

Molecular characterisation

Working Group "Molecular characterisation of transgenic plants": Geert Angenon, Marc Boutry, Marc De Loose, Jacques Dommes, Godelieve Gheysen, Pierre Van Cutsem, Nathalie Verbruggen; Secretariat: Adinda De Schrijver

1 Introduction

Following the comparative approach the characterisation of the genetic modification allows to identify at molecular level the differences between the genetically modified crop and its conventional counterpart. In terms of the safety assessment these data are important to characterise the introduced or modified trait and to check if unintended effects of the genetic modification have taken place, for instance as a consequence of possible rearrangements of the DNA at the site of insertion. The (extent of) expressions of (new) proteins in the genetically modified crop should also be assessed to identify the possible occurrence of toxicants and allergens, which have to be evaluated along with the information described in chapters III and IV.

2 Information on the presentation of data

Any relevant scientific articles and reports referred to should accompany the dossier. The quality of the experimental data should also be sufficient to verify clearly any statements made by the applicant and the notifier should submit all sequences in electronic format.

3 Description of the genetic material used for the transformation

For each of the vectors used for transformation, the notifier has to provide a detailed map including the genetic elements listed hereafter, indicating their location, order and orientation in the vector, and the position of relevant restriction sites.

For each of the vectors used for transformation, a list (and a table summarizing name, position and brief description) should be provided of all genetic elements, including coding and non-coding sequences (e.g. origins of replication, T-DNA borders of *Agrobacterium*, bacterial transposable elements, promoters). For each of these elements the notifier should include:

- A **description** of the genetic element, or a **citation** where the genetic element was isolated and characterized (completed with an accession number in a publicly available database).
- The **portion** and **size** of the genetic element that was inserted in the vector, and its **location** in the vector.
- Information about its **source**. The scientific and the common or trade name of the donor organism should be given. The history of use of the

donor organism (or that of relevant elements thereof) and its relevance to risk assessment should be described, indicating whether the donor organism is responsible for any disease or injury to plants or other organisms (e.g. produces toxicants, allergens, pathogenicity factors or irritants).

- Information on whether the genetic element itself is coding for or involved in the production of proteins responsible for **disease or injury** to plants or other organisms (e.g. a toxicant, allergen, pathogenicity factor or irritant).
- Information about the molecular, biochemical and physiological **properties of its products**, as known in the donor organism and aimed at in the transgenic plant.

For direct transformation methods, data should be provided on how the part(s) of the vector(s) used for the transformation was purified and indicate how purity was assessed.

4 Description of the transformation method

The transformation protocol should be **described** and relevant **references** should be provided for the transformation method. In case of direct transformation pure DNA has to be used, implying the absence of carrier DNA.

For *Agrobacterium*-mediated transformation, the **strain** designation of the *Agrobacterium* used during the transformation process should be provided, and it should be indicated if and described how the Ti/Ri plasmid based vector was **disarmed**.

For transformation methods that involve the use of **helper plasmids**, these plasmids should be described in detail.

5 Description of the transgene loci

Experimental data should be included revealing the **number of sites** where (part of the) DNA used for the transformation is inserted and it should be indicated whether it is located in the **nucleus, mitochondria or chloroplasts**. For allopolyploid plants it should be indicated into **which parental genome** these insertions have occurred. The methods that were used should be described and their **sensitivity** should be assessed.

For each of the insertion sites the notifier should provide:

- The sequence of the entire insert and of both flanking regions (about 500 basepairs, proven to correspond to plant DNA using appropriate methods). The different genetic elements should be delineated and any rearrangements on a schematical representation of the transgene locus should be indicated. A list of all the genetic elements and rearrangements should be provided and, for each of these, its position and origin has to be indicated and its integrity to be evaluated.
- An examination of the **nature of the flanking sequences** using up-to-date bioinformatics tools (database searches, prediction models, etc.).

- An evaluation of the presence and functionality of **novel chimaeric open reading frames** applying up-to-date bioinformatics tools (database searches and application of prediction models designed to reveal the presence of open reading frames, searches for immunologically and toxicologically relevant similarity, etc.) on the complete sequence. If a chimaeric open reading frame is detected that extends beyond the region, more of the flanking region should be sequenced until the putative end of this open reading frame is reached.

6 Transcript and protein characterisation

The expression of all open reading frames identified in section 5 should be analysed. The methods that were used for the expression analysis should be described and their sensitivity assessed.

- For open reading frames intended to be expressed in the transgenic plant, data should be provided on the levels and the spatial and temporal specificity of expression at the protein level. In case the purpose of the transformation is to alter the expression of endogenous genes (e.g. by antisense constructs, ribozymes, or via the mechanism of RNA silencing), data should be included on the expression of the target.
- For all other genes present on the DNA used for the transformation and (partly) inserted in the genome of the transgenic plant, data should be provided on the levels and tissue specificity of expression at the transcript and/or protein level, unless it can be demonstrated that the necessary regulatory sequences for expression are not linked to the open reading frame, or unless the open reading frame is linked to a non-plant promoter for which it can be demonstrated or for which reference can be provided that it is not functional in the plant, or unless it can be otherwise demonstrated that the open reading frame is not expressed. If transcription occurs, it should be determined whether the transcript is translated.
- For novel chimaeric open reading frames identified in section 5, data should be provided on the levels and tissue specificity of expression at the transcript and/or protein level. If transcription occurs, it should be determined whether the transcript is translated.

The **properties** of the **proteins** that are expressed in the plants or the target proteins of which the expression level has been altered should be described.

If there has been a DNA modification that affects the amino acid sequence of the plant expressed protein, the **modified amino acid sequence** must be provided. It should be indicated whether the modifications are known or expected to result in changes in the properties of the protein.

On a case-by-case basis, data on protein stability in the cell and the environment may be required.

7 Inheritance and stability

Statistically significant data should be provided that demonstrate the inheritance pattern and the stability of the sequences inserted, as well as data that demonstrate

the stability of expression of all proteins that are intended to be expressed in the plant or of the target proteins of which the expression level has been altered.

For plants which are either infertile or for which it is difficult to produce seed (e.g. vegetatively propagated male-sterile potatoes, plants with long sexual generation times such as trees), statistically significant data should be provided to demonstrate that the transgene trait is stably maintained and expressed during vegetative propagation.

8 Detection and identification

The sequence of a **primer pair** should be provided which enables the unequivocal identification of the transformation event, as well as a detailed **protocol** for its use for identification, detection and quantification purposes.

Reference transgenic and control **material** should be provided at the time of deposition of the dossier.