Chapter 6

A particular case of novel food: Genetically modified organisms

Virginia García-Cañas, Alejandro Cifuentes

Laboratory of Foodomics, Institute of Food Science Research, CIAL (CSIC)
Nicolas Cabrera 9, 28049 Madrid, Spain.

ABSTRACT

The rapid progress of recombinant DNA technology has opened new prospects in the development of novel foods and food ingredients, including those containing genetically modified organisms (GMOs). These new transgenic products have brought about a considerable demand for analytical methods able to detect, characterize and/or quantify GMOs along the food chain. The current status and future challenges in the development, characterization and detection of GMOs in foods are discussed in this chapter describing the different methodologies proposed so far based on profiling or target analysis. Advantages and limitations derived from the application of these methodologies are highlighted and discussed. Special emphasis is given to the potential of omics technologies to study these novel foods.

Keywords: genetically modified organisms, transgenic food, PCR, qPCR, microarray, foodomics, transcriptomics, proteomics, metabolomics.
1. Introduction

2. Controversial safety issues and legislation on GMOs

3. Strategies used for the analysis of GMOs
   3.1 Target-based strategies
      3.1.1 Screening methods for GMOs detection in food
      3.1.2 Quantification of GMOs in food
   3.2 Profiling strategies
      3.2.1 Transcriptomics
      3.2.2 Proteomics
      3.2.3 Metabolomics

4. Conclusions and future outlooks
1. INTRODUCTION

Genetic engineering (or recombinant DNA technology) allows selected individual gene sequences to be transferred from an organism into another and also between non-related species. The rapid progress of this technology has opened new prospects in the development of novel foods and food ingredients\(^1\). The organisms derived from genetic engineering are termed genetically modified organisms (GMOs). A transgenic food is defined as a food that is derived from or contains GMOs.

The fast adoption of genetic engineering in agriculture has led to the production of transgenic crops such as soybean, maize, wheat, rice, cotton, potato, canola and tobacco that include benefits in industrial processing and agronomic productivity. Among the modifications, tolerance to herbicide\(^2\) and resistance to insects and disease\(^3\) are the predominant traits in current commercialized genetically modified (GM) crops. In the past decade, over 144 GMOs, representing 24 crops, have been approved by regulatory agencies in different countries. Furthermore, this number is expected to rise, including a second generation of GMOs with nutritionally enhanced traits, such as, for instance, plants enriched in β-carotene\(^4\), vitamin E\(^5\), or omega-3 fatty acids\(^6\) that could likely enter the market in the near future\(^7\).

In spite of its important economic potential, recombinant DNA technology has become highly controversial, not only within the scientific community but also in the public sector since its beginning more than three decades ago\(^8\). The main controversial issues focus on four areas: concerns about potential harm to human health\(^9-11\), environmental concerns\(^12,13\), ethical concerns interferences with nature and individual choice\(^14\), and concerns related to patent issues\(^15,16\) that are next discussed.
2. CONTROVERSIAL SAFETY ISSUES AND LEGISLATION ON GMOS

Regardless of the presumed accuracy of recombinant DNA technology for genetic modification, possible unintended effects that derive from the genetic transformation might occur. Unintended effects go beyond the primary expected effects of the genetic modification, and represent statistically significant differences in a phenotype compared with an appropriate phenotype control\textsuperscript{17}. In some cases, unintended effects might be potentially linked to secondary effects of gene expression in a way that could be somehow explained from our current knowledge of plant biology and metabolism, or from the function of a transgene or the site of its integration in the genome\textsuperscript{18,19}. On the other side, some other unintended effects might be associated with unexpected transformation-induced mutations (deletions, insertions, rearrangements, etc.) that occur during the transformation and tissue-culture stages of GMO development\textsuperscript{20-24}. The unintended effects are difficult to predict or explain without the thorough characterization of the plant at the molecular level. Such effects could also be observed if the changes result in a distinct phenotype, including compositional alterations. Thus, unintended effects represent a significant source of unpredictability that might have an impact on human health and/or the environment\textsuperscript{25}.

The development and use of GMOs for food applications are issues of intense debate and public concern that have led to the establishment of strict regulations concerning different aspects of GMOs, including risk assessment, marketing, labeling and traceability in the European Union and other countries. A general leading strategy has been based on the assumption that traditional crop-plant varieties currently on the
market that have been consumed for decades have gained a history of safe use, and, therefore, they can be used as comparators for the safety assessment of new GM crop varieties derived from established plant lines. Although this concept, typically referred as “substantial equivalence” or “comparative safety assessment”, has been adopted for the current safety assessment of GM foods in several countries, the approval procedure of these novel foods differs across national jurisdictions. As a consequence, this lack of international harmonization among regulations has generated an “asynchronous approval” of GMOs worldwide.

As for authorization of GMOs, labeling and traceability requirements are also quite heterogeneous between the legal frameworks in different countries. For instance, labeling of foodstuffs may be voluntary or mandatory, and the specific thresholds set for labeling vary between countries. In the particular case of the European Union, the Regulation 1829/2003 establishes that any food containing more than 0.9% GM content has to be labeled as such, provided that the presence of this GM ingredient is adventitious or technically unavoidable. For non-authorized GM ingredients, the threshold is set at 0.5%, provided that the source GMO has passed the first stages of approval.26

3. STRATEGIES USED FOR THE ANALYSIS OF GMOs

Appropriate methodologies able to cope with the analysis of these novel foods are required in order to verify the compliance with the requirements imposed by the legislation regarding GMOs. Such analytical tools should: (1) enable the specific identification and accurate determination of GMOs content in foods for labeling
compliance; and (2) facilitate comprehensive compositional studies of GMOs in order to effectively investigate the potential adverse effects on the human health, including the existence (or not) of unintended modifications. In order to face both analytical challenges, two methodologically different analytical approaches have been proposed, i.e., target analysis and profiling. Next sections will be focused on the last developments and advances made in the areas of target analysis and profiling for the study of GMOs, discussing some of their advantages and drawbacks.

### 3.1 Target-based strategies

The analysis of targeted compounds is helpful to study the primary or intended effect of the genetic modification. In some cases, the interest might be focused on the insertion and the expression of the new transgene; subsequently, the analysis is directed towards the detection of specific DNA, mRNA, or proteins (i.e., target analytes). Also, with the goal to study the intended effect induced by the genetic modification at the metabolite level, target analysis might also focus on the detection of a limited selection of metabolites that are involved in altered biochemical/physiological pathways in the GMO. Moreover, the application of targeted analysis to characterize a number of constituents, including macro- and micronutrients, antinutrients, and natural toxins in food crops, has also been proposed as a tool for comparative safety assessments of a GM crop with its traditional counterpart\textsuperscript{17,27}. Nevertheless, numerous concerns have been raised about the use of such targeted analytical approaches to compare the composition of GM crops with their traditional counterparts. It has been pointed out that this approach is biased\textsuperscript{28}, and presents many limitations, such as the possible occurrence of unknown toxicants and anti-nutrients, particularly in food-plant species with no history of (safe) use\textsuperscript{29}. Moreover, although a few studies have identified unintended
effects with targeted approaches\textsuperscript{30-32}, this strategy might restrict the possibilities to detect other unpredictable effects that could result directly or indirectly from the genetic modification.

On the contrary, the analysis of target molecules has been the strategy of choice for the detection, identification and quantification of GMOs and GM-based materials in food samples\textsuperscript{33-37}. In target analysis, two types of macromolecules, specific for the genetic modification, have been used in order to reveal the presence of GM-based material in foods: proteins and DNA. Proteins specific for a given GMO are usually detected by immunoassay. Some commercially available immunoassays, based on the use lateral-flow strips, have proven to be useful for the analysis of raw materials for in-field applications. However, its performance is greatly affected by most food processing technologies, and also by the expression levels of the transgene, which depend on the physiological state of the plant or tissue. In addition to the difficulties for multiplexing this technique, its evolution has been limited by the higher costs of developing and producing specific antibodies in contrast to oligonucleotides used in DNA-based techniques\textsuperscript{38}.

In contrast to proteins, DNA presents higher thermal stability, it is present in most biological tissues, and the fact that the genetic modification affects primarily DNA sequence, makes it a more suitable target for GMO detection. Most DNA-based detection methods for GMOs rely on the use of polymerase chain reaction (PCR) to detect, identify and quantify GMOs in food. PCR in its different formats has been established as the prevailing technique for GMO detection and traceability due to its specificity, sensitivity and the fact that it allows a rapid and relatively low-cost analysis.
However, many factors can affect sensitivity and specificity of PCR-based methods, such as quality of DNA, sample processing, equipment and chemicals. The ability of PCR to amplify specific DNA sequences in a complex DNA extract will depend, to a great extent, on the integrity, quantity and purity of the DNA extract. Integrity of DNA is affected by several factors during food processing and DNA isolation (e.g., variations of pH and temperature, existence of nucleases)\textsuperscript{39}. Chemical modifications, like depurination, and the mean size of the DNA molecules, influence the minimal size of the target sequences that can be amplified. In addition, contaminants from the food matrix or from the chemicals used for DNA isolation can inhibit the PCR reactions. These limiting factors define the amplificability of target DNA sequences by PCR-based methods, and are considered critical issues for GMO analysis in highly processed and complex food samples\textsuperscript{40}. Taking into consideration the differences in type, composition and degree of processing of foods, DNA extraction protocols must be developed and applied on a case-by-case basis.

\textbf{3.1.1 Screening methods for GMOs detection in food}

The number of approved GM crops and the extension areas where they are cultivated are steadily increasing around the globe. At present, about 90 novel GMOs are in advanced stages of the development, authorization, or commercialization process and may enter the market in the near future\textsuperscript{41}. In this context, screening methods based on DNA analysis have become essential to minimize the analytical effort to rapidly assess whether or not a sample under investigation is likely to contain GM-derived materials. Ideally, screening methods should provide a global snapshot of the transgenic elements present in a given food sample allowing a further pre-selection of more specific analysis for identification and quantification.
Numerous screening methods are based on the amplification and detection of DNA sequences found in as many different GMOs as possible\textsuperscript{42-46}. The two sequences more frequently used for this purpose are the promoter, P-35S, from cauliflower mosaic virus (CaMV) and the nopaline synthase gene terminator, T-nos, from \textit{Agrobacterium tumefaciens}. Alternatively, some genes which encode a certain trait, such as herbicide tolerance, e.g., phosphinothricin N-acetyltransferase (\textit{bar} and \textit{pat}) and 3-phosphoshikimate 1-carboxyvinyltransferase (EPSPS), an insecticide, e.g., δ-endotoxins (cry genes), or others can also be suitable targets for screening methods. However, as the number of new GMOs grows, as well as the novel phenotypic characters and transcription control regions are included in their production, the screening approach based on a very limited number of targets is not longer effective\textsuperscript{47}. Consequently, screening methods need to be as comprehensive as possible to face the complexity of this topic. Comprehensive GMO screening requires a high degree of parallel tests or multiplexing. Parallel examination of screening elements, also known as “matrix” or “pattern” approach, enables combinatorial coverage of genetic targets with different degree of specificity in order to increase the probability of detecting any possible GMO\textsuperscript{48}. Certain combinations of plant markers with screening targets can, in some cases, be very specific for certain GMOs; however, unambiguous identification is generally obtained using event-specific elements\textsuperscript{49}. Targets for event-specific identification are those sequences that cover the border of the transgenic insert and the plant genome.

Multiplex detection-based strategies are becoming preferred for GMO routine detection in food and feed\textsuperscript{50} in order to keep GMO analysis within the practicable limits of routine
analysis, minimizing cost and time per analysis. Among the molecular techniques available for simultaneous detection, multiplex PCR has been the most widely explored for GMO analysis. This technique involves the simultaneous amplification of more than one target sequence per reaction by mixing multiple primer pairs with different specificities in the same reaction. However, the application of multiplex PCR is limited by several constraints. First, multiplex PCR-based methods are more susceptible to non-specific product amplification or cross-amplification reactions than conventional PCR since several primer pairs are added to the reaction mixture and consequently, the risk of having false positive signals might be increased \textsuperscript{51-53}. Second, small differences in amplification efficiencies for the different primer pairs results in different amplification rates for the different targets. Then, some target sequences are preferentially amplified owing to the exponential nature of PCR, leaving other amplicons undetectable, which compromises the sensitivity and increases the risk of having false negatives \textsuperscript{54}. In the last years, sensitive and highly efficient capillary gel electrophoresis-laser induced fluorescence (CGE-LIF) methods have been proposed as alternative to agarose gel electrophoresis to overcome risks associated to false positive and negative detection in multiplex amplifications \textsuperscript{55-59}. An additional shortcoming of multiplex PCR, particularly important for its application to screening analysis, is the lack of flexibility for further modifications of the amplification system, as for instance, the incorporation of extra primer pairs for the detection of additional target sequences \textsuperscript{60}.

In the last years, much interest has been focused on the development of alternative amplification techniques for detection of multiple GMO targets in food samples. Thus, innovative ligation-based approaches have shown good potential for multiplexing. Ligation-based techniques combine a ligation step, required for specificity, and an
amplification step, required for sensitivity. The latter step is often performed using a single pair of primers, which favors equal amplification rates for all target sequences. Ligation-dependent probe amplification (LPA) combined with CGE-LIF has been applied to the simultaneous detection of DNA from MON810 maize and Roundup Ready soybean in a single reaction\textsuperscript{61}. The technique does not amplify the target genomic DNA, but is rather based on the amplification of products resulting from the ligation of bipartite hybridization probes using universal amplification primers. Further studies on the application of LPA-CGE-LIF to GMO detection have demonstrated its good multiplexing capabilities, allowing the simultaneous detection of several DNA target sequences\textsuperscript{62,63}. Chaouachi \textit{et al.} (2008) used a more sophisticated strategy based on the same ligation-amplification principle and a commercial CGE genotyping system that was adapted to simultaneously detect 48 short sequences from taxa endogenous reference genes, GMO constructions, screening targets, construct-specific, and event-specific targets, and from donor organisms\textsuperscript{53}. Also, Multiplex Ligation-Dependent Genome Dependent Amplification (MLGA) technique has been developed in combination with CGE-LIF to simultaneously detect several GM lines of maize\textsuperscript{64} (see Figure 1). In contrast to the aforementioned ligation techniques, MLGA is based on the ligation of genomic DNA instead of probe molecules, and a single specific probe is required for each target. Other approaches toward simultaneous detection of GMOs include the combined use of multiplex PCR with Ligase Chain Reaction (LCR) to improve the sensitivity in polyacrylamide gel electrophoresis with fluorescent scanning detection\textsuperscript{65}.

Multiplexing of real-time PCR methods has been also explored by a strategy that exploits the different spectral properties of a variety of fluorochromes to discriminate
amplified products. In spite of the known restricted capabilities of this approach for multiplexing due to the limited number of potential dyes and the overlaps of fluorescence spectra, sensitive detection can be achieved using a suitable combination of dyes\textsuperscript{45}. Following this approach, a screening kit based on hydrolysis (5′-nuclease) probes was developed to simultaneously detect four common screening targets in 22 GMOs\textsuperscript{66}. Alternative strategies have exploited the high-throughput benefits of the commonly applied 96-well plate formats for multiparallel screening analysis using real-time PCR platforms. Thus, a ready-to-use plate has been configured to contain lyophilized primers and probes for the individual detection of targets allowing the simultaneous identification of 39 GM events in one run by the use of event specific primers and probe combinations\textsuperscript{67}. Similarly, the Combinatory SYBR\textsuperscript{®}GREEN qPCR Screening (CoSYPS) is a flexible screening alternative that has been developed to be implemented adapting the matrix approach to the 96-well plate qPCR format\textsuperscript{68}. Here, the interpretation of the analytical results relies on the combination of two analytical parameters of each real-time PCR method: the $C_t$ value of the amplification (an estimate of the target abundance as it will be discussed later) and the $T_m$ value of the amplified product (used for authentication). The application of real-time PCR for screening purposes presents the additional advantage that results are directly obtained from the instrument software avoiding further manipulations of PCR products for subsequent analysis, a step that entails the main risk in terms of laboratory contamination.

Nowadays, DNA microarray technology has the greatest capabilities for multiplexing when nucleic acids are analyzed. Microarray represents the leading trend for developing screening methods for GMO detection. In general, microarray technology is conducted by previous amplification of target DNA sequences by multiplex PCR, and the resulting
amplification products are labeled and hybridized on oligonucleotides immobilized on a solid support in predefined locations. Once hybridization is complete, samples are washed for further image capturing and data analysis. Theoretically, the fluorescent or colorimetric signal of derivatized-nucleic acids bound to any probe is a function of their concentration. The relative signal for each sequence is extracted and transformed to a numeric value\textsuperscript{69,70}. In the last years, several microarray methods, most of them based on low-density formats, have been reported for screening analysis with different GMO coverage\textsuperscript{71-73}. To date, a commercial microarray system, DualChip®GMO V2.0 (Eppendorf Array Technologies, Namur, Belgium), combines three multiplex PCR methods with a colorimetric detection and a scanning system (Silverquant system) to detect up to 30 different screening targets, covering more than 80% of all GMOs known at this moment\textsuperscript{74}.

Although microarrays are already comparable to real-time PCR in terms of sensitivity and reproducibility, most of the tests developed with microarrays are qualitative. Moreover, one of the major limitations of microarrays is the high cost of the technology, which requires a significant investment in equipment and consumables. Besides, the need of a previous DNA amplification step prior hybridization onto microarray chip for the most common microarray platforms limits the real high-throughput and quantitative capabilities of this technology\textsuperscript{75}. Further developments in this area include more efficient and robust pre-amplification strategies, such as Whole Genome Amplification\textsuperscript{76}, Padlock Probe Ligation technique\textsuperscript{77}, and Nucleic Acid Sequence-Based Amplification, NAIMA\textsuperscript{78} (see Figure 2), which might provide with true multiplexing and quantitative capabilities to microarray platforms for GMO detection.
3.1.2 Quantification of GMOs in food

Verification of the compliance with regulations on labeling requires the quantification of the amount of GMO per ingredient (defined as a “taxon”, e.g. soy, maize, etc.) to verify whether GMO content exceeds a threshold level (above 0.9% GMO per ingredient). As discussed above, its high sensitivity has made PCR the technique of choice in order to qualitatively detect the presence of GMOs in foods. However, inherent features of the PCR, together with differences in the composition of the samples, constrain its use for accurate quantification of GMOs in food samples. In this regard, quantitative real-time PCR has emerged as the established method due to its specificity, wide dynamic range, speed, easy use and its power to accurately quantify small traces of GM derivatives in processed material. The main singularity of real-time PCR is that the amount of amplicon generated can be monitored and quantified after each amplification cycle. This is achieved by the use of fluorescent dyes, which can interact with double-stranded DNA, or sequence-specific oligonucleotides. The main parameters used to quantify the target sequence by real-time PCR are the threshold fluorescence signal, set as a given value statistically significant above the noise, and the threshold cycle (Ct), defined as the cycle number at which the fluorescence surpasses the threshold. In order to make the results reproducible, the threshold value is chosen in the exponential phase of the amplification. Under ideal conditions, Ct value is inversely proportional to the amount of target sequence at the beginning of the reaction. This allows the generation of standard curves, by using reference materials, at which the Ct numbers obtained from given samples are calculated. Generally, real-time PCR assays are also directed at target molecules in the range 80-150 bp, which makes them very suitable for use in the analysis of processed foods. For GMO quantification, both targets
for the endogenous gene and the transgene should be of similar length in order to obtain similar PCR efficiencies for both targets. The percentage of GMO is calculated from a combination of two absolute quantification values: one for the specific GMO and the second one for the endogenous reference gene.

In real-time PCR, the sensitivity is affected by the nature and concentrations of the reagents needed for the amplification, and especially by the type of the fluorogenic system (FS) used to generate the fluorescence signals. About 20 different FSs have been developed for real-time PCR assays although not all have been tested for GMO analysis. A commonly used FS is SYBR-Green I, an intercalating dye that provides fluorescent signal upon interaction with any double-stranded DNA. Although the use of this reagent is more cost-effective compared to the alternative primer- or probe-based FSs, SYBR-Green I has several disadvantages, including preferential binding to specific DNA sequences, inhibition of amplification reaction at high concentrations, and the need to perform melting or dissociation experiments after the amplification in order to discard non-specific amplifications. The most reliable results for quantification using real-time PCR are obtained by simultaneously performing the detection and confirmation of the new products generated after each new PCR cycle, using probes or primers labelled with fluorescent dyes. The interaction of the probe with the DNA in a specific moment of the PCR cycle induces an increase of the quantum yield. The most common form used to attain the fluorescence signal, is an oligonucleotide that contains both a fluorophore (reporter) that emits fluorescence when excited and a quencher that shields this emission when it is close to the fluorophore (up to 10–100 Å). As a consequence of either the 5’-3’ exonuclease activity of the DNA polymerase (TaqMan, MGB and LNA probes) or specific interactions between the amplicon and the probe.
Molecular Beacons), a physical separation between reporter and quencher takes place, and therefore, an increase in the fluorescence intensity is detected. Many different real-time PCR assays have been published for the quantification of GMOs targeting event-specific sequences. Also, novel fluorescent primer chemistries, namely, Lux, Plexor and AmpliFluor, are a relatively inexpensive alternative to the mentioned fluorescent probes. In a series of reports, the performance of SYBR-Green I, and different fluorogenic primers and probes has been compared in terms of efficiency in PCR amplification and limits of detection and quantification, along with their applicability, evaluated by their practicability and cost. In general, limits of detection values were comparable between chemistries, ranging from 2 to 20 amplifiable target copies. However, careful design and optimization are essential to obtain good performance.

One critical step of any quantitative analytical procedure is the calibration of the system, which is performed by using certified reference materials (CRMs) in which the amount of the analyte of interest is known. However, CRMs are not available for the majority of GMOs. The utility of CRMs is also compromised by the fact that they are prepared from unprocessed raw materials. Quantification based in the direct comparison of results obtained from processed samples and those from CRMs should be taken with care, because matrix effects and the loss of integrity of the DNA molecules during processing could bias the estimation of the actual GMO content of the samples. Novel synthetic DNA standards, including genomic and plasmidic DNA molecules that contain the event-specific and reference sequences of interest, have been also proposed as calibrators in real-time PCR methods for GMO quantification.

Another limitation of GMO quantification methods in food products is related to the
complex zygosity of certain crops. For instance, corn endosperm is triploid whereas embryo and pericarp are diploid and haploid, respectively. In food industry, the transformation of corn grain differs depending on the type of final food product under production. In some cases, endosperm fraction is used for the elaboration of food, after removal of the embryos and seed coat. Therefore, estimation bias is expected when measured GMO content of those products is obtained using CRMs produced from the whole grain.40

3.2 Profiling strategies

The study of biological systems, such as GMOs, entails high complexity and restricts the applicability of target analysis. These difficulties corroborate the need for new and more powerful analytical approaches to study such complexity for comparative safety assessment, and to increase the opportunities to detect unintended effects. As an alternative approach to target analysis, the European Food Safety Authority (EFSA) has recommended the development and use of profiling technologies with the potential to extend the coverage of comparative analyses of GMOs.100 However, it is well-known the impossibility to detect all compounds found in a GMO (or in any other organism) in a single analysis. In consequence, multiple analytical techniques are combined to provide analytical coverage of genes, proteins and metabolites. In this context, Foodomics, defined as a new discipline that studies the food and nutrition domains through the application of advanced omics technologies in order to improve consumers well-being and confidence,101,102 could play an important role in the investigation of GMOs. Thus, Foodomics can provide valuable information about compounds profiling based on genomic/transcriptomic/proteomic and/or metabolomic analysis, information
that could be essential for GMOs traceability and characterization, for the detection of unintended effects and/or the development of new transgenic foods\textsuperscript{103}. A scheme of this approach can be seen in Figure 3, discussing next the different profiling strategies that can be applied.

3.2.1 Transcriptomics

Regarding transcriptomics, for years, the expression of individual genes has been determined by quantification of mRNA with Northern blotting. This classical technique has gradually been replaced by more sensitive techniques such as real-time PCR. It has to be considered that both techniques can only analyse gene expression for a limited number of genes per analysis. This can be very useful to monitor the up- or down-regulation of a given gene for a specific problem. However, it is an important limitation in situations in which the potential up- or down-regulated genes are unknown, since it only provides the analysis of a reduced number of genes. On the other hand, the global analysis of gene expression profiling may offer better opportunities for the comprehensive study of the transcriptome in GMOs. For instance, gene expression microarray has shown to be a valuable profiling method to assess possible unintended effects of genetic transformation in plants. With this technology detailed information has also been obtained on non-targeted effects of transgenes in several plant crops including potato, rice, wheat and maize. In these cases, the genetic modification did not considerably alter overall gene expression, falling within the range of natural variation of the plant varieties\textsuperscript{104,105}, supporting the possibility of producing transgenic plants that are substantially equivalent to non-transformed plants at transcriptomic level.

Although microarray is currently the technique of choice for profiling RNA populations
under different conditions, the new features of next-generation sequencers have
stimulated the development of new techniques that have expanded their applications for
example, to comprehensively map and quantify transcriptomes, for which Sanger
sequencing would not have been economically or logistically practical before\textsuperscript{106,107}.
These novel techniques for transcriptomics have been termed RNA-Seq methods and
are still under active development and evaluation in multiple laboratories for RNA
profiling. They may represent a good alternative for the future comprehensive study of
GMOs at genome and transcriptome level.

\textbf{3.2.2 Proteomics}

Proteomic analysis has become a key technology for the study of differentially
expressed proteins in transgenic food and food ingredients. Two-dimensional gel
electrophoresis (2-DGE), followed by image analysis, and MS (typically MALDI-TOF-
MS) or different variants of LC-MS configure the so-called bottom-up approach. 2-
DGE provides the highest protein-resolution capacity with a low-instrumentation cost.
This strategy has been applied to compare protein profiles of GMOs, including insect-
resistant GM maize\textsuperscript{108,109}, wheat with improved functional properties\textsuperscript{110,111}, tomatoes
with a genetically added resistance to virus and insect attacks\textsuperscript{112,113} and potatoes
showing a delayed sprouting process and modified cell wall structure\textsuperscript{114,115} versus their
corresponding unmodified lines. The comparison of protein profiles of GMOs with
those obtained from the unmodified lines often does not reveal more significant
differences than those observed between different non-modified cultivars/genotypes. In
some cases, the differential expression of the proteins in the GMO is considered
predictable and it can be explained by the result of the genetic modification. However, it
has been highlighted the importance of knowing the extent of natural variation in the
proteome of plants grown under a range of different environments to avoid any mis- or over-interpretation of the results.

In 2-DGE, besides the technical limitations to separate highly hydrophobic, extreme isoelectric point or high molecular weight (MW) proteins, one of the major sources of error is the gel-to-gel variation that makes difficult an exact match of spots in the image-analysis process. Differential in gel electrophoresis (DIGE) can help to circumvent the gel-to-gel irreproducibility for comparative proteomics by loading different samples labeled with ultrahigh-sensitive fluorescent dyes, typically Cy5 and Cy3, in the same gel\textsuperscript{116}. DIGE has been used to compare the proteomes of wild-type cultivars with two GM pea lines expressing α-amylase inhibitor from the common bean\textsuperscript{117} (see Figure 4). Proteins from individual excised spots were digested with trypsin and the resulting peptides were analyzed with LC-ESI-QTOF-MS. Approximately, 600 proteins with MW ranging from 15 to 100 kDa and isoelectric points between 3 and 10 were resolved in the gels. In that study, the gel images for the analysis of one of the GM peas displayed 66 spots showing significant changes. In addition to changes in the abundance of these proteins, complementary analyses suggested post-transcriptional and post-translational modifications of endogenous proteins. Recently, Brandao et al. (2010) have also emphasized the importance of optimizing the parameters that influence the comparisons of the protein map after different gel runs, including those parameters involved in image acquisition\textsuperscript{118}. Using a strictly controlled routine for image analysis of 2-D gels, a maximum of 79\% of spot match was achieved when GM soybean proteome was compared to the corresponding non-modified soybean line.
LC methods with UV or fluorimetric detection have been developed for the protein profiling of GMOs\textsuperscript{119-121}. During the method development stage, different types of columns, including perfusion and monoliths were tested. These methodologies in combination with data analysis using multivariant methods were applied to the identification and quantitation of GMOs. In a different report, protein profiles of transgenic MON810 maize lines has been compared to those obtained from their corresponding unmodified cultivar using LC-ESI-IT-MS in order to investigate possible differences\textsuperscript{122}. The analyses revealed spectral signals that seemed to be very similar between GM lines and the unmodified ones.

CE-ESI-MS has been also applied for the analysis of the intact zein-proteins fraction from three different GM maize cultivars and their corresponding isogenic lines\textsuperscript{123}. A comparative study of two different mass analyzers, namely, TOF and IT, was carried out. Results showed similar sensitivity and repeatability for both instruments; however, CE-ESI-TOF-MS provided better results with regard to the number of identified proteins. A comparison of the protein profiles obtained by CE-ESI-TOF-MS did not show significant differences between the GM lines and their non-modified counterpart. A novel profiling CE-ESI-TOF-MS method, based on shotgun-proteomics strategy, was developed for the investigation of unintended effects in GM soybeans\textsuperscript{124}. In this approach, protein digestion was performed without any prefractionation/separation of the proteome. In this study, several parameters affecting the separation and detection of peptides were studied during the optimization stage of the method. Using this method, a total of 151 peptides were automatically detected for each soybean line (see Figure 4). The comparative analysis showed not differences between the peptide profiles obtained from GM soybean and its conventional counterpart.
### 3.2.3 Metabolomics

The study of metabolome is aimed at the identification and quantification of all small molecules in an organism. Metabolomics, within the frame of GMO analysis, might indicate whether intended and/or unintended effects have taken place as a result of genetic modification\(^{27}\). However, a single method enabling complete metabolome analysis does not exist. Metabolites encompass a wide range of chemical species with divergent physicochemical properties. In addition, the relative concentration of metabolites in a cell or tissue can range from the millimolar to the picomolar level. Accordingly, high sensitivity and resolution are the most relevant parameters to consider for the selection of an appropriate method for comprehensive metabolomic analysis\(^{125}\). In last years, MS-based techniques have shown to offer wide possibilities to evaluate GM crops based on their metabolic profiles, as demonstrated through the large number of applications that use GC-MS, LC-MS, CE-MS, or MS as a stand-alone technique.

GC-MS is one of the most reported analytical tools to study the metabolome of GMOs in the literature. This technique provides high separation efficiency, reproducibility and allows the analysis of primary metabolites such as amino acids, organic acids, and sugars by employing chemical derivatization. In a pioneer work, Roessner et al. (2000) applied GC-MS to characterize the metabolic composition of transgenic potato tubers with modified sugar or starch metabolism\(^{126}\). The identification of 77 out of 150 compounds detected by GC-MS provided valuable information regarding the altered metabolic pathways and unexpected changes in the levels of some compounds in the transgenic tubers. In a separate report, the GC-MS analysis of GM potato tubers with
altered sucrose catabolism indicated an increased level of amino acids\textsuperscript{127}. Further works of the same group demonstrated the suitability of GC-MS in combination with data-mining tools (e.g., principal components analysis (PCA) and hierarchical clustering) to discover differences that enable the discrimination of the transgenic potato and tomato lines from the respective non-modified lines\textsuperscript{128,129}.

Catchpole et al. (2005) used two MS-based techniques to obtain complementary data regarding the compositional similarities/differences between transgenic potato designed to contain high levels of inulin-type fructans and its conventional counterpart\textsuperscript{130}. Initially, flow-injection analysis (FIA) ESI-MS was used to analyze 600 potato extracts. Data sets were analyzed by principal component analysis (PCA) to identify top-ranking ions for genotype identification. Then, further GC-TOF-MS profiling of more than 2000 tuber samples provided complementary data covering 242 individual metabolites (90 positively identified, 89 assigned to a specific metabolite class, and 73 unknown). In a separate report, the unintended effects in insect-resistant GM rice have been also investigated by Zhou et al. (2009) by means of the combined application of GC-flame ionization detection (FID) and GC-MS\textsuperscript{131}. In this case, however, GC-MS was exclusively used to identify certain important compounds after GC-FID profiling.

Volatile aroma compounds, such as aldehydes and alcohols, are secondary metabolites influenced by a number of variables, including genetic makeup and abiotic factors. GC-MS has demonstrated to be a valuable tool for profiling aroma compounds in transgenic fruits and vegetables. Thus, volatile fraction of GM raspberries with added resistance to virus attack has been investigated using GC-MS\textsuperscript{132}. The quantitative study of 30 selected compounds belonging to various chemical classes (e.g., alcohol, aldehyde,
ketone, ester and terpene) did not show significant differences between the GM line and the wild-type. Similarly, the investigation of aroma compounds profiles of four lines of GM cucumber and their unmodified lines using GC-MS only showed quantitative differences\textsuperscript{133}. The combinations of GC-MS with several selective extraction methods using supercritical fluids or accelerated solvents have been also explored for the investigation of unintended effects in GMOs\textsuperscript{134,135}. These techniques have been applied to extract selectively amino acids and fatty acids from soybean and maize for subsequent profiling and quantification.

LC-MS has been used for the study of flavonoids in transgenic rice and wheat\textsuperscript{25,136}. Also, a novel LC-MS method has been developed for the profiling of stilbenes, a specific class of polyphenols, in transgenic tomato overexpressing a grapevine gene that encoded the enzyme stilbene synthase\textsuperscript{137}. Using this methodology, differences in the concentration of rutin, naringenin, and chlorogenic acid were detected when transgenic tomatoes were compared to the control tomato lines. The combined use of LC-MS with GC-MS has demonstrated to improve the description of the metabolome status of GMOs. In this regard, differences in some phenolic compounds and volatile secondary metabolites that belong to the classes of monoterpenes, C12-norisoprenoids, and shikimates were detected using LC-ESI-IT-MS and GC-MS for the comparative analysis of GM grapevine lines with the unmodified control\textsuperscript{138}. Although reversed-phase is the most frequent mode used in LC-MS metabolite profiling in GMO analysis, other suitable modes have proven to be useful. For example, the levels of the major carbon metabolites in transgenic rice have been determined using a HILIC phase for the separation in LC-ESI-MS/MS\textsuperscript{139}. In a recent paper, LC-ESI-Q-MS has also been applied to the study the effect of gene encoding a feedback insensitive α-subunit of anthranilate
synthase expression on the metabolic profile of GM rice with increased tryptophan (Trp) content\textsuperscript{140}. Different plant tissues were analyzed and then, metabolic profile data were analyzed using different statistical methods to determine the peaks that characterize the difference between GM rice and the unmodified counterpart. Results obtained in the study also indicated that the concentration of Trp changes in a time-dependent manner showing a tissue-dependent profile of accumulation.

CE-MS has demonstrated to be suitable for the analysis of a wide range of analytes including ionic and polar thermolabile compounds, being considered complementary to LC-MS and GC-MS. High efficiency, analysis speed and resolution are characteristics of CE-MS, requiring moreover little sample pretreatment. On the other hand, moderate sensitivity is often achieved due to the minute sample volumes injected in CE-MS. The potential of CE-MS for metabolic profiling of GMOs has already been demonstrated. For instance, novel methods for metabolite profiling of GM maize and soybean, have been investigated in the last years\textsuperscript{141,142}. Thus, CE-ESI-TOF-MS was used to determine statistically significant differences in metabolic profile of varieties of conventional and insect-resistant GM maize\textsuperscript{141}. Treatment of the CE-MS data using PCA indicated some statistically significant differences in the levels of L-carnitine and stachydrine between conventional and GM maize. A similar CE-ESI-TOF-MS methodology was developed for the comparative analysis of metabolic profiles from GM soybean (glyphosate resistant) and its corresponding unmodified parental line\textsuperscript{142}. In that study, over 45 different metabolites, including isoflavones, amino acids and carboxylic acids were tentatively identified. Differences in metabolic profiles of both lines were more evident on the concentration of three free amino acids (proline, histidine and asparagine) while a metabolite tentatively identified as 4-hydroxi-L-threonine disappeared in the transgenic
soybean compared to its parental non-transgenic line. A chiral CE-ESI-TOF-MS method has also been developed to study differences in the chiral amino acid profile between varieties of conventional and herbicide-tolerant transgenic soybean showing some quantitative variations.

Fourier transform ion-cyclotron MS (FT-ICR-MS) has already been used as a powerful analytical platform for metabolomic studies in GMOs. This technique has been used also in combination with CE-TOF-MS for the metabolomic profiling of six varieties of maize, three GM insect-resistant lines and their corresponding isogenic lines. The FT-ICR-MS data obtained in both positive and negative ESI mode were uploaded into MassTRIX server in order to identify maize specific metabolites annotated in the KEGG database. Multivariate analysis of data by partial least squares-discriminant analysis (PLS-DA) showed the most discriminative masses that contribute to differentiate the GMO samples from their isogenic lines. Such discriminant m/z values were uploaded in MassTRIX showing the total number of compounds identified and present in Z. mays (see Figure 5). Despite of the good mass resolution and accuracy offered by FT-ICR-MS, certain compounds could not be unequivocally identified, since FT-ICR-MS cannot differentiate isomers having the same molecular formula, so that electrophoretic mobilities and m/z values provided by CE-TOF-MS were used to confirm the identity of various isomeric compounds.

4. CONCLUSIONS AND FUTURE OUTLOOKS

The emergent market of novel foods and food products containing GMOs including the so-called second generation GMOs will maintain the need for new analytical methods allowing their qualitative (target or profiling) and quantitative analysis. In order to
provide the control organisms with the appropriate tools, and given the path of the evolution of the GMO market, future research in this field will keep focusing on methods able to reliably detect, characterize and quantify GMOs. Moreover, these methodologies will pass through validation process in order to be confidently used by enforcement and commercial laboratories. In this context, both target and profiling strategies are expected to keep playing a definitive role.

Besides, more work will also be needed to characterize natural variability of crops to make easier the identification of any unintended effect or GM crop. The definition of common standardized experimental protocols is a major challenge in omics strategies for which Foodomics can be the right framework. Unifying analytical platforms and protocols will allow the comparison of experiments performed in laboratories worldwide. In addition, much effort is needed to integrate proteomics and metabolomics with genomics data. This integration will involve a vast quantity of collaborative work to compare and share data within the scientific community.

ACKNOWLEDGEMENTS

This work was supported by Projects AGL2008-05108-C03-01, 200970I083 and CONSOLIDER INGENIO 2010 CSD2007-00063 FUN-C-FOOD (Ministerio de Educación y Ciencia).

REFERENCES

1. Petit, L., Pagny, G., Baraige, F., Nignon, A.C., Zhang D. Characterization of


21. Fitch, M.M., Manshardt, R.M., Gonsalves, D., Slightom, J.L., Sanford, J.C. Virus resistant papaya plants derived from tissues bombarded with the coat protein gene of
papaya ringspot virus, Biotechnology 1992, 10 (11), 1466-1472.


27. Shepherd, L.V.T., McNicol, J.W., Razzo, R., Taylor, M.A., Davies, H.V. Assessing the potential for unintended effects in genetically modified potatoes perturbed in


33. García-Cañas, V., Cifuentes, A., González, R. Detection of genetically modified organisms in foods by DNA amplification techniques. Critical Reviews in Food Science
and Nutrition 2004, 44 (6), 425-436.


51. Peano, C., Bordoni, R., Gulli, M., Mezzelani, A., Samson, M.C., De Bellis, G.,


56. Nadal, A., Coll, A., La Paz, J.L., Esteve, T., Pla, M., A new PCR-CGE (size and


71. Rudi, K., Rud, I., Holck, A. A novel multiplex quantitative DNA array based PCR (MQDA-PCR) for quantification of transgenic maize in food and feed. Nucleic Acids Research 2003, 31 (11), e62.

72. Germini, A., Rossi, S., Zanetti, a., Corradini, R., Fogher, C., Marachelli, R. Development of a peptide nucleic acid array platform for the detection of genetically


75. Dobnik, D., Morisset, B., Gruden, K. NAIMA as a solution for future GMO diagnostics challenges. Analytical and Bioanalytical Chemistry 2009, 396 (6), 2229-2233.


95. Taverniers, I., Van Bockstaele, E., De Loose, M. Cloned plasmid DNA fragments as
calibrators for controlling GMOs: different real-time duplex quantitative PCR methods. 

96. Taverniers, I., Windels, P., Vaitilingom, M., Milcamps, A., Van Bockstaele, E., Van 
    den Eede, G., De Loose, M. Event-Specific Plasmid Standards and Real-Time PCR 
    Methods for Transgenic Bt11, Bt176, and GA21 Maize and Transgenic GT73 Canola. 

97. Mattarucchi, E., Weighardt, F., Barbati, C., Querci, M., Van den Eede, G. 
    Development and applications of real-time PCR standards for GMO quantification 
    based on tandem marker plasmids. European Food Research & Technology 2005, 221 
    (3-4), 511-519.

98. Wang, S., Li, X., Yang, L, Shen, K., Zhang, D. Development and in-house 
    validation of a reference molecule pMIR604 for simplex and duplex event-specific 
    identification and quantification of GM maize MIR604. European Food Research & 

99. Lievens, A, Bellocchi, G., De Bernardi, D., Moens, W., Savini, C., Mazzara, M., 
    Van den Eede, G., Van den bulcke, M. Use of pJANUS™-02-001 as a calibrator 
    plasmid for Roundup Ready soybean event GTS-40-3-2 detection: an interlaboratory 
    trial assessment. Analytical and Bioanalytical Chemistry 2010, 396 (6), 2165-2173.

100. European Food Safety Agency. Guidance document of the scientific panel on 
    genetically modified organisms for the risk assessment of genetically modified plants
and derived food and feed. EFSA Communications Department, Parma, Italy. 2006.


4 (1), 193-200.


118. Brandao, A.R., Barbosa, H.S., Arruda, M.A.Z. Image analysis of two-dimensional gel electrophoresis for comparative proteomics of transgenic and non-transgenic
soybean seeds, Journal of Proteomics 2010, 73 (8), 1433-1440.


FIGURE LEGENDS

Figure 1. CGE-LIF analysis of MLGA reactions of maize DNA sample containing 1% MON810, 1% MON863 and 1% GA21 maize DNA. Figure shows the electrophoregram obtained before (A) and after (B) optimization of the ligation conditions. Redrawn from ref.64.

Figure 2. An example of a microarray scanning image after the hybridization of the screening triplex NAIMA product performed on Mon863 (10%) reference material78. Reproduced with permission.

Figure 3. Ideal Foodomics platform to analyze genetically modified foods.

Figure 4. DIGE analysis of pea seed proteins. Upper panel: non-transgenic seed proteins (labeled with Cy5, red) and transgenic (Cy3, green). Lower panel: Coomassie-stained gels corresponding to the analytical gels in the upper panel117. Copyright Wiley-VCH Verlag GmbH & Co. KGaA. Reproduced with permission.

Figure 5. PLS-DA model with six different maize varieties analyzed by FT-ICR-MS. Samples (A, C and E) non-modified maize lines; (B, D and F). The score scatter plot underlines a different pattern for the transgenic (blue) and isogenic lines (red). The different properties of the discriminative masses (represented in blue and red in the loading plot) are investigated with MassTRIX. Redrawn from ref. 147.