

## 5. Characterization of the genetic modification(s)

### Molecular analysis of the recombinant-DNA insert

Characterization of a recombinant-DNA plant at the molecular level is performed to provide information about the composition and integrity of the inserted DNA, the number and genomic location of the single or multiple sites of insertion, and the level of expression of the introduced protein(s) over time and in different tissues and environments.

As explained in the Section 4, the process of recombinant-DNA plant production may result in a transformed plant that contains a single insert or multiple inserts present in one or several locations in the host plant genome.

Regulatory authorities examine the information on the integrity and copy number of the inserted DNA in recombinant-DNA plants. Biotechnologists usually seek to minimize the copy number and size of the inserted DNA in recombinant-DNA plants to ease the regulatory process by producing fewer genetic changes that require assessment. However, recombinant-DNA plants containing multiple copies of the inserted DNA are not necessarily less “safe” than comparable plants containing only a single copy<sup>10</sup>.

**CODEX GUIDELINE PARAGRAPH 30.** In order to provide clear understanding of the impact on the composition and safety of foods derived from recombinant-DNA plants, a comprehensive molecular and biochemical characterization of the genetic modification should be carried out.

**CODEX GUIDELINE PARAGRAPH 31.** Information should be provided on the DNA insertions into the plant genome; this should include:

- A) the characterization and description of the inserted genetic materials;
- B) the number of insertion sites;
- C) the organisation of the inserted genetic material at each insertion site including copy number and sequence data of the inserted material and of the surrounding region, sufficient to identify any substances expressed as a consequence of the inserted material, or, where more appropriate, other information such as analysis of transcripts or expression products to identify any new substances that may be present in the food; and
- D) identification of any open reading frames within the inserted DNA or created by the insertions with contiguous plant genomic DNA including those that could result in fusion proteins.

Knowledge of the genomic locations in which the transgene(s) have been inserted in the plant genome is necessary to assess if existing genes or regulatory sequences have been affected by the insertion, which may result in altered gene expression patterns and, hence, plant phenotype. To assess whether new protein molecules could be produced from the integration of inserted DNA, DNA sequence-based bioinformatics analyses are used to determine the presence of open reading frames (ORFs) in and around the DNA insert.

An open reading frame is a part of a gene that is transcribed to produce RNA. Bioinformatics analysis is usually focused on both the newly introduced ORFs present in the DNA insert itself and the potential presence or creation of new ORFs produced from the random insertion of DNA into existing ORFs in the plant genome.

A detailed molecular characterization of the recombinant-DNA may be able to address issues related to possible positional effects that lead to variable gene expression, multiple character changes (pleiotropic

<sup>10</sup> One example of an “event” containing a high transgene copy number is in a line of canola (*Brassica napus*; event 23-198, 23-18-17) approved by the Canadian Government, which was developed by introducing a thioesterase encoding gene from the California bay tree (*Laurus nobilis*) to increase levels of lauric acid (12:0) and, to a lesser extent, myristic acid (14:0). The original transformation event 23 was estimated to contain 15 copies of the gene, at five independent genetic loci, as shown by Southern blot and segregation analyses.

effects) arising from the DNA insertion, or gene silencing resulting from overexpression of the inserted DNA. However, in the absence of other empirical data, such molecular analyses are unlikely to predict unforeseen effects on the concentrations of key nutrients, antinutrients or endogenous toxins. Thus, additional compositional analyses are preformed.

Where the result of the modification is the expression of a novel protein, the plant material is characterized with respect to the biochemical composition and functionality of the new gene product(s). Several methods are used to verify and measure the expression of the introduced traits in a recombinant-DNA plant. For novel protein-derived traits, serological techniques are frequently used. Such techniques (e.g. Western immunoblotting or enzyme-linked immunosorbent assay [ELISA]) are used to identify the presence of the transgene product and to quantify its level in the sampled material. If the newly inserted trait is one that does not result in the expression of a new or modified protein<sup>11</sup> but, for instance, results in antisense RNA sequences, other techniques (e.g. Northern blotting) are used to measure transcript production.

In addition to the direct biochemical characterization of the inserted trait, regulatory authorities usually assess studies of the recombinant-DNA plant grown under various conditions. Such studies can show that the intended trait is expressed at the desired life stage of the plant cultivar, and that expression is as expected and is stable over environments and plant generations.

The overall concentration of novel proteins expressed in recombinant-DNA plant tissues is low, often less than 0.1 percent on a dry weight basis. Biosafety studies, such as acute toxicity testing (chapter 6), that require relatively large amounts of material are often not feasible using the protein purified from plant tissue. Instead, these studies normally make use of protein purified from bacterial expression systems. In such cases, it is necessary to demonstrate the functional equivalence (in terms of physiochemical properties and biological activities) of the proteins purified from the two sources<sup>12</sup>.

Refer to the Codex Guideline paragraph 33, for each introduced trait, the expected expression pattern and stability of inheritance is usually demonstrated using data from field trials collected over several seasons and geographical locations. The genomic stability of the insert is usually shown by Southern blotting of DNA extracted from plant material sampled over several seasons and locations. Similarly, stable expression of the inserted DNA is shown by quantification of the corresponding protein or protein activity.

#### **CODEx GUIDELINE PARAGRAPH 32.**

Information should be provided on any expressed substances in the recombinant-DNA plant; this should include:


- A) the gene product(s) (e.g. a protein or an untranslated RNA);
- B) the gene product(s)' function;
- C) the phenotypic description of the new trait(s);
- D) the level and site of expression in the plant of the expressed gene product(s), and the levels of its metabolites in the plant, particularly in the edible portions; and
- E) where possible, the amount of the target gene product(s) if the function of the expressed sequence(s)/gene(s) is to alter the accumulation of a specific endogenous mRNA or protein.

<sup>11</sup> For example the FlavrSavr™ tomato, which contains an antisense sequence corresponding to the polygalacturonase encoding gene.

<sup>12</sup> When equivalence is demonstrated based on serological cross-reactivity between the plant and bacterial proteins, it is important to use antisera (either polyclonal or monoclonal antibodies) that have been well characterized with respect to their specificity.

#### **CODEx GUIDELINE PARAGRAPH 33.** In addition, information should be provided:

- A) to demonstrate whether the arrangement of the genetic material used for insertion has been conserved or whether significant rearrangements have occurred upon integration;
- B) to demonstrate whether deliberate modifications made to the amino acid sequence of the expressed protein result in changes in its post-translational modification or affect sites critical for its structure or function;
- C) to demonstrate whether the intended effect of the modification has been achieved and that all expressed traits are expressed and inherited in a manner that is stable through several generations consistent with laws of inheritance. It may be necessary to examine the inheritance of the DNA insert itself or the expression of the corresponding RNA if the phenotypic characteristics cannot be measured directly;
- D) to demonstrate whether the newly expressed trait(s) are expressed as expected in the appropriate tissues in a manner and at levels that are consistent with the associated regulatory sequences driving the expression of the corresponding gene;
- E) to indicate whether there is any evidence to suggest that one or several genes in the host plant has been affected by the transformation process; and
- F) to confirm the identity and expression pattern of any new fusion proteins.



The application of modern profiling technology, such as DNA/RNA microarrays, proteomics, gas chromatography coupled to mass spectrometry (GC-MS) or liquid chromatography coupled to nuclear magnetic resonance (HPLC-NMR), has the potential to broaden the data available for the safety assessment. Sensitive profiling methods may provide indications of minor or major changes at the genome level in mRNA expression or protein production, and/or changes at the level of metabolism. These broad, non-targeted approaches, which do not require prior knowledge of hypothesized changes in levels of particular plant constituents to guide the choice of method, could be of particular interest for foods derived from recombinant-DNA plants modified through the insertion of multiple genes, such as plants with nutritional or health-promoting characteristics (see also the chapter on Evaluation of metabolites).

The utility and applicability of these non-targeted techniques for generating data for risk assessment purposes need further exploration, in particular with respect to establishing and validating the relevance to food safety of any observed changes. One of the major challenges in using these techniques is that observed differences may not be easily distinguishable from natural variations (baseline fluctuations in several thousand variables) in biochemical composition due to the properties of different varieties, the stage of plant development and the health status of the plant, and environmental influences and variations in growth conditions. Profiling methods are not yet suitable for routine risk assessment purposes because the observed variation in profiles cannot be routinely linked to specific biosafety considerations. Further description of baseline ranges, cost reduction, and development and validation of methods are needed.

### Randomly generated plant transformation events

The transgene is generally integrated into the host chromosome(s) upon successful application of transformation processes such as the *Agrobacterium*-mediated or biolistics (microprojectile bombardment) methods. Some insertions occur in regions of the plant genome that are not involved in any obvious function, in which case the transgene may express the novel protein as expected without causing unintended change in other plant traits.

When the random insertion occurs in a region of the plant genome that is involved in genome regulation, transcription or protein production, the insertion may lead to unintended plant phenotypes. Each of the plants recovered after the transformation process that is carrying the integrated DNA represents a unique gene transfer “event”.

Because insertion of the transgene into the host plant genome occurs randomly, a large number of transformed plants are usually produced initially, each containing single or multiple copies of the transgene. Subsequent small-scale cultivation and selection-based screening will remove unintended phenotypes possessing unwanted traits and/or multiple copy insertion “events” and preserve the most suitable phenotypes for further characterization and further rounds of selection-based breeding to obtain elite cultivars.

### Transgene detection using event-specific primers

Two DNA primers (each 20–30 bases long) with nucleotide sequences complementary to the DNA inserted into the recombinant-DNA plant are generally employed in a polymerase chain reaction (PCR) to detect the presence of a transgene. If both of the PCR primers are complementary to the transgene sequence, then all plant varieties and species that carry the same transgene will show the PCR amplification product, irrespective of the location of the insertion in the plant genome. However, it is possible to distinguish among the different insertion “events” of the same transgene in the same plant cultivar by designing the primer pair appropriately.

Event specificity is based on using a primer pair of which one primer is complementary to the plant genomic region adjacent to the point of insertion of the transgene, and the other primer is complementary to a region within the transgene. These primers are known as “event-specific” primers. This primer pair will only amplify a specific insertion “event” because the process of DNA insertion into plants is effectively random. Therefore, each insertion of DNA will take place at random in the plant genome and will lead to that insertion having unique flanking regions of plant DNA.

The use of event-specific primers is necessary for identifying a particular transformation event among other events carrying the same gene in the same host variety or other varieties of the same crop species. Hence, access to sequence information for the flanking regions of the integration site of the inserted DNA is necessary so that regulatory authorities can conduct event-specific monitoring of recombinant-DNA plants. Due to the large variety of plant cultivars harbouring the same transgene, monitoring of recombinant-DNA plants is typically done in two steps. Step one, which is PCR-based, determines the presence of frequently used transgene constructs, and if this is positive, a second-step (also PCR-based) is performed, which employs event-specific primers.

For examples of the use of event-specific primers, see the validated methods published online by the European Commission’s Joint Research Centre: <http://gmo-crl.jrc.it/default.htm>

## Extent of refinement at the current level of the technology

Unintended changes can result from the random insertion of DNA sequences into the plant genome, which may cause modifications in the expression of existing genes, or activation of silent genes, possibly resulting in elevated levels of native or new toxins in the food. It is emphasized that the occurrence of unintended effects is not specific to the application of recombinant-DNA technology in plants, as it also occurs in classic plant breeding. In breeding practice, backcrossing and selection based on morphology, yield, crop quality, insect/disease resistance, etc. identify lines with unwanted characteristics that are discarded<sup>13</sup>. Similarly, during the development of recombinant-DNA plants, modified lines that do not meet the expected agronomic, safety and quality requirements will be discarded, resulting in the elimination of many unintended effects from the tissue culture or DNA insertion process<sup>14</sup>.

A limitation in the current application of recombinant-DNA technology in plants is the inability to direct the insert DNA (transgene) into a specific genomic location. Further developments in the technology leading to the option to specifically target the DNA insertion to particular genomic regions may eliminate unintended effects such as positional effects on transgene expression and the influence of the insert on plant genome expression ●

<sup>13</sup> Reports of unintended effects that may affect human health are rare, and include examples such as low yields in barley or maize, high content of furanocoumarins in celery, and high glycoalkaloid content in potatoes.

<sup>14</sup> Examples of unintended effects that have been observed in recombinant-DNA plants are potatoes with abnormal tuber tissue or with reduced glycoalkaloid content, soybeans with higher lignin content, and rice with increased Vitamin B6 content or higher levels of certain carotenoid derivatives.