Annual Plant Reviews (2011) **43**, 181–217 http://onlinelibrary.wiley.com
doi: 10.1002/9781405199544.ch7

Chapter 7

CROPS AND TASTY, NUTRITIOUS FOOD – HOW CAN METABOLOMICS HELP?

Derek Stewart[1], Louise V.T. Shepherd[1], Robert D. Hall[2,3,4] and Paul D. Fraser[5]

[1] *Scottish Crop Research Institute, Plant Products and Food Quality Programme, Mylnefield, Invergowrie Dundee, DD2 5DA Scotland, UK*
[2] *Plant Research International, Wageningen University and Research Centre (Wageningen-UR), PO Box 16, 6700 AA Wageningen, The Netherlands*
[3] *Centre for BioSystems Genomics, PO Box 98, 6700 AB Wageningen, The Netherlands*
[4] *Netherlands Metabolomics Centre, Einsteinweg 55, 2333 CC Leiden, The Netherlands*
[5] *School of Biological Sciences, Royal Holloway University of London, Egham, Surrey, TW20 0EX, UK*

Abstract: Food quality, security and safety have become topics of considerable recent interest. With a rapidly growing world population, entailing an ever-expanding requirement for food, and with the global consumer making higher and better-informed demands on our crop and food producers, much attention is being given to how we can meet all these growing needs. Crop growers and food processors alike are already looking at state-of-the-art technologies such as metabolomics as a source of new inroads into the generation of detailed knowledge on the biochemical composition of crop products and how they change during transport, storage and industrial processing. All the steps in the food production chain, from the moment the plant breeder makes the first cross to the canning factory delivering to the supermarket distributor, have potentially significant influence on the quality of the final product the consumer places on the kitchen table. Metabolomics of crop plants and plant products is already being widely applied and is generating new information on the complexity of food composition. In this chapter, some of the main subjects of recent research, concerning both fresh and processed materials, are covered. Emphasis has been laid on technological

Annual Plant Reviews Volume 43, Biology of Plant Metabolomics, First Edition, edited by Robert Hall
© 2011 Blackwell Publishing Ltd.

c07 BLBK354-Hall October 29, 2010 21:16 Trim: 234mm×156mm Series: APR Char Count=

advances and how we approach food improvement strategies as well as the role of metabolomics is also finding a role in food safety analysis.

Au: Please check the sense in text "as well as the role of metabolomics is also finding a role in food safety analysis". Should it be changed to read "as well as the role of metabolomics in food safety analysis"? Please confirm.

Keywords: crop plants; processed foods; genetically modified organisms; substantial equivalence; food quality

7.1 Every food chain begins with plants

Food is fundamental to our existence and survival. Fresh food products are important sources of essential nutrients, vitamins, etc. and indeed, are generally experienced as being pleasurable to eat. The global population, and especially those living in the Western world, is also eating increasing amounts of processed food materials. In fact, even much of the 'fresh' food we eat has undergone some kind of limited processing in terms of cutting, packaging, reduced oxygen storage, etc. (Hall, 2006a; Hall *et al.*, 2008, 2010). There is now an increasing demand both from industry and consumer alike, for improved knowledge about the nutritional quality of these foods and how we might design new strategies for enhancing/maintaining food quality to meet current expectations. Clearly there is a broad opportunity here for metabolomics to play a role. Indeed, there are already many examples of different metabolomics approaches having been successfully employed to extend our knowledge of the biochemical composition of a variety of processed crop materials and how this can be used in manifold ways, e.g. to further our understanding of the effect of food processing, to develop new strategies for assessing food quality, to identify biomarkers for food adulteration, etc. Furthermore, metabolomics is also already being applied to find out what happens to components in our food when they enter the human body and are subjected to external (gut) and internal (blood, liver, etc.) environments. We all know how eating asparagus rapidly leads to a more 'fragrant' urine (Mitchell, 2001) but the metabolic and catabolic processes behind this phenomenon are still poorly understood. The same is true of the digestion, uptake and metabolism of most of our food components. Here also, metabolomics is giving us unprecedented new insights into the integral complexity and highly interactive nature of food component digestion and, in this regard, metabolomics is really reflecting a change in scientific philosophy regarding how we approach such complex biological questions.

7.2 Potato and tomato – both fresh and processed

These two crops represent two of the world's top four most consumed vegetables. Furthermore, they are two of the most widely grown crops globally, with varieties having been bred to grow in a wide range of climatic conditions with highly varying temperatures, degrees of humidity, soil composition, day

length, etc. Considering their importance, it is therefore not surprising that both have already been the subject of extensive metabolomic investigations.

7.2.1 Potato metabolomics

The importance of potato, the 3rd most important global food crop, has meant that there has been a commensurate effort at the metabolome level with respect to development, end use, agricultural practices, etc. In particular, the development of potato *per se,* via the introgression of different metabolite contents and diversities, with a view to developing new varieties has been addressed by groups working with wild species collections. For example, the Commonwealth Potato Collection (Anon, 2010) – 1500 accessions of about 80 wild and cultivated potato species – was sub-sampled and analyzed by gas chromatography-mass spectrometry (GC-MS) focused metabolomics (Davies, 2006). These analyses clearly showed that specific taxonomic groups segregated on the basis of both non-polar and polar metabolites (e.g. amino acids). Interestingly, the alternative approach of metabolite fingerprinting using direct infusion-mass spectrometry (DI-MS; positive ion mode) was able to differentiate accessions and taxonomic classifications but this time, due to the sensitivities of the technology, the components driving the segregations were identified as mass ions associated with specific glycoalkaloids with some groups dominated by demissine, others by commersonine, α-tomatine and dehydro-demissine, or by α-solanine and α-chaconine.

More detailed studies on narrower subsets were undertaken by Dobson *et al*. (2008, 2010) who also employed GC-MS based metabolomics to analyze 29 genetically diverse potato cultivars and landraces (Dobson *et al*., 2008). This study had an interesting construction, employing 27 tetraploid cultivars and landraces, comprising 20 × *Solanum tuberosum* ssp. *Tuberosum* (16 with known introgression of a variety of useful traits from a variety of wild species, and four with no introgressed disease resistance), 7 Chilean landraces, as well as two diploid cultivars (*Solanum phureja).* GC-MS polar metabolomics highlighted several accessions with high specific metabolites (sugars and amino acids) whilst the corresponding non-polar analyses showed lesser discrimination and these were largely based on minor fatty acids. The data did highlight however, that although the variation among the cultivars and landraces was not great, sometimes there was considerable variation among field replicates.

In a more recent study by Dobson *et al*. (2010), confined to four potato species (Andigena, Phureja, Stenotomum and Tuberosum), the data showed that there was a large range in levels of metabolites, including those such as asparagine, fructose and glucose. This could have a significant bearing in the research effort towards the generation of low acrylamide-forming potatoes and indeed has been supported by the recent research of McCann *et al*. (2010) whose study into wild potatoes and acrylamide forming potential, showed

that there were species-specific relationships between sugar profiles, specific sugars and potato chip colour.

Others have utilized the DI-MS approach and Beckmann *et al*. (2007) employed this, along with GC-MS, to assess compositional differences in potato cultivars. Like Dobson *et al*. (2008, 2010) they found that the flow infusion electrospray ionization mass spectrometry (FIE-MS) approach suggested that large differences existed between tubers of individual cultivars. The associated GC-MS data also identified changes in the metabolites closely associated with established quality traits of potato. For example, levels of the amino acids isoleucine, tyrosine and phenylalanine were higher in certain cultivars. These amino acids are associated with flavour/aroma, post cooking blackening and bruising.

As stated earlier, the global importance of potato has meant that it has been the focus of an intense research effort on many fronts and the advent of genetic modification technologies was no exception and latterly, these approaches have been accompanied by associated metabolomic analysis to support the transformations and assessment of the potential for unintended effects. A GC-MS study by Roessner *et al*. (2001) on a genetically modified (GM) potato with altered sucrose catabolism detected ∼90 metabolites including sugars, sugar alcohols, amino acids, organic acids and several miscellaneous compounds. They showed that the specific transformation was accompanied by metabolites associated with several metabolic pathways increasing in tandem in the GM tuber compared to the wild type. Conversely, nine metabolites were shown to be reduced below detectable limits in the GM tubers.

A different analytical approach to the analysis of GM and wild-type potato was taken by Defernez *et al*. (2004) who employed Nuclear Magnetic Resonance (NMR) and Liquid Chromatography (LC)-MS to analyze about 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). Interestingly, the metabolite-related changes accompanying the GM event were small in comparison to those between the two parent varieties with both Principle Components Analysis (PCA) and individual compound ANalysis Of VAriance (ANOVA) supporting this. This finding is not isolated as the combined GC-MS and flow injection-MS study by Catchpole *et al*. (2005) used GC-time-of-flight (ToF)-MS and FIE-MS to provide a comprehensive comparison of total metabolites in field-grown potato genetically modified to induce fructan biosynthesis. With the exception of the predicted intended effects of up-regulated fructans and their expected derivatives, the levels of metabolites detected were very similar in the GM and its control. Importantly, metabolite levels in the GM lines fell within the range of the five non-GM commercial cultivars used as reference material. In fact, the major finding from the study was the large variation in the metabolite profile between the five conventional cultivars that overrode the differences between GM and the associated wild type parent.

One of the benefits of a metabolomics approach is the breadth of coverage and this is exemplified in the work of Parr *et al*. (2005) who, during the assessment of potato tubers for compositional changes occurring after genetic modifications to different metabolic pathways, identified kukoamine A, a spermine alkaloid, and related compounds, in wild-type tubers. Although this class of compounds had previously been reported in other species (*Lycium chinense* [Funayama *et al*., 1995], *Aphelandra tetragona* [Hedberg *et al*., 1996] and *Iochroma cyaneum* [Sattar *et al*., 1990]) this was the first report in a major food species. Indeed following this, the components were subsequently detected in other Solanaceae such as tomato (*Lycopersicon esculentum*) and tobacco (*Nicotiana sylvestris)* (Parr *et al*., 2005). This unexpected discovery in Solanaceous species highlights the potential and utility of metabolomics approaches since clearly these metabolites would have remained undiscovered had a targeted approach been employed.

7.2.2 Fresh tomatoes

No article on metabolomics analysis of crop plants would be complete without attention being given to the tomato. Here however, we shall just give a short summary of the main findings as this topic has been extensively been covered elsewhere (de Vos *et al*., 2010a). Tomato has become the model fruit crop for many biological studies and is a favorite for many prominent groups working on the development and application of metabolomics technologies to forward our understanding of important aspects of the fruits in terms of development, ripening process, functional genomics analysis, metabolic engineering etc (Bovy *et al*., 2007, 2010; Schijlen *et al*., 2008; Fernie & Schauer, 2009; Osorio *et al*., 2009; de Vos *et al*., 2010b). Taste is of course a key issue (Tikunov *et al*., 2005, 2010; van den Heuvel *et al*., 2008) but the presence and potential health-promoting effects of anti-oxidant (poly) phenolic components have been the topic of major research efforts (Rein *et al*., 2006; Butelli *et al*., 2008; Gonzali *et al*., 2009).

Studies utilizing GC-MS, LC-MS and NMR have been employed to assess metabolite changes and variation in (*S. lycopersicum*) fruits. Significantly, Moco *et al*. (2006) published a Metabolome Tomato Database (MoTo DB) dedicated to LC-MS based metabolomics of tomato fruit (Grennan, 2009) and this has been accompanied by other groups such as the one at Cornell (http://ted.bti.cornell.edu/cgi-bin/TFGD/metabolite/home.cgi) who have also developed Plant MetGenMAP, which facilitates the identification of changes in pathways and biological processes from 'omic (including metabolomic) data. The MoTo DB was based on literature research into metabolites reported to be present in tomato fruit from both wild and cultivated varieties as well as transgenic tomato plants. It has subsequently been expanded following metabolomic analysis of a representative tomato fruit sample made by combining fruits of 96 different tomato cultivars producing ripe red, orange-coloured beef, round or cherry fruits at different stages

of ripening (Tikunov *et al*., 2005). In addition to this, a selection of purple-skinned fruits was also analyzed for anthocyanins, which are known only to occur in certain varieties (Jones *et al*., 2003) or in transgenic plants (Mathews *et al*., 2003; Butelli *et al*., 2008). Schauer *et al*. (2005) further expanded the metabolite knowledge on wild species of tomato in a comparative GS-MS study of both wild and cultivated tomato wherein they found that changes in the metabolite contents of the fruit were identified in the wild species that are potentially important with respect to stress responses, as well as in metabolites of nutritional importance. For example, the wild species generally exhibited lower levels of dehydroascorbate, L-ascorbate, succinate and threonate but selected species showed elevated levels of citramalate chlorogenate, fumarate and salicylate etc (all in comparison to the cultivated species).

Furthermore, the developmental analysis of tomato, via metabolic engineering, has exploited the opportunities offered by metabolomics with Fraser *et al*. (2007a, 2009) highlighting metabolite-metabolite correlations associated with relative changes following overexpression of Psy-1 (a carotenoid biosynthesis phytoene synthase) compared to the wild type. This showed that there were multiple metabolite correlations going beyond the expected 'within-chemical class' ones, with many primary metabolite-isoprenoid/carotenoid correlations being identified. Mintz-Oron *et al*. (2008) further refined this to specific tissues (peel and flesh) and used a combined GC-MS and ultra performance (UP)LC-ToF-MS approach to study tomato fruit development. They found that 100 metabolites were enriched in the peel tissue during development including flavonoids, glycoalkaloids and amyrin-type pentacyclic triterpenoids as well as polar metabolites associated with cuticle and cell wall metabolism and protection against photo-oxidative stress. The combined approach, and the inclusion and correlation with associate transcriptomic data, suggested that the formation of cuticular lipids preceded phenylpropanoid and flavonoid biosynthesis.

Moco *et al*. (2008) comprehensively followed this up for ripe tomato fruit using NMR and LC-ToF-MS and PCA approaches to characterize 50 different tomato cultivars, including cherry, beef and round types of fruit. These studies highlighted major sugar differences between the cherry and the beef/round tomatoes. The beef and round tomatoes could not be segregated suggesting that they were metabolically similar (at point of sampling). A similar broad scale metabolic analysis was undertaken by Ursem *et al*. (2008) who used GC-MS to analyze 94 tomato genotypes with the aim of applying network analysis of correlations between abundances of metabolites across the genotypes to elucidate the biological basis of organoleptic variation in tomato. They identified several metabolites that exhibited a wide abundance range and high heritability, such as 2-methoxyphenol and methylsalicylate, whilst other metabolites such as 2-methylbutanal, 3-methylbutanol, 2-methylbutanol, 2-isobutylthiazol and phenylacetaldehyde, although high heritable exhibited a less expansive abundance range. Subsequent network analysis of correlations of the metabolomic data showed largely logical connections with, for example, methoxyphenol and methylsalicylate being strongly connected and

the phenolic metabolites (phenylethanol and phenylacetaldehyde) and the (iso)leucine derived metabolites (2- and 3-methylbutanol, 2-methylbutanal) similarly showing paired correlations.

A different approach to metabolite analysis was taken by Fraser *et al*. (2007b) who applied matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI/ToF-MS) to tomato, predominantly to focus on carotenoid metabolism and diversity. They showed that this approach was useful for the rapid identification and quantification (by isotope dilution) of carotenoids present in crude extracts from plant tissues and whole cells and facilitated simultaneous semi-quantitative determination of carotenoid metabolites (m/z values) in crude plant extracts. As with other metabolomic approaches the MALDI/ToF-MS lent itself to multivariate analysis and subsequent segregation of genotypes. One of the clear advantages of this approach was the ability to characterize and quantify carotenoids that have generally proved recalcitrant via LC-MS-based approaches.

Metabolomics has also been employed to determine the effects of abiotic stress on tomato. Bauer *et al*. (1997) highlighted changes in the amino acid pools, specifically asparagine, glutamine and glutamic acid of water-stressed tomato. More recently, Sánchez Pérez *et al*. (2009) used a chemometric strategy based on multivariate curve resolution and alternating least-squares (MCR-ALS) applied to LC-MS three-way data arrays to assess the impact of carbofuran application, a carbamate pesticide, on tomato. They showed that pesticide treatment produced altered metabolites reflecting physiological stress.

As highlighted above, GM has been utilized to tease apart the biochemical pathways in tomato (Fraser *et al*. 2007a, 2007b, 2009) and the consequence of the transformation on tomato fruit has been the focus of several studies. Le Gall *et al*. (2003) used NMR to study metabolite changes in hydroponically glasshouse-grown GM tomatoes (with overexpressed flavonols) compared to the controls. They found that subsequent PCA analysis showed separation of the samples into discrete groups – GM, control, and according to ripening stage. An analogous approach was taken by Noteborn *et al*. (2000), using also LC-MS, who found that there was a large number of significant differences (100–200 metabolites) between GM lines and controls in two series of modified tomatoes. In one of the GM lines, the Cry1Ab5 protein from *Bacillus thuringiensis* was expressed, which showed a differential level of 100–200 metabolites over the three years that the crops were grown compared to the wild type. Interestingly when the data from all three years were combined no significant differences at all were obtained suggesting that the environmental, year-on-year (seasonal) influence exerted a much greater effect than any consequences from the GM event.

7.2.3 Tomato puree – a model for the food processing industry?

Next to being a fresh product, tomato fruits are also very widely used as processed foodstuff both in terms of tinned/pasteurized materials and as a fruit puree concentrate. These products are very widely eaten, in particular

in many European countries, where they are almost a common denominator of the daily diet (Capanoglu *et al*., 2010). The processing of fresh tomatoes into tomato puree involves a number of washing, chopping, filtration, evaporation and Pasteurization steps, all of which can be expected to have some (as yet unknown) effect on the biochemical composition of the end product. In the first major study, untargeted high performance (HP) LC-electrospray ionization (ESI)-quadrupole(Q)ToF-MS has been used to follow this process from field to can in order to determine which steps in this process have the greatest impact on the metabolic composition, particularly with regard to the key anti-oxidant groups, carotenoids and flavonoids. Results revealed that major changes occurred, usually, but not always, to the detriment of the anti-oxidant content (Capanoglu *et al*., 2008). Perhaps the most significant observation was that the pulverization process proved to be inadequate to fully destruct the epidermis with the consequence that the removal of this skin fraction during the filtration step, resulted in a concomitant loss of the majority of the phenolic components that are known to be solely concentrated in this tissue in wild-type tomato fruits (Bovy *et al.*, 2002). Additional studies, some involving metabolomics, others involving more targeted approaches, have also shown that essentially, processed tomato products are biochemically, significantly different than their fresh starting materials (for a complete review of the literature see, Capanoglu *et al*., 2010). Furthermore, subsequent cooking in the home is also a major factor in determining what we eventually ingest and this component in the entire process, farm-to-fork, is still pretty much a black box regarding what enters our digestive system and hence also deserves more detailed attention. Metabolomics in the home may be the next step!

7.3 Grain crops

7.3.1 The cereals

Surprisingly cereals, unlike the Solanaceous crops potato and tomato, have had a limited exposure to metabolomic analysis. Baker *et al*. (2006) used NMR to analyze the metabolite changes in three field-grown transgenic wheat genotypes expressing additional high molecular-weight subunit genes and the corresponding parental lines (including a null (azygous) transformant) all cultivated at two sites. They found, as with the tomato studies of Noteborn *et al*. (2000), that site (environment) and not GM was the dominant factor in metabolite changes. However, separation of the transgenic and parental lines was observed predominantly due to increased levels of maltose and/or sucrose in the transgenic line (B73-6-1, highest expressor), and, to a lesser extent, to differences in free amino acids. GC-MS analysis of material from one of the growth years corroborated the site-related amino acid changes

with significant differences in acidic amino acids (glutamic, aspartic) and their amine equivalents (glutamine, asparagine). Conversely the same lines showed elevation in proline and γ-aminobutyric acid (GABA) at the alternative environment. Obert *et al*. (2004) also failed to find any significant differences between grain field plots of herbicide-resistant and control lines using a combined GC and LC approach.

At the more applied end, Beleggia *et al*. (2009) undertook a combined GC-MS and GC-static headspace solid-phase micro-extraction (GC-HS-SPME) metabolomic approach to study the interplay between metabolite and volatile (organoleptic) components in semolina and pasta obtained from four durum wheat cultivars (*Triticum durum* Desf., cvs. PR22D89, Creso, Cappelli, Trinakria). The correlations between cooked pasta volatiles and semolina metabolites demonstrated that the flavour of the end product may significantly differ depending on the durum wheat cultivar employed.

An initial foray into the potential impact of climate change of the wheat metabolome was reported by Levine *et al*. (2008) who assessed the impact of sub-, optimal- and supra-CO_2 concentrations on the wheat metabolome during development. Their GC and LC-MS analyses revealed that both [CO_2] and physiological age exert an impact with plants grown under high [CO_2] exhibiting metabolite profiles similar to those of plants grown under ambient [CO_2]. More specifically, elevated [CO_2] promoted the accumulation of secondary metabolites (flavonoids) progressively to a greater extent as plants became mature.

A greater metabolomics effort has been focused toward maize, with targeted studies of maize kernels highlighting the influence of genetic background and growing season (Ridley *et al*., 2004; Reynolds *et al*., 2005), developmental stage (Seebauer *et al*., 2004) and environment and agricultural practice (Harrigan *et al*., 2007a, 2007b) on the natural variability of metabolites.

Maize metabolite biodiversity has recently been reported by Röhlig *et al*. (2009) who employed a multiple fractionation approach to separate maize grain fatty acid methyl esters and hydrocarbons (fraction I), free fatty acids, alcohols and sterols (fraction II), sugars and sugar alcohols (fraction III), acids, amino acids and amines (fraction IV). These were analyzed by GC-MS yielding 300 distinct analytes of which 167 could be identified. Analysis of the metabolite variation of the grain, from four maize cultivars, at differing developmental stages, grown for three years, showed that all factors contributed to differences. Significantly when the data for all fractions per cultivar were combined, the cultivars could be distinguished by PCA but when the combined data from all three growing seasons were co-analyzed the cultivars could no longer be distinguished as being different but rather, the collated data did segregate according to growth year. The data therefore indicate a more pronounced impact of growing season corroborating previous findings in potato and tomato. Furthermore PCA analysis of the metabolite variation, centred on one variety grown for three years at four locations, showed that there was within-year segregation of location but this differed year on year.

7.3.2 Rice metabolomics

Rice is world food crop number one. Rice grains are the staple food of a huge proportion of the world's population, primarily located in the developing world. In S.E. Asia alone, rice provides about 75% of total calorific intake (Garris *et al.*, 2005; Fitzgerald & Hall 2008; Hall *et al*., 2008; Fitzgerald *et al*., 2009). For many, rice is the main source of daily macro-nutrients and micro-nutrients and, as such, has been the topic of much research into means not just to enhance crop yield but also grain quality. The latter can be defined both in terms of nutritional quality and also, quality related to flavour. The former is of fundamental nutritional importance and the latter has major influence on cultural preference for particular rice varieties (e.g. Jasmine versus Basmati). Furthermore, the qualities of so-called, fragrant rice varieties also determines their market value and hence, import revenues, thus contributing significantly to the GNP of many of the poorer Asian countries. Nevertheless, in each case, both traits concerned directly relate to metabolite content and it is therefore not surprising that metabolomics is already actively being considered, and employed, as an advanced tool in rice research. Rice is often chosen as the model grain crop, not just due to its societal importance, but also, due to the advantages associated with its small genome size (Garris *et al*., 2005). Indeed, rice was chosen by Kind *et al*. (2009) for a technical study on the value of available plant metabolomic datasets. They concluded that these cannot yet be used with confidence to predict how large a plant metabolome actually is. In this regard, much more uniformity of practice and combination of efforts is needed before we will gain a true and complete picture of the richness of the individual plant metabolome.

Metabolomic analyses of rice foliage have already been initiated and methods have been reported for the use of both Capillary Electrophoresis (CE)-MS and CE-PDA for the analysis of 88 key (mainly primary) water-soluble metabolites. Results revealed that on quantification, levels were comparable to those reported for tobacco but were about 10× lower than is typical for Arabidopsis (Sato *et al.*, 2004). Later, this work was continued and the same approaches were used to perform time-resolved metabolomics of rice leaf metabolic content over a 24-hour light/dark period (Sato *et al.*, 2008). Here, the synchronous dynamics of 56 primary metabolites was followed simultaneously at 1-hour intervals. Unsupervised statistical clustering clearly separated the light/dark cycle and the authors predict that hierarchical clustering of a correlation coefficient matrix could help identify enzymatic bottlenecks regulating metabolic networks under specific environmental conditions.

Methods for rice grain metabolomics have also been published despite the technological challenges of such materials, which contain about 90% starch. Both 1D GC-MS and 2D GCxGC-MS approaches have been developed for brown (unpolished) rice grains by Kusano *et al*. (2007). Here, using a collection of 68 rice genotypes, chosen to represent 90% of the known rice DNA polymorphism, GC-based analyses of derivatized extracts followed by multivariate unsupervised PCA and supervised PLS-DS analyses revealed

discriminatory metabolite/metabolic profiles between contrasting rice types. The clear added value of GCxGC for higher resolution was emphasized. Shu *et al.* (2008) also used GC-MS to analyze derivatized extracts of germinating rice seedlings. Profiles of a broad spectrum of lipophilic and hydrophilic compounds were followed over a 96-hour period and the 'extensive metabolic dynamism' observed clearly demonstrated the major time-dependent shifts and causal connectivity in seed metabolism occurring during this critical phase in the plant life cycle. Methods based on GC-MS/GC-FID (flame ionization detector) have also been used by Zhou *et al.* (2009) to investigate potentially unintended metabolic consequences of transgenesis. Examining polar grain extracts did reveal significant changes in a number of key primary metabolites in the transgenic lines as compared to non-transgenic, wild type controls.

A problem associated with nutritive value in cereals has been the anti-nutrient phytic acid that inhibits iron uptake. Attempts to develop crops (e.g. rice, maize, barley, wheat) with lowered contents of the anti-nutrient phytic acid have been described by Raboy (2007). Low phytic acid crop mutants are typically selected on the basis of their altered levels of inorganic phosphorous (P_i). With respect to rice, metabolomics has been employed to characterize two *lpa* rice mutants (*Os-lpa*-XS110-1 and *Os-lpa*-XS110-2), generated by γ-irradiation of the corresponding wild-type rice (Xiushui 110) and grown at five field trial sites in China in 2005/2006. The mutant *Os-lpa*-XS110-1 showed a significant segregation from the associated wild type due to the polar metabolite profiles in this mutant. This was less evident for the other mutant/wild type comparison but the rice lines were well differentiated by growth location (Frank *et al*., 2009). More detailed analysis at the compound level of mutation-derived variation showed that these were largely accounted for by methyl pentadecanoate, galactose, raffinose, *myo*-inositol and phosphate, the last two being key components in the phytic acid biosynthetic pathway (Frank *et al*, 2007)

Fragrant rices (such as the well – known Jasmine and Basmati varieties) have also already been the subject of preliminary metabolomics analyses. Suitable methods for the extraction and concentration of natural volatiles, using for example SPME and TENAX trapping, followed by GC-MS based methods for separation/detection have been developed (Verhoeven *et al*., 2010). Early results have shown that different fragrant rice genotypes are readily separated using simple PCA or hierarchical clustering and the importance and relevance of this trait to rice-producing countries has been emphasized (Hall *et al*., 2008). Rice fragrance is a factor of key relevance to market value but is not lacking political sensitivity (Hall *et al.*, 2008; Fitzgerald *et al*., 2009). Metabolomics can play a key role here in helping us to define better what we mean by rice quality in terms of fragrance and to identify those compounds of greatest importance in determining positive and negative sensory characteristics. A more targeted breeding strategy specific for fragrance characteristics should therefore become possible.

c07 BLBK354-Hall October 29, 2010 21:16 Trim: 234mm×156mm Series: APR Char Count=

7.4 Soft fruit metabolomics

The underpinning metabolic components characterizing fruit quality, including nutrition bioactivity and safety, and hence public purchase and consumption, are significantly diverse encompassing simple sugars acids, amino acids, carotenoids and simple through to complex polyphenolics to name but a few of the chemical classes. Adding to this complexity is further subclass diversity with the polyphenols generally being described as comprising anthocyanins, flavonols, (iso)flavones, flavanones, catechins, ellagitannins, cinnamates and hydroxyl benzoic acids and stillbenes (Pietta *et al.*, 2003; D'Archivio *et al.*, 2007; Mullen *et al.*, 2007). In addition, these subclasses are further populated by differential levels and pattern of polyphenol polymerization, glycosylation, methylation and acylation (Clifford, 2000; Reed *et al.*, 2005; Xie & Dixon, 2005; Prior & Wu, 2006).

Hand in hand with phytochemical diversity is a broad dynamic range. For example, in fruit, the total anthocyanin content can be as low as to be virtually undetectable in fruit such as banana, but can reach levels of 2–10 mg/g fresh weight in blackcurrant, raspberry, blueberry and the lesser researched fruits such as choke berry and elderberry (Clifford, 2000). Similarly, other metabolites impacting upon organolepsis (sugars and organic acids), nutrition (vitamins C, A etc) and putative bioactive components (flavonoids and ellagotannins) also display similar levels of variation (Anon 2003; Beekwilder *et al.*, 2005, 2008; Anon 2009a, 2009b).

These factors *in toto* have meant that there have been only limited attempts to apply true metabolomics, i.e. an untargeted study of metabolite changes either by GC-MS, LC-MS, NMR, etc., to fruit and these have been limited to melon (Biais *et al.*, 2009a, 2009b), raspberry (Stewart *et al.*, 2007; McDougall *et al.*, 2008) and strawberry (Fait *et al.*, 2008). The approach taken by Biais *et al.* (2009a, 2009b) on melon has been one of establishing, within these large fruit, spatial variation in primary metabolites and using a cross comparative approach wherein the metabolomics data generated via both ^{1}H NMR and GC-ToF-MS systems were mined for the metabolite trends at a spatial level and compared using independently performed PCA and multi-block hierarchical PCA (HPCA), respectively. In general the analytical systems reported on the same primary metabolites and yielded similar metabolite spatial trends. A confirmation of this cross comparability was provided by a correlation-based superblock HPCA for direct comparison of both analytical data sets. Indeed it is with the HPCA approach that advances were evident allowing different source data sets, with different levels of sensitivity to be confidently cross-compared thereby extending the validity of the multi-analytical approach to metabolomics. For melon at least this has been extended to determine the underlying factors impacting upon shelf life and associated spoilage via hypoxia-related fermentation.

For strawberries, fruit development has been studied using a combined GC-MS and UPLC-QToF-MS approach (Fait *et al.*, 2008), which reported not

only commonalities in metabolite trends via reporting on primary metabolism but also extended the analysis into a key sector of fruit quality by reporting on the variation on 105 secondary metabolites including phenylpropanoid derivatives both with respect to spatial and developmental presence and changes. This is a significant step beyond the state-of-the-art, which has generally confined itself to reporting on changes in specific chemical classes such as flavonoids (Wang *et al*., 2003; Panico *et al*., 2009), amino acids (Keutgen & Pawelzik, 2008) accompanying stress or biodiversity (Capocasa *et al*., 2008).

The application of metabolomics to study trait inheritance or the influence of the environment on primary and secondary changes is very much in its infancy with respect to fruit. This approach has been hampered, at least with respect to fruit breeding, by the sheer numbers of samples (distinct lines and replication) to be analyzed in a standard segregating cross. Methods to manage this have been developed by Stewart *et al*. (2007) and McDougall *et al*. (2008) who have truncated standard LC-MS to give a short column method that is closer to DI-MS. Their study, employing the same segregating raspberry cross 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. The MS data generated from short column direct infusion MS (SC-DI-MS) was subject to PCA analysis and this showed that year-on-year variation was perhaps the key driver of metabolite variation at least over the period studied. Interestingly the data for one of the years showed a clear environmental segregation the source of which can be traced back to the weather during the fruit development period, being distinctly dry and that the soils on each site exhibited differential water retention abilities.

Consideration of each year's fruit and environment as distinct experiments showed, following reanalysis, that chemical class segregation was evident across the populations. A wealth of polyphenols was characterized following comparison to standards but the most evident amongst these were the following: cyanidin 3-glucoside, cyanidin 3-sophoroside, cyanidin 3-glucosylrutinoside cyaniding 3-rutinoside, pelargonidin 3-sophoroside, pelargonidin 3-glucosylrutinoside and quercetin acetylrutinoside. An apparent feature of the associated PCAs was the clean segregation between the cyanidin-3-sophoroside and cyanidin-3-rutinoside associated groups. This is extremely informative and means that the SC-DI-MS approach could facilitate the rapid identification of (screen for) plant progeny showing relatively elevated levels of these compounds, thereby potentially allowing targeted breeding, e.g. 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*. (2008) and is currently being expanded, via collaboration with other groups, into strawberry and blackcurrant breeding. The metabolomic correlation with sensory scoring is also in progress.

c07 BLBK354-Hall October 29, 2010 21:16 Trim: 234mm×156mm Series: APR Char Count=

Clearly the application of metabolomics to soft fruit *per se* has a long way to go. However, the speed and range of the technologies, in combination with the development of statistical approaches that will facilitate metabolite and sensory cross comparison, mean that there is every likelihood that it will prosper and develop in this crop sector.

7.5 Metabolomics and our most important beverages – coffee, tea and wine

Perhaps some of the most heavily processed plant products that we consume are the plant-based beverages. Products such as tea and coffee go through extensive treatments, post-harvest, which are not only known, but have been specifically designed to generate a product with a specific chemical composition. Processes such as fermentation and roasting ensure that the chemical profile of the final product is wholly different from the freshly harvested leaves or beans. The conversion of grapes through juice, into wine, is less intrusive in not involving, e.g. heating steps. However, once again, the final product has undergone extensive chemical recompositioning before it reaches the bottle and indeed, this continues during subsequent storage.

7.5.1 Coffee metabolomics

Coffee is consumed, generally as a hot beverage, worldwide. Coffee extracts have also found their way into the food and confectionary industry as flavouring agents. Despite huge popularity, widespread use and a long history as a stimulant, we still lack a great deal of knowledge of those factors important in determining the quality of the coffee drink and how these relate back to genetic or environmental perturbance and industrial practices. Such knowledge is essential if we are to improve our understanding of determinant factors driving coffee quality and uniformity. Metabolomics approaches have already been seen as a potentially valuable route to rectify this hiatus in our knowledge.

Coffee is generally prepared from the beans of the shrub *Coffea arabica*. Once ripe, the berries are harvested, usually by local farmers after which they have to be processed, roasted and blended to obtain the desired taste and quality characteristics. In this process, two key components are therefore recognized, pertaining to the pre and post-harvest phases. Using a more targeted metabolic profiling approach, Joët *et al.* (2009) aimed to reconstruct the main metabolic pathways in Coffea seeds related to the main seed storage compounds. Contrasting patterns of sugars and chlorogenic acids accumulation revealed the complexity and dynamic nature of seed development. Both compound groups are known to have major influences on the sensory properties of the final product.

c07 BLBK354-Hall October 29, 2010 21:16 Trim: 234mm×156mm Series: APR Char Count=

As with other processed products such as wine and tea, where sensory issues are also crucial, both the genotype and the source of the beans are of great importance. Regarding the source of the starting materials, geographical location and its related climatic conditions (rainfall, sunlight, soil composition, etc.) are known to play an influential role. However, as with many processed products, what happens after harvesting can be equally or even, more important to the end product. Most metabolomic investigations on coffee to date, have actually concentrated on this phase of the production process. Prior to roasting, raw, 'green' coffee beans, once harvested, must be processed, during which the seeds are separated from the fruit wall (dehulling). In Brazil, for example, this can be done either through wet or dry processing, the choice of which is often down to the region or even individual farmer. During wet processing, the fruit is kept moist and after a short fermentation period the fruit wall can be removed before the isolated beans are dried in the sun. In dry processing, the entire fruits are dried after which the dry hulls are mechanically removed. As might be anticipated, both processes are prone to local differences resulting in differences in final quality and as the two processing steps are dramatically different, contrasts in composition are inevitable. Selmar and Bytof (2007), using both biochemical and gene expression analyses, showed that during wet processing coffee seeds actually begin the germination process while in dry processing, the seed rapidly undergo stress responses. These differences concomitantly result in detectable biochemical changes such as a significantly increased level of GABA following wet processing. The authors conclude that the 'peculiarities' of wet and dry processing have a significant biochemical effect on the final product and its distinctive coffee characteristics and thus deserves further investigation.

In an attempt to understand better the drying process, Borém *et al.* (2007) followed the changes in cell structure and plasma membrane integrity in arabica beans, which were then later correlated with metabolite profile changes observed to occur simultaneously during this process (de Vos *et al.*, 2007, 2010). Using SEM/TEM (Scanning Electron Microscopy/Transmission Electron Microscopy), suboptimal drying conditions were shown to be accompanied by membrane breakdown and the loss of sub-cellular organization and it was predicted that this could be related to loss of quality and the occurrence of off-flavours in these beans. Both GC-MS and LC(QToF)MS were subsequently used in non-targeted metabolomics approaches to reveal both the changes occurring during coffee bean processing and drying and how these may relate to the conditions used (de Vos *et al.*, 2007, 2010). PCA treatment clearly separated the wet and dry treatments and results revealed key differences in glucose and fructose (but not sucrose) and in certain organic acids. Taking this work even further, Lindinger *et al.* (2010) have been able to show the power of a non-targeted chemometric approach, based on different analytical methods (PTR-MS (proton transfer reaction) and GC-ToF-MS) to detect a novel marker for off-flavour in coffee. By first performing sensory analysis with a trained panel, combined with the choice of well-defined coffee

hor: After styling of the reference list, reference "de Vos *et al.*, nd "de Vos *et al.*, 2010c" has resulted into "de Vos *et al.*," and "de Vos *et al.*,". Please specify in the text citation "de Vos *et al.*," whether it is 2010a "or 2010b

preparations, in this case robusta coffee from Ecuador (*C. canephora*), ethylformate was identified as a head-space marker for roasted coffee of inferior quality.

While the above work is generally still preliminary and much more can still be expected from coffee metabolomics, the story of course does not end there. While quality studies and sensory analyses will reveal more about the type and origin of quality differences, the fate of the coffee metabolites as they enter the body and are taken up and metabolized is already also a topic of investigation. The stimulatory effect of caffeine is of course known, but coffee is a tremendously rich source of potentially healthy chlorogenic acids and their derivatives. Allard *et al.* (2008) employed CE-ESI-ToF-MS to follow the fate of coffee and tea metabolites after human consumption by monitoring the biochemical profiles of the urine from 13 individual volunteers. Results revealed highly significant differences as a result of beverage intake and that the MS spectra revealed 'hot spots' representing groups of discriminatory molecules deserving further analysis and identification. Stalmach *et al.* (2009) performed a similar study and used HPLC-PDA-MS to follow the uptake and metabolism of chlorogenic acids-enriched coffee for a period of 24 hours following ingestion. Results clearly revealed that uptake of key chlorogenic acids is extensive and may take place either in the small or the large intestine where, in the latter, they can also first be metabolized by the gut flora before entering the blood. The dynamic complexity of the process is considerable and one can still only speculate as to the timing and location of the modifications observed and their potential biological relevance. Two biomarkers for the consumption of even small amounts of coffee were identified in urine – dihydrocaffeic acid-3-*O*-sulfate and feruloylglycine. Both these molecules are not present in the coffee itself but are likely to be the result of *in vivo* metabolism by gut microflora and/or liver detoxification reactions. How exactly these emerging patterns bear relationship to the underlying molecular and biochemical mechanisms will be the topic of much metabolomics research for years to come.

7.5.2 Tea metabolomics

With tea (*Camellia sinensis*) being the most widely consumed non-alcoholic beverage on the planet, it is perhaps not surprising that tea has already also been the topic of several metabolomics studies. For example, Thomas *et al.* (2006) used metabolic profiling to characterize somaclonal variants derived from *in vitro* culture in relation to their quality aspects. However, few papers have focused on the starting materials. Several have been dedicated to the potential health benefits of tea as a rich source of bioactive antioxidants, and on the effects of tea consumption and the fate of tea metabolites after they enter the human body. Once again, metabolomics is revealing not only the complexity of the plant materials but also the dynamics of tea metabolites once they enter our blood system via the digestive tract.

Tea is grown in about 30 countries and two main products are predominantly made – green tea, where young leaves are simply rolled, steamed to reduce oxidation and then dried and black tea. For black tea, the leaves are gently crushed to break down cellular compartmentalization after which they are left to 'ferment' for 1–2 hours at ambient temperature (in tea–producing countries, often 25–35°C). This mixing of the cell contents, and subsequent incubation, is known to result in extensive oxidation and the conversion of simple phenolics into complex condensed tannins. Consequently, green and black teas are significantly different not only in appearance but also in chemical composition and their association with potential health benefits is the cause of much debate and associated research. In an early study, del Rio *et al.* (2004) used HPLC-ESI-MS^n to compare black and green tea extracts. A wide range of standard phenolic and alkaloidal reference compounds were used to identify key discriminatory metabolites. All the identified flavan-3-ols in black tea were at significantly higher levels and the total flavan-3-ol content was about 45 times higher than in green tea. In contrast, gallic acid was 20 times higher in green tea. Several theaflavins were detected and identified in black tea, which, as expected, were undetectable in green tea extracts as these are known to be formed during the fermentation step (van Dorsten *et al.*, 2006). All in all, this early metabolomics study demonstrated the value of MS^n as a metabolite identification tool for even trace levels of key flavonols in foods although the complexity of positional (sugar) substitutions still requires NMR for exact structural determination.

So far, most metabolomic research has focused on green tea, which has had the greatest association with potential health benefits. In an early metabolite profiling study, Le Gall *et al.* (2004) used ^{1}H NMR to profile 191 green teas, including 17 of the jasmine type. Discrimination based on chemometric analysis in relation to source of origin or quality, was only partially successful although some samples were readily separable from related ones. Furthermore, some samples were so distinctive that possible biomarkers for use in authentication were proposed. Fukusaki's group in Japan has performed extensive studies on a range of selected green tea samples (Ikeda *et al.*, 2007; Pongsuwan *et al.*, 2007, 2008). Interestingly, in this case the teas chosen had already been subjected to extensive quality analysis as part of a professional tea tasting contest. FT-NIR (Fourier transform-near infra red) spectroscopy was used by Ikeda *et al*. (2007) to profile 13 contrasting green tea samples, selected from a set of 53, and on the basis of the results a reliable, quality-prediction model was proposed for what appears to be a rapid analytical method. Subsequently GC-ToF-MS and UPLC-ToF-MS were used as alternative metabolomic profiling methods. GC-ToF-MS analysis of the full set of 53 professionally tasted teas, followed by chemometrics treatment of the data, enabled both sample separation and qualification. Discriminatory metabolites potentially linked to quality were also identified (Pongsuwan *et al*., 2007). Similarly, LC-MS methods were also found to be successful and again, a quality prediction model was generated and tested for reliability (Pongsuwan

et al., 2008). Based on the panel's sensory quality scores, a number of biomarkers for quality, identified with the help of authenticated standards, were proposed. These included a number of characteristic catechins. Finally, ^{1}H NMR analyses on the same set of 53 taste-scored samples have also been performed and again, similar conclusions could be drawn (Tarachiwin *et al*., 2007). Unfortunately, little cross-comparison between methods seems to have been carried out and no attempt seems to have been made to fuse the data from these same samples, which could have provided an even deeper insight into the extent of sample differences and enable the development of a more extensive, multiplex metabolomics – based predictor model for the quality of green tea.

Inevitably, considering tea is a rich source of dietary antioxidants putatively linked to health benefits, research has already also begun into the fate of these metabolites after ingestion. van Dorsten *et al.* (2006) treated 17 healthy male volunteers to green and black tea preparations and a caffeine control in a full scale, cross-over design experiment. The tea consumption was equivalent to 12 cups of tea/day. Subsequent, sequential urine and plasma metabolomics using high resolution ^{1}H NMR allowed the uptake, metabolism and excretion of the tea polyphenols to be followed in time. Results clearly revealed a ready separation of all samples based on treatment. Catabolism products were detected and alterations to other indigenous metabolites (e.g. blood sugar was reduced following consumption of both green and black teas) were observed. Evidence for gut microfloral involvement in tea metabolism and its general impact on human (energy) metabolism were also presented. These authors aim to generate a mechanistic understanding of the effect of tea flavonoids in humans and their effect on intermediary metabolism and these first results clearly demonstrate the complexity of the dynamics of the process and the potential added value of a metabolomics-based approach.

7.5.3 Grapes and wine

If there is one food product above all others where the biochemical profile is immediately linked to the quality and hence often, value, it is wine. The many nuances of bouquet and taste are legendary and for generations, the vintage and *terroir* (region of production associated with geographical location, soil type, sunlight, etc.) are often the determinants of quality and hence, price. It is, perhaps surprising that, to date, few metabolomic studies have yet focused on this product or the grapes used in its production. Perhaps there is a worry that some of the romance will be lost should we finally be able to define wine purely in terms of its precise biochemical composition, which is ultimately, what the consumer (or investor!) is purchasing. Nevertheless, it is to be expected that metabolomics will be widely used in the wine industry and will help with many aspects of quality control in production, in relation to authentication issues, storage regimes, etc. Indeed, the term wine-omics (Anon, 2008) has already been coined and the link between GC-MS, LC-MS

and NMR approaches to help define previously almost indefinable terms such as 'body' has already been made. As with many processed products, the complexity is already clear, even without the intervention of metabolomics, so the challenge is considerable. As well as genotype (grape variety) and *terroir*, the manner of fermentation and the strain(s) of yeast used, the subsequent manner of barreling and storage will all play a part in determining the composition of the final product and its sensory properties and hence, overall quality. That the art of wine–making may be taken, through metabolomics, to a higher level (Anon, 2008) is nevertheless, perhaps not a premise welcomed by all.

One pioneering French group published the first preliminary metabolomics study in 2005 (Periera *et al*., 2005). ^{1}H NMR was used to profile grape skin and pulp from a number of grape varieties from four Appellations in the Bordeaux region with contrasting *terriors*. Using relatively straightforward PCA on the data obtained it proved readily possible to separate the samples being compared in relation to skin vs pulp, variety vs *terrior*, variety vs variety, etc. presumably reflecting the high complexity of the mixtures and their strong correlation with source differences. PCA revealed the importance of differences in several major metabolite groups, including the sugars, amino acids and (poly)phenols as being mainly responsible for sample discrimination. Additional analyses have subsequently been performed recently by a S. Korean/Danish consortium in a short series of papers (Son *et al.*, 2008, 2009; Lee *et al.*, 2009). Here also, ^{1}H NMR was used to investigate differences between wines of the same grape variety (e.g. Cabernet Sauvignon) obtained from different world regions (Australia, France, California; Son *et al.*, 2008). Results showed that following PCA and PLS-DA, the wines were readily separable and that loading plots revealed many compounds, some known and some unknown, which were discriminatory. Wines from different grape varieties were also compared and key differences were attributed to primary metabolites such as lactate, glucose, glycerol, 2,3-butane diol and several secondary phenolic compounds. Unfortunately, in this preliminary study, all samples used had been bought from a local shop and consequently, as no metadata was available, it is not possible to attribute reliably the differences observed to specific components of the production process. However, subsequently, the same authors performed a more controlled study where they collected grapes of the variety Muscat Bailey A from a number of *terriors*, and produced wines under controlled conditions. Here, the authors were then able to relate profile differences to location and potentially, also climatic effects. High sunlight for example, known already to increase the sugar content was also found to be associated with reduced levels of other primary metabolites such as malate, citrate and certain amino acids (Son *et al*., 2009). Follow-up research on Meoru wines, again using ^{1}H NMR, allowed vintage (2006 vs 2007) and vineyard differences to be readily discerned and again, many common primary metabolites proved to be discriminatory in complementary PCA loading plots (Lee *et al.*, 2009). Proline levels and certain

phenolic compounds were clearly different in 2006 and 2007 wines and this was putatively linked to the key seasonal differences experienced in these two years. Clearly, these investigations are just preliminary but the potential of metabolomics has already been proven. With so much hanging on the importance of seasonal quality differences and the relevance of aspects such as *terrior* and variety (blends), a future for metabolomics in this industry is assured. As with tea and coffee, the metabolism of major wine components such as the polyphenols has also already become the subject of metabolomics research (Grün *et al.*, 2008). Once again, the fate of such compounds is seen to be determined by a combination of gut flora and human metabolism resulting in a complex interaction of biochemical and enzymatic reactions leading to extensive modification and catabolism of the metabolites ingested.

7.6 Food product contamination and adulteration

The food industry, particularly where it concerns processed foods, is regularly faced with issues of products being adulterated or contaminated with components other than those that should be present. So long as certain components are cheaper than other related ones, the possibility of bulking – up products with the cheaper alternatives and still claiming the full price is for some, too tempting. For importers and traders, it is however, often difficult to detect instances of adulteration (Hall *et al*., 2005, Hall, 2006a, 2006b). Does a bag of coffee really come from Brazil, or does a bottle of fruit juice really contain 100% pure orange? Here also is a potential field of application for metabolomics where, once again, the identification of specific marker compounds for particular products could lead to a relatively cheap and simple detection system for authenticity (Hall, 2006a). An early example, is the detection of cheaper vegetable oils having been used to adulterate expensive virgin olive oils by using DI-MS in analyses lasting just a couple of minutes (Goodacre *et al*., 2003). Shortly after, Reid *et al.* (2004) used SPME-GC together with chemometrics to establish a method to detect the presence of apple pulp used to bulk up strawberry fruit purees. Volatile compounds (about 35) were found to be significantly different in the adulterated samples and using GC-MS, three of these were identified to be hexanoic acid, 2-hexenal and α-farnesene, which are known key apple aroma components. Detection limits were however disappointing as the method did not prove reliable when adulteration levels fell below 25%. Le Gall *et al*. (2004) also speculated on the potential of using ^{1}H NMR to identify biomarkers indicative of the type and source of green teas as did Schaneberg *et al*. (2003) for the sourcing of Ephedera botanicals used in alternative medicines. Alternatively, the value of using of omics approaches, including metabolomics, to detect contaminating aflatoxins in various foodstuffs has been speculated upon by Bhatnagar *et al.* (2008) as an alternative to the current laborious targeted methods.

That there is great demand for automated adulteration/quality control tests is clearly evident from the SGF (Spin Generated Fingerprint) Profiling™ system developed by one of the world's main NMR manufacturers (Rinke, 2008). This ^{1}H NMR based system has been developed specifically for the food industry and its monitoring bodies for use in the automated detection of the adulteration/purity/authenticity of source of commercial fruit juices. The SGF profiling method is simple and rapid and has been developed to the extent that, being backed up by a major spectral database of >3000 fruit juices, the manufacturers are able to detect down to 10% adulteration of orange juice by mandarin juice (which is cheaper) and determine if additives (such as extra sugars) have been used. The database is also already sufficiently extensive so that it is now also possible to predict reliably the country of origin of particular fruit products (Schütz *et al.*, 2008). It is envisaged that many similar applications will be developed both for adulteration monitoring as well as quality control in the food processing industry.

7.7 Metabolite profiling technologies used to evaluate crop safety

Metabolomics has been defined as the 'comprehensive analysis of all metabolites present in an organism', (Fiehn, 2002). Currently this objective represents an impossible task, and in many cases, is inappropriate and unnecessary. As a consequence, four classifications of metabolomic analysis have emerged. Firstly 'targeted analysis', which relates to the quantitative determination of a limited number of key compounds. 'Metabolite profiling', refers to the analysis of a specific pathway or metabolite groups. The third category 'metabolomics', is the exhaustive determination of metabolites in an extract from an organism. Finally 'metabolite fingerprinting', relates to a characterized profile of an extract/organism in which peak identification is not essential. Many of these terms have become interchangeable throughout the literature and virtually all these approaches have been evaluated and applied to safety assessment of novel foodstuffs.

7.7.1 The generation and standardization of the biological material

Before embarking on the determination of chemical composition for assessing substantial equivalence, it is essential to standardize both the biological and analytical system adopted, so that phenotypic variation between samples can be determined accurately. Typically 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. an azygous or empty vector line) 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.

7.7.2 Evaluation of novel foodstuffs using targeted metabolite profiling

Decades of food analysis has revealed a set of key known, and well-characterized, metabolites essential for quality and health attributes. Therefore, an unbiased semi-quantitative method that determines numerous compounds is typically inappropriate in this instance. Instead, an efficient extraction procedure and focused analytical method providing optimal detection as well as quantification is the objective. Elegant metabolite profiling procedures using GC-ToF-MS and LC-MS/MS for volatiles contributing to aroma and taste (Tikunov *et al*., 2005) as well as phenolics that confer health and colour traits have been described (Moco, *et al*., 2006; Stewart *et al*., 2007).

A major class of pigments typically present in fruits and vegetables that are in part responsible of health and colour traits is the carotenoids. These pigments are essential dietary components for humans. β-carotene is the most potent precursor of vitamin A, while other carotenoids reduce the risk of incidence of age-related diseases such as macular degeneration (e.g. zeaxanthin and lutein) and prostate cancer (e.g. lycopene) (Fraser & Bramley, 2004). In addition to their health benefits, carotenoids confer colour to many food products. The hydrophobic and thermoliable nature of carotenoids prevents separation by GC-MS. Alternatively, HPLC has become the method of choice for carotenoid separation. Both reverse-phase C_{18} and normal phase silica stationary phases have been used for this purpose. The mobile phases typically used are methanol or acetonitrile containing modifiers such as water or ethyl acetate for reverse phase C_{18} systems, while normal-phase columns use hexane based mobile phases with ethyl acetate as a modifier. In these instances the systems are typically optimized to a specific class of carotenoids for example, normal phase columns are mainly used for the separation of xanthophylls. More recently C_{30} reverse-phase columns have been utilized to profile a range of carotenoids with diverse polarities as well as numerous other isoprenoids such as tocopherols. The C_{30} reverse-phase matrix is also ideal for the separation of geometric isomers (Fraser *et al*., 2000).

c07 BLBK354-Hall October 29, 2010 21:16 Trim: 234mm×156mm Series: APR Char Count=

Modern mass spectrometry has been one of the principal contributing factors to the development of metabolite profiling, however the hydrophobic nature, sensitivity to light, heat, oxygen, acid and in some cases alkali, precludes routine detection of carotenoids by MS due to poor and differential ionization. The number of conjugated double bonds, the nature of the cyclic end groups and oxygen moieties present in the carotenoid molecule give rise to characteristic UV/VIS spectra. The ability of in-line photodiode array detectors (PDA) to record absorbance simultaneously across the whole spectrum makes them ideal for carotenoid identification. In addition, the use of electrochemical array detection is gaining ground exhibiting significantly increased sensitivity for both hydrocarbon (β-carotene and α-carotene) and oxygenated (lutein and zeaxanthin) carotenoids. Co-chromatography and comparison of spectral characteristics with authentic standards enable conclusive identification. Carotenoid standards can in some cases be purchased commercially. It is however, often necessary to purify the compounds from known biological sources, and compare these to their properties documented in the literature. Quantitation of carotenoids separated by HPLC can be achieved by the construction of dose–response curves prepared from authentic standards. For accurate determination, it is advantageous to prepare a curve for each carotenoid and record the chromatographic area at the λ_{max} for each. If an authentic standard is unavailable, a carotenoid with similar chromatographic properties and λ_{max} can be used. Non-endogenous carotenoids can be used as internal standards and relative quantification can be performed. This approach is not as accurate as the use of dose–response curves and recovery can be affected by the matrix. The internal standards are also useful for the normalization of chromatographic retention times. HPLC coupled to PDA detectors is the method of choice when analyzing carotenoid pigments, which overcomes their lack of amenability to routine MS (Fraser *et al*., 2007b). The use of C_{30} reverse phase columns means that a robust profiling method can be used that will identify all pigments within the pathway in a simultaneous chromatographic run. Such approaches have become an ideal means of evaluating changes occurring in a key class of metabolites, essential for conferring quality attributes in foods and have been applied successfully to the GM crops.

7.7.3 Evaluation of novel foodstuffs using metabolomic and chemical fingerprinting

As described above the application of metabolite profiling to food quality typically involves focused analysis of specific classes of compounds. This is not the objective when assessing substantial equivalence of novel crops/foodstuffs. Fundamentally the techniques must be able to detect perturbations in metabolites that are unrelated by intuitive biological knowledge related to intended manipulation. For such analysis metabolomic or chemical

fingerprinting procedures utilizing GC-MS, NMR, MALDI-ToF/MS and DI-MS have been evaluated.

Reports now exist where GM varieties of the food crops tomato (Le Gall *et al.*, 2003), potato (Defernez *et al.*, 2004), pea (Charlton *et al.*, 2004) and wheat (Baker *et al.* 2006) have been assessed for substantial equivalence using a variety of technologies (^{1}H-NMR, GC-MS, DI-MS and MALDI-ToF/MS). Among these techniques, ^{1}H-NMR has been used successfully with potato, pea, tomato and wheat. In these studies less than 50 metabolites were identified and quantified. As outlined earlier, comparative to MS techniques, NMR is less sensitive and has a low resolution, which limits detection of low abundance metabolites. The hardware involved is also expensive and not routine in public analyst laboratories. DI-MS (Catchpole *et al.*, 2005) and MALDI-ToF/MS (Fraser *et al.*, 2007b) have been used to differentiate between GM and non-GM potato and tomato varieties respectively. Potentially these procedures are likely to be used as fingerprinting approaches as identification of *m/z* signals can be ambiguous without incorporation of chromatographic behaviour into the analysis. In addition, quantification can also be affected by ion-suppression if crude extracts are used.

GM tomato (Roessner-Tunali *et al.*, 2003), potato (Catchpole *et al.*, 2005) and wheat (Baker *et al.*, 2006) varieties have all been differentiated from their parent backgrounds and appropriate controls using GC-MS approaches to determine chemical composition. To date the GC-MS analysis provides the most comprehensive coverage of identified metabolites. The compounds identified include sugars, sugar phosphates, organic acids, fatty acids, polyols and some terpenoids, and in total about 120 metabolites can be identified in one chromatographic separation. However, one of the most frustrating aspects is the presence of numerous unknown chromatographic components of which many are metabolites.

Traditionally when evaluating metabolomics data generated on novel foods, multivariate principal component analysis is the method of data analysis routinely used. In virtually all cases the experimental approaches were able to differentiate varieties using PCA scattered plots, clustering individually according to genotype. However, the difference between varieties (e.g. GM and non-GM) was very small. In conclusion, these studies indicated that the overall difference in metabolite composition resulting from the intended manipulation was not greater than the transformation process solely.

More recently, a number of software solutions have been developed that enables the changes in metabolites to be overlaid onto biochemical pathways (www.Biosynlab.com; Thimm, *et al.* 2004). In this way the sectors of metabolism affected can be clearly differentiated. With the advent of more Systems Biology based approaches the potential exists to integrate different omics-based datasets for a given crop variety (Thimm *et al.*, 2004)). This will enable correlation analysis to be performed and eventually a more predictive modelling approach developed for the assessment of substantial equivalence.

7.7.4 Metabolomics in the development and evaluation of GM crops

Genetically modified (GM) crops have tremendous potential to improve the quality of life and reduce environmental impact. For example GM technology can generate crops that require less herbicide and pesticide intervention, reduce water and nutrient usage and contain multiple dietary acquired health promoting chemicals. In the US, GM crops are now well established within their agricultural system. However, the European consumer is presently not prepared to accept foodstuffs produced by the technology, thus preventing commercialization in Europe. The main concerns of GM crops relates to the presence of foreign DNA, environmental/ecological implications (e.g. effect on native species), potential unintended effects on chemical composition, which could lead to elevated or novel toxins and allergenic material as well as altered nutritional content. An alternative to GM breeding is the development of genetically defined breeding populations in which new biodiversity has been introduced. The utilizations of molecular markers with these populations will speed-up conventional breeding and the transfer of QTLs (Quantitative Trait Loci) to elite varieties. However, the regions of DNA introgressed are presently large and this can result in gene drag and associated detrimental traits that may have adverse effects on human health.

Presently, in order for novel foods GM or non-GM, to be accepted into the market place, they must be considered substantially equivalent. The concept of substantial equivalence works upon the characteristics of the novel crop being comparable to an existing food/crop with a history of safe use. The approach has been developed in collaboration with international agencies such as the Organization for Economic Co-ordination and Development (OECD; Anon, 1993) and the United Nations World Health Organization/Food and Agricultural Organization (FAO/WHO, 1991, 2000). The comparator used in the case of GM material is usually the parent background to which genetic manipulation has occurred. Typically three scenarios of substantial equivalence can be considered; (i) the novel food is equivalent to an accepted traditional foodstuff, in which case no further testing is needed, (ii) the novel food is equivalent to the traditional counterpart except for intended differences, in this case safety criteria will be focused on these known differences and (iii) the novel food is different in many respects and there are no known counterparts, in this instance extensive safety assessment will be carried out.

Traditionally the degrees of substantial equivalence are based upon targeted compositional chemical analysis and include major nutrients and toxicants. Concerns have been raised with respect to the targeted (and limited) nature of the chemical analyses used in these evaluations. It is clear that such technologies cannot take into account the possibility of unintended effects resulting directly or indirectly from the action of the transgene inserted or its effects at the biochemical level. Recently metabolite profiling/metabolomic technologies have been evaluated and adopted within the risk assessment

c07 BLBK354-Hall October 29, 2010 21:16 Trim: 234mm×156mm Series: APR Char Count=

of novel foods as a means of evaluating unintended effects on the chemical composition.

7.7.5 Non-targeted approaches and detection of unintended effects

A fuller evaluation of the compositional variation of raw crop plant materials and downstream products will emerge through the development of comparative metabolomics databases that can be expanded and evolved by the international community. This information can be used to benchmark any measured differences between a particular crop against the extent of 'acceptable' variation within the framework of a history of safe use of the crop species in question. There is an ongoing debate over the potential value of much broader scale, more unbiased analytical approaches including metabolomics in risk assessment, which, through the quantity of data they generate, may help to (a) identify effects which could stimulate the need for further risk assessment and (b) reduce the level of uncertainty that unintended effects have occurred. Most of this debate has clearly focused on GM crops but it is already clear that significant natural variation exists within crop gene pools, accentuated by interactions with the prevailing environment.

Metabolomics clearly has much to offer in developing new insights into the regulation of plant metabolism but it must be recognized that the technology has limitations. The plant kingdom may contain between 90,000 and 200,000 metabolites (Dixon & Strack, 2003), although, for a single species, the number may approach a few thousand (the estimate for Arabidopsis is about 5000). Thus full coverage of the metabolome is a real challenge. Data analysis is also challenging as the technologies produce vast datasets. Various data mining approaches are used, e.g. cluster analysis and PCA, to assist the researcher identify non-random patterns that can be further explored (possibly using targeted analytical approaches). A number of initiatives have looked towards developing standards for metabolomics data in addition to a range of technology-specific and general data formats (see Hardy & Taylor, 2007; Davies, 2009 and references therein).

7.8 The future importance of metabolomics in crop research

Metabolomics has emerged to become one of the key tools in all areas of biology, essentially starting with phytochemists then latterly into human diseases, nutrition, drug discovery etc. More recently, as the ability to sequence plant and crop genomes via next generation sequencing has almost become common place and relatively inexpensive (Varshney *et al*., 2009), the requirement to correlate this data with detailed and quantitative pheno(chemo)typic data has become a requirement and has seen a significant ramping up of

complementary 'omics efforts under the banner of systems biology to bridge the genotype-to-phenotype gap (Fiehn, 2002).

Already we are beginning to see these highly detailed complementary analytical approaches being applied in our most common crops such as potato (Lehesranta *et al*., 2005; Shepherd *et al*., 2006; Lehesranta *et al*., 2006; van Dijk *et al*., 2009; Shepherd *et al*., 2010), tomato (Hoekenga, 2008; Matsukura *et al*., 2008; Barone *et al*., 2009; Gavai *et al*., 2009; Plechakova *et al*. 2009; Sánchez Pérez *et al.*, 2009), tomato (Hoekenga, 2008; Matsukura *et al*., 2008; Barone *et al*., 2009; Gavai *et al*., 2009; Plechakova *et al*. 2009; Sánchez Pérez *et al.*, 2009) and to a lesser degree soft fruit such as raspberry (2004 2004, 2009; Mazzitelli *et al*., 2007; Stewart *et al.*, 2007; McDougall *et al*. 2008). The crops addressed by this unified approach will undoubtedly broaden as the approaches become more common place and the technologies, and associated data handling software, more accessible.

Metabolomics *per se* also has a key position in addressing the current and future problems surrounding crop and food production: safety, (enhanced) nutritive value sustainability, food security and climate change. Many of these issues are of general importance and ideally, require broad, multidisciplinary efforts to tackle them. Thankfully, several are already being addressed under international, multi-partner projects. For, example the recently completed projects NOFORISK (Quantitative risk assessment strategies for novel foods; http://www.scri.ac.uk/research/ppfq/foodquality/foodsafety/noforisk) and SAFEFOODS (Promoting Food Safety through a New Integrated Risk Analysis Approach for Foods; http://www.safefoods.nl), both EU-FP6 funded projects, had metabolomics at their core as high throughput and detailed analytical approaches to be assessed as a platform for inclusion as part of a risk assessment process for novel foods (in these cases GMO). The metabolomics effort developed as part of SAFEFOODS was utilized to assist another FP6 project QualityLowInputFood (http://www.qlif.org/) whose aim was to improve quality, ensure safety and reduce cost along the organic and 'low input' food supply chains through research thereby initiating the utility of metabolomics in agricultural sustainability research.

More recently the projects DEVELONUTRI (Development of High Throughput Approaches to Optimise the Nutritional Value of Crops and Crop-Based Foods; http://www.develonutri.info) and META-PHOR (Metabolomic Technology Applications for Plants, Health and Outreach; http://www.meta-phor.eu/), 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. DEVELONUTRI, meanwhile is focused on employing state-of-the-art and emergent metabolomic technologies to potato, tomato and wheat (durum and bread) crop generation and the assessment that the post-harvest processing chain has on nutritive value and the global metabolite pool. Such projects shall prove essential in helping us move forward and

jump to the next level and such technology-driven project clearly can touch on all the hot topics such as food security, nutritive value and food safety.

Of the newer crop metabolomics projects, the EU Interreg IVb project ClimaFruit (Futureproofing the North Sea Berry Industry; www.climafruit.com) is addressing what is seen to be a key issue for the North Sea berry industry: climate change and sustainability. Within this project, metabolomics will be used to elucidate the impact of specific elements of climate change ([CO_2] and temperature) and sustainability (water and nutrient use efficiency, carbon foot print, etc.) on fruit development and quality with a view to feeding this back, with a matched functional genomics effort, into breeding programmes.

In conclusion, metabolomics evidently has gained its place at the centre of crop and food research. As our knowledge of, and ability to apply, metabolomics in these areas increases, the utility of the technology will increase accordingly and we will see it become one of the 'must have' technologies for crops and food research in the near future.

Acknowledgements

RDH acknowledge funding from the EU in the form of the plant metabolomics project META-PHOR (FOOD-CT-2006-036220), part of the EU Framework VI initiative; and from the Netherlands Genomics Initiative (Centre for BioSystems Genomics and the Netherlands Metabolomics Centre). DS, LVTS and PF acknowledge EU funding for DEVELONUTRI (FOOD-CT-2006-036296), EU-SOL (FOOD-CT-2006-016214), NOVORISK (FOOD-CT-2004-506387), ClimaFruit (Interreg IVb-NorthSEa Region Programme), the Scottish Government's Rural and Environment Research and Analysis Directorate (RERAD) and the Biotechnology and Biological Sciences Research Council.

References

Allard, E., Bäckström, D., Danielsson, R. *et al.* (2008) Comparing capillary electrophoresis-mass spectrometry fingerprints of urine samples obtained after intake of coffee, tea, or water. *Analytical Chemistry* **80**, 8946–8955.

Anon (1993) Safety evaluation of foods derived by modern biotechnology: concept and principles. OECD, Organisation for Economic Cooperation and Development. Paris, France.

Anon (2003) USDA Database for the flavonoid content of selected foods. http://www.nal.usda.gov/fnic/foodcomp/Data/Flav/flav.pdf. Accessed on 4 November 2009.

Anon (2008) Wine-omics. *Nature* **455**, 699.

Anon (2009a) USDA compositional data base. http://www.ars.usda.gov/SP2UserFiles/Place/12354500/Data/SR22/reports/sr22fg09.pdf. Accessed on 4 November 2009.

Anon (2009b) Fineli® - Finnish food composition database. http://www.fineli.fi./ <τπ/> Accessed on 4 November 2009.

Anon (2010) The commonwealth potato collection. http://germinate.scri.sari.ac.uk/germinate_cpc/app/index.pl. Accessed on 27 January 2010.

Baker, J.M., Hawkins, N.D., Ward, J.L. *et al.* (2006) A metabolomic study of substantial equivalence of field-grown genetically modified wheat. *Plant Biotechnology Journal* **4**, 381–392.

Barone, A., Di Matteo, A., Carputo, D. *et al.* (2009) High-throughput genomics enhances tomato breeding efficiency. *Current Genomics* **10**, 1–9.

Bauer, D., Biehler, K., Fock, H. *et al.* (1997) A role for cytosolic glutamine synthetase in the remobilization of leaf nitrogen during water stress in tomato. *Physiologia Plantarum* **99**, 241–248.

Beckmann, M., Enot, D.P., Overy, D. *et al.* (2007) Representation, comparison and interpretation of metabolome fingerprint data for total composition analysis and quality trait investigation in potato cultivars. *Journal of Agricultural and Food Chemistry* **55**, 3444–3451.

Beekwilder, J., Hall, R.D. and De Vos, C.H.R. (2005) Identification and dietary relevance of antioxidants in raspberry. *BioFactors* **23**, 197–205.

Beekwilder, J., Van Der Meer, I.M., Simic, A. *et al.* (2008) Metabolism of carotenoids and apocarotenoids during ripening of raspberry fruit. *BioFactors* **33**, 57–66.

Beleggia, R., Platani, C., Spano, G. *et al.* (2009) Metabolic profiling and analysis of volatile composition of durum wheat semolina and pasta. *Journal of Cereal Science* **49**, 301–309.

Bhatnagar, D., Rajasekaran, K., Payne, G.A. *et al.* (2008) The 'omics' tools: genomics, proteomics, metabolomics and their potential for solving the aflatoxin contamination problem. *World Mycotoxin Journal* **1**, 3–12.

Biais, B., Allwood, J.W., Deborde, C. *et al.* (2009a). 1H NMR, GC-EI-TOFMS, and data set correlation for fruit metabolomics: application to spatial metabolite analysis in melon. *Analytical Chemistry* **81**, 2884–2894.

Biais, B., Beauvoit, B., William Allwood, J. *et al.* (2009b). Metabolic acclimation to hypoxia revealed by metabolite gradients in melon fruit. *Journal of Plant Physiology* ~~(in press)~~.

Borém, F.M., Marques, E.R. and Alves, E. (2007) Ultra-structural analysis of drying damage in parchment Arabica coffee endosperm cells. *Biosystems Engineering* **99**, 62–66.

Bovy, A, de Vos, R., Kemper, M. *et al.* (2002) High-flavonol tomatoes resulting from the heterologous expression of the maize transcription factor genes LC and C1. *The Plant Cell* **14**, 2509–2526.

Bovy, A., Schijlen, E.G.W.M. and Hall, R.D. (2007). Metabolic engineering of flavonoids in tomato (*Solanum lycopersicum*): the potential for metabolomics. *Metabolomics* **3**, 399–412.

Bovy, A.G., Gomez-Roldan, V. and Hall, R.D. (2010) Strategies to optimize the flavonoid content of tomato fruit. In: Santos-Buelga, C., Escribano-Bailon, M-T. and Lattanzio, V. (eds.) *Recent Advances in Polyphenol Research. The Handbook of Polyphenols*. pp. 138–162. ~~(In Press).~~

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

Author: Please provide updated status in reference "Biais et al. (2009b)"

Author: Please provide updated status in reference "Bovy et al. (2010)"

Capanoglu, E., Beekwilder, J, Boyacioglu, D. *et al.* (2008) Changes in antioxidant and metabolite profiles during production of tomato paste. *Journal of Agricultural and Food Chemistry* **56**, 964–973.

Capanoglu, E., Beekwilder, J, Boyacioglu, D. *et al.* (2010) The effect of industrial food processing on potentially health-beneficial tomato antioxidants. *Critical Reviews in Food Chemistry* (in press).

Capocasa, F., Scalzo, J., Mezzetti, B. *et al.* (2008) Combining quality and antioxidant attributes in the strawberry: The role of genotype. *Food Chemistry* **111**, 872–878.

Catchpole, G.S., Beckmann, M., Enot, D.P. *et al.* (2005) Hierarchical metabolomics demonstrates substantial compositional similarity between genetically modified and conventional potato crops. *Proceedings of the National Academy of Science* **102**, 14458–14462.

Charlton, A., Allnutt, T. and Holmes, S. (2004). NMR profiling of transgenic peas. *Plant Biotechnology Journal*, **2**, 27–35.

Clifford, M.N. (2000). Anthocyanins – nature, occurrence and dietary burden. *Journal of the Science of Food and Agriculture* **80**, 1063–1072.

D'Archivio, M., Filesi, C., Di Benedetto, R. *et al.* (2007). Polyphenols, dietary sources and bioavailability. *Annual Ist Super Sanita* **43**, 348–361.

Davies, H.V. (2006). Metabolomics: Applications in functional biodiversity analysis in potato. *International Solanaceae Conference and Solanaceae Genomics Network, Madison, Wisconsin*, 23–27 July 2006.

Davies, H.V. (~~2009~~). A role for "omics" technologies in food safety assessment. *Food Control* ~~(in press). DOI:10.1016/j.foodcont.2009.03.002.~~

Defernez, M., Gunning, Y.M., Parr, A.J. *et al.* (2004). NMR and HPLC-UV profiling of potatoes with genetic modifications to metabolic pathways. *Journal of Agricultural and Food Chemistry* **52**, 6075–6085.

de Vos, C.H.R., Borém, F.M., Bouwmeester, H.J. *et al.* (2007) Untargeted metabolomics as a novel tool in coffee research. In: *Proceedings 21st ASIC Conference, Montpellier 2006*. pp. 125–132.

de Vos *et al.* (2010a) Metabolomics as a novel approach to study post-harvest processing of green coffee beans (submitted).

de Vos, C.H.R., Hall, R.D. and Moing, A. (2010b) Metabolomics of a model fruit: tomato. In: Hall, R.D. (ed.) *Biology of Plant Metabolomics*. Wiley Blackwell.

del Rio, D., Stewart, A.J. and Mullen, W. (2004) HPLC-MS^n analysis of phenolic compounds and purine alkaloids in green and black tea. *Journal of Agricultural and Food Chemistry* **52**, 2807–2815.

Dixon, R.A. and Strack, D. (2003) Phytochemistry meets genome analysis, and beyond. *Phytochemistry* **62**, 815–816.

Dobson, G., Shepherd, T., Verrall, S.R. *et al.* (2008). Phytochemical diversity in tubers of potato cultivars and landraces using a GC-MS metabolomics approach. *Journal of Agricultural and Food Chemistry* **56**, 10280–10291.

Dobson, G., Shepherd, T., Verrall, S.R. *et al.* (2010). A metabolomics study of cultivated potato (solanum tuberosum) groups andigena, phureja, stenotomum, and tuberosum using gas chromatography−mass spectrometry. *Journal of Agricultural and Food Chemistry* ~~(In Press)~~

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

Author: Please provide updated status in reference "Capanoglu et al. (2010)"

Author: Please provide updated status in reference "Davies (2009)"

Author: Please provide the updated status in reference "de Vos et al. (2010a)"

Author: Please provide publisher location in reference "de Vos et al. (2010b)"

Author: Please provide updated status in reference "Dobson et al. (2010)"

FAO/WHO (Food and Agriculture Organization/World Health Organization) (2000). *Safety Aspects of Genetically Modified Foods of Plant Origin*. Geneva, Switzerland.

Fernie, A.E. and Schauer, N. (2009) Metabolomics-assisted breeding: a viable option for crop improvement? *Trends in Genetics* **25**, 39–48.

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

Fitzgerald, M.E. and Hall, R.D. (2008) Illuminating the gap. *Rice Today*, April–June; 38–39.

Fitzgerald, M.E., McCouch, S. and Hall, R.D. (2009) Not just a grain of rice: the quest for quality. *Trends in Plant Science* **14**, 133–138.

Frank, T., Meuleye Seumo, B., Miller, A., Shu, Q.Y. and Engel, K.-H. (2007) Metabolite profiling of two low phytic acid (*lpa*) rice mutants. *Journal of Agricultural and Food Chemistry* **55**, 11011–11019.

Frank, T., Nörenberg, S. and Engel, K.-H. (2009) Metabolite profiling of two novel low phytic acid (*lpa*) soybean mutants. *Journal of Agricultural and Food Chemistry* **57**, 6408–6416.

Fraser, P.D. and Bramley, P.M. (2004) The biosynthesis and nutritional uses of carotenoids. *Progress in Lipid Research* **43**, 228–265.

Fraser, P.D., Pinto, M.E.S., Holloway, D.E. and Bramley, P.M. (2000) Application of high-performance liquid chromatography with photodiode array detection to the metabolic profiling of plant isoprenoids. *Plant Journal* **24**, 551–558.

Fraser, P.D., Enfissi, E.M.A., Halket, J.M. *et al.* (2007a) Manipulation of phytoene levels in tomato fruit: effects on isoprenoids, plastids and intermediary metabolism. *Plant Cell* **19**, 3194–3211.

Fraser, P.D., Enfissi, E.M.A., Goodfellow, M. *et al.* (2007b) Metabolite profiling of plant carotenoids using the matrix-assisted laser desorption ionization time-of-flight mass spectrometry. *The Plant Journal* **49**, 552–564.

Fraser, P.D., Enfissi, E.M.A. and Bramley, P.M. (2009) Genetic engineering of carotenoid formation in tomato fruit and the potential application of systems and synthetic biology approaches. *Archives of Biochemistry and Biophysics* **483**, 196–204.

Funayama, S., Zhang, G. and Nozoe, S. (1995) Kukoamine B, a spermine alkaloid from Lycium chinense. *Phytochemistry* **38**, 1529–1531.

Garris, A.J., Tai, T.H., Coburn, J. *et al.* (2005) Genetic structure and diversity in *Oryza sativa* L. *Genetics* **169**, 1631–1638.

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

Gonzali, S., Mazzucatom A. and Perata, P. (2009) Purple as a tomato: towards high anthocyanin tomatoes. *Trends in Plant Science* **14**, 237–241.

Goodacre, R., Vaidyanathan, S., Dunn, W.B. *et al.* (2003) Chemometric discrimination of unfractionated plant extracts analysed by electrospray mass spectrometry. *Phytochemistry* **62**, 859–863.

Graham, J., Smith, K., MacKenzie, K. *et al.* (2004) The construction of a genetic linkage map of red raspberry (Rubus idaeus subsp. idaeus) based on AFLPs, genomic-SSR and EST-SSR markers. *Theoretical and Applied Genetics* **109**, 740–749.

Graham, J., Hackett, C.A., Smith, K. *et al.* (2009) Mapping QTLs for developmental traits in raspberry from bud break to ripe fruit. *Theoretical and Applied Genetics* **118**, 1143–1155.

Grennan, A.K. (2009) MoTo DB: a metabolic database for tomato. *Plant Physiology* **151**, 1701–1702.

Author: Please provide name of the publisher in refernece "FAO/WHO (Food and Agriculture-Organization/World Health Organization) (2000)"

Author: Please provide the updated status in reference "Hall et al. (2010)"

Grün, C.H., van Dorsten, F.A., Jacobs, D.M. *et al.* (2008) GC-MS methods for metabolic profiling of microbial fermentation products of dietary polyphenols in human and in vitro intervention studies. *Journal of Chromatography B* **871**, 212–219.

Hall, R.D. (2006a) Plant metabolomics: from holistic hope, to hype, to hot topic. *New Phytologist Tansley Review* **169**, 453–468.

Hall, R.D. (2006b) Food metabolomics: META-PHOR. A new European research initiative. *AgroFOOD Industry Hi-Tech* **18**, 13–16.

Hall, R.D., De Vos, R.C.H. and Ward, J.L. (2010) Plant metabolomics applications in the Brassicaceae: added value for science and industry. *Acta Horticulturae* (in press).

Hall, R.D., de Vos, C.H.R., Verhoeven, H.A. *et al.* (2005) Metabolomics for the assessment of functional diversity and quality traits in plants. In: Vaidyanathan, S., Harrigan, G.G. and Goodacre, R. (eds.) *Metabolic Profiling*. Kluwer Acad. Publishers, Dordrecht, NL, pp. 31–44.

Hall, R.D., Brouwer, I.D. and Fitzgerald, M.A. (2008) Plant metabolomics and its potential application for human nutrition. *Physiologia Plantarum* **132**, 162–175.

Hardy, N.W. and Taylor, C.F. (2007) A roadmap for the establishment of standard data exchange structures for metabolomics. *Metabolomics* **3**, 243–248.

Harrigan, G.G., Stork, L.G., Riordan, S.G. *et al.* (2007a) Metabolite analyses of grain from maize hybrids grown in the United States under drought and watered conditions during the 2002 field season. *Journal of Agricultural and Food Chemistry* **55**, 6169–6176.

Harrigan, G.G., Stork, L.G., Riordan, S.G. *et al.* (2007b) Impact of genetics and environment on nutritional and metabolite components of maize grain. *Journal of Agricultural and Food Chemistry* **55**, 6177–6185.

Hedberg, C., Hesse, M. and Werner, C. (1996) Spermine and spermidine hydroxycinnamoyl transferases in *Aphelandra tetragona*. *Plant Science* **113**, 149–156.

Hoekenga, O.A. (2008) Using metabolomics to estimate unintended effects in transgenic crop plants: problems, promises, and opportunities. *Journal of Biomolecular Technology* **19**, 159–166.

Ikeda, T., Kanaya, S., Yonetani, T. *et al.* (2007) Prediction of Japanese green tea ranking by Fourier Transform Near-InfraRed reflectance spectroscopy. *Journal of Agricultural and Food Chemistry* **55**, 9908–9912.

Joët, T., Laffargue, A., Salmona, J. *et al.* (2009) Metabolic pathways in tropical dicotyledonous albuminous seeds: *Coffea arabica* as a case study. *New Phytologist* **182**, 146–162.

Jones, C.M., Mes, P. and Myers, J.R. (2003) Characterisation and inheritance of the *Anthocyanin fruit (Aft)* tomato. *Journal Heredity* **94**, 449–456.

Keutgen, A.J. and Pawelzik, E. (2008) Contribution of amino acids to strawberry fruit quality and their relevance as stress indicators under NaCl salinity. *Food Chemistry* **111**, 642–647.

Kind, T., Scholz, M., Fiehn, O. (2009) How large is the metabolome? A critical analysis of data exchange practices in chemistry. *PLoS ONE* **4**(5), e5440.

Kusano, M., Fukushima, A., Kobayashi, M. *et al.* (2007) 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. *Journal of Chromatography B* **855**, 71–79.

Le Gall, G., Colquhoun, I.J., Davis, A.L. *et al.* (2003) Metabolite profiling of tomato (Lycopersicon esculentum) using 1H NMR spectroscopy as a tool to detect potential unintended effects following a genetic modification. *Journal of Agricultural and Food Chemistry* **51**, 2447–2456.

Le Gall, G., Colquohoun, I.J., Defernez, M. (2004) Metabolite profiling using ^{1}H NMR spectroscopy for quality assessment of green tea, *Camellia sinensis* L. *Journal of Agricultural and Food Chemistry* **52**, 692–700.

Lee, J-E., Hwang, G-S., Van Den Berg, F. *et al.* (2009) Evidence of vintage effects on grape wines using ^{1}H NMR – based metabolomic study. *Analytica Chimica Acta* **648**, 71–76.

Lehesranta, S.J., Davies, H.V., Shepherd, L.V. *et al.* (2005) Comparison of tuber proteomes of potato varieties, landraces, and genetically modified lines. *Plant Physiology* **138**, 1690–1699.

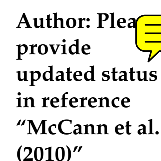
Lehesranta, S.J., Davies, H.V., Shepherd, L.V. *et al.* (2006) Proteomic analysis of the potato tuber life cycle. *Proteomics* **6**, 6042–6052.

Levine, L.H., Kasahara, H., Kopka, J. *et al.* (2008) Physiologic and metabolic responses of wheat seedlings to elevated and super-elevated carbon dioxide. *Advances in Space Research* **42** 1917–1928.

Lindinger, C., Pollien, P., de Vos, R.C.H. *et al.* (2010) Identification of a quality marker of the fermented off-note in coffee by a non-targeted chemometric approach. *Journal of Agricultural and Food Chemistry* **57**, 9972–9978.

Mathews, H., Clendennen, S.K., Caldwell, S.K. *et al.* (2003). Activation tagging in tomato identifies a transcriptional regulator of anthocyanins biosynthesis, modification, and transport. *Plant Cell* **15**, 1689–1703.

Author: Plea provide updated status in reference "McCann et al. (2010)"



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

Mazzitelli, L., Hancock, R.D., Haupt, S. *et al.* (2007) Co-ordinated gene expression during phases of dormancy release in raspberry (*Rubus idaeus* L.) buds. *Journal of Experimental Botany* **58**, 1035–1045.

McCann, L.C., Bethke, P.C., Simon, P.W. (2010) Extensive variation in fried chip color and tuber composition in cold-stored tubers of wild potato (solanum) germplasm. *Journal of Agricultural and Food Chemistry* (In Press).

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

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

Mitchell, S.C. (2001) Food idiosyncrasies: beetroot and asparagus. *Drug Metabolism and Disposition* **29**, 539–543.

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

Moco, S., Forshed, J., De Vos, R.C.H. *et al.* (2008) Intra- and inter-metabolite correlation spectroscopy of tomato metabolomics data obtained by liquid chromatography-mass spectrometry and nuclear magnetic resonance. *Metabolomics* **4**, 202–215.

Mullen, W., Marks, S.C. and Crozier, A. (2007) Evaluation of phenolic compounds in commercial fruit juices and fruit drinks. *Journal of Agricultural and Food Chemistry* **55**, 3148–3157.

Noteborn, H.P.J.M., Lommen, A., van der Jagt, R.C. *et al.* (2000) Chemical fingerprinting for the evaluation of unintended secondary metabolic changes in transgenic food crops. *Journal of Biotechnology* **77**, 103–114.

Obert, J.C., Ridley, P.W., Schneider, R.W. *et al.* (2004). The composition of grain and forage from glyphosate tolerant wheat MON 71800 is equivalent to that of conventional wheat (Triticum aestivum L.). *Journal of Agriculture and Food Chemistry* **52**, 1375–1384.

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

Panico, A.M., Garufi, F., Nitto, S. *et al.* (2009). Antioxidant activity and phenolic content of strawberry genotypes from Fragaria x ananassa. *Pharmacological Biology* **47**, 203–208.

Parr, A., Mellon, F., Colquhoun, I. *et al.* (2005). Dihydrocaffeoyl polyamines (kukoamine and allies) in potato (*Solanum tuberosum*) tubers detected during metabolite profiling. *Journal of Agricultural and Food Chemistry* **53**, 5461–5466.

Periera, G.E., Gaudillere, J-P., van Leeuwen, C. *et al.* (2005) ^{1}H NMR and chemometrics to characterise mature grape berries in four wine – growing areas in Bordeaux, France. *Journal of Agricultural and Food Chemistry* **53**, 6382–6389.

Pietta, P., Minoggio, M. and Bramati, L. (2003) Plant polyphenols: Structure, occurrence and bioactivity. *Studies in Natural Products Chemistry* **28**, 257–312.

Plechakova, O., Tranchant-Dubreuil, C., Benedet, F. *et al.* (2009) MoccaDB - an integrative database for functional, comparative and diversity studies in the Rubiaceae family. *BMC Plant Biology* **29**, 123.

Pongsuwan, W., Fukusaki, E., Bama, T. *et al.* (2007) Prediction of Japanese green tea ranking y gas chromatography/mass spectrometry-based hydrophilic metabolite fingerprinting. *Journal of Agricultural and Food Chemistry* **55**, 231–236.

Pongsuwan, W., Bamba, T., Harada, K. *et al.* (2008) High-throughput technique for comprehensive analysis of Japanese green tea quality assessment using ultra-performance liquid chromatography with time-of-flight mass spectrometry (UPLC/TOF MS). *Journal of Agricultural and Food Chemistry* **56**, 10705–10708.

Prior, R.L. and Wu, X. (2006) Anthocyanins: structural characteristics that result in unique metabolic patterns and biological activities. *Free Radical Research* **40**, 1014–1028.

Raboy, V. (2007) Seed phosphorus and the development of low-phytate crops. In: Turner, B.L., Richardson, A.E., Mullaney, E.J. (eds.) *Inositol Phosphates. Linking Agriculture and the Environment*, pp. 111–132. CAB International, Wallingford, Oxfordshire, UK.

Reed, J.D., Krueger, C.G. and Vestling, M.M. (2005) MALDI-TOF mass spectrometry of oligomeric food polyphenols. *Phytochemistry* **66**, 2248–2263.

Reid, L.M., O'Donnell, C.P. and Downey, G. (2004) Potential of SPME-GC and chemometrics to detect adulteration of soft fruit purées. *Journal of Agricultural and Food Chemistry* **52**, 421–427.

Rein, D., Schijlen, E., Kooistra, T. *et al.* (2006) Transgenic flavonoid tomato intake reduces C-reactive protein in human C-reastive protein transgenic mice more than wild-type tomato. *The Journal of Nutrition* **136**, 2331–2337.

Reynolds, T.L., Nemeth, M.A., Glenn, K.C., Ridley, W.P., Astwood, J.D. (2005) Natural variability of metabolites in maize grain: differences due to genetic background. *Journal of Agricultural and Food Chemistry* **53**, 10061–10067.

Ridley, W.P., Shillito, R.D., Coats, I. et al.. (2004) Development of the International Life Sciences Institute Crop Composition Database. *Journal of Food Composition and Analysis* **17**, 423–438.

Rinke, P. (2008) Successful application of SGF-profiling. *New Food* **1**, 18–23.

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

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

Röhlig, R.M., Eder, J. Engel, K.-H. (2009) Metabolite profiling of maize grain: differentiation due to genetics and environment. *Metabolomics*, DOI: 10.1007/s11306-009-0171-5.

Sánchez Pérez, I., Culzoni, M.J. *et al.* (2009) Detection of unintended stress effects based on a metabonomic study in tomato fruits after treatment with carbofuran pesticide. Capabilities of MCR-ALS applied to LC-MS three-way data arrays. *Analytical Chemistry* **81**, 8335–8346.

Sato, S., Soga, T., Nishioka, T. *et al.* (2004) Simultaneous determination of the main metabolites in rice leaves using capillary electrophoresis mass spectrometry and capillary electrophoresis diode array detection. *The Plant Journal* **40**, 151–163.

Sato, S., Arita, M., Soga, T. *et al.* (2008) Time-resolved metabolomics reveals metabolic modulation in rice foliage. *BMC Systems Biology* **2**, 51.

Sattar, E.A., Glasl, H., Nahrstedt, A. *et al.* (1990) Hydroxycinnamic acid amides from *Iochroma cyaneum*. *Phytochemistry* **29**, 3931–3933.

Schaneberg, B.T., Crockett, S., Bedir, E. *et al.* (2003) The role of chemical fingerprinting application to *Ephedra*. *Phytochemistry* **62**, 911–918.

Schauer, N., Zamir, D., Fernie, A.R. (2005) Metabolic profiling of leaves and fruit of wild species tomato: a survey of the Solanum lycopersicum complex. *Journal of Experimental Botany* **56**, 297–307.

Schijlen, E.G.W.M., Beekwilder, J., Hall, R.D. *et al.* (2008) Boosting beneficial phytochemicals in vegetable crop plants. *CAB Reviews* **3**(25), 1–21.

Schütz, B., Humpfer, E., Mörtter, M. *et al.* (2008) SGF Profiling™ – NMR-based fruit juice quality control. www.brucker.com/sgf.

Seebauer, J.R., Moose, S.P., Fabbri, B.J. *et al.* (2004) Amino acid metabolism in maize earshoots. Implications for assimilate preconditioning and nitrogen signaling. *Plant Physiology* **136**, 4326–4334.

Selmar, D. and Bytof, G. (2007) A review on the metabolic processes taking place in coffee beans during processing and their implication for modern coffee research. In: *Proceedings 21st ASIC Conference*, Montpellier, 2006, pp. 423–433.

Shepherd, L.V., McNicol, J.W., Razzo, R. *et al.* (2006) Assessing the potential for unintended effects in genetically modified potatoes perturbed in metabolic and developmental processes. Targeted analysis of key nutrients and anti-nutrients. *Transgenic Research* **15**, 409–425.

Shepherd, L.V.T., Alexander, C.A., Sungurtas, J.A. *et al.* (2010) Metabolomic analysis of the potato tuber life cycle. *Metabolomics* ~~(in press), DOI: 10.1007/s11306-009-0183-1 (online).~~

Shu, X-L., Frank, T., Shu, Q-Y. *et al.* (2008) Metabolite profiling of germinating rice seeds. *Journal of Agricultural and Food Chemistry* **56**, 11612–11620.

Author: Please provide updated status in reference "Shepherd et al. (2010)"

Son, H-S., Kim, K.M., van den Berg, F. *et al.* (2008) ^{1}H NMR – based metabolomic characterisation of wines by grape varieties and production areas. *Journal of Agricultural and Food Chemistry* **56**, 8007–8016.

Son, H-S., Hwang, G-S., Kim, K.M. *et al.* (2009) Metabolomic studies on geographical grapes and their wines using ^{1}H NMR analysis coupled with multivariate statistics. *Journal of Agricultural and Food Chemistry* **57**, 1481–1490.

Stalmach, A., Mullen, W., Barron, D. *et al.* (2009) Metabolite profiling of hydroxycinnamate derivatives in plasma and urine after the ingestion of coffee by humans: identification of biomarkers of coffee consumption. *Drug Metabolism and Disposition* **37**, 1749–1758.

Stewart, D., McDougall, G.J., Sungurtas, J. *et al.* (2007) Metabolomic approach to identifying bioactive compounds in berries: advances toward fruit nutritional enhancement. *Molecular Nutrition and Food Research* **51**, 645–651.

Tarachiwin, L., Ute, K., Kobayashi, A. *et al.* (2007) ^{1}H NMR based metabolomic profiling in the evaluation of Japanese tea quality. *Journal of Agricultural and Food Chemistry* **55**, 9330–9336.

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

Thomas, J., Kumar, R.R. and Mandal, A.K.A. (2006) Metabolite profiling and characterization of somaclonal variants in tea (*Camellia* spp.) for identifying productive and quality accession. *Phytochemistry* **67**, 1136–1142.

Tikunov, Y., Lommen, A. and de Vos, C.H.R. *et al.* (2005) A novel approach for nontargeted data analysis for metabolomics: large-scale profiling of tomato fruit volatiles. *Plant Physiology* **139**, 1125–1137.

Tikunov, Y., de Vos, R.C.H., Gonzalez-Paramas, A.M. *et al.* (2010) A role for differential glycoconjugation in the emission of phenylpropanoid volatiles from tomato fruit discovered using a metabolic data fusion approach. *Plant Physiology* (in press).

Ursem, R., Tikunov, Y., Bovy, A. *et al.* (2008) A correlation network approach to metabolic data analysis for tomato fruits. Euphytica, **161**, 181–193.

van den Heuvel, T., Renes, R.J., Gremmen, B. *et al.* (2008) Consumers' images regarding genomics as a tomato breeding technology: "maybe it can provide a more tasty tomato". *Euphytica* **159**, 207–216.

van Dijk, J.P., Cankar, K., Scheffer, S.J. *et al.* (2009) Transcriptome analysis of potato tubers–effects of different agricultural practices. *Journal of Agricultural and Food Chemistry* **25**, 1612–1623.

van Dorsten, F.A., Daykin, C.A., Mulder, T.P.J. *et al.* (2006) Metabolomics approach to determine metabolic differences between green tea and black tea consumption. *Journal of Agricultural and Food Chemistry* **54**, 6929–6938.

Varshney, R.K., Nayak, S.N., May, G.D. *et al.* (2009) Next-generation sequencing technologies and their implications for crop genetics and breeding. *Trends in Biotechnology* **27**, 522–530.

Verhoeven, H.A., Jonker, H., de Vos, R.C.H. *et al.* (2010) Solid phase micro-extraction (SPME) GC - MS analysis of natural volatile components in melon and rice. In: Hardy, N.G. and Hall, R.D. (eds.) *Methods for Plant Metabolomics*. Humana Press (in press).

Wang, S.Y, Bunce, J.A. and Maas, J.L. (2003) Elevated carbon dioxide increases contents of antioxidant compounds in field-grown strawberries. *Journal of Agricultural and Food Chemistry* **51**, 4315–4320.

Author: Please provide updated status in reference "Tikunov et al. (2010)"

Author: Please provide updated status in reference "Verhoeven et al. (2010)"

Xie, D.Y. and Dixon, R.A. (2005) Proanthocyanidin biosynthesis–still more questions than answers? *Phytochemistry* **66**, 2127–2144.

Zhou, J., Ma, C., Xu, H. *et al.* (2009) Metabolic profiling of transgenic rice with *cryIAc* and *sck* genes: an evaluation of unintended effects at metabolic level by using GC-FID and GC-MS. *Journal of Chromatography B* **877**, 725–732.