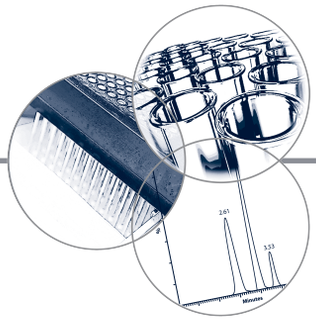

Metabolomics: a second-generation platform for crop and food analysis

The combined factors of financial and food security, a rapidly increasing population and the associated requirement for food generated sustainably in a changing environment have brought food swiftly to the top of most government agendas. The consequence of this is that we need to produce more food at an equivalent or higher quality with lower inputs. These aims are achievable using conventional breeding, but not in the required timelines, and thus state-of-the-art genetic and analytical technologies are coming to the forefront. The concept of metabolomics, underpinned by mainstream (GC–MS, LC–MS, NMR) and specialist (e.g., MALDI-TOF-MS) analytical technologies addressing broad chemical (class) targets and dynamic ranges, offers significant potential to add significant value to crop and food science and deliver on future food demands. Metabolomics has now found a home in the food analytical toolbox with raw material quality and safety the major quality areas, although, as we will show, it is translating beyond this into food storage, shelf-life and post harvest processing.

Louise VT Shepherd[1], Paul Fraser[2] & Derek Stewart[†1]

[1]Plant Products & Food Quality Program, The James Hutton Institute, Mylnefield, Invergowrie, Dundee DD2 5DA, Scotland, UK
[2]School of Biological Sciences, Royal Holloway University of London, UK
[†]Author for correspondence:
Tel.: +44 1382 568 517
Fax: +44 1382 568 503
E-mail: Derek.Stewart@scri.ac.uk

The need for metabolomics

Food research is undergoing a renaissance with the requirement for a greater understanding of how our food is produced, its origins and the changes associated with the multiple, and often highly specialized, post-harvest processes, which include milling [1], climatic modulation [2,3], steaming [4], frying/baking [5], extrusion [6] and modified/controlled atmosphere storage [7,8]. At all stages of these post-harvest processes, and indeed during the crop developmental and ripening stages, the inherent (bio)chemistries change, often with a modification in one component directly affecting another. This has major consequences throughout the food chain not least with respect to food safety [9–11].

Altering the metabolite and protein chemistries and contents of a crop will translate down the chain and impact upon taste, texture, aroma, functionality, bioactivity, nutrition, shelf-life and safety. To characterize and quantify these changes requires a detection system that can go beyond the first generation of analysis targeted at specific compounds and/or chemical classes, and can cover a diverse chemical range. This is the research and analytical space that metabolomics occupies and, with the paradigm shift towards untargeted analysis, suggests that a second generation of analysis is emerging.

Metabolomics, essentially defined by Fiehn as "a comprehensive analysis in which all the metabolites of an organism are identified and quantified" [12] is as good a definition as any. Of course, when we consider foods as they are eaten, this becomes less organism- and more product-specific. However, the advances being made via the multitude of metabolomic technologies including LC–MS, GC–MS, NMR and Fourier transform-infrared (FT–IR) spectrometry, are significantly expanding our knowledge of small molecule interplay and cross talk. This should allow food processors to design, modify and refine food chain processes (raw material to product) whilst facilitating elucidation of the consequences of these changes.

To best show how metabolomics continues to transform crop and food research, we have chosen selected crops that are either staple foods (e.g., potato, wheat and rice) and/or are exemplars where real advances in the application of metabolomics technologies have been made. Furthermore, the crops chosen form part of a food chain, and often require various processing regimes with accompanying metabolite changes; for example, wine [13], beer [14–16], sourdough bread [17] and potato [4]. To cover all plant and crop species is well outside the remit, or indeed space available, of this review, and we acknowledge that others may have equally been chosen. The following sections should give the reader a feel for what is achievable in crop and food analysis and the potential this has for the future, with detailed examples of fundamental through to applied applications of metabolomics with respect to food and agriculture [18], with a special attention to food safety.

FUTURE SCIENCE part of fsg

Key Term

Metabolomics: Comprehensive analysis in which all the metabolites of an organism are identified and quantified.

The opening gambit: extraction & analysis

Food and its associated raw materials, here plant and crops, are by their nature heterogeneous. Indeed, one of the strong arguments to be made for the application of metabolomics in crop and food science is that the associated technologies, especially when used in concert, comprise of a broad range of detection systems covering the majority of the chemistries (metabolites and process-derived components) seen in food (Figure 1). This in turn facilitates the assessment, quantification and interpretation of these variations within a biological framework. This has meant that significant effort has gone into eliminating as much non-sample-related variations as is possible, leading to the drafting of standard reporting requirements for metabolomics experiments with minimum information about a metabolomics experiment (MIAMET) [19] and a data model for plant metabolomics known as architecture for metabolomics (ArMet) [20]. Attempts to construct a more overarching approach have also been initiated [21–24]. These aim to define, as far as possible, the path that a metabolomic experiment must follow to maximize hypothesis testing, as well as the extraction and interpretation of useful information. A key point from many of these, and associated reports, is the requirement to record metadata [24,25] including growth, harvest and storage conditions of sample tissue, sample extraction and instrument parameters. The impact of several of these factors on the metabolome will be discussed later at the crop-specific level.

Strictly adhered to protocols for sample harvesting, extraction and measurement are key to limiting data variability. Coordinated sampling time is crucial for metabolome variability, particularly when analyzing fruit experiments where maturation, ripening [26,27] or post-harvest storage [28–30] is accompanied by significant metabolite changes. A more extreme example of the influence of temporal sampling was reported for *Ginkgo biloba* by Wang *et al.* [31] who reported significant changes in the leaf metabolome when harvested at sunrise and sunset. More in-depth studies of the global changes in the metabolome accompanying diurnals shifts and the associated metabolite interplay were discussed, albeit in the model plant *Arabidopsis*, by Morgenthal *et al.* [32] and Müller-Linow *et al.* [33] who found that the day and night metabolome (and associated networks) were significantly different and not necessarily similar to those predicted from genome-based information.

Similarly, to obtain a viable snapshot of the metabolome, metabolism must be stopped and followed by efficient extraction and analysis. This has been detailed by several authors covering metabolism, with fewer considering plants [34–38]. Furthermore, some have attempted to create catch-all extraction methods for potato [37].

The precise nature of the extraction protocol is dictated by the ultimate analytical platform. NMR (in its multiple dimensional 1D–3D guises) is perhaps the easiest and, following successful exploitation in mammalian metabolomics [39,40], is now widely used in plant and crop metabolomics (see later). NMR has both advantages and disadvantages. The latter rests with the relative insensitivity, and ability, to rapidly distinguish the complex metabolome in plant and crop systems. However, as an unbiased approach with data mining and subsequent statistical analysis, it has more than proved its worth as a metabolomics screening tool in crop and food science (see the crop sections later). Furthermore, the application of 2D [41,42], 3D [43] and higher dimensional [44] approaches offers the ability to resolve structural and temporal metabolome changes.

The MS-based systems linked to GC and LC have generally become the approach of choice in plant and crop metabolomics, as highlighted in the following sections, due to several factors. These technologies have been developed from the standard MS approaches that have generated significant databases, such as the National Institute of Standards and Technology (NIST) Chemistry WebBook with 20,000 electron impact (EI) mass spectra [201] and the commercially available Wiley Registry of Mass Spectral Data containing approximately 660,000 (of which approximately 70,000 are duplicates) compounds [202]. Indeed, there are now many crop- and plant-specific mass spectral databases, with selected examples highlighted in Table 1.

The application of chromatography to metabolomics analysis greatly expands the ability to resolve and characterize compounds, but co-elution and peak-drift still remains a confounding problem. This has been handled using several software approaches. At the chromatographic level, Chrompare is a software package that uses a univariate peak-by-peak approach to the manual and automated correction of retention times and responses of GC- and LC-derived data [45]. Approaches utilizing both the chromatographic and mass spectral data have been developed, and are being regularly applied with the automated

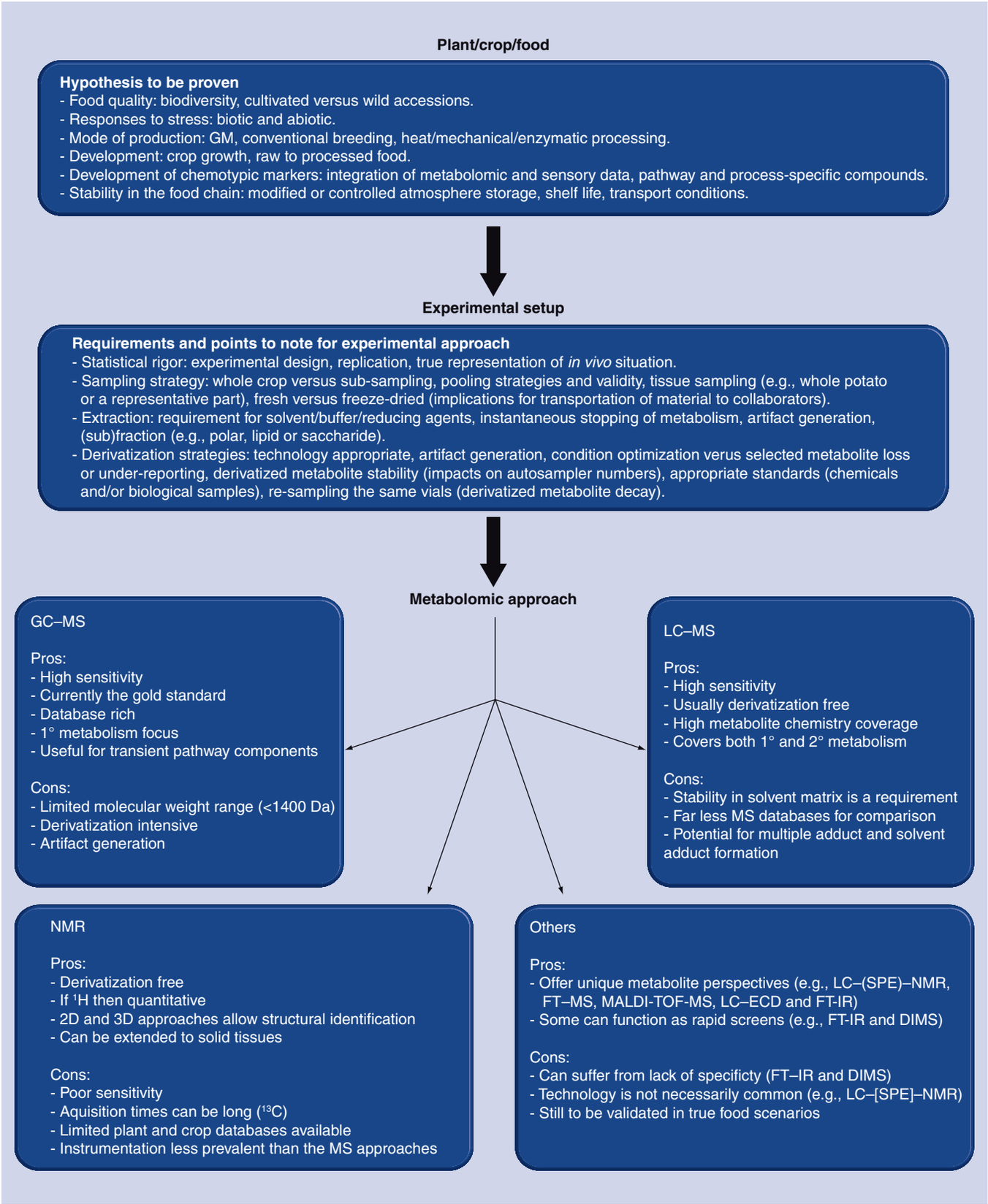

Figure 1. The application of metabolomics to crop and food analysis.
GM: Genetically modified; FT-IR: Fourier transform infrared.

Table 1. Selected examples of metabolite mass spectra databases with a focus on plants and crops.

Database	Technologies	Organisms	Comments	Refs
The Golm Metabolome Database	GC–MS and GC–TOF-MS		General plant- and crop-derived data. Also access to a browseable protocol and methods page	[222]
RIKEN MSn spectral database for phytochemicals (ReSpect) MS/MS spectral tag phytochemical library	GC–MS and LC–MS	Non specific	Data derived from multiple sources including crops. All searchable for cross-comparative purposes	[223,224]
Oliver Fiehn Lab	GC–TOF-MS, LC–Q-MS, CE-MS and LC–MS	Arabidopsis mutant lines	Metabolomic and targeted analysis generated from Arabidopsis mutant lines	[225]
TrichOME: a comparative 'omics database for plant trichome	GC–MS	TrichOME hosts GC–MS data sampled from two cultivars of *Medicago sativa* and *Humulus lupulus*	Linked to transcriptomic and expressed sequence tag analysis	[226]
MoTo DB: a metabolic database for tomato	GC–MS and LC–MS	*Solanum lycopersicum*	Searchable by mass	[227]
MassBank	Multiple MS sources ± chromatography	Multiple sources and standards	High precision and accurate mass spectra of primary metabolites and secondary metabolites	[228]
Metlin KEGG	Q-TOF-MS		Metabolite database for metabolomics, containing over 25,000 structures; also linked to the KEGG	[229,230]
Spectral Database for Organic Compounds	ESI-MS. Also has data derived from FT-IR, ^{1}H and ^{13}C-NMR, laser Raman and electron spin resonance spectroscopies	Nonspecific	Not crop specific but contains a wealth of spectral data for primary and secondary metabolites	[231]

KEGG: Kyoto Encyclopaedia of Genes and Genomes.

mass spectral deconvolution and identification system (AMDIS) developed by NIST [203] and MetAlign [46–47,204] – free-to-download software packages that extract pure (background-free) component mass spectra from highly complex GC- and LC–MS data files, respectively, and uses these purified spectra for a search in a mass spectral library. In general, the main metabolomics technology providers have similarly focused software as part of their own on-board package: XCalibur (Thermo), MassHunter (Agilent) and Chroma/MassLynx (Waters). However, for the metabolomics researcher the problem of mass spectral data redundancy remains a potentially complicating issue. For example, from one compound one may get several peaks associated with the generation of multiple ions from a single compound via adduct formation (Na^{+}, $NH4^{+}$), in-source reaction and dimerization. All of these need not be used for metabolomic analysis and a reduction in the number of those chosen can be established by first establishing a correlation between the multiple ions. Werner *et al.* discussed approaches to doing this using free or commercially available software packages to detect MS peaks differing by exact masses, which correspond to established, typical compositional changes that are reported to occur during the generation of adduct or product ions [48]. Alternatively, an algorithmic–based Random Forest approach wherein the individual variables in complex metabolomic data sets are ranked for significance, thus reducing redundancy to describe the perceived changes, has been outlined with some success for *Arabidopsis* [49] and potato [50].

Once robust data acquisition has been performed, the interpretation almost invariably requires statistical intervention to tease significance and relevance. Due to the multicomponent nature of metabolomic data, various multivariate statistical analytical approaches have been employed with **principal component analysis** (PCA), partial least squares-discriminant analysis (PLS-DA) and orthogonal projections to latent structure-discriminate analysis (OPLS-DA) amongst the commonest methods used to tease apart the data [51–53]. The application of statistical rigor must start earlier, and this is elegantly brought together with respect

Key Terms

Principal component analysis: Mathematical procedure that transforms a number of (possibly) correlated variables into a (smaller) number of uncorrelated variables, called principal components.

Genetic modification: The use of modern biotechnology techniques to change the genes of an organism.

to plant and crop metabolomics by Jansen *et al.* who describe a step-by-step description of a multivariate data analysis, starting from the experiment through interpretation to publication [54].

To obtain a truly representative feel as to how these points are pertinent to plant, crop and food metabolomics the reader should consult the excellent books edited by Nikolau *et al.* [55] and, in particular, the recent one by Hall [56].

Much of the groundwork described above, including the rules of engagement with respect to experimental design and statistical rigor analysis, have been very successfully developed at the academic plant biology level, many of which have been referenced here. However, it is only recently that this approach and accompanying rigor has being adopted into the food sector and although there is crossover, for example at the raw material (crop) end, the food products sector is relatively new to metabolomics utilization. The following sections have been framed with this in mind.

Solanaceous crops

Potato and tomato are by far the most common crops of this family (although other members include peppers and eggplant), with potato currently ranked as the third most commonly consumed crop worldwide. These crops are grown across the globe and are therefore exposed to a broad range of climatic conditions and environments, all of which impact on the plant metabolome.

For potato, significant effort has gone into exploiting metabolomics to mine for different metabolite contents and diversities in wild species collections, with a view to developing new varieties and products. For example, the Commonwealth Potato Collection [205], comprising 1500 accessions of approximately 80 wild and cultivated potato species, was analyzed by GC–MS based metabolomics [57–59]. This showed that taxonomic segregation was evident based on both nonpolar and polar metabolites (e.g., amino acids). Furthermore, a more rudimentary approach, eliminating the metabolite chromatography aspect, using positive-mode direct infusion–MS (DI–MS) differentiated accessions and taxonomic classifications. This was predominantly due to mass ions associated with specific glycoalkaloids with some groups dominated by demissine, others by commersonine, α-tomatine and dehydrodemissine or by α-solanine and α-chaconine. Of course, this approach is very dependent on the susceptibility of the metabolites to accrue and hold charges, with the alkaloids being particularly good at this. By mining a subset of 29 genetically diverse potato cultivars and landraces in more detail, Dobson *et al.*'s GC–MS approaches highlighted several accessions with high levels of specific metabolites (sugars and amino acids) related to end products organolepsis and safety, which could be useful for introgression into cultivated varieties [58,59]. Dobson *et al.* [59] and Beckmann *et al.* [50] took this further to show that, in some instances, metabolomics highlighted wide ranges of, amongst others, asparagine, fructose and glucose – the proposed precursors in the raw material responsible for the generation of the toxin acrylamide in fried potato products [60]. Furthermore, variations in amino acids were identified and these were associated with flavor/aroma, after-cooking blackening and bruising.

Potato metabolomics has flourished when applied in tandem with **genetic modification** (GM) technologies, largely to support the transformations and assessment of the potential for unintended effects. Roessner *et al.* applied a GC–MS approach to GM potato with altered sucrose catabolism and, following detection of ~90 metabolites including sugars, sugar alcohols, amino acids, organic acids and several miscellaneous compounds, they showed that the specific GM event was accompanied by metabolites associated with several metabolic pathways increasing (and some decreasing) in tandem compared with the wild type [61].

An alternative metabolomics approach, this time NMR and LC–MS, was taken by Defernez *et al.* who found that an analysis of 40 GM lines and controls, belonging to four groups of samples (derived from cv. Record or cv. Desirée, and modified in primary carbon metabolism, starch synthesis, glycoprotein processing or polyamine/ethylene metabolism), highlighted that the metabolite-related changes accompanying the GM event were not significant in comparison with those between the two parent varieties, besides those of the targeted event [62]. Indeed, this supported the combined GC–TOF-MS and flow injection (FI)–MS study of Catchpole *et al.* who drew similar conclusions in a metabolomic analysis of field-grown potatoes genetically modified to induce fructan biosynthesis [63]. The major finding from this study was the large variation in the metabolite profile between the five conventional cultivars, which overrode the differences between GM and the associated wild-type parent.

The innate untargetedness of metabolomics, besides that of the detection and separation limitations, is an attractive feature both at the scientific and applied level. This is highlighted well by the study of several distinct GM potatoes (and controls) by Parr *et al.* [64] wherein they identified kukoamine A, a spermine alkaloid, and related compounds in wild-type tubers. Subsequent to this discovery, these compounds have been detected in other Solanaceae such as tomato (*Lycopersicon esculentum*) and tobacco (*Nicotiana sylvestris)*, all as a result of the potential and utility of metabolomics.

The other popular Solanaceous crop, tomato, has also benefited from intense metabolomics effort, in many cases due to its function as a model for fruit and the associated biological processes by functional genomics analysis and metabolic engineering [65–66]. Beyond this are crop and product organolepsis [67] and biofunctionality [68], both of which have attracted metabolomics studies. For example, Deborde *et al.* used a ^{1}H-NMR approach to follow the changes in organoleptic and nutritional quality of greehouse-grown tomato fruit and highlighted that nutrient solution recycling had very little effect on fruit composition [69].

Tomato metabolomics – addressing biodiverse, GM and cultivated material – has been served by GC–TOF-MS, LC–MS and NMR, with the resultant data used to construct metabolic databases [70,206]. As a consequence of this, visualization and analysis packages such as the online Plant MetGenMAP [207] have been developed to rationalize and visualize metabolomic (and associated transcriptomic) changes.

As with potato, the combination of GM and metabolomics has been a boon to plant developmental biologists, with Fraser *et al.* highlighting metabolite–metabolite correlations associated with relative changes following overexpression of a carotenoid biosynthetic gene [71,72]. Indeed, metabolomic analysis showed that there were multiple metabolite correlations going beyond the expected 'within-chemical class' ones, with many primary metabolite–isoprenoid/carotenoid correlations being identified. This was taken further, and at a tissue-specific level, by Mintz-Oron *et al.* [73] who, using a combined GC–MS and ultraperformance (UP)LC–MS approach, found that 100 chemically diverse metabolites, including flavonoids, glycoalkaloids and amyrin-type pentacyclic triterpenoids amongst others, were enriched in the peel tissue during development. These data were then combined with corresponding transcriptomic data to facilitate the temporal biosynthesis of many compound classes during tomato cuticle and cell wall metabolism. Bino *et al.* [74] also used a combined unbiased GC–MS and LC–MS approach to elucidate the changes associated with mutation in the high pigment-2 dark green (hp-2dg) and showed that, compared with the nonmutant control there were many significant changes, particularly in metabolites known for their antioxidant or photo-protective activities, suggesting a role in response to and management of light stress.

A different approach to tomato metabolomics was taken by Fraser *et al.* [75] who, by analyzing tomato using MALDI-TOF-MS, showed that this technology could be used for rapid characterization, identification and quantification (by isotope dilution) of carotenoids present in a host of plant tissues and whole cells. Another alternate metabolomic approach was reported by Pérez *et al.* [76] who applied ^{1}H high resolution magic angle spinning (HRMAS) NMR to tomato development, thus allowing the simultaneous detection of polar and nonpolar metabolite changes in intact fruit tissues to be assessed.

Tomato, like potato, is used in a processed form (e.g., puree or paste) and in some European countries forms a common component in the daily diet [77]. The whole process from fruit to paste was reported on by Capanoglu *et al.* [78] who, by using an untargeted LC–Quadropole(Q)–TOF-MS approach highlighted that, for example, insufficient pulverization of the fruit failed to fully destruct the epidermis. This led to the removal of this skin fraction during the filtration step with the concomitant loss of the majority of the skin-resident health beneficial phenolic components. Furthermore, a combined ^{1}H-NMR/OPLS–DA approach was successfully applied to establishing authenticity/provenance of triple concentrated tomato paste, despite the different tomato cultivars and ripening stages employed to obtain the final product [79].

Cereals

It is surprising that given the ubiquity of cereals in the world's diet they have been comparatively poorly served with regard to metabolomic studies, although recent reviews have highlighted the potential in some areas such as rice [80] and the outputs from large EU cereal studies such as HealthGrain [81]. NMR, GC–MS and LC–MS approaches have been used to study changes

in the wheat metabolome accompanying GM events [82,83], environmental influence and climate change in the form of modified CO_2 [84].

Interestingly, metabolomics has been brought smartly into the processed food arena with a combined GC–MS and GC–static headspace solid-phase microextraction (GC–HS-SPME) metabolomic study [6], which established correlations between the raw matrix and cooked volatile components in durum wheat varieties and highlighted the influence of the wheat variety on end-product (pasta) flavor.

Maize has also been the focus of metabolomics, with the targets being the influence of genetic background and growing season [85], developmental stage [86], and environment and agricultural practice [87] on the metabolome. More recently, Skogerson *et al.* [88] applied a GC–TOF-MS approach to tease out the relative contributions of genetic and environment maize grain metabolite diversity and reported that **genotype**-related diversity was significant with the fold-range of genotype mean values of identified metabolites covering the range 1.5- to 93-fold.

Rice is the world's most commonly consumed cereal (and crop). As a result, this key source of energy and nutrition has been subject to much research relating to both yield and grain (nutritional) quality, but less so at the metabolome level. However, Kusano *et al.* [89] used both GC–MS and GC × GC–MS approaches to mine rice metabolome biodiversity and highlight variations that can be exploited via breeding. One such variation rests with the antinutrient phytic acid, which inhibits iron uptake, and this has been addressed via the generation of low phytic acid crop mutants by γ-irradiation of the corresponding wild-type rice (Xiushui 110). Metabolomics of the corresponding rice grown in several environments in China highlighted a significant segregation from the associated wild type due to the polar metabolite profiles driven by methyl pentadecanoate, galactose, raffinose, *myo*-inositol and phosphate, the last two being key components in the phytic acid biosynthetic pathway [90].

Mochida *et al.* combined ^{1}H-NMR metabolomics and restriction fragment length polymorphism (RFLP) markers to assess and correlate biodiversity in a selection of globally diverse rice accessions with limited success [91]. They highlighted that there was no correlation between global metabolic diversity and the phylogenetic relationships (based on RFLP markers), but that localized relationships between shifts and markers were evident and significant.

Fruit

Fleshy fruit offer up a further level of chemical complexity for the metabolomics researcher since the metabolites characterizing fruit quality, including nutrition bioactivity and safety, and hence public purchase and consumption, are significantly diverse. They encompass simple sugars, organic acids, amino acids, carotenoids and simple through to complex polyphenolics, to name but a few of the chemical classes. In addition, the polyphenols themselves are subdivided into anthocyanins, flavonols, (iso)flavones, flavanones, catechins, ellagitannins, cinnamates and hydroxyl benzoic acids and stillbenes [92]. Furthermore, additional complexity can also derive from differential levels of polymerization, glycosylation, methylation and acylation [93].

This metabolic diversity is matched by a broad dynamic range with the total anthocyanin content in fruit species covering the range from 0 mg/g in some fruit to 2–10 mg/g fresh weight in blackcurrant, raspberry, blueberry and the lesser researched fruits such as choke berry and elderberry [93]. Similarly, other metabolites responsible for organolepsis (sugars and organic acids), nutrition (vitamins C and A) and putative bioactive components (flavonoids and ellagitannins) also display similar levels of variation [208–210].

These confounding factors have meant that although there is a proliferation of targeted analytical studies, there have only been a few true fruit metabolomic studies, (i.e., an untargeted study of metabolite changes either by GC–MS, LC–MS or NMR), and these have been limited to comparatively few fruits such as apple [29,94], melon [95], raspberry [96,97], strawberry [98,99] and grape [100–103]. Rudell *et al.* used a combined matrix and volatiles metabolomics approach to study the effects of superficial temperature stress manifestations (scald) [29]. They found that sesquiterpene oxidation products, such as conjugated trienols, 6-methyl-5-hepten-2-one and 6-methyl-5-hepten-2-ol, were presymptomatic of subsequent scald manifestations and spoilage. Rudell *et al.* also took a metabolomic approach to elucidating the impact of prestorage UV–white light irradiation and subsequent cold storage on the apple peel metabolome [94]. They identified that irradiation was attributed with significant changes in diverse primary and secondary metabolic pathways including, for example, those associated with ethylene synthesis, acid metabolism and flavonoid pigment synthesis,

Key Term

Genotype: The genetic makeup, as distinguished from the physical appearance, of an organism or a group of organisms.

Key Terms

Phenotype: The physical characteristics (including chemical makeup) of an organism that result from its genetic code (genotype).

Substantial equivalence: The assessment of novel foods, including foods derived from a GM crop that relies on the use of a food generally recognized as safe as a comparator.

With the aim of determining changes associated with shelf-life and spoilage, Biais *et al.* used a cross-comparative ^{1}H-NMR and GC–TOF-MS approach to establish a spatial variation in primary metabolites in melon and found that, following multivariate statistical analysis, the analytical systems reported on the same primary metabolites and yielded similar metabolite spatial trends [95].

Fait *et al.* reported some groundbreaking work in strawberry development using a combined GC–MS and UPLC–QTOF-MS approach to facilitate primary and secondary metabolite temporal and spatial mapping, including phenylpropanoid derivatives metabolism [98]. This approach was a significant advance beyond previous approaches reported for strawberry, which were largely confined to compound class targeted analysis.

Metabolomics is proving its worth in the area of plant breeding where there is the requirement to simultaneously characterize and quantify multiple chemical **phenotypes** (chemotypes). Since many of the key organoleptic, nutritive and bioactive components in food crops are small, primary and secondary metabolites they fall within the detection limits of the LC–MS- and GC–MS-based metabolomic platforms; as a result, this is being used to analyze the progeny of generated plant populations in greater detail than was thought possible 15 years ago [104]. The significantly large progeny within such populations, allied with the requirement for replication and machine controls and standards, means that further refinement of the existing metabolomics approaches have been undertaken, in particular with respect to soft fruit. Stewart *et al.* [97] and McDougall *et al.* [96] have truncated standard LC–MS to give a short column method that is closer to DI–MS. PCA analysis of the MS-derived data from the same segregating raspberry cross-grown on two distinct environments (one a low-input [fertilizer] site with minimal standard agronomic management and the other one classified as a high health site with standard and regular inputs of fertilizer and agronomic management) showed clear differences in global metabolite changes with year-on-year variation the key driver of metabolite variation. Deeper data mining identified environment-related segregation and within a single environment further subsegregation into chemical classes with, for example, clean segregation between the cyanidin-3-sophoroside and cyanidin-3-rutinoside associated lines. This is extremely informative and means that the short-column DI–MS (SC-DI–MS) approach offers significant promise in facilitating the rapid identification of (screening for) plant progeny showing relatively elevated levels of these compounds, thereby potentially allowing targeted breeding of, for example, cyanidin-3-rutinoside enhanced raspberries. This approach has subsequently been validated as a 'near-quantitative' approach, for (poly)phenolic metabolites at least, by McDougall *et al.* [96], and is currently being expanded, via collaboration with other groups, into strawberry, blueberry and blackcurrant breeding.

Furthermore, the opportunities for using this approach to mine for metabolite pathways, and therefore genes associated with responses to climate change and other key agronomic issues such as sustainability with respect to water and fertilizer, are enormous and an area where metabolomics will surely score significantly in the future.

The next logical step is the correlation of this data with quantitative sensory scoring to allow virtual point-of-harvest scoring at the metabolite level for ultimate end-product sensory and quality attributes. This has been explored for selected fruit – the transition of grape to wine a classic example – with the studies of Hufnagel and Hofmann [105] and Cuadros-Inostroza *et al.* [106] attempting to elucidate and define the quality attributes of wine, with the latter using UPLC–FT-ion cyclone resonance (ICR)–MS to reproducibly classify (or distinguish) attributes such as variety, origin and vintage, as well as underpinning quality criteria. The volatile components have also been a focus with an elegant FT–ICR–MS study by Liger-Belair *et al.* [107] into 'champagne fizz', which identified differences between the bulk and aerosol (fizz) components with monounsaturated fatty acids and isoprenoids such as β-damascenone precursors.

Not surprisingly the opportunities afforded using metabolomic approaches to analyze fruit products have been numerous with exploitation in the areas of provenance authentication and adulteration detection in wine [108–111] and fruit juices [112,113], whilst the more (semi-) solid products such as jams/purees have been probed for authenticity using infrared approaches [114–116], and more recently NMR [117].

Food safety

It has been repeatedly outlined above that there are many factors that impact upon metabolite content and composition, and untoward changes in these, can impact deleteriously on food safety

either in the raw crops material or in the processed food. For example, the current furore surrounding the generation of acrylamide in fried potato and baked cereal products has identified reducing sugars and asparagine as the predominant precursors in the raw material. Ways to reduce these are now being sought [118], with metabolomics an intimate part of the research toolbox since their reduction will undoubtedly have other metabolic consequences. Indeed, metabolomics is now being viewed as a platform technology as part of a risk and safety assessment with the potential to assess, for example, unintended metabolite changes in novel foodstuffs [119]. Readers should note that in this review we will only consider crop- and plant-derived food, but they are urged to read the excellent article by Cevallos-Cevallos *et al.* who succinctly outlines metabolomics in food, encompassing nonplant food and food microbiology [120].

By and large this approach has adopted the **substantial equivalence** paradigm adopted in GM versus non-GM assessment by the Organization for Economic Cooperation and Development (OECD) [121] and Food and Agricultural Organization of the United Nations (FAO)/WHO [122] with the use of existing crops and food with a history of safe use acting as comparators, reference and/or benchmarks against which the novel crop or food (and its metabolome) is judged [123]. Indeed, metabolomics is being viewed by some as displaying the potential to become part of a risk assessors technology toolbox [9].

The use of metabolomics in this application is vital as the ability to measure variation is implicit in safety and risk assessment. Typically for crops, biological variation and the growth or preparation stages are the main source of variation; analytical variation is often minimal in comparison. Growth plots should be randomized and the adequate number of controls interspersed in order to minimize intra- and inter-plant variability. In the case of GM crops the appropriate controls (e.g., near-isogenic or 'empty vector' lines) must be included. Harvesting of plant tissue should ideally occur at the same daily time point and the tissue from all samples should represent an identical developmental stage. Optimal storage and preparation conditions need to be determined at all stages of analysis. Most procedures require extraction of metabolites from the matrix, therefore homogenization must be performed. Incomplete homogenization can be a major cause of variation, thus it is essential that the material is homogenized into a homogeneous solution to minimize intra-sample variation. Consideration of these parameters is essential to the overall metabolomic outputs and must not be ignored.

The application of metabolomics to food safety follows a long history of targeted analysis that has identified metabolites essential for quality and health attributes and those with detrimental actions. Therefore, unbiased semiquantitative methods (such as the various MS-based approaches) that determine numerous compounds may be inappropriate in some instances, but these technologies, in particular the MS-based ones, are becoming increasingly quantitative if accompanied with appropriate calibration. Alternate quantification systems used in tandem with detection systems (e.g., MS or NMR) displaying wide dynamic ranges, such as coulimetric detectors [124], are increasingly gaining acceptance. Indeed, the coulimetric detectors offer other advantages. When used in series before MS detection, they can quantify what may be unknown compounds and in the process oxidize the compound (as part of the detection reaction) generating oxo-compounds that are more amenable to MS ionization and characterization [125]. This approach has only been used in a few cases, but a recent study by Hájek *et al.* [126] of phenolic and flavone natural antioxidants in beer, identified that this combined detection system offers major advantages with respect to limits of quantification and baseline drift.

More commonly, however, when metabolomic analysis is employed in a food safety remit it is comparative rather than quantitative, requiring that the techniques must be able to detect perturbations in metabolites that are unrelated by intuitive biological knowledge related to intended manipulation [127]. For such analysis metabolomic or chemical fingerprinting procedures utilizing GC–MS, NMR, MALDI-TOF-MS and DI–MS have been evaluated. However the quantitative aspect is a definite requirement if the approaches are also to identify metabolite changes along with multiple pesticide residues [30,128].

A range of technologies (^{1}H-NMR, GC–MS, DI–MS and MALDI-TOF-MS) have been used for substantial equivalence of GM varieties of the food crops: tomato [129], potato [62], pea [130] and wheat [82]. As part of this, ^{1}H-NMR, although quantitative, displayed a low sensitivity and specificity with less than 50 metabolites identified

and quantified. Furthermore, NMR can be a routine technology in food safety and testing laboratories [113]. The relative costs, sensitivity thresholds and ability to distinguish multiple components more easily have seen the MS-based technologies becoming increasingly more prevalent.

DI–MS [63] and MALDI-TOF-MS [75] have been used to differentiate between GM and non-GM potato and tomato varieties, respectively. These approaches, used to generate mass ion fingerprints, do suffer some disadvantages via ion suppression if crude extracts are used. However, as a first-pass screen they can be effective and very rapid, particularly DI–MS.

It is safe to say that for substantial equivalence, GC–MS is still considered the 'gold standard' and has been used to differentiate GM tomato [131], potato [63] and wheat [82] varieties from their parent backgrounds and appropriate controls. The compounds identified include sugars, sugar phosphates, organic acids, fatty acids, polyols and some terpenoids, and in total approximately 120 metabolites can be identified in one chromatographic separation. However, often one of the most frustrating aspects is the presence of numerous unknown metabolites. GC–MS 'wins' here since existing mass spectra databases are much more richly populated for GC–MS analysis than the other approaches (LC–MS, MALDI-TOF-MS and FT-MS), thereby offering a greater change of at least tentative structural characterization.

Traditionally, when evaluating metabolomics data generated on novel foods, multivariate PCA is the data analysis method used routinely. Generally, such is the depth and quantity of the data generated via metabolomic approaches (e.g., MS or NMR) that subsequent PCA analysis will segregate the novel food, plant or crop from its comparators even if the variation associated with this segregation is very small (>1%). However, in virtually all the studies targeted at GM plant/crop material, the difference between varieties, the GM(s) and its parent line was small and often indicated that the overall difference in metabolite composition resulting from the intended manipulation was not greater than the transformation process alone, natural variability within the crop [132] or that attributable to the environmental differences [133].

The integration of the metabolomic outputs from food safety assessments are increasingly being viewed from a biological perspective, and with the advent of a number of software solutions such as MAPMAN [134], the changes accompanying the GM event are being assessed at the biochemical pathway level [134,211]. In this way the sectors of metabolism affected can be clearly differentiated, and with the advent of Systems Biology-based approaches [135] the potential exists to integrate different 'omics-based datasets for a given crop variety using the aforementioned MAPMAN platform [134]. This will enable correlation analysis to be performed and eventually a more predictive modeling approach developed for the assessment of substantial equivalence.

Future perspective

The future for metabolomics in crops and food research is extremely positive. The recent paradigm shift in biology research towards systems biology [136] shifts the research foci squarely onto the shoulders of the 'omic technologies, with metabolomics a key part since it characterizes crop metabolic end points, particularly where food is concerned. Such has been the success of metabolomics, and the ease (and relatively low cost) with which plant and crop genomes can be sequenced via next-generation sequencing [137], that it has almost become *de rigueur* that both are undertaken in unison to bridge the genotype-to-phenotype gap [12]. Indeed, the approach is being extended beyond the crop confines into food and human diseases [138], nutrition [139] and bioefficacy in drug discovery [140–142].

This unification of 'omics is becoming more commonplace in research approaches to our most common crops such as potato [143–145], tomato [146–148], and to a lesser degree soft fruits such as raspberry [96–97,149]. Clearly, as this approach and postacquisition data handling becomes more commonplace and utilitarian, the number of crops addressed will undoubtedly broaden.

The world is a changing environment and metabolomics has a significant role to play in addressing current and future problems surrounding crop and food production: safety, (enhanced) nutritive value sustainability, food security and climate change. In fact, several of these issues are being addressed via collaborative international efforts utilizing the whole gamut of metabolomics technologies. More recently, the completed Quantitative Risk Assessment Strategies for Novel Foods (NOFORISK) [212] and Promoting Food Safety Through a New Integrated Risk Analysis Approach for Foods (SAFEFOODS) [213] projects, both EU-FP6-funded projects, had metabolomics (LC–MS, GC–MS and NMR) at their core as high-throughput analytical platforms in

the role as risk assessment processes for novel foods. In addition, the metabolomics platform in SAFEFOODS was extended to assist another FP6 project: QualityLowInputFood [214], with the aim of assessing the impact of organic and 'low-input' agricultural regimes on food safety and nutritive value.

More recently, the Development of High Throughput Approaches to Optimize the Nutritional Value of Crops and Crop-Based Foods (DEVELONUTRI) [215] and Metabolomic Technology Applications for Plants, Health and Outreach (META-PHOR) [216] projects, sister EU-FP6-funded projects, have metabolomics as their primary approach to look at a number of issues in specific crops. META-PHOR focuses on developing innovative metabolite profiling and identification technologies for the detailed characterization of broccoli, rice and melon. Meanwhile, DEVELONUTRI is focused on employing state-of-the-art and emergent metabolomic technologies to potato, tomato and wheat (durum and bread) crop generation, as well as assessment of teh effect that the post-harvest processing chain has on nutritive value and the global metabolite pool. Similarly, in the USA and

Executive summary

The need for metabolomics

- The paradigm shift to systems biology required analytical platforms that report on multiple metabolites simultaneously, and consequentially these have centered on GC–MS, LC–MS and NMR, but other detection technologies such as FT-IR have been applied.
- The adoption and utility of transcriptomics and, to a lesser extent, proteomics in plant and crop science has necessitated the corresponding development of high-throughput and robust metabolite analysis.

Solanaceous crops (e.g., potato & tomato)

- Solanaceous crops have been applied to mine metabolomic diversity, as well as the raw to processed food chain, with the aim of identifying new sources of, and changes in, organoleptic characteristics.
- MS and NMR approaches have been applied to identify planned and any unplanned changes in metabolism following genetic modification.
- These crops have been the subject of studies exploring the potential of emergent metabolomic approaches such as MALDI-TOF-MS and UPLC–MS.
- These crops are the subjects of open-use metabolomic databases, and this is an area ripe for further development.

Cereals

- Application to cereals and associated products has been limited.
- Metabolomics was used to correlate durum wheat metabolite composition to pasta volatile diversity and organoleptic behavior.
- 2D GC × GC–MS was used to mine for phytochemical diversity.

Fruit

- DI–MS metabolomics was applied to fruit breeding populations yielding quality and nutritional data in a fraction of the normal time.
- Combined matrix (LC–MS) and volatile (GC–MS) approaches were combined to elucidate the (bio)chemical consequences of post-harvest storage procedures.

Food safety

- LC–MS, NMR and, in particular, GC–MS metabolomics are being considered as fundamental parts of food safety risk assessment strategies.
- To date, all metabolomic approaches have highlighted that the changes associated with genetic modification are less than those seen when comparing associated non-GM wild and cultivated species.
- The application of coulometric detection both in tandem and parallel to LC–MS are broadening the dynamic range and coverage of the latter.

Future perspective

- The combination of metabolomics with proteomics and/or transcriptomics will yield deeper insights into plant, crop and food production and safety.
- Key to the progression and utility of metabolomics for crops and food will be the expansion of open access LC–MS, GC–MS and NMR databases, which should also include a commonly annotated approach to unknowns, and thereby facilitate the reduction in the metabolite identification bottleneck.
- New methods to visualize metabolomics outputs within a biosynthetic or food process framework are emerging and these will enhance interpretation and utility.
- Correlated mining of metabolomic and full genome sequence data are set to be the next big area in crop and food science.

Canada there are several crop- and food-focused initiatives that have metabolomics at their heart with tomato [217], grape [218,219] and tree fruit [220] exemplars. These projects are only the start of further projected metabolomics efforts in the areas of food security, nutritive value and food safety.

Climate change is increasingly becoming a dominant issue in all matters pertaining to food [150,151] and this is being addressed to some degree by metabolomics in new crop projects, such as the EU Interreg IVb project ClimaFruit (Future Proofing the North Sea Berry Industry [221]). In this project, metabolomics will be used to characterize the impact of specific elements of climate change ([CO_2] and temperature) and sustainability (water and nutrient use efficiency) on fruit development and quality with a view to feeding this back, with a matched functional genomics effort, into breeding programs.

Finally, the field of metabolomics continues to advance at the technical level with many of the companies now providing FT–ICR–MS systems for detection, thereby allowing detection at sub-parts per million levels in conjunction with high mass accuracy and multiple fragmentation characterization. Furthermore, the variability often inherent in biological samples and sampling means that roboticized sample preparation, where extraction and preparation error can be vastly reduced, is one that would benefit from attention, for example, when applied to metabolome analysis in plant and crop breeding.

Metabolomics is on the cusp of realizing its full potential in the plant arena. The plunging costs of next-generation genome sequencing technologies [152] is providing massive amounts of data that, when matched with the associated metabolomic datasets, will allow a true understanding of the translation of gene-to-end point (metabolite) biochemistry to be understood and ultimately exploited.

Financial & competing interests disclosure

The authors acknowledge EU funding for DEVELONUTRI (FOOD-CT-2006-036296), EU-SOL (FOOD-CT-2006-016214), BrainHealthFood (FP7-SME project number 222503), ClimaFruit (Interreg IVb-North Sea Region Program), EU-Berry (EU-FP7), and from The Research Council of Norway (Metabolic Profiling of Rubus), The Scottish Government Rural and Environment Research and Analysis Directorate (RERAD) and The Biotechnology and Biological Sciences Research Council (BBSRC). The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Bibliography

Papers of special note have been highlighted as:
▪ of interest
▪▪ of considerable interest

1 Hemery Y, Rouau X, Lullien-Pellerin V, Barron C, Abecassis J. Dry processes to develop wheat fractions and products with enhanced nutritional quality. *J. Cereal Sci.* 46, 327–347 (2007).

2 Coombe BG. Influence of temperature on composition and quality of grapes. *Acta Hort.* 206, 23–35 (1987).

3 Moretti CL, Mattos LM, Calbo AG, Sargent SA. Climate changes and potential impacts on postharvest quality of fruit and vegetable crops: a review. *Food Res. Int.* 43, 1824–1832 (2010).

4 Morris WL, Shepherd T, Verrall SR, McNicol JW, Taylor MA. Relationships between volatile and nonvolatile metabolites and attributes of processed potato flavor. *Phytochemistry* 71, 1765–1773 (2010).

5 Oruna-Concha MJ, Bakker J, Ames JM. Comparison of the volatile components of two cultivars of potato cooked by boiling, conventional baking and microwave baking. *J. Sci. Food Agric.* 82, 1080–1087 (2002).

6 Beleggia R, Platani C, Spano G, Monteleone M, Cattivelli L. Metabolic profiling and analysis of volatile composition of durum wheat semolina and pasta. *J. Cereal Sci.* 49, 301–309 (2009).

7 Tomás-Barberan FA, Ferreres F, Gil MI. Antioxidant phenolic metabolites from fruit and vegetables and changes during postharvest storage and processing. *Studies Nat. Prod. Chem.* 23, 739–795 (2000).

8 Odriozola-Serrano I, Soliva-Fortuny R, Martín-Belloso O. Effect of minimal processing on bioactive compounds and color attributes of fresh-cut tomatoes. *LWT Food Sci. Technol.* 41, 217–226 (2008).

9 Kuiper HA, Kleter GA, Noteborn HP, Kok EJ. Substantial equivalence – an appropriate paradigm for the safety assessment of genetically modified foods? *Toxicology* 181–182, 427–431 (2002).

10 Cellini F, Chesson A, Colqhoun I *et al.* Unintended effects and their detection in genetically modified crops. *Food Chem. Toxicol.* 42, 1089–1125 (2004).

11 Cevallos-Cevallos JM, Reyes-De-Corcuera JI, Etxeberria E, Danyluk MD, Rodrick GE. Metabolomic analysis in food science: a review. *Trends Food Sci. Tech.* 20, 557–566 (2009).

12 Fiehn O. Metabolomics – the link between genotypes and phenotypes. *Plant Mol. Biol.* 48, 155–171 (2002).

▪▪ **The potential and integration of metabolomics in plant breeding.**

13 Rochfort S, Ezernieks V, Bastian SEP, Downey MO. Sensory attributes of wine influenced by variety and berry shading discriminated by NMR metabolomics. *Food Chem.* 121, 1296–1304 (2010).

14 Almeida C, Duarte IF, Barros A, Rodrigues J, Spraul M, Gil AM. Composition of beer by ^{1}H NMR spectroscopy: effects of brewing site and date of production. *J. Agric. Food Chem.* 54, 700–706 (2006).

15 Rodrigues JA, Barros AS, Carvalho B, Brandão T, Gil AM, Silva Ferreira AC. Evaluation of beer deterioration by gas chromatography–mass spectrometry/multivariate analysis: a rapid tool for assessing beer composition. *J. Chromatogr. A*. 1218, 990–996 (2011).

16 Frank T, Scholz B, Peter S, Engel K-H. Metabolite profiling of barley: influence of the malting process. *Food Chem*. 124, 948–957 (2011).

17 Weckx S, Van der Meulen R, Maes D *et al.* Lactic acid bacteria community dynamics and metabolite production of rye sourdough fermentations share characteristics of wheat and spelt sourdough fermentations. *Food Microbiol*. 27, 1000–1008 (2010).

18 Dixon RA, Gang DR, Charlton AJ *et al.* Applications of metabolomics in agriculture. *J. Agric. Food Chem*. 54, 8984–8994 (2006).

19 Bino RJ, Hall RD, Fiehn O *et al.* Potential of metabolomics as a functional genomics tool. *Trends Plant Sci*. 9, 418–425 (2004).

20 Jenkins H, Hardy N, Beckmann M *et al.* A proposed framework for the description of plant metabolomics experiments and their results. *Nat. Biotechnol*. 22, 1601–1606 (2004).

21 Fiehn O, Kristal B, Ommen B *et al.* Establishing reporting standards for metabolomic and metabonomic studies: a call for participation. *OMICS* 10, 158–163 (2006).

22 Castle LA, Fiehn O, Kaddurah-Daouk R, Lindon JC. Metabolomics standards workshop and the development of international standards for reporting metabolomics experimental results. *Brief. Bioinform*. 7, 159–165 (2006).

23 Morrison N, Bearden D, Bundy J *et al.* Standard reporting requirements for biological samples in metabolomics experiments: environmental context. *Metabolomics* 3, 203–210 (2007).

24 Bais P, Moon SM, He K *et al.* PlantMetabolomics.org: a web portal for plant metabolomics experiments. *Plant Physiol*. 152, 1807–1816 (2010).

25 Fiehn O, Wohlgemuth G, Scholz M *et al.* Quality control for plant metabolomics: reporting MSI-compliant studies. *Plant J*. 53, 691–704. (2008).

26 Kovács K, Fray RG, Tikunov Y *et al.* Effect of tomato pleiotropic ripening mutations on flavor volatile biosynthesis. *Phytochemistry* 70, 1003–1008 (2009).

27 Ali F, Maltese F, Fortes AM, Pais MS, Choi YH, Verpoorte R. Monitoring biochemical changes during grape berry development in Portuguese cultivars by NMR spectroscopy. *Food Chem*. 124, 1760–1769 (2011).

28 Pelayo C, Ebeler SE, Kader AA. Postharvest life and flavor quality of three strawberry cultivars kept at 5°C in air or air + 20 kPa CO_2. *Postharvest Biol. Tec*. 27, 171–183 (2003).

29 Rudell DR, Mattheis JP, Hertog ML. Metabolomic change precedes apple superficial scald symptoms. *J. Agric. Food Chem*. 57, 8459–8466 (2009).

▪ **Good combination of food matrix and volatile metabolomics.**

30 Picó Y, La Farré M, Segarra R, Barceló D. Profiling of compounds and degradation products from the postharvest treatment of pears and apples by ultra-high pressure liquid chromatography quadrupole-time-of-flight mass spectrometry. *Talanta* 81, 281–293 (2010).

31 Wang M, Lamers RJAN, Korthout HAJA *et al.* Metabolomics in the context of systems biology: bridging traditional chinese medicine and molecular pharmacology. *Phytother. Res*. 19, 173–182 (2005).

32 Morgenthal K, Wienkoop S, Scholz M, Selbig J, Weckwerth W. Correlative GC–TOF-MS based metabolite profiling and LC–MS based protein profiling reveal time-related systemic regulation of metabolite-protein networks and improve pattern recognition for multiple biomarker selection. *Metabolomics* 1, 109–121 (2005).

33 Müller-Linow M, Weckwerth W, Hütt MT. Consistency analysis of metabolic correlation networks. *BMC Syst. Biol.* 1(44), (2007).

34 Dunn WB, Ellis D. Metabolomics: current analytical platforms and methodologies. *Trends Analyt Chem*. 24, 285–294 (2005).

35 Fukusaki E, Kobayashi A. Plant metabolomics: potential for practical operation. *J. Biosci. Bioeng.* 100, 347–354 (2005).

36 Seger C, Sturm S. Analytical aspects of plant metabolite profiling platforms: current standings and future aims. *J. Proteome Res*. 6, 480–497 (2007).

37 Shepherd T, Dobson G, Verrall SR *et al.* Potato metabolomics by GC–MS: what are the limiting factors? *Metabolomics* 3, 475–488 (2007).

38 Dunn WB. Current trends and future requirements for the mass spectrometric investigation of microbial, mammalian and plant metabolomes. *Physical Biol*. 5, 011001 (2008).

39 Dunn WB, Broadhurst DI, Atherton HJ, Goddacre R, Griffin JL. Systems level studies of mammalian metabolomes: the roles of mass spectrometry and nuclear magnetic resonance spectroscopy. *Chem. Soc. Rev*. 40(1), 387–426 (2010).

40 Zhang S, Nagana Gowda GA, Ye T, Raferty D. Advances in NMR-based biofluid analysis and metabolite profiling. *Analyst* 135, 1490–1498 (2010).

41 Palama TL, Menard P, Fock I *et al.* Shoot differentiation from protocorm callus cultures of *Vanilla planifolia* (Orchidaceae): proteomic and metabolic responses at early stage. *BMC Plant Biol*. 10, 82 (2010).

42 Ludwig C, Viant MR. Two-dimensional J-resolved NMR spectroscopy: review of a key methodology in the metabolomics toolbox. *Phytochem. Anal*. 21, 22–32 (2010).

43 Chikayama E, Suto M, Nishihara T, Shinozaki K, Hirayama T, Kikuchi J. Systematic NMR analysis of stable isotope labeled metabolite mixtures in plant and animal systems: coarse grained views of metabolic pathways. *PLoS One* 3, e3805 (2008).

44 Eisenreich W, Bacher A. Advances of high-resolution NMR techniques in the structural and metabolic analysis of plant biochemistry. *Phytochemistry* 68, 2799–2815 (2007).

45 Frenzel T, Miller A, Engel K-H. A methodology for automated comparative analysis of metabolite profiling data. *Eur. Food Res. Technol.* 216, 335–342 (2003).

46 Vorst O, de Vos CHR, Lommen A *et al.* A nondirected approach to the differential analysis of multiple LC–MS-derived metabolic profiles. *Metabolomics* 1, 169–180 (2005).

47 De Vos RCH, Moco S, Lommen A, Keurentjes JJB, Bino RJ, Hall RD. Untargeted large-scale plant metabolomics using liquid chromatography coupled to mass spectrometry. *Nat. Protoc*. 2, 778–791 (2007).

48 Werner E, Heilier JF, Ducruix C, Ezan E, Junot C, Tabet JC. Mass spectrometry for the identification of the discriminating signals from metabolomics: current status and future trends. *J. Chromatogr. B* 871, 143–163 (2008).

49 Enot DP, Beckmann M, Overy D, Draper J. Predicting interpretability of metabolome models based on behavior, putative identity, and biological relevance of explanatory signals. *Proc. Natl Acad. Sci. USA* 103, 14865–14870 (2006).

50 Beckmann M, Enot DP, Overy D, Draper JH. Representation, comparison and interpretation of metabolome fingerprint data for total composition analysis and quality trait investigation in potato cultivars. *J. Agric. Food Chem.* 55, 3444–3451 (2007).

51 Goodacre R, Vaidyanathan S, Dunn WB, Harrigan GG, Kell DB. Metabolomics by numbers: acquiring and understanding global metabolite data. *Trends Biotechnol*. 22, 245–252 (2004).

52 Holmes E, Antti H. Chemometric contributions to the evolution of metabonomics: mathematical solutions to characterizing and interpreting complex biological NMR spectra. *Analyst* 127, 1549–1157 (2002).

53 van der Greef J, Smilde AK. Symbiosis of chemometrics and metabolomics: past, present, and future. *J. Chemometr*. 19, 376–386 (2005).

54 Jansen JJ, Smit S, Hoefsloot HC, Smilde AK. The photographer and the greenhouse: how to analyze plant metabolomics data. *Phytochem. Anal*. 21, 48–60 (2010).

▪▪ **Excellent and detailed description of the main concepts in multivariate data analysis and their application in plant and crop science.**

55 *Concepts in Plant Metabolomics*. Nikolau BJ, Wurtele ES (Eds). Springer, The Netherlands (2007).

56 Hall RD. Plant metabolomics in a nutshell: potential and future challenges. In: *Annual Plant Reviews, Volume 43, Biology of Plant Metabolomics*. Hall RD (Ed.). Wiley, UK 1–24 (2011).

57 Davies HV. Metabolomics: applications in functional biodiversity analysis in potato. Presented at: *The International Solanaceae Conference and Solanaceae Genomics Network.* Madison, WI, USA, 23–27 July 2006.

58 Dobson G, Shepherd T, Verrall SR *et al.* Phytochemical diversity in tubers of potato cultivars and landraces using a GC–MS metabolomics approach. *J. Agric. Food Chem.* 56, 10280–10291 (2008).

59 Dobson G, Shepherd T, Verrall SR *et al.* A metabolomics study of cultivated potato (*Solanum tuberosum*) groups Andigena, Phureja, Stenotomum, and Tuberosum using gas chromatography–mass spectrometry. *J. Agric. Food Chem*. 58, 1214–1223 (2010).

60 McCann LC, Bethke PC, Simon PW. Extensive variation in fried chip color and tuber composition in cold-stored tubers of wild potato (*Solanum*) germplasm. *J. Agric. Food Chem.* 58, 2368–2376 (2010).

61 Roessner U, Luedemann A, Brust D *et al.* Metabolic profiling allows comprehensive phenotyping of genetically or environmentally modified plant systems. *Plant Cell* 13, 11–29 (2001).

62 Defernez M, Gunning YM, Parr AJ, Shepherd LVT, Davies HV, Colquhoun IJ. NMR and HPLC–UV profiling of potatoes with genetic modifications to metabolic pathways. *J. Agric. Food Chem.* 52, 6075–6085 (2004).

63 Catchpole GS, Beckmann M, Enot DP *et al.* Hierarchical metabolomics demonstrates substantial compositional similarity between genetically modified and conventional potato crops. *Proc. Natl Acad. Sci. USA* 102, 14458–14462 (2005).

▪ **Good example of the utility of multivariate statistical analysis in plant metabolomics.**

64 Parr A, Mellon F, Colquhoun I, Davies HV. Dihydrocaffeoyl polyamines (kukoamine and allies) in potato (*Solanum tuberosum*) tubers detected during metabolite profiling. *J. Agric. Food Chem.* 53, 5461–5466 (2005).

65 Fernie AE, Schauer N. Metabolomics-assisted breeding: a viable option for crop improvement? *Trends Genet.* 25, 39–48 (2009).

▪▪ **Outstanding review of the potential and integration of metabolomics in plant breeding.**

66 Osorio S, Tohge T, Fernie AR. Application of metabolomic profiling for identifying valuable traits in tomato. *CAB Reviews* 4, 024 (2009).

67 van den Heuvel T, Renes RJ, Gremmen B, van Woerkum C, van Trijp H. Consumers' images regarding genomics as a tomato breeding technology: 'maybe it can provide a more tasty tomato'. *Euphytica* 159, 207–216 (2008).

68 Butelli E, Titta L, Georgio M *et al.* Induced anthocyanin biosynthesis in purple fruit with enhanced antioxidant, dietary and health-inducing properties. *Nat. Biotechnol.* 26, 1301–1308 (2008).

69 Deborde C, Maucourt M, Baldet P *et al.* Proton NMR quantitative profiling for quality assessment of greenhouse-grown tomato fruit. *Metabolomics* 5, 183–198 (2009).

70 Moco S, Bino RJ, Vorst O *et al.* A liquid chromatography–mass spectrometry-based metabolome database for tomato. *Plant Physiol.* 141, 1205–1218 (2006).

71 Fraser PD, Enfissi EMA, Halket JM *et al.* Manipulation of phytoene levels in tomato fruit: effects on isoprenoids, plastids, and intermediary metabolism. *Plant Cell* 19, 3194–3211 (2007).

72 Fraser PD, Enfissi EMA, Bramley PM. Genetic engineering of carotenoid formation in tomato fruit and the potential application of systems and synthetic biology approaches. *Arch. Biochem. Biophys.* 483, 196–204 (2009).

73 Mintz-Oron S, Mandel T, Rogachev I *et al.* Gene expression and metabolism in tomato fruit surface tissues. *Plant Physiol.* 147, 823–851 (2008).

74 Bino RJ, de Vos RCH, Lieberman M *et al.* The light-hyperresponsive high pigment-2dg mutation of tomato: alterations in the fruit metabolome. *New Phytol*. 166, 427–438 (2005).

75 Fraser PD, Enfissi EMA, Goodfellow M, Eguchi T, Bramley PM. Metabolite profiling of plant carotenoids using the matrix-assisted laser desorption ionization time-of-flight mass spectrometry. *Plant J.* 49, 552–564 (2007).

▪ **Good example of an alternate analytical approach to crop and food metabolites.**

76 Pérez EMS, Iglesias MJ, Ortiz FL, Perez IS, Galera MM. Study of the suitability of HRMAS NMR for metabolic profiling of tomatoes: application to tissue differentiation and fruit ripening. *Food Chem*. 122, 877–887 (2010).

77 Capanoglu E, Beekwilder J, Boyacioglu D, De Vos RCH, Hall RD. The effect of industrial food processing on potentially health-beneficial tomato antioxidants. *Crit. Rev. Food Chem.* 50, 919–930 (2010).

78 Capanoglu E, Beekwilder J, Boyacioglu D, Hall RD, De Vos R. Changes in antioxidant and metabolite profiles during production of tomato paste. *J. Agric. Food Chem.* 56, 964–973 (2008).

79 Consonni R, Cagliani LR, Stocchero M, Porretta S. Triple concentrated tomato paste: discrimination between Italian and Chinese products. *J. Agric. Food Chem*. 57, 4506–4513 (2009).

80 Fitzgerald MA, McCouch SR, Hall RD. Not just a grain of rice: the quest for quality. *Trends Plant Sci*. 14, 133–139 (2009).

81 Shewry PR, Ward JL. Future prospects for the analysis of bioactive components in cereal grain. *Cereal Foods World* 55, 71–75 (2010).

82 Baker JM, Hawkins ND, Ward JL *et al* A metabolomic study of substantial equivalence of field-grown genetically modified wheat. *Plant Biotech. J.* 4, 381–392 (2006).

83 Stamova BS, Roessner U, Suren S, Laudencia-Chingcuanco D, Bacic A, Beckles DM. Metabolic profiling of transgenic wheat over-expressing the high molecular-weight Dx5 glutenin subunit. *Metabolomics* 5, 239–252 (2009).

84 Levine LH, Kasahara H, Kopka J *et al.* Physiologic and metabolic responses of wheat seedlings to elevated and super-elevated carbon dioxide. *Adv. Space Res.* 42, 1917–1928 (2008).

85 Reynolds TL, Nemeth MA, Glenn KC, Ridley WP, Astwood JD. Natural variability of metabolites in maize grain: differences due to genetic background. *J. Agric. Food Chem.* 53, 10061–10067 (2005).

86 Seebauer JR, Moose SP, Fabbri BJ, Crossland LD, Below FE. Amino acid metabolism in maize earshoots: implications for assimilate preconditioning and nitrogen signalling. *Plant Physiol.* 136, 4326–4334 (2004).

87 Harrigan G, Stork LG, Riordan SG *et al.* Impact of genetics and environment on nutritional and metabolite components of maize grain. *J. Agric. Food Chem.* 55, 6177–6185 (2007).

88 Skogerson K, Harrigan GG, Reynolds TL *et al.* Impact of genetics and environment on the metabolite composition of maize grain. *J. Agric. Food Chem*. 58, 3600–3610 (2010).

89 Kusano M, Fukushima A, Kobayashi M *et al.* Application of a metabolomic method combining one-dimensional and two-dimensional gas chromatography–time-of-flight/mass spectrometry to metabolic phenotyping of natural variants in rice. *J. Chromatogr. B* 855, 71–79 (2007).

▪ **Sophisticated application of 1D and 2D metabolomics.**

90 Frank T, Meuleye Seumo B, Miller A, Shu QY, Engel K-H. Metabolite profiling of two low phytic acid (*lpa*) rice mutants. *J. Agric. Food Chem.* 55, 11011–11019 (2007).

91 Mochida K, Furuta T, Ebana K, Shinozaki K, Kikuchi J. Correlation exploration of metabolic and genomic diversity in rice. *BMC Genomics* 10, 568 (2009).

92 Pietta P, Minoggio M, Bramati L. Plant polyphenols: structure, occurrence and bioactivity. *Studies Nat. Prod. Chem.* 28, 257–312 (2003).

93 Clifford MN. Anthocyanins – nature, occurrence and dietary burden. *J. Sci. Food Agric.* 80, 1063–1072 (2000).

94 Rudell DR, Mattheis JP, Curry EA. Prestorage ultraviolet-white light irradiation alters apple peel metabolome. *J. Agric. Food Chem*. 56, 1138–1147 (2008).

95 Biais B, Allwood JW, Deborde C *et al.* ^{1}H NMR, GC–EI-TOF-MS, and data set correlation for fruit metabolomics: application to spatial metabolite analysis in melon. *Anal. Chem.* 81, 2884–2894 (2009).

96 McDougall G, Martinussen I, Stewart D. Towards fruitful metabolomics: high throughput analyses of polyphenol composition in berries using direct infusion mass spectrometry. *J. Chromatogr. B* 871, 362–369 (2008).

97 Stewart D, McDougall GJ, Sungurtas J, Verrall S, Graham J, Martinussen I. Metabolomic approach to identifying bioactive compounds in berries: advances toward fruit nutritional enhancement. *Mol. Nutrit. Food Res.* 51, 645–651 (2007).

98 Fait A, Hanhineva K, Beleggia R *et al.* Reconfiguration of the achene and receptacle metabolic networks during strawberry fruit development. *Plant Physiol.* 148, 730–750 (2008).

99 Zhang J, Wang X, Yu O *et al.* Metabolic profiling of strawberry (*Fragaria x ananassa Duch.*) during fruit development and maturation. *J. Exp. Bot*. 62(3), 1103–1118 (2011).

100 Figueiredo A, Fortes AM, Ferreira S *et al.* Transcriptional and metabolic profiling of grape (*Vitis vinifera L.*) leaves unravel possible innate resistance against pathogenic fungi. *J. Exp. Bot*. 59, 3371–3381 (2008).

101 Grimplet J, Cramer GR, Dickerson JA, Mathiason K, Van Hemert J, Fennell AY. VitisNet: 'omics' integration through grapevine molecular networks. *PLoS One* 4, 8365 (2009).

102 Ali K, Maltese F, Choi YH, Verpoorte R. Metabolic constituents of grapevine and grape-derived products. *Phytochem. Rev.* 9, 357–378 (2010).

103 Pereira GE, Gaudillere JP, Pieri P *et al.* Microclimate influence on mineral and metabolic profiles of grape berries. *J. Agric. Food Chem*. 54, 6765–6775 (2006).

104 Steinfath M, Strehmel N, Peters R. Discovering plant metabolic biomarkers for phenotype prediction using an untargeted approach. *Plant Biotechnol. J.* 8, 900–911 (2010).

105 Hufnagel JC, Hofmann T. Quantitative reconstruction of the nonvolatile sensometabolome of a red wine. *J. Agric. Food Chem*. 56, 9190–9199 (2008).

106 Cuadros-Inostroza A, Giavalisco P, Hummel J, Eckardt A, Willmitzer L, Pena-Cortés H. Discrimination of wine attributes by metabolome analysis. *Anal. Chem*. 82, 3573–3580 (2010).

107 Liger-Belair G, Cilindre C, Gougeon RD *et al.* Unraveling different chemical fingerprints between a champagne wine and its aerosols. *Proc. Natl Acad. Sci. USA* 106, 16545–16549 (2009).

▪ **interesting example of applying high-end metabolomics to a practical food scenario.**

108 Ogrinc N, Kosir IJ, Spangenberg JE, Kidric J. The application of NMR and MS methods for detection of adulteration of wine, fruit juices and olive oil. A review. *Anal. Bioanal. Chem*. 376, 424–430 (2003).

109 Hong-Son S, Ki MK, Van den Berg F. ^{1}H nuclear magnetic resonance-based metabolomic characterization of wines and grapes by varieties and production areas. *J. Agric. Food Chem*. 56, 8007–8016 (2008).

110 Hong-Son S, Geum-Sook H, Ki MK *et al.* Metabolomic studies on geographical grapes and their wines using ^{1}H NMR analysis coupled with multivariate statistics. *J. Agric. Food Chem*. 57, 1481–1490 (2009).

111 Rochfort S, Ezernieks V, Bastian SEP, Downey MO. Sensory attributes of wine influenced by variety and berry shading discriminated by NMR metabolomics *Food Chem*. 121, 1296–1304 (2010).

112 Cuny M, Vigneau E, Le Gall G, Colquhoun I, Lees M, Rutledge DN. Fruit juice authentication by ^{1}H NMR spectroscopy in combination with different chemometrics tools. *Anal. Bioanal. Chem*. 390, 419–427 (2008).

113 Spraul M, Schütz B, Humpfer E *et al.* Mixture analysis by NMR as applied to fruit juice quality control. *Magn. Reson. Chem*. 47, S130–S137 (2009).

114 Wilson RH, Slack PT, Appleton GP, Sun L, Belton PS. Determination of the fruit content of jam using Fourier transform infrared spectroscopy. *Food Chem.* 47, 303–308 (1993).

115 Defernez M, Katherine EK, Wilson RK. Use of infrared spectroscopy and chemometrics for the authentication of fruit purees. *J. Agric. Food Chem*. 43, 109–111 (1995).

116 Kemsley EK, Holland JK, Wilson RH. Detection of adulteration of raspberry purées using infrared spectroscopy and chemometrics. *J. Agric. Food Chem.* 44, 3864–3870. (1996).

117 Schütz B, Humpfer E, Mörtter A, Schäfer H, Spraul M. SGF Profiling™ – NMR based fruit juice quality control. *Bruker BioSpin Technical Report* 10/08 T12224 (2008).

118 Halford NG, Muttucumaru N, Curtis TY *et al.* Genetic and agronomic approaches to decreasing acrylamide precursors in crop plants. *Food Addit. Contam*. 24(Suppl. 1), 26–36 (2007).

119 Rischer H, Oksman-Caldentey K-M. Unintended effects in genetically modified crops: revealed by metabolomics? *Trends Biotechnol*. 24, 102–104 (2006).

120 Cevallos-Cevallos JM, Reyes-De-Corcuera JI, Etxeberria E. Metabolomic analysis in food science: a review. *Trends Food Sci. Tech*. 20, 557–566 (2009).

121 Organization for Economic Cooperation and Development. *Report of the OECD Workshop on the Toxicological and Nutritional Testing of Novel Foods.* Aussois, France, 5–8 March 1997.

122 Food and Agriculture Organization of the United Nations, Rome (2000). *Safety Aspects of Genetically Modified Foods of Plant Origin. Report of a Joint FAO/WHO Expert Consultation on Foods Derived from Biotechnology.* Geneva, Switzerland, 29 May–2 June 2000.

123 Perry JP, Ter-Braak CJF, Dixon PM *et al.* Statistical aspects of environmental risk assessment of GM plants for effects on nontarget organisms. *Environ. Biosafety Res.* 8, 65–78 (2009).

124 Ferruzzi MG, Sander LC, Rock CL, Schwartz SJ. Carotenoid determination in biological microsamples using liquid chromatography with a coulometric electrochemical array detector. *Anal. Biochem.* 256, 74–81 (1998).

125 Gamache PH, Meyer DF, Granger MC. Metabolomic applications of electrochemistry/mass spectrometry. *J. Amer. Soc. Mass Spectrom.* 15, 1717–1726 (2004).

126 Hájek T, Skeríková V, Cesla P, Vynuchalová K, Jandera P. Multidimensional LC × LC analysis of phenolic and flavone natural antioxidants with UV-electrochemical coulometric and MS detection. *J. Sep. Sci.* 31, 3309–3328 (2008).

127 Beale MH, Ward JL, Baker JM. Establishing substantial equivalence: metabolomics. *Methods Mol. Biol.* 478, 289–303 (2009).

128 Young SK, Kim JH, Li QX. Metabolomics in pesticide toxicology. In: *Hayes' Handbook of Pesticide Toxicology (Third Edition)*. Krieger R (Ed.). Academic Press, Maryland Heights, USA 627–643 (2010).

129 Le Gall G, Colquhoun IJ, Davis AL, Collins GJ, Verhoeyen ME. Metabolite profiling of tomato (Lycopersicon esculentum) using ^{1}H NMR spectroscopy as a tool to detect potential unintended effects following a genetic modification. *J. Agric. Food Chem.* 51, 2447–2456 (2003).

130 Charlton A, Allnutt T, Holmes S *et al.* NMR profiling of transgenic peas. *Plant Biotechnol. J.* 2, 27–35 (2004).

131 Roessner-Tunali U, Hegemann B, Lytovchenko A *et al.* Metabolic profiling of transgenic tomato plants overexpressing hexokinase reveals that the influence of hexose phosphorylation diminishes during fruit development. *Plant Physiol.* 133, 84–99 (2003).

132 Cellini F, Chesson A, Colquhoun I *et al.* Unintended effects and their detection in genetically modified crops. *Food Chem. Toxicol.* 42, 1089–1125 (2004).

133 Barros E, Lezar S, Anttonen MJ *et al.* Comparison of two GM maize varieties with a near-isogenic non-GM variety using transcriptomics, proteomics and metabolomics. *Plant Biotechnol. J.* 8, 436–451 (2010).

134 Thimm O, Bläsing O, Gibon Y *et al.* MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. *Plant J.* 37, 914–939 (2004).

▪ **Excellent interface for visualizing metabolomic changes in a biological framework.**

135 Yin X, Struik PC. Modeling the crop: from system dynamics to systems biology. *J. Exp. Bot.* 61, 2171–2183 (2010).

136 Appleby N, Edwards D, Batley J. New technologies for ultra-high throughput genotyping in plants. *Methods Mol. Biol.* 513, 19–39 (2009).

137 Varshney RK, Nayak SN, May GD, Jackson SA. Next-generation sequencing technologies and their implications for crop genetics and breeding. *Trends Biotechnol.* 27, 522–530 (2009).

138 Madsen R, Lundstedt T, Trygg J. Chemometrics in metabolomics – a review in human disease diagnosis. *Anal. Chim. Acta.* 659, 23–33 (2010).

139 Hall RD, Brouwer ID, Fitzgerald MA. Plant metabolomics and its potential application for human nutrition. *Physiol. Plantarum.* 132, 162–175 (2008).

140 van der Kooy F, Maltese F, Choi YH, Kim HK, Verpoorte R. Quality control of herbal material and phytopharmaceuticals with MS and NMR based metabolic fingerprinting. *Planta Med.* 75, 763–775 (2009).

141 Lan K, Jia W. An integrated metabolomics and pharmacokinetics strategy for multicomponent drugs evaluation. *Curr. Drug Metab.* 11, 105–111 (2010).

142 Yuliana ND, Khatib A, Choi YH, Verpoorte R. Metabolomics for bioactivity assessment of natural products. *Phytother. Res.* 25, 157–166 (2011).

143 Lehesranta SJ, Davies HV, Shepherd LVT *et al.* Proteomic analysis of the potato tuber life cycle. *Proteomics* 6, 6042–6052 (2006).

144 van Dijk JP, Cankar K, Scheffer SJ *et al.* Transcriptome analysis of potato tubers-effects of different agricultural practices. *J. Agric. Food Chem.* 25, 1612–1623 (2009).

145 Shepherd LVT, Alexander CA, Sungurtas JA, McNicol JW, Stewart D, Davies HV. Metabolomic analysis of the potato tuber life cycle. *Metabolomics* 6, 274–291 (2010).

146 Matsukura C, Aoki K, Fukuda N *et al.* Comprehensive resources for tomato functional genomics based on the miniature model tomato micro-tom. *Curr. Genomics.* 9, 436–443 (2008).

147 Barone A, Di Matteo A, Carputo D, Frusciante L. High-throughput genomics enhances tomato breeding efficiency. *Curr. Genomics* 10, 1–9 (2009).

148 Gavai AK, Tikunov Y, Ursem R *et al.* Constraint-based probabilistic learning of metabolic pathways from tomato volatiles. *Metabolomics* 5, 419–428 (2009).

149 Graham J, Hackett CA, Smith K, Woodhead M, Hein I, McCallum. Mapping QTLs for developmental traits in raspberry from bud break to ripe fruit. *Theoret. Appl. Gen.* 118, 1143–1155 (2009).

150 Beddington J. Food security: contributions from science to a new and greener revolution. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 365, 61–71 (2010).

151 Godfray HC, Beddington JR, Crute IR *et al.* Food security: the challenge of feeding 9 billion people. *Science* 327, 812–818 (2010).

152 Rounsley SD, Last RL. Shotguns and SNPs: how fast and cheap sequencing is revolutionizing plant biology. *Plant J.* 6, 922–927 (2010).

■ Websites

201 A searchable compounds spectral database. http://webbook.nist.gov (Accessed 20/02/11).

202 660,000 high-resolution spectra. http://eu.wiley.com/WileyCDA/Section/id-403828.html (Accessed 27/07/10).

203 Freely available and sophisticated software for GC–MS data interpretation from the National Institute of Standards and Technology. www.amdis.net (Accessed 27/07/10).

204 A software program for the preprocessing and comparison of full scan nominal or accurate mass LC–MS and GC–MS data. www.metAlign.nl (Accessed 27/07/10).

205 The Commonwealth Potato Collection. http://germinate.scri.sari.ac.uk/germinate_cpc/app/index.pl (Accessed 27/07/10).

206 A set of target tomato metabolites collected from public domains. http://ted.bti.cornell.edu/cgi-bin/TFGD/metabolite/home.cgi (Accessed 07/12/2010).

207 A visualization and analysis package that allows one to identify changed pathways and enriched gene ontology terms from gene expression and/or metabolite profile data. http://bioinfo.bti.cornell.edu/cgi-bin/MetGenMAP/home.cgi (Accessed 07/12/2010).

208 USDA database for the flavonoid content of selected foods. www.nal.usda.gov/fnic/foodcomp/Data/Flav/flav.pdf (Accessed 07/12/2010).

209 USDA compositional database. http://www.nal.usda.gov/fnic/foodcomp/Data (Accessed 07/12/2010).

210 Fineli®. Finnish food composition database. www.fineli.fi (Accessed 17/05/2010).

211 A software tool for the analysis of large metabolic datasets. www.Biosynlab.com (Accessed 10/10/2010).

212 Description of the EU project NOFORISK – qualitative risk assessment strategies for novel foods. www.scri.ac.uk/research/ppfq/foodquality/foodsafety/noforisk (Accessed 22/02/2011).

213 Project website for the EU FP6 project SAFEOODS – promoting food safety through a new integrated risk analysis approach for foods. The application of 'omics was an element to be explored in a food safety perspective. www.safefoods.nl (Accessed 22/02/2011).

214 Project website for the EU FP6 project QLIF – integrated project qualitylowinputfood. 'Omics were used selectively to assess the impact of organic versus conventional agricultural approaches. www.qlif.org (Accessed 22/02/2011).

215 Project website for the EU FP6 project DEVELONUTRI – development of high throughput approaches to optimize the nutritional value of crops and crop-based foods. Metabolomics was explored as a viable way to describe changes in the food chain for potato, tomato and wheat crops-to-products. www.develonutri.info (Accessed 22/02/2011).

216 Project website for the EU FP6 project Meta-Phor – metabolomics for plants health and outreach. The project used metabolomics to generate data on broccoli, melon and rice with the aim of feeding this back to breeding programs for nutritional enhancement. www.meta-phor.eu (Accessed 22/02/2011).

217 Functional genomics of tomato fruit quality: bridging the gap between QTLs and gene, NSF (2010–2012). www.nsf.gov/awardsearch/showAward.do?AwardNumber=0923312 (Accessed 21/02/2011).

218 Grape metabolomic profiles and natural defenses against biotic stress, USDA project number: 5302–22000–22008–016 (2006–2011). www.ars.usda.gov/research/projects/projects.htm?ACCN_NO=410344 (Accessed 22/02/2011).

219 Grape and wine genomics, genome Canada (2009–2012). www.genomecanada.ca/medias/PDF/EN/Grape-and-Wine-Genomics.pdf) (Accessed 21/02/2011).

220 Genetic, metabolic and physiological factors regulating deciduous tree fruit quality USDA project number: 5350–43000–43006–000 (2010–2015). www.ars.usda.gov/research/projects/projects.htm?ACCN_NO=419848) (Accessed 22/02/2011).

221 Project website for the EU Interreg IVb project ClimaFruit – future proofing the North Sea berry fruit industry in times of climate change. www.climafruit.com (Accessed 22/02/2011).

222 The Golm Metabolome Database. http://csbdb.mpimp-golm.mpg.de/csbdb/gmd/gmd.html

223 RIKEN MSn spectral database for phytochemicals. http://spectra.psc.riken.jp/

224 MS/MS spectral tag (MS2T) viewer. http://prime.psc.riken.jp/lcms/ms2tview/ms2tview.html

225 Metabolomics Fiehn laboratory. http://fiehnlab.ucdavis.edu/Metabolite-Library-2007

226 TrichOME: a comparative omics database for plant Trichome. MS experiments for metabolite profiles. www.planttrichome.org/trichomedb/metabolite_profile.jsp

227 Applied Bioinformatics, Plant Research International, Wageningen UR. http://appliedbioinformatics.wur.nl/moto

228 Mass Bank. High resolution MS database. www.massbank.jp

229 Scrips centre for metabolomics. METLIN: metabolite and MS/MS database. http://metlin.scripps.edu

230 Kyoto Encyclopedia of Genes and Genomes. www.genome.jp/kegg

231 National Institute of Advanced Industrial Science and Technology, Japan. Spectral Database for Organic Compounds. http://riodb01.ibase.aist.go.jp/sdbs/cgi-bin/direct_frame_top.cgi