Metabolomics Vol. 1, No. 2, April 2005 (© 2005)
DOI: 10.1007/s11306-005-4433-6

Evaluation of automated electrospray-TOF mass spectrometry for metabolic fingerprinting of the plant metabolome

W. B. Dunn[a], S. Overy[b], and W. P. Quick[b,*]

[a]*Department of Chemistry, University of Manchester, PO Box 88, Manchester, M60 1QD, England, UK*
[b]*Department of Animal and Plant Sciences, University of Sheffield, Western Bank, Sheffield, S10 2TN, England, UK*

Received 16 February 2005; accepted 10 March 2005

Metabolic fingerprinting is increasingly employed in microbial and plant metabolomics. Identification and evaluation of analytical factors that influence mass spectra produced with automated electrospray time of flight mass spectrometry to support metabolic fingerprinting are described. Instrument resolution of 4000 (FWHM) at mass 200 Da provided detection of ions of the same nominal mass but different monoisotopic masses. Complex mass spectra were obtained from polar extracts of tomato fruit in positive and negative ion mode. These spectra consist of metabolite ions (molecular, adduct and fragment) and those derived from the extraction medium, largely in the form of $[M+H]^+$, $[M-H]^-$, $[M+Na]^+$, $[M+K]^+$, $[2M+H]^+$, $[M+Cl]^-$ and $[2M-H]^-$. Ionisation suppression reduced sensitivity, although its effect was consistent for a wide range of metabolite concentrations. Variability in ion signal intensity was lower in analytical (2.2–30.1%) compared to biological (within fruit 9.6–27.6%; between-fruit 13.2–34.4%) replicates. The method is applicable to high throughput metabolic fingerprinting and, with accurate mass measurements, is able to provide reductions in data complexity and preliminary identification of metabolites.

KEY WORDS: *L. esculentum*; *L. pennellii*; Electrospray Time-of-Flight Mass Spectrometry (ES-TOF-MS); metabolic fingerprinting; metabolomics.

1. Introduction

Within the 'omics' community a wide array of analytical technologies are available to enhance our knowledge of microbial, plant and mammalian biological systems. The metabolome (Goodacre *et al.*, 2004; Harrigan and Goodacre, 2003) is a complex mixture of chemical species with broadly differing chemical and physical properties which presents limitations to all analytical technologies. Instrumentation utilized in measurement of the metabolome includes mass spectrometry (Roessner, 2000; Soga *et al.*, 2003; Wilson *et al.*, 2005), nuclear magnetic resonance spectroscopy (NMR) (Lindon *et al.*, 2000; Solanky *et al.*, 2003; Viant *et al.*, 2003), Fourier transform infra red spectroscopy (Johnson *et al.*, 2003; Kaderbhai *et al.*, 2003), Raman spectroscopy (Jarvis and Goodacre, 2004) and liquid chromatography with electrochemical detection (Kristal *et al.*, 1998). A number of different analytical strategies are employed (Harrigan and Goodacre, 2003; Goodacre *et al.*, 2004) ranging from the study of a few metabolites common to a specific metabolic pathway or chemical class such as amino acids (metabolite targeted analysis, metabolic profiling) to the comprehensive analysis of hundreds of chemically diverse metabolites (metabolite profiling, metabolic fingerprinting, metabolomics, metabonomics). Currently no single technology can provide non-biased detection of all species present in a metabolome and in future applications it is envisaged that a number of different technologies will be employed to fulfil this objective.

Metabolite profiling involves separation of metabolites, using gas chromatography (Roessner-Tunali *et al.*, 2003), liquid chromatography (Plumb *et al.*, 2003) or capillary electrophoresis (Soga *et al.*, 2003), prior to quantitative and sensitive detection and with metabolite identification by comparison to mass spectral/retention index libraries (Wagner *et al.*, 2003). However, for large studies sample preparation and analysis are both time consuming and laborious. Typical runtimes of 10–20 min provides detection of more than 900 metabolite peaks (O'Hagan, 2005; Plumb *et al.*, 2004; Soga *et al.*, 2003). To overcome difficulties of long runtimes when analysing large sample sets metabolic fingerprinting can be employed. This involves analysis with no chromatographic separation and minimal sample preparation, though generally at the expense of lower sensitivity and a reduced ability to identify metabolites. The objective is to provide high-throughput (1–3 min analysis time) 'global' analyses with the objective of screening samples to enable discrimination between samples of differing biological origin or status.

Electrospray Ionisation Time-of-Flight Mass Spectrometry (ESI-TOF-MS) can be employed and is generally

*To whom correspondence should be addressed.
E-mail: p.quick@sheffield.ac.uk

referred to as Direct Injection (or infusion) Mass Spectrometry (DIMS). Electrospray ionisation is amenable to metabolites containing polar moieties in their molecular structure. Ionisation provides minimal fragmentation of molecular ions and a less complex mass spectrum when compared to the mass spectra of electron impact ionisation of multi-component samples. Although mass spectra are still complex the presence of molecular ions enhances the ability to identify metabolites. Therefore chromatographic separation of complex samples is not required (Schroder, 1996), although some degree of ionisation suppression can be expected. A range of commercially available Time-of-Flight (TOF) instruments can provide a greatly improved mass resolution (FWHM = 5000+ at mass 500 Da) compared to the unit resolution of traditional quadrupole instruments. The higher resolving power allows detection of metabolites of different monoisotopic masses but the same nominal mass, for example, malic acid (MW = 134.0215 Da) and deoxyribose (MW = 134.0576 Da). Additionally, accurate mass measurements can be performed with mass errors of less than 5 ppm using standards of known monoisotopic mass to compensate for small temporal changes in mass calibration accuracy. The standards can be added to the analysis solution or preferentially introduced via a second electrospray probe in a dual ESI source to eliminate possible ionisation suppression effects in quantitative analyses (Wolff *et al*., 2001). Preliminary metabolite identification can be performed by the calculation of molecular formulae of molecular mass equivalent to the accurate monoisotopic masses determined, though further more targeted identification techniques should be subsequently employed.

Early applications of DIMS in the 1990's involved high-throughput screening of clinical samples (blood and urine) by tandem mass spectrometry (Rashed, 2001; Chace *et al*., 2003) for targeted analyses of metabolites indicative of metabolism disorders, such as fatty acid oxidation deficiencies (Carpenter and Wilcken, 1999), organic acidemias (Chace *et al*., 2001) and post-mortem detection of metabolic disorders that resulted in sudden infant death syndrome (Rashed *et al*., 1995). Microbial applications have included the chemotaxonomic discrimination of crude fungal extracts of *Pencillium* cultures (Smedsgaard and Frisvad, 1996), the guidance of culture conditions during batch processes of actinomycetes (Zahn *et al*., 2001) and bacterial characterisation of cell free extracts of *Escherichia coli*, *Bacillus* spp and *Brevibacillus laterosporus* (Vaidyanathan *et al*., 2002) and intact microorganisms of *Escherichia coli* and *Bacillus cereus* (Goodacre *et al*., 1999), where discrimination to sub-species level was achieved. Metabolic footprints has also been used to discriminate between *Escherichia coli* wild-type and tryptophan mutant strains (Kaderbhai *et al*., 2003) and further work discriminated between different physiological states and single gene knockout mutants of *Saccharomyces cerevisiae* (Allen *et al*., 2003). Plant metabolomics studies have included the characterisation of olive oils and of vegetable oils used in its adulteration (Goodacre *et al*., 2002) and discrimination of different photoperiods of *Pharbitis nil* by analysis of the leaf sap (Goodacre *et al*., 2003). Also characterization of medicinal plant extracts (Mauri and Pietta, 2000) and propolis resins (Sawaya *et al*., 2004) by detection of a wide range of secondary metabolites have been reported, with subsequent metabolite identification by tandem mass spectrometry. Finally, Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (FTICR-MS) provides many advantages over other mass spectrometers in terms of sensitivity, resolving power and metabolite identification (Brown, 2005), although for many metabolomics laboratories is prohibitively expensive. Non-targeted metabolome analysis of strawberry fruits and tobacco flowers has been undertaken and showed its potential by identification of metabolites differing between developmental stages of strawberry fruit (Aharoni, 2002).

Experimental and instrumental parameters greatly influence the mass spectra of biological samples. Reliable metabolic fingerprints therefore require robust standardisation and verification. Thus rigorous standards now employed for analysis of microarray experiments (Minimum Information About a Microarray Experiment – MIAME) required for publication in peer-reviewed journals (Brazma *et al*., 2001) should be established for metabolomics (Bino *et al*., 2004; Jenkins *et al*., 2004). In this manuscript the analytical factors influencing mass spectral variability for ESI-TOF-MS are identified and evaluated for reference to future applications, using data obtained with polar extracts from green tomato fruit. A new method of data processing, which does not require specialist software, is also presented and an application showing the differences between cultivated tomato (*Lycopersicon esculentum*) and the wild tomato relative (*Lycopersicon pennellii*) is described. Tomato was chosen as a typically complex biological sample and human foodstuff that would benefit from metabolic fingerprinting using electrospray-TOF instrumentation.

2. Materials and methods

2.1. Chemicals

All chemicals used in this study were of Analytical Grade purity and were purchased from Sigma Aldrich (Dorset, UK).

2.2. Sample extraction and preparation

Samples were taken from the pericarp of mature green tomato fruits of cultivated tomato (*L. esculentum*)

and the wild tomato (*L. pennellii*). Samples were immediately frozen in liquid nitrogen after harvesting and were extracted using a revised method (Valle *et al.*, 1998). 0.1 g of frozen pericarp was ground in 750 μl of an extraction medium comprising buffer (50 mM 4-(2-hydroxyethyl)-1-piperazine-ethanesulfonic acid [HEPES], pH 8.5)/chloroform/methanol (8/20/47 v/v/v), and left on ice for 30 min. Polar metabolites were extracted twice with 400 μl water (polar extract) and non-polar metabolites were retained in chloroform (non-polar extract); only the polar extracts were analysed in this study. Extracts were stored at −80°C until analysis. For positive ion mode studies formic acid was introduced to a final concentration of 0.25% (v/v).

2.3. ESI-TOF mass spectrometry

ESI-TOF-MS analyses were performed with a LCT mass spectrometer operating with a 3.6 GHz time to digital conversion (Waters Ltd, Manchester, UK). MassLynx data system (version 3.5) running under Windows NT on an IBM compatible PC provided instrument control, data acquisition and data processing. The mass spectrometer was operated at a resolution of 4000 (FWHM) at mass 200 in positive (ES+) and negative (ES−) ion modes at capillary voltages of 2800 V (ES+) and 2500 V (ES−), extraction cone at 3 V and sample cone at 20 V. A RF lens voltage of 75 V was chosen to allow detection of masses from 50–800 Da. Source and desolvation temperatures were maintained at 150°C and 120°C, respectively, with desolvation and nebuliser gases flow rates operated at 400 and 100 L/h, respectively. Spectra were collected in centroid mode in the mass range 50–800 Da at a rate of one spectrum/second (0.95 s scan time, 0.05 s interscan delay) and 180 spectra were summed over a period of 3 min before further data analysis and calculations. All responses in this paper are reported as ion counts per seconds (cps). All data from accurate mass measurements were exported as centroid data without smoothing and no background subtraction was performed. Mass calibration was performed daily in the range 50–800 with a range of amino acids and PEG 200–600 in ES+ mode and with organic acids and PEG diacid 600 in ES− mode. The use of amino acids and organic acids improved the accuracy of mass calibration in the mass range 60–200 Da, where most metabolites were detected. A Lockspray™ interface was used to allow automated accurate mass measurements. Leucine enkephalin (5 ng/μl in 50/50 acetonitrile/water) was employed during Lockspray operation as an internal mass reference. Instrument sensitivity and resolution were assessed each day to ensure they met minimum resolution and sensitivity criteria using leucine enkephalin (10 ng/μl in 50/50 acetonitrile/water).

Initial studies (typical mass spectra, ion fragmentation, ion suppression, metabolite identification) used a syringe pump (Razel, Connecticut, USA) at a flow rate of 20 μl/minute to introduce samples directly into the mass spectrometer and employed the Lockspray™ interface for accurate mass measurements. Sample introduction was later automated using a Waters 2690 Separations Module combining a HPLC pump and an autosampler (Waters, Hemel Hempstead, UK) for comparison of tomato species without operation of the Lockspray™ interface. A flow rate of 125 μl/min and injection volume of 25 μl was used. A 3 metre length of PEEK tubing (0.005 inch id; Supelco, Dorset, UK) was used to provide a constant back pressure at the low flow rate employed.

2.4. Raw data processing

Peak lists (accurate mass to 4 decimal places vs. ion counts) were transferred from MassLynx data system to Microsoft Excel (Microsoft Corp, USA) as text files. To minimise loss of low intensity metabolite peaks a noise threshold was not set. Instead, accurate masses from three replicate analyses of the same sample were automatically compared and if the standard deviation of the accurate masses determined for three analytical replicates was less than 0.006 Da the mean accurate mass and mean response of the three replicate analyses were automatically calculated and pasted to a second table. Further refinement of the methodology has been presented (Overy *et al.*, 2005). Small differences in determined accurate masses between replicates occurred due to short-term drifts in mass calibration and errors in accurately defining the peak apex of low response peaks, from which the accurate mass is determined. The mass difference of 0.006 Da was selected after review of experimental data to ensure that low response metabolites were detected and processed forward while also ensuring elimination of noise related peaks during the data reduction process.

2.5. Tentative metabolite identification

In order to provide preliminary identification of metabolites, a list of over 700 metabolites that by studying metabolic pathways are known to be present in plant and microbial metabolomes, was compiled from a range of sources including the KEGG database (http://www.genome.ad.jp/kegg/kegg2.html). The monoisotopic masses of molecular ions likely to be present in plant extracts as $[M+H]^+$, $[M-H]^-$, $[M+NH_4]^+$, $[M+Na]^+$ or $[M+K]^+$ ions were calculated and tabulated by metabolite name vs. monoisotopic mass for each molecular ion; for example lactic acid, $[M+H]^+$, 91.0395. The metabolite identification table is available on request. Metabolites were identified by searching, using the "lookup" function in Microsoft Excel, of determined accurate masses against monoisotopic masses of known metabolites.

3. Results and discussion

3.1. Typical mass spectra

DIMS is used to study microbial and plant metabolomes which are composed of many hundreds of metabolites. Therefore the choice of tomato fruit as a biological system to evaluate metabolic fingerprinting by direct injection ES-TOF-MS is appropriate as samples are complex with hundreds of metabolites to detect in any one extract.

Typical mass spectra of polar extracts in ES+ and ES− modes are shown in Figure 1. The mass spectra (before data processing) are characterised by 3000–4000 peaks (over 30% of which were in the mass range 50–200 Da), ion counts ranging from 1–1500 counts per second (cps), and peak widths of 0.03–0.05 Da. The majority (85%) of the peaks have intensities of less than 100 cps (75% less than 50 cps), with a smaller number of relatively intense peaks (greater than 300 cps). Detected peaks are likely to comprise metabolite-related ions (molecular, adduct and fragment), solvent and extraction medium ions and chemical noise. The analysis of microbial samples exhibits similiar characteristics, though with the exception of fewer metabolite peaks being detected (unpublished data).

Some studies reduce the initial volume of data generated from TOF instruments prior to further data processing by summation of responses for peaks that lie within a broad mass range, typically 1 Da (i.e. all responses for 99.501–100.500 are summed into mass bin 100) (Goodacre *et al*., 2002). The disadvantage of summing responses of different mono-isotopic peaks is the loss of molecular identity. For example, in this study the peak for the amino acid GABA at 104.0711 Da was resolved from that of a solvent peak (104.1102 Da). Small changes in the response for GABA were detected in accurate mass mode despite the presence of the larger response (up to 100 times greater) for the solvent peak. Study of the summed responses did not provide the indication of changes in the response at mass 104 Da as changes were minimal when compared to the larger variation in response measured for the solvent peak. Applying accurate mass TOF mass spectrometry, as detailed here, enables a separate data processing tool to be used to reduce the volume of data and yet maintain maximum biological information. For the mass range 65–800, when 2, 3, and 6 analytical replicate analyses of a fruit extract are processed as described above, the number of peaks common to all replicates is reduced from 3624 to 1125, 538 and 486, respectively. In comparison the application of unit mass resolution mass spectrometers, or equivalent data processing, will produce a maximum of 750 peaks in the mass range 50–800, many of which will have no measured ion intensity with signal-to-noise ratio greater than 3 when compared to a blank analysis. In our experience, 3 replicate analyses provided optimum identification of the maximum number of peaks whilst minimizing instrument time. The methodology outlined here identifies and discards those peaks that represent noise or metabolite peaks of insufficient response to provide accurate mass determination (which can be up to 90% of all peaks detected in a single analysis). A separate method of data processing has been described by Hansen and Smedsgaard and involves the application of a matching algorithm to compare high resolution mass spectra of microbial metabolomes without mass alignment or binning of data (Hansen and Smedsgaard, 2004).

3.2. Metabolite identification

The application of instruments working at resolutions greater than nominal mass can provide preliminary identification of metabolites by calculation of molecular formulae with molecular weights equivalent to accurate masses determined. This allows identification of metabolites of the same nominal but different monoisotopic mass, for example, glutamine (146.0691 Da) and lysine (146.1055 Da) could be experimentally resolved with the instrument employed. However, metabolites with accurate mass differences of generally less than 0.02 Da will not be mass resolved and in these cases the mass accuracy for accurate mass determination of either metabolite will be compromised. In this study we have made use of a commercially available TOF instrument to provide preliminary identification of metabolites, which can be verified by further targeted analysis by chromatographic-mass spectrometry, FTICR-MS or NMR techniques.

A metabolite identification table detailing over 700 metabolites and their monoisotopic masses has been compiled, and used to identify candidate metabolites in plant extracts. Table 1 details metabolites initially identified in tomato fruit extracts on the basis of the match of experimental accurate masses to the theoretical monoisotopic masses (with an associated mass error of less than ± 0.003 Da). Ongoing work employing GC-MS is providing verification of a number of these metabolites. Increasing the mass error to a larger value (± 0.006 Da) increased the number of metabolites initially identified and these frequently had a response of less than 50 cps. In negative ion mode many of the increased number of metabolites identified were organic acids which would be expected to be detected in negative ion mode. The peaks of greatest intensity were clearly identified as the HEPES buffer, and these undoubtedly negatively influenced sensitivity and hence detection of low abundance metabolites within the sample. The influence of buffer

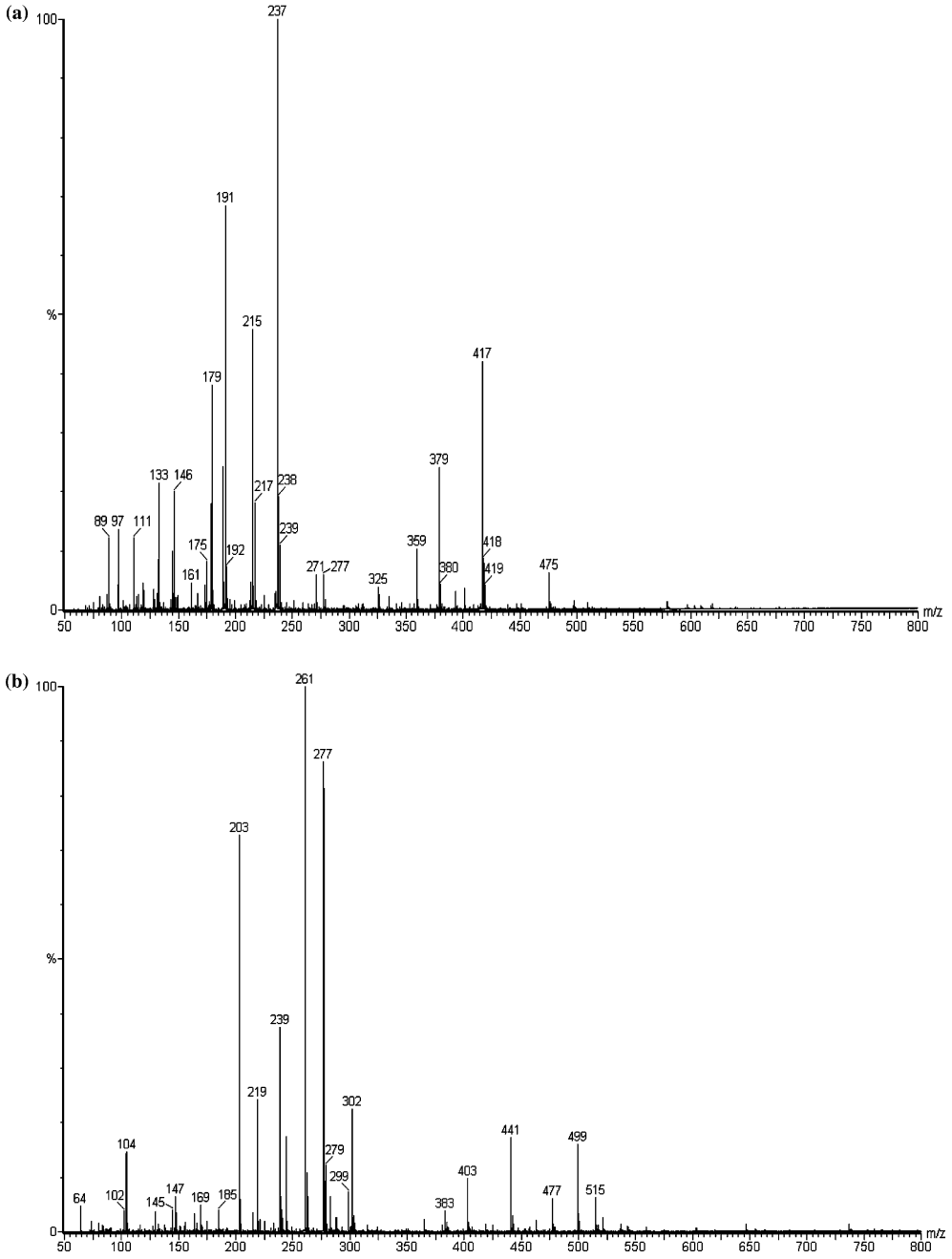

Figure 1. (a) Typical ES− mass spectrum for polar extract of green tomato (*L. esculentum*) fruit. Major identifiable peaks: 179 (hexose sugars, $[M-H]^-$), 191 (citric/iso-citric acid, $[M-H]^-$), 215 (hexose sugars, $[M+Cl]^-$), 237 (HEPES buffer, $[M-H]^-$), 475 (HEPES buffer, $[2M-H]^-$). (b) Typical ES+ mass spectrum for polar extract of green tomato (*L. esculentum*) fruit. Major identifiable peaks: 147 (glutamic acid, $[M+H]^+$), 203 (hexose sugars $[M+Na]^+$), 219 (hexose sugars, $[M+K]^+$), 239 (HEPES buffer, $[M+H]^+$), 261 (HEPES buffer, $[M+Na]^+$), 277 (HEPES buffer, $[M+K]^+$).

type and concentration on mass profiles is worthy of further study though HEPES buffer was used to ensure sample stability during the extraction process.

Compounds were detected as one or more of the following: $[M+H]^+$, $[M-H]^-$, $[M+Na]^+$, higher concentration components as $[2M+H]^+$, $[2M-H]^-$, or as fragment ions. The presence of relatively high salt contents in tomato resulted in ionisation of a number of metabolites as salt adducts and therefore metabolites can produce protonated and salt adduct ions in a mass spectrum. Metabolites can also be detected in one or both ion modes; for example, hexose sugars (such as glucose and fructose) were detected as $[M-H]^-$ and $[M+Cl]^-$ in negative ion mode, and $[M+Na]^+$ in positive ion mode. The ratio of responses of different ions for a metabolite (for example glucose and associated isomers in ES+ and ES− ion modes) varied by up to 30% between different biological samples. This variation is most likely caused by differences in the salt content of different samples. Other compounds, such as

Table 1
Metabolites preliminarily identified by accurate mass determination with mass error of ≤0.003 Da from ES+ and ES− analysis of polar extracts of green tomato (*L. esculentum*) fruit. The theoretical mass of the molecular ion is shown, together with the observed mass range for a minimum of 10 different polar extracts. The standard deviation (sd) of measurements is shown

Metabolite	Ion	Mean ion intensity (cps) (1sd, $n = 3$)	Theoretical mass (Da)	Measured masses (Da) (1sd, $n = 10$)
Propionic acid	$[M-H]^-$	74.3 (8.3)	73.0290	73.0278 (0.0021)
Glycine	$[M+H]^+$	7.6 (1.3)	76.0398	76.0414 (0.0018)
Pyruvic acid	$[M-H]^-$	457 (28.5)	87.0082	87.0098 (0.0020)
Lactic acid	$[M-H]^-$	484 (36.6)	89.0239	89.0229 (0.0020)
Alanine/β-alanine	$[M+H]^+$	10.3 (1.5)	90.0555	90.0559 (0.0021)
Phosphoric acid	$[M-H]^-$	1107 (87)	96.9691	96.9698 (0.0018)
GABA	$[M+H]^+$	28.0 (3.9)	104.0711	104.0723 (0.0017)
Serine	$[M+H]^+$	13.6 (3.5)	106.0504	106.0489 (0.0016)
Proline	$[M+H]^+$	9.6 (1.5)	116.0711	116.0706 (0.0019)
Succinic acid	$[M-H]^-$	47.7 (5.8)	117.0188	117.0195 (0.0017)
Valine	$[M+H]^+$	9.0 (1.2)	118.0868	118.0878 (0.0019)
Pyroglutamic acid	$[M+H]^+$	8.0 (1.1)	130.0504	130.0516 (0.0019)
Leucine/Isoleucine	$[M+H]^+$	10.3 (1.5)	132.1024	132.1027 (0.0021)
Malic acid	$[M-H]^-$	1639 (75.9)	133.0137	133.0149 (0.0017)
Aspartic acid	$[M+H]^+$	22.0 (2.8)	134.0453	134.0462 (0.0018)
Glutamine	$[M+H]^+$	15.3 (2.1)	147.0769	147.0763 (0.0019)
Glutamic acid	$[M+H]^+$	12.0 (1.4)	148.0610	148.0602 (0.0019)
Methionine	$[M+H]^+$	7.6 (1.3)	150.0589	150.0596 (0.0019)
Histidine	$[M+H]^+$	132.0 (12.8)	156.0773	156.0775 (0.0016)
2-aminoadipic acid	$[M-H]^-$	20.6 (2.8)	160.0610	160.0622 (0.0019)
Phenylalanine	$[M+H]^+$	15.0 (1.7)	166.0868	166.0873 (0.0019)
Ascorbic acid	$[M-H]^-$	56.3 (6.0)	175.0243	175.0246 (0.0018)
Arginine	$[M+H]^+$	7.0 (1.2)	175.1195	175.1188 (0.0021)
Hexose sugars	$[M-H]^-$	755 (62.6)	179.0556	179.0539 (0.0017)
Tyrosine	$[M+H]^+$	70.3 (6.7)	182.0817	182.0811 (0.0018)
Citric acid	$[M-H]^-$	1600 (140)	191.0192	191.0179 (0.0018)
Glucoheptonic acid	$[M-H]^-$	16.0 (2.4)	225.0611	225.0603 (0.0020)
Glucose-1-phosphate and isomers	$[M-H]^-$	26.0 (4.3)	259.0219	259.0226 (0.0019)
Sucrose and isomers	$[M-H]^-$	128 (10.7)	341.1084	341.1076 (0.0019)

organic acids, were detected in one ion mode only. To reduce the bias of the technique, detection in both ion modes should be performed. Metabolites present at both low and high concentrations were detected over 2–3 orders of ion intensity (5–1500 cps). Further improvements to the metabolite identification table will be undertaken to allow a larger number of unknown peaks to be initially identified. An intrinsic limitation of the technique is that isomers which have the same monoisotopic mass cannot be separated and will require some form of additional sample fractionation. The ability for metabolites to form molecular and adduct ions can result in the presence of a single metabolite more than once in a peak list. The ability to detect multiple ions for a metabolite allows further confidence in metabolite identification. However to ensure the metabolite is not present in a processed peak list more than once, automated searching can be used to search for mass differences of ±17.0265 (NH_4), ±21.9820 (Na), ±37.9559 (K) and ±33.9611 (Cl) between ions detected. As these monoisotopic masses are unique (there are no other molecular formulae that can be calculated with these masses within 20 ppm mass accuracy), if a match is found it can be said with confidence that two or more ions are present for one metabolite and the adduct ion can be removed from the peak list.

3.3. Fragmentation of molecular ions

In order to evaluate fragmentation of metabolites under the experimental conditions employed, a range of 53 single component metabolite standards (amino acids, organic acids, carbohydrates, sugar alcohols, amines, phosphorylated compounds) were dissolved in water/methanol (1/1 v/v) and analysed in ES+ and ES− modes at varying cone voltages (5–40 V). The number and intensity of mass ion peaks were determined. Voltages below 20 V resulted in very little metabolite fragmentation (fragment ions were detected with responses less than 5% of the molecular ion response), while voltages above 30 V resulted in unacceptable fragmentation to multiple fragment ions for some metabolites studied. Therefore an optimised cone voltage of 20 V was used in all further analyses of tomato fruit extracts in both positive and negative ion modes to provide minimal fragmentation and optimal sensitivity (data not shown). Even at 20 V carbohydrates showed significant fragmentation in

both ES+ and ES− ion modes by, for example, the loss of CH_2O fragments (loss of mass 30 Da) in ES− ion mode. Resulting fragment ions were of the same monoisotopic mass as carbohydrates of lower molecular weight, for example fragmentation of sucrose to hexose and lower mass ions. Therefore, while it is possible to detect differences in carbohydrate metabolism between samples, it is not possible to say which carbohydrates contribute to the difference and further targeted analysis would be required.

3.4. Ionisation suppression

A major consideration with the analysis of multi-component mixtures by electrospray mass spectrometry is ionisation suppression or enhancement, otherwise known as matrix effects (King *et al.*, 2000). The mechanism is as yet not fully understood, though the major controlling factors appear to be observed during ion formation in the liquid phase, droplet formation from the Taylor cone and droplet desolvation and fission. In the multi-component samples analysed in this study, the presence of ionisation suppression is probable. However the implication of ionisation suppression for detecting differences between samples has not previously been assessed.

Initial experiments were performed to compare the addition of single component metabolite solutions to either solvent only or plant extracts. Separate aliquots of solvent (85/15 water/methanol) and plant extracts were spiked with single component standard solutions containing one of thirteen metabolites. In both ionisation modes the response for metabolites in single component, solvent only solutions was greater than in multi-component extraction solvent solutions (by up to a factor of 4 for amino acids in ES+ ion mode). This clearly demonstrates the occurrence of ionisation suppression in complex tomato extracts, and most likely other samples. Therefore to enable representative comparison of biological samples there should be minimal changes in the overall metabolic composition of the biological sample being compared.

Further studies employed spiking aliquots of plant extracts with one multi-component standard solution containing 13 metabolites (tryptophan, leucine, proline, GABA, alanine, pyruvic acid, oxalic acid, fumaric acid, succinic acid, malic acid, ascorbic acid, citric acid, glucose) to provide a number of solutions with final concentrations in the range 2–753 μmol. These amounts represent the measured concentration range of these metabolites in plant extracts. Figure 2 shows the relationship between peak intensity and the concentration of standard added to a tomato fruit extract. In ES+ mode the relationship between concentration and response showed linearity typically over the range 1–753 μmol (correlation coefficients 0.983–0.996), while in ES− mode linearity (correlation coefficient 0.981–0.994) was observed typically over a 1–100 μmol concentration range. Although the relationship between response and concentration was non-linear at higher concentrations (data not shown) there was still an increase in signal amplitude with increasing metabolite concentration and hence relative changes in metabolite abundance between samples can be detected up to millimolar concentrations. Loss of linearity at higher concentrations, for most metabolites studied, was likely caused by the onset of MCP detector saturation where the time interval between consecutive single ions reaching the detector is less than the required detector electronics recovery time between consecutive ion events (detector dead-time). Here the relationship between number of ions reaching the detector and detector response becomes nonlinear. The instrument company quotes the onset of detector saturation as 300 cps. The presence of the ionic HEPES buffer at millimolar concentrations may have a greater influence on ionisation suppression than the μmol concentrations of individual metabolites found in the extracts and, since HEPES is present in all samples, may well have a stabilising influence on ionisation suppression. As a precaution, the use of relative abundance of ions as a measure of individual metabolite concentration should be avoided in samples that have considerable differences in the concentration of different metabolites present or in the sample matrix composition.

3.5. Analytical and biological precision

A range of metabolites identified in fruit extracts were chosen to assess analytical and biological variation between samples. We typically observed that the signal-to-noise ratio was improved with increasing data collection time over a period of one to 5 min (data not shown). The number of metabolites that could be identified from their accurate mass also increased with increasing data collection time; results showed that 37, 54 and 59 metabolites were identified for 1, 3 and 5 min analysis times, respectively. Metabolites were initially identified on the basis of the match of experimental accurate masses (mass error less than ±0.003 Da) to the theoretical monoisotopic masses of the metabolite. To ensure a large number of metabolites were identified while maintaining high throughput (short sample duration) a compromise was made and an analysis time of 3 min was chosen.

The ability to observe differences between different biological samples requires that the variation resulting from endogenous biological differences be greater than analytical variation (resulting from extraction and analysis methods). Three green tomato fruit of similar developmental stage from the same plant were harvested. Three replicate extractions of each fruit were

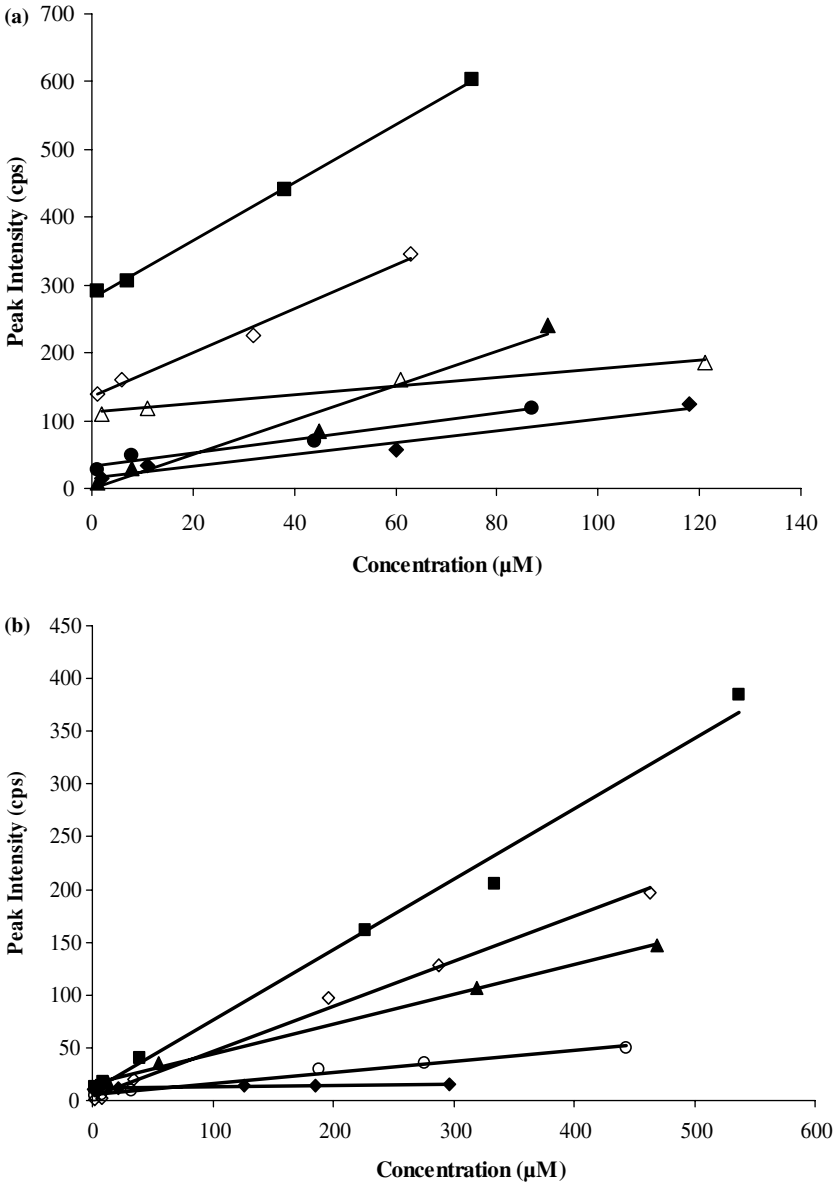

Figure 2. Relationship between concentration of metabolite standard added to a plant extract and molecular ion intensity. (a) ES−; open circle pyruvate, open triangle oxalate, closed circle fumarate, open triangle oxalate, closed square malate, open diamond ascorbate. (b) ES+; open circle alanine, open diamond proline, closed triangle GABA, closed diamond aspartate, closed square leucine.

performed and each extract was analysed in triplicate. Both analytical and biological variation was assessed. Table 2 shows the relative standard deviation (RSD) values for 16 metabolites, again chosen to show the effect across the range of ion intensities experimentally observed. Analytical variability was observed in the range 2.2–30.1% (though many observations were less than 10%) whereas the biological variability was substantially greater (within fruit variation 9.6–27.6%, average between fruit variation 13.2–34.4%). In summary the variability caused by biological differences is greater than the analytical variability for 13 of 16 metabolites, and where analytical variability is greater the ion counts are less than 50 cps. These data show similar trends of variability as described for GC-MS analyses (Roessner, 2000).

3.6. Automated analyses and data improvement

Initially, syringe infusion pumps were used to introduce samples into the mass spectrometer. More recently, we have fully automated the system using a HPLC pump and autosampler. Accurate mass measurement requires continuous addition of a standard with known accurate mass to correct for short-term drifts in mass calibration. With the instrument employed this is normally introduced via a second electrospray probe (Lockspray^{TM}). Since no chromatography was employed we added a known standard as an internal mass reference to each sample so to eliminate the requirement of two electrospray probes. Although for quantitative chromatographic analyses this is not advised, as samples studied are complex the addition of one extra component will not

Table 2
Analytical and biological variation of selected metabolites in extracts of three fruit taken from the same plant. Values for analytical and biological variation are standard deviations shown as percentages of the mean ion intensity (cps) for three replicates

Metabolite	Mean ion intensity (cps) for analytical replicates	Analytical Variation ($n = 3$)	Within fruit Variation ($n = 3$)	Between fruit Variation ($n = 3$)
Glycine	9.2	30.1	22.3	17.1
Asparagine	10.2	2.7	9.6	15.2
Glutaric acid	14.8	8.4	15.6	17.2
Aspartic acid	19.0	3.5	15.3	21.9
Glycerol	19.7	29.2	26.9	15.1
Hydroxybenzoic acid	23.2	6.6	16.1	16.4
Succinic acid	45.0	6.3	13.6	15.1
Glycolic acid	121.6	14.0	25.3	34.4
Histidine	136.5	7.2	12.6	19.8
Sucrose	136.6	2.7	23.4	13.2
Erythrose	225.8	2.4	23.9	24.2
Lactic acid	445.8	2.4	9.9	17.8
Pyruvic acid	472.7	7.3	14.5	27.5
Glucose	798.2	2.2	27.6	19.4
Malic acid	1569	4.1	15.3	20.2
Citric acid	1589	2.7	11.9	15.6
Range		2.2–30.1	9.6–27.6	15.1–34.4

influence sensitivity or ionisation suppression. In ES+ mode VAL-TYR (monoisotopic mass 281.1501 Da) at a concentration of 10 μg/ml in solution was used and in ES− mode the ^{13}C isotope peak of the HEPES buffer (monoisotopic mass 238.0943 Da) at a concentration of 5 mM in solution was used. Accurate mass measurements were corrected for short-term mass calibration drift using the measured and theoretical monoisotopic masses for the internal reference mass, after transfer of data to Microsoft Excel, using the following calculation

$$\text{Corrected mass} = \text{Experimentally measured mass} \times \text{Correction factor}$$

Where the correction factor is defined as

$$\frac{\text{Theoretical monoisotopic mass of internal mass reference standard}}{\text{Experimentally determined accurate mass of internal mass reference standard}}$$

Currently we are operating the instrument as a screening tool to provide metabolic fingerprints, performing up to 252 analyses per day, equivalent to 84 biological samples analysed per day in triplicate. We use a 3 min analysis time to aid metabolite identification though a one minute analysis time can be used to increase sample throughput but with a loss of biological information. In comparison, GC-TOF-MS is a metabolite profiling tool with typical analysis times of 20 min, resulting in 72 analyses per day (24 biological samples analysed in triplicate). This clearly shows the benefits of the application of ESI-TOF to high-throughput screening analyses. However, chromatography-mass spectrometry systems do provide quantification and improved metabolite identification over non-chromatographic techniques. Metabolite concentrations detected for GC-MS (less than 1.0 μmol/g fresh weight) (Roessner, 2000) are experimentally lower than those for the technique described in this paper (greater than 10 μmol/g fresh weight). However this technique is not a quantification tool but a screening tool, since differences observed are depicted by changes in ion responses observed.

The robustness of high throughput techniques is an important consideration. In this study the source was cleaned every 4–5 days (every 1000 analyses) to prevent loss of sensitivity seen when longer periods between cleaning were used. Also the internal reference mass VAL-TYR was monitored over two days (two batches of 15 h runtime each, 504 analyses in total) to determine reproducibility of signal. The mean and relative standard deviation for all analyses was 1.85E + 5 and 6.33%, respectively (response summed over 180 scans). The variation observed is lower than the validation criteria set before the experiment of 10% and is lower than the biological variation observed. The mean and standard deviation of measured responses for ten analyses at start (analyses 1–10), middle (analyses 250–259) and end (analyses 490–499) of the period monitored were statistically compared using Student's t-test. No significant difference was observed between any two sets of data ($p < 0.01$).

3.7. Initial comparison of two tomato species

The analytical and data processing methodology described above was applied to analyse species-specific differences in metabolite profiles of tomato fruit. Metabolic fingerprinting of two tomato species, cultivated *L. esculentum* and wild tomato *L. pennellii*, was performed. Polar extracts of three separate fruit were analysed in triplicate in ES+ and ES− modes with a 3-minute analysis time using the automated method detailed. Figure 3 shows some of the metabolic differences observed between the two species. Metabolites were chosen to reflect the range of experimentally observed ion intensities and the range of metabolites tentatively identified. Data for 13 metabolites are shown, including carbohydrates, organic acids and amino acids. Significant differences were observed in the amounts of 10 of the 13 selected metabolites. This clearly demonstrates the ability of this technique to distinguish species-specific metabolic profiles. Further

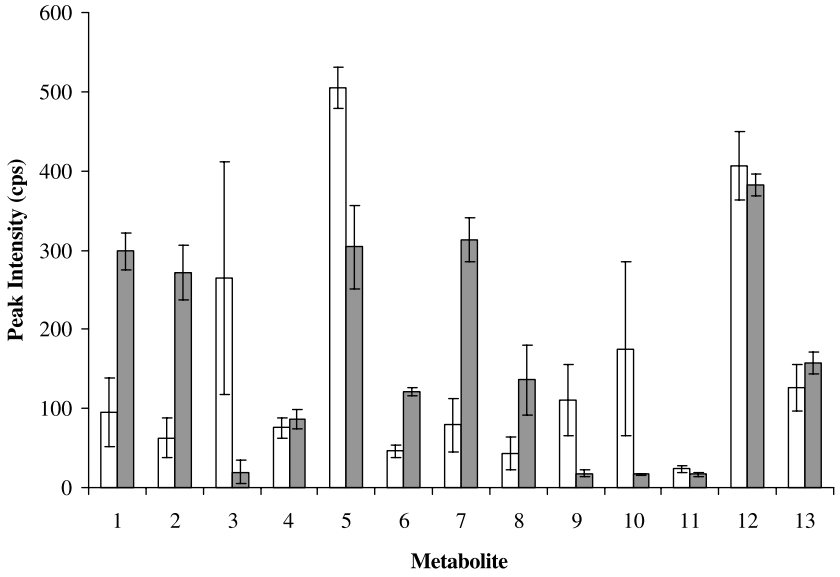

Figure 3. Peak intensity for 13 selected metabolite ions measured in each of three fruit extracts of two tomato species; *Lycopersicon esculentum* white fill and *L. pennellii* grey fill; 1 malic acid, 2 citric acid, 3 GABA, 4 C4 sugars, 5 hexoses, 6 pyruvic acid, 7 fumaric acid, 8 ascorbic acid, 9 valine, 10 leucine and isoleucine, 11 asparagine, 12 glutamine, 13 tyrosine. For clarity, the responses for 3–8 are increased by a factor of 10, and those for 9–13 increased by a factor of 50. Values are ion intensity (cps), calculations employed the summed ion intensity for 180 scans and are presented as the means of three replicate extracts ±1 standard deviation.

work is ongoing to provide metabolic fingerprinting of a large range of tomato lines, and specifically crosses between wild and cultivated varieties with the aim of introducing novel metabolic traits. In comparison to univariate data analysis methods employed here, multivariate methods combined with larger datasets will enhance data analysis and interpretation as shown recently (Overy *et al*., 2005).

4. Concluding remarks

The complexity of samples analysed by DIMS provides certain analytical and data processing issues to be considered. Mass spectral variability and complexity are influenced by ion fragmentation in the source region, ionisation suppression and the composition of the sample matrix, including salt content. However, it should be noted that the range of manufacturer's instruments available can be expected to influence the mass spectra in different ways and therefore validation of different instrumentation should be performed. Analytical variability generally was observed to be less than that from biological variability. However with control of these factors the technique discussed provides a rapid screening method (250 samples per day) with higher throughput than chromatography-mass spectrometry techniques available. Preliminary metabolite identification can be undertaken though with the intention of further identification using targeted approaches such as gas chromatography-mass spectrometry or tandem mass spectrometry without chromatographic separation. The work discussed here highlights the requirements for rigorous standards and control of experimental procedures and instrumental settings to be employed within the metabolomics field.

Acknowledgements

The authors would like to acknowledge financial support for this work from the BBSRC (Grant Number 50/D14495).

References

Aharoni, A., RicDe Vos, C.H., Verhoeven, H.A., Maliepaard, C.A., Kruppa, G., Bino, R. and Goodenowe, D.B. (2002). Nontargeted metabolome Analysis by Use of Fourier Transform Ion Cyclotron Mass Spectrometry. *OMICS* **6**, 217–234.

Allen, J., Davey, H.M., Broadhurst, D., Heald, J.K., Rowland, J.J., Oliver, S.G. and Kell, D.B. (2003). High-throughput classification of yeast mutants for functional genomics using metabolic footprinting. *Nat. Biotechnol.* **21**, 692–696.

Bino, R.J., Hall, R.D., Fiehn, O., Kopka, J., Saito, K., Draper, J., Nikolau, B.J., Mendes, P., Roessner-Tunali, U., Beale, M.H., Trethewey, R.N., Lange, B.M., Wurtele, E.S. and Sumner, L.W. (2004). Potential of metabolomics as a functional genomics tool. *Trends Plant Sci.* **9**, 418–425.

Brazma, A., Hingamp, P., Quackenbush, J., Sherlock, G., Spellman, P., Stoeckert, C., Aach, J., Ansorge, W., Ball, C.A., Causton, H.C., Gaasterland, T., Glenisson, P., Holstege, F.C.P., Kim, I.F., Markowitz, V., Matese, J.C., Parkinson, H., Robinson, A., Sarkans, U., Schulze-Kremer, S., Stewart, J., Taylor, R., Vilo, J. and Vingron, M. (2001). Minimum information about a microarray experiment (MIAME) - toward standards for microarray data. *Nat. Genet.* **29**, 365–371.

Brown, S.C., Kruppa, G. and Dasseux, J-L (2005). Metabolomics applications of FT-ICR mass spectrometry. *Mass Spectrom. Rev*. **24**, 223–231.

Carpenter, K.H. and Wilcken, B. (1999). Neonatal diagnosis of long-chain 3-hydroxyacyl-CoA dehydrogenase deficiency and implications for newborn screening by tandem mass spectrometry. *J. Inherit. Metab. Dis*. **22**, 840–841.

Chace, D.H., DiPerna, J.C., Kalas, T.A., Johnson, R.W. and Naylor, E.W. (2001). Rapid diagnosis of methylmalonic and propionic acidemias: Quantitative tandem mass spectrometric analysis of propionylcarnitine in filter-paper blood specimens obtained from newborns. *Clin. Chem*. **47**, 2040–2044.

Chace, D.H., Kalas, T.A. and Naylor, E.W. (2003). Use of tandem mass spectrometry for multianalyte screening of dried blood specimens from newborns. *Clin. Chem*. **49**, 1797–1817.

Goodacre, R., Heald, J.K. and Kell, D.B. (1999). Characterisation of intact microorganisms using electrospray ionisation mass spectrometry. *FEMS Microbiol. Lett*. **176**, 17–24.

Goodacre, R., Vaidyanathan, S., Bianchi, G. and Kell, D.B. (2002). Metabolic profiling using direct infusion electrospray ionisation mass spectrometry for the characterisation of olive oils. *Analyst* **127**, 1457–1462.

Goodacre, R., Vaidyanathan, S., Dunn, W.B., Harrigan, G.G. and Kell, D.B. (2004). Metabolomics by numbers: acquiring and understanding global metabolite data. *Trends Biotechnol*. **22**, 245–252.

Goodacre, R., York, E.V., Heald, J.K. and Scott, I.M. (2003). Chemometric discrimination of unfractionated plant extracts analyzed by electrospray mass spectrometry. *Phytochemistry* **62**, 859–863.

Hansen, M.E. and Smedsgaard, J. (2004). A new matching algorithm for high resolution mass spectra. *J. Am. Soc. Mass Spectrom*. **15**, 1173–1180.

Harrigan, G.G. and Goodacre, R. (2003). *Metabolic Profiling: Its Role in Biomarker Discovery and Gene Function Analysis*. Kluwer Academic Publishers, London.

Jarvis, R.M. and Goodacre, R. (2004). Ultra-violet resonance Raman spectroscopy for the rapid discrimination of urinary tract infection bacteria. *FEMS Microbiol. Lett*. **232**, 127–132.

Jenkins, H., Hardy, N., Beckmann, M., Draper, J., Smith, A.R., Taylor, J., Fiehn, O., Goodacre, R., Bino, R.J., Hall, R., Kopka, J., Lane, G.A., Lange, B.M., Liu, J.R., Mendes, P., Nikolau, B.J., Oliver, S.G., Paton, N.W., Rhee, S., Roessner-Tunali, U., Saito, K., Smedsgaard, J., Sumner, L.W., Wang, T., Walsh, S., Wurtele, E.S. and Kell, D.B. (2004). A proposed framework for the description of plant metabolomics experiments and their results. *Nat. Biotechnol*. **22**, 1601–1606.

Johnson, H.E., Broadhurst, D., Goodacre, R. and Smith, A.R. (2003). Metabolic fingerprinting of salt-stressed tomatoes. *Phytochemistry* **62**, 919–928.

Kaderbhai, N.N., Broadhurst, D.I., Ellis, D.I., Goodacre, R. and Kell, D.B. (2003). Functional genomics via metabolic footprinting: monitoring metabolite secretion by Escherichia coli tryptophan metabolism mutants using FT-IR and direct injection electrospray mass spectrometry. *Comp. Funct. Genom*. **4**, 376–391.

King, R., Bonfiglio, R., Fernandez-Metzler, C., Miller-Stein, C. and Olah, T. (2000). Mechanistic investigation of ionization suppression in electrospray ionization. *J. Am. Soc. Mass Spectrom*. **11**, 942–950.

Kristal, B.S., Vigneau-Callahan, K.E. and Matson, W.R. (1998). Simultaneous analysis of the majority of low-molecular-weight, redox-active compounds from mitochondria. *Anal. Biochem*. **263**, 18–25.

Lindon, J.C., Nicholson, J.K., Holmes, E. and Everett, J.R. (2000). Metabonomics: Metabolic processes studied by NMR spectroscopy of biofluids. *Concepts Magn. Reson*. **12**, 289–320.

Mauri, P. and Pietta, P. (2000). Electrospray characterization of selected medicinal plant extracts. *J. Pharmaceut. Biomed. Anal*. **23**, 61–68.

O'Hagan, S., Dunn, W.B., Brown, M., Knowles, J. and Kell, D.B. (2005). Closed-loop, multiobjective optimisation of analytical instrumentation: gas-chromatography-time-of-flight mass spectrometry of the metabolomes of human serum and of yeast fermentations. *Anal. Chem*. **77**, 290–303.

Overy, S.A., Walker, H.J., Malone, S., Howard, T.P., Baxter, C.J., Sweetlove, L.J., Hill, S.A. and Quick, W.P. (2005). Application of metabolite profiling to the identification of traits in a population of tomato introgression lines. *J. Exp. Bot*. **56**, 287–296.

Plumb, R., Castro-Perez, J., Granger, J., Beattie, I., Joncour, K. and Wright, A. (2004). Ultra-performance liquid chromatography coupled to quadrupole-orthogonal time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom*. **18**, 2331–2337.

Plumb, R., Granger, J., Stumpf, C., Wilson, I.D., Evans, J.A. and Lenz, E.M. (2003). Metabonomic analysis of mouse urine by liquid-chromatography- time of flight mass spectrometry (LC-TOFMS): detection of strain, diurnal and gender differences. *Analyst* **128**, 819–823.

Rashed, M.S. (2001). Clinical applications of tandem mass spectrometry: ten years of diagnosis and screening for inherited metabolic diseases. *J. Chromat. B-Anal. Technol. Biomed. Life Sci*. **758**, 27–48.

Rashed, M.S., Ozand, P.T., Bennett, M.J., Barnard, J.J., Govindaraju, D.R. and Rinaldo, P. (1995). Inborn-Errors of Metabolism Diagnosed in Sudden-Death Cases by Acylcarnitine Analysis of Postmortem Bile. *Clin. Chem*. **41**, 1109–1114.

Roessner, U., Wagner, C., Kopka, J., Trethewey, R.N. and Willmitzer, L. (2000). Simultaneous analysis of metabolites in potato tuber by gas chromatography-mass spectrometry. *Plant J*. **23**, 131–142.

Roessner-Tunali, U., Hegemann, B., Lytovchenko, A., Carrari, F., Bruedigam, C., Granot, D. and Fernie, A.R. (2003). Metabolic profiling of transgenic tomato plants overexpressing hexokinase reveals that the influence of hexose phosphorylation diminishes during fruit development. *Plant Physiol*. **133**, 84–99.

Sawaya, A., Tomazela, D.M., Cunha, I.B.S., Bankova, V.S., Marcucci, M.C., Custodio, A.R. and Eberlin, M.N. (2004). Electrospray ionization mass spectrometry fingerprinting of propolis. *Analyst* **129**, 739–744.

Schroder, H.F. (1996). Polar organic pollutants from textile industries in the wastewater treatment process - Biochemical and physicochemical elimination and degradation monitoring by LC-MS, FIA-MS and MS- MS. *Trac-Trends Anal. Chem*. **15**, 349–362.

Smedsgaard, J. and Frisvad, J.C. (1996). Using direct electrospray mass spectrometry in taxonomy and secondary metabolite profiling of crude fungal extracts. *J. Microbiol. Methods* **25**, 5–17.

Soga, T., Ohashi, Y., Ueno, Y., Naraoka, H., Tomita, M. and Nishioka, T. (2003). Quantitative metabolome analysis using capillary electrophoresis mass spectrometry. *J. Proteome Res*. **2**, 488–494.

Solanky, K.S., Bailey, N.J.C., Beckwith-Hall, B.M., Davis, A., Bingham, S., Holmes, E., Nicholson, J.K. and Cassidy, A. (2003). Application of biofluid H-1 nuclear magnetic resonance-based metabonomic techniques for the analysis of the biochemical effects of dietary isoflavones on human plasma profile. *Anal. Biochem*. **323**, 197–204.

Vaidyanathan, S., Kell, D.B. and Goodacre, R. (2002). Flow-injection electrospray ionization mass spectrometry of crude cell extracts for high-throughput bacterial identification. *J. Am. Soc. Mass Spectrom*. **13**, 118–128.

Valle, E.M., Boggio, S.B. and Heldt, H.W. (1998). Free amino acid composition of phloem sap and growing fruit of Lycopersicon esculentum. *Plant Cell Physiol*. **39**, 458–461.

Viant, M.R., Rosenblum, E.S. and Tjeerdema, R.S. (2003). NMR-based metabolomics: A powerful approach for characterizing

the effects of environmental stressors on organism health. *Environ. Sci. Technol.* **37**, 4982–4989.

Wagner, C., Sefkow, M. and Kopka, J. (2003). Construction and application of a mass spectral and retention time index database generated from plant GC/EI-TOF-MS metabolite profiles. *Phytochemistry* **62**, 887–900.

Wilson, I.D., Plumb, R., Granger, J., Major, H., Williams, R. and Lenz, E.M. (2005). HPLC-MS-based methods for the study of metabonomics. *J. Chromatogr. B-Anal. Technol. Biomed. Life Sci.* **817**, 67–76.

Wolff, J.C., Eckers, C., Sage, A.B., Giles, K. and Bateman, R. (2001). Accurate mass liquid chromatography/mass spectrometry on quadrupole orthogonal acceleration time-of flight mass analyzers using switching between separate sample and reference sprays. 2. Applications using the dual-electrospray ion source. *Anal. Chem.* **73**, 2605–2612.

Zahn, J.A., Higgs, R.E. and Hilton, M.D. (2001). Use of direct-infusion electrospray mass spectrometry to guide empirical development of improved conditions for expression of secondary metabolites from actinomycetes. *Appl. Environ. Microbiol.* **67**, 377–386.