Chapter 20
Metabolomics and Metabolic Profiling: Investigation of Dynamic Plant-Environment Interactions at the Functional Level

Dominik Skoneczny, Paul A. Weston, and Leslie A. Weston

1 Introduction

Sessile plants routinely face challenges associated with environmental extremes or neighbouring competitors, and have therefore developed mechanisms that allow them to withstand constant exposure to these diverse abiotic and biotic stressors. In some cases, the response to plant stress can be manifested on demand by the plant (so-called inducible responses), while other responses are expressed constitutively and are available at all times to counter the stressor. Thus, it can be said that environment shapes a plant's physiology and, in turn, also impacts the functioning of ecosystems (Wittstock and Gershenzon 2002; Arbona et al. 2013; Weston et al. 2015). Interactions between plants, their competitors, and the environment are always dynamic and as a result often difficult to characterize. It is therefore not surprising that recent studies of such complex interactions have utilised a multitude of advanced techniques for experimentation, and have eventually led to an enhanced understanding of the physiological basis for these interactions.

Plant response to environmental stress or stimuli frequently occurs rapidly, sometimes within seconds to minutes, by triggering biochemical pathways that can be measured in this brief time frame. Such rapid and specific changes *in planta* can be challenging to detect and quantify (Ye et al. 2013; Schuman and Baldwin 2016). Other environmental stimuli trigger broader response patterns causing secondary changes *in planta* resulting in, for example, differential resource allocation. Such stress responses can be monitored at the gene level by studying the plant's transcriptome as assessed by gene expression. Following transcription, translation of gene

D. Skoneczny · P. A. Weston · L. A. Weston (✉)
Graham Centre for Agricultural Innovation, Charles Sturt University, Wagga Wagga, NSW, Australia

School of Agricultural and Wine Sciences, Charles Sturt University, Wagga Wagga, NSW, Australia
e-mail: lewestonn@csu.edu.au

A. M. Sánchez-Moreiras, M. J. Reigosa (eds.), *Advances in Plant Ecophysiology Techniques*, https://doi.org/10.1007/978-3-319-93233-0_20

products into functional or non-functional proteins occurs, producing what is referred to collectively as the plant's proteome. However, regulation of most environmental interactions in the plant system occurs at the metabolite level. Metabolites clearly play a major role in elicitation of plant response to the environment and can act as mediators, signal molecules, defence metabolites or regulators of certain cellular functions. This complete and functional set of metabolites within a living plant is referred to as the plant metabolome (Rochfort 2005; Weston et al. 2015).

Plants are generally successful in combating or adapting to moderate levels of environmental stress, herbivory and pathogen infestation (Akula and Ravishankar 2011; Ryalls et al. 2016). Stimuli known to trigger well-coordinated defence responses include, but are not limited to, temperature, photoperiod, drought, as well as herbivores, pathogens or neighbouring plant species (Lee 2002; Franceschi et al. 2005). Following herbivory, some plants exhibit increased production of defence metabolites that deter the specific herbivore. These host-specific responses require recognition of the herbivore at the chemical level and the ability to rapidly produce defence metabolites (Schuman and Baldwin 2016). Such complex biochemical interactions can now be studied over time both qualitatively and quantitatively using an analytical approach known as metabolomics.

When performed in a non-targeted manner, metabolomics refers to the systematic study of all metabolites in a living organism; the plant's metabolome consists of both primary and secondary metabolites. In the plant, metabolomics can be performed at various organismal levels including study of a single cell, a tissue, an organ or the entire organism. In addition, it can be performed on matrices associated with plants, including soil from the rhizosphere containing water-soluble metabolites, or the headspace surrounding a plant containing plant-produced volatiles (Dunn and Ellis 2005; Rochfort 2005; Weston et al. 2015). Metabolomics is often performed to monitor metabolite flux over time and aims to quantify those metabolites associated with a particular biochemical response at a specific point in time (Kim and Verpoorte 2010).

Although metabolomics is performed using multiple techniques, the steps in a metabolic profiling analysis should be performed with the highest level of care and precision to successfully assess variation in metabolite concentrations and presence over time, and also to detect those less stable metabolites often present in trace quantities. Since the biochemistry of living plants is dynamic and occurs in association with other living organisms including the plant's living microbiome, sample preparation for metabolomics must first involve the termination or quenching of all biochemical processes in the system under study, followed later by extraction, separation and detection of all key metabolites present. Once detection has been performed successfully, the complex data set is then analysed using selected software packages and chemometrics approaches.

Analytical techniques commonly used in metabolomics studies generally involve liquid or volatile samples prepared for separation using liquid or gas chromatography (LC or GC, respectively) coupled to various types of mass spectrometry (MS) (Roessner and Bacic 2009; Weston et al. 2015). In addition, nuclear magnetic resonance (NMR) spectroscopy instrumentation can also be used for detection

(Smolinska et al. 2012). Recently, MALDI TOF/MS (matrix assisted laser desorption and ionisation coupled to time of flight mass spectrometry) has been used directly on solid samples for intensive protein and fat metabolomics analyses (Fuchs et al. 2010).

This chapter explores the most common applications of metabolomics in plant ecophysiology research by providing an overview of typical instrumentation and workflows used by plant scientists along with a discussion of the experimental outcomes of such studies. It is important to remember that this field is rapidly advancing and new technological improvements to equipment, techniques and data processing occur on a yearly basis.

Metabolomics requires the knowledge of separation science and analytical techniques, thus scientists who do not specialise in these areas can find suitable collaborators in well-equipped regional or national research hubs. A deeper understanding of organic chemistry, biochemistry, chromatography, bioinformatics and statistics will allow for successful experimentation and data analysis.

2 Metabolomics Tools

A plant's metabolome typically consists of upwards of thousands of metabolites that vary with species and cultivar as well as with phenology in both composition and abundance. Due to enormous chemical diversity present in a typical plant extract, no single analytical method can detect all metabolites present. As a result, metabolomics is generally performed using a variety of platforms. In contrast, analysis of a subset of metabolites (metabolic profiling) can be accomplished with a single platform, which results in reduced number and diversity of metabolites detected in a sample, allowing for a more simplified analysis (Hill and Roessner 2014; Weston et al. 2015).

Integrated platforms, with a diversity of separation techniques and instruments, have frequently been used to provide the most comprehensive analysis of metabolomes. Chromatography [LC, GC, or CE (capillary electrophoresis)] coupled with MS or NMR analysis of the entire sample (Table 20.1) are the most common high-throughput techniques employed to study a suite of low molecular weight (<1500 Da) constituents. LC/MS is frequently used in the analysis of compounds ranging in polarity, whereas GC/MS is most commonly employed for the analysis of volatile and derivatized metabolites.

LC or GC/CE/MS usually require extensive sample preparation including filtration in contrast to NMR/MS, which can be carried out with minimal sample preparation. NMR is typically only used for the analysis of mixtures of metabolites in high abundance due to its limited sensitivity (Zhang et al. 2012). NMR can also be used in some cases for detection of compounds that are labile or less stable, in contrast to typical GC or LC techniques (Smolinska et al. 2012).

Sample preparation for LC or GC/CE/MS is highly specific. Samples must conform to meet the requirements of each instrument employed for metabolomics

Table 20.1 Comparison and characteristics of mass spectrometry (MS) and nuclear magnetic resonance (NMR) spectroscopy for metabolomics applications (Pan and Raftery 2007; El-Aneed et al. 2009; Yan et al. 2017)

Platform	Characteristics	Considerations
Mass spectrometry (MS)	Sensitive (LOD 0.5 nM)	Sample preparation and separation techniques are based on polarity of constituents
	Sample preparation required (purification)	
	Challenging quantification	Samples need to be derivatized and/or prepared in a solvent
	Limited structural information	
	Time of analysis depends on separation technique (15–40 min/sample)	High throughput but slower than NMR
	Destructive to samples	Untargeted analysis of large numbers of metabolites at low concentrations
	Detection of >500 metabolites per run	
	Small sample volume 1–100 μL	
	Less expensive instrumentation	Targeted analysis with prior knowledge of compounds of interest or analytical standard
	Matrix-Assisted Laser Desorption Ionization (MALDI) MS allows for analysis of sample co-crystalized with a solid matrix	
		MALDI can be used in analysis of solids and large biomolecules (up to 500,000 Da, including proteins and DNA). In metabolomics typically used to study lipids and glycoconjugates
Nuclear magnetic resonance spectroscopy (NMR)	Less sensitive than MS (LOD 0.5 μM)	Broad-based analyses
	Limited sample pre-preparation	Samples in various solvent matrices
	Quantification easy and precise	
	Limited or no sample preparation needed	High throughput technique for metabolites in high concentration
	Rich structural information	
	Rapid analysis (2–3 min/sample)	Useful for non-targeted analysis of metabolites present in high abundance
	Non-destructive analysis of precious samples	
	Detection of 40–200 metabolites per run	
	High reproducibility	
	Large sample volume required (0.1–0.5 mL)	
	Higher cost of equipment	

analyses (e.g. high acid content can damage HPLC systems; unfiltered samples can block HPLC columns) and should be suitable for the separation technique selected (e.g. GC samples may require derivatization if analysis is performed on less volatile mixtures of constituents).

Table 20.2 Comparison of separation techniques coupled to mass spectrometers

	Gas chromatography	Liquid chromatography	Capillary electrophoresis
Sample preparation	Extensive sample preparation	Sample purification required	Minimum sample preparation
	Derivatization of non-volatile compounds	Sample volume 1–15 µL	Sample volume 1–20 nL
	Sample volume ~1 µL	Sample in solvent mixture	Sample in aqueous phase
	Sample in solvent mixture/inert gas		
Type of analytes	Volatiles and compounds that volatilise after derivatization	Liquid samples (solubilized solids) compatible with mobile phase	Ionic and very polar metabolites
			Small sample volume
			Faster than GC/LC
	Thermally stabile compounds	Polar and non-polar compounds	Lower resolution than normal-phase LC but generally results in detection of different metabolites
	Non-polar low MW compounds	Mainly secondary metabolites	
	Mainly primary but some secondary metabolites associated with fragrance and flavour		
Application	Availability of multiple libraries, rapid identification, high reproducibility	Lower reproducibility, lack of large reproducible spectral databases	Lower reproducibility, lack of large reproducible spectral databases
	Carbohydrates, amino acids, organic acids, sugars, oils, terpenoids	Phenolics, alkaloids, glucosinolates, terpenoids, etc.	

In general, all samples should be relatively free from salts, residues, proteins and other contaminants that might cause interference during the analysis. Once prepared, certain liquid or gaseous samples can be injected directly into the MS interface or analysed using NMR; however, most complex sample matrices will need to be separated by a chromatographic method prior to analysis. The separation technique selected (Table 20.2) depends on various physical and chemical properties of analytes under study. Solids can be directly profiled using certain LC/MS interfaces and also by matrix-assisted laser desorption ionization (MALDI) MS. During MALDI/MS, samples are dissolved in a solution or frozen and applied onto a solid matrix and dried to form crystals that are bombarded with a laser beam to allow ion formation using matrix as a mediator (Fuchs et al. 2010). Additionally, desorption electrospray ionization (DESI) mass spectrometry allows for profiling of metabolites directly from solid surfaces. This ionization technique allows for profiling of metabolites in situ and creates a metabolic profile as an image of the solid under study (Claude et al. 2017).

2.1 Mass Spectrometry

Mass spectrometry is currently the most widely applied microanalytical technique in metabolomics as it provides high sensitivity and resolution while allowing for analysis of a wide range of metabolites (Table 20.1). Mass spectrometry also allows for quantitative and qualitative analyses as it has the capacity to generate structural information. It is generally highly reproducible for quantitative analysis and allows for a simultaneous determination of the relative abundance of hundreds of metabolites in a single run. Ideally it is linked to a chromatographic separation prior to analysis but direct-injection MS can also be used for some applications as well as for profiling of solids (Dunn and Ellis 2005; Zhang et al. 2012; Weston et al. 2015).

For all MS analyses, analytes must be ionized and present in the gaseous state in order to be detected. When performing MS, movement of ions within the ionisation chamber is regulated by modification of electromagnetic fields. Analytes are ionized in the MS interface using a variety of ionization sources including electrospray ionisation (ESI) or atmospheric-pressure chemical ionization (APCI) before introducing them into the *m/z* analyser (Fig. 20.1). To ensure free movement and lack of contamination, the MS is always operated under high vacuum. The MS platform generally consists of one (single MS) or two mass analysers (tandem MS or MS-MS), with the latter able to provide additional structural information and therefore more precise quantification.

In an MS-MS experiment, additional ion pre-selection and collision-induced dissociation (CID) take place for clear determination of the fragmentation pattern of a selected ion. The resulting mass spectra display abundance of fragments covering a range of mass-to-charge ratios *(m/z),* which can be used to infer the molecular structure of metabolites. Neutral molecules and molecules that have not ionized are generally not detected in the MS (Watson and Sparkman 2007).

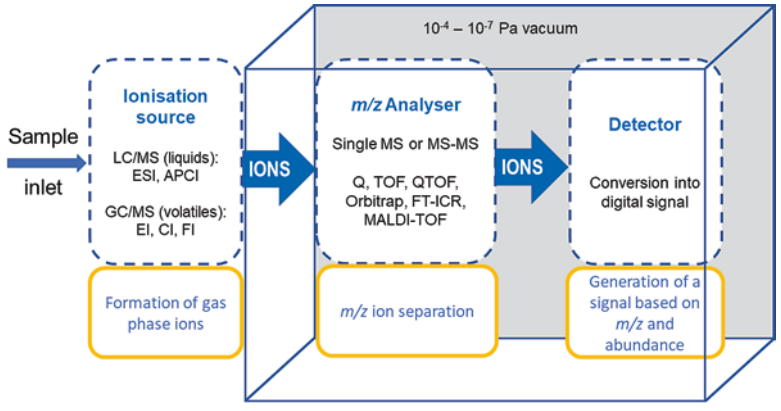

Fig. 20.1 General and simplified representation of a mass spectrometer platform typically used for metabolomics and metabolic profiling analysis. (Schematic representation based on Watson and Sparkman 2007)

2.2 Nuclear Magnetic Resonance

NMR spectroscopy is becoming of increasing importance in the field of metabolomics as sensitivity of NMR instrumentation increases. This analytical method is well suited for untargeted metabolomic analysis because it is non-selective and can often provide highly reproducible results and structural information, depending on concentration of analytes in the sample matrix. In contrast to LC/MS, hundreds of low molecular weight metabolites can be detected simultaneously in a single run with very limited sample preparation. NMR can also be used in vivo with samples without any sample preparation, such as a newly poured sample of wine studied for composition. NMR has been employed for metabolite fingerprinting as well as metabolic profiling, and since it is non-destructive, sample recovery is possible following analysis. Workflows for NMR are typically automated (Smolinska et al. 2012; Zhang et al. 2012), permitting the method to be used for larger sample sets depending on the platform; in some cases, up to 500 samples can be processed per day. The major drawback of NMR is its overall low sensitivity, limiting its utility to samples where profiling compounds of higher abundance is desirable (Pan and Raftery 2007).

During NMR experimentation, magnetic fields are applied to samples and nuclei of atoms with an odd atomic number such as ^{1}H or ^{13}C gain what is referred to as nuclear spin. At the same time, applied radio frequencies allow nuclei to reach high-spin energies, and the radiation generated during the relaxation of the magnetic field is then detected and compared to that of reference atoms. Commonly, ^{1}H or proton NMR is used in metabolomics studies as the majority of metabolites contain ^{1}H atoms or protons (Dunn and Ellis 2005).

2.3 Application of NMR and MS for Metabolomics Approaches for the Study of Plant Response to Stress

In a study using *Arabidopsis* as a model system, Jänkänpää et al. (2012) studied the impact of light intensity on dynamic changes in the *A. thaliana* metabolome using GC/QTOF MS. Results indicated that more than 70 compounds fluctuated in concentration with varying light intensity when plants were grown under controlled conditions. Over 30 metabolites, including amino acids, lipids and carbohydrates, were later identified to be responsive to light intensity. When plants were moved from controlled environment growth chambers (artificial light) to natural light conditions in the field, changes in the leaf metabolome were observed within a 3-day period, and pronounced shifts in metabolism were noted within 4 h following transition to the field.

A metabolomics approach was also used to evaluate the nutritional value of a fresh vegetable product (Maldini et al. 2015) under varying levels of light. This study used LC ion mobility-QTOF to investigate changes in the metabolome of

broccoli (*Brassica oleracea*) sprouts under alternating light/dark conditions. In this study, increased light intensity was associated with increased chlorophyll biosynthesis as well as elevated levels of phytosterols, lipids, fatty acids and carotenoids.

Metabolomics can also prove useful for assessment of pharmaceutics and plant-produced medicinals. For example, alkaloid biosynthesis in the commercial poppy plant was assessed using of FT-ICR MS (Fourier-transform ion cyclotron resonance MS) in cell cultures. Metabolic differences between elicitor-treated and control cell cultures were observed, and targeted analysis of benzylisoquinoline alkaloid biosynthesis in opium poppy (*Papaver somniferum*) was performed with a focus on morphine, codeine and sanguinarine production. Approximately 1000 metabolites were annotated in the study, including compounds previously uncharacterized. ^{1}H NMR was also performed for comparative purposes and resulted in annotation of a greater number of metabolites in contrast to LC/MS QToF methods employed. Following exposure to an elicitor, changes in metabolism of cell suspension cultures were noted within hours following treatment (Hagel and Facchini 2008).

The impact of drought and climate change on the shoot and root metabolism in two invasive perennial grasses, *Holcus lanatus* (Yorkshire fog) and *Alopecurus pratensis* (meadow foxtail), were evaluated using LC/MS QToF with the LTQ Orbitrap XL and ^{1}H NMR. Over 850 metabolites were detected; however only 55 metabolites were positively identified. Shoot extracts were more chemically complex than root extracts, a typical observation in plant metabolomics. Simultaneous exposure to high temperature and drought resulted in differential responses in comparison to exposure to individual stressors, suggesting that the plant response to stress is complex and is mediated by multiple biochemical pathways (Gargallo-Garriga et al. 2015).

3 Metabolomics Pipeline: From Harvest to Data Analysis

As suggested previously, metabolomics workflows (the steps involved in processing and analysing samples and their resulting data files as part of a metabolomics analysis) are dependent on sample type, analytical instrumentation and desired outputs. However, metabolomics experiments should be designed to minimize sources of external variation while achieving sample-to-sample uniformity. When considering metabolomics and metabolic profiling analyses, four main steps—sample preparation, separation, detection and chemometrics analysis—are generally undertaken (Fig. 20.2). These approaches are explored in greater detail in Kim and Verpoorte (2010) and Hill and Roessner (2014).

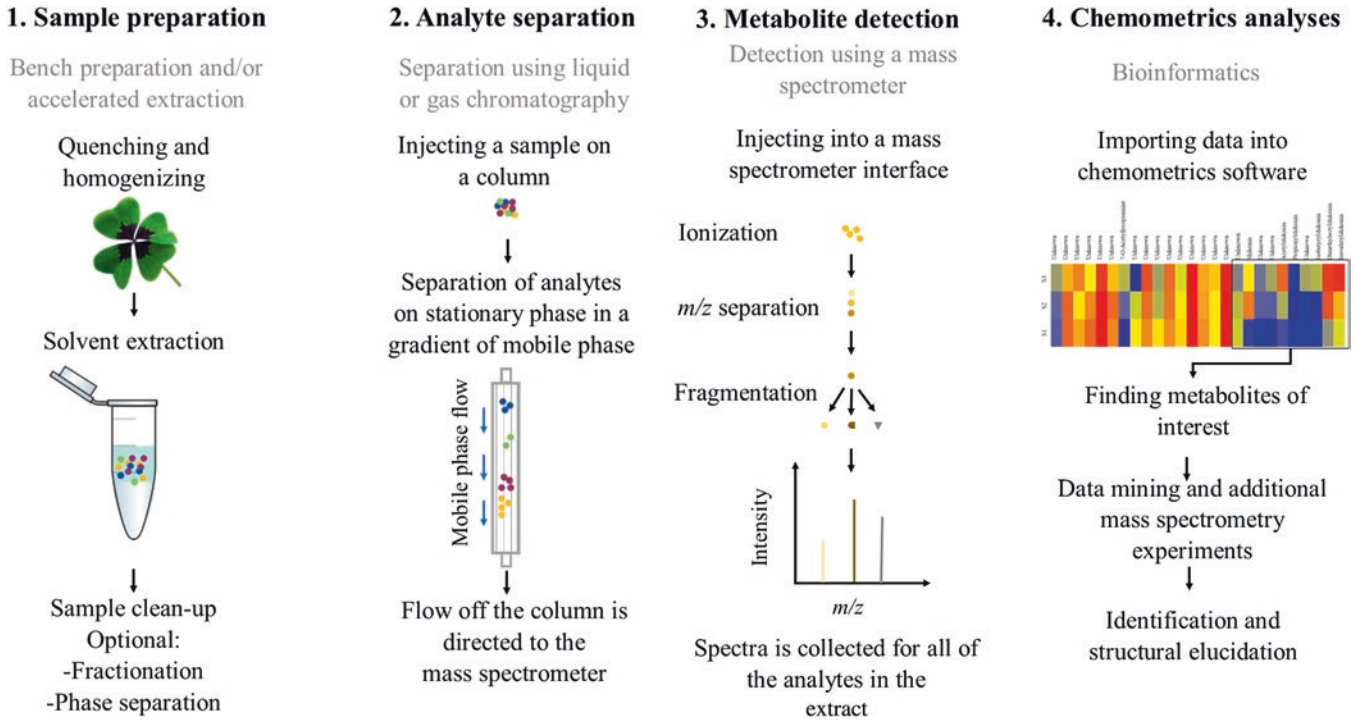

Fig. 20.2 Metabolomics pipeline used for identification and profiling of differentially regulated compounds. Although every step in this process can be modified, the general workflow remains the same

3.1 Sample Preparation

3.1.1 Harvest

It should be noted that plant-based experiments often require sampling at multiple time points as metabolite levels fluctuate due to circadian rhythms in higher plants, as well as with changes in plant phenology. The impacts of biotic and abiotic stress exposure on plants can therefore add complexity to a standard time-course experiment. Plant harvest should be performed similarly for all samples and processing of plant tissues should occur rapidly following harvest. Most protocols involve placing plants on ice, in an −80 °C freezer or freeze-drying tissues directly using liquid nitrogen to prevent biochemical changes post-harvest (Hill and Roessner 2014; Kim and Verpoorte 2010; Weston et al. 2015).

3.1.2 Quenching and Homogenizing

To conserve the sample and quench active biochemical processes in living tissues, plant tissue may be snap-frozen in liquid nitrogen and then stored at −80 °C until extraction. Several studies indicate the importance of drying using freeze-drying and homogenizing prior to extraction which can be performed using mortar and pestle or cryogenic mill (Hill and Roessner 2014). Protocols must be optimized for tissue type and/or compounds of interest in metabolic profiling to ensure stability of thermolabile and/or light-sensitive metabolites. In some cases, only fresh tissue should be processed immediately after harvest as freeze-drying can result in degradation of certain natural products such as porphyrins. For more stable metabolites, tissues can be processed to dryness at room temperature followed by grinding before extraction.

3.1.3 Extraction

Extraction is a critical step in sample preparation and varies with the choice of solvent, temperature and duration. Solvent extraction is most commonly employed but microwave-assisted extraction and supercritical fluid extraction (Kim and Verpoorte 2010) can also be used. The process of solvent extraction can be automated to increase uniformity and speed for high throughput (e.g. using BUCHI Speed Extractor) (Skoneczny et al. 2015; Weston et al. 2015).

3.1.4 Final Sample Preparation

Sample clean-up is the final step prior to analysis, and ensures compatibility of the sample matrix with selected instrumentation. For most NMR instruments, filtration and appropriate solvent choice will ensure sample homogeneity. Gas chromatography for non-volatile compounds such as amino acids, organic acids, fatty acids and

sugars including sugar derivatives typically requires derivatization prior to chromatography to ensure volatility for separation. For liquid samples separated using liquid chromatography, samples are often prepared prior to separation by performing protein precipitation (Kim and Verpoorte 2010).

3.2 Separation of Analytes

Although NMR, MALDI/MS and direct infusion MS do not require separation preceding the analysis, most metabolic profiling experimentation involves a separation step for accurate identification and profiling. Analytes can be further separated using LC, GC or CE as summarized in Table 20.2 and previously described by numerous authors (Hill and Roessner 2014; Ramautar and de Jong 2014; Weston et al. 2015; García et al. 2017; Yan et al. 2017)

3.3 Metabolite Detection

Analytes are detected according to properties of the instrumentation used; thus, accuracy and resolution are instrument dependent. In general, QToF mass spectrometers used in metabolomics studies provide accurate mass (MS) and fragmentation pattern (MS-MS) for each molecule under study, which can later be compared to information available for known compounds in chemical libraries such as NIST or METLIN. In contrast to results obtained with LC and CE/MS, GC/MS fragmentation patterning is more robust and consistent for various instruments and therefore identification of constituents is generally more reliable.

3.4 Non-targeted Metabolomics

Non-targeted approaches aim to evaluate the presence and abundance of as many metabolites as possible in an often complex biological matrix or system. Unfortunately, many plant metabolites are uncharacterized which complicates identification of a large number of plant metabolites. In general, most holistic studies performed with plant extracts or tissues are not able to identify the majority of metabolites contained within samples and instead focus on metabolites that are significantly up- or down-regulated between treatment groups.

Various mass spectrometers and their application for use in non-targeted metabolomics and metabolic profiling studies have been previously described (Dunn and Ellis 2005; Watson and Sparkman 2007; El-Aneed et al. 2009; Viant and Sommer 2013). The most commonly used instrumentation for non-targeted studies include single and tandem MS along with quadrupole time-of flight (QTOF), time-of-flight

(TOF), Fourier-transformed ion cyclotron resonance (FT-ICR) and linear ion trap or orbitrap mass spectrometers. These platforms generally offer high mass accuracy and resolution required for metabolite identification and, more importantly, fast data acquisition and multiple fragmentation options required for complex metabolomics experiments.

Analytical standards are typically employed for confirmation of identification of metabolites. Existing databases can also be employed for annotation or tentative identification of metabolites in cases where standards are not available. Annotation should be performed according to minimum reporting standards such as those presented by the Metabolomics Standards Initiative (Sumner et al. 2007).

3.5 Data Analysis

Data is analysed following acquisition in situ using software packages that are provided by instrument vendors or are available on-line. Some useful software packages are available free of charge. Most of the data analysis workflows include data pre-processing, data pre-treatment and univariate and multivariate statistical analysis (Martínez-Arranz et al. 2015). Depending on the software, data obtained from samples separated via LC is typically deconvoluted; major peaks detected, integrated and aligned; and within-batch and between-batch normalization and baseline correction performed. Several software packages also allow complex metadata analyses across multiple datasets (Hill and Roessner 2014).

Large datasets from metabolomics studies generally require multivariate data analysis techniques, specifically the simultaneous analysis of more than one parameter or variable. Unsupervised classification methods such as principal component analysis (PCA), hierarchal cluster analysis (HCA) and supervised methods based on partial least square (PLS) regressions common when analysing metabolomics data. Such methods allow for identification of metabolites that are significantly up- or down-regulated between treatments, and may include comparisons among various cultivars, treatments or stressors. Selected metabolites are then compared to metabolites documented in databases and libraries, some of which are publicly available e.g. KEGG, METLIN, or NIST, or in-house libraries. The METLIN library as of 2018 currently contains over 100,000 metabolite entries, and most have a complete mass spectrum for referencing [31.10.2017] (Hill and Roessner 2014).

Final structural confirmation of investigated metabolites requires more detailed MS/MS experimentation on existing datasets and, ideally, comparison to analytical standards.

3.6 Bioassay-Driven Data Analysis

Results generated from chemometric analysis of metadatasets are often difficult to interpret. For those studying biological activity or toxicity, it is critical to employ bioassay-driven identification of key constituents associated with activity. In this

case, one may use a statistical method to establish the relationship between abundance of molecular features and bioactivity, which is powerful and offers greater insights into biological activity of metabolic features. Such evaluations are now being utilised in advanced metabolomics and pathways analyses, but are currently not frequently undertaken due to the complexity of these analyses. However, in our opinion, these bioassay-driven associations have proven invaluable for discovery of families of compounds and biomarkers associated with biological activity.

3.7 Determination of the Metabolites

Identification of specific metabolites associated with bioactivity can often be accomplished using a number of multivariate statistical techniques. Partial least squares (PLS) regression is one technique suitable for the task of associating a large number of independent variables (in this case, metabolites) with bioactivity. The output from PLS regression, however, may be difficult to interpret because the relationship between all of the molecular features and bioactivity is revealed. Those entities that have a high degree of association with activity and are present in reasonable abundance are of greatest interest; however, the percentage of variance in activity that is associated with each of these entities is also of interest.

Stepwise linear regression can also be employed to deduce metabolite association with activity. Rather than trying to fit the abundance of all metabolites in the dataset with bioactivity, stepwise linear regression measures the degree of association between each metabolite and bioactivity, and incrementally adds metabolites with the highest degree of association to a predictive model. The output of this technique is a series of models that progressively account for much of the variation between metabolite abundance and bioactivity. In addition, the output indicates the degree to which variation is explained by a particular set of metabolites (r^2) and the statistical significance of the regression. Ideally, only a limited number of metabolites for a given study will result in a regression with a suitably high value of r^2.

4 Case Studies and Emergining Approaches in Plant Metabolomics

Rapid development of new strategies and platforms in metabolomics has recently triggered more insightful and novel research into plant metabolism. Below we present three case studies featuring various platforms to study plant metabolites produced in plants under stress, or those plants experiencing interactions with insects or other plants during plant/insect and plant/plant interactions. In addition, we have summarised a number of recent studies that exemplify a range of approaches to the study of plant metabolomics in Table 20.3.

Table 20.3 Recent research projects exemplifying current approaches to the study of plant metabolomics

Aim of study	Instrumentation	Citation
Comparison of suppressive and allelopathic metabolites in several wheat cultivars for selection of weed suppressive cultivars. Identification of 14 benzoxazinoids in plant tissues and after exudation into the rhizosphere	Targeted analysis using Applied Biosystems Q Trap LC/MS (Nærum, Denmark) and LC-QToF	Mwendwa et al. (2016)
Metabolomics of root exudates that foster the dialogue between belowground herbivores, nematodes, and microbial communities and other plants. Identification of up to 103 metabolites in root exudates. Understanding their role and degradation processes	Untargeted analysis of metabolites in root exudates using a variety of platforms including NMR and GC/LC-ToF	van Dam and Bouwmeester (2016) and Zhu et al. (2016)
Mass spectrometry imaging in plant tissue. Precise profiling of metabolites *in situ* with limited sample preparation. Mapping of the distribution of metabolites in cells or even organelles	Matrix assisted laser desorption ionisation (MALDI), desorption electrospray ionisation (DESI) and secondary ion mass spectrometry	Heyman and Dubery (2016)
Estimation and quantification of microbial metabolites in soils. Understanding their cycling in water-extractable organic matter fraction and decomposition	Untargeted analysis using Agilent GC/MS system (California, USA)	Swenson et al. (2015)
Metabolic profiling as a tool allowing for characterization of plant genotypes. This phenotyping approach can be applied to better understand genetically modified plant systems. Authors compared sucrose metabolism across four cultivars of potatoes tuber	Targeted GC/MS (Thermo-Quest, Manchester, UK)	Roessner et al. (2001)

4.1 Case study 1: Untargeted Metabolomics to Study Plant Response to Infestation by Weaver Ants Using GC/MS

4.1.1 Background

An interesting mutualistic interaction between plants and ants has been explored using a non-targeted metabolomics approach. It was previously noted that plants can provide nutrition or shelter for ants while ants aggressively defend the plant (Metlen et al. 2009).

More detailed insight into this mutualistic interaction was obtained from a study of weaver ants, *Oecophylla smaragdina,* and their association with cultivated coffee, *Coffea arabica*. Vidkjær et al. (2015) measured the nutritive effects of ant faeces on coffee plants; plants that hosted a colony of weaver ants were compared to

control plants under glasshouse-controlled conditions. Previous reports had indicated that weaver ants could improve plant health and crop yield. Considering that some plants are able to take up nitrogen from urea through the leaves, the authors hypothesised that metabolic changes in *Coffea arabica* would occur due to nutrient transfer following infestation by ants.

4.1.2 Methodology

Leaves from control and treated plants were collected every 2 weeks and snap-frozen in liquid nitrogen (quenching) and stored at −80 °C until extraction (stable storage).

Samples were then extracted by shaking crushed leaves in a mixture of solvents. Dried extracts were derivatized within 24 h prior to GC analysis. Samples were analysed in duplicate in randomized order using gas chromatography (Agilent 7890A, Santa Clara, CA) coupled to Leco Pegasus HT 4010 time-of-flight mass spectrometer (St. Joseph, MI). In addition, total carbon and nitrogen were analysed.

4.1.3 Data Analysis

GC/MS data was deconvoluted and annotated using BinBase metabolomics database. Univariate and multivariate data analyses were conducted as well as principal component analysis (PCA) and partial least square discriminant analysis (PLS-DA) with rigorous validation of the models (Vidkjær et al. 2015).

4.1.4 Results and Discussion

This research provided insight into metabolic changes in *C. arabica*. Over 500 molecular features were identified, of which 96 were annotated across all samples. Plants that hosted ant colonies were observed to have increased levels of several fatty acids. In addition, an increase in total nitrogen levels was observed in colonized plants, and several nitrogen-containing metabolites originating from the phenylpropanoid biosynthetic pathway were also up-regulated, suggesting exposure to potential biotic stress. Results suggested that increased nitrogen levels were a result of assimilation of additional nitrogen through leaves due to the presence of ants and ant faeces. The up-regulation of the phenylpropanoid biosynthetic pathway in response to exposure to biotic stress may have led to increased production of secondary metabolites providing additional defence against potential predators (Vidkjær et al. 2015).

4.2 *Case study 2: Metabolic Profiling: Identification and Profiling of Naphthoquinones Under Environmental Stress in the Roots of* Echium plantagineum *Using LC/MS MS*

4.2.1 Background

Echium plantagineum is an invasive weed species in Australia that produces bioactive allelochemicals known as shikonins (naphthoquinones) in the periderm of living roots. Production of naphthoquinones was greater in field-grown plants collected from hotter and drier climates in Australia (Weston et al. 2013). To further understand the role of naphthoquinones in plant defence and their production in response to environmental factors, a platform for identification and profiling of shikonins was developed (Skoneczny et al. 2017).

4.2.2 Methodology

Echium plantagineum plants were exposed to drought, simulated herbivory and different temperature regimes in controlled condition experimentation. Roots were harvested and peeled periderm was extracted in ethanol. Composite sampling was performed initially to study the chemical diversity present in the root periderm. Rhizosphere soil was also sampled for shikonins using polydimethylsiloxane (PDMS) tubing.

Extracts were analyzed using UPLC (Agilent Infinity 1200) coupled with a QTOF mass detector (Agilent 6530). Negative ionization mode was selected for detection of shikonins in periderm extracts and MS-MS experimentation was performed to observe fragmentation patterns for individual and related shikonins present in the extracts (Fig. 20.3). The optimized method with tentatively identified nine compounds was then applied to samples collected from controlled condition experiments.

Data was processed using Mass Hunter Software and Mass Profiler Professional chemometrics software (Agilent Technologies, Santa Clara, CA, USA), (Skoneczny et al. 2017).

4.2.3 Results and Discussion

Nine red-pigmented naphthoquinones, also known as shikonins, were tentatively identified based on comparison of their MS-MS spectra to spectra of known standards and others previously reported (Fig. 20.3). Additionally, polymeric shikonins were detected but due to their structural complexity, complete elucidation of all structures was not performed (Skoneczny et al. 2017). Three shikonins were

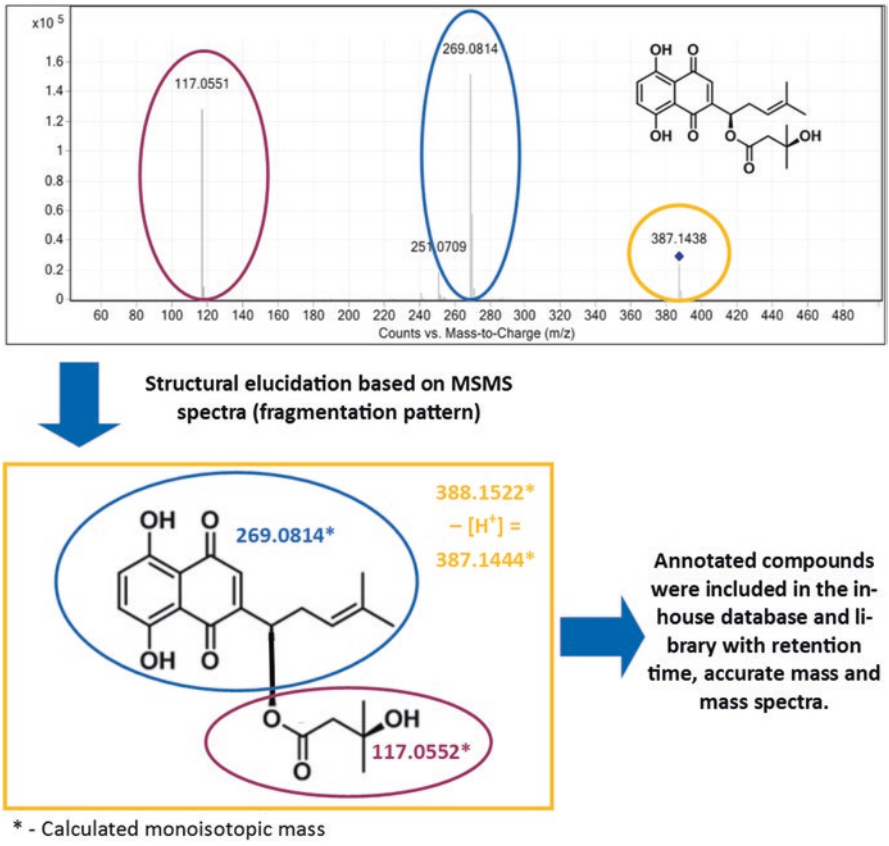

Fig. 20.3 Identification of metabolites by the analysis of the MS-MS spectra. In the case of shikonins, several analytical standards were obtained and permitted a comparison of generated spectra. Identified compounds mainly differed in side chain structure (purple) but all produced a characteristic 269.0814 ion, also present in analytical reference standards

identified from the soil extract suggesting that shikonins were exuded from living roots into the rhizosphere (Zhu et al. 2016).

Shikonins were also profiled in stress-treated plants. Up-regulation of the shikonin biosynthetic pathway was observed within 6 h after plant exposure to simulated herbivory and within days following exposure to increased temperature and water withholding. Temperature clearly impacted production of total shikonins under controlled environment conditions; production of the individual metabolite shikonin, a precursor to other shikonin metabolites, was significantly elevated following exposure to high temperature growth regimes (Fig. 20.4).

High abundance of shikonins was found to be correlated with red pigmentation of periderm and its extracts. Metabolites potentially associated with the shikonin biosynthetic pathway contributed to the clustering of differently coloured root

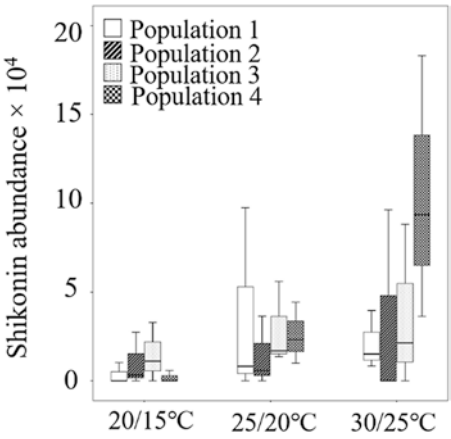

Fig. 20.4 Relative abundance of the individual metabolite, shikonin, in four *Echium plantagineum* populations grown in controlled conditions in different temperature regimes, measured 3 weeks after the beginning of the treatment

extracts in the principal component analysis. This study showed potential changes in secondary metabolism of differently coloured roots (Fig. 20.5).

In general, our studies showed that naphthoquinone production, as monitored through the production of specific shikonins, is rapidly elicited in the roots of healthy plants following exposure to a variety of biotic and abiotic stressors including neighbouring plants, pathogens, herbivory, water withholding and high temperatures.

4.3 Case study 3: Identification of Thrips: Resistant Senecio *Hybrids Using NMR*

4.3.1 Background

Thrips (*Frankliniella occidentalis*) are a global insect pest of food crops; feeding by thrips leads to reduced growth and yields. However, some plants are resistant to thrips due to production of toxic metabolites in their foliage. A study by Leiss et al. (2009) evaluated *Senecio* species F2 hybrids (susceptible and resistant) for their metabolite composition using an untargeted metabolomics approach.

4.3.2 Methodology

Over 30 different F2 hybrids of *S. jacobaea* and *S. aquaticus* were grown in tissue culture and then evaluated for their resistance to thrips by profiling both young and older leaves of individual plants. Thrips-resistant and -susceptible plants were further evaluated using a metabolomics approach performed with a 600 MHz Bruker DMX-600 NMR spectrometer (Bruker, Karlsruhe, Germany).

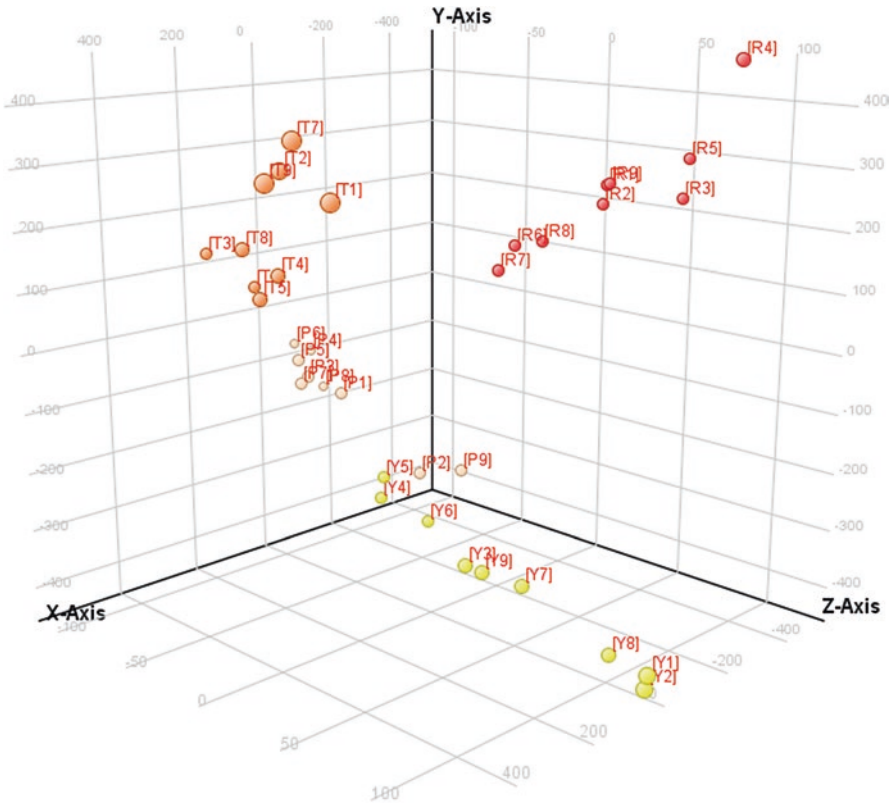

Fig. 20.5 Principal component analysis (PCA) of 36 root periderm extracts of *Echium plantagineum* with varying colouration from pale to red due to the variable presence of a group of naphthoquinones collectively referred to as the shikonins. A total of 162 entities (possible metabolites) were used in the analysis and contributed to the separation of coloured extracts as shown in the PCA plot above. This analysis clearly revealed that differential colouration of the samples is associated with variation in the shikonin biosynthetic pathway

4.3.3 Results and Discussion

NMR proved useful for the determination of secondary metabolite composition in *Senecio* and allowed the identification of variation in metabolic profiles between thrips-resistant and -susceptible hybrids. Resistant plants accumulated significantly greater quantities of toxic pyrrolizidine alkaloids (jaconine and jacobine *N*-oxide) and kaempferol glucoside. Interestingly, higher concentrations of defence metabolites were present in young leaves, resulting in less thrips damage. These findings are consistent with literature noting greater resistance in young leaves and the presence of constitutively expressed pyrrolizidine alkaloids in both *Senecio* species (Leiss et al. 2009).

5 Metabolomics in Systems Biology and Functional Genomics

Understanding the relationship between a living plant's functional phenotype (in this case the metabolites produced) and genotype is currently a hot topic in plant science. Focused studies are now providing strong evidence for the role of specific genes and their impacts on biological function and regulation of plant response to biotic and abiotic stresses (Bino et al. 2004; Saito and Matsuda 2010). In addition, a greater understanding of plant stress response and cellular physiology can be achieved by integrating transcriptomics, proteomics and metabolomics using various computation approaches. However, a key challenge in performing such studies is to create a unified network of genes, transcripts, proteins and metabolites through the integration and evaluation of sets of "-omics" data (Bunnik and Le Roch 2013).

Genes are now frequently annotated by correlating available transcripts with expressed metabolites (Schauer and Fernie 2006). Plants exposed to various stressors can be studied with respect to gene expression associated with metabolic response. Using this approach, gene function can be further predicted and associated metabolic responses can be elucidated (Saito and Matsuda 2010). Despite the complexity of such systems-based experimentation, global databases have been created which can be used to identify proteome, transcriptome and metabolome responses for specific plants under various environmental conditions (Hagel and Facchini 2008; Saito and Matsuda 2010). Such data sets can be especially useful for evaluation of biosynthetic pathways and metabolite flux.

6 Conclusions

A plant's metabolome is a complex and dynamic compilation of both identified and unknown metabolites. Metabolomics is a process that enables identification and quantification of key plant metabolites using a variety of methods and analytical instrumentation. Recent advances in analytical instrumentation and bioinformatics and the integration of multiple "omics" platforms now allows for advanced characterization of biosynthetic pathways and the study of their regulation. Thus, metabolic profiling and metabolomics approaches can be particularly useful for investigation of plant responses to stress.

Current challenges for those performing plant metabolomics include the paucity of metabolite databases containing identified plant metabolites, streamlining the processing of massive datasets generated in metabolomic studies, and bridging the gaps between laboratories employing various platforms for metabolomics on a global scale. Although metabolomics is of increasing importance in the plant sciences, costs associated with intensive metabolite analyses are often higher than those for other "omics" platforms. However, as we have shown in this review,

metabolomics currently offers unparalleled opportunities to investigate the interactions between plants and their environment at a functional level.

Additional Information

The authors have attempted to provide an overview of common techniques, instrumentation and methods used in metabolomics studies in plant ecophysiology experimentation. However, as both metabolomics and chemometric analysis can be performed using a variety of platforms and approaches, not all were able to be described in detail in this review.

References

Akula R, Ravishankar GA (2011) Influence of abiotic stress signals on secondary metabolites in plants. Plant Signal Behav 6:1720–1731

Arbona V, Manzi M, Ollas C, Gómez-Cadenas A (2013) Metabolomics as a tool to investigate abiotic stress tolerance in plants. Int J Mol Sci 14:4885

Bino RJ, Hall RD, Fiehn O, Kopka J, Saito K, Draper J, Nikolau BJ, Mendes P, Roessner-Tunali U, Beale MH, Trethewey RN, Lange BM, Wurtele ES, Sumner LW (2004) Potential of metabolomics as a functional genomics tool. Trends Plant Sci 9:418–425

Bunnik EM, Le Roch KG (2013) An introduction to functional genomics and systems biology. Adv Wound Care 2:490–498

Claude E, Jones EA, Pringle SD (2017) DESI Mass Spectrometry Imaging (MSI). In: Cole LM (ed) Imaging mass spectrometry: methods and protocols. Springer New York, New York, pp 65–75

van Dam NM, Bouwmeester HJ (2016) Metabolomics in the rhizosphere: tapping into belowground chemical communication. Trends Plant Sci 21:256–265

Dunn WB, Ellis DI (2005) Metabolomics: current analytical platforms and methodologies. Trends Anal Chem 24:285–294

El-Aneed A, Cohen A, Banoub J (2009) Mass spectrometry. Review of the basics: electrospray, MALDI, and commonly used mass analyzers. Appl Spectrosc Rev 44:210–230

Franceschi VR, Krokene P, Christiansen E, Krekling T (2005) Anatomical and chemical defenses of conifer bark against bark beetles and other pests. New Phytol 167:353–376

Fuchs B, Süß R, Schiller J (2010) An update of MALDI-TOF mass spectrometry in lipid research. Prog Lipid Res 49:450–475

García A, Godzien J, López-Gonzálvez A, Barbas C (2017) Capillary electrophoresis mass spectrometry as a tool for untargeted metabolomics. Bioanalysis 9:99–130

Gargallo-Garriga A, Sardans J, Pérez-Trujillo M, Oravec M, Urban O, Jentsch A, Kreyling J, Beierkuhnlein C, Parella T, Peñuelas J (2015) Warming differentially influences the effects of drought on stoichiometry and metabolomics in shoots and roots. New Phytol 207:591–603

Hagel J, Facchini P (2008) Plant metabolomics: analytical platforms and integration with functional genomics. Phytochem Rev 7:479–497

Heyman HM, Dubery IA (2016) The potential of mass spectrometry imaging in plant metabolomics: a review. Phytochem Rev 15:297–316

Hill CB, Roessner U (2014) Advances in high-throughput untargeted LC-MS analysis for plant metabolomics. Future Science eBook Series “Advanced LC-MS applicafions for metabolomics”. Future Science Group, London

Jänkänpää HJ, Mishra Y, Schröder WP, Jansson S (2012) Metabolic profiling reveals metabolic shifts in Arabidopsis plants grown under different light conditions. Plant Cell Environ 35:1824–1836

Kim HK, Verpoorte R (2010) Sample preparation for plant metabolomics. Phytochem Anal 21:4–13

Lee CE (2002) Evolutionary genetics of invasive species. Trends Ecol Evol 17:386–391

Leiss K, Choi Y, Abdel-Farid I, Verpoorte R, Klinkhamer PL (2009) NMR metabolomics of thrips (*Frankliniella occidentalis*) resistance in *Senecio* hybrids. J Chem Ecol 35:219–229

Maldini M, Natella F, Baima S, Morelli G, Scaccini C, Langridge J, Astarita G (2015) Untargeted metabolomics reveals predominant alterations in lipid metabolism following light exposure in broccoli sprouts. Int J Mol Sci 16:13678–13691

Martínez-Arranz I, Mayo R, Pérez-Cormenzana M, Mincholé I, Salazar L, Alonso C, Mato JM (2015) Data in support of enhancing metabolomics research through data mining. Data Brief 3:155–164

Metlen KL, Aschehoug ET, Callaway RM (2009) Plant behavioural ecology: dynamic plasticity in secondary metabolites. Plant Cell Environ 32:641–653

Mwendwa JM, Weston PA, Fomsgaard I, Laursen BB, Brown WB, Wu H, Rebetzke G, Quinn JC, Weston LA (2016) Metabolic profiling for benzoxazinoids in weed-suppressive and early vigour wheat genotypes. In: Conference proceedings of the 20th Australasian weeds conference, pp. 353–357. Weeds Society of Western Australia

Pan Z, Raftery D (2007) Comparing and combining NMR spectroscopy and mass spectrometry in metabolomics. Anal Bioanal Chem 387:525–527

Ramautar R, de Jong GJ (2014) Recent developments in liquid-phase separation techniques for metabolomics. Bioanalysis 6:1011–1026

Rochfort S (2005) Metabolomics reviewed: a new "omics" platform technology for systems biology and implications for natural products research. J Nat Prod 68:1813–1820

Roessner U, Bacic A (2009) Metabolomics in plant research. Aust Biochem 40:9–11

Roessner U, Luedemann A, Brust D, Fiehn O, Linke T, Willmitzer L, Fernie AR (2001) Metabolic profiling allows comprehensive phenotyping of genetically or environmentally modified plant systems. Plant Cell Online 13:11–29

Ryalls JMW, Moore BD, Riegler M, Johnson SN (2016) Above–belowground herbivore interactions in mixed plant communities are influenced by altered precipitation patterns. Front Plant Sci 7:345

Saito K, Matsuda F (2010) Metabolomics for functional genomics, systems biology, and biotechnology. Annu Rev Plant Biol 61:463–489

Schauer N, Fernie AR (2006) Plant metabolomics: towards biological function and mechanism. Trends Plant Sci 11:508–516

Schuman MC, Baldwin IT (2016) The layers of plant responses to insect herbivores. Annu Rev Entomol 61:373–394

Skoneczny D, Weston PA, Zhu X, Gurr GM, Callaway RM, Weston LA (2015) Metabolic profiling of pyrrolizidine alkaloids in foliage of two *Echium* spp. invaders in Australia—a case of novel weapons? Int J Mol Sci 16:26721–26737

Skoneczny D, Weston PA, Zhu X, Gurr GM, Callaway RM, Barrow RA, Weston LA (2017) Metabolic profiling and identification of shikonins in root periderm of two invasive *Echium* spp. Weeds in Australia Molecules 22:330

Smolinska A, Blanchet L, Buydens LMC, Wijmenga SS (2012) NMR and pattern recognition methods in metabolomics: from data acquisition to biomarker discovery: a review. Anal Chim Acta 750:82–97

Sumner LW, Amberg A, Barrett D, Beale MH, Beger R, Daykin CA, Fan TW-M, Fiehn O, Goodacre R, Griffin JL (2007) Proposed minimum reporting standards for chemical analysis. Metabolomics 3:211–221

Swenson TL, Jenkins S, Bowen BP, Northen TR (2015) Untargeted soil metabolomics methods for analysis of extractable organic matter. Soil Biol Biochem 80:189–198

Viant MR, Sommer U (2013) Mass spectrometry based environmental metabolomics: a primer and review. Metabolomics 9:144–158

Vidkjær NH, Wollenweber B, Gislum R, Jensen K-MV, Fomsgaard IS (2015) Are ant feces nutrients for plants? A metabolomics approach to elucidate the nutritional effects on plants hosting weaver ants. Metabolomics 11:1013–1028

Watson JT, Sparkman OD (2007) Introduction to mass spectrometry: instrumentation, applications, and strategies for data interpretation. Wiley, Chichester, p 862

Weston P, Weston L, Hildebrand S (2013) Metabolic profiling in *Echium plantagineum*: Presence of bioactive pyrrolizidine alkaloids and napthoquinones from accessions across southeastern Australia. Phytochem Rev 12(1–7):831e837

Weston LA, Skoneczny D, Weston PA, Weidenhamer JD (2015) Metabolic profiling: an overview – new approaches for the detection and functional anlysis of biologically active secondary plant products. J Allelochem Interact 2:15–27

Wittstock U, Gershenzon J (2002) Constitutive plant toxins and their role in defense against herbivores and pathogens. Curr Opin Plant Biol 5:300–307

Yan D, Afifi L, Jeon C, Trivedi M, Chang HW, Lee K, Liao W (2017) The metabolomics of psoriatic disease. Psoriasis 7:1–15

Ye M, Song Y, Long J, Wang R, Baerson SR, Pan Z, Zhu-Salzman K, Xie J, Cai K, Luo S (2013) Priming of jasmonate-mediated antiherbivore defense responses in rice by silicon. PNAS 110:3631–3639

Zhang A, Sun H, Wang P, Han Y, Wang X (2012) Modern analytical techniques in metabolomics analysis. Analyst 137:293–300

Zhu X, Skoneczny D, Weidenhamer JD, Mwendwa JM, Weston PA, Gurr GM, Callaway RM, Weston LA (2016) Identification and localization of bioactive naphthoquinones in the roots and rhizosphere of Paterson's curse (*Echium plantagineum*), a noxious invader. J Exp Bot 67:3777–3788