USE OF MASS SPECTROMETRY FOR METABOLITE PROFILING AND METABOLOMICS

Joel Gummer[1], Michael Banazis[2], Garth Maker[3], Peter Solomon[4], Richard Oliver[1] and Robert Trengove[2,5]*

[1]Australian Centre for Necrotrophic Fungal Pathogens, Western Australian State Agricultural Biotechnology Centre, Murdoch University, WA 6150

[2]Separation Science and Metabolomics Laboratory, Murdoch University, WA 6150

[3]School of Pharmacy, Murdoch University, Murdoch, WA 6150

[4]Plant Cell Biology, Research School of Biology, Australian National University, ACT 0200

[5]Metabolomics Australia, Murdoch University, WA 6150

*Corresponding author: r.trengove@murdoch.edu.au

Introduction

The metabolome of a biological system refers to the complement of all low molecular weight (<1,500 daltons) metabolites in that system (**Fig. 1**). As biological changes in a system are thought to be amplified at the level of the metabolome, metabolites have been coined 'the canaries of the genome'. Metabolomics refers to the quantitative analysis of the metabolome. Whilst the measurement and quantification of individual or small numbers of metabolites is well established in biochemistry, metabolomics differs from more targeted analyses in the number of classes of metabolites being detected, the range of analytical techniques being employed and the need for advanced signal processing and bioinformatics tools.

Different organisms are likely to contain variable numbers of metabolites. For example, well-characterised prokaryotic systems, such as *E. coli*, are estimated to contain approximately 750 metabolites (1). On the other hand, individual eukaryotic cells may contain between 4,000 and 20,000 metabolites (2), while estimates of all metabolites in the plant and fungal kingdoms, which are characterised by having complex secondary metabolism, range into the hundreds of thousands (3). The number of metabolites in specific cell, tissue and biofluid samples of metazoan organisms may also vary markedly. For example, the Human Metabolome Project (http://www.hmdb.ca/) has identified and quantified 6,826 metabolites in human tissues and biofluids. Of these, 3,970 have been identified in serum, while other biofluids, such as urine and cerebrospinal fluid, contain a comparatively simpler composition (472 and 360 metabolites, respectively) (4).

In common with some other '-omics' approaches, metabolomics employs and is highly dependent on diverse analytical approaches (summarised in **Fig. 2**), including mass spectrometry (MS), nuclear magnetic resonance spectroscopy (NMR) and Fourier Transform infrared spectroscopy. Of these approaches, MS-based techniques have developed most rapidly and are increasingly being deployed in metabolomics analyses (**Table 1**). This article provides a short overview of MS-based metabolomics and provides a starting point for scientists considering exploiting this rapidly emerging field.

Sample Preparation

The workflow for metabolomics experiments is summarised in **Fig. 3**. Considerable attention must be directed to ensuring that samples are adequately processed at the point of sampling to ensure all metabolic reactions are rapidly quenched and that any leakage of metabolites from cellular systems is minimised. A wide variety of different sample quenching and metabolite extraction conditions have been developed and there is a need to optimise these for every new system under study. Metabolite extraction conditions will also vary depending on the focus of the study (i.e., polar or apolar metabolites) and the type(s) of instrument(s) being used in subsequent analysis (5). However, as the aim of most metabolomics experiments is to analyse as much of the metabolome as possible, extracts tend to be relatively crude and contain other components (e.g., salts, metals) that can

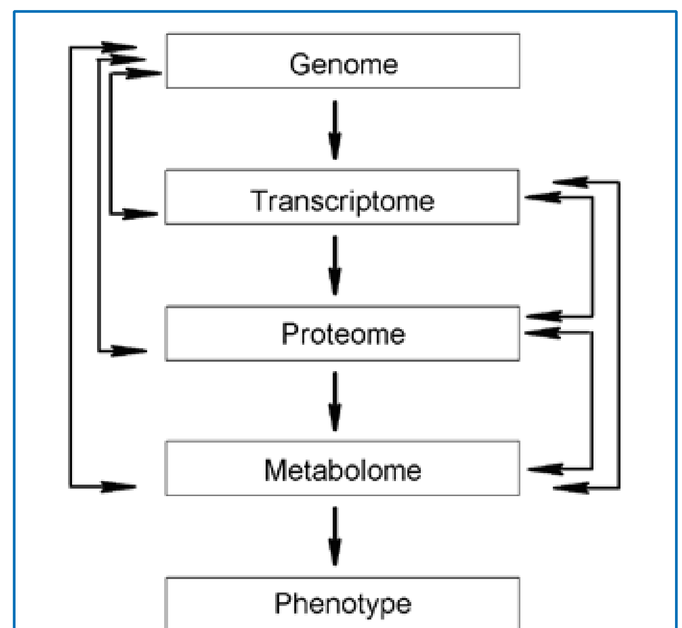


Fig. 1. '-Omics' organisational relationship. General schematic of the '-omic' organisation where the flow is from genes to transcripts to proteins with metabolites the link to phenotype.

affect instrument performance and introduce substantial analytical variability. It is therefore critical that appropriate quality assurance/quality control (QA/QC) protocols are included in all metabolomics protocols (see below). Similarly, consideration must be given to the number of replicates required for each analysis.

GC-MS Instrumentation

Gas chromatography-mass spectrometry (GC/MS) is one of the most common instrument platforms to be used in metabolomics experiments. GC-MS instruments using linear quadrupole analysers have been available for decades providing a robust technology that is amenable to automation. The identification of a wide range of primary metabolites (often after derivitisation) is greatly facilitated by the high resolution of capillary GC, the reproducible fragmentation of metabolites in the mass spectrometer and the ready availability of large mass spectral libraries (6). Recent developments in GC-MS have resulted in improvements in both the GC and MS capabilities of this platform and a move towards the use of high mass accuracy/ high mass resolution instruments. The requirement for high throughput has led to the use of nominal GC-time-of-flight (TOF)-MS with much faster scan rates. High scan rates allow rapid temperature gradient programs, resulting in shorter run times and increased sensitivity. Alternatively, the combination of 2-dimensional GC (GCxGC) with TOF-MS has resulted in the development of very high resolution fast MS that can be used to detect more metabolites than is possible using single quadrupole (Q), TOF and ion trap mass analysers. Recently, GC instruments interfaced with ultra-high resolution Fourier-transform ion cyclotron resonance mass spectrometry (FT-ICR-MS) systems have been developed and introduced commercially. Thus a number of GC-MS configurations are available. While these instruments are often lumped together as one platform, each system provides niche capabilities.

Table 1. Comparison of mass analyser capabilities used in metabolomics experiments.

Chromatographic Technique	Mass Analyser	MS/MS	High Resolution	Strengths	Weaknesses
GC	Q	No	No	Robust Commercial libraries	
	QQQ	Yes	No	Low limits of detection in MS/MS mode	Sensitivity in full scan
	Ion trap	Yes	No		Dynamic range
	TOF Low Resolution	No	No	Very high scan rates	
	TOF High Resolution	No	Yes	Accurate mass	
	QTOF	Yes	?	Accurate mass?	Awaiting development
	FT-ICR-MS	Yes	Yes	Accurate mass Stoichiometry Highest mass resolution	Scan speed
GCxGC	TOF	No	No	Very high compound resolution Excellent limits of detection	
LC	Q	No	No	Robust Commercial libraries	
	QQQ	Yes	No		
	Ion trap	Yes	No	Low limits of detection in MS/MS mode	Sensitivity in full scan
	TOF	No	No/Yes		
	QTOF	No	No		
	Orbitrap	Yes	Yes		
	FT-ICR-MS	Yes	Yes	Accurate mass Stoichiometry Highest mass resolution	
LCxLC	TOF	No	Yes	Polar and nonpolar analytes in a single run	
CE	IT	Yes	No	Anions, neutrals and cations can be analysed in a single analysis run	
	TOF	No	Yes	Accurate mass	
	QTOF	Yes	Yes	Accurate mass	

LC- and CE-MS Instrumentation

One of the drawbacks of many GC-MS metabolomics analyses is the need to derivitise metabolites before analysis. In contrast, many classes of polar metabolites can be analysed directly by liquid chromatography (LC)-MS without derivitisation. LC systems interfaced with TOF mass analysers are now commonly used in metabolomics analyses, delivering high throughput, and high mass resolution analysis capability with mass accuracy approaching single digit ppm. Recently developed instruments also allow rapid polarity switching between positive and negative mode within a single run, reducing the need for multiple runs and cost per sample.

LC-MS linear quadrupole, triple quadrupole (QQQ), QTrap and ion trap mass analysers have also been utilised for global and targeted metabolomics, but may be limited by mass accuracy and mass resolution in identifying metabolites. However, the use of triple quadrupole and QTrap mass analysers in various selective ion scanning modes (precursor ion scanning, neutral loss and multiple reaction monitoring) can be used to detect specific metabolites or metabolite classes with high sensitivity and are particularly useful for targeted metabolomic analysis (see Meikle *et al.* article on page 12 of this Special Technical Feature for more information). In the area of targeted neonatal analysis, the use of tandem mass spectrometry approaches has provided high throughput, high sensitivity and selectivity. These analyses can be carried out without any chromatographic separation and can screen for many health conditions simultaneously.

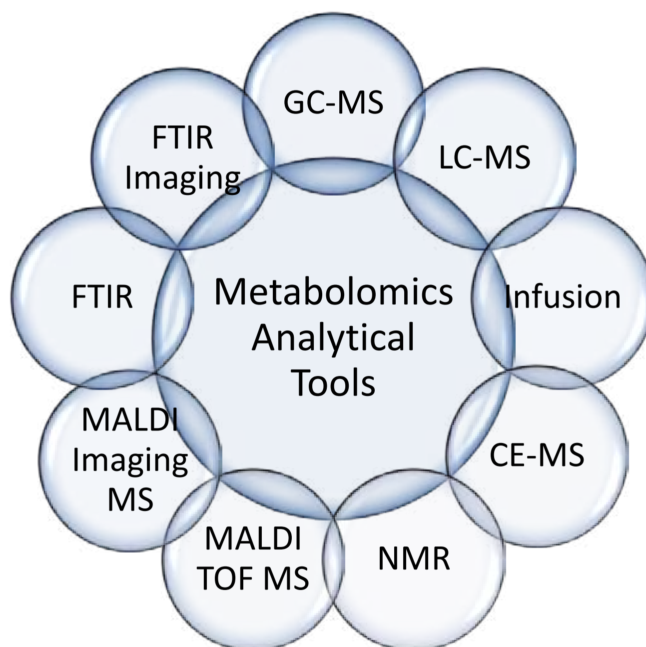
High mass resolution systems, taking advantage of the TOF or Q-TOF mass analyser geometry, or very high mass resolution Orbitrap or FT-ICR mass analysers have recently been used for untargeted or global metabolomics analyses. The TOF and Q-TOF systems have the advantage of high mass resolution and high scan rates, providing high throughput metabolomics systems when coupled with ultra high pressure liquid chromatography (UPLC) systems. FT-ICR mass analysers have reported resolutions as high as 1,000,000 depending on the field of the magnet, although mass resolutions between 100,000 and 250,000 are more typical. It is important to stress that choice of mass analyser will be dictated by many factors, including whether the focus is on targeted vs. untargeted analysis, high vs. low sample throughput, high or low mass resolution and cost of analysis. Comprehensive metabolomic analysis will generally require the use of multiple mass analysers.

Capillary electrophoresis (CE)-MS offers a complementary approach to LC-MS for analysing anions, cations and neutral particles in a single run. Metabolites can be analysed directly without derivitisation and the chromatographic resolution and sensitivity of CE is very high. However, CE is less frequently used for metabolomic analyses than LC-MS with the exception of a number of key metabolomics groups in Japan. This is partly attributable to the technical challenges of maintaining CE systems and the poor retention time reproducibility of CE.

Direct Infusion

LC-MS-based instruments can be operated in direct infusion mode with no chromatographic separation for measurement of the total mass spectrum for the mixture. The infusion can be performed with either the LC autosampler or with an offline syringe pump. Ion trap, TOF, Q-TOF, Orbitrap and FT-ICR-MS mass analysers have been used with this mode of sample delivery. This approach relies totally on the mass analyser to resolve compounds and as such sacrifices the ability to resolve isobaric metabolites such as leucine or isoleucine. The key advantage of direct infusion analysis is the potential for automated high-throughput sample analysis with both low and high mass resolution mass analysers.

Fig. 2. Metabolomics analytical tools.
The analytical tools used for metabolomics analyses: GC-MS (gas chromatography-mass spectrometry); LC-MS (liquid chromatography-mass spectrometry); infusion (direct infusion of sample into an electrospray ionisation [ESI] source, atmospheric pressure chemical ionisation [APCI] source, or atmospheric pressure photo ionisation [APPI] source with no chromatographic separation); CE-MS (capillary electrophoresis mass spectrometry); NMR (nuclear magnetic resonance spectroscopy – may be ^{1}H or ^{13}C, although ^{1}H is much more commonly used); MADLI-TOF-MS (matrix assisted laser desorption/ionisation time of flight mass spectrometry); MALDI imaging MS (direct MALDI mass spectrometry of tissue or surfaces); FTIR (Fourier Transform infrared spectroscopy); and FTIR imaging (direct FTIR on surfaces and tissue).



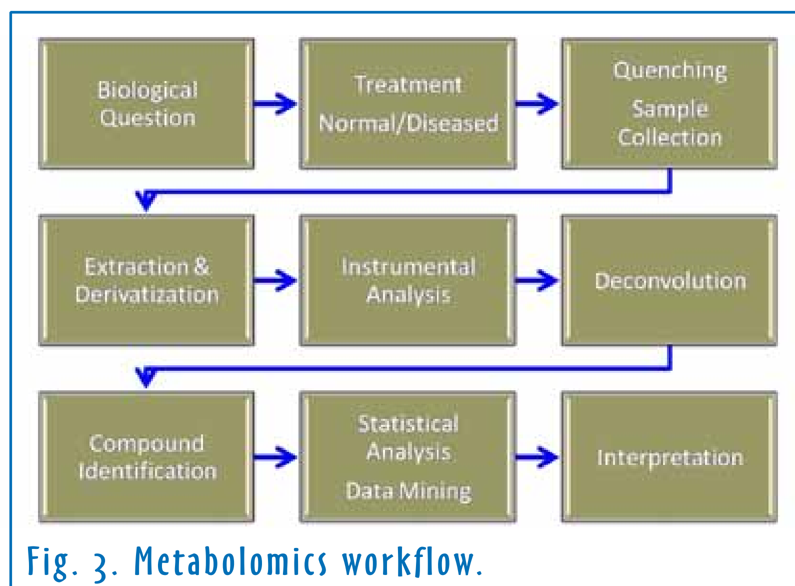


Fig. 3. Metabolomics workflow.

Mass Spectral Imaging Metabolomics

Metabolomic analyses of animal or plant tissues is often complicated by the fact that the tissue may comprise many cell types, each with distinct metabolomes. Information on this spatial heterogeneity is lost when tissue samples are extracted to isolate metabolites. As a result, there is increasing interest in using mass spectral imaging (MSI) of frozen tissue sections to start to map intratissue variations in metabolite levels. MSI has already been developed to detect peptides and proteins in tissue sections, pioneered by Richard Caprioli and collaborators (7). Recent MSI instruments can be used to detect masses in the low molecular weight metabolite m/z range (8). MSI shows promise in allowing the metabolite distribution to be determined 'in tissue' for a single tissue section and 'in organ' from tissue section reconstruction. A key issue in the use of MSI for metabolomics is the ability to decouple the ionisation process from the mass analysis process to remove the potential interference from matrix and matrix clusters. Whilst the Orbitrap and FT-ICR mass analysers have this decoupling by design, other mass analyser geometries can also be configured to minimise matrix interference. As with all analysis systems and techniques, the sample preparation (sectioning of tissue and application of matrix) is critical to the success of the MSI experiment.

Quality Control

All mass spectrometric methods require the implementation of stringent QA/QC procedures to ensure that differences in biological samples are not confounded by alterations in instrument performance. Typically, analysis of QA/QC mixtures (containing metabolites related to or identical to those in the biological samples) are regularly interspersed between analytical runs, and the signal to noise, peak symmetry, and the resolution of closely eluting metabolites compared across the sequence of samples. Some QA/QC mixtures may contain in excess of 40 compounds. Recently, QA/QC mixtures have been generated by pooling all the samples to be analysed (9,10). These pooled QA/QC samples have the advantage that all metabolites in the analysis can potentially be used as a test compound for QA/QC purposes. The latter approach has been used successfully across many mass spectrometry platforms, including GC-MS, LC-MS and UPLC-MS.

Metabolome Coverage

As with proteomic analyses, most current metabolomic analyses only detect and quantitate a subfraction of metabolites in complex biological mixtures. Moreover, comparative analyses using a variety of different analytical techniques suggests that there may sometimes be little overlap in the types of metabolites detected by each method. This was particularly evident in a recent analysis of human cerebral spinal fluid (CSF) by a consortium of groups involved in the Human Metabolome Project (11). Using a combination of NMR, GC-MS and LC-FTMS, these investigators detected 22% (308 metabolites) of all metabolites reported in the literature to be present in CSF. High resolution NMR, GC-MS and LC-MS detected 77%, 60% and 23% of the validated metabolites, respectively. Thus each technique provided a biased view of the metabolome, highlighting the need for a range of sample preparation methods and a combination of both global and targeted analytical methods in analysis of the total metabolome in normal and affected/diseased systems.

Getting Started

As with other '-omics' approaches, careful experimental design and execution is critical in obtaining metabolomics data that will yield useful biological insights. Researchers need to understand the limitations of the sample preparation and analysis techniques before they start their experiments. Collaborating with groups who are actively engaged in metabolomics studies is highly recommended in order to fast track the establishment of methodologies and to tailor the relevant analytical platform to the biological system under investigation.

Acknowledgements

Robert Trengove would like to thank Dr Jon Hall and Dr Patrick Jeanville from Varian Inc for technique complementarity discussions and Dr Steve Fisher from Agilent Technologies for metabolomics methodology and hardware discussions.

References

1. Nobeli, I., Ponstingle, H., Krissinel, E.B. and Thornton, J.M. (2003) *J. Mol. Biol.* **334**, 697-719
2. Fernie, A.R., Trethewey, R.N., Krotzky, A.J., and Willmitzer, L. (2004) *Nature Rev. Mol. Cell Biol.* **5**, 763-769
3. Fiehn, O. (2002) *Plant Mol. Biol.* **48**, 155-171
4. Wishart, D.S., Knox, C., Guo, A.C., *et al.* (2009) *Nucleic Acids Res.* **37**, D603-D610
5. Villas-Bôas, S.G., and Bruheim, P. (2007) *Anal. Biochem.* **370**, 87-97

References continued on page 16

References continued from page 8

6. Roessner, U., Luedemann, A., Brust, D., Fiehn, O., Linke, T., Willmitzer, L., and Fernie, A.R. (2001) *Plant Cell* **13**, 11-29
7. Chaurand, P., Schwartz, S.A., and Caprioli, R.M. (2004) *Anal. Chem.* **76**, 86A-93A
8. Cornett, D.S., Frappier, S.L., and Caprioli, R.M. (2008) *Anal. Chem.* **80**, 5648–5653
9. Sangster, T., Major, H., Plumb, R., Wilson, A.J., and Wilson, I.D. (2006) *Analyst* **131**, 1075-1078
10. Sangster, T., Wingate, J.E., Burton-Lye, Teichert, F., and Wilson, I.D. (2007) *Rapid Commun. Mass Spectrom.* **21**, 2965-2970
11. Wishart, D.S., Lewis, M.J., Morrissey, J.A., *et al.* (2008) *J. Chromatogr. B* **871**, 164-173