Challenges for Metabolomics as a Tool in Safety Assessments

George G. Harrigan and Bruce Chassy

Regulatory Product Characterization and Safety Center, Monsanto Company, St. Louis,
USA

1. Introduction

Agriculture's ability to supply an abundance of nutritious foods and feeds to nourish the world's growing population faces serious challenges (Foresight, 2011). In order to meet these challenges, plant breeders will be required to continuously improve agricultural productivity as well as enhance food and feed quality. In recent years, the development of methods for the direct introduction of new traits to produce transgenic varieties – also known as GM crops – has proven to be a powerful tool in the hands of breeders. In most countries, however, GM crops are subjected to rigorous pre-market regulatory assessments that require numerous laboratory and field studies and which consume time and resources (Kalaitzandonakes et al., 2007).

Comprehensive compositional analyses represent a key component of the pre-market safety evaluations of GM crops (Harrigan, et al., 2010). These analyses typically include the measurement of levels of key nutrients such as protein, storage oil, fiber, amino acids, fatty acids, vitamins, as well as crop-specific metabolites such as gossypol and cyclopropenoid fatty acids in cotton or isoflavones in soybean. The Organization of Economic Cooperation and Development (OECD) has produced a series of consensus documents that identify key analytes in a number of major crop varieties (http://www.oecd.org). These documents carefully review the composition and uses for each crop and identify those components that contribute to nutritional or functional food or feed value as well as components that might confer health-beneficial, health-protective, or harmful effects (e.g. allergens, anti-nutrients, and potential toxicants). The large-scale compositional studies performed as part of regulatory assessments must follow internationally accepted guidelines. These are outlined in detail by Codex Alimentarius (Codex Alimentarius, 2008) and OECD. In most cases, these studies are typically conducted under Good Laboratories Practice (GLP), a practice that places a high premium on documentation and reconstructability of data, method validation and personnel training, and a requirement for professionally staffed Quality Assurance Units.

The fact that different crops produce foods or feeds with differing compositions, along with the fact that human and animal diets vary greatly in their consumption of these crops, means that each crop plays a unique role in diet and health. Most plant foods in the human diet make significant contributions to the total intake of just a few macro- and

micronutrients (Senti and Rizek, 1974; Chassy, 2010). It is therefore important to assure that no changes have occurred that would lower the dietary intake of an essential nutrient; on the other hand, large changes in the content of one or more nutrients in a crop that supplies an infrequently consumed food, one which is consumed in small amounts in the diet, or one which is not an important source of that nutrient in the diet, are of no health consequence and will have no adverse effect on health (Chassy, 2010).

The identification and analysis of a key set of relevant metabolites is often referred to a "targeted" compositional analysis. Analyses utilize quantitative assays and the overall approach allows the generation of data that is easily interpretable from a nutrition and food/feed safety aspect. Furthermore, since the small molecule metabolite pool in seed is of low abundance relative to macromolecular components, measurement of macronutrients approximates the total seed biomass. For example, the small molecule metabolite pool in corn grain is only ~5% of the total biomass (corn is dominated by starch, fiber, protein, and fat). Anti-nutrient components in grain such as phytic acid and raffinose (which represent much of the small molecule metabolite pool) are measured in regulatory assessments. Other small molecules metabolites can be included if they are an intended endpoint of compositional or nutritional modification. Otherwise analytical measurement of the metabolites that constitute this pool, mainly ubiquitous free amino acids, sugars, and organic acids), is of little value owing to the extreme sensitivity of metabolite levels to environmental influences and the negligible contribution they make to safety and nutritional content (Herman et al., 2009; Skogerson et al., 2010, Harrigan et al., 2007).

In fact, levels of all crop compositional components are influenced markedly by environment (Harrison and Harrigan, 2011; Harrigan, et al., 2010; Zhou et al., 2011a, 2011b). To illustrate, as far back as 1983, it was noted that "The concentration of the isoflavones vary from [soybean] variety to variety, and there are also differences when the same variety is grown in different locations" (Eldridge and Kwolek, 1983). Given the extensive scientific literature on isoflavone variability, it was unsurprising that Gutierrez-Gonzalez et al. (2009) recently concluded that "The range of values of isoflavones is overwhelming, even for homozygous genotypes growing in the same year and location, which greatly complicates genetic studies." This is true for almost all crop compositional components as evidenced by challenges in enhancing nutritional quality in staple crops through conventional approaches. Figure 1 illustrates the type of variability than can be observed for metabolites such as isoflavones.

The use of multiple geographically separate sites is required in regulatory assessments to allow compositional studies across a wide range of environmental conditions. Indeed, information on compositional variation in conventional crops with respect to their responsiveness to environmental factors is necessary to provide context to evaluations of new GM crops. Studies incorporating four to five replicated field sites utilizing randomized complete block designs with three blocks per comparator are typical in regulatory assessments, although the European Food Safety Authority (EFSA) has recently mandated a minimum of eight replicated sites utilizing randomized complete block designs with four blocks (EFSA, 2011).
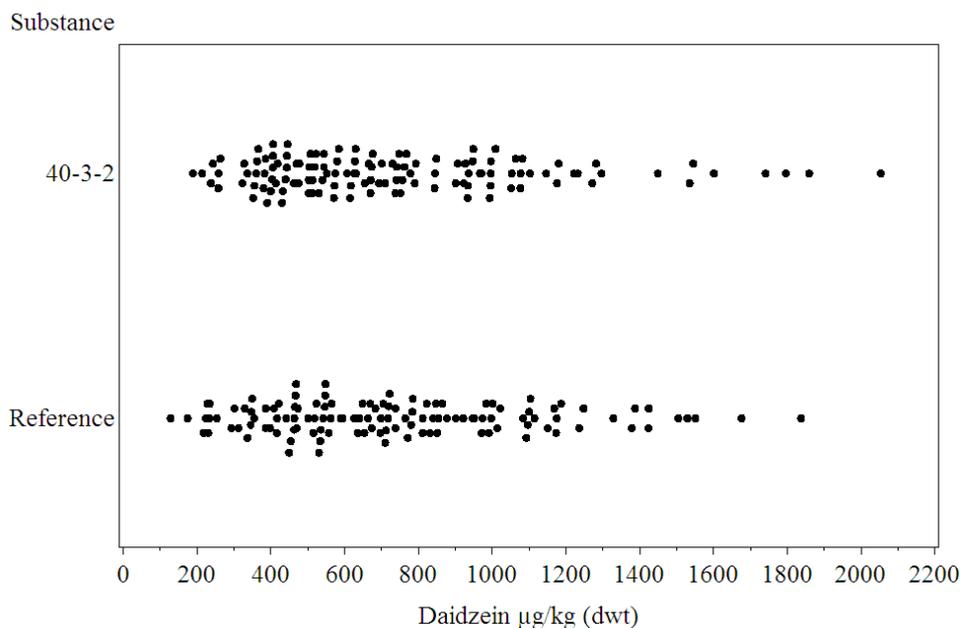
Substance

Fig. 1. An overview of variability in isoflavone levels. Datapoints show daidzein values from an analysis of GM (40-3-2) and conventional reference comparators from a total of nine (2001-2009) growing seasons. A total of 112 unique GM varieties were assessed (Zhou et al., 2011b). This type of information presents context to any GM-non-GM pairwise comparison and would be a required component of any metabolomic assessments.

Results to date from these large-scale compositional studies have generally demonstrated that the effect of transgene insertion is significantly less than the impact of environmental or germplasm variation on conventional crops (Harrigan et al., 2010). This has allowed some to question the relevance and design of compositional assessments. One review, for example, suggests that "the current complexity and resource requirements for compositional studies on transgenic crops containing input traits are not justified by a commensurate understanding of safety" (Herman et al., 2009).

Despite continued confirmation that conventional breeding and environmental variation contribute to compositional variability more so than transgene insertion (Ricroch et al., 2011), and the resource-intensiveness of the large-scale studies currently required for regulatory assessments, there remains some interest in the application of profiling technologies to compare GM and conventional crops. These are often posited in terms of "gap-filling" (Heineman et al., 2011) or "case-by-case" (Davies, 2010) evaluations. It is also perceived by many (e.g., Kok et al., 2008) that measurement of "primarily the low-molecular weight molecules" is more relevant to safety than proteomic or transcriptomic profiling due to a closer relationship to "the plant phenotype and nutritional and toxicological characteristics". This potential advantage of metabolic profiling could be extended as an improvement over, for example, measurements of gross levels of protein, fat, and fibers, key nutritional but essentially "safe and inert" components of food. It is

noteworthy that Kok et al. (2008) define metabolomics as the "generation of profiles of secondary metabolites" whereas most metabolic profiling experiments to date have focused on primary metabolites. It has also been suggested that untargeted profiling techniques are unbiased while "targeted" compositional analysis is biased. Finally, advocates of metabolomic profiling have suggested that such an approach can detect potentially deleterious totally novel metabolites that would have been missed by "targeted" analysis, although it should be noted that many profiling technologies require standards of known identity to accurately identify and measure specific metabolites thus limiting this potential advantage. In addition, in examples where a new traditionally bred plant variety has caused toxic effects, this has been attributable to increased levels of well-known toxicants (Chassy, 2010).

Profiling technologies have confirmed on a case-by-case basis the compositional "equivalence" of GM crops to their conventional near-isogenic comparators (Ricroch et al., 2011). Profiling technologies are, however, unlikely to provide immediately interpretable data in safety assessments that would provide added value to, or otherwise enhance, rigorously quantitative assessments of known nutrients and anti-nutrients that comprise foodstuffs. In the case of metabolic profiling, this can be directly attributable to i) the intrinsically safe nature of food itself, ii) inconsistencies in metabolite coverage versus quantitative capabilities afforded by different data acquisition technologies, iii) the ubiquitous and innocuous nature of small molecule metabolites identified in profiling as well as extreme variability in metabolite levels even within homozygous genotypes, and iv) the "chasm" between the large number of data generated in profiling experiments and the ability to interpret them in a way that is meaningful to nutrition and food safety. We now expand on these observations and further emphasize that a clear distinction between "substantial equivalence" and food safety should be promoted.

2. Key challenges for the omics

1. *Domesticated crops have been selected to serve human needs and have an extensive history of safe consumption. Extensive information on levels of nutrients and crop-specific antinutrients is available. These can be measured through highly quantitative assays to provide interpretable data of direct relevance to food nutrition and safety.*

Of over 250,000 plant species, only 7000 are considered as foodstuffs (Khoshbahkt and Hammer, 2008), and even fewer, 150, supply over 90% of all plant food. Three major crops, *i.e.* maize, wheat, and rice, supply over 66%. Choices made in crop domestication and breeding have enabled food and feed qualities that serve human needs. Numerous path changes between wild and domesticated plants are known and include *e.g.* loss of spontaneous shattering of seed head on ripening and greater uniformity of seed ripening and germination, both of which facilitate human harvest. A key path change however is reduction and even loss of toxic compounds in the edible parts of domesticated crops. In other words, human selection has resulted in crop phenotypes and compositions that distinguish domesticated varieties from their natural counterparts, are more suited to human diets and needs, and are safer and more nutritious. Interestingly, one review (Jones, 1998) asked "Why are so many food plants cyanogenic?" and concluded that "Cyanogenesis by plants is not only a surprisingly effective chemical defence against casual herbivores, but it is also easily overcome by careful pre-ingestion food processing, this latter skill being

almost exclusive to humans." In other words, because "cyanogenic plants are surprisingly well protected from herbivory and yet can be readily detoxified by food processing, … early farmers fortuitously chose these plants above all others for cultivation."

Of course, many modern foodstuffs are still associated with "ancestral" secondary metabolites that may confer nutritional or safety concerns at elevated levels. Classic examples include glycoalkaloids in potato (NIEHS, 1998), β-N-oxalyl-L-α,β-diaminopropionic acid (ODAP) in *Lathyrus sativus* (Bell, 2003), psoralens in celery (Beier and Oertli, 1983), and gossypol in cotton (Sunlkumar et al., 2006). Targeted measurement of these components as opposed to broad-based compositional screening is recommended by Herman et al. (2009); in other words, compositional assessments should focus on molecules explicitly associated with safety concerns. This is consistent with the observation that in the very few examples where a new plant variety has caused toxic effects it has been attributable to well-know toxicants associated with conventionally bred crops and not to a hitherto undetected metabolite (Chassy, 2010).

It is noteworthy that such targeted assessments could easily facilitate a partnership with omics researchers conducting semi-targeted profiling on pathways associated with toxic metabolites to support both early development and commercialization of nutritionally enhanced products. Such a partnership could, at least in principle, mitigate the current regulatory burden imposed on new GM crops (Graff et al., 2009; Potrykus, 2010) and promote the application of omics within modern agricultural biotechnology.

2. *Information on compositional variation in conventional crops with respect to their responsiveness to environmental and genetic factors is necessary to provide context to evaluations of new GM crops. The need to assess natural variation is also true for metabolomics yet little information on the impact of conventional breeding on metabolite profiles is available. The inconsistent coverage of metabolites offered through different data acquisition platforms may provide challenges in establishing a coherent literature in this area.*

Ironically, as mentioned earlier, continued confirmation that conventional breeding, environment, and germplasm contribute to compositional variation more than transgene insertion has coincided with increased interest in the use of 'omics technologies. This paradox is compounded by the fact that results from these technologies have only further highlighted the equivalence of GM crops to their conventional counterparts and reaffirmed the substantial effect of environment and germplasm on compositional and biochemical variability (see Ricroch et al., 2011). Although there are complexities in the interpretation of data generated through modern profiling technologies (Broadhurst and Kell, 2006; Lay et al., 2006) including the fact that the data is not quantitative and there is no standardized framework for comparisons, the lack of variation between GM crops and their conventional comparators at the transcriptomic, proteomic, and metabolomic level has, nonetheless, been independently corroborated. These profiling evaluations extend to a wide range of plants including wheat (Baker et al., 2006; Gregersen et al., 2005; Ioset et al., 2007), potato (Catchpole et al., 2005; Defernez et al., 2004; Lehesranta et al., 2005), soybean (Cheng et al., 2008), rice (Dubouzet et al., 2007; Wakasa et al., 2006), tomato (Le Gall et al., 2003), tobacco, *Arabidopsis* (Kristensen et al., 2005), and *Gerbera* (Ainasoja et al., 2008).

As with the compositional studies reported above, results from many of the 'omics studies emphasize the need to understand natural variation in levels of endogenous metabolites in providing biological context to pair-wise differences in any recorded profiles (see Figure 1). Levels of compositional components are sensitive to environmental conditions. This has been established for, for example, protein and oil in key crops (Panthee et al., 2005; Lam et al., 2010). Protein levels in soybean seed generally average ~40% dry weight (dwt), with values reported in the USDA soybean germplasm collection, for example, ranging from 34.1 to 56.8% dwt (Wilson, 2004). In a recent meta-analysis of environmental effects on soybean composition, Rotundo and Westgate (2009) observed that water stress, temperature, and/or nitrogen supply all affected protein levels measured in mature seed.

Variability is even greater for lower abundance small molecule metabolites. Vitamin E (α-tocopherol) is typically only a minor component in soybean but is known to be important in maintaining oxidative stability of soybean oil. Levels in soybean seed are affected by environment and germplasm. For example, Britz et al. (2008) showed a greater than 2-fold variation in levels across three locations in the U.S. over a period of four years. Levels in soybean seed harvested from six different locations in Eastern Canada over a single year ranged from 0.87 to 3.32 mg/100g dwt (Seguin et al. 2009). Seguin et al. (2010) point out that environmental factors associated with variability in vitamin E levels include drought, temperature, and even crop management systems. The "overwhelming variability" of isoflavones was mentioned in the introduction (see Figure 1). As will be discussed later, this "overwhelming variability" can be considered to apply to levels of small molecule metabolites in harvested seed and grain of most crops.

Encouragingly, many comparative profiling studies on GM and non-GM crops have been designed to include at least one element of genotypic or environmental variability. This is exemplified in the following two examples, both of which reaffirm the need to provide biological context to pairwise-differences between two comparators.

In Baker et al. (2006) NMR-based metabolic profiles of three GM wheat varieties and the corresponding parents were generated. The incorporated transgenes encoded high-molecular weight subunits of the storage protein, glutenin. The wheat varieties were grown at two different sites over three different growing seasons (1999 -2001). Differences between the GM and parental lines were within the same range as the differences between the control lines grown on different sites and in different years. Analogous to the approach adopted in targeted compositional analyses adopting OECD recommendations, this study emphasized the importance of data from multiple years and multiple sites and that environmental variation influences metabolome composition.

In Catchpole et al. (2005) two GM potato varieties modified in fructan chemistry were grown over two different seasons (2001, 2003). Metabolic profiles of the GM and five conventional crops were generated using flow-injection MS (FIE-MS), GC-MS, and LC-MS. These demonstrated that differences between the GM and conventional potatoes were due to the intended metabolic changes, but aside from these targeted changes, the GM crops were "substantially equivalent to traditional cultivars". A major finding recognized by the authors was the large variation in the metabolic profiles of the conventional crops and, as such, the study emphasized the importance of understanding genotypic variability in assessments of compositional changes.

An often overlooked aspect of the Catchpole et al. (2005) paper is their demonstration that levels of glycoalkaloids (α-chaconine and α-solanine) were normal in the GM potatoes, a result that is easily interpretable from a food and feed perspective. Our understanding of nutrients and anti-nutrients forms the basis of attempt to modify crops through conventional breeding or agricultural biotechnology. It has allowed crops to be developed by conventional breeding that are deliberately non-equivalent to their parental progenitors in a wide range of nutritional (and agronomic) characteristics.. As Rischer et al. (2006) point out "For centuries, conventional plant breeding programs have produced new traits, higher yields and improved quality. However, little attention has been paid to metabolic changes occurring in successive generations. The issue has gained importance only recently in the context of defining thresholds for safety assessments of GM crops." It is not immediately obvious why these hitherto neglected metabolites should now be at the center of such attention. Indeed, there are few studies on small molecule metabolite changes in crops where macro-molecular composition has been deliberately changed through conventional breeding (e.g. high oil and high protein corn, high oil soybean).

Catchpole et al. (2005) in their demonstration of the compositional equivalence of GM potatoes to conventional lines also remark on the large metabolite variation in conventional potato as follows; "These significant differences [between conventional cultivars] were never sought as desired traits in traditional breeding programs, and overall composition has not given cause for public safety concerns". Overall, however, experimental designs that will both account for natural variation and have enough power to identify differences that can be attributed to transgene insertion will offer opportunities to maximize the value of omics technologies as tools in plant breeding and the development of new crops.

3. *Metabolomics offers opportunities to generate data on large numbers of metabolites. Most of these metabolites will be low in abundance and levels will be highly variable. They are also more likely to include central (and hence ubiquitous) metabolites such as sugars, organic acids, and free amino acids; metabolites that are not immediately associated with safety or nutritional relevance.*

Compositional assessments of new foodstuffs generally focus on the article of commerce, most typically harvested seed or grain. This material is generally characterized by high levels of starch, protein, fat, and fibers, with the small metabolite pool being low in abundance. For example, approximately 95-98% of maize grain is comprised the aforementioned materials; the small metabolite pool in grain, is of low abundance (~2-5% of grain biomass) and its levels are highly dependent on changes in the macromolecular pool. Soybean seed is comprised 40% protein, 20% fat, and 15% fiber. The residual 15% is comprised mainly of sugars (e.g. sucrose, raffinose, stachyose, glucose, galactose, fructose) of which the principal two, raffinose and stachyose, are measured in regulatory assessments. The fact that the small molecule metabolite pool in seed or grain is low abundance and influenced by levels of the major macromolecular nutrients accounts for its extensive variability (Skogerson et al., 2010; Harrigan et al., 2007).

Skogerson et al. (2010) sought to assess genetic and environmental impacts on the metabolite composition of corn grain. Their data acquisition technology (gas

chromatography-mass spectrometry) measured 119 identified metabolites including free amino acids, free fatty acids, sugars, organic acids, and other small molecules in a range of corn hybrids derived from 48 inbred lines crossed against two different tester lines (from the C103 and Iodent heterotic groups) and grown at three locations in Iowa (Table 1). Different metabolic phenotypes were clearly associated with the two distinct tester populations. Overall, grain from the C103 lines contained higher levels of free fatty acids and organic acids, whereas grain from the Iodent lines were associated with higher levels of amino acids and carbohydrates. In addition, the fold-range of genotype mean values [composed of six samples each (two tester crosses per inbred × three field sites)] for identified metabolites ranged from 1.5- to 93-fold with sugars and polyols being particularly variable. Interestingly, some grain metabolites showed a non-normal distribution over the entire corn population, which could, at least in part, be attributed to large differences in metabolite values within specific inbred crosses relative to other inbred sets.

Metabolite class	No. of analytes	Affected by Tester[a]	Affected by Location[b]
free amino acids	26	14	2
sterols, amines, and others	17	6	1
organic acids	17	6	0
free fatty acids and related metabolites	17	5	0
sugar alcohols	18	5	0
mono-, di-, and trisaccharides	16	1	0
sugar acids	8	0	1

[a]This indicates a statistically significant difference (p<0.0001) between hybrids derived from a cross with one tester (C103 heterotic group) versus another tester (Iodents heterotic group) [b]This indicates a statistically significant difference (p<0.0001) across the three locations in this study

Table 1. Variation in Metabolites due to Genotype or Environmental Variation

In an analogous report on the same samples, Harrigan et al. (2007) concluded that, given such variability, measurement of the small metabolite pool, was unlikely to prove useful to a comparative assessment of GM crops unless a given metabolite was an intended nutritional or toxicological endpoint. In fact, it is not immediately obvious how the data generated in Skogerson et al. (2011) could be used to determine which hybrids in this study were the safest.

In its report in 2004 the US National Research Council made pointed remarks about this disconnect as summarized in the following quotes. ".. severe imbalances between highly advanced analytical technologies and limited ability to interpret the results and predict health effects that result from the consumption of food that is genetically modified" and "….inherent difficulties, however, in identifying all of the constituents detected in profiling methods or understanding the activity and potential biological consequence of all genes in an organism severely limit the usefulness of these methods for predictive purposes.." Unable to bridge this gap, many profiling proponents make an assumption of safety on the non-GM comparator and consider statistical differences to equate with unintended effects. This tendency is described later.

4. *Another challenge in establishing a coherent literature on the impact of conventional and other approaches to breeding on natural variability in metabolite as well as determining a framework to establish nutritional meaning from metabolite analysis is the differential coverage of metabolites offered through the numerous data acquisition platforms available to omics researchers.*

As described in numerous articles on metabolomics, (*e.g.* Goodacre et al., 2004; Rischer and Oksman-Caldentey, 2006; Kusano et al., 2011) the large physico-chemical diversity of small molecule metabolites renders comprehensive metabolomic profiling through a single data acquisition technology impossible. A range of technologies associated with different detection capabilities (metabolite coverage and sensitivity), precision, resolution, throughput and reproducibility are now extensively deployed by the research community. Nuclear magnetic resonance spectroscopy (NMR), gas-chromatography mass spectrometry (GC-MS), liquid-chromatography (LC)-MS utilizing different ionization modes, Fourier-transform MS, and capillary electrophoresis (CE)- MS have all been applied in comparative assessments of GM and non-GM crops. MS approaches predominate over NMR analyses given their greater sensitivity and coverage; however this advantage does come at the expense of quantitation (i.e. MS would need an internal standard for every metabolite to be quantitated) and with a large number of unidentified MS signals in any metabolite profile. Whilst it has been suggested that untargeted profiling techniques are unbiased, it is clear that selection of a specific data acquisition technology *is* a bias and that this type of analytical bias would need to be justified by pre-specified experimental hypotheses. This justification would be critical in a Regulatory environment.

Recognizing inherent limitations for any given data acquisition technology Kusano et al. (2011) applied a multi-platform approach to an evaluation of transgenic tomato. These authors used a combination of GC-MS, LC-MS, and CE-MS with each technology covering distinct metabolite classes. Free amino acids, sugars and organic acids were covered by GC-MS, larger molecules (e.g. flavonoids) by LC-MS whereas CE-MS measured specific cations and anions. Overall, the data generated 175 unique identified metabolites but a total of 1460 with "no or imprecise metabolite annotation." Of the identified metabolites, only 56 were observed in at least two platforms. A total of 261 peaks showed no correlation with experimental factors (transgene, cultivar, tissue type) and had to be removed from statistical analyses.

It is worth pointing out that two studies that assessed the metabolic profiles of grain from GM maize containing the *Cry1ab* gene and that utilized the same data acquisition platform (NMR) differed in their conclusions on the impact of transgene insertion on levels of free amino acids (Manetti et al., 2006; Piccioni et al., 2009). Manetti et al. (2006) reported that the GM crop included higher levels of sugars (glucose, sucrose, meliobiose), GABA, glutamine, and succinate and decreased levels of alanine, asparagine, and choline. Piccioni et al. (2009) reported lower levels of all amino acids, lower sugars, and lower succinate (and other organic acids). Piccioni et al. (2009) were also able to report on metabolites observed in their NMR profiles but absent in those of Manetti et al. (2006). Key design differences between the two studies include different parental lines, different growth conditions, and sample extraction protocols.

Levandi et al. (2008) utilized CE-MS to compare levels of 27 metabolites in three different GM maize lines also containing the Cry1ab gene. Some of these metabolites (e.g., certain free amino acids, choline, GABA) were also recorded in the NMR platforms of Mannetti et al

(2006) and Piccioni et al. (2009). No consistent association of these metabolites with the GM trait when assessed over all three GM lines was observed, a conclusion in line with the combined results of Mannetti et al. (2006) and Piccioni et al. (2009). Several of the metabolites reported by Levandi et al. (2006) are more typically associated with other taxonomic groupings, for example, graveolin (*Ruta graveolus*, Rutaceae) and lunarine, (*Lunaria annua*, Brassicaceae). The assignment of peaks to metabolites not typically associated with a genus or family would almost certainly require extensive validation in a Regulatory environment.

Leon et al. (2009) utilized FT-MS on the same samples assessed by Levandi et al. (2008). This allowed coverage of 5500 mass signals of which approximately 1000 could be assigned an elemental composition. Those elemental compositions could be associated (through MasstTRIX) with specific metabolic pathways (KEGG); these associations are referred to as "isomeric hits". This approach would identify any differences in GM and non-GM metabolic profiles, especially where an elemental composition could be assigned, to be tentatively associated with biochemical differences. Overall, it was shown that a greater number of isomeric hits in pathways such as arachidonic acid metabolism, free amino acid metabolism, purine metabolism, and folate biosynthesis were associated with the GM samples. A list of 33 possible compounds that could distinguish the GM and non-GM varieties was generated, of which 12 could be confirmed in an orthogonal assay (CE-MS). The authors then indicated that only four of these could be considered as potential GM "biomarkers"; L-carnitine, apigenidin, 5, 6-dihydroxyindole, and one unidentified metabolite. There is little further literature on levels of these metabolites in maize and the association of these metabolites as GM biomarkers is almost certainly premature. Further, the interpretability of the Levandi et al. (2009) approach is not at all clear; there are fewer isomeric hits associated with inositol phosphate metabolism, yet levels of phytic acid have been well-established to be near-identical in GM and non-GM maize. The association of isomeric hits for bile acid biosynthesis, which is not typically associated with plant metabolism, is also difficult to interpret.

In summary, different metabolic profiling platforms applied to similar biological questions will yield non-overlapping solutions. This is due to differential metabolite coverage (even within similar data acquisition technologies) and is compounded both by the number of unidentified signals observed in current metabolite profiles and, in some cases, "identification" of metabolites not previously known to be biosynthetically associated with the plant species or genus in question.

3. Equating statistical equivalence with biosafety

Predetermined criteria would need to be established for any study protocol, data acquisition steps or statistical analyses utilized in a safety assessment. As alluded to earlier, sampling from multiple replicated field sites would be required. Discussions on the number of replicates required to generate meaningful results from omics experiments are available as well as on the potential for bias and over-fitting (Broadhurst and Kell, 2006; Goodacre et al., 2007). Here we focus on the routine misinterpretation of "statistical significance" (Goodman, 2008) and the tendency to associate statistically significant differences between GM and a non-GM comparator as, minimally, an unintended effect, and often to imply the statistical difference raises a question about the safety of a newly evaluated crop.

As indicated earlier, compositional assessments of GM crops involve direct comparisons of levels of key nutrients and anti-nutrients in the new crop variety to those of a near-isogenic conventional comparator. Statistical evaluations of the compositional data have typically utilized classical frequentist significance testing. There are, however, several features of significance hypothesis testing that impact its application to compositional comparisons between crops with different agronomic qualities (Lecoutre, et al., 2001). Berger (1985), for example, stated, "We know from the beginning that the point null hypothesis is almost certainly not exactly true, and that this will always be confirmed by a large enough sample. What we are really interested in determining is whether or not the null hypothesis is approximately true." There are many factors that impact crop composition, *including agronomic traits we seek to modify through plant breeding*, (e.g. Scott et al., 2006; Uribelarrea et al., 2004; Dornbos and Mullen, 1992; Hymowitz et al., 1972; Wilcox and Shibles, 2001; Yin and Vyn, 2005) and any compositional changes that accompany enhanced agronomic quality may confound interpretation of results generated through significance testing.

Statistical significance is used only as a first step in comparative assessments. The interpretation of statistical significance from a p-value, *the probability of an observed result or a more extreme result occurring if the null hypothesis were true*, does not imply biological significance (Goodman, 2008). Statistically significant differences do not imply large differences between GM and conventional comparators or that these comparators can be easily distinguished from a biological perspective. In fact, the power of the experimental designs (multiple highly replicated field trials) adopted in current compositional assessments allows statistical significance to be assigned even where there are very small difference in mean values of a given component but where the distribution of component values overlap extensively. As such, significance approaches must be accompanied with further data analysis encompassing discussion of magnitudes of differences, assessments of component ranges, and the sensitivity of component values to environmental factors such as location. This is consistent with the recommendation by Codex Alimentarius (2008, Ch. 44) that "The statistical significance of any observed differences should be assessed in the context of the range of natural variations for that parameter to determine its biological significance." It is further consistent with observations of high variability in crop composition recorded in the scientific literature. The current scientific consensus is that, in most if not all cases, statistically significant differences between GM and near-isogenic conventional controls represent modest and nutritionally meaningless differences in magnitude. For example, a recent review of studies on GM crop composition showed that over 99% of all nutrient and antinutrients comparisons, where significant differences at the 5% level ($\alpha=0.05$) in mean values were observed, had a relative magnitude difference less than 20%. These differences are considerably less than the range of values attributable to germplasm and environmental factors (Harrigan et al., 2010).

Most metabolic profiling experiments utilize significance testing and Rischer and Oksman-Caldentey (2006) refer to unintended effects as "effects which represent a statistically significant difference (e.g. in chemical composition of the GM plant compared with a suitable non-GM plant)" although they acknowledge that such differences would have to be evaluated in the context of natural variability. One review that endorses the use of omics in safety assessments suggests that "the amount of variation from genetic engineering should be small (~3%)." (Heineman et al., 2011). Whilst this particular number is unrealistic since it

falls well within the natural variability of metabolite levels and is even less than typical experimental error, setting a universal threshold for relative magnitude of differences as a trigger for further safety assessments of GM crops has been considered. In 2000, the Nordic Council of Ministers recommended that if a component in a GM crop differed from the conventional control by ±20% in relative magnitude, additional analyses of the GM crop were warranted (cited in Hothorn and Oberdoerfer, 2006). This concept was refined to account for the nutritional relevance of a component and the experimental precision of its measurement (Hothorn and Oberdoerfer, 2006). Threshold ranges for GM components were suggested as follows; 0.833-1.20 of the conventional control for "nutritionally very relevant" components (minerals, vitamins, anti-nutrients, bioactives, essential amino acids, and fatty acids), 0.769-1.30 for "relevant" (non-essential amino and fatty acids), and 0.667-1.50 for components of "less relevance" (proximates, fiber). Suggestions for the use of limits and triggers of this kind have been criticized for their failure to fully account for the role and contributions of the specific crop in the human diet; and with GM crops in particular since they are often not eaten as such but are used as a source of macronutrients such as oil, starch and protein (Chassy, 2008; Chassy, 2010). As noted previously, most plant foods in the human diet make significant contributions to the total intake of just a few macro- and micronutrients and therefore even large compositional changes in a single crop plant might produce little impact on the nutritional value of the overall diet. Chassy (2010) has observed that composition cannot be viewed in isolation since the composition of the diet is far more important than the composition of a single variety of a single crop. Strictly numerical approaches have not been adopted in compositional studies and there is no reason they would be relevant to profiling experiments.

At least one profiling study has attempted to apply statistical equivalence testing but again falls prey to the dubious association of equivalence with safety. Kusano et al. (2011) compared a GM-tomato (a miraculin protein expressor) to not only to the parental line but to a panel of conventional reference varieties. The statistical design (described by the authors as a proof-of-safety test) involved comparing the difference between test and control and the determining whether these differences fell within equivalence limits established by the reference varieties. However such a design makes more of a statement about the selection of the reference substances and the control to which the GM-trait is introgressed, and not about the effect of transgene insertion; the same test-to-control differences can be equivalent or non-equivalent contingent on whether a limited or diverse range of genotypes is available. The overall conclusion from the study however was that "miraculin over-expressors are remarkably similar to the control line".

In summary, there are no defined data analyses strategies currently being consistently applied to profiling data that would facilitate interpretability of data.

4. Conclusion

There are clearly divergent views about the utility of 'omics sciences in food safety assessments. This paper has discussed some of the reasons metabolic profiling technologies are, however, unlikely to provide immediately interpretable data in safety assessments that would otherwise enhance rigorously quantitative assessments of known nutrients and anti-nutrients that comprise foodstuffs. Indeed, it is not clear to the present authors that any new types of data are in fact necessary to judge GM or other foods as safe. We are also unaware

of any "gaps" in our compositional knowledge that might compromise safety and in fact, our current understanding of plant anti-nutrients and toxicants, allows GM solutions to enhancing food safety (e.g. Sunilkumar et al., 2006). The last 25 years of research on GM plants and 15 years of commercial experience planting GM crops without harm or incident suggest that no difference in safety that would require further analysis exists between GM and crops bred by other strategies. All breeding induces genetic changes and these changes give rise to transcriptomic, proteomic and metabolomic alterations.

We consider that metabolic profiling could increase its value in food safety science as well as in the development of nutritionally enhanced crops as follows;

1. *Improved compositional analysis*. One potential target for future research could be to develop metabolic screening methods that afford a comprehensive compositional assessment in a single suite of determinations rapidly and at lower cost than traditional targeted analysis. It is known that the metabolites in a cell form a large, complex and interconnected network; one possible approach would be elucidation of key metabolic compound whose determination might provide insight into the global concentrations of numerous other metabolites. If such a validated analytical method could be developed it would great aid research and development and would be particularly valuable in assessments of nutritionally enhanced crops where changes in a specific pathway are sought. However, metabolomic technologies are not able to supply this kind of analysis and data.

2. *Detection of novel toxicants*. Targeted analysis is inherently incapable of assessing levels of metabolites that are not selected (targeted) for analysis. Proponents of metabolic profiling have argued that profiling might detect the emergence of previously unknown novel toxicants presumably created by the breeding process. However, the abundance of a few macro-components (protein, fiber, carbohydrate, lipids) and numerous minor metabolites leaves little compositional "space" for novel toxicants. If wholly new molecules were created by the spontaneous evolution of a new pathway or pathways necessary for its biosynthesis, the chances that sufficient quantities would be present to exert an adverse effect are small indeed. Perhaps this is why such effects have not yet been observed by science or why coherent hypotheses as to how a novel toxicant would be generated by a specific breeding process appear to be sparse in the literature.

3. *Detection of unintended effects*. Proponents of metabolic profiling often suggest that a profile itself may be an indicator that unintended changes had occurred. Methods to draw safety conclusions based on differences in metabolic profiles do not yet exist, and certainly as we have discussed above, no reason to assume that differences in profiles imply a safety concern; in fact, by any objective measure, there is no such technique as metabolomic profiling. What we have today is a series of distinct and emerging powerful scanning techniques each of which surveys a slightly different molecular landscape with variable degrees of resolution. Clearly, the number of metabolites present in crops is very large and the power of targeted metabolic profiling will become increasingly useful in analyzing the chemical complexity of prospective commercial releases as they progress through initial research and development phases.

Metabolomics is an expanding and exciting field of research. The rapidly expanding scope of the metabolomic profiling technologies tempts us to test their applicability to a wide

array of analytical challenges. We have, on the other hand, a long history of safe experience with plant breeding. We know that many unintended changes take place in plant breeding, however, these are almost without exception innocuous. There is no reason to believe that GM breeding should require any new or different data set than other forms of breeding.

It seems clear to the present authors that there is no role for metabolic profiling in food safety assessment. We agree that modern targeted metabolic profiling technologies can rapidly identify pathway perturbations and, if judiciously applied and interpreted, might enhance food safety science, although traditional analytical methods can still be used to assess if changes in pathways and metabolite pools have occurred. If incorporated into the early selection stages of a prospective new trait targeted metabolic profiling may greatly aid in the selection of metabolites that need to be considered during the compositional phase of a risk assessment. To quote Larkin and Harrigan (2007) "However, it should be self-evident that GM crops ought not to be considered a single monolithic class that is either good or bad for the economy, agriculture or the environment. Each novel crop should be considered on its own merits and demerits. If we ever get to that point we will have achieved something positive out of the GM controversy." It is our hope that colleagues will take this as a challenge to further metabolic profiling in the advancement of food safety and nutritional enhancement of crops.

5. Acknowledgements

6. References

Ainasoja, M. M., Pohjala, L. L., Tammela, P.S. M., Somervuo, P. J., Vuorela, P. M. & Teeri, T. H. 2008. Comparison of transgenic Gerbera hybrida lines and traditional varieties shows no differences in cytotoxicity or metabolic fingerprints. Transgenic Res, 17, 793-803.

Baker, J. M. Hawkins, N. D. Ward, J. L., Lovegrove, A., Napier, J. A., Shewry, P. R., & Beale, M. H. 2006. A metabolomic study of substantial equivalence of field-grown genetically modified wheat. Plant Biotechnol J, 4, 381-92.

Beier, R. C. & Oertli, E. H. 1983. Psoralen and other linear furocoumarins as phytoalexins in celery. Phytochem, 22, 2595-97.

Bell, A. E. 2003. Nonprotein amino acids of plants: Significance in medicine, nutrition, and agriculture. J Agric Food Chem, 51, 2854–65.

Berger, J.O. 1985. Statistical Decision Theory and Bayesian Analysis. 2nd edition: Springer-Verlag.

Britz, S. J., Kremer, D. F., & Kenworthy, W. 2008. Tocopherols in soybean seeds; Genetic variation and environmental effects in field-grown crops. J Am Oil Chem Soc, 85, 931-36.

Broadhurst, D. I. & Kell, D. B. 2006. Statistical strategies for avoiding false discoveries in metabolomics and related experiments. Metabolomics, 2, 171-96.

Catchpole, G. S., Beckmann, M., Enot, D. P., Mondhe, M., Zywicki, B., Taylor, J., Hardy, N., Smith, A., King, R. D., Kell, D. B., Fiehn, O. & Draper, J. 2005. Hierarchical

metabolomics demonstrates substantial compositional similarity between genetically modified and conventional potato crops. Proc Natl Acad Sci U S A, 102, 14458-62.

Chassy, B., Egnin, M., Gao, Y., Glenn, K., Kleter, G. A., Nestel, P., Newell-Mcgloughlin, M., Phipps, R. H. & Shillito, R. 2008. Recent developments in the safety and nutritional assessment of nutritionally improved foods and feeds. Comp Rev Food Sci Food Safety, 7, 50-113.

Chassy, B. 2010. Can 'omics inform a food safety assessment? Reg Toxicol Pharmacol,58, S62-70.

Cheng, K. C., Beaulieu, J., Iquira, E., Belzile, F. J., Fortin, M. G., & Stromvik, M. V. 2008. Effect of transgenes on global gene expression in soybean is within the natural range of variation of conventional cultivars. J Agric Food Chem, 56, 3057–67.

Codex Alimentarius. 2008. Guideline for the conduct of food safety assessment of foods derived from recombinant-DNA plants. CAC/GL 45-2003.

Davies, H. 2010. A role for "omics" technologies in food safety assessment. Food Control, 21, 1601-10.

Defernez, M., Gunning, Y. M., Parr, A. J., Shepherd, L. V., Davies, H. V. & Colquhoun, I. J. 2004. NMR and HPLC-UV profiling of potatoes with genetic modifications to metabolic pathways. J Agric Food Chem, 52, 6075-85.

Dornbos, D. & Mullen, R. 1992. Soybean seed protein and oil contents and fatty acid composition adjustments by drought and temperature. J Am Oil Chem Soc, 69, 228-31.

Dubouzet, J. G., Ishihara, A., Matsuda, F., Miyagawa, H., Iwata, H. & Wakasa, K. 2007. Integrated metabolomic and transcriptomic analyses of high-tryptophan rice expressing a mutant anthranilate synthase alpha subunit. J Exp Bot 58, 3309-21.

Eldridge, A. C. & Kwolek W. F. 1983. Soybean isoflavones: effect of environment and variety on composition. J Agric Food Chem, 31, 394-6.

EFSA. 2011. EFSA Panel on Genetically Modified Organisms (GMO); Scientific opinion on guidance for risk assessment of food and feed from genetically modified plants. EFSA Journal, 9, 2150 [37pp]. Available online: www.efsa.europa.eu/efsajournal.htm

Foresight. The Future of Food and Farming (2011) Final Project Report. The Government Office for Science, London.

Goodacre, R., Vaidyanathan, S., Dunn, W. R., Harrigan, G. G. & Kell, D. B. 2004. Metabolomics by numbers-Acquiring and understanding global metabolite data. Trends Biotechnol, 22, 245-52.

Goodacre, R., Broadhurst, D., Smilde, A. K., Kristal, B.S., Baker, D. J., Beger, R., Bessant, C., Connor, S., Capuani , G., Craig, A., Ebbels, T., Kell, D. B., Manetti, C., Newton, J., Paternostro, G., Somorjai, R., Sjostrom, M., Trygg, J. & Wulfert, F. 2007. Proposed minimum reporting standards for data analysis in metabolomics. Metabolomics, 3, 231-41.

Goodman, S. 2008. A dirty dozen: Twelve p-value misconceptions. Sem Hematol, 45, 135-40.

Graff, G. D., Zilberman, D. & Bennett, A. B. 2009. The contraction of agbiotech product quality innovation. Nat Biotechnol, 27, 702-4.

Gregersen, P. L., Brinch-Pedersen, H. & Holm, P. B. 2005. A microarray-based comparative analysis of gene expression profiles during grain development in transgenic and wild type wheat. Transgenic Res, 14, 887-905.

Gutierrez-Gonzalez, J. J., Wu, X., Zhang, J., Lee, J. D., Ellersieck, M., Shannon, J. G., Yu, O., Nguyen, H. T. & Sleper, D.A. 2009. Genetic control of soybean seed isoflavone content: importance of statistical model and epistasis in complex traits. Theor Appl Genet, 119, 1069-83.

Harrigan, G. G. & Harrison, J. M. 2011. Assessing compositional variability through graphical analysis and Bayesian statistical approaches; case studies on transgenic crops. Biotechnol Genet Eng Rev, 28, 1-18.

Harrigan, G. G., Stork, L. G., Riordan, S. G., Reynolds, T. L., Ridley, W. P., Masucci, J. D., Macisaac, S., Halls, S. C., Orth, R., Smith, R. G., Wen, L., Brown, W. E., Welsch, M., Riley, R., Mcfarland, D., Pandravada, A. & Glenn, K. C. 2007. Impact of genetics and environment on nutritional and metabolite components of maize grain. J Agric Food Chem, 55, 6177-85.

Harrigan, G. G., Lundry, D., Drury, S., Berman, K., Riordan, S. G., Nemeth, M. A., Ridley, W. P. & Glenn, K. C. 2010. Natural variation in crop composition and the impact of transgenesis. Nat Biotechnol, 28, 402-4.

Heineman, J. A., Kurenbach, B. & Quist, D. 2011. Molecular profiling - a tool for addressing emerging gaps in the comparative risk assessment of GMOs. Environ Int, 37, 1285-93.

Herman, R. A., Chassy, B. M. & Parrott, W. 2009. Compositional assessment of transgenic crops: an idea whose time has passed. Trends Biotechnol, 27, 555-7.

Hothorn, L. A. & Oberdoerfer, R. 2006. Statistical analysis used in the nutritional assessment of novel food using the proof of safety. Reg Toxicol Pharmacol, 44, 125-135.

Hymowitz, T., Collins, F. I., Panczner, J. & Walker, W. M. 1972. Relationship between the content of oil, protein, and sugar in soybean seed. Agron J, 64, 613-16.

Ioset, J. R., Urbaniak, B., Ndjoko-Ioset, K., Wirth, J., Martin, F., Gruissem, W., Hostettmann, K. & Sautter, C. 2007. Flavonoid profiling among wild type and related GM wheat varieties. Plant Mol Biol, 65, 645-54.

Kok, E. J., Keijer, J., Kleter, G. A. & Kuiper, H. A. 2008. Comparative safety assessment of plant-derived foods. Regul Toxicol Pharmacol, 50, 98-113.

Kristensen, C. Morant M. Olsen C. E., Ekstrøm, C. T., Galbraith, D. W., Møller, B. L., & Bak, S. 2005. Metabolic engineering of dhurrin in transgenic Arabidopsis plants with marginal inadvertent effects on the metabolome and transcriptome. Proc Natl Acad Sci U S A, 102, 1779-84.

Jones, D.A. 1998. Why are so many food plants cyanogenic? Phytochem, 47, 155-62.

Kalaitzandonakes, N., Alston, J.M. & Bradford, K. J. 2007. Compliance costs for regulatory approval of new biotech crops. Nat Biotechnol, 25, 509-11.

Khoshbahkt, K. & Hammer, K. 2008. How many plant species are cultivated? Genet Resour Crop Evol, 55, 925-28

Kusano, M., Redestig, H., Hirai, T., Oikawa, A., Matsuda, F., Fukushima, A., Arita, M., Watanabe, S., Yano, M., Hiwasa-Tanase, K., Ezura, H. & Saito, K. 2011. Covering chemical diversity of genetically-modified tomatoes using metabolomics for objective substantial equivalence assessment. PLoS, 6, e16989.

Lam, H-M., Xu, X., Liu, X., Chen, W., Yang, G., Wong F. L., Li, M.-W., He, W., Qin, N., Wang, B., Li, J., Jian, M., Wang, J., Shao, G., Wang, J., Sun , S. S.-M., & Zhang, G. 2010. Resequencing of 31 wild and cultivated soybean genomes identifies patterns of genetic diversity and selection. Nature Gen, 42, 1053-59.

Larkin, P. & Harrigan, G. G. 2007. Opportunities and surprises in crops modified by transgenic technology: metabolic engineering of benzylisoquinoline alkaloid, gossypol and lysine biosynthetic pathways. Metabolomics, 3, 371-82.

Lay Jr., J.O, Borgmann, S., Liyanage, R. & Wilkins, C.L. 2006. Problems with the ''omics''. Trends Anal Chem, 25, 1046-56.

Le Gall, G., Colquhoun, I. J., Davis, A. L., Collins, G. J. & Verhoeyen, M. E. 2003. Metabolite profiling of tomato (*Lycopersicon esculentum*) using ^{1}H NMR spectroscopy as a tool to detect potential unintended effects following a genetic modification. J Agric Food Chem, 51, 2447-56.

Lecoutre, B., Lecoutre, M.-P. & J. Poitevineau, J. 2001. Uses, abuses and misuses of significance tests in the scientific community: Won't the Bayesian choice be unavoidable? Int Stat Rev, 69, 399-417.

Lehesranta, S. J., Davies, H. V., Shepherd, L. V., Nunan, N., Mcnicol, J. W., Auriola, S., Koistinen, K. M., Suomalainen, S., Kokko, H. I. & Karenlampi, S. O. 2005. Comparison of tuber proteomes of potato varieties, landraces, and genetically modified lines. Plant Physiol, 138, 1690-9.

Leon, C., Rodriguez-Meizoso, I., Lucio, M., Garcia-Canas, V., Ibanez, E., Schmitt-Kopplin, P. & Cifuentes, A. 2009. Metabolomics of transgenic maize combining Fourier transform-ion cyclotron resonance-mass spectrometry, capillary electrophoresis-mass spectrometry and pressurized liquid extraction. J Chromatog A, 1216, 7314-23.

Levandi, T., Leon, C., Kaljurand, M., Garcia-Canas, V. & Cifuentes, A. 2008. Capillary electrophoresis time-of-flight mass spectrometry for comparative metabolomics of transgenic versus conventional maize. Anal Chem, 80, 6329-35.

Manetti, C., Bianchetti, C., Casciani, L., Castro, C., Di Cocco, M. E., Miccheli, A., Motto, M. & Conti, F. 2006. A metabonomic study of transgenic maize (*Zea mays*) seeds revealed variations in osmolytes and branched amino acids. J Exp Bot, 57, 2613-25.

NIEHS. 1998. National Institute of Environmental Health Sciences; α-Chaconine [20562-03-2] and α-solanine [20562-02-1]. Review of toxicological literature. http://ntp-server.niehs.nih.gov/index.cfm?objectid=6F5E930D-F1F6-975E-7037ACA48ABB25F4

NRC. 2004. National Research Council: Safety of Genetically Engineered Foods. Approaches to Assessing Unintended Health Effects. The National Academies Press, Washington, D.C.

Panthee, D. R., Pantalone, V. R., West, D. R., Saxton A. M., & Sams, C. E. 2005. Quantitative trait loci for seed protein and oil concentration and seed size in soybean. Crop Sci, 45, 2015-22.

Piccioni, F., Capitani, D., Zolla, L. & Mannina, L. 2009. NMR metabolic profiling of transgenic maize with the Cry1A(b) gene. J Agric Food Chem, 57, 6041-49.

Potrykus, I. 2010. Regulation must be revolutionized. Nature, 466, 561

Ricroch, A. E., Berge, J. B. & Kuntz, M. 2011. Evaluation of genetically engineered crops using transcriptomic, proteomic, and metabolomic profiling techniques. Plant Physiol, 155, 1752-61.

Rischer, H. & Oksman-Caldentey, K.-M. 2006. Unintended effects in genetically modified crops: revealed by metabolomics? Trends Biotechnol, 24, 102-4.

Rotundo, J.L. & Westgate, M. E. 2009. Meta-analysis of environmental effects on soybean seed composition. Field Crops Res, 110, 147-56.

Scott, M. P., Edwards, J. W., Bell, C. P., Schussler, J. R. & Smith, J. S. 2006. Grain composition and amino acid content in maize cultivars representing 80 Years of commercial maize varieties. Maydica, 51, 417-23.

Seguin, P., Turcotte, P., Tremblay, G., Pageau, D., & Liu, W. 2009. Tocopherols concentration and stability in early maturing soybean genotypes. Agron. J, 101, 1153-59.

Seguin, P., Tremblay, G., Pageau, D. & Liu, W. 2010. Soybean tocopherol concentrations are affected by crop management. J Agric Food Chem, 58, 5495-5501.

Senti, F.R. & Rizek, R. R.1974. An overview of GRAS regulations and the effect from the viewpoint of nutrition, In Hanson C. H. (ed) CSSA Special Publication. Vol. 5. SSA, Madison, WI.

Skogerson, K., Harrigan, G. G., Reynolds, T. L., Halls, S. C., Ruebelt, M., Iandolino, A., Pandravada, A., Glenn, K. C. & Fiehn, O. 2010. Impact of genetics and environment on the metabolite composition of maize grain. J Agric Food Chem, 58, 3600-10.

Sunilkumar, G., Campbell, L. M., Puckhaber, L., Stipanovic, R. D. & Rathore, K. S. 2006. Engineering cottonseed for use in human nutrition by tissue-specific reduction of toxic gossypol. Proc Natl Acad Sci U S A, 103, 18054-59.

Uribelarrea, M., Below, F. E. & Moose, S.P. 2004. Grain composition and productivity of maize hybrids derived from the Illinois protein strains in response to variable nitrogen supply. Crop Sci, 44, 1593-1600.

Yin, X.H. & Vyn, T. J. 2005. Relationships of isoflavone, oil, and protein in seed with yield of soybean. Agron J, 97, 1314-21.

Wakasa, K., Hasegawa, H., Nemoto, H., Matsuda, F., Miyazawa, H., Tozawa, Y., Morino, K., Komatsu, A., Yamada, T., Terakawa, T. & Miyagawa, H. 2006. High-level tryptophan accumulation in seeds of transgenic rice and its limited effects on agronomic traits and seed metabolite profile. J Exp Bot, 57, 3069-78 .

Wilcox, J.R. & Shibles, R. M. 2001. Interrelationships among seed quality attributes in soybean. Crop Sci, 41, 11-4.

Wilson, R. F. 2004. Seed composition, In H.R. Boerma and J.E. Specht (ed.) Soybeans: Improvement, production and uses, 3rd ed. ASA, CSSA, SSA, Madison, WI.

Zhou, J., Berman, K. H., Breeze, M. L., Nemeth, M. A., Oliveira, W. S., Braga, D. V., Berger, G. U. & Harrigan, G. G. 2011a. Compositional variability in conventional and glyphosate-tolerant soybean varieties grown in different regions in Brazil. J Agric Food Chem, 59, online

Zhou, J., Harrigan, G. G., Berman, K. H., Webb, E. G., Klusmeyer, T. H. & Nemeth, M. A. 2011b. Stability of the compositional equivalence of grain from insect-protected corn and seed from herbicide-tolerant soybean over multiple seasons, locations and breeding germplasms. J Agric Food Chem, 59, 8822-28.
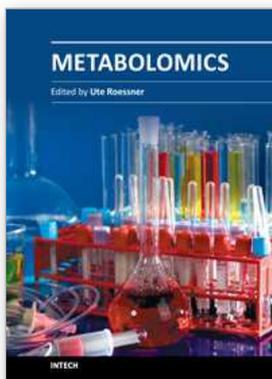
Metabolomics

Edited by Dr Ute Roessner

Metabolomics is a rapidly emerging field in life sciences, which aims to identify and quantify metabolites in a biological system. Analytical chemistry is combined with sophisticated informatics and statistics tools to determine and understand metabolic changes upon genetic or environmental perturbations. Together with other 'omics analyses, such as genomics and proteomics, metabolomics plays an important role in functional genomics and systems biology studies in any biological science. This book will provide the reader with summaries of the state-of-the-art of technologies and methodologies, especially in the data analysis and interpretation approaches, as well as give insights into exciting applications of metabolomics in human health studies, safety assessments, and plant and microbial research.

INTECH

open science | open minds