NMR spectroscopy and chemometrics-based analysis of grapevine

Ali, K.

CHAPTER 3

Metabolomics: Platforms, applications, and perspectives for grapes and wine research

Kashif Ali, Young Hae Choi, and Robert Verpoorte

Natural Products Laboratory, Institute of Biology, Leiden University, The Netherlands

Abstract

Plant metabolomics is the technology which aims towards providing unbiased and comprehensive qualitative and quantitative snapshots of the plant metabolites. Many technologies have been developed for the metabolomic studies but gas or liquid chromatography in combination with mass spectrometry and nuclear magnetic resonance spectroscopy are considered as the most useful tools to drive secondary metabolites analysis. The objective of this review is providing basic knowledge regarding the current platforms available for the metabolomics studies with their application in grapevine research. Grape and wine, both globally important food products, have been extensively studied using this approach. Achievement of various objectives like metabolic engineering, stress physiology, quality control, identification of novel metabolites, and prediction of bioactive ingredients in grape and wine is also presented, followed by a brief account on future prospects of this approach in the area of grape and wine research.

Introduction

With the scientific knowledge on the biochemistry of plants exponentially increasing, it becomes clear that plants are capable to produce a wide range of compounds of complex and diverse structures and functions. Each cell type in an organ has its own decided metabolism, necessary to perform vital functions. Many stress factors, like local environment, temporal and seasonal changes, herbivores, and pathogen attack, have to be dealt with and as a response plants have developed a complex system of metabolites. These compounds, known as 'secondary metabolites', are either common to nearly all plants or can be specific to the level of genera or even species. These secondary metabolites are known to play key functions in plant physiology which relate to adaptation of plants to their environment (Lewinsohn et al. 2001), the quality of food (color, taste, aroma), colors and pigments of ornamental plants (Dixon 2001), resistance against pests and diseases, attraction of pollinators and building of symbiotic relations with micro-organisms (Harborne 2001).

It is roughly estimated that the total number of plant metabolites, including primary and secondary metabolites, is in the range of 200,000 and in a individual plant there might

be as many metabolites as genes i.e. in the order of 30,000 (Oksman-Caldentey and Inzé 2004; Verpoorte et al. 2008). Their complete analysis can be used to define plants at the level of genotype and/or phenotype and consequently, this complete insight into the metabolic composition of plant tissues can be utilized for many purposes (Jander et al. 2004). Regardless of extensive knowledge of the metabolic processes, it is crucial to perform unbiased metabolite analysis to define the biochemical functions of plant metabolic pathways (Trethewey et al. 1999). After the establishment of technologies for high-throughput DNA sequencing (genomics), gene expression analysis (transcriptomics), and protein analysis (proteomics), the remaining functional genomics challenge is that of *metabolite analysis*. To understand the term 'metabolomics' and other related terms that emerged from this like 'metabonomics', 'metabolic profiling', and 'metabolic fingerprinting', a short description of these terminologies is presented in Table 1. It is important for the appropriate choice of the tools in the biological experiments to understand each method and know their limitations.

Table 1. Descriptions of some useful metabolomics related terminologies as presented by Dettmer et al. (2007) and Nicholson and Lindon (2008).

Terms	Description
Metabolome	The complete complement of small molecules present in an organism.
Metabolomics	The technology geared towards providing an essentially unbiased, comprehensive qualitative and quantitative overview of all the metabolites present in an organism.
Metabonomics	A non plant term generally defining the technology used to measure quantitatively the metabolic composition of body fluids following a response to pathophysiological stimuli or genetic modification.
Metabolic profiling	Quantitative analysis of set of metabolites in a selected biochemical pathway or a specific class of compounds. This includes target metabolite analysis i.e. analysis of a very limited number of metabolites, e.g. single analytes as precursors or products of biochemical reactions.
Metabolic fingerprinting	High throughput qualitative screening of the metabolic composition in an organism or tissue with the primary aim of sample comparison and discrimination analysis. Generally no attempt is initially made to identify the metabolites present. All steps from sample preparation, separation and detection should be rapid and as simple as is feasible. Often used as a forerunner to metabolic profiling.
Metabolic footprinting	Fingerprinting analysis of metabolites that are excreted by cells (extra cellular) to the culture medium.

Biological sciences have experienced remarkable advances over the past decade. The most striking newly emerged field is 'systems biology', generally considered as the study of the interactions between the components of different biological systems. The different 'omics' are the building blocks of systems biology and comprehensive knowledge of this 'omics cascade' is the basis of this new approach (Figure 1). As indicated, metabolomics is the endpoint of this cascade and closest to the organism's phenotype. The broad analytical coverage of metabolites produced by an organism in response to a stress can lead to better understanding of biochemical mechanism in complex systems. The 'omics' technologies as unbiased methods should be the basis of public databases that can be used for data mining. For sequences of genes and proteins this has already been achieved but for metabolites, this is increasingly difficult in this sequence due to the physical and chemical diversity of these compounds.

The numerous uses of grapes, especially in the production of wine, make it one of the most important fruit crops around the world. The quality of grapevine, as with most plants, mainly depends on its metabolites, the production of which is especially sensitive to external conditions. In particular, the chemical diversity of grapevine is mostly affected by secondary metabolites. Recently a detailed account on grape and wine chemistry, its impact on human health, and its role in biotic and abiotic stresses has been extensively reviewed (Ali et al. 2010). Grapes and wine, as vital food commodities, have attracted many researchers for the metabolites profiling studies in recent years. Dependent on the objective, nearly all the major analytical tools in metabolomics have been used to study various aspects of grape and wine. In grapes, these studies are mainly focused on characterization of different cultivars, monitoring berry development, and studying different biotic and abiotic stress responses. Wine chemistry is generally regarded as one of most complex as it involve so many factors like grape cultivars, vintage, fermentation process, yeast strain, storage, and many more. Many studies targeted to explore the influence of these factors on metabolic classification of different wines and their sensory attributes has been recently published (discussed later).

Here we will discuss the platforms which so far have been developed for wine and grapes and an analysis will be made of what platforms seems most suited for a public database and widespread use for data mining. This review is an attempt to represent a

brief overview regarding the key platforms used in different studies with a short discussion on extraction procedures and data analysis in metabolomics. Special emphasis on the applications of these platforms in grapes and wine research along with future prospects of this approach is also presented.

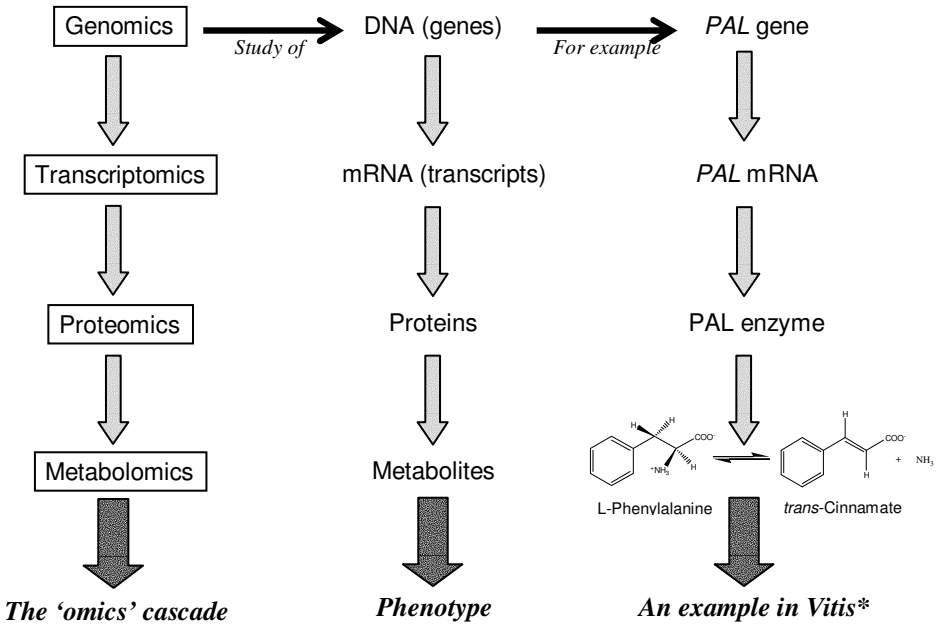

Figure 1. The 'omics' cascade comprised of genomics, transcriptomics, proteomics, and metabolomics which are the building blocks of systems biology approach. The transcriptomics, proteomics, and metabolomics can provide comprehensive information and better understanding about the response of biological system towards different stresses (phenotyping). *The transcription, translation, and function of phenylalanine ammonia-lyase (PAL) gene is a nice example and the GenBank® ID for *PAL* gene, mRNA, and enzyme in *Vitis* is 100233012, DQ887093.1, and ABL74865.1, respectively.

Platforms for grapes and wine metabolomics

To gain a comprehensive and complete overview of the entire metabolic complement of a plant sample in a single analysis is, and should be, the ultimate goal in metabolomics. Since plant metabolomics studies deal with a very large number of metabolites, occuring in different concentrations, with different stability, and diverse chemical properties, this qualitative and quantitative analysis of *'all'* metabolites in a given sample is impossible with the current methodologies. Chemical analysis techniques applied to metabolomics should be unbiased, rapid, and reproducible, while requiring only simple sample preparation (Verpoorte et al. 2010). Different platforms are

available for metabolome analysis including high performance liquid chromatography, gas chromatography-mass spectrometry, liquid chromatography-mass spectrometry, capillary electrophoresis-mass spectrometry, fourier transform ion cyclotron-mass spectrometry, and nuclear magnetic resonance, among them gas or liquid chromatography-mass spectrometry and nuclear magnetic resonance are most widely used. Since no single technology is 'for all', there is a payoff between different technologies in terms of sensitivity, high throughput, robustness, quantitation analysis, and suitability for specific chemical class of metabolites. Nevertheless, a carefully chosen analytical method can be an excellent initial strategy for gaining a first impression of a metabolic profile which can ultimately used to identify key biochemical leads to further or more focused studies (Hirai et al. 2004; 2005).

Grapes and wine chemistry is dominated by very different classes of metabolites and ranges from primary metabolites, like sugars, amino acids, and organic acids, to secondary metabolites, like cinnamates, benzoates, flavonols, flavan-3-ols, and condensed tannins. Many of the above metabolomics platforms have been used in metabolomics based studies on grapes and wines. Studies on grapes are subjugated by NMR since many reports are available for NMR based metabolomics of grapes. On the other hand, wine has been nearly equally studied using NMR or MS based approaches, mainly depending on the objective of the research. A detailed account related to the techniques used and the objectives of grape and wine metabolomics studies will be presented in the later sections of this review. First a crucial part of any metabolomics study, i.e. sample preparation and metabolite extraction, will be discussed.

Sample preparation and metabolite extraction

In metabolomics, a reliable extraction of metabolites from the biological matrices is necessary for an accurate snap shot of the metabolome. The total solubility parameter and constituent partial solubility parameter are widely used for selecting a solvent in the case of pure organic compounds but extraction in metabolomics represents a different scenario. Although there are already remarkable broad metabolic profiles available, even in a single analysis (Kikuchi et al. 2004; von Roepenack et al. 2004; Tohge et al. 2005), the complete coverage is always a daunting task due to the existence of a wide array of compounds at different concentrations and polarities. This makes a single, high

throughput extraction procedure to extract the whole range of metabolites impossible. There are many factors, like solvent, time, temperature, pH, energy, solubility, and dissolution rate, which influence the extraction process. All current extraction and detection techniques, irrespective of their high level of sophistication, have unavoidable built-in bias towards certain metabolite groups (Hall et al. 2005). No single extraction/detection technique is therefore able to cover the desired broad metabolomic picture and parallel technologies are required to achieve this goal. Careful selection of suitable combinations of extraction, separation and detection protocols can lead to a rapid build-up of complementary biochemical data on the composition of biological samples (Bino et al. 2005; Broeckling et al. 2005; Hirai et al. 2004; Sato et al. 2004).

Based on sample chemistry and aim of the research, many extraction protocols for metabolomics studies have been published and some are widely accepted, offering different advantages but also having limitations (Lisec et al. 2006; De Vos et al. 2007; Kruger et al. 2008; Kim et al. 2010a). For samples with special characteristics like high sugar content (grape berries, for example), the analytical method needs to be combined with a more specific sample preparation (the use of solid phase extraction, for instance) in order to detect minor compounds. Large-scale sample preparation can result in significant time differences between the moment of extraction and measurement of the different samples and even with -80°C storage this can be reflected in the results obtained. This adds an undesired extra complication during data analysis. The use of robotics for sample preparation has proven a valuable approach (Lindon et al. 2004). Robotics offers opportunities for high throughput, parallel, uniform and reproducible extraction; however, widespread use for plants has not yet been established. Other extraction and sample preparation pitfalls, and how to avoid them, are described in detail by Dunn and Ellis (2005) and Kim and Verpoorte (2010c), while the risks of artifact formation during extraction have been reviewed by Maltese et al. (2009).

Many platforms are being used for the high throughput analysis of plant metabolites, but vary by their sensitivity. In the following section we discuss the most widely used platforms for the metabolomics studies with their pros and cons along with a brief account on the key applications of these platforms in the area of grapes and wine research.

Mass spectrometry (MS)

Mass Spectrometry is considered as a primary detection method of choice for metabolomics due to its sensitivity, speed and broad application. Many papers have shown its suitability for metabolite detection in complex matrices (Fiehn et al. 2000; Fiehn 2002). Gas chromatography (GC) or liquid chromatography (LC) is commonly used for metabolite separation prior to MS detection. These combinations of different separation techniques with MS along with their applications have been extensively discussed by Dettmer et al. (2007). Following sections present brief accounts on each of these techniques and their key applications in grapes and wine analysis.

Gas chromatography–mass spectrometry (GC-MS)

This technique is the most popular and widely applied method in metabolomics. This popularity is mainly due to robustness of both separation and detection along with the availability of some excellent metabolites identification tools. This technique combines high sensitivity and resolution with a reproducible fragmentation pattern of the separated molecules. Application of two-dimensional GC-MS has resulted in the improvement of resolution (Dallüge et al. 2003; Vial et al. 2011). GC-MS is the principal technique for separation and detection of metabolites that are naturally volatile at temperatures up to 250 °C (e.g. alcohols, monoterpenes and esters) at the cost of thermolabile compounds. The technology can also be applicable to groups of nonvolatile, polar (mainly primary) metabolites, such as amino acids, sugars and organic acids, by converting these into volatile and thermostable compounds through chemical derivatization. These derivatized samples can then be analyzed by GC-MS and detailed information on many of the key primary metabolites in plants can be obtained in a single chromatographic process (Roessener et al. 2001; Desbrosses et al. 2005). However, the extent of derivatization or incomplete derivatization can cause the problem of more than one peak for the same compound (Ryan and Robards 2006). Though comparison between chromatograms of identical peaks is possible but absolute quantitation further requires calibration curves. Another limitation of this technique is that complex plant secondary metabolites, like phenolic glycosides, cannot be analyzed by gas chromatography.

In the area of grapes and wine research, GC-MS based metabolic profiling has been widely applied. In wine, aroma depends up on many factors which include grape cultivar, yeast, fermentation process, and storage. Wine aroma comes from three different sources: primary, associated with grape aroma; secondary, developed during must fermentation; and tertiary, appears during wine aging. Volatiles and aroma compounds, in grapes and specially wines, have been a major target for GC-MS based metabolite analysis (Vilanova et al. 2007, Sánchez-Palomo et al. 2005). This approach is also found effective in the analysis of fatty acids, and terpenes in grape must and wines (Yunoki et al. 2005). Some of recent applications of GC-MS based profiling of grapes and wine are listed in Table 2.

Liquid chromatography-mass spectrometry (LC-MS)

It is a very important and versatile technology in metabolomics, capable to facilitate the analysis of several large groups of secondary metabolites of plant tissues without any chemical derivatization of the metabolites (Verhoeven et al. 2006). Advances in chromatographic technologies (like ultra performance liquid chromatography) together with advances in column chemistry (like hydrophilic interaction chromatography and long monolithic columns) resulted in a significantly improved separation potential. The technology is inherently restricted to molecules, which can be ionized, either as positively or negatively charged ions, before moving through the MS. The wide range of analytes in terms of molecular weight and polarity along with precise molecular weight determination are certainly the strong points of LC-MS (Gobey et al. 2005; Looser et al. 2005; Sumner 2003). An authorative review on LC-MS in plant metabolomics has been recently published (Allwood and Goodacre 2010). Unlike GC-MS, very few mass spectral libraries are available for LC-MS and this is a key topic being given considerable attention today.

Phenolics are mostly the primary target for the LC-MS based metabolic profiling in grapes and wine. They are a complex group of metabolites particularly contributing to the characteristics of wines. They are also well known for their contribution to pigmentation of different organs along with their role in different biotic and abiotic stresses. In wine they usually are associated with appearance, taste, mouth-feel, fragrance, and antimicrobial activity (Kennedy 2008). Wine phenolics may arise from

the grapes (skins and seeds), and/or may be the products of yeast metabolism. Wine phenolics can be classified into three categories which include flavonoids, stilbenoids, and simple phenolics. Flavonoids are synthesized from the combination of the shikimic and the polyketide pathway and are found primarily in the berry skins and seeds of the grapes. They can be further divided into several sub-groups. The most common flavonoids in wine are flavonols (quercetin, kaempferol, myricetin etc.), flavan-3-ols (catechin and epicatechin), and anthocyanins. Polymerization of polyhydroxy flavan 3-ol units, (+)-catechin and (-)-epicatechin, and their gallate esters produces oligomers and polymers called proanthocyanidins (often referred to as procyanidins). Many studies on grapes or wine using an LC-MS based approach, target specific categories of phenolics like stilbenes (Wang et al. 2002), anthocyanins (Wang et al. 2003; Kosir et al. 2004), and proanthocyanidins (Wu et al. 2005), or phenolic in general (Borbalán et al. 2003). Some recent examples are listed in Table 2, with respective objectives, and analytical details.

Capillary electrophoresis- mass spectrometry (CE-MS)

The recent development of capillary electrophoresis (CE) as an alternative separation technology is growing in popularity particularly when combined with MS for extra selectivity and sensitivity (Soga et al. 2003). High-resolution chromatographic separation and sensitive detection of water soluble extracts make a strong combination suitable for the analysis of a diverse range of charged primary and secondary metabolites (Sato et al. 2004). Derived from CE, capillary electro chromatography (CEC) is another promising separation technique. It uses LC or has monolithic stationary phases hence a hybrid of liquid chromatography and capillary electrophoresis. The combination of CEC with MS, the interfaces used, and different bio-analytical applications like analysis of proteins, peptides, amino acids, saccharides, has been reviewed by Klampfl (2004).

Food and beverages is the major field of application of CE and CEC based metabolites profiling. A substantial number of papers on this approach has been published and have been reviewed extensively (Klampfl et al. 2000; Simó et al. 2003; Mato et al. 2005). Several publications on CE- or CEC-MS based profiling method development and method validation for the analysis of organic and amino acids are available (Vandaveer

et al. 2002; Liu and Cheng 2002; Mori et al. 2003; Powell et al. 2005; Saito et al. 2007; Polesello and Valsecchi 1999). The organic and amino acids contents in grapes and, specially, wine are considered very critical for quality. The amino acids not only contribute to wine taste and appearance (Hernendez-Orte et al. 2002) but also have an influence on aromas during the maturing process (Escudero et al. 2000). Organic acids composition in wine and grape juice is of high importance as it influences the organoleptic properties, is involved in control of microbiological growth, and is a critical parameter in wine stabilization. These acids originate directly from the grapes or are formed in processes like alcoholic fermentation, oxidation of ethanol, and malolactic fermentation. Compounds like tartarate, malate, citrate (from grape berry) along with succinate, oxalate, fumarate, citrate (from fermentation) are known to influence the pH of wine. In the field of grape and wine research, the application of CE based methods seems limited to the analysis of organic acids (Mato et al. 2006; Mardones et al. 2005) and a few recent papers are listed in Table 2.

Fourier Transform Ion Cyclotron Resonance-Mass Spectrometry (FT-MS)

A relatively new technique known as Fourier transform ion cyclotron-mass spectrometry (FT-MS) is capable of non-targeted metabolic analysis and is suitable for rapid screening of similarities and dissimilarities in large collections of biological samples, e.g., plant mutant populations (Cooper and Marshall 2001). After a pause following the first paper on this topic by Aharoni et al. (2002), which explain the metabolic changes during strawberry ripening and to differentiate transgenic and non-transgenic plants, more recent applications are emerging for the phenotypical studies (Hirai et al. 2004; 2005; Murch et al. 2004; Tohge et al. 2005; Brown et al. 2005). This technology requires specialized skills and equipments, not easily accessible to most researchers, and the inability to separate structural isomers which have identical mono-isomeric masses is still seen as a significant limitation to its application. Nearly no literature is available on the application of this methodology in the area of grape and wine research. Though a very interesting publication on non-targeted approach to metabolite identification in wine using FT-MS has been recently published (Gougeon et al. 2009).

Table 2. Application of MS-based platforms in metabolomics studies related to grapes and wine.

Platform	Objectives	Sample type	References
GC-MS	Volatile and aroma compounds profiling	Wine	Falcão et al. 2008; Sarrazin et al. 2007
		Grapes	Pedroza et al. 2010; Parker et al. 2007; Ruberto et al. 2008; Malherbe et al. 2009
	Fatty acids and Terpenes	Grapes	Wenguang et al. 2007
		Wine	Peña et al. 2005
LC-MS	Stilbenes analysis	Grapes	Wang et al. 2002
		Wine	Buiarelli et al. 2007
		Leaves	Jean-Denis et al. 2006; Pezet et al. 2003
	Phenolics profiling	Grapes	Jacob et al. 2008
		Wine	Jaitz et al. 2010; Bravo et al. 2006
	Anthocyanins analysis	Grapes	Mazzuca et al. 2005
	Proanthocyanidins analysis	Wine	González-Manzano et al. 2006
CE- or CEC-MS	Organic acids analysis	Grape and wine	Mato et al. 2007
		Wine	Bianchi et al. 2005
FT-MS	Volatiles and phenolics profiling	Wine	Gougeon et al. 2009

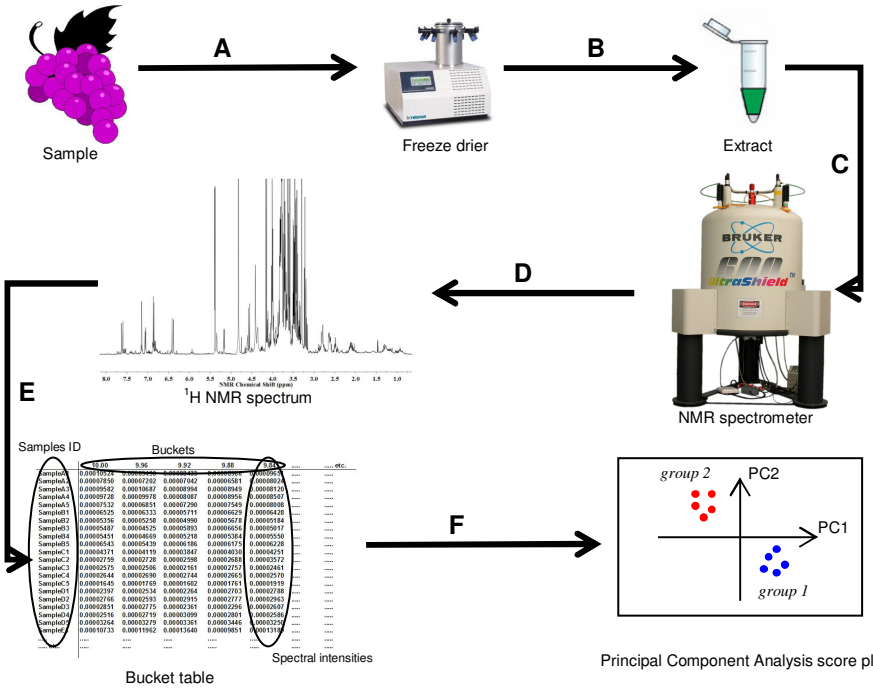

Figure 2. Schematic diagram of major steps involved NMR-based metabolic profiling performed in authors' laboratory. **A:** Samples are ground with mortar pestle and liquid nitrogen and dried in freeze drier. **B:** Completely dried samples are extracted with solvents and the supernatant is collected after centrifugation. **C:** The extract (~ 800 µl) is subjected to NMR analysis. **D:** The ^{1}H NMR and/or 2D dimensional NMR spectra are acquired for metabolites identification. **E:** The ^{1}H NMR spectra are processed and exported to a bucket file which contains buckets of fix width (e.g. δ 0.04) and the spectral intensities in each bucket. **F:** The bucket data is used for different multivariate data analyses methods to highlight the metabolic differences or similarities among the samples.

Nuclear Magnetic Resonance (NMR)

Nuclear magnetic resonance spectroscopy has been widely applied in plant metabolomics and is the first choice for medical metabonomics. Nuclear magnetic resonance analysis is a favored choice for the major metabolites (Defernez et al. 2004; Choi et al. 2004a; Liang et al. 2006) as it often criticized for low sensitivity. The sensitivity of NMR is determined among others by the time of accumulation of the spectra. With a standard 500 MHz NMR, a spectrum of an extract of 50 mg dry weight plant material can be obtained in about 10 minutes (128 scans). Higher field strength, cryoprobes, and microtubes have greatly contributed to the shorter time needed to record a spectrum in NMR spectroscopy and it clearly beats the chromatographic

methods. Also the amount needed in most type of experiments is not limiting. So sensitivity is more related to the dynamic range and the overlapping of signals.

Nuclear magnetic resonance has some unique advantages over chromatography and MS based methods. First, it is a more uniform detection system and can directly be used to identify and quantify metabolites, even *in vivo*. The most promising features of NMR are its non destructive nature, simple sample preparation, in relatively shorter time or even direct measurement of samples e.g. urine. Another major advantage of NMR is that quantification is easy for all compounds as with a single internal standard all the detected metabolites can be quantified without the need of calibration curves for each single compound as signal intensity is only dependent on molar concentration of the compound. Since nearly no sample pretreatment is required in NMR spectroscopy, the inherent properties of the sample are well kept. The non selectiveness of NMR makes it an ideal tool for the profiling of a broad range metabolites (Dixon et al. 2006). NMR has been already demonstrated to be a robust method and unaffected by instrumental and experimental factors as it is the case in other analytical methods. Continuous improvements in instrumentation design may lead to increasing popularity of this approach and a full overview of the current potentials and limitations of NMR has been provided by Ratcliffe and Shachar-Hill (2005) and recently by Schripsema (2010). The schematic workflow of NMR based metabolomics in authors' laboratory is shown in Figure 2.

Nuclear magnetic resonance is apparently the method of choice in grapes and wine research as quite a number of studies have been published. The non selective nature of NMR made it suitable for the analysis of diverse classes of metabolites ranging from amino acids, organic acids, sugars, cinnamates, benzoates, flavonols, and flavan-3-ols, both in grapes and wine. In wine research, NMR spectroscopy has been used for the identification of novel metabolites (Baderschneider and Winterhalter 2000; Kosir and Kidric 2001), analysis of the fermentation process (Kirwan et al. 2008), metabolic characterization (Yuan-Yuan et al. 2007), and determination of minor compounds (Kosir and Kidric 2002). Grapes are also analyzed by NMR in order to investigate the microclimate influence (Pereira et al. 2006a), and monitoring the biochemical changes during berry ripening (Ali et al. 2011). Recent key applications of NMR spectroscopy in grapes and wine metabolomics are also listed in Table 3.

Data analysis

A storm of metabolic, protein and expression profiling activity is now challenging the biologists with 'data'. For relatively old 'omics', the problems of data storage and accessibility to other groups have been mostly resolved. Databases for the genomics like 'GenBank' (www.ncbi.nlm.nih.gov/genbank/), transcriptomics like 'GEO' (www.ncbi.nlm.nih.gov/geo/), and for proteomics database like 'UniProt' (www.uniprot.org) is available for data mining and a universal way of storing data is formalized. Unfortunately, so far, this has not happened yet in the case of metabolomics. In the case of MS based metabolic profiling, great efforts have been made by the groups in Germany and USA in making their own databases of metabolites analyzed by GC-MS and LC-MS, respectively, which are available online for free (http://gmd.mpimp-golm.mpg.de and http://fiehnlab.ucdavis.edu). So far very few like this have happened (e.g. http://hmdb.ca or www.liu.se/hu/mdl/main/) for NMR. Some commercial databases are available for NMR but they are very expensive and certainly out of reach for most of the working groups around the globe. As the current analytical capabilities in metabolomics increase, there will be a need for a central repository with standardized format to store and share data among different working groups.

Apart from data storage, another big challenge in metabolomics is analysis of data. Since metabolomics studies generate too large data to interpret manually, multivariate data analysis methods are used for data interpretation. It is very important for the researcher to use the correct methods of data analysis in order to interpret experimental results. Principal component analysis (PCA) and partial least squares-discriminant analysis (PLS-DA) are the two most widely applied regression methods used to reduce the multidimensionality of the metabolomics data (Ali et al. 2009; 2011) (Figure 3). Principal component analysis (PCA) is considered as a primary tool in metabolomics and helps to understand possible differences between samples. It is an unsupervised method hence the clustering or separation of samples is purely due to similarities or differences, respectively, among all the samples. Partial least squares-discriminant analysis (PLS-DA) is considered as the second step of metabolomics studies. It is a supervised method in which samples are classified into different groups on the basis of creating a set of dummy *Y*-variables or classes. Many good reviews on multivariate data

analyses in metabolomics are available (Goodacre et al. 2004; van der Greef and Smilde 2005). A very recent review explains the basics of data analysis in plant metabolomics (Jansen et al. 2010).

Table 3. Some recent applications of nuclear magnetic resonance (NMR) spectroscopy in grape and wine research.

Sample type	Description	Target metabolites	Reference
Leaves	Metabolic characterization of different cultivars	Amino acids, sugars, organic acids, phenolics	Ali et al. 2009
Berries	Vintage and soil effects on grape berry skin	Amino acids, organic acids	Pereira et al. 2006b
	Characterization of skin and pulp of the berries from different areas	Amino acids, sugars, and organic acids	Pereira et al. 2005
Berries and Wine	Effects of environmental factors on harvested grapes and their wines	Minerals, organic acids, phenolics	Son et al. 2009a
Must and Wine	Analysis of wine fermentation process and yeast strain behavior	Sugars, organic acids, amino acids	Son et al. 2009b
Wine	Identification of novel compounds	Benzoates, cinnamates, flavonoids, lignans	Baderschneider and Winterhalter, 2001
	Analysis of wine fermentation process	Amino acids, and organic acids	López-Rituerto et al. 2009
	Sensory attributes analysis	Sugars, organic acids	Rochfort et al. 2010
	Metabolic characterization	Amino acids, sugars, organic acids	Son et al. 2008
	Analysis of vintage effects	Organic acids, phenolics	Lee et al. 2009
	Metabolic classification based on variety, region, and vintage	Phenolics	Anastasiadi et al. 2009

Applications of metabolite analysis

Plant metabolite analysis has a potentially broad field of applications and this is reflected by the variation in studies emerging over the last few years. As well as being a valuable tool for fundamental science and particularly systems biology type of studies, examples are already emerging where metabolomics is being used in applied situations concerning, for example, crop quality characteristics (Hall et al. 2005) or identifying potential biochemical markers to detect product contamination and adulteration (Reid et

al. 2004). A number of recent publications have shown how metabolomics contributes to our understanding of plant metabolism and its role for plant survival. The major metabolomics platforms and their key applications along with the target metabolites are shown in Figure 4. Recent advancement of plant metabolomic applications in general, and for grapes and wine specifically, are discussed in the following sections.

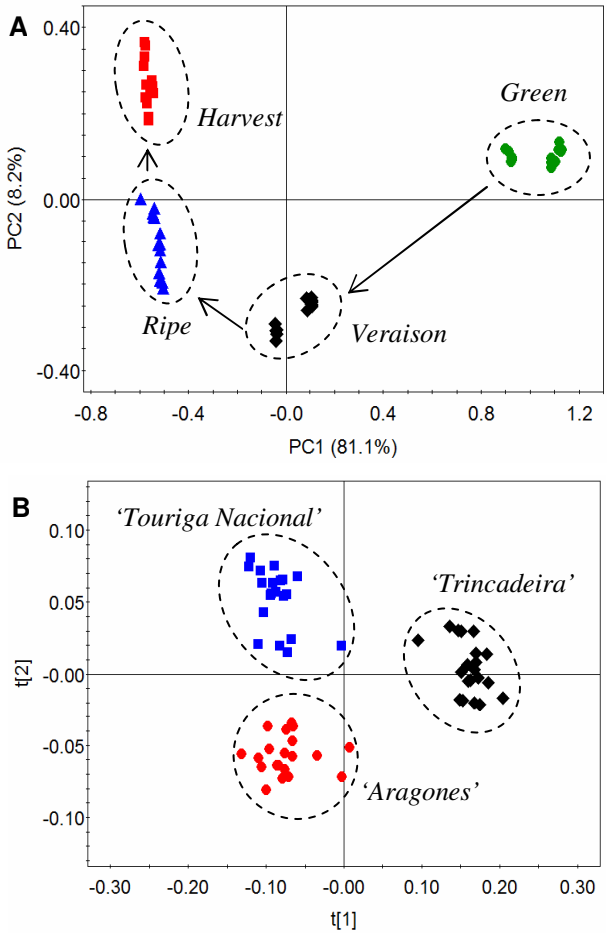

Figure 3. Principal component analysis (A) and partial least squares-discriminant analysis (B) for monitoring grape berry growth and cultivar differentiation, respectively. The score plot (A) shows the clustering of samples from different stages of berry ripening i.e. from green to harvest. The score plot (B) shows the differentiation of three grape cultivars with tight clustering of samples derived from the same cultivar (Adapted from Ali et al. 2011).

Aid to metabolic engineering

The use of the term 'engineering' implies precise understanding of the system that is being modified. Plant metabolic engineering had some spectacular successes but also numerous failures because plant metabolic engineering is in fact a difficult task (Fernie et al. 2002; Carrari et al. 2003; Lessard et al. 2002). The key to improve the success rate is to improve the understanding of the pathways that are subject to engineering. This can be achieved with comprehensive metabolic analyses. A typical metabolic-engineering approach sets out with a focus on a particular metabolic intermediate or final product, such as starch, amino acids, flavonoids, terpenoids, or alkaloids. An assessment of the targeted biosynthetic pathway and the identification of rate-limiting step(s) should be performed. These steps are then subjected to engineering through the over expression of the endogenous gene (cisgenic) or a gene from another plant or organism (transgenic) that encodes the target enzyme. This approach has been successful in a number of cases (Hughes and Shanks 2002; Verpoorte and Memelink 2002; Galili et al. 2002; DellaPenna 2001).

Metabolite profiling technologies are increasingly being adopted to support metabolic engineering projects. In the area of grapes and wine research, the better understanding of metabolic response, via metabolomics, towards infection, abiotic stress, and sensory attributes really help biotechnologists to develop stress and disease resistant cultivars, with increased productivity, sustainability, and taste. Grapevine (*V. vinifera*) has proved to be reluctant to genetic modification and only a few successful attempts have reported. Targets like increased pathogen tolerance (Vidal et al. 2006), resistance to abiotic stress (Tsvetkov et al. 2000), and improved quality factors (Kobayashi et al. 2001) have now been achieved using this approach.

Understanding stress response

Metabolomics is being increasingly used for understanding the cellular phenotypes in response of various types of stresses, both biotic and abiotic. In a study of sulfur deficiency response, a general metabolic readjustment was found (Nikiforova et al. 2005). These observations together with those of Hirai et al. (2005) are likely to advance the field of nutritional stress response further. Metabolite analysis has also been applied in case of cold stress response (Cook et al. 2004). Almost 325 metabolites were up regulated in cold-treated *Arabidopsis* ecotype. In case of biotic stress, several

biosynthetic pathways were found to be involved in plant defense response by utilizing combined metabolomics and transcriptomics approaches (Kant et al. 2004).

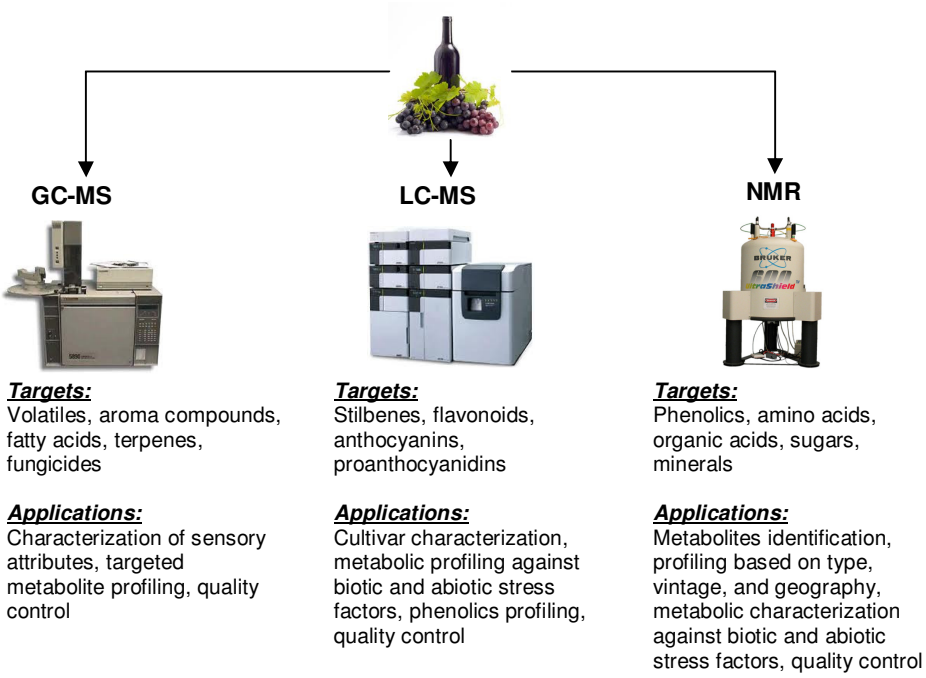

Figure 4. Three most widely applied platforms in grape and wine metabolic profiling studies i.e. GC-MS, LC-MS, and NMR spectroscopy. The major target classes of compounds and few key applications to grape and wine research are also mentioned.

Metabolic profiling is currently providing the best platform to study the stress response in plants. Chemical elicitors' effect, pathogen stress, and climate or vintage effect, are the most studied phenomena in grapes and wine research. The effect of chemical elicitors like methyl jasmonate (Belhadj et al. 2006) and Aluminium chloride (Borie et al. 2004) have been studied and characterized by induced production of grapevine specific phytoalexins and defense related proteins. The metabolic response towards the pathogen infection has also been extensively studied mostly using a targeted approach towards the synthesis or accumulation of stilbenes. Analyses of phenolics, organic acids, amino acids, minerals are the aim of many studies focused on the influence of vintage

on grape (Pereira et al. 2006a; 2006b) and wine chemistry (Son et al. 2009a; Lee et al. 2009; Anastasiadi et al. 2009).

Identification of new metabolites

Metabolomics can be used to identify uncommon plant metabolites. In one case, GC-MS method was used for qualitative and quantitative detection of 150 compounds in *Arabidopsis*. Fifteen uncommon plant metabolites were also identified after these compounds (Fiehn et al. 2000). Selected ecotypes and mutants studies showed the importance of this method for functional genomics studies. The number of metabolites that can be identified in a single analysis have been growing and more than 500 compounds can be identified using FT-MS technique (Hirai et al. 2005).

Nuclear magnetic resonance spectroscopy is the best choice for the studies targeted to identify new or novel metabolites in plant system due to many unique features as previously discussed. Since normally crude extracts are analyzed in metabolomics, identification of new compounds is a daunting task due to highly complex spectra or chromatogram. Application of liquid:liquid fractionation provides a better alternative and by using this approach several novel benzoates, cinnamates, flavonoids, stilbenoids, and lignans, have been isolated and characterized in Riesling wine (Baderschneider and Winterhalter 2000; 2001).

Classification and quality control

Recently, phytomedicines have acquired great importance in the drug industry. To improve the accuracy and consistency of control of phytomedicines worldwide, new analytical methods for their more strict standardization have been employed. Metabolic profiling has proven to be an efficient tool for the quality control and authentication of phytomedicines and is applied to many important medicinal plants like *Artemisia annua* (Ma et al. 2008), *Angelica acutiloba* (Tianniam et al. 2008), and *Panax notoginseng* (Chan et al. 2007). The application of MS and NMR-based metabolomics in the area of quality control has been recently reviewed (van der Kooy et al. 2009; Gilard et al. 2010). Regarding grapes and wine, several studies have been published related to quality control. Pesticide application is a common practice in vineyards across the globe. Both GC-MS (Abreu et al. 2006a; Savant et al. 2010) and LC-MS (Pico et al. 2007; Abreu et

al. 2006b; Venkateswarlu et al. 2007) have been proved to be effective in analyzing pesticide residues in grapes and wines. Metabolites responsible for taste and other sensory attributes in wine have also been characterized by GC-MS (Cuadros-Inostroza et al. 2010) and NMR (Rochfort et al. 2010) based metabolic profiling.

Classification of different plant species and/or cultivars based on metabolic contents is another application of metabolic fingerprinting. Metabolic characterization of plant species, like *Ilex* and *Ephedra* (Kim et al. 2005; Kim et al. 2010b), and cultivars, like *Cannabis sativa* (Choi et al. 2004b), have been successfully done using NMR spectroscopy. The same approach was found very effective in classification of different *Vitis* species (Ali et al. 2009) and grapevine cultivars (Figueiredo et al. 2008) differ in resistance capabilities towards fungal infection.

Biomarkers and bioactivity

The chemicals present in plants generally have an *in vivo* function related to their biochemical structure. These bioactivities can be of great relevance regarding how plants are important for humans. Some components may have potential health-promoting properties (e.g. Chinese medicines, plant infusions with anticancer activity, antioxidant activity, etc.). The use of various supervised multivariate data analyses methods, like projections to latent structures (PLS), is an integral component of such studies. Several studies showed the analysis of extracts from, e.g., *Hypericum perforatum* (Roos et al. 2004), *Artemisia annua* (Bailey et al. 2004), *Citrus grandis* (Cho et al. 2009), and *Galphimia glauca* (Cardoso-Taketa et al. 2008) for the identification of marker compounds responsible for pharmacological activities using NMR spectroscopy in combination with chemometrics methods.

Grapes and wine are known for many pharmacological activities, mainly due to high antioxidant properties of their metabolic contents (Ali et al. 2010; Overman et al. 2010). Recently a report on grape powder extracts' affects on tumor necrosis factor α (TNFα) on human adipocytes culture has been published (Chuang et al. 2010a). Our lab determined the TNFα inhibition activity in different grape cultivars at different stages of their development, and in the different wine types from different vintages. In both cases, we use NMR spectroscopy and chemometric methods to identify the active ingredients. Several phenolics, which were previously reported to have antioxidant, anti-

inflammatory, and anti-TNFα production activity (Chuang et al. 2010b; Baowen et al. 2010; Manna et al. 2000) (quercetin and resveratrol, etc.) were successfully identified using this approach (unpublished data) (Figure 5).

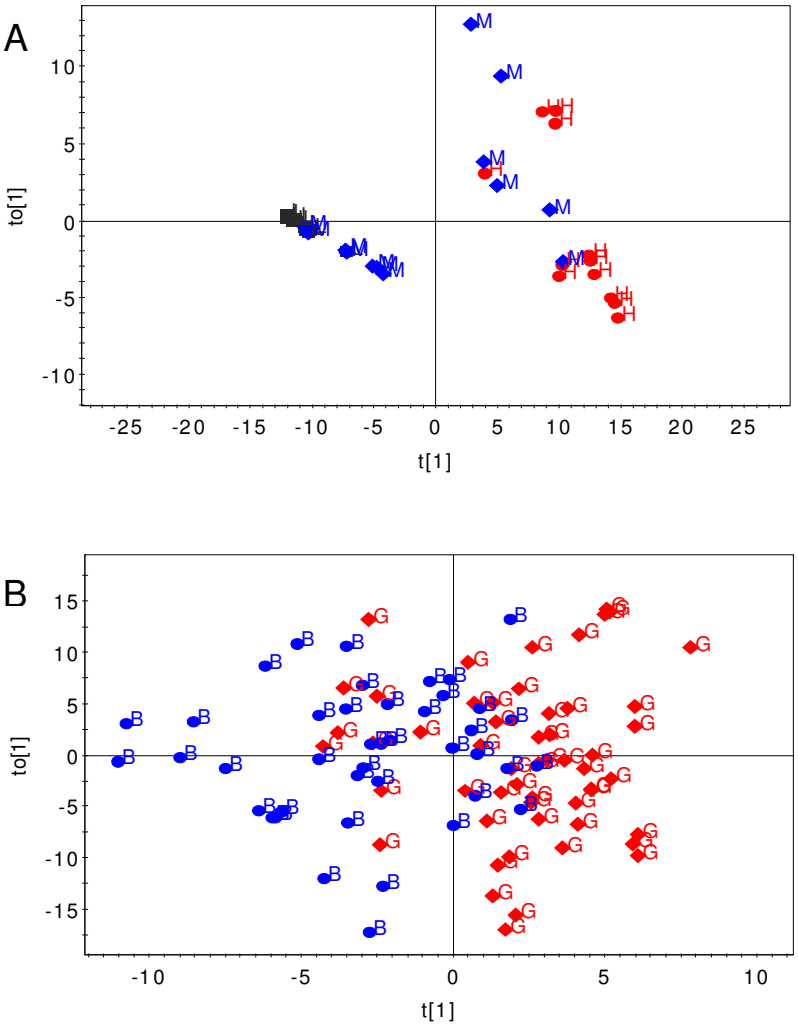

Figure 5. The application of bidirectional orthogonal projection to latent structures (O2PLS-DA) in combination with NMR spectroscopy for the prediction of metabolic ingredients in crude extracts responsible for anti-TNFα activity (A) and wine taste (B). The score plot 'A' present samples with low (L), medium (M), and high (H) activity while 'B' shows wine samples of bad (B) and good (G) taste (unpublished data).

Future prospects and conclusion

So far in metabolomics, the aim is to enhance our capabilities in detection and identification of all metabolites in a living system. Recent advancements in the analytical platforms enable us to increase the number of detectable metabolites drastically but still we need tools for metabolite identification, to enable us to give a more comprehensive coverage of the metabolome. Metabolomics can provide a full-scale discovery platform and can work complementary to other '-omics' to advance our understanding of plant systems. Still a lot more is to be done in developing central data storage not only for the storage of metabolomics data from different analytical platforms and to share this data with the other working groups for data mining.

An area where metabolomics approaches will prove crucial in the coming era is to better define plant genotypes and to relate gene to phenotype. Considering our current capacity for complete genome sequencing, this is where we have to proceed to link gene sequences with functions. The interest in metabolomics as a large-scale assessment of gene expression is greatly accelerating. To accomplish this, it is essential that multiple and parallel approaches for comprehensive analyses can be incorporated. Metabolomic profiling technologies are increasingly being adopted to support metabolic engineering projects. These technologies will become a standard methodology in the future. The wider application of available metabolomic profiling technologies will increase our understanding of metabolic networks by identifying correlations and links between different metabolites, and will facilitate the process of hypothesis generation. To extract the biological information contained in the large metabolomics datasets, bioinformatics tools capable of managing massive experimental data sets and processing them to yield biological knowledge will be imperative.

In the area of grapes and wine, considerable work has been done and some ambitious targets have been achieved using metabolomic profiling approach but this is certainly not the end. Wine is considered as one of the most chemically complex materials with an ever changing chemistry even after fermentation. Related to wine, many aspects are still needed to be explored including the chemistry of fermentation, the grape cultivar, the yeast strain, storage, and bioactivities. Analysis of sensory attributes in wine is also gaining much interest and the metabolic profiling concept can provide a good chance to understand these complex chemical phenomena of one of the most important beverages

of the world. Grapes, as an important fruit, have been subjected to many metabolomics-based studies and resulting in some very useful information on grapevine chemistry. Understanding resistance against certain stress factors, berries ripening, and nutritional and pharmacological values, are definitely areas where metabolomics can largely contribute as these physiological processes mainly relate to certain metabolites. With comprehensive knowledge of metabolic networks, the engineering of cultivars with increased resistance towards various stresses, and with higher nutritional and medicinal importance can be expected.