Scientific requirements for GMO risk assessment and future perspectives

Project responsible
Antonio Fernandez Dumont and Claudia Paoletti

Contributors
Stefania Favilla, Silvia Federici, Michele Ardizzone, Andrea Gennaro, Jose Gomez, Anna Lanzoni and Konstantinos Paraskevopoulos

INDEX

1. The Scientific requirements for GMO Risk Assessment: food and feed safety

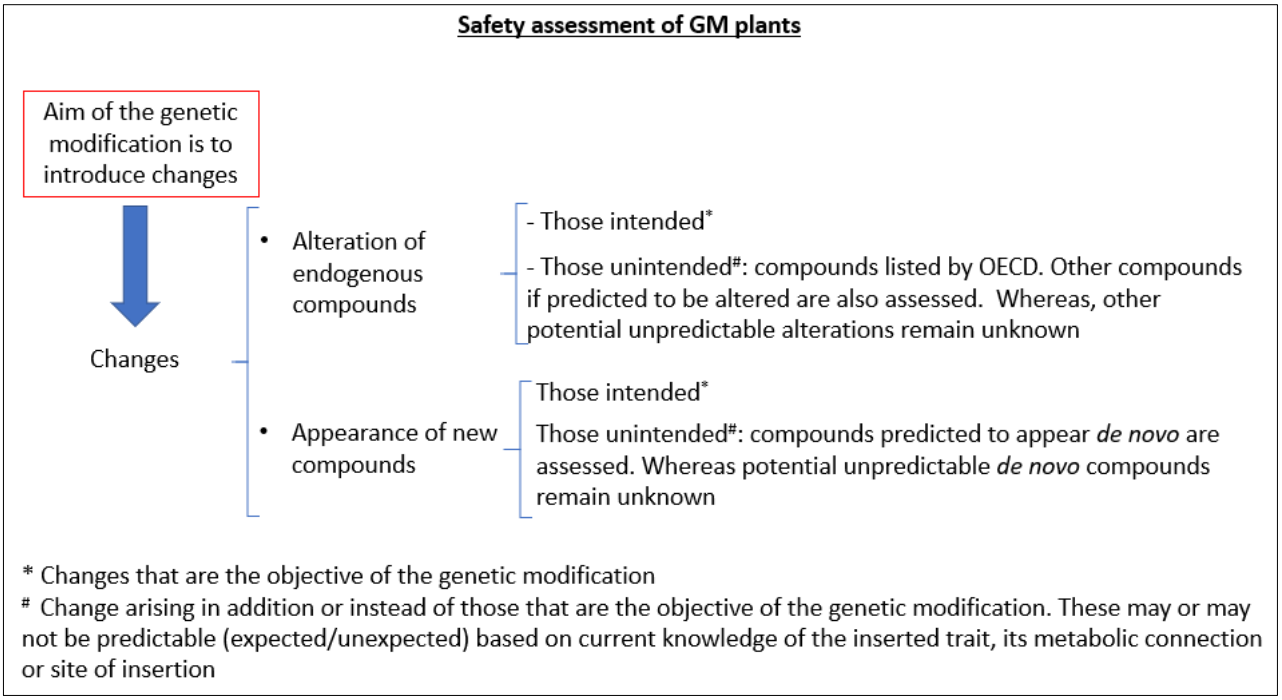
1.1 Introduction

This document focuses on the scientific requirements that the EU has in place for the risk assessment of GM plants, which are regularly updated to ensure compliance with scientific developments. As it will become clear reading, GM plant risk assessment requirements in the EU are in line with internationally agreed approaches and procedures and Codex Alimentarius principles (Codex Alimentarius, 2009).

A request for authorisation of a GM plant on the EU market starts when an applicant prepares an application (also called technical dossier) to be submitted for safety evaluation. The application must define the scope of the authorisation request (see below for further details) and must include studies and data to identify and evaluate any possible adverse effects of the GM plant, ultimately necessary to demonstrate its safety for humans, animals and the environment. Given that environmental risk assessment considerations are out of the scope of the present document, below we focus exclusively on the food and feed safety of applications submitted under the frame of Reg. (EC) 1829/2003.

Annex II of Reg. (EU) 503/2013 provides a detailed overview of the different components of the risk assessment of GM plants and derived food and feed, submitted within the framework of Regulation (EC) No 1829/2003. The technical basis used to develop Reg. (EU) 503/2013 was the EFSA guidance document '*for risk assessment of food and feed from genetically modified plants*' (EFSA, 2011a). Both Annex II of Reg. (EU) 503/2013 and EFSA guidance document are valuable references where additional detailed information in support of the present document can be found.

In line with Codex Alimentarius principle (Codex Alimentarius, 2009) pillar of the EU risk assessment is the comparative approach, according to which the safety of a GM plant can be established by comparison with a non-GM comparator that can be deemed as safe: if it can be demonstrated that the two have similar characteristics, the GM plant can be considered as safe as the non-GM comparator. If changes are observed, these are considered indicators of potential effects which are then evaluated in the subsequent steps of the risk assessment process. Therefore, GM plants' safety is not pursued in absolute terms, rather established relatively to what is widespread, used, consumed and considered safe. From a broader perspective, a plant is genetically modified to introduce a change which, at the present state of technical possibilities, can be either the production of a new compound (e.g. a protein not normally expressed in the crop, conferring resistance to a herbicide) or the alteration of an endogenous one (e.g. the so called 'nutritionally-enhanced' crops in which, for example, the fatty acid profile is modified to enhance their nutritional characteristics). However, new endpoints can be introduced (or endogenous ones altered) also unintentionally, because of possible effects which were not the objective of the genetic modification, but that may happen in complex biological systems like plants. The extent to which these unintended effects can be predicted/anticipated strictly depends on the degree of knowledge of the inserted trait -i.e. the objective of the genetic modification - its metabolic connection(s) and its site of insertion within the genome. The scheme below illustrates the possible scenarios.



The information that must be provided in a GM plant application in the EU includes: 1) the molecular characterization of the genetic modification, 2) the comparative analysis of the agronomic, phenotypic characteristics and compositional endogenous endpoints of the GM plant with respect to its non-GM comparator; 3) the toxicological assessment of the protein newly expressed in the GM plant, of any new constituent other than the newly expressed protein and of the altered compositional endogenous endpoints detected in the comparative analysis; 4) the allergenicity assessment of the protein newly expressed in the GM plant, of any new constituent other than the newly expressed protein and of the altered compositional endogenous endpoints detected in the comparative analysis; 5) the nutritional assessment of the GM food and feed; 6) the dietary exposure to the GM food and feed; 7) when appropriate, a post market monitoring proposal.

Below the information/data required for GM plant applications in the EU are schematically summarised. Each topic is addressed in further detail Sections 2.2 – 2.8.

Molecular Characterization	• Identifies the actual nucleotide sequence of the genetic modification (what is inserted). • Identifies the exact location in the recipient genome where the new genetic material is integrated (where it is inserted). • Verifies if other genetic material besides the one intended for insertion - e.g. part(s) of the DNA molecule (such as a 'plasmid') used to generate the GM plant - has been unintentionally integrated in the recipient plant genome. • Evaluates the genetic stability of the inserted genetic material within the recipient plant genome, i.e. if it is stably inherited in subsequent generations. • Assesses if and to what extent the newly inserted genetic material is expressed in the recipient plant. • Assesses the molecular and biochemical properties of the product (e.g. of the newly expressed protein) resulting from the expression of the inserted genetic material (e.g. the amino acid sequence or the biological function of the newly expressed protein).

	• Analyses the amino acid sequence of all potential peptides that could be unintentionally produced as a result of the genetic material insertion (i.e. open reading frames) for similarity to known allergenic and toxic proteins.
Comparative Assessment of endogenous endpoints	Identifies possible changes in endogenous compositional endpoints, agronomic and phenotypic characteristics between the GM plant and its non-GM comparator. Such changes are considered indicators of possible effects of the genetic modification. The difficulties to identify unintended effects are well known (see Section 2.3; Fernandez and Paoletti, 2018). To this aim, the EU has put in place an approach requiring the simultaneous application of two complementary statistical tests to evaluate changes observed with respect to the non-GM comparator (test of difference) taking into account the natural variability of each measured endpoint estimated from non-GM commercial reference varieties (test of equivalence). Comparative assessment conclusions are based on the integrated evaluation of the outcomes of these two tests: a difference between the GM plant and its comparator for a given endpoint is considered relevant if the value of that endpoint in the GM plant falls outside the distribution of values of the reference varieties – i.e. the interval identified by the equivalence test.
Toxicological Assessment of new and endogenous endpoints	Addresses the impact on human and animal health of the observed changes: the protein newly expressed in the GM plant, any new constituent other than the newly expressed protein and any relevant change in endogenous compositional endpoints identified during the molecular characterisation and/or the comparative assessment. Toxicological assessment conclusions are based on an integrated evaluation of the observed changes to establish the overall safety of GM food and feed.
Allergenicity Assessment of new and endogenous endpoints	Addresses the impact on human and animal health of the observed changes: the protein newly expressed in the GM plant, any new constituent other than the newly expressed protein and any relevant change in endogenous compositional endpoints identified during the molecular characterisation and/or the comparative assessment. Allergenicity assessment conclusions are based on an integrated evaluation of the observed changes to establish the overall safety of GM food and feed.
Nutritional Assessment of new and endogenous endpoints	Demonstrates that the food and feed derived from the GM plant is nutritionally equivalent to that derived from its non-GM comparator, addressing therefore the possible nutritional impact of the protein newly expressed in the GM plant, any new constituent other than the newly expressed protein and any relevant change in endogenous compositional endpoints identified during the molecular characterisation and/or the comparative assessment.

Dietary Exposure Assessment	Characterizes the potential risk associated to the consumption of GM plants based on expected intake levels of a GM food and/or feed obtained from representative consumption data, reflecting specific food and feed habits of the human and animal population under consideration.
Post Market Monitoring	Possible changes in GM food and feed intake patterns that may happen after marketing are determined by post-market monitoring activities. Post-market monitoring is needed whenever changes in levels of relevant constituents and/or in the overall dietary intake patterns of GM food and/or feed affecting dietary exposure estimates can be expected, as these may influence the outcome of the nutritional safety assessment *a posteriori*.

1.2 Molecular Characterization

Purpose is to characterize the molecular profile of the GM plant on the basis of a dataset which includes: data concerning the insertion of the genetic material within the genome of the recipient plant, data on the product derived from the expression of the genetic material inserted in the recipient plant (usually a protein), as well as data/information on the characteristics of the donor (the source of genetic material) and of the recipient (the plant genetically manipulated). In addition, the similarity to known allergenic and toxic proteins is assessed for all the potential peptides that could be unintentionally produced as a result of the genetic material insertion (i.e. open reading frames).

The data concerning the insertion of the new genetic material in the recipient plant genome, allows answering the following questions:

- *What is the structure of the inserted genetic material and its location(s) in the recipient plant genome?* This information can be obtained by several molecular methods such as Southern blot analysis and nucleotide sequencing. Southern blot analysis can be used to: a) demonstrate that the genetic material has been integrated in the plant genome; b) identify the number of genomic locations where the insertion occurred; as well as c) determine the number of times (how many copies) the genetic material was inserted in each genomic location. In addition, Southern blot analysis allows the detection of genetic material other than the one intended for insertion that may have been unintentionally integrated in the receiving genome, e.g. portions of the DNA molecule (such as a 'plasmid') used to generate the GM plant. The assessment of unintended genetic material insertion(s) is a key aspect of the molecular characterization, which can be also addressed by applying sequencing methods (EFSA, 2018a). Sequencing methods are applied to determine the exact nucleotide sequence of the inserted genetic material and of the two genomic regions flanking the insertion. These two sequence elements (inserted genetic material and genomic flanking regions) constitute what is normally called the 'GM event'.
  In silico methods are used to annotate all the genetic elements comprising the GM event and to assess if there is a likely impact on the function of known or predictable endogenous plant genes by the insertion of the new genetic material.

- *Is the insertion of the new genetic material stable?* The structure of the newly inserted genetic material, its position in the recipient plant genome and its expression in the recipient plant need to be stably

inherited from one generation to the next. This is verified collecting data from GM plants across several subsequent generations, typically at least five.

The data concerning the product (e.g. the newly expressed protein) derived from the expression of the newly inserted genetic material, allows answering the following questions:

- *Does the expression of the newly inserted genetic material result in the intended product and, if so, what are the levels of this product in the GM plant?* The methodology applied to provide the experimental evidence needed must be specific, reproducible and reliable (EFSA, 2018b).

- *What are the biochemical and molecular properties of the product (e.g. newly expressed protein) derived from the expression of the newly inserted genetic material?* The information documenting the molecular profile of the product present in the GM plant includes information on its biochemical (e.g. what is the molecular mass of the produced protein), structural (e.g. what is the amino acid sequence of the produced protein) and functional (e.g. what is the enzymatic activity) properties.

The assessment of all the potential peptides that could be unintentionally produced as a result of the genetic material insertion starts from the identification of all open reading frames (ORFs) present within the inserted genetic material sequence, within the sequence junction regions of the inserted genetic material and the recipient plant genome. An ORF can be defined as any nucleotide sequence consisting of a string of codons in the same 'reading frame', uninterrupted by the presence of a 'stop' codon. Then *in silico* methods, using up to date databases publicly available, allow to analyse the extent of amino acid sequence similarity between these potentially produced peptides and known allergenic or toxic proteins.

1.3 Comparative Assessment

Purpose of the comparative assessment is to identify changes in the composition, agronomic performance and phenotypic characteristics in the GM plant (and derived food and feed) with respect to its non-GM comparator. These changes, if found, are further assessed with respect to potential impact on human and animal health.

Clearly, the selection of appropriate non-GM comparators is central to the correct application of the comparative approach. Codex Alimentarius (Codex Alimentarius, 2009) indicates that the non-GM comparator should be selected according to its genetic similarity to the GM line under assessment (the genetically closer the better), but it also recommends that the non-GM comparator should have a history of safe use. The experience gained over 20 years of GM plant risk assessment indicates that it is difficult for a non-GM comparator to fulfil both expectations, i.e. be a near-isogenic line and have a history of safe use, because often the lines that are genetically modified are laboratory lines with no history of use, that alone history of safe use (Fernandez and Paoletti, 2018). In the EU, the EFSA developed specific guidance to address this issue (EFSA, 2011a and EFSA, 2011b) describing the criteria for selecting the most suitable comparator for GM plant risk assessment under different scenarios. In particular the EFSA (EFSA, 2011b) clarifies the possible confusion among the various terms that have been used synonymously in the pertinent literature to describe non-GM comparators used in the risk assessment of GM plants: these include, among others, the terms 'control', 'non-GM comparator' and 'conventional counterpart'. The EFSA recommends (EFSA, 2011b) the use of the term 'conventional counterpart' only when referring to a non-GM comparator that:

i) in the case of vegetatively propagated crops, is the non-GM isogenic line;
ii) in the case of crops that are propagated sexually, is the non-GM genotype with a genetic background as close as possible to that of the GM plant.

In all other cases, i.e. all cases in which the comparative assessment includes genotypes which do not fit the definition of conventional counterpart provided above, the term 'comparator' should be used. For simplicity, in the present document we only use the term 'comparator' as all what is described would apply not only to the 'conventional counterpart', but also to the broader category of 'comparator'.

EFSA comparative assessment approach requires the inclusion not only of a non-GM comparator, but also of non-GM commercial varieties – i.e. reference varieties - which can be securely assumed to have a history of safe use, since consumers are regularly exposed to varieties commercially available. The inclusion of reference varieties serves two purposes: first it allows comparing the GM plant to non-GM varieties with proven history of safe use; second it allows measuring the variation of the endpoints selected for the comparative assessment among non-GM varieties, providing a quantitative estimation of their variability which, in itself, provides an effective baseline for risk assessment decision-making. Certainly, variability estimates based on a limited number of commercial varieties are conservative as they can capture only a portion of the natural variation, but this is a portion to which consumers are exposed and, as such, already accepted as safe. This approach is incorporated into Regulation (EU) 503/2013 and it outstands as the first regulatory effort to quantify the degree of natural variability of compositional, agronomic and phenotypic characteristics.

The other central element for the correct application of the comparative assessment is the design of comparative field trial studies and the statistical analysis of the collected data. In the EU minimum requirements for the design of field trials have been established (EFSA, 2010a; van der Voet *et al.*, 2011; Regulation (EU) 503/2013) to: 1) ensure harmonization across GM plant applications with the concurrent assessment of the GM plant, its non-GM comparator and the non-GM reference varieties in the same study under the same conditions to minimize confounding effects; and 2) maximize statistical power to detect unintended effects.

1.3.1 Design of comparative field trial studies

The design of the field trial is described in detail in the EFSA pertinent guidance documents (EFSA, 2010a; EFSA, 2011a; EFSA, 2015) and in Reg. (EU) 503/2013. Overall, a comparative field trial study must include at least eight sites, characterised by environmental conditions (soil characteristics, climatic conditions, day length, soil moisture, biotic factors such as presence/absence of pests) representative of the environments in which the GM crop is expected to be grown. Clearly, the selection of sites is closely related and interdependent to the selection of the line used to test the genetic transformation because, in most situations, a given GM line is not appropriate to be grown in all the possible environments in which the introduction of the GM event may elicit an agronomic/commercial interest. To partially mitigate the impossibility of testing all possible varieties, the use of a variety/line with high adaptability would be preferable as it would allow testing a wider range of environmental conditions. The selection of the non-GM reference varieties to be included in the field trial follows a similar logic: among the available commercial varieties those most appropriate/suitable for the environmental conditions and customary agronomic practices of the chosen sites should be included in the field trial study. *A priori* standard guidance for such decision cannot be made given the intrinsic complexity of the topic and the specificity of

the considerations to be made on a case-by-case basis. Therefore, explicit justification of the criteria followed to make this educated decision must be provided in each application/dossier.

The field trial can be conducted in a single year or spread over multiple years. Replication should never be less than four at each site. The reference-varieties may vary between sites and at least six different reference varieties must be included overall in the field trial. All test materials (the GM plant, its comparator and the reference-varieties) should be randomised to plots at each site, usually in a complete randomised or randomised block design.

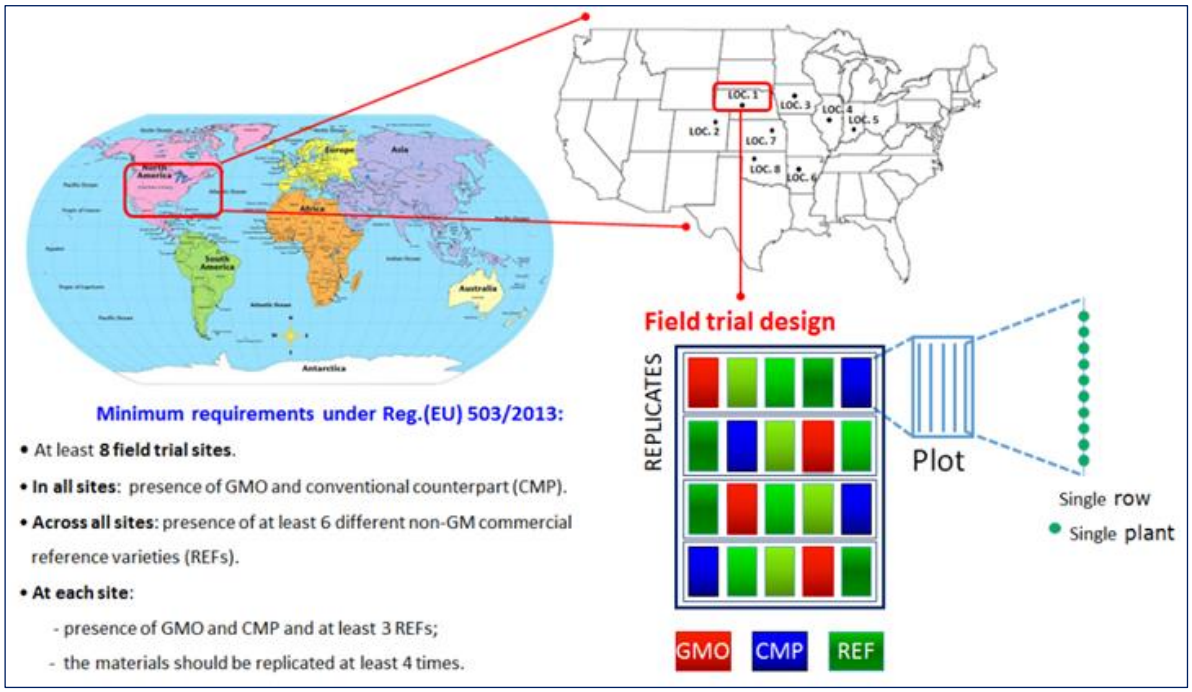


Figure 2.1: EU minimum requirements for comparative field trial studies design

1.3.2 Selection of endpoints to be measured

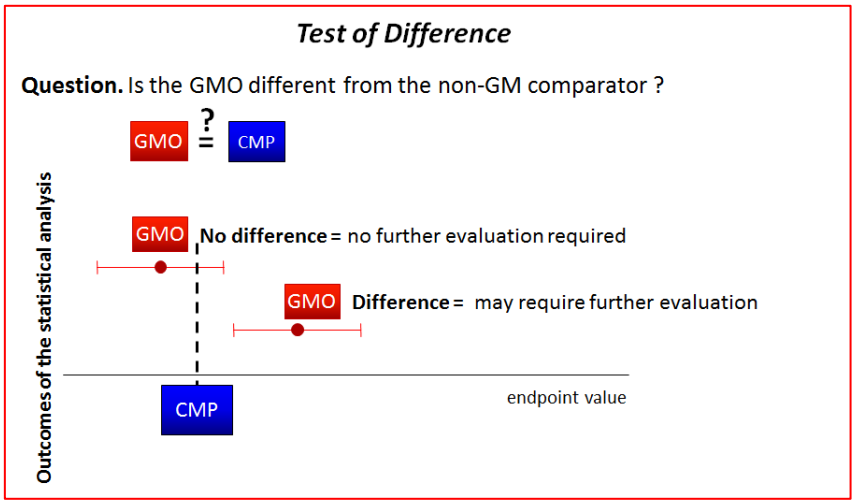
According to EU requirements (EFSA, 2011a; Reg(EU) 503/2013), the compositional endogenous endpoints to be measured in comparative studies must always include all the endpoints listed in the available respective crop-specific OECD consensus documents (http://www.oecd.org/chemicalsafety/biotrack/consensus-document-for-work-on-safety-novel-and-foods-feeds-plants.htm). Additional endpoints may be needed on a case-by-case basis, depending upon the type of genetic modification, the crop being modified, or any other specific factor to be considered during the risk assessment. In addition, a range of agronomic and phenotypic endpoints are also measured to characterise the GM plant's biology and performance, normally covering characteristics such as plant vigour, growth and development, morphology, yield, pest and disease susceptibility. However, the specific list of agronomic and phenotypic endpoints to be considered in each field trial should be compiled considering the biological characteristics of the crop and the trait intentionally introduced into the GM plant.

1.3.3 Data Analysis

The statistical approach chosen for the data analysis – i.e. the simultaneous application of the test of difference and the test of equivalence (EFSA, 2010a; EFSA, 2011a) – enables to evaluate the relevance of

the differences observed between the GM plant and its non-GM comparator considering endpoints' variability. In particular:

The test of difference (see EFSA, 2010a for technical statistical details) is used to verify whether the GM plant, apart from the introduced genetic modification(s), is different from its non-GM comparator.

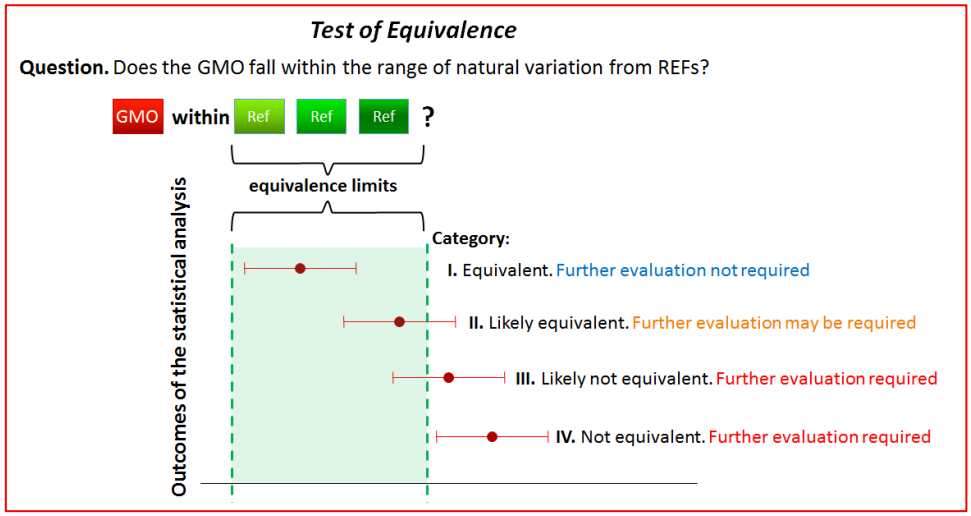


Simplified illustration of the test of difference. Shown are the mean of the GM plant (circle), the confidence limits (whiskers) for the difference between the GM plant and its comparator and a vertical dotted line indicating zero difference. The two possible outcomes are:

- *No difference: the confidence interval bar overlaps with the line of no-difference. The null hypothesis of no difference cannot be rejected, and the appropriate conclusion is that there is insufficient evidence that the GM plant and its comparator differ.*

- *Difference: the confidence interval bar does not overlap with the line of no-difference. The null hypothesis of no difference must be rejected, and the appropriate conclusion is that the GM plant is different from its comparator.*

The test of equivalence (see EFSA, 2010a for technical statistical details) is used to verify whether the agronomic, phenotypic and compositional characteristics of the GM plant fall within the variability estimated from the set of non-GM reference varieties included in the field trial.



Simplified illustration of the test of equivalence. Shown are the mean of the GM plant (circle), the confidence limits (whiskers) for the difference between the GM plant and its comparator and the vertical lines indicating equivalence limits (for test of equivalence). Four categories i - iv are identified. In category:

i. *The null hypothesis of non-equivalence is rejected in favour of equivalence (both confidence limits lie between the equivalence limits and the null hypothesis of non-equivalence is rejected. The appropriate conclusion is that the GM plant is equivalent to the set of non-GM reference varieties).*

ii. *The mean of the GM plant lies between the equivalence limits, but the confidence interval bar overlaps at least one of the equivalence limits on the graph. Non-equivalence cannot be rejected. The appropriate conclusion is that equivalence between the GM plant and the set of non-GM reference varieties is more likely to be the case than lack of equivalence. Further evaluation may be required.*

iii. *The mean of the GM plant lies outside the equivalence limits, but the confidence interval bar overlaps with at least one of the equivalence limits. Non-equivalence cannot be rejected, and the appropriate conclusion is that equivalence between the GM plant and the set of non-GM reference varieties is less likely to be the case than lack of equivalence. Further evaluation is required.*

iv. *Both confidence limits lie outside the equivalence limits. The appropriate conclusion is that the evidence analysed here demonstrates non-equivalence between the GM plant and the set of non-GM reference varieties. Further evaluation is required.*

The conclusions of the comparative assessment indicate if the agronomic, phenotypic and compositional characteristics of the GM plant differ from those of the non-GM comparator and of the reference varieties. If relevant changes are observed – i.e. differences where the value of a given endpoint in the GM plant falls outside the equivalence interval defined by the reference varieties - these are evaluated, as appropriate, in the subsequent steps of the risk assessment process. This approach is rooted in the concept of substantial equivalence, initially developed by OECD in 1993 and pillar of Codex Alimentarius risk assessment strategy, which establishes that a GM-derived food can be considered as safe as a conventional food if it can be demonstrated that its characteristics and composition are similar to those of the conventional food, apart from the change(s) introduced with the genetic modification which is(are) independently assessed following a weight-of-evidence approach.

1.3.4 Effect of processing

Before entering the core of the food and feed safety assessment, the possible effect of processing is assessed to verify whether the characteristics of products derived from the GM plant are likely to be modified with respect to those of products derived from non-GM varieties of the same crop. The focus is mainly on product content, quality and purity. When assessing the effect of processing the outcome of the comparative assessment is the key starting point: if the composition of the GM plant is similar to that of the non-GM comparator, it can be assumed that processing will not result in products different from those of non-GM plant varieties, i.e. the processing outcomes will not be affected by the traits introduced in the GM plant. If instead the comparative assessment identified changes, it must be evaluated if and how processing will be affected by such changes.

1.4 Toxicology

The toxicological assessment verifies whether the intended change and relevant unintended changes identified during the preceding steps of the risk assessment (molecular characterization and comparative assessment) may have adverse effects on human and animal health. In line with Codex Alimentarius, the toxicological assessment in the EU focuses on the protein(s) newly expressed in the GM plant, on any new constituent other than the newly expressed protein(s), and on endogenous compositional endpoints showing altered levels. The toxicological assessment of these should be evaluated considering their anticipated intake via the consumption of genetically modified food and feed (see Section 2.7.1).

Assessment of newly expressed protein(s) - in accordance with Codex Alimentarius (2009), the assessment of newly expressed protein(s) is based on the available knowledge of the protein(s) with respect to its/their source, function/activity and history of consumption, as well as on studies conducted *in silico*, *in vitro* and *in vivo* which are selected on a case-by-case basis. The relevance of these studies and the criteria followed to select, execute and interpret them are addressed in Section 2.6 of this document.

Assessment of new constituents other than the newly expressed protein(s) - the safety assessment of new constituents other than protein(s) is conducted selecting, on a case-by-case basis, the most appropriate study(ies) with respect to the nature and type of novel constituent(s) to be assessed, accounting for specific needs, considerations and peculiarities of each situation.

Assessment of endogenous compositional endpoints – alterations of endogenous compositional endpoints beyond their degree of natural variability (see Section 2.3 on comparative assessment) are assessed with respect to their safety for human and animal health, taking into account the available knowledge on their toxicological profile and the anticipated exposure via consumption of GM food and feed.

The need for additional toxicity studies is established based on the outcome of the assessment of newly expressed protein(s), new constituent other than proteins and altered level of endogenous endpoints mentioned above. Therefore, the battery of studies for the toxicological assessment is not established *a priori*, rather decided *de novo* each time to account for specific needs, considerations and peculiarities. Notwithstanding, available standardised test protocols for the different types of toxicological studies that are considered appropriate on a case-by-case basis should always be used to maximize reliability of results and findings. A summary of these protocols is provided in EFSA guidance (EFSA, 2011a).

Finally, a 90-day feeding study in rodents is required in the EU whenever remaining uncertainties identified during the risk assessment need to be resolved. However, given the difficulty of defining with sufficient precision the level of uncertainty of the risk assessment process that would trigger the request for a 90-day feeding study, the submission of such a study is currently mandatory under Reg. (EU) 503/2013 for all GM plant application containing a single event (for an explanation on the distinction between single and stack event please see Section 3.1). The 90-day feeding study design (EFSA, 2011c) is required to be in line with OECD TG 408 (OECD, 1998) and further considerations on the design of 90-days feeding studies in GM plant risk assessment are provided in EFSA technical report (EFSA, 2014).

The conclusions of the toxicological assessment indicate if the intended change of the genetic modification and if relevant unintended changes identified in preceding parts of the safety assessments impact human and animal health taking into account the anticipated intake of the genetically modified food and feed (see Section 2.7.1).

1.5 Allergenicity

The allergenicity assessment verifies whether the intended change and relevant unintended changes identified during the preceding steps of the risk assessment (molecular characterization and comparative assessment) impact the allergenic profile of the GM plant. In line with Codex Alimentarius, the allergenicity assessment in the EU focuses on the protein(s) newly expressed in the GM plant, on any new constituent other than the newly expressed protein(s) and on endogenous compositional endpoints showing altered levels.

Assessment of newly expressed protein(s) - in accordance with Codex Alimentarius (2009), since there is currently no single test that on its own provides sufficient information on the allergenic capacity of a protein, a weight-of-evidence approach considering all information available on the protein under assessment is followed. In 2009, Codex Alimentarius set the basis of information required for the assessment. Subsequently, EFSA endorsed Codex principles, provided further detailed justification of such requirements and a description of their practical implementation (EFSA, 2010b). More recently, EFSA elaborated additional clarifications and considerations (EFSA, 2011a; EFSA, 2017a). Summarising ten years of scientific developments in allergenicity assessment is beyond the purpose of the present document, however the essence of the debate centred around the following key considerations: currently it is not fully known what makes a protein an allergen, as it still has to be elucidated if it is a single feature shared among different proteins or not. To further complicate the topic, exposure to the allergen alone is not enough to provoke an allergic reaction, as specific environmental conditions are necessary for the allergic reaction to happen in predisposed individuals. Because of the intrinsic complexity of the mechanisms behind allergy diseases and because of the current lack of knowledge, the only currently available option for allergenicity assessment is to follow a weight-of-evidence approach, built upon sound and up-to-date scientific knowledge.

The assessment of newly expressed proteins can be divided into IgE-mediated and non-IgE-mediated adverse immune reactions. Similarly to the toxicity assessment, the battery of studies needed for the assessment of the newly expressed proteins is based on the available knowledge of the protein with respect to its source, function or activity and history of consumption, as well as on studies conducted *in silico*, *in vitro* and *in vivo* which are selected on a case-by-case basis. The relevance of these studies and the criteria followed to select, execute and interpret them are addressed in Section 2.6 of this document.

Assessment of new constituents other than the newly expressed protein(s) – the safety assessment of new constituents other than protein(s) is conducted on a case-by-case basis, depending on the information available on such new constituent regarding its allergenic potential or its capacity to stimulate the immune system.

Assessment of endogenous compositional endpoints – alterations of endogenous compositional endpoints beyond their degree of natural variability (see Section 2.3 on comparative assessment) are assessed for their possible allergenic relevance. Among these, the most relevant ones are the proteins since most known allergic reactions are triggered by proteins, their breakdown products, or peptides. Thus, the evaluation of potential increases in the endogenous allergenicity of a GM plant relatively to that of its comparator is requested when the plant receiving the introduced gene is considered allergenic, as in the case of soybean (OECD, 2012). For such assessment, relevant known endogenous allergens are included in

the array of compositional endpoints to enable the detection of possible chances in the allergenicity of the GM plant.

The conclusions of the allergenicity assessment indicate if the intended change of the genetic modification and if relevant unintended effects identified in preceding parts of the safety assessments impact the allergenic profile of the GM plant with respect to that of the non-GM comparator, taking into account the anticipated intake of the genetically modified food and feed (see Section 2.7.1).

1.6 Role of *in silico*, *in vitro* and *in vivo* studies in toxicity and allergenicity assessment of proteins

The assessment of newly expressed protein(s) is based on knowledge on the protein as well as on *in silico*, *in vitro* and *in vivo* studies. Whereas specific *in silico* and *in vitro* degradation studies are always required, other *in vitro* and *in vivo* studies are selected on a case-by-case basis depending on the knowledge on the protein and the outcome of the previous studies which are used to calibrate if and what additional studies are required. *In silico* studies normally compare the amino acid sequence of the newly expressed protein(s) with that of known toxins and allergens, including known celiac disease peptide sequences, to identify possible similarities in need of specific toxicological and/or allergenicity assessment. *In vitro* studies mainly focus on *in vitro* degradation to assess protein resistance to proteolytic enzymes such in the case of the pepsin resistance test, extensively used in allergenicity because of the causative relationship between resistance to pepsin-degradation and allergenicity first hypothesised (Astwood *et al.*, 1996) and then explored in larger details by Thomas and collaborators (2004). Pepsin-resistance is the study currently requested in the EU (Regulation (EU) 503/2013) and by Codex Alimentarius (2009).

Additional studies such as specific serum screening and cell-based or animal studies are requested on a case by case basis. In line with Codex Alimentarius (2009) and EFSA guidance document (2011a), specific serum screening is requested in cases where: i) the source of the transgene is considered allergenic; and ii) there are relevant hits identified from the bioinformatics analysis, regardless of the source allergenic potential. Nevertheless, it cannot be forgotten that serum-screening has limitations due to the intrinsic variability between allergic individuals and the difficulty of obtaining well-characterised sera. Furthermore, when no reliable information on the safety of the newly expressed protein(s) is available, an *in vivo* 28-day oral toxicity study in rodents is required.

1.7 Nutritional assessment

Purpose of the nutritional assessment is to establish whether the food and feed derived from the GM plant are nutritionally equivalent to those derived from their non-GM comparator(s), addressing therefore the possible nutritional impact of the intended change as well as the relevant unintended changes identified. The nutritional considerations provided below focus only on human nutrition because contrary to animals' diets, which are generally controlled and established on the basis of feed products *a priori* balanced to meet nutritional requirements, human diets are not fixed and food products are for the vast majority not balanced, at least not yet.

First, the biological role of the endpoint object of the genetic modification as well as that of the altered endogenous endpoints is explored and the existence of Dietary Reference Values (DRVs) for these endpoints is verified. DRVs indicate the amount of a compound/nutrient which must be consumed on a regular basis to maintain health in an otherwise healthy individual/population (EFSA, 2017b). In parallel,

the contribution of the crop to the total intake of such compound and the magnitude and direction (increase/decrease) of the changes observed in the comparative assessment are also assessed. In most cases, these pieces of information allow conclusive judgement on the nutritional impact. However, when DRVs are present and the contribution of the crop to the intake of the compound is relevant, a more thorough nutritional assessment is appropriate. This more in-depth nutritional assessment is performed replacing the nutrient contents reported in the EFSA composition database[1] for foods derived from the crop under assessment with the values reported for the GM crop in the application (values obtained from the comparative compositional analysis). A 100% replacement scenario – i.e. the consumption of foods is assumed to be 100% GM-food – is adopted to ensure a conservative approach, maximizing consumer protection. Attention must be paid not only to the portion of the population that may be below the recommended nutritional intakes specified by the DRVs, but also to the portion that may exceed tolerable upper intake level (UL) based on the changes identified for the compound/nutrient during the comparative assessment.

If it is not possible to conclude on the nutritional equivalence between foods derived from the GM crop and those derived from the non-GM crop, further nutritional studies are needed, such as comparative growth studies with young rapidly growing animal species (i.e. broiler chicks as animal model for non-ruminants; lambs for ruminants; or other rapidly growing species).

1.7.1 Exposure assessment

Dietary exposure estimates how much of a compound a population consumes. To estimate dietary exposure two types of data are required: occurrence or concentration data providing information on the amount of a particular compound in different food commodities, and consumption data informing on the intake of these food commodities. By combining these data and considering the body weight of the subjects, dietary exposure is estimated.

In GMO risk assessment the concentrations of the different endpoints are - in most of the cases - determined in raw primary commodities (e.g. in maize grains rather than in popcorn) as these represent the main point of entry of the GM material in the food industrial production chains. Nevertheless, most consumption data available in the EFSA Comprehensive European Food Consumption Database refer to processed commodities hampering the direct combination of consumption data with concentration data. Consequently, some assumptions are needed to estimate the concentration of the endpoints in processed GM commodities. First, recipes are used to identify and quantify the crop-derived ingredients used to elaborate the relevant processed foods; second, reverse yield factors are applied to convert ingredients amount into raw primary commodities amount, allowing a direct link between concentration data and consumption data. For newly expressed proteins, it is also assumed that no protein loss occurs during processing, except in few cases - such as oil - where the concentration of newly expressed proteins is negligible.

An alternative approach, occasionally used to estimate dietary exposure to newly expressed proteins, is the ratio of total protein content between processed food and raw primary commodity. The concentration of newly expressed proteins reported in raw primary commodities is then multiplied by this ratio to estimate their levels in the processed commodities. This approach considers that all proteins present in processed

[1] https://www.efsa.europa.eu/en/data/food-composition

commodities are derived from the GM crop and that no losses of newly expressed proteins occur during processing.

Human dietary exposure assessment considers the duration of the exposure based on the outcome of the hazard characterization - i.e. acute or chronic hazard. Chronic (long-term) exposure represents average daily exposure over years or the entire lifetime, while acute (short-term) exposure usually covers a period of 24 hours. The type assessment (chronic/acute) determines how the concentration and consumption datasets are used. The assessment focuses on the average population (average consumers), as well as on those consumers who may eat large amounts of one or more foods (high consumers). Likewise, attention should be paid to groups potentially more vulnerable due to their age (e.g. infants), consumption habits (e.g. vegetarians), life status (e.g. pregnancy), etc. Dietary exposure estimates can be presented for the whole population (represented by those participating in the survey) or for "consumers only", the second scenario being more appropriate to cover small population groups that consume regularly particular foods.

Dietary exposure is an essential component of the risk assessment of GM food, and it is used to conclude on the safety - absence of health concerns - of GM food constituents through two different scenarios. The first scenario follows the typical paradigm of risk assessment, where the dietary exposure to a compound is combined with the outcome of the hazard identification/characterization to conclude on the risks (risk characterization). However, appropriate toxicological studies are needed to identify and characterize the hazards of the relevant endpoints linked to short- and long-term consumption. The second scenario is based on what is known as comparative dietary exposure assessment, typically applied for the assessment of newly expressed proteins: when protein(s) similar or identical to the newly expressed protein(s) under evaluation are naturally present in one or more habitually consumed foods, the dietary exposure through the consumption of these foods is compared to that through the consumption the GM-derived foods, under an assumed 100% replacement scenario. Clearly, this scenario can be applied only when no evidence of safety concerns exists (history of safe consumption) for the proteins present in the conventional foods. This approach can also be used with other endpoints (e.g. NAA, lectins) when their levels are altered in the GM-crop, but there are evidences that humans are exposed to similar or higher concentrations through the consumption of different conventional foods.

Overall, the dietary exposure assessment follows a conservative approach where different assumptions are made in order to maximise consumer protection. This introduces uncertainties in the estimations. As an example, the absence of specific consumption data for GM-derived foods impels a 100% replacement scenario where all the consumption data of the relevant commodities are assumed to be GM food. The identification of all sources of uncertainties associated with dietary exposure estimations is pivotal to continuously improve the reliability of overall risk estimations and to inform on the strengths and limitations of the assessment.

1.8 Post Market Monitoring (PMM)

The need for a post-market monitoring (PMM) of the use of GM food and feed is established on a case-by-case basis. The pre-market risk assessment concludes on the safety of the GM crop and on the nutritional equivalence to its non-GM comparator, based on scientific evidence. This scientific evidence must be considered when addressing the potential need for a PMM plan. In particular, four aspects are fundamentally important to decide whether a PMM plan is needed: 1) possible changes in dietary exposure**,** 2) identification of new hazards, 3) new knowledge on already identified hazards (e.g. new

toxicological studies), and 4) availability of new information on tolerable upper intake levels (UL)/dietary reference values (DRV) for nutrients. Any of these aspects may affect the conclusions of the pre-market risk assessment which, in a worst-case scenario, can change from absence of risks to health/nutritional concerns. It is worthy underlining that changes in the dietary exposure to a particular GM component may easily occur either because of changes in consumption, or because of changes in the levels of that component in foods. The latter can result from changes in the composition of the crop (e.g. protein expression levels) or in food processing techniques affecting the levels of the components in processed products. Dietary exposure may also vary because new food commodities became available after the pre-market risk assessment (e.g. protein isolates) and/or changes in consumption habits over time.

Overall, post-market monitoring (PMM) is a tool complementing the pre-market safety assessment according to specific risk-based considerations, conceived to ensure the highest level of protections to consumers. When considered necessary, after the authorisation, the PMM must be implemented and its outcome must be reported and monitored regularly.

1.9 Concluding Remarks

In conclusion, the EU has developed a rigorous and detailed framework for the risk assessment of GM plants, in line with Codex Alimentarius principles and internationally agreed risk assessment procedures. The EFSA continues to monitor and to contribute to the scientific advances of risk assessment and updates its guidance documents accordingly, strengthening cooperation with other national organizations experienced in risk assessment of GM food and feed, and supporting EU risk managers in ensuring a forward looking legislation with respect to scientific requirements, approaches and methodologies.

2. Additional considerations

2.1 Stacked events

A new promising frontier of biotechnological development is the production of genetically modified plants containing more than one transformation events, the so called 'stacked events' or, more simply 'stacks'. These can be produced either by conventional crossing of parental GM plant lines containing one or more events without the involvement of additional genetic transformations, or with other biotech methodologies such as multiple gene cassettes, co-transformation, re-transformation. Stacks are becoming increasingly important and are at the centre of attention of biotech industry as they allow combining multiple specific traits into one plant. At the time when Codex Alimentarius principles for GM plant risk assessment were first established (Codex Alimentarius, 2003) and then revised (Codex Alimentarius, 2009), even the most forward-thinking scientific mind could not imagine that just few years later stacks would have been a concrete reality.

In the EU, GM plants containing multiple events must be risk assessed like GM plants containing single events. The risk assessment requirements are similar, even though a novel layer of complexity linked to the possible occurrence of interactions among the transformation events combined in the same genome is considered.

2.1.1 Stacked events: specific considerations for their risk assessment

The risk assessment of a GM plant containing multiple events has triggered - and continues to spark - intense discussion in the relevant technical *fora* because it is still being disputed whether the information necessary for the risk assessment of stacks can or cannot be extrapolated from singles already risk assessed. The EU has settled the issue requiring the independent risk assessment of all the events (i.e. GM plants containing single events) combined in the stack before authorising a stack on the EU market. The resolution of this diatribe at the global level is out of the remit of this document. Below we provide a short overview of the challenges that stacks pose to GM plant risk assessment, with the objective of triggering the reader's interest. Certainly, we do not have the ambition of addressing all the complexity intrinsic to the topic.

The overall principle followed for risk assessing stacks in the EU is the same as the one used to assess a GM plant containing one event (single) illustrated earlier in this document. The primary concern when assessing a stack is to establish that the combination of events is stable and does not result in interactions raising safety concerns as compared to singles. At the molecular characterization level, it is necessary to determine first, whether the inserts' structures are retained and stable when combined in the receiving genome; second, whether the expression of the introduced genes and their products is stable or not. The structure and stability of the inserts in the stack can be verified by direct comparison with the structure and stability of the inserts in the corresponding single events across multiple generations. Instead establishing stability of expression is more challenging: in principle this can be verified comparing the levels of the proteins newly expressed in the stack with their levels in the respective singles. If expression levels observed in the stack are linear function of what was observed in the singles, stability of expression can be safely assumed. However, this is not as simple as it may appear. It is well known that the genomic background of a plant as well as environmental growing conditions, influence the expression levels of the plant genes, including those of newly introduced genes, as in the case of GM plant varieties. Consequently, conclusions on the

stability of expression levels through comparisons between singles and stacks across different genetic backgrounds, possibly grown under different environmental conditions, need careful attention and have limited extrapolation relevance, since other relevant influencing factors are unavoidably confused.

At the comparative assessment level, the first challenge posed by stacks is the definition of the criteria to select appropriate non-GM comparators. For stacks produced by conventional breeding the first choice is the non-GM comparator, similarly to what is done for singles and described in Section 2.3. If a non-GM comparator is not available, any set of GM plants, previously risk assessed, that include - between them - all the events stacked in the GM plant under assessment and no others can be used for the comparative assessment. This set of GM plants may include either parental GM lines, or GM plants containing the single events. Such flexibility in the selection of the non-GM comparator testing material(s) allows to circumvent the possible lack of a non-GM comparator, while ensuring appropriate analysis of potential interactions. If a stack is obtained by re-transformation of an existing GM line, similarly to what is done for singles, the first choice is the non-GM comparator. If a non-GM comparator does not exist, either the negative segregant or the recipient GM plant, which must have been risk assessed previously, can be used for the comparative assessment. If a stack is produced by inserting, in a non-GM line, a single cassette with multiple genes or sequences which will modify gene expression, it is expected that the insert will occur at a single locus. Therefore, independent segregation of the elements of this cassette is not likely. Such a stack can be treated as a single also with respect to the criteria followed to select a suitable non-GM comparator.

All the other aspects of the comparative field trials - design, statistical analysis and interpretation - described in Sections 2.3.1, 2.3.2 and 2.3.3 are directly applicable also to stacks.

The possible occurrence of interactions among the transformation events combined in the same genome is at the root of the school of thinking requiring the assessment of stacks. Such interactions could be either synergistic or antagonistic, depending on the nature of the events involved, and can occur at multiple levels. At the level of the stack composition, where the simultaneous presence of two or more newly expressed gene products, e.g. proteins, or of two or more altered endogenous endpoints result in an interaction that may raise safety concerns in human or animals, when consumed as food or feed. At the cellular level in the stack, where the new gene product, e.g. the newly expressed protein, act on the same pathway or on interlinked metabolic pathways, possibly affecting the level of cell components or even give rise to new components. At the genomic level, where genes may interact and suppress the expression of a certain gene at the transcriptional or translational level, resulting in gene silencing.

Detecting all possible interactions among events within a stack remains challenging: interactions can be detected if the tools/methodologies used for the risk assessment have been designed to capture them. Although this may seem a trivial logic, it necessary implies a hypothesis driven frame. However, the extent to which interactions can be hypothesised strictly depends on the degree of knowledge of the inserted traits, the objectives of the transformation events combined in a single hosting genome and their metabolic connections. Our limited understanding of all these factors intrinsically also limits our capacity to detect interactions: the more our understanding of such biologically systems advance, the larger the likelihood to detect interactions and reduce uncertainty.

2.2 GM animals

No application related to GM animals was ever submitted in the EU under Regulation (EC) No. 1829/2003 or Directive 2001/18/EC. However, scientific developments suggest that submissions of applications for the

authorization of GM animals may happen for different animal species in the close future. As a proactive measure, the EFSA developed comprehensive risk assessment guidelines to be used by applicants and risk assessment bodies to evaluate the safety of food and feed products derived from GM animals. The animals taken into consideration were husbandry animals, fish, crustaceans and molluscs. Insects and other invertebrates were not considered, except for honeybees. The variety of animal species covered allowed to develop only high-level guidance, whilst specific requirements for the safety assessment of food and feed derived from particular GM animal species, or even GM animal lines are to be determined on a case-by-case basis.

At present, in the EU there are two guidance documents for the risk assessment of GM animals. One, focusing on the environmental risk assessment of GM animals, was published in May 2013 (EFSA, 2013). Since environmental issues are out of the scope of this document, we do not elaborate further and refer readers to the original document. The other, published in January 2012 (EFSA, 2012), outlines the data required and the methodology to be followed for the risk assessment of food and feed derived from animals genetically modified to express new heritable traits. Food and feed derived from animals with introduced non-heritable traits are not covered. Below we summarise the highlights of such guidance document and we refer the reader to the original text for further details (EFSA, 2012).

As for GM plants, the first step of the food and feed safety assessment of GM animals and derived food and feed products is a comprehensive molecular characterisation, assessing the structure and expression of the insert(s) and the stability of the expressed trait(s), followed by the comparative safety assessment of the phenotypic and compositional characteristics of the GM animal. The implicit assumption allowing the application of the comparative approach is that food and feed from conventionally bred animals have a history of safe use and can serve as a baseline for the risk assessment of food and feed derived from GM animals. Such an assumption is neither unreasonable nor surprising: traditionally, the health status of a food and feed producing animal has always been considered an important indicator of the safety of the derived food and feed. The practice of allowing only animals with known and acceptable health and welfare status to enter the human and animal food/feed supply is an essential step for ensuring safe food and feed. Therefore, the most important component in the risk assessment of GM animals is an extensive comparative analysis of their phenotypic characteristics, including health and physiological parameters.

Alterations in health and welfare of a GM animal can be identified through clinical observations and examinations to detect deviations from normal health and behaviour. Alterations in the phenotype can be identified through a comparative analysis of visual characteristics, feed intake, growth performance, developmental characteristics, digestive and reproductive capacity, disease resistance and health (including immunological) parameters. Alterations in the composition of specific parts of a GM animal or in products such as milk and eggs, compared with the same parts and products of its non-GM comparator, can be identified by measurements of constituents representing components of important metabolic pathways in the organism, taking into account natural variation. The components should include key nutrients (i.e. proximates, macronutrients and micronutrients) and bioactive compounds, when relevant. The outcome of the comparative analysis further structures the subsequent assessment procedure, which may include specific safety and nutritional testing.

The subsequent steps of the risk assessment are logically equivalent to those for GM plants: the toxicological and allergenicity assessments, which address the impact on human and animal health of the relevant changes observed in the GM animal and/or derived food and feed during the molecular

characterisation and/or the comparative assessment; the nutritional assessment which evaluates whether food and feed derived from a GM animal are nutritionally equivalent to those derived from traditionally-bred (non-GM) animals. Finally, post market monitoring plans are recommended whenever changes in levels of relevant constituents and/or in the overall dietary intake patterns of GM animal food and/or feed affecting dietary exposure estimates can be expected.

3. Future Developments

The ultimate goal of safety assessment outline by Codex Alimentarius (2003-2009) is first to define the likelihood for an adverse event to happen upon consumption of a novel food derived from genetically modified organisms and second, to provide acceptance/rejection criteria for final decision. Considering that the purpose of a genetic modification - regardless of the technique used may this be transgenesis, mutagenesis, conventional breeding, new breeding techniques or any other type, assisted or not by humans - is to introduce changes into existing plants (animals) varieties, the focus of Codex Alimentarius guidelines (2003-2009) is precisely their safety assessment. As explained in this document, changes might involve either the alteration of the levels of endogenous endpoints or the de novo appearance of new compounds. These changes have been classified into intended/ unintended and predictable/unpredictable, depending on our level of understanding (Codex Alimentarius, 2003). Independently of their classification, the comparative approach philosophy imposes the assessment of all change as any of them may pose a hazard and different approaches, are proposed and used to identify changes introduced by genetic modifications. This does not help the consistency and harmonization needed to facilitate trades in an increasingly global market. Experts continue to advocate different approaches, targeted or untargeted, depending on their level of risk acceptance. Future developments should focus on defining the best comparative approach possible for the assessment of changes (unintended and intended) introduced by genetic modifications, and better exploiting the potentiality of post market monitoring.

3.1 Assessment of changes introduced unintentionally

The challenges linked to the assessment of changes introduced unintentionally are clearly described by Codex Alimentarius (2009), which acknowledges the practical impossibility that one individual test can detect all of them. As a pragmatic solution, Codex Alimentarius recommends analysing 'key components' that, in their totality, allow estimating the likelihood of adverse effects on humans. To overcome limitations of current approaches for the safety assessment of changes introduced unintentionally, Fernandez and Paoletti (2018) have previously defined three outstanding needs: i) Building up consensus on the history-of-safety-use concept; ii) harmonizing criteria to select appropriate conventional counterparts; and iii) improving endpoint selection on the basis of their representativeness, relevance and variability. The development of alternatives to the current paradigm established by the Codex Alimentarius can only start once international consensus on these three main aspects is achieved, as these impact directly how to frame the comparative approach with respect to the boundaries set by conventional breeding. Let's elaborate this further. The overall strategy of the risk assessment of GMOs is performed in a comparative manner and confined within the limits of what is currently known, improperly named as "natural variability". The estimation of natural variability has been the obsession of risk assessment bodies around the world and an aspect of controversial discussions. This is because understanding endpoints variability patterns is at the same time a necessity for decision making during the process of risk assessment, and a struggle because of the large number of varieties present on the market at any given time, and the indefinite number of conditions under which such varieties can be grown. To this end, databases can serve as reference points to identify and assess the relevance of changes observed in the context of known 'natural variability'. The development of more robust databases against which new GMOs can be compared is of crucial importance.

The collection of databases currently available covers several domains, e.g. compositional analysis, known allergenic and toxic proteins, horizontal gene transfer. Within each of these domains several databases are

available, evidencing the lack of consensus of what information a database should include, thus creating confusion and misunderstanding across risk assessors around the world. The underlying reason for the lack of international acceptance criteria for database composition is the unavailability of commonly accepted inclusion criteria. The construction of a database recording the compositional profile of commercial crop varieties on a crop-by-crop basis, ensuring comparability of information and minimization of confounding effects (i.e. constructed with clearly defined and stringent entrance criteria), would be an effective step to properly frame natural variability estimates in the risk assessment process. Such a database would need regular maintenance as crop variability estimates should capture the speed of development of new varieties with unique degrees and patterns of variability. Because credibility and transparency are key values of a reliable risk assessment, the development of inclusion criteria requires cooperation between the international scientific community and its stakeholders. A schematic logic flow for the establishment of an international database has been proposed recently (Figure 4.1). Although developed for crops compositional analysis, the principles described are applicable to any other domain requiring high-quality database design, maintenance and curation following transparent and accepted inclusion criteria.

-In light of this discussion, if we consider again that the purpose of the comparison defined by Codex Alimentarius is to establish that substances either nutritionally important, or affecting food safety have not been altered in a manner that would have an adverse impact on human health, we cannot ignore the challenge that future products going beyond the limits established by conventional breeding will impose. These are cases for which readily available comparators do not exist. A concrete example? The risk assessment of crops that have undergone substantial intentional changes in their compositional profile, that is, in one or more endogenous endpoints normally used as benchmarks to assess overall substantial equivalence to a comparator, is already raising conceptual difficulties.

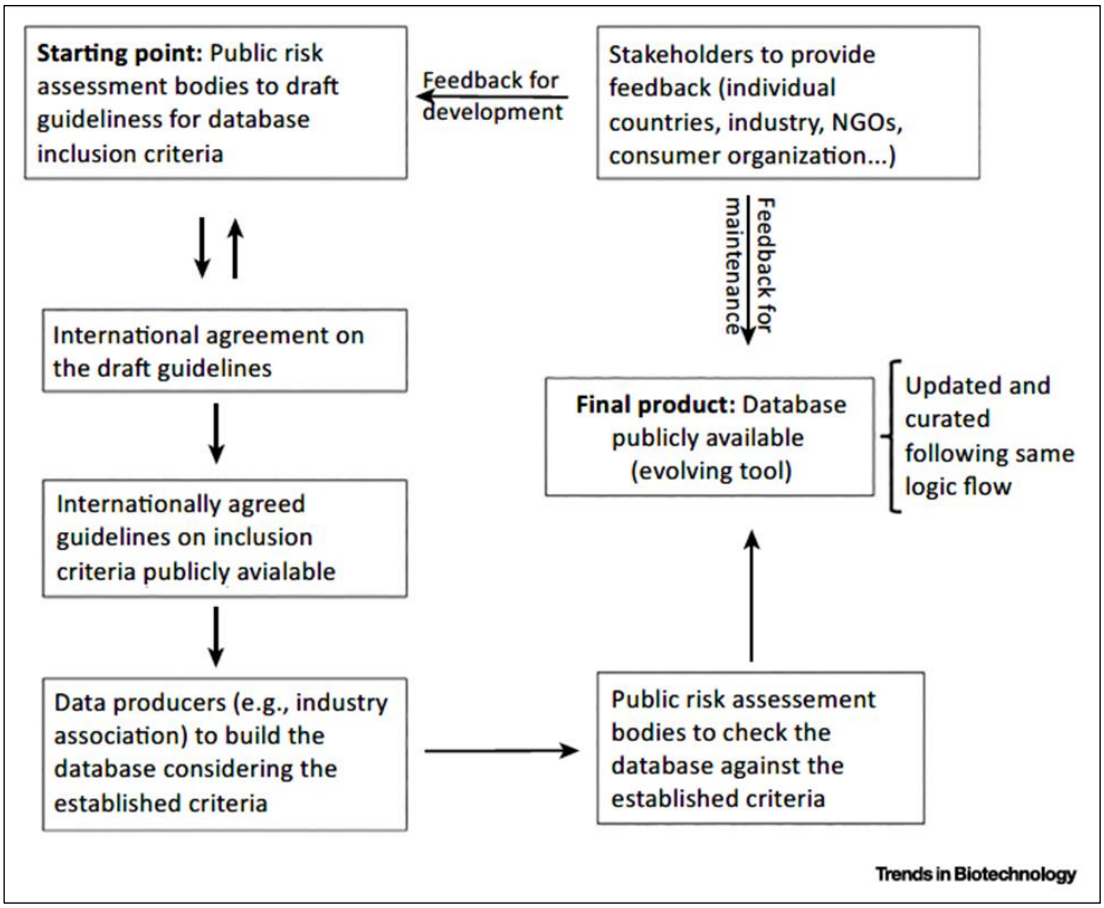


Figure 4.1: Schematic logic flow for the establishment of an international database. From Fernandez and Paoletti 2018.

3.2 Assessment of changes introduced intentionally

Intended changes normally alter compositional endpoints or introduce new constituents *de novo* which at the current stage are mainly newly expressed proteins, even though the alteration of endogenous compositional endpoints is also possible and is already triggering industrial/market interests. The latter raises difficulties to the application of Codex Alimentarius principle in its traditional delineation. The comparative approach is confronted with the struggle of identifying suitable comparators in a situation where, by definition, a customary comparator does not exist. Also, under these circumstances, understanding compositional endpoint variability patterns could prove itself an asset to explore alternatives. Codex Alimentarius indirectly foresaw the problem when established that "*Although it is preferable to use a scientifically determined upper level of intake of a specific nutrient or related substance, when no such value has been determined, consideration may be given to an established history of safe use for nutrients or related substances that are consumed in the diet if the expected or foreseeable exposure would be consistent with those historical safe levels*". However, it remains unclear how first a history of safe use for a specific nutrient or substance consumed in the diet can be established and second, how this can be translated into historical quantifiable safe levels (Fernandez and Paoletti, 2018).

The safety assessment of newly expressed proteins is an important element in the overall GMO risk assessment since adverse effects from exposure to proteins can be life-threatening as in the case of anaphylactic reactions. The assessment of proteins safety relies on information of different nature which, all together, provide the cumulative body of evidence necessary to estimate risk (the so-called weight of evidence approach). The larger the body of evidence, the more reliable the predication and the smaller the associated uncertainty. Current international guidelines for the safety assessment of proteins date back to 2003 (Codex Alimentarius, 2003) and borrow the classical principles and methodologies developed to assess 'chemicals' Since 2003, the limits of applying an approach developed for simple chemicals to complex biological molecules like proteins have become increasingly apparent. Furthermore, food safety assessments should always rely on methods/evidences at the cutting edge of science, set in guidelines under continuous update. For example, in 2017 the EFSA reviewed allergenicity assessment and comprehensively addressed novel key aspects linked to non-IgE-mediated immune adverse reactions to foods (EFSA, 2017a). Scientific advances and the evolution of risk assessment tools should drive the re-thinking of protein safety assessment strategy upon which a global consensus should be achieved. Some relevant elements in such respect are the following:

- Development of more robust bioinformatic approaches: current approaches searching for similarities to known toxins and allergens are either not internationally recognized, or date back to 2001 (FAO/WHO, 2001). New approaches proposed in the scientific literature, including the development of predictive modelling tools as recently done in the case of celiac disease (EFSA, 2017a) should be discussed with the common objective of complementing or replacing old methodologies.

- Development of more predictive *in vitro* studies. Recently a refined *in vitro* protein digestion protocol better reflecting the fate of the protein in the gastrointestinal tract has been developed (EFSA GMO Panel, 2017).At the same time, the usefulness of the classical pepsin resistance test for the risk assessment of proteins embedded in Codex Alimentarius is being questioned by new scientific developments in the area (Bøgh and Madsen, 2016; EFSA, 2017a; Fernandez *et al.*, 2019, 2020; Verhoecks *et al.*, 2019, 2020). Developments as these are of particular relevance to ensure the proactive incorporation of up-to-date technology in the risk assessment process. In addition, the possibility of

using cell-based studies to integrate and possible replace the more traditional animal studies is a need deserving special attention.

- Development of more targeted *in vivo* approaches, of crucial relevance in the view of the 3Rs philosophy (Directive 2010/63/EU) when dealing with animal models but requiring long-term activity and articulated discussion. The possible development of novel *in vivo* approaches implies a deep revision of the use of animal models for protein safety, starting from a common understanding and recognition of what parameters are most informative and should be considered under different scenarios. Finally, it should be discussed how to make use of human data for the safety assessment which are becoming readily available (EFSA, 2019). Finally, future strategies should also consider the impacts on proteins of food processing and matrices, role of different routes of exposure and assumptions and uncertainties derived from such assessment.

3.3 Relevance of post market monitoring for both intended and unintended changes

Development of a solid strategy to link the outcome of the safety assessment, including its uncertainties and limitations, with the triggering for post market monitoring (PMM) plans. Currently, PMM is requested when the verification of conditions of use of GM products is needed, or when monitoring of their consumption is deemed necessary. However, properly designed PMM plans could provide valuable additional information on other aspects, including GMO traceability after-market release, and more accurate consumption data down to the branded/product level. PMM could also provide effective iterative feedback to better inform and continuously improve the pre-market risk assessment strategy on the basis of relevant post-marketing observations. Finally, newly developed monitoring systems can become a valuable tool to sample an ever-changing reality and examine the extent to which pre-market assumptions embedded into risk assessment methodologies remain valid/acceptable over time or not, providing a system to censor safety standard reliability (Fernandez and Paoletti, 2020).

4. Reference

Astwood JD, Leach JN, Fuchs RL, 1996. Stability of food allergens to digestion *in vitro*. Nature Biotechnology, 14(10):1269-73. DOI: 10.1038/nbt1096-1269

Bøgh KL and Madsen CB, 2016. Food Allergens: Is There a Correlation between Stability to Digestion and Allergenicity? Critical Reviews in Food Science and Nutrition, 56(9):1545-1567. DOI: 10.1080/10408398.2013.779569

Codex Alimentarius Commission, 2003. Principles for the Risk Assessment of foods derived from modern biotechnology. CAC/GL 44-2003. Available online: http://www.fao.org/input/download/standards/10007/CXG_044e.pdf

Codex Alimentarius Commission, 2009. Foods Derived from Modern Biotechnology (2nd edition). World Health Organization (WHO) & Food and Agriculture Organization of the United Nations (FAO). Available online: http://www.fao.org/3/a-a1554e.pdf

Decision 69/414/EEC of the Council of 13 November 1969 on setting up a Standing Committee for Foodstuffs. Official Journal L291, 19/11/1969, pp9-10. Available online: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:31969D0414

Directive 2001/18/EC of the European Parliament and of the Council of 12 March 2001 on the deliberate release into the environment of genetically modified organisms and repealing Council Directive 90/220/EEC - Commission Declaration. Official Journal L 106, 17/04/2001, pp1-39. Available online: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32001L0018

Directive 2010/63/EU of the European Parliament and of the Council of 22 September 2010 on the protection of animals used for scientific purposes. Official Journal L 276, 20/10/2010, pp33-79. Available online: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32010L0063

EFSA GMO Panel Working Group on Animal Feeding Trials, 2008. Safety and nutritional assessment of GM plants and derived food and feed: The role of animal feeding trials. Food and Chemical Toxicology, 46: S2–S70. DOI: 10.1016/j.fct.2008.02.008

EFSA GMO Panel (EFSA Panel on Genetically Modified Organisms), 2010a. Statistical considerations for the safety evaluation of GMOs. EFSA Journal 2010, 8(1):1250, 59pp. DOI: 10.2903/j.efsa.2010.1250

EFSA GMO Panel (EFSA Panel on Genetically Modified Organisms), 2010b. Scientific Opinion on the assessment of allergenicity of GM plants and microorganisms and derived food and feed. EFSA Journal 2010; 8(7):1700, 168 pp. DOI: 10.2903/j.efsa.2010.1700

EFSA GMO Panel (EFSA Panel on Genetically Modified Organisms), 2011a. Scientific Opinion on Guidance for risk assessment of food and feed from genetically modified plants. EFSA Journal 2011, 9(5):2150, 37pp. DOI: 10.2903/j.efsa.2011.2150

EFSA GMO Panel (EFSA Panel on Genetically Modified Organisms), 2011b. Guidance document on Selection of Comparators for the Risk Assessment of GM Plants. EFSA Journal 2011, 9(5):2149, 20pp. DOI: 10.2903/j.efsa.2011.2149

EFSA Scientific Committee, 2011c. Guidance on conducting repeated-dose 90-day oral toxicity study in rodents on whole food/feed. EFSA Journal 2011, 9(12):2438, 21 pp. DOI: 10.2903/j.efsa.2011.2438

EFSA Panels on GMO (Genetically Modified Organisms) and AHAW (Animal Health and Welfare), 2012. Scientific Opinion on the Guidance on the risk assessment of food and feed from genetically modified animals and animal health and welfare aspects. EFSA Journal 2012, 10(1):2501, 43pp. DOI: 10.2903/j.efsa.2012.2501

EFSA GMO Panel (EFSA Panel on Genetically Modified Organisms), 2013. Guidance on the environmental risk assessment of genetically modified animals. EFSA Journal 2013, 11(5):3200, 190pp. DOI: 10.2903/j.efsa.2013.3200

EFSA (European Food Safety Authority), 2014. Explanatory statement for the applicability of the Guidance of the EFSA Scientific Committee on conducting repeated-dose 90-day oral toxicity study in rodents on whole food/feed for GMO risk assessment. EFSA Journal 2014, 12(10):3871, 25pp. DOI: 10.2903/j.efsa.2014.3871

EFSA GMO Panel (EFSA Panel on Genetically Modified Organisms), 2015. Guidance on the agronomic and phenotypic characterisation of genetically modified plants. EFSA Journal 2015, 13(6):4128, 44pp. DOI:10.2903/j.efsa.2015.4128

EFSA GMO Panel (EFSA Panel on Genetically Modified Organisms), 2017a. Naegeli H, Birch AN, Casacuberta J, De Schrijver A, Gralak MA, Guerche P, Jones H, Manachini B, Messean A, Nielsen EE, Nogue F, Robaglia C, Rostoks N, Sweet J, Tebbe C, Visioli F, Wal J-M, Eigenmann P, Epstein M, Hoffmann-Sommergruber K, Koning F, Lovik M, Mills C, Moreno FJ, van Loveren H, Selb R and Fernandez Dumont A. Guidance on allergenicity assessment of genetically modified plants. EFSA Journal 2017, 15(5):4862, 49pp. DOI: 10.2903/j.efsa.2017.4862

EFSA (European Food Safety Authority), 2017b. Dietary reference values for nutrients: Summary report. EFSA supporting publication, 2017:e15121, 92pp. DOI: 10.2903/sp.efsa.2017.e15121

EFSA GMO Panel (EFSA Panel on Genetically Modified Organisms), 2018a. Casacuberta J, Nogue F, Naegeli H, Birch AN, De Schrijver A, Gralak MA, Guerche P, Manachini B,Messean A, Nielsen EE, Robaglia C, Rostoks N, Sweet J, Tebbe C, Visioli F, Wal J-M, Moxon S, Schneeberger K, Federici S, Ramon M, Papadopoulou N and Jones H, 2018. Scientific Opinion on the technical Note on the quality of DNA sequencing for the molecular characterisation of genetically modified plants. EFSA Journal 2018, 16(7):5345, 11 pp. DOI: 10.2903/j.efsa.2018.5345

EFSA (European Food Safety Authority), 2018b. Paraskevopoulos K, Ramon M, Dalmay T, du Jardin P, Casacuberta J, Guerche P, Jones H, Nogué F, Robaglia C, Rostoks N. Explanatory note on the determination of newly expressed protein levels in the context of genetically modified plant applications for EU market authorisation. EFSA supporting publication 2018:EN-1466, 13pp. DOI: 10.2903/sp.efsa.2018.EN-1466

EFSA (European Food Safety Authority), 2019. Parenti MD, Santoro A, Del Rio A, Franceschi C. Literature review in support of adjuvanticity/immunogenicity assessment of proteins. EFSA supporting publication 2019:EN-1551, 68pp. DOI:10.2903/sp.efsa.2019.EN-1551

FAO/WHO (Food and Agriculture Organization of the United Nations/World Health Organization), 2001. Evaluation of Allergenicity of Genetically Modified Foods. Report of a Joint FAO/WHO Expert Consultation on Allergenicity of Foods Derived from Biotechnology. Food and Agricultural Organization/World Health Organization. Rome, Italy. Available online: http://www.fao.org/fileadmin/templates/agns/pdf/topics/ec_jan2001.pdf

Fernandez A and Paoletti C, 2018. Unintended effects in genetically modified food/feed safety: a way forward. Trends in Biotechnology, 36(9):872-875. DOI:10.1016/j.tibtech.2018.03.005

Fernandez A, Mills ENC, Koning F, Moreno FJ, 2019. Safety assessment of immune-mediated adverse reactions to novel food proteins. Trends in Biotechnology, 37(8):796-800. DOI: 10.1016/j.tibtech.2019.03.010

Fernandez A, Mills ENC, Koning F, Moreno FJ, 2020. Allergenicity Assessment of Novel Food Proteins: What Should Be Improved? Trends in Biotechnology, in press: 1947. https://doi.org/10.1016/j.tibtech.2020.05.011

Fernandez A and Paoletti C, 2020. What is unsafe food? Change of perspective. Nature Food, under review.

OECD (Organisation for Economic Cooperation and Development), 1993. Safety evaluation of foods derived by modern biotechnology. Concepts and principles. Available online: http://www.oecd.org/env/ehs/biotrack/41036698.pdf

OECD (Organisation for Economic Co-operation and Development), 1998. Repeated Dose 90-Day Oral Toxicity Study in Rodents. Test Guideline No. 408, Guidelines for the Testing of Chemicals, OECD, Paris. Available online: https://www.oecd-ilibrary.org/environment/test-no-408-repeated-dose-90-day-oral-toxicity-study-in-rodents_9789264070707-en

OECD (Organisation for Economic Cooperation and Development), 2012. Revised consensus document on compositional considerations for new varieties of soybean [*Glycine max* (L.) Merr.]: key food and feed nutrients, antinutrients, toxicants and allergens. Series on the Safety of Novel Foods and Feeds No. 25. Available online: http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono(2012)24&doclanguage=en

Regulation (EC) No 1829/2003 of the European Parliament and of the Council of 22 September 2003 on genetically modified food and feed. Official Journal L 268, 18/10/2003, pp1-23. Available online: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32003R1829

Regulation (EU) No 503/2013 of the Commission of 3 April 2013 on applications for authorisation of genetically modified food and feed in accordance with Regulation (EC) No 1829/2003 of the European Parliament and of the Council and amending Commission Regulations (EC) No 641/2004 and (EC) No 1981/2006. Official Journal L157, 08/06/2013, pp1-48. Available online: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32013R0503

Thomas K, Aalbers M, Bannon GA, Bartels M, Dearman RJ, Esdaile DJ, Fu TJ, Glatt CM, Hadfield N, Hatzos C, Hefle SL, Heylings JR, Goodman RE, Henry B, Herouet C, Holsapple M, Ladics GS, Landry TD, MacIntosh SC, Rice EA, Privalle LS, Steiner HY, Teshima R, Van Ree R, Woolhiser M, Zawodny J, 2004. A multi-laboratory evaluation of a common in vitro pepsin digestion assay protocol used in assessing the safety of novel proteins. Regulatory Toxicology and Pharmacology, 39(2):87-98. DOI:10.1016/j.yrtph.2003.11.003

Van der Voet H, Perry J, Amzal B, Paoletti C, 2011. A statistical assessment of differences and equivalences between genetically modified and reference varieties. BMC Biotechnology, 11-15. DOI: 10.1186/1472-6750-11-15

Verhoecks K, Bøgh KL, Dupont D, Egger L, Gadermaier G, Larré C, Mackie A, Menard O, Adel-Patient K, Picariello G, Portmann R, Smit J, Turner P, Untersmayr E, Epstein MM, 2019. The relevance of a digestibility evaluation in the allergenicity risk assessment of novel proteins. Opinion of a joint initiative of COST action ImpARAS and COST action INFOGEST. Food and Chemical Toxicology, 129:405-423. DOI:10.1016/j.fct.2019.04.052

Verhoeckx K, Bøgh KL, Constable A, Epstein MM, Hoffmann Sommergruber K, Holzhauser T, Houben G, Kuehn A, Roggen E, O'Mahony L, Remington B, Crevel R, 2020. COST Action 'ImpARAS': What Have We Learnt to Improve Food Allergy Risk Assessment. A Summary of a 4 Year Networking Consortium. Clinical and translational allergy. May 18;10:13. doi: 10.1186/s13601-020-00318-x.