analysis and visualization engine. *Curr. Protoc. Bioinform*. **2012**, *37*, 14–11.

85. Xia, J.; Psychogios, N.; Young, N.; Wishart, D.S. MetaboAnalyst: A web server for metabolomic data analysis and interpretation. *Nucleic Acids Res*. **2009**, *37*, 652–660.

86. Carroll, A.J.; Badger, M.R.; Millar, A.H. The Metabolome Express Project: enabling web-based processing, analysis and transparent dissemination of GC/MS metabolomics datasets. *BMC Bioinformatics* **2010**, 11, 1–13.

87. Zhou, B.; Wang, J.; Ressom, H.W. MetaboSearch: Tool for mass-based metabolite identification using multiple databases. *PLoS ONE* **2012**, 7, e40096.

88. Wanichthanarak, K.; Fan, S.; Grapov, D.; Barupal, D.K.; Fiehn, O. Metabox: A toolbox for metabolomic data analysis, interpretation and integrative exploration. *PLoS ONE* **2017**, *12*, e0171046.

89. Lommen, A.; Kools, H.J. MetAlign 3.0: Performance enhancement by efficient use of advances in computer hardware. *Metabolomics* **2012**, *8*, 719–726.

90. Kastenmüller, G.; Römisch-Margl, W.; Wägele, B.; Altmaier, E.; Suhre, K. metaP-server: A web-based metabolomics data analysis tool. *BioMed Res. Int.* **2010**, *2011*, 1–7.

91. Daly, R.; Rogers, S.; Wandy, J.; Jankevics, A.; Burgess, K.E.; Breitling, R. MetAssign: Probabilistic annotation of metabolites from LC–MS data using a Bayesian clustering approach. *Bioinformatics* **2014**, *30*, 2764–2771.

92. Ruttkies, C.; Schymanski, E.L.; Wolf, S.; Hollender, J.; Neumann, S. MetFrag relaunched: Incorporating strategies beyond in silico fragmentation. *J. Cheminform*. **2016**, *8*, 3.

93. Lei, Z.; Li, H.; Chang, J.; Zhao, P.X.; Sumner, L.W. MET-IDEA version 2.06; improved efficiency and additional functions for mass spectrometry- based metabolomics data processing. *Metabolomics* **2012**, *8*, 105–110.

94. Rojas-Chertó, M.; van Vliet, M.; Peironcely, J.E.; Van Doorn, R.; Kooyman, M.; Te Beek, T.; Van Driel, M.A.; Hankemeier, T.; Reijmers, T. MetiTree: A web application to organize and process high-resolution multi-stage mass spectrometry metabolomics data. *Bioinformatics* **2012**, *28*, 2707–2709.

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

96. Cui, Q.; Lewis, I.A.; Hegeman, A.D.; Anderson, M.E.; Li, J.; Schulte, C.F.; Westler, W.M.; Eghbalnia, H.R.; Sussman, M.R.; Markley, J.L. Metabolite identification via the madison metabolomics consortium database. *Nat. Biotechnol.* **2008**, *26*, 162

97. Menikarachchi, L.C.; Cawley, S.; Hill, D.W.; Hall, L.M.; Hall, L.; Lai, S.; Wilder, J.; Grant, D.F. MolFind: a software package enabling HPLC/MS-based identification of unknown chemical structures. *Anal. Chem.* **2012**, *84*, 9388–9394.

98. Mistrik, R., Lutisan, J., Huang, Y., Suchy, M., Wang, J., Raab, M. mzCloud: a key conceptual shift to understand 'Who's Who' in untargeted metabolomics. In Metabolomics Society Conference, Glasgow, **2013**, pp. 1–4.

99. Draper, J.; Enot, D.P.; Parker, D.; Beckmann, M.; Snowdon, S.; Lin, W.; Zubair, H. Metabolite signal identification in accurate mass metabolomics data with MZedDB, an interactive m/z annotation tool utilising predicted ionisation behaviour 'rules'. *BMC Bioinformatics* **2009**, *10*, 1–16.

100. Pluskal, T.; Castillo, S.; Villar-Briones, A.; Orešič, M. MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics* **2010**, *11*, 1–11.

101. Naumkin, A.V.; Kraut-Vass, A.; Gaarenstroom, S.W.; Powell, C.J. NIST Standard Reference Database 20. **2012**.

102. Sakurai, T.; Yamada, Y.; Sawada, Y.; Matsuda, F.; Akiyama, K.; Shinozaki, K.; Hirai, M.Y.; Saito, K. PRIMe update: innovative content for plant metabolomics and integration of gene expression and metabolite accumulation. *Plant Cell Physiol.* **2013**, *54*, e5.

103. Pérez-Alonso, M.M.; Carrasco-Loba, V.; Pollmann, S. (2018). Advances in Plant Metabolomics. *Annu. Rev. Plant. Biol.* **2018**, 557–588.

104. Kanehisa, M.; Furumichi, M.; Tanabe, M. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* **2017**, *45*, 353–361.

105. Chae, L.; Kim, T.; Nilo-Poyanco, R. Genomic signatures of specialized metabolism in plants. *Science* **2014**, *344*, 510–513.

106. Caspi, R.; Foerster, H.; Fulcher, C.A. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Res.* **2008**, *36*, 623–631.

107. Mueller, L.A.; Zhang, P.; and Rhee, S.Y. AraCyc: a biochemical pathway database for Arabidopsis. *Plant Physiol.* **2003**, *132*, 453–460.

108. Jewison, T.; Su, Y.; Disfany, F.M. SMPDB 2.0: big improvements to the small molecule pathway database. *Nucleic Acids Res.* **2014**, *42*, 478–484.

109. Xia, J.; Wishart, D.S. MSEA: a web-based tool to identify biologically meaningful patterns in quantitative metabolomic data. *Nucleic Acids Res.* **2010**, *38*, 71–77.

110. Karnovsky, A.; Weymouth, T.; Hull, T. Metscape 2 bioinformatics tool for the analysis and visualization of metabolomics and gene expression data. *Bioinformatics* **2012**, *28*, 373–380.

111. Kohl, M.; Wiese, S.; and Warscheid, B. Cytoscape: software for visualization and analysis of biological networks. *Methods Mol. Biol.* **2011**, *696*, 291–303.

112. Garcia-Alcalde, F.; Garcia-Lopez, F.; Dopazo, J. Paintomics: a web based tool for the joint visualization of transcriptomics and metabolomics data. *Bioinformatics* **2011**, *27*, 137–139.

113. Neuweger, H.; Persicke, M.; Albaum, S.P. Visualizing post genomics data-sets on customized pathway maps by ProMeTra-aeration-dependent gene expression and metabolism of *Corynebacterium glutamicum* as an example. *BMC Syst. Biol.* **2009**, *3*, 82.

114. Grapov, D.; Wanichthanarak, K.; Fiehn, O. MetaMapR: pathway independent metabolomic network analysis incorporating unknowns. *Bioinformatics* **2015**, *31*, 2757–2760.

115. Smoot, M.E.; Ono, K.; Ruscheinski, J.; Wang, P.L.; Ideker, T. Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* **2011**, *27*, 431–432.

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

117. Letunic, I.; Yamada, T.; Kanehisa, M.; Bork, P. iPath: interactive exploration of biochemical pathways and networks. *Trends Biochem. Sci.* **2008**, *33*, 101–103.

118. Thimm, O.; Bläsing, O.; Gibon, Y.; Nagel, A.; Meyer, S.; Krüger, P.; Selbig, J.; Müller, L.A.; Rhee, S.Y.; Stitt, M. MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. *Plant J.* **2004**, *37*,914–939.

119. Chagoyen, M.; Pazos, F. MBRole: enrichment analysis of metabolomic data. *Bioinformatics* **2011**, *27*, 730–731.

120. Ara, T.; Enomoto, M.; Arita, M.; Ikeda, C.; Kera, K.; Yamada, M.; Nishioka, T.; Ikeda, T.; Nihei, Y.; Shibata, D.; Kanaya, S. Metabolonote: a wiki-based database for managing hierarchical metadata of metabolome analyses. *Front. Bioeng. Biotechnol.* **2015**, *3*, 38.

121. Schreiber, F.; Colmsee, C.; Czauderna, T.; Grafahrend-Belau, E.; Hartmann, A.; Junker, A.; Junker, B.H.; Klapperstück, M.; Scholz, U.; Weise, S. MetaCrop 2.0: managing and exploring information about crop plant metabolism. *Nucleic Acids Res.* **2012**, *40*, 1173-1177.

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

123. Kankainen, M.; Gopalacharyulu, P.; Holm, L.; Orešič, M. MPEA—metabolite pathway enrichment analysis. *Bioinformatics* **2011**, *27*, 1878–1879.

124. Elliott, B., Kirac, M., Cakmak, A., Yavas, G., Mayes, S., Cheng, E., Wang, Y., Gupta, C., Ozsoyoglu, G.; Meral Ozsoyoglu, Z. PathCase: pathways database system. *Bioinformatics* **2008**, *24*, 2526–2533.

125. Mlecnik, B.; Scheideler, M.; Hackl, H.; Hartler, J.; Sanchez-Cabo, F.; Trajanoski, Z. PathwayExplorer: web service for visualizing high-throughput expression data on biological pathways. *Nucleic Acids Res.* **2005**, *33*, 633–637.

126. Junker, B.H.; Klukas, C.; Schreiber, F. VANTED: a system for advanced data analysis and visualization in the context of biological networks. *BMC Bioinformatics* **2006**, **7**, 1–13.

127. Kelder, T.; van Iersel, M.P.; Hanspers, K.WikiPathways: building research communities on biological pathways. *Nucleic Acids Res.* **2012**, *40*, 1301–1307.

128. Niederbacher, B.; Winkler, J.; Schnitzler, J. Volatile organic compounds as non-invasive markers for plant phenotyping. *J. Exp. Bot.* **2015**, *66*, 5403–5416

129. Razzaq, A.; Saleem, F.; Kanwal, M.; Mustafa, G.; Yousaf, S.; Imran Arshad, H.M.; Hameed, M.K.; Khan, M.S.; Joyia, F.A. Modern trends in plant genome editing: An inclusive review of the CRISPR/Cas9 toolbox. *Int. J. Mol. Sci.* **2019**, *20*, 4045.