SCIENTIFIC OPINION

Draft guidance on the agronomic and phenotypic characterisation of genetically modified plants[1]

EFSA Panel on Genetically Modified Organisms (GMO)[2, 3]

European Food Safety Authority (EFSA), Parma, Italy

ABSTRACT

This document provides guidance for the agronomic and phenotypic characterisation of genetically modified (GM) plants. It aims to provide a comprehensive and harmonised approach for the agronomic and phenotypic characterisation of GM plants, and to ensure the best use of those data for the comparative analysis and food/feed and environmental risk assessment. Specific recommendations are given on: the representativeness of the selected sites of field trials; the representativeness of the selected test materials and their quality; the experimental design of field trials; the selection and description of (ecologically) more relevant endpoints; data analysis; and risk assessment implications.

KEY WORDS

Comparative analysis, field trial design, invasiveness, persistence, receiving environments, representativeness, unintended effects

[1] On request from EFSA, Question No EFSA-Q-2013-00606 adopted on 10 September 2014.

[2] Panel members: Salvatore Arpaia, Andrew Nicholas Edmund Birch, Andrew Chesson, Patrick du Jardin, Achim Gathmann, Jürgen Gropp, Lieve Herman, Hilde-Gunn Hoen-Sorteberg, Huw Jones, Jozsef Kiss, Gijs Kleter, Martinus Lovik, Antoine Messéan, Hanspeter Naegeli, Kaare Magne Nielsen, Jaroslava Ovesna, Joe Perry, Nils Rostoks, Christoph Tebbe Correspondence: gmo@efsa.europa.eu.

[3] Acknowledgement: The Panel wishes to thank the members of the Working Group on the agronomic and phenotypic characterisation of genetically modified plants: Salvatore Arpaia, Hans Christer Andersson, Paolo Bàrberi, Lammert Bastiaans, Tom de Jong, Thomas Frenzel, Jürgen Gropp, Huw Jones, Antoine Messéan, Joe Perry, Angelo Porta Puglia and Geoffrey Squire and the hearing expert: Jeremy Sweet for the preparatory work on this scientific opinion and EFSA staff: Hermann Broll, Yann Devos, Andrea Gennaro and Claudia Paoletti for the support provided to this scientific opinion.

Suggested citation: EFSA GMO Panel (EFSA Panel on Genetically Modified Organisms), 2014. Draft guidance on the agronomic and phenotypic characterisation of genetically modified plants. EFSA Journal 201Y;volume(issue):NNNN, 57 pp. doi:10.2903/j.efsa.201Y.NNNN

Available online: www.efsa.europa.eu/efsajournal

21 **SUMMARY**

22 A summary will be provided after the public consultation.

23

TABLE OF CONTENTS

BACKGROUND AS PROVIDED BY EFSA

Genetically modified organisms (GMOs) and derived food and feed products are subject to a risk analysis and regulatory approval before they can enter the market in the European Union (EU). In this process, the role of the European Food Safety Authority (EFSA) is to independently assess and scientifically advice risk managers on any possible risk that the use of GMOs may pose to human and animal health and the environment. EFSA's scientific advice is elaborated by its GMO Panel (referred to hereafter as EFSA GMO Panel) with the scientific support of specific working groups and EFSA scientists.

The main focus of EFSA in the field of GMOs lies in the evaluation of authorisation applications for the marketing of GMOs (referred to hereafter as GMO applications) and the development of risk assessment guidelines.

The EFSA GMO Panel developed several guidelines for the risk assessment of GMOs, as well as on specific aspects of their risk assessment (see Devos et al., 2014 for a comprehensive overview). These guidelines assist applicants in the preparation and presentation of their applications by describing elements and data requirements for the risk assessment of GMOs.

Key guidelines developed by the EFSA GMO Panel and used for the evaluation of risk assessments supplied by applicants as part of their GM plant applications, are those on the risk assessment of food and feed from GM plants (EFSA, 2011a) and on the environmental risk assessment of GM plants (EFSA, 2010a).

- The EFSA GMO Panel guidelines on the risk assessment of food and feed from GM plants (EFSA, 2011a) outline the principles of the risk assessment of food and feed containing, consisting or produced from GM plants, and provide definitions of the different steps and objectives of the risk assessment process. These guidelines incorporate previously issued guidelines on specific aspects, such as: (1) the selection of comparators for the risk assessment of GM plants and derived food and feed (EFSA, 2011b); and (2) statistical considerations for the safety evaluation of GMOs (EFSA, 2010b).

- The EFSA GMO Panel guidelines on the environmental risk assessment of GM plants (EFSA, 2010a) describe the principles to be followed when assessing potential effects of GM plants on the environment, and provide the scientific rationale underpinning the necessary data requirements for a comprehensive environmental risk assessment.

Although the abovementioned guidelines address the generation, analysis and interpretation of agronomic and phenotypic data used in support of the comparative analysis of GM plants, the EFSA GMO Panel and national risk assessment bodies have identified the need for further guidance, in order to provide a more comprehensive and harmonised approach for the agronomic and phenotypic characterisation of GM plants. Therefore, the EFSA GMO Panel decided to develop its guidelines on the agronomic and phenotypic characterisation of GM plants further.

TERMS OF REFERENCE AS PROVIDED BY EFSA

To establish a new working group as part of a self-task activity, in order to supplement and specify the EFSA GMO Panel requirements for the agronomic and phenotypic characterisation of GM plants. During the development of these specific guidelines, EFSA shall consult EU Member States and all relevant stakeholders via an online public consultation. Scientific comments received during these consultations shall be considered when finalising the guidelines.

ASSESSMENT

The EFSA GMO Panel regularly reviews its risk assessment guidelines in the light of experience gained, technological progress and scientific development, and identified the need to develop its guidelines on the agronomic and phenotypic characterisation of GM plants further.

1. INTRODUCTION

The risk assessment of GM plants includes a comparative analysis in which appropriate methods are used to compare the GM plant with its appropriate comparator (OECD, 1993; FAO/WHO, 2000; Kok and Kuiper, 2003; Codex Alimentarius, 2007; EFSA, 2010a, 2011a). The underlying assumption of the comparative approach is that traditionally cultivated non-GM crop plants have gained a history of safe use for consumers and animals, and familiarity for the environment. This approach therefore enables to evaluate the relevance of potential risks posed by a GM plant and derived food and feed products in the context of the impacts posed by its comparator, by assessing whether intentionally and unintentionally modified properties of the GM plant alter the level of risk or give rise to additional risks.

In the context of the food and feed safety assessment, the comparative analysis requires the simultaneous application of two complementary tests: a test of difference, to identify possible differences between the GM plant and its appropriately selected comparator; and a test of equivalence to assess whether the characteristics of the GM plant fall within the range of natural variation estimated from a set of conventional non-GM reference varieties (referred to hereafter as non-GM reference varieties) with a history of safe use (EFSA, 2010b, 2011a; van der Voet et al., 2011).

The comparative analysis typically relies on data from three sources: molecular, agronomic/phenotypic and compositional characterisation of the GM plant[4]. This document focuses on the agronomic and phenotypic characterisation of the GM plant, but also considers the interplay between agronomic and phenotypic data, the other data sources and the various areas of concern considered in the risk assessment.

Data on agronomic and phenotypic characteristics are usually obtained from field trials performed in multiple sites under conditions that should be representative of receiving environments (REs) where the GM plant may be grown (EFSA, 2011a; EU, 2013). The field trials can be conducted in a single year, or spread over multiple years (EU, 2013). To reliably estimate the natural variation of plant characteristics, non-GM reference varieties must be included in the field trials, along with the GM plant and its comparator(s).

Typically, a range of plant characteristics (endpoints) are measured in the agronomic and phenotypic field trials, covering plant vigour, growth and development, morphology, yield, crop characteristics, pest and disease susceptibility and fertility (i.e., Horak et al., 2007; Nickson, 2008; Garcia-Alonso, 2010; Raybould et al., 2010, 2012). These are selected on a case-by-case basis, depending on the plant species, the trait(s), and the intended uses of the GM plant, which are determined by the scope of the GMO application. Seed and pollen characteristics of the GM plant are also assessed, but typically under controlled instead of field conditions.

The agronomic and phenotypic dataset can also provide useful information on the quality of the field trials used to produce material for the compositional analyses when those data are generated in the same field trials. Moreover, they can help to explain the cause of differences observed in specific compositional endpoints. Through the detection of intended and unintended changes in the GM plant, the agronomic and phenotypic characterisation also informs the assessment of the persistence and invasiveness potential of the GM plant, and gives indications of potentially altered biotic and abiotic interactions.

[4] For GM plant applications for cultivation, the EFSA GMO Panel also requires data on interactions of the GM plant and its comparator with non-target organisms.

1.1. Objectives

Based on an analysis of the agronomic and phenotypic data recently supplied by applicants as part of their GM plant applications and relevant scientific literature, the EFSA GMO Panel identified aspects of the agronomic and phenotypic characterisation of GM plants that deserve greater attention. These include among others:

- the representativeness of the selected sites of field trials;
- the representativeness of the selected test materials and their quality;
- the experimental design of field trials;
- the selection and description of (ecologically) more relevant endpoints;
- the data analysis; and
- environmental risk assessment implications.

For the above mentioned aspects specific recommendations are developed in this document, in order to: (1) provide a more comprehensive and harmonised approach for the agronomic and phenotypic characterisation of GM plants; and (2) ensure the best use of the agronomic and phenotypic data to inform the food/feed and environmental risk assessments.

(1) Heterogeneity in agronomic and phenotypic datasets supplied by applicants has been observed. This heterogeneity is not unexpected, because such datasets are generated by different applicants and cover different plant species and trait combinations. Moreover, there are no internationally agreed standards on how to perform the agronomic and phenotypic characterisation of GM plants. Therefore, the EFSA GMO Panel intends to provide guidance to support the harmonisation of the agronomic and phenotypic characterisation of GM plants. This should ensure the generation of comprehensive datasets that can be analysed in a standardised manner.

(2) Additional recommendations are developed to warrant the best use of agronomic and phenotypic data for the risk assessment. The EFSA GMO Panel also explored whether and how agronomic and phenotypic datasets can be extended to generate data more informative for the environmental risk assessment.

1.2. Scope

This document provides guidance for the agronomic and phenotypic characterisation of GM plants, and will assist applicants in the generation, analysis and interpretation of the agronomic and phenotypic dataset submitted as part of their GM plant applications in the frame of Regulation (EC) No 1829/2003 It supplements the general recommendations on the agronomic and phenotypic characterisation of GM plants outlined in the EFSA GMO Panel guidelines on the risk assessment of food and feed from GM plants (EFSA, 2011a), the environmental risk assessment of GM plants (EFSA, 2010a) and the Implementing Regulation (EU) No 503/2013 (EU, 2013).

The guidelines apply to the most common plant species and trait combinations for which GM plant applications have been submitted so far in the frame of Regulation (EC) No 1829/2003. GM plant applications cover mostly maize, cotton, soybean, and – to a lesser extent – oilseed rape, potato and sugar beet. The predominant traits are resistance to insect pests and tolerance against certain herbicidal active substances, but they also include modified composition (e.g., altered fatty acid profile), tolerance to drought or reduced amylose content.

A growing fraction of GM plant applications concern so called "stacked" events, i.e. GM plants in which two or more single events have been combined by conventional crossing, in order to introduce several traits into one plant. Therefore, the guidelines address both GM plants containing single and stacked events.

Even though the scope of most GM plant applications submitted in the frame of Regulation (EC) No 1829/2003 is limited to import and processing for food and feed uses or for industrial uses and does not include cultivation in the EU, both categories of intended uses are covered here.

In the next decade, EFSA may be faced with a number of challenges that may require the adjustment and revision of its current risk assessment guidelines or the development of new risk assessment strategies. Therefore, the EFSA GMO Panel will consider the possibility to complement its guidelines with specific recommendations on additional plant species (e.g., perennials, trees), topics (e.g., sampling, with the aim to harmonise sampling approaches for generation of data in support of the risk assessment of GM plants[5]), and technological developments (e.g., phenotyping platforms), as needed.

[5] See internal mandate proposed by EFSA to the GMO Unit for a procurement on the development of reliable sampling approaches for GM plants (http://registerofquestions.efsa.europa.eu/roqFrontend/questionLoader?question=EFSA-Q-2013-00904; Contract number OC/EFSA/GMO/2013/04, http://www.efsa.europa.eu/en/tendersawarded/tender/ocefsagmo201304.htm).

2. SELECTION OF SITES AND TEST MATERIALS

2.1. Suitability and representativeness

In order to ensure that the comparative analysis is appropriate for GM plants that will be grown and/or imported, the Implementing Regulation (EU) No 503/2013 states a number of conditions in its Chapter 1.3.2.1(b). These conditions relate to: (i) the suitability of a site; (ii) the suitability of a variety for a site (see Section 2.1.1, below, for further details of the definitions of varieties); and (iii) the representativeness of the set of sites for REs, as a whole. Here, *suitability* refers to the appropriateness (under optimal or sub-optimal conditions) of either the site selected or the variety to be grown at a particular site; *representativeness* refers to whether the totality of the sites used in the field trials captures sufficiently the variability between the possible totality of REs deemed suitable.

The first three conditions stated in the Implementing Regulation (EU) No 503/2013 are that: "the different sites selected for the field trials shall reflect the different meteorological and agronomic conditions under which the plant is to be grown [*suitability and representativeness of sites*]; the choice shall be explicitly justified. The choice of non-GM reference varieties shall be appropriate for the chosen sites [*suitability of varieties*] and shall be justified explicitly. if the sites cover a restricted range of growing conditions, then the applicant shall replicate the field trials over more than one year [*representativeness of sites*]". The fourth condition is that: "*each field trial shall be replicated at a minimum of eight sites, chosen to be representative of the range of likely REs where the plant is to be grown* [*representativeness of sites*]".

There are several factors that influence how suitable a site is to grow a particular variety of the crop in question. These include, but are not restricted to: soil characteristics, climatic conditions (average temperature over a particular period, average rainfall during the growing season, maximum expected wind speed close to harvest, etc.); day length; soil moisture and fertility; biotic factors (presence/absence of pests and natural enemies); earliness of flowering; etc.

Applicants are therefore requested to demonstrate that the varieties used for the field trials are suitable, and that the sites used are both suitable and representative of the REs, using the protocols developed below.

2.1.1. Receiving environments

As defined in the EFSA GMO Panel guidelines for the environmental risk assessment of GM plants (EFSA, 2010a), REs are the environments in which the GM plant(s) may be grown. Current EFSA GMO Panel guidelines already request that a description of REs of field trials is submitted, but further guidance is needed to make it possible to assess the suitability and representativeness of the selected sites. The extent of REs that should be considered in the risk assessment could be restricted by the specific trait introduced: certain traits, such as disease or insect resistance, are primarily designed for crops growing in specific areas where the targeted disease or pest is present.

REs are determined and influenced by three components (see Figure 3 of EFSA, 2010a): the GM plant (e.g., plant species, genetic transformation(s) and intended uses(s)), the geographical zone (characterised by, for example, e.g., climate, altitude, soil, water, flora, fauna, habitats); the management systems (e.g., land use and production systems, other cultivated GM plants, cultivation practices, integrated and other pest management, non-production activities and nature conservation activities typically associated with the crop).

These components can differ significantly between geographical regions and result in biotic and abiotic interactions that shall be considered by applicants when considering REs for planning field trials to generate agronomic and phenotypic (and possibly also compositional) data. It is also recommended that applicants pay special attention to the selection of the particular line(s), which is/are the recipient of the event for testing, as well as likely crop management practices associated with

those lines when considering the suitability (Section 2.1.3, below) and representativeness (Section 2.1.4, below) of sites for selection. Due to the diversity of REs into which a GM plant may be grown, the following sections: discuss site and variety selection for field trials; provide guidance on the necessary information for an adequate assessment of whether selection is appropriate; and give a framework to assess the suitability and representativeness of field trials.

2.1.2. The extent of receiving environments

The total worldwide set of regions in which a particular crop is typically grown, irrespective of the variety, is termed here 'RE-plant'. This definition excludes areas where the plant might be spilled during transit or import.

Within each of such regions, there are sub-regions where typically the event could be present if it were bred into a suitable genetic background. Except in special cases (see below), the totality of such sub-regions cannot exceed the extent of 'RE-plant'. The extent of this set of sub-regions is termed here 'RE-event'.

The GM plant assessed in the GM plant application consists of the event introgressed into one or few particular line(s). The set of areas in which this/these plant(s) typically could be grown is clearly smaller than the area where the event could be present, and is termed here 'RE-line'. The GM line tested is generally unlikely to be suitable for all possible REs for the event and experimental datasets in applications are therefore usually limited to a restricted range of REs.

The distinction between these different REs is illustrated by the map in Figure 1. A special case occurs when the trait of the GM plant is designed to alter the extent of the REs, as, for example, would be the case for drought tolerance. Then, the extent of RE-line and RE-event may overlap with that of RE-plant, without being completely contained by it, and/or may contain RE-plant in extent, since the GM plant may be grown in the future in areas currently unsuitable for the crop.

Figure 1: Geographical distribution of a hypothetical GM plant, restricted for simplicity to the continents of North and South America. The extent of RE-plant is shown by the area enclosed within the red ellipses; similarly for RE-event (blue ellipses) and RE-line (green ellipses). Usually, Extent(RE-plant) ≥ Extent(RE-event) ≥ Extent(RE-line), as depicted. However, the GM trait concerned may alter the area within which the plant may be grown; an example might be drought tolerance. In that case, the green and blue ellipses might overlap the red ellipse without being completely contained within it, or may themselves completely contain the red ellipse.

2.1.3. The assessment of site suitability

To assess the suitability of a site for field trials, it is necessary to determine whether the site is contained within RE-event and RE-line. As illustrated by Figure 2, this would be determined by confirming that all the asterisks representing site locations were within the areas denoting the RE-line.

There are several factors that influence how suitable a site may be for a particular variety. Applicants should assess suitability explicitly listing all these factors and possibly quantifying them to demonstrate a matching between the requirements of the GM line plus its comparator and the attributes of each selected site.

The submitted information describing REs could also take the form of maps similar to Figures 1 and 2 and/or a table of factors in a qualitative, or preferably quantitative format including relevant information for each site: soil characteristics; climatic (average temperature over a particular period, average rainfall during the growing season, maximum expected wind speed close to harvest, etc.); day length; soil moisture and fertility; biotic factors (presence/absence of pests and natural enemies) and earliness of flowering (see Sections 4.1 and 4.2).

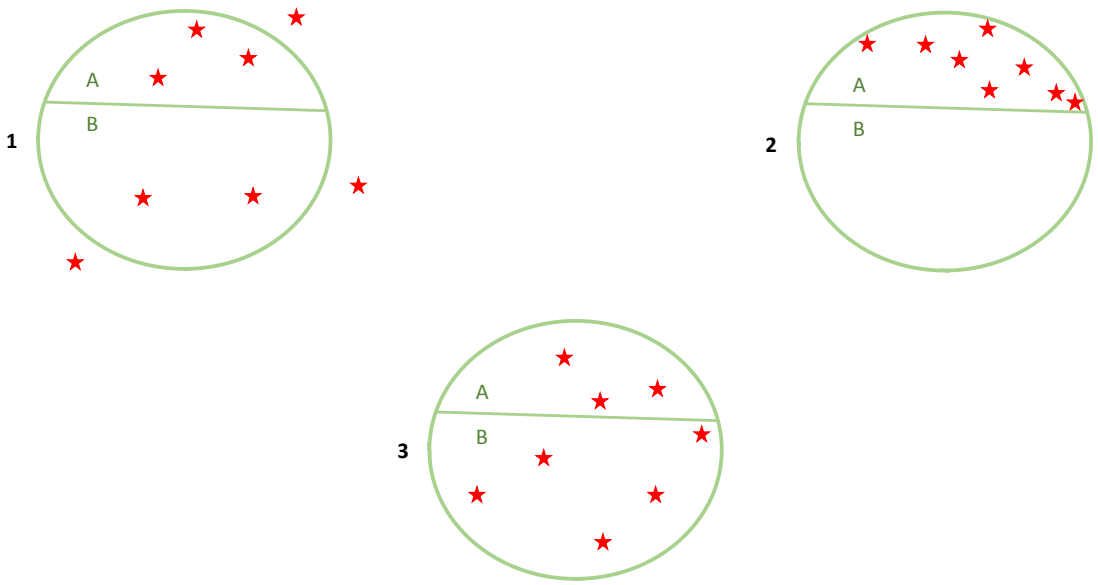


Figure 2: Suitability and representativeness illustrated by the extent of RE-line (the green ellipse in North America from Figure 1, above). Smaller area A, in the northern part of RE-line, is assumed to have heavy soils; larger area B, in the southern half, to have light soils. Sites for field trials are depicted by the red asterisks. In 1, some sites are outside the extent of RE-line and are deemed by definition unsuitable for the specific GM line under consideration and its comparator. In 2, all sites are suitable, but no site is located on a light soil, so the sites as a whole are unrepresentative. In 3, all the sites are suitable and the sites are distributed proportionately to the areas of A and B, so they are also representative (Garcia-Alonso et al., 2014).

2.1.4. The assessment of site representativeness

The assessment of the representativeness of a set of candidate sites for field trials requires a determination of whether that set adequately represents the variability in possible suitable sites defined by RE-line. Field trials need not necessarily to be distributed geographically across all possible RE-line areas. However, if they are located in a limited geographical area (e.g., the Midwest USA), then applicants should demonstrate that they are also representative of other possible geographical areas of the RE-line. As an example, it is possible that all the sites selected might be suitable, but be so similar in soil type, that they would fail to represent other soil types appropriate for the GM plant and its comparator. Hence, if the GM plant and its comparator could be grown equally well on clay, loam and sandy soils, but field trials included only sandy sites, potential effects manifested most strongly on clay sites might be underestimated or not identified (see for example, Figure 2).

Representativeness should be assessed, wherever possible using a quantitative approach based on the set of factors[6] identified by the applicant to assess suitability, above (and see Section 4.2, below). Existing data on crop production worldwide or in the EU, as well as information from variety registration, if available, are of specific relevance to demonstrate representativeness.

It is recommended that applicants discuss the representativeness of the selected sites not just with regard to RE-line, but also with regard to RE-event, by considering the likely REs which will be applicable after commercialisation, once the event is introgressed into different genetic backgrounds.

[6] These include but are not restricted to: climatic (average temperature over a particular period; average rainfall during the growing season; maximum expected wind speed close to harvest, etc.); soil moisture and fertility; biotic (presence/absence of pests and natural enemies); earliness of flowering; etc.

In particular, as there is a broad range of genetic variability in most crop varieties any major differences between the extent of RE-line and RE-event should be identified and the likely consequences considered. If RE-line is not sufficiently representative of RE-event in terms of soil, climate, pest profiles, and/or management, then the applicant should consider using more than one genetic background, to ensure that the variability of RE-event is sufficiently addressed by field trials.

2.1.5. The assessment of non-GM reference varieties suitability

Applicants should demonstrate that the non-GM reference varieties used at each site are well adapted to the local environmental/agronomic conditions and thus appropriate for the selected sites. The methodologies to assess suitability are outlined in Section 2.1.3. If one or more varieties are well adapted to a range of local environmental and agronomic conditions and thus appropriate for more than one trial site, then applicants should preferentially include them in the non-GM reference varieties used to allow the presence of the same variety(s) in the across site comparison (see Section 6.2). Information may be supplemented from recommended lists of commercial varieties and variety registration data.

2.2. Spatial and temporal representativeness

As outlined in the EFSA GMO Panel guidelines on the risk assessment of food and feed from GM plants (EFSA, 2011a), "*the different sites selected for the trials should be representative of the range of receiving environments where the crop will be grown, thereby reflecting relevant meteorological, soil and agronomic conditions; the choice should be explicitly justified*". Therefore, applicants should implement sufficient numbers of sites to ensure that a representative portion of the environmental variability in the field trials is captured. Even if there is no variability in REs, "*each field trial shall be replicated at a minimum of eight sites, chosen to be representative of the range of likely REs where the plant is to be grown. The field trials may be conducted in a single year, or spread over multiple years*" (EU, 2013).

As indicated in the Implementing Regulation (EU) No 503/2013, "*the different sites selected for the field trials shall reflect the different meteorological and agronomic conditions under which the crop is to be grown; the choice shall be explicitly justified*". Therefore, applicants are requested to describe and justify the criteria used for the sites selection, to provide location-specific environmental data to allow a detailed assessment of the selected sites (see Sections 2.1 and 4 for further details).

Agrometereological data generated *a posteriori* are important to confirm that the selection of the sites was appropriate and thus successful in capturing environmental heterogeneity or not. If environmental heterogeneity was not sufficiently captured, then applicants should consider the inclusion of additional sites in the field trials dataset and analyse the complete data package with the appropriate statistical model.

3. QUALITY OF STARTING MATERIALS USED AS TEST MATERIALS

The EFSA GMO Panel guidelines on the selection of comparators for the risk assessment of GM plants and derived food and feed (EFSA, 2011a) specify how to select an appropriate comparator. Here, the EFSA GMO Panel complements its previous guidelines by providing recommendations on how applicants should demonstrate the quality of the starting material used as test material in the comparative studies.

The starting material (mostly seeds) should be of high quality, and as homogeneous as possible, in order to minimise confounding factors in the comparative assessment. Seed quality in terms of purity (including presence/absence of events) and health should be documented. The OECD schemes for the varietal certification of seed moving in international trade promote the use of agriculture seed of consistently high quality, and set the minimal requirements for the production of basic and certified seeds to achieve this goal. Applicants are recommended to adhere to the relevant minimum requirements for the production of basic and certified seeds outlined in OECD schemes (OECD, 2012, 2014) when generating seeds to be used in comparative studies. Applicants are invited to demonstrate that they complied with the relevant minimal requirements. The abovementioned recommendations apply to all test materials.

For all performed measurements, applicants shall adhere to the principles of sampling described in Section 5.

3.1. Seed production conditions

The seed used in the comparative studies should be as homogeneous as possible in terms of origin, year of production and production conditions, and ensure sufficient seed quality (purity, health see subsequent sections below). It is recommended to produce the seeds of the GM line and comparator under the same environmental and cultivation conditions, and to harvest and store them under similar conditions. If applicants want to use the same seed lots over multiple seasons, then they should ensure that the seed storage conditions do not induce an additional source of variability in the starting materials with the aim to preserve as much as possible the seed germination capacity and health.

Seed coating and any other pre-treatment should be homogeneously applied to the GM line and its comparator. Any difference in treatments between GM line and its comparator should be justified.

Applicants should strive to follow the abovementioned requirements for the non-GM reference varieties as well. Major deviations in terms of environmental and agricultural conditions of production, seed treatments, seed coating or any pre-treatments applied to the non-GM reference varieties as compared with the GM line and its comparator should be justified in the appropriate section of the study, where applicable.

The possible risk of any soil-borne diseases should be limited as much as possible, as they could subsequently be transmitted to the harvested seed that will serve as starting material in the comparative studies. To avoid any contamination by seed-borne diseases transmitted from volunteer plants, such plants should be controlled timely in production sites. If any previous crops in the crop rotation would have made a production site unsuitable for the abovementioned reasons, then adequate remedial measures must be taken (e.g., soil disinfestation). Likewise, production sites for the generation of starting materials should as far as possible be isolated from all sources of seed-borne diseases (including seed-borne virus infection and wild plants that might serve as a source of inoculum).

Plants from which the starting materials are derived shall be inspected at least once at appropriate growth stage(s) by specially trained personnel, in order to check that all the minimum requirements defined above are complied with, and that the occurrence of any seed-borne diseases is at the lowest possible level.

3.2. Seed purity

To guarantee a proper comparison of test materials and limit confounding factors, the genetic homogeneity of the defined seed lots used for any of the comparative studies should be maximised.

Besides the standard practices applied to assess the level of seed purity, applicants shall perform a phenotypic assessment for the presence of undesired seeds in seed lots, and the reporting of any observed contamination (if any). Also, the presence of plants belonging to different species or to interspecific hybrid plants, detected during plant growth, shall be reported (e.g., presence of one or more *Gossypium hirsutum* plants in a *G. barbadense* plot).

The adventitious presence of any non-GM varieties, or any events other than the one under test, should be limited in the GM line(s). The comparator and the non-GM reference varieties should be free from any contamination. Applicants should demonstrate that the test materials are as homogeneous as possible, consider the likely sources of contamination, and should provide adequate evidence to support their statement. Relevant documents supporting the purity of the starting materials should be supplied. In addition, applicants are asked to characterise the purity of the test materials, and to describe the implemented sampling approach, sample preparation (grinded seeds, single seed, etc.) and the tools used to characterise seed purity (such as PCR, qPCR, ELISA) of the test materials[7]. Applicants can use validated detection methods, recognised by international institutions, whenever available. In addition to the event under assessment, other events which have been introgressed into the same species should also be considered by applicants. The identification of event(s) in addition to the one(s) present in the GM line under assessment, should account for the event(s) cultivated in the proximity during seed production, as well as any potential additional source of contamination when handling the plant material (e.g., packaging, storage, labelling).

If genetic impurities are detected in the test materials, the applicant shall discuss the implications on the results of the agronomic and phenotypic characterisation of the GM plant, and any other comparative study performed with these materials. The adventitious presence of undesired events or any other genetic material might result in the rejection of the comparative study, as the proper comparison between the event under investigation and the appropriate comparator cannot be guaranteed.

3.3. Seed health and germination capacity

Seeds used as test material in comparative studies should be of adequate phytosanitary quality, and should have similar germination capacity. Applicants are therefore requested to check the health, germination and vigour of the seeds prior to their use in the comparative studies[8].

Seed health testing along with the knowledge of physiologic seed parameters (germination, vigour) delivers useful information on the phytosanitary quality of the starting materials, and may elucidate causes of poor germination or poor establishment in the field.

Applicants are encouraged to follow the rules on seed health testing defined in ISTA (2014). If other protocols are followed, then applicants should justify this deviation by providing detailed information on the appropriateness of the adopted testing approach.

Seed health test results may indicate the necessity to treat the seed in order to eradicate seed-borne pathogens or to reduce the risk of disease transmission. If pests/pathogens are detected and an

[7] The adventitious presence of undesired genetic material is to be expressed in percentage related to mass fraction (m/m) of GM material in case quantitative event-specific PCR methods are used. If other than quantitative PCR methodologies were used, then the identified levels should be converted into 'percentage related to mass fraction (m/m) of GM material'.

[8] These parameters should be reported as % germinated seeds for seed germination rate, as percentage of infected/contaminated seeds for seed health and as % of normal seedlings to express seed vigour.

effective seed treatment against them exists, then it is recommended to apply it, unless the incidence of the pathogen is below threshold values known to cause significant field crop damage.

Seeds should always be tested for germination (according to ISTA rules or other recognised procedures; see above), as this enables to assess whether the tested seeds have the same germination capacity. If incidences of seed-borne pathogens/pests above damaging thresholds are observed in seed health testing of untreated seed, then seed germination should be assessed also on a seed-treated sample. Seed vigour testing should be done according to ISTA rules or other recognised procedures.

4. DESIGN OF FIELD TRIALS

This section defines the necessary information that applicants should supply to ensure an accurate description and organisation of their field trials. This information will enable a critical appraisal of the agronomic and phenotypic data provided in support of the comparative analysis. The requirements listed in this section apply to all the sites in field trials.

4.1. Data generation for the characterisation of GM plants

The EFSA GMO Panel strongly recommends that applicants generate protein expression[9] and compositional data in the same field trials as those performed for the agronomic and phenotypic characterisation of the GM plant. When these datasets derive from the same field trials, they can deliver useful information on the quality of the trials. Moreover, the variability in protein expression is informative to analyse possible differences in responses of the GM plants (in terms of agronomic/phenotypic and compositional characteristics) between sites. In specific cases, agronomic and phenotypic data can also help to explain the differences observed in specific compositional endpoints between the GM plant and its comparator[10].

4.2. Description of the receiving environment of field trials

The Implementing Regulation (EU) No 503/2013 and the EFSA GMO Panel guidelines on the risk assessment of food and feed and derived products (EFSA, 2011a) require applicants to describe several aspects of the RE as part of their GM plant applicants. The Implementing Regulation states that "*field trials should be adequately described, giving information on important parameters such as management of the field before sowing, date of sowing, soil type, herbicide use, climatic and other cultivation/environmental conditions during growth and time of harvest, as well as the conditions during storage of the harvested material*". The EFSA GMO Panel therefore lists relevant aspects of REs necessitating a comprehensive and accurate description, as this information will enable to assess the suitability and representativeness of the selected sites.

The field trials should be adequately described by providing detailed information on: (i) their geographical location and size; (ii) agrometeorological data; (iii) soil type and soil characteristics; (iv) cropping history; (v) crop management; and (vi) post-harvest storage conditions for harvested materials to be used for further testing. These aspects are elaborated further in the subsections below.

4.2.1. Location of field trials and their size

Applicants are requested to specify the geographical location where field trials have been performed, specifying their geographical coordinates (in terms of latitude, longitude and altitude). Moreover, a map of the graphical position of the selected sites is to be supplied, and the size of the sites is to be specified.

4.2.2. Agrometereological data

Agrometereological data provide fundamental information needed for the description of the REs and the identification of extreme weather conditions. Moreover, they are considered useful for understanding plant development in relation to specific environmental conditions. Therefore, the following information should be provided in electronic format:

- Daily maximum, minimum and average air temperature (°C);
- Daily precipitation (mm);
- Daily maximum, minimum and average air relative humidity (%); and

[9] This recommendation primarily applies to GM plants containing single events.
[10] For example, a GM plant exhibiting altered pest resistance might have altered levels of anti-nutrients or toxicants, or other types of bioactive constituents (see also Section 5.2).

- Exceptional weather conditions (e.g., drought, frost, hail or wind).

Measurement of the following additional parameters is encouraged:

- Daily global radiation (MJ m^{-2}) or daily evaporation (mm)[11]; and
- Daily maximum, minimum and average wind speed (m s^{-1}).

In addition to the data described above, applicants are also required to report descriptive statistics (such as mean and standard error on a weekly basis) throughout the cropping period (from sowing to harvest) in the GM plant application, and to produce a combined thermopluviometric graph for each site, as this facilitates the comparison of trends in agrometeorological data across sites. In addition, meteorological abiotic stressors during the growing season should be clearly reported (e.g., drought, frost, hail or wind).

If one or more of the selected sites are not equipped with an on-site weather station, then data generated from a nearby weather station can be used. In that case, the geographical coordinates of this station (see Section 4.2.1), and the distance (km) between the site of the station and that of field trial should be given.

If available, historical agrometeorological data of the selected trial sites should be provided.

4.2.3. Soil type and mean soil characteristics

Applicants should give a detailed description of the soil type and mean soil characteristics for each of field trial sites. The soil description should reflect field conditions before sowing.

The soil type should be described according to the World Reference Base for Soil Resources (IUSS, 2007)[12].

Individual and mean values from an appropriate and representative number of soil samples taken at each locality should be given for the below soil characteristics:

- Particle size analysis (% clay, % silt and % loam);
- Soil Organic Carbon (%);
- pH;
- Total Nitrogen (‰);
- Extractable Phosphorus (mg P_2O_5/kg);
- Exchangeable Potassium (mg K_2O/kg);
- Cation Exchange Capacity (CEC; cmol+/kg); and
- Total soil carbonates (%).

Any other locally important soil characteristic should also be analysed and reported, such as conditions leading to abiotic stress (e.g., soils with low water infiltration capacity or high boron concentration). All soil characteristics should be measured according to internationally recognised standard methods for soil analysis.

[11] Daily potential evaporation can be estimated from global radiation.
[12] http://www.fao.org/fileadmin/templates/nr/images/resources/pdf_documents/wrb2007_red.pdf

The analysis of the soil samples will provide helpful indications on the level of heterogeneity between the selected sites.

4.2.4. Cropping history

Each field trial site should represent a typical RE for the GM plant, and – to assess their representativeness – relevant descriptors should be considered by applicants (see Section 2). Descriptors for the occurrence of biotic stressors (such as pests, diseases, nematodes, etc.) should be reported for the typical cropping systems in the surrounding areas. In addition, the crops grown during the previous three years should be documented.

4.2.5. Post-harvest conditions

Applicants should describe the conditions under which the harvested/sampled materials have been stored, indicating: (i) the type of storage facilities; (ii) storage temperature; (iii) relative humidity during storage, if relevant; (iv) storage duration; and (v) any other environmental parameters relevant for storage of the specific product. If materials are transported for analysis to facilities distant from the place of storage, then the conditions during transport should also be described (see EFSA, 2011a for further details on sample storage).

4.3. Crop management

Applications should supply the following information for each field trial site:

- Main tillage: (i) type, (ii) depth, and (iii) date;
- Seedbed preparation: (i) type, (ii) depth, and (iii) dates;
- Sowing or planting: (i), date, (ii) method, (iii) number of seeds or vegetative plant parts per m^2, (iv) number of rows per plot, (v) inter-row distance (cm), and (vi) thinning;
- Fertilisers: for each fertiliser application detailed information on: (i) type (element), (ii) method (e.g., on soil, vegetation), (iii) rate (kg fertiliser/ha), (iv) date, and (v) growth stage of crop (BBCH scale);
- Insecticides, fungicides, herbicides, other pesticides, other agrochemicals: for each application detailed information on: (i) date, (ii) type, (iii) target, (iv) active ingredient(s), (v) name of commercial product, (vi) rate (g active ingredients/ha), (vii) application method (e.g., soil, on vegetation, other), (viii) water volume (L/ha), (ix) adjuvants (type and rate), and (x) growth stage of the crop (BBCH scale). These details should be provided also for agrochemicals (e.g., herbicides) applied pre-sowing;
- Irrigation: specify (i) when, (ii) how, and (iii) amount;
- Harvest: (i) type, (ii) moisture (%) of harvested plant part, (iii) date, and (iv) growth stage of crop (BBCH scale); and
- Any other crop-specific management interventions (post-emergence hoeing or cultivation, ridging/hilling, etc.): (i) type, (ii) date, and (iii) growth stage of crop (BBCH scale).

To facilitate registration of information related to application of agrochemicals, applicants are required to fill in a table.

When considering crop protection management measures at each site, applicants should adhere to the principles of Integrated Pest Management (IPM). IPM means "*careful consideration of all available plant protection methods and subsequent integration of appropriate measures that discourage the development of populations of harmful organisms and keep the use of plant protection products and*

other forms of intervention to levels that are economically and ecologically justified" (EC, 2009). This should enable a sufficient level of pest and disease control while allowing the assessment of responses of the GM plant and its comparators to biotic pressure (see also Section 5.1.14.2).

4.3.1. Herbicide regimes in the case of GM herbicide tolerant plants

Due to the diversity of current weed management strategies and the numerous additional options that GM herbicide tolerant (HT) plants give to farmers, the EFSA GMO Panel developed more guidance on which herbicide regimes can be applied in field trials used for the agronomic and phenotypic characterisation of GM plants.

4.3.1.1. Herbicide regimes

The Implementing Regulation (EU) No 503/2013 indicates that three treatments should be compared in field trials with GMHT plants. It states that "*in the case of herbicide tolerant genetically modified plants and in order to assess whether the expected agricultural practices influence the expression of the studied endpoints, three test materials shall be compared: the genetically modified plant exposed to the intended herbicide; the conventional counterpart treated with conventional herbicide management regimes; and the genetically modified plant treated with the same conventional herbicide management regimes*".

- (A) *The conventional counterpart treated with conventional herbicide regime*: All field trials, including those performed for the comparative assessment of GM plants, require maintenance according to local needs (fertilisation, tillage, pesticide treatment, etc.), in order to establish suitable agronomic conditions. This maintenance may also include herbicide treatments. A conventional herbicide regime is the programme of herbicide applications commonly used by farmers in a given location throughout the growing period of a crop with standard herbicides, i.e. not being the object of the genetic transformation in the GM plant, and may include pre-sowing, pre-emergence and post-emergence treatments. It is expected that conventional herbicide regimes differ between locations and farmers, reflecting different soil, crop history, weed, environmental and usage conditions. In the case of low weed pressure, favourable climatic conditions and/or alternative weed control options (e.g., mechanical weeding), no herbicide use could be expected at some locations.

- (B) *The GMHT plant treated with the same conventional herbicide regimes as the conventional counterpart*: At each field trial site, some or all GMHT plants are treated with the same conventional herbicides and management regimes as those applied on the conventional counterpart (see A, above). The comparison between (A) and (B) allows a direct comparison of the GM plant and its comparator under the same conventional herbicide regimes, and enables to detect unintended effects arising from the genetic transformation.

- (C) *The GMHT plant exposed to the intended herbicide*: Depending on whether the GM plant is tolerant to single or multiple herbicidal active ingredients, the use of the intended herbicide(s) can occur in various ways (as detailed below).

    - Over-the-top option: The GMHT plant is treated with the intended herbicide(s) in addition to the full conventional herbicide regimes included in the field management;

    - Full-replacement option: The GMHT plant is treated with the intended herbicide(s) only; and

    - Partial-replacement option: The GMHT plant is treated with the intended herbicide(s) in addition to a subset of conventional herbicides used in the field management.

  In order to be representative of local REs (crop management, weed pressure and climate), the third treatment (C) may differ across location sites.

In the case of GM plants that are tolerant to multiple herbicidal active substances, the intended herbicides can all be applied either in sequence or in a tank mix, or only in partial combination with the abovementioned conventional herbicide treatments.

4.3.1.2. Recommendations

For each site, applicants are requested to supply information specifying precisely which combinations of herbicides were applied to each of the test materials, including also potential non-chemical weed management interventions (e.g., pre-sowing treatments, post-emergence hoeing). A scientific rationale justifying the selection of herbicide regimes is to be provided.

If the agronomic/phenotypic and compositional data are generated in the same field trials, then applicants should also account for potential implications of herbicide regimes on the compositional analysis when selecting suitable herbicide regimes.

As indicated in Section 4.3.1.1, a range of different herbicide regimes (and other crop management practices) can be applied to GMHT plants. Plausible options for weed management for GM plants with tolerance to a single or multiple herbicidal active ingredients are listed below. These represent possible ways of herbicide application that could be chosen, taking into account the potential negative interactions between herbicides that may occur in case of a tank mix application.

Options for GM plants with tolerance to a single herbicidal active substance include:

- (i) GMHT plants treated with the intended herbicide in addition to the full conventional herbicide regimes included in the field management ("over-the-top" treatment);
- (ii) GMHT plant treated with the intended herbicide only (full replacement of the conventional herbicide regime);
- (iii) GMHT plant treated with the intended herbicide in addition to a subset of conventional herbicides used in the field management (partial replacement of the conventional herbicide regime); and
- (iv) GMHT plant with no herbicide application at a limited number of location sites.

Options for GM plants with tolerance to multiple herbicidal active substances include:

- (i) GMHT plants treated with all the intended herbicides, either in sequence or in a tank mix, in addition to conventional herbicide regimes included in the field management (full additional treatment);
- (ii) GMHT plant treated with a subset of the intended herbicides, either in sequence or in a tank mix, in addition to the full conventional herbicide regimes included in the field management (partial additional treatment);
- (iii) GMHT plant treated with all the intended herbicides but with no conventional herbicide, either in sequence or in a tank mix (full replacement of the conventional herbicide regime);
- (iv) GMHT plant treated with a subset of the intended herbicides, either in sequence or in a tank mix, in addition to a subset of the herbicides used in conventional herbicide regimes included in the field management (partial replacement of the conventional herbicide regime with partial treatment);
- (v) GMHT plant treated with a subset of the intended herbicides but with no conventional herbicide, either in sequence or in a tank mix (full replacement of the conventional herbicide regime with partial treatment); and

- (vi) GMHT plant with no herbicide application.

In some cases, applicants might consider adding additional weed control treatments, if deemed necessary as to ensure representativeness and allow a proper assessment of GMHT plants.

All the herbicide use strategies listed above are considered suitable. Applicants shall specify the selected options and substantiate the scientific rationale behind their selection.

4.3.1.3. Information on weed treatment conditions

Besides the applied weed treatments, applicants should provide detailed information on weed treatment conditions, including the application date, growth stage, rate, the inclusion of adjuvants.

4.4. Experimental design

A graphical/visual representation of the design of the field trial covering the spatial distribution and size of the blocks and the plots, the distances among them, as well as the presence of buffer rows and outside border rows should be supplied

4.4.1. Plot size and shape

Applicants should describe (see Sections 4.2.1 and 4.3) and justify the choice of the plot size by accounting for the biological characteristics and common agricultural practices of the plant species. Plots should be of a size and shape sufficient to ensure the inclusion of a representative number of plants and the collection of reliable observations/samples also in its inner part. Plot size in field trials can be affected by a number of factors (crop, agronomical practices, equipment, etc.), but the minimal acceptable surface, according to the EFSA GMO Panel, is 25 m^2.

4.4.2. Inter-plot distances and buffer/guard rows

The selected distances between the plots and the inclusion of buffer/guard rows should be described (see Section 4.3) and justified by applicants.

5. AGRONOMIC AND PHENOTYPIC ENDPOINTS

As specified in the Implementing Regulation (EU) No 503/2013, "*the applicant shall provide a comparison between the genetically modified plant and its conventional counterpart. This comparison shall enable the applicant to identify unintended effects resulting from the genetic modification and shall address also plant biology and agronomic traits, including common breeding parameters (such as yield, plant morphology, flowering time, day degrees to maturity, duration of pollen viability, response to plant pathogens and insect pests, sensitivity to abiotic stress)*".

This section provides guidance on the selection of agronomic and phenotypic endpoints and data collection, and explores their relevance for risk assessment. Some agronomic and phenotypic endpoints are mandatory, and should always be considered for the agronomic and phenotypic characterisation of GM plants, while the relevance of other endpoints is case-specific. For the case-specific endpoints applicants shall decide on a case-by-case basis whether those shall be measured, and they shall provide a scientific rationale justifying their inclusion or exclusion.

For all endpoints and performed measurements, applicants shall follow the generic sampling process defined in the standard DS 3077 Horizontal – Representative Sampling, which is based on the Theory of Sampling (TOS). Documentation of sampling representativeness is required for each relevant endpoint to guarantee valid analytical outcomes as a basis for reliable risk assessment (e.g., sampling processes need to be structurally correct to fulfil TOS' accuracy requirements). Composite sampling is a general necessity and needs a sufficient number of increments to reduce the effective sampling variance (sampling precision). Sample representativeness can only be guaranteed when both accuracy and precision demands are fulfilled (see DS-3077, 2013). Applicants need to document all applied sampling extraction and sampling preparation steps, including mass reduction to final analytical aliquot size (if required) are in compliance with these specifications. References within the standard give detailed explanation to all required aspects of sampling, of which Esbensen et al. (2012a,b) and Minkkinen et al. (2012) are particularly relevant for the GM case.

The use of internationally agreed units of measures (Thompson and Taylor, 2008) and of growth scales (BBCH, 2001) is advocated to ensure a more comprehensive interpretation of agronomic and phenotypic data and agricultural management practices.

5.1. Mandatory agronomic and phenotypic endpoints

The EFSA GMO Panel established a minimum set of agronomic and phenotypic endpoints that are deemed mandatory to allow a proper comparative analysis. Table 1 summarises these mandatory endpoints for cotton, maize, oilseed rape and soybean, and provides information on the entity to be studied, the methodology to be followed, the unit to be used, and the recommended growth stage(s) for which the necessary data should be gathered. Although this information is plant species-specific, generic requirements are given, which apply to all the listed plant species and which should be adapted to other crops/species on a case-by-case basis.

Table 1: Overview of mandatory endpoints to be considered when performing field trials for the agronomic and phenotypic characterisation of GM soybean, maize, cotton and oilseed rape (specific considerations for other crops are given in the core text of the document, and are not included in this table). The reported endpoints should be measured in all the field trials and on the whole set of test materials including the selected reference varieties. Some of the listed mandatory endpoints are described in Section 3.3.

Phenologic phase	Endpoint	Measure[a]	Unit	Soybean	Maize	Cotton	Oilseed rape
				Growth stages[b]			
Pre-sowing	See Section 3.3	-	-	00-08			
Establishment	Early stand count	M	m^{-2}	12-13	11-14	11-13	11-13
	Ground cover I	V	%	2x	16-18	23-25	23-25
	Herbicide injury	V	%	After each HT			
Reproductive	10% flowering[c]	V	Days	61			
	50% flowering[d]	V	Days	65			
	Ground cover II	V	%	65	36-38	65	65
	End of flowering[e]	V	Days	69	67	69	69
	Lodging[f]	V	#/%	89	89	89	80-89
	Final stand count	M	m^{-2}	89	87-89	89	89
Maturity	Plant height	M	cm	89	36-38	69-89	71-89
	Days to maturity	V	Days	89	87	89	89
	Seed loss[g]	M	#	89			
	Fruit count[h]	M	#	89			
	Seed per fruit[i]	M	#	99			
Harvest	Seed moisture	M	%	99			
	100/1000 seed weight	M	g	99			
	Yield	M	g^{m-2}	99			
All phases after plant emergence	Biotic interactions	V	-	11-99			
	Abiotic interaction	V	-	11-99			

(a): M: Measurement; V: visual estimation.
(b): Growth stage in accordance with BBCH scale of the corresponding crop. Indicates when the measurement/assessment of the endpoint should be conducted.
(c): In case of maize: stamens in the middle of tassel visible.
(d): In case of maize: upper and lower parts of tassel in flower.
(e): In case of maize: male flowering completed.
(f): In case of maize stalk lodged and root lodged plants should be counted separately and one value provided.
(g): In case of soybean: shattering; for maize dropped ear count; for oilseed rape: pod shattering.
(h): In case of soybean and oilseed rape: pod count; for maize ears per plant; for cotton: bolls per plant.
(i): In case of soybean and oilseed rape: number of seeds per pod; for maize number of seed per ear; for cotton: number of seed per boll.

5.1.1. Stand count

5.1.1.1. Why

Early stand count provides information on the establishment of plants, while late stand count indicates plant survival to maturity and its capacity to complete the vegetative and reproductive phases. Data generated for these endpoints will deliver helpful information on emergence and early survival and subsequent death of plants before the final stand density is achieved.

5.1.1.2. How

It is advisable to count the number of plants within a given area or distance (e.g., a row length, a quadrate of stated area). If there are large numbers of plants per plot, then the quantification should be

conducted by counting subsections of the entire plot (i.e., transects of 1m length in two or more rows, $1m^2$ quadrats, etc.).

The same sampling method and area for measuring stand count should be used at each different growth stage, in order to be in a position to make suitable comparisons.

5.1.1.3. When

Stand count is best measured at the seedling phase when establishment is completed (after BBCH10), and before any intervention in the field trial (e.g., first application of any post-emergence herbicide or thinning).

In the case of overwintering plants, an additional measurement should be conducted in spring to assess overwinter plant losses caused by factors such as pests, chilling and waterlogging.

5.1.1.4. Measurement unit

Stand counts should be quantified as the total number of plants per m^2.

5.1.1.5. Recommendations

Unless thinning is the normal practice, it is discouraged and should only be applied in cases of high heterogeneity in the initial stand counts between plots. When applied, it should be described in detail, including a description of the time of execution, the method of removal of the plants and the target plant density. The decision to perform thinning must be based only on plant density, and not on phenotypic characteristics of the plants. A correct execution of thinning is required to ensure a reliable comparative assessment of the agronomic and phenotypic endpoints to be measured later during the vegetative and reproductive phases of growth, and hence yield and other yield components.

5.1.2. Crop cover

5.1.2.1. Why

The solar radiation intercepted by crops is a primary driver of productivity. When intercepted, radiation is accumulated over the growing season; it is strongly related to the dry matter produced by crops – the total cumulative intercepted radiation accounting for differences between crop types and years.

5.1.2.2. How

Solar radiation intercepted by crops can be measured by solarimeters positioned above and below the crop canopy, but crop percent cover is a useful surrogate. Measurement by solarimeters is impractical for estimating interception for many fields, as might be needed in a large scale survey like field trials. Equally impractical are the destructive sampling needed for leaf area index.

However, percentage ground cover, estimated visually, is a useful surrogate (Steven et al., 1986). Generally, a 0.5×0.5 m quadrat is placed over the crop, and the amount of ground covered by the crop is estimated visually.

5.1.2.3. When

Ground cover should be measured twice during the season. Ideal timings to make the observations include: early vegetative growth (before stem extension) when the plant is still under low competition for resources; and the start of reproductive phase (first flower) and before maturity. These phases can be identified between BBCH16-35 and BBCH61, respectively.

5.1.2.4. Measurement unit

Visual estimation of percentage crop cover (ranging from 0% to 100%, or fraction ranging from 0 to 1).

5.1.3. Herbicide injury

5.1.3.1. Why

Herbicide injury helps to assess whether there are differences in sensitivity to the applied herbicides between the test materials. In the case of GMHT plants, sensitivity to the intended herbicide treatment(s) would also be considered through this endpoint. If observed, herbicide injury may explain differences observed in other agronomic and phenotypic endpoints (such as initial stand count).

5.1.3.1. How

Visual estimation of leaf chlorosis, necrosis or rolling, or any other symptoms that are indicative of herbicide injury.

5.1.3.2. When

The assessment of potential herbicide injury must be executed after herbicide treatments in the period when symptoms should be visible.

5.1.3.3. Measurement unit

Percent of leaf surface injured averaged over the entire plot, with short description of symptoms..

5.1.3.4. Recommendations

Besides herbicide injury, any differential injury caused by applied agrochemicals other than herbicides (e.g., pesticides or growth regulators) should also be recorded and reported.

5.1.4. Flowering

5.1.4.1. Why

Differences in flowering indicate heterogeneity in the flowering characteristics of the plants, and will provide useful information on their reproductive development.

5.1.4.2. How

Flowering is to be estimated visually at the plot level.

5.1.4.3. When

Visual observations should be carried during the reproductive phase of the plant life cycle.

5.1.4.4. Measurement unit

Days from planting when flowers occur on 10%, 50% and 90% of the plants.

5.1.5. Lodging

5.1.5.1. Why

Lodging is an important agronomic characteristic that indicates the standing power of the plant and crop. It also gives an indication on whether the plants grew and developed normally (for example, they were not unnaturally spindly) and the level of heterogeneity between the test materials. Lodging affects flowering, reduces photosynthetic capabilities of the plant and hence affects carbohydrate assimilation. It can also contribute to uneven maturity, high moisture content, loss of grain quality and

seed drop before harvest. Therefore, if lodging occurs, it may explain observed differences in other agronomic and phenotypic endpoints (such as maturation).

Lodging can also interfere with the normal development and maturation of adjacent plants occurring in the same field trial.

5.1.5.2. How

To visually estimate the proportion of lodged plants per plot and the lodging degree from the vertical.

5.1.5.3. When

Lodging should be measured at plant maturity, before harvest.

5.1.5.4. Measurement unit

Percentage of lodged plants in a plot (to nearest 10%) and mean angle (e.g., horizontal to 20°; 30° to 70°; slight angle down to 70°).

5.1.6. Plant height

5.1.6.1. Why

Information on plant height will give general information on the development and the growth of the test materials, and their heterogeneity. Plant height also serves as a general indicator of vegetative vigour and plant biomass.

5.1.6.2. How

A representative number of plants, per plot are to be measures from the soil surface to the plant tip.

5.1.6.3. When

Measurements are to be performed at flowering, at stage BBCH69.

5.1.6.4. Measurement unit

Distance (cm) from the base (soil level) of the plant to the uppermost terminal meristem.

5.1.6.5. Recommendations

The total number of sampled plants per plot is to be justified and reported, as well as the individual values measured.

5.1.7. Maturity

5.1.7.1. Why

Maturation is a general characteristic of the plant, and delivers information on its life cycle and whether it developed normally. Observed differences between the GM plant and its comparator, and the non-GM reference varieties may confirm the suitability or not of the test materials used in field trials.

5.1.7.2. How

Visual estimation of fruit/kernel/seed maturity.

5.1.7.3. When

Maturation is to be measured when the plants reach full physiological maturity (BBCH89).

5.1.7.4. Measurement unit

Days from planting when mature fruits/kernel/seeds occur on 90% of the plants.

5.1.7.5. Recommendations

If it is not possible to reach physiological maturity of the crop (for instance, owing to the use of desiccants, swathing), then applicants should characterise the level of maturity (BBCH) at the time of the intervention.

5.1.8. Seed loss[13]

5.1.8.1. Why

Seed loss has implications for two aspects of characterisation[14]. First, seed loss may be a feature of the plant or induced by local conditions. Either way it causes an underestimate of plant production and associated errors in fruit number, seed number, seed per fruit and mean seed weight. Second, it provides an indicator of the potential of the plant and its traits to persist in the soil. All propagules should be considered: for example, loss by tubers and true seed should be considered for a plant such as potato,

5.1.8.2. How

To count the number of dropped ears/cobs/seed per plot, or shattered pods per specific area of plot or the number of tubers not harvested. In some instances, 'seed rain' trays may be positioned beneath the plots to collect falling or shattered seed (e.g., oilseed rape).

5.1.8.3. When

Observation should ideally be made at plant maturity and harvest. Estimates should be made before and after harvest because the mechanism of loss during these periods may differ. For example, variation in maturity date may affect loss before harvest, while the propensity of pods/cobs to shatter during mechanical agitation will only be picked up by a sample during or after harvest.

5.1.8.4. Measurement unit

Total number of mature dropped cobs or seeds per plot, or total number of shattered pods per specific lengths or areas of the plot. Measurements should be converted to per unit area (m^2).

5.1.8.5. Recommendations

For some crops, as oilseed rape or potato, the measurement of seed loss can be challenging due to the natural seed loss during inspection or incomplete tubers removal during harvest. This may lead to an overestimation of losses and, in those cases, alternative approaches can be proposed by applicants. For example, in case of oilseed rape, seed loss can be measured trough the sampling of pods at different intervals up the main stem, and then assess those for losses (so no plants are removed) or alternatively, leaving trays on the soil to collect the dropped seeds just before harvest.

[13] The term varies between plant species. It refers to shattering in the case of soybean and oilseed rape, and dropped ears in the case of maize.

[14] Harvested parts that are not seed, such as tubers, should be treated as seed for this purpose.

5.1.9. Fruit count[15]

5.1.9.1. Why

The number of fruits per plant is an important yield component that provides an indication of the plant fertility.

5.1.9.2. How

To count the number of fruits from a representative sample of plants in the plot.

5.1.9.3. When

The number fruits per plant is to be assessed at maturity (BBCH89).

5.1.9.4. Measurement unit

Number of fruits per plant, scaled up with plant stand count to number of fruits per m^2.

5.1.9.5. Recommendations

The total number of sampled plants per plot is to be justified and reported, as well as the individual values measured.

5.1.10. Seeds per fruit[16]

5.1.10.1. Why

Seeds per fruit is important yield component.

5.1.10.2. How

To count the number of seeds in pods/ears/bolls calculated from a representative sample of fruits in the plot.

5.1.10.3. When

Seed count is to be assessed after harvest on a representative sample of pods/ears/bolls harvested at maturity (BBCH89).

5.1.10.4. Measurement unit

Number of seeds per pods/ears/bolls.

5.1.10.5. Recommendations

Seeds per fruit shall be directly measured and not calculated using relevant data gathered for other endpoints. The sampling strategy shall be specified, including the fruit selection criteria (if any; primary/secondary branch; top bottom or middle of the plant; etc.).

[15] Number of pods per plant in the case of soybean and oilseed rape; number of ears per plant in the case of maize; and number of bolls per plant for cotton. (In the case of potato, this character can also be reported as the number of true fruit produced seed)

[16] Number of seeds per pod for soybean and oilseed rape; number of seed per ears for maize; and number of seed per boll for cotton. (In the case of potato, this character can also be reported as the number of tubers).

5.1.11. Seed moisture at harvest

5.1.11.1. Why

The dry matter content of harvested products varies with crop variety, soil and weather. National yield statistics are derived by converting on-farm bulk yield (with a specified moisture content measured for example in a combine harvester) to standard moisture content for comparison between regions and years. In field trials, the yield should be expressable as dry matter using a measured value of percent moisture content.

5.1.11.2. How

"Weight loss on drying" based on the entire harvest or representative sample. Specification of the methodology applied should be provided. Alternative methodology such as moisture meter can be applied on a representative sample of the entire harvest.

5.1.11.3. When

Immediately after harvest (BBCH99) and before any processing (e.g., natural or artificial air drying, storage).

5.1.11.4. Measurement unit

Water as a percentage of fresh harvested weight. The remainder is the dry matter.

5.1.12. Seed weight

5.1.12.1. Why

Seed weight is another important yield component, often strongly related to fruit and seed number. If a crop is source-limited (i.e., there are more seeds than the available plant assimilate can fill), then a trade-off commonly exists between seed number and seed weight – generally, the lesser the number of seeds per plant, the larger the seeds. However, if a crop is sink-limited (i.e., there are fewer seed sites than can be filled by the available assimilate), then seed weight is more stable. The relation between seed weight and number is an important attribute which is often more discriminating between varieties than is total seed weight.

5.1.12.2. How

To weigh and count a representative sample of seeds.

5.1.12.3. When

Seed weight is to be assessed after harvest on a representative sample of pods/ears/bolls harvested at the plant growth stage BBCH99.

5.1.12.4. Measurement unit

Mass (g) of 100 or 1000 seeds corrected for moisture content (seed index: g/100 or 1000 seeds).

5.1.13. Yield

5.1.13.1. Why

Yield is a common parameter based on which the overall quality of the field trial and the agronomic performance of the test materials can be determined. Moreover, yield is related to several other agronomic and phenotypic endpoints. Therefore, differences observed in yield may be attributed to differences observed in other endpoints.

990 5.1.13.2. How

991 To weigh the harvested plant material for a representative area or the plot.

992 5.1.13.3. When

993 Yield is to be determined at harvest (BBCH99).

994 5.1.13.4. Measurement unit

995 Mass (g or kg) of the harvest material per sampled plot surface (m^2) corrected for moisture content.

996 5.1.13.5. Recommendations

997 Buffer/guard rows should be excluded when quantifying yield of the sampled plant material. In field
998 trials, the plant material should be as free as possible from soil (e.g., for potato) and should be given in
999 the standard form for the species. It should be clearly stated, for example, whether the yield of cereals
1000 refers to pure grain, grain and cob, or grain plus attached husk.

1001 ***5.1.14. Plant responses to biotic stressors***

1002 5.1.14.1. Why

1003 Crop varieties generally differ intrinsically in their susceptibility and response to biotic stresses
1004 independently of any difference due to genetic modification. Therefore, it is necessary to estimate
1005 whether the plants differ in susceptibility to the activity of specific arthropods and/or plant diseases.
1006 Moreover, an early phytosanitary assessment may help confirming the health condition of the starting
1007 material (seed/tubers), and may detect the presence of pathogens, pests in the soil and their pattern of
1008 distribution throughout the testing field.

1009 5.1.14.2. How

1010 An excessive use of plant protection products may result in a "clean" experimental field, but that may
1011 impair a thorough evaluation of pest-plant interactions. Similarly, the lack of pest management
1012 measures in the experimental field may lead to the presence of excessive stress on plants, which does
1013 not reflect normal agronomic practices. It is therefore advisable to follow local pest management
1014 practices and consider the principles of Good Agricultural Practices and IPM, in order to keep the
1015 treatments with pesticides to the minimum required to contain the level of disease/infestation below an
1016 acceptable level of disease.

1017 *Arthropods*: The necessary survey or monitoring programs in field trials will depend on the biology of
1018 the herbivorous arthropods and plant phenology, and therefore will have to be set on a case-by-case
1019 basis.

1020 Since it will be impossible to consider all potentially exposed arthropods, relevant pest species to be
1021 surveyed should be identified at an early stage of the field experiments considering the expected/actual
1022 presence of plant pests locally. Problem formulation will thereby serve as a relevant tool to choose
1023 suitable arthropod species. The following criteria should be considered when selection relevant
1024 arthropod species to be surveyed in field trials:

1025 - *Feeding mode*: A first requirement for the selection of relevant species is that it feeds on the
1026 plant (such as sap sucker, foliage chewer, miner/stem borer, gall former, root feeder, flower
1027 feeder, seed/grain feeder). Species using the plants as shelter and reproduction sites are not
1028 necessarily directly feeding on the plant, and therefore not optimal to assess plant-arthropod
1029 stressor interactions. Applicants are asked to consider the mode of feeding/feeding guilds of
1030 the arthropods when selecting the relevant species. Ideally, the different abovementioned
1031 categories of feeding mode should be covered, ranging from foliage chewers to sap suckers, to
1032 root feeders, where relevant.

- *Biological and economic relevance*: The biological and economic relevance of the species is to be considered too, as the abundance and frequency of the species may lead to populations exceeding economic injury levels. Measuring biological features of these species will enable to foresee unintended changes with potentially relevant economic consequences.

Plant diseases: Disease incidences or severity per plot is to be estimated visually. All relevant disease should be considered during the survey.

Appearance of symptoms in the field depends on the interaction between the host, the pathogen and the environment, and may vary from year to year. It is therefore necessary to prioritise the observations and to make an appropriate selection of plant diseases requiring investigation based on prior experiences. The major diseases known to pose a threat to crops in field trials should be considered by applicants. The species that will be surveyed should cover diverse categories of etiological agents (viruses, bacteria, fungi, etc.) and/or mechanisms of pathogenesis (biotrophic, necrotrophic). Moreover, for non-systemic diseases, the different organs (leaves, stem/stalk/trunk, roots, flowers, fruit) or plant parts (e.g., aerial part vs. subterranean parts) which can be affected by the disease should be considered.

Diseases known to have a severe biological/economical impact should be prioritised in the selection process. Yet, attention should also be paid to relatively minor or 'emerging' diseases (Anderson et al., 2004; Giraud et al., 2010), and to important diseases in the RE.

The applicant must provide a scientific rationale justifying the selection of diseases to be studied in field trials. Expert knowledge and supporting scientific literature on the various crop diseases, such as the information available in the CABI Crop Protection Compendium (CABI, 2014) should be accounted for when selecting relevant species used. Likewise, applicants are asked to provide specifications on the survey methodology used.

5.1.14.3. When

Arthropods: Overall, more than one sampling occasion is necessary to infer patterns of herbivore infestation. The most relevant information necessary to determine the timing of survay is the knowledge of the biological cycle of the species and the pest-plant interactions for each insect instar. Voltinism is another very relevant feature that should determine (on the basis of the local agro-environmental conditions) the optimal timing for survey. In addition, plants may be more/less susceptible to herbivory during plant development (e.g., some maize varieties are naturally resistant to corn borers in their early growth phases). In addition, repeated samplings will also enable to estimate when relevant the presence/incidence of natural enemies of the chosen herbivore species.

Plant diseases: The timing for scoring disease level(s) and the appropriate number of observations to be conduct depend on each plant-pathogen interaction, and should be defined on a case-by-case basis, accounting for the disease progress in each testing location. Disease survey should be conducted according to plant phenology and pathogen epidemiology to ensure the generation of appropriate data on plant response to infestation.

For plants which product is an organ growing partially or totally under the ground (e.g., beet roots, potato tubers) an additional assessment of disease must be done on a representative, cleaned sample of the subterranean organs immediately after harvest.

Applicants are asked to justify their sampling/observation strategy, and to report the plant growth stages at which disease scorings were performed.

1075 5.1.14.4. Measurement unit

1076 *Arthropods*: The presence and incidence of arthropod pests can be estimated in different ways. While a
1077 specific discussion for each species is beyond the scope of the present document, it is advisable to use
1078 one of the two following approaches:

1079 - Measure of abundance of herbivores when and where relevant on plants via visual
1080 observations (directly or after removing plant parts or the organisms living on it, e.g. via
1081 sucking devices or knock-down systems). This way of comparing abundances of arthropods
1082 needs to be discussed in the light of the possible relevance of the data collected (e.g., compare
1083 the different population levels detected with existing action thresholds for integrated pest
1084 management (IPM) principles).

1085 In more cases, damage classes, percent of damaged leaves, fruits, etc. or infestation level over the
1086 plots (% of damaged plants) can be used as an estimate plant response to biotic stressors It is
1087 recommended to use these endpoints where a relevant scientific literature and expertise already exists.

1088 5.1.14.5. Recommendations

1089 Biotic interactions in the field trials are regularly monitored during the entire crop cycle, in order to
1090 ensure that appropriate crop protection measures are taken timely in order to avoid the build up of
1091 populations that could overexploit the crop. These measures should effectively reduce the pest
1092 pressure to comparable levels in the different selected sites for the field trial, thereby reducing the
1093 incidence of targeted arthropods and the potential plant damage caused by them.

1094 Any anomaly observed in the response of plants towards biotic stressors must be reported (Garcia-
1095 Alonso et al., 2014).

1096 ***5.1.15. Plant responses to abiotic stress***

1097 5.1.15.1. Why

1098 Crop plants generally experience abiotic stresses of one form or another at various times during their
1099 growth. Stresses include low and high temperature, salinity, hail, dryness, nutrient deficiency and
1100 waterlogging. If GM plants have traits specifically intended to reduce the susceptibility of the plant to
1101 a defined stressor, then substantial field trials would be required as part of a much broader risk
1102 assessment (see Section 5.2). These trials might be carried out across sites over a natural gradient of
1103 the stressor or through local manipulation of the environment. For GM plants that are not intended to
1104 have traits making them less susceptible to stresses, then measurements should be made as part of
1105 standard agronomic and phenotypic assessment, in order to look for any unintended, differential
1106 response between the GM plant and its comparators to the abiotic environment.

1107 5.1.15.2. How

1108 In addition to the mandatory endpoints that can affected by the occurrence of specific biotic stressors
1109 (such as yield and other yield components), applicants should record visually as standard practice, any
1110 damage to leaves, stems and reproductive structures related to abiotic stressors. The range and type of
1111 possible visible damage are very wide and varied, but might include leaf scorch and discolouring, leaf
1112 mottling, patterning due to nutrient deficiency or effects of herbicide, weakness or imbalance of stem
1113 causing lodging or a fractured canopy, failure of flowers to open fully and damage to reproductive
1114 organs (stamens and stigmas) preventing effective pollination. Measurements at three scales are
1115 useful, each by rapid visual assessment: they are the average proportion of the field plot (i.e., of all
1116 plants) that is affected; the proportion of organs on damaged plants that is affected; and the proportion
1117 of the damaged organ that is affected. Damage to crops in field trials occurs as part of the general
1118 responses of crops to environment and should be expected. However, the measurements should be able
1119 to assess whether the GM plant differs from the comparators in these visible symptoms of stress.

5.1.15.3. When

Assessments should be made throughout growth, typically at the same time as morphological measurements – e.g., at the vegetative stage before flowering, during full flower and when the crops are maturing before harvest. For crops that overwinter, a second vegetative score should be made after winter to note wind-chill or ice-damage.

5.1.15.4. Measurement unit

Units will vary with the type of damage. Typical examples at the three scales indicated above are: % of plot affected; % plant organs affected on damaged plants; and % of single organ affected (e.g., 10% of a plot is affected by leaf scorch; on damaged plants, 20% of leaves are affected; on damaged leaves, 50% of the surface is browned due to drought). Alternatively, damage can be defined semi-quantitatively by, for example, high, medium and low, but these categories should be defined.

5.1.15.5. Recommendations

Applicants should record damage due to abiotic stress routinely and typically at the same time as other phenotypic assessments. Given the wide range and type of damage, the method of assessment and descriptors or units of damage are not prescribed but should be justified by the applicant. Damage is assessed at plot, plant and organ scales and the likely causes defined. The aim is to assess whether the GM plant differs from its comparators in response to abiotic stress experienced during the trial.

5.2. Additional endpoints to be considered on a case-by-case basis

Besides the mandatory agronomic and phenotypic endpoints mentioned in Section 5.1, applicants should consider the inclusion of other additional relevant endpoints on a case-by-case basis using the outcomes of problem formulation. Problem formulation enables one to identify missing information and scientific uncertainties that may limit the assessment, and thus helps the selection of relevant endpoints for which informative data should be gathered. This process will mainly depend on the type of the genetic modification. For example, it can be important to assess the efficacy of male sterility and restorer genes (e.g., barstar/barnase) in the case of sterile GM plants. Likewise, GM plants having traits specifically intended to reduce susceptibility of the plant to a defined (a)biotic stressor may necessitate the consideration of additional endpoints (Sections 5.1.14 and 5.1.15).

If the objective of the genetic modification is to alter the composition of the plant, then it is recommended to measure additional agronomic and phenotypic endpoints associated with this intended compositional change in field trials. For example, pest susceptibility could be a relevant endpoint for a GM plant exhibiting an altered level of an anti-nutrient or toxicant, or dormancy and survival characteristics in the case of a GM plant with modified seed/tuber composition (see Section 6.4).

There is also the possibility that the outcomes of the compositional analysis indicate the necessity to perform additional field trials, in order to measure additional agronomic and phenotypic endpoints which are known to be correlated with the compositional endpoints showing differences (see Section 6.4).

In the frame of the persistence and invasiveness assessment, the decision tree presented in Section 7.1 (and Appendix A) should be used to identify under what circumstances the set of mandatory endpoints shall be extended with additional ones. For the case-specific endpoints applicants shall decide on a case-by-case basis whether those shall be measured. A scientific rationale justifying their inclusion or exclusion for the agronomic and phenotypic characterisation of the specific GM plant under assessment shall be provided.

Irrespective of the outcomes of the problem formulation, applicants may decide to evaluate other additional plant characteristics (e.g., pollen characteristics, flower (corolla) colour, number of flowers per plant, or other fruit characteristics). In the latter case, applicants should supply the gathered data,

as those may add to the weight of evidence approach followed for the identification of unintended changes in GM plants.

In the below sections, specific recommendations pertaining to the assessment of pollen and seed characteristics are given.

5.2.1. Pollen characteristics

5.2.1.1. Why

In specific cases, the genetic modification is designed to target specific pollen characteristics. Depending on the plant species, trait and the intended uses, altered pollen characteristics may affect pollen viability and thus exposure. Therefore, further characterisation of pollen characteristics is required for those GM plants for which increased pollen viability is hypothetised.

If significant differences in seed set (see Section 5.1.10) or yield (see Section 5.1.13) are observed between the GM plant and its comparator based on the data gathered in field trials, which cannot be related to observed changes in correlated endpoints (such as stand count), then the assessment of pollen characteristics is required, even for GM plants for which the genetic modification was not designed to target specific pollen characteristics.

Applicants can also evaluate pollen characteristics as part of the assessment of unintended changes in the GM plant, irrespective of the outcomes of the problem formulation.

5.2.1.2. How

Based on the source of the problems and the nature of the hypothesis under test, applicants should assess pollen viability and morphology *in vitro*.

Pollen characteristics should be investigated on a representative sample of pollen collected from the GM plant, its comparator and at least three non-GM reference varieties in at least three sites (representing contrasted REs) of the field trials performed for the agronomic and phenotypic characterisation of the GM plant.

Owing to the insensitivity of some staining techniques to measure pollen viability, it is advisable to directly measure pollen germination capacity. Staining allows a preliminary test assessing pollen grain maturity, but does not directly measure germination capacity (or pollen viability). Staining therefore cannot be reliably used as a standalone to estimate pollen viability, but can supplement tests measuring pollen germination capacity.

5.2.1.3. When

In order to maximise pollen viability and germination, pollen should be sampled at the time of anthesis.

5.2.1.4. Measurement unit

Percentage of viable (germinated) pollen grains (with suitable tube length) over a representative number of pollen grains in a given period of time. Pollen morphology expressed as diameter of a representative sample of pollen grains can supplement the previous measurements.

5.2.1.5. Recommendations

Identical sampling conditions shall be followed for all test materials; any deviation from the sampling protocol should be reported.

5.2.2. Seed characteristics

5.2.2.1. Why

In specific cases, the genetic modification is designed to target specific seed characteristics (e.g., altered oil content). Depending on the plant species, trait and the intended uses, altered seed characteristics may affect the dormancy potential and thus the persistence and invasiveness of the GM plant (Section 7.1). Therefore, further characterisation of seed characteristics is required for those GM plants for which increased seed dormancy is hypothetised (see decision tree presented in Section 7.1.1 and Appendix A).

Applicants can also evaluate seed characteristics as part of the assessment of unintended changes in the GM plant, irrespective of the outcomes of the problem formulation.

5.2.2.2. How

Based on the source of the problems and the nature of the hypothesis under test, applicants should assess the seed dormancy potential under controlled conditions and seed survival under field conditions as part of the agronomic and phenotypic characterisation (see Step 4 in Appendix A for further details).

- Seed dormancy potential tested under controlled conditions:
    - Seed germination testing in Petri dishes under optimal conditions, in order to establish the fraction of seeds with primary dormancy;
    - Temperature-dependent germination testing (the lower the temperature requirement, the more seeds will immediately germinate before secondary dormancy can develop); and
    - Seed germination testing of seeds with induced secondary dormancy (to examine the fraction of seeds that can acquire secondary dormancy (dark dormancy)) (Schatzki et al., 2013).
- Seed survival tested under field conditions:
    - Seed germination testing of seeds buried in the field at 10-20 cm depth for one year, or other methods to estimate seed survival in the soil (Schatzki et al., 2013); and
    - Survey of volunteers in the subsequent years in the field trials.

Seed characteristics should be investigated on a representative sample of seeds collected from the GM plant, its comparator and at least three non-GM reference varieties in at least three sites (representing contrasted REs) of the field trials performed for the agronomic and phenotypic characterisation of the GM plant.

5.2.2.3. When

F_2 seed should be sampled when the plants reach full physiological maturity (BBCH89).

5.2.2.4. Measurement unit

The percentage of (normal vs. abnormal) germinated seed, dead seed, viable non-dormant seed and dormant seed over a representative number of seeds, and the number of volunteer plants per m^2.

5.2.2.5. Recommendations

Identical sampling conditions shall be followed for all test materials; any deviation from the sampling protocol should be reported.

6. DATA ANALYSIS

In this section, recommendations are provided for the analysis of agronomic and phenotypic data. Moreover, clarifications are given for specific aspects of the data analysis requirements, already addressed in its existing guidance documents (EFSA 2010b, 2011a)

6.1. Data submission

The Implementing Regulation (EU) No 503/2013 states that "*raw data should be provided in all cases and be in a suitable electronic format*".

Data for the agronomic and phenotypic endpoints are recorded either as single values per plot or as multiple values per plot (e.g., plant height or seed count per plant). In the latter case, applicants typically calculate the mean and standard error for each plot, but do not supply the single values recorded. Single values are useful to assess the heterogeneity and appropriateness of the test materials. Therefore, complete raw datasets – to the scale of individual plant measurements – should be provided in a suitable electronic format for all sites and replicates included in the field trial.

The raw data should be structured in a format suitable for a direct analysis with the EFSA GMO Statistical software[17][18]. If a different statistical software is used, then the complete code as well as all the specific files required for the correct execution of the analysis should be provided.

In designing the field trials, the likelihood that some sites and/or replicates may not produce data of appropriate quality for a range of reasons, both within and outside the control of the applicant, exists.

In line with the requirements outlined in the Implementing Regulation (EU) No 503/2013 and EFSA guidelines (2010a, 2010b), the implemented field trials should ensure that a representative portion of the environmental variability in the field trial design is captured. Therefore, applicants shall plan a sufficient number of sites and replicates to ensure fulfilment of the representativeness of field trials. In any case applicants shall provide data collected at least from eight sites replicated four times.

If some sites and/or replicates, initially planned in the field trials, are disregarded applicants shall explicitly describe the reasons behind the (complete or partial) exclusion. The EFSA GMO Panel considers this acceptable provided that there is an explicit and clear rationale and that the minimum data requirements are fulfilled even after the data removal.

If field trials and/or replicates were discarded during the growing season, then the part of the dataset gathered before the unexpected event(s) can be retained. In all cases, applicants are requested to submit all available data for the discarded sites and/or replicates (including those of poor quality or the ones not harvested). These data may be informative for the assessment of agronomic and phenotypic characteristics of the GM plant.

6.2. Statistical analysis

Specific requirements for the analysis of field trials data, including agronomic and phenotypic data, are provided in the EFSA guidelines on statistical considerations for the safety evaluation of GMOs (EFSA, 2010b) and the guidance document on risk assessment of food and feed from GM plants (EFSA, 2011a).

As discussed in van der Voet et al. (2011), there may be situations in which the test of equivalence cannot be carried out because of zero or very small variance between the non-GM reference varieties. The practical experience gained during the assessment of recent GM plant applications, identified several of these situations. This may occur by chance or because of some unexpected event, such as

[17] http://www.efsa.europa.eu/en/gmo/gmoanalysissoftware.htm

[18] The annex within the supporting documentation of the EFSA GMO statistical software provides detailed description of the requested information.

specific climatic conditions minimising variance between the non-GM reference varieties. If this occurs for several uncorrelated endpoints, it may indicate that the non-GM reference varieties and/or REs did not sufficiently capture the range of natural variation. Should this occur, applicants may need to submit additional data to ensure sufficient representativeness (see Section 2.1.4).

The practical experience gained during the assessment of recent GM plant applications also indicates the sporadic occurrence of an outcome type not foreseen in the guidance documents on the risk assessment of food and feed from GM plants (EFSA, 2011a). This corresponds to the case of no difference between the GM and its comparator, and of no equivalence with the non-GM reference varieties. This might occur by chance for a single or few endpoint(s) if a very large number of endpoints is tested. However, if several uncorrelated endpoints show this outcome, then it may indicate that the GM plant and its comparator derive from varieties with characteristics not present in the non-GM reference varieties and, consequently, the test material may not be appropriately chosen.

6.3. Analysis of Genotype*Environment interactions

As both intended and unintended effects of the genetic modification may become apparent under specific environmental conditions, special attention should be paid to genotype*environment (G*E) interactions, as recommended in Section 2.1 of EFSA (2010b). Therefore, the EFSA GMO Panel requires that the analysis of the agronomic and phenotypic dataset shall test for possible G*E interactions. Such an approach is recommended for all types of data analysed in the comparative assessment (i.e., compositional analysis).

The Implementing Regulation (EU) No 503/2013 states that "*in the case of significant difference and/or lack of equivalence for any particular endpoint, further statistical analysis shall be carried out to assess whether there are interactions between any of the test materials and site, possibly using a simple standard ANOVA approach*".

If significant G*E interactions are identified, then a per-site analysis should be carried out to give input into further assessment of the experimental data. Applicants are requested to assess whether site differences are related to specific characteristics of REs (e.g., whether differences between the GM plant and its conventional counterpart only occur in a specific soil type across sites), and to analyse their implications in terms of risk assessment.

6.4. Correlated endpoints

The endpoints used to support agronomic and phenotypic characterisation of GM plants cover plant development along the whole growing season, from seeding to harvesting, characterising the development of the GM plant at different stages. Many of the endpoints are correlated (e.g., initial plant stand and final plant stand) and thus not independent. Therefore, significant differences observed in specific endpoints may be related. For example, if yield differences are observed, applicants are requested to provide a yield component analysis, in order to analyse the relationship between yield components such as the number of pods vs. final plant density, number of seeds per area unit vs. plant density, and seed size vs. number of seeds per unit area.

In specific cases, agronomic and phenotypic data can help to explain the differences observed in specific compositional endpoints between the GM plant and its comparator. As unexpected agronomic and phenotypic differences may only be detected in the course of a field trial, real-time monitoring and corresponding preliminary statistical analysis of agronomic and phenotypic endpoints (see Section 5) is recommended. Depending on the outcome of the real-time approach, the spectrum of compositional endpoints to be assessed in the compositional analysis should be updated so that potential interrelationships can be considered. A comprehensive assessment of the entirety of agronomic and phenotypic data may only be feasible after completion of the field trial. Therefore, representative samples should be retained enabling supplementary compositional analyses on a case-by-case basis.

7. RELEVANCE OF AGRONOMIC AND PHENOTYPIC DATA FOR ENVIRONMENTAL RISK ASSESSMENT

This section clarifies to which extent the agronomic and phenotypic dataset can inform the environmental risk assessment of GM plants (EFSA, 2010a). The EFSA GMO Panel considers that the data generated for the agronomic and phenotypic characterisation of GM plants are informative for the persistence and invasiveness assessment of GM plants (see Section 7.1). As for the other areas of environmental concern outlined in Directive 2001/18/EC, the EFSA Panel is of the opinion that the agronomic and phenotypic data have limited value to assess interactions with target and non-target organisms:

- *Interaction of the GM plant with target organisms*: The monitoring of unexpected field damage caused by the target pest is part of the insect resistance monitoring programmes for *Bt*-plants implemented by applicants. However, owing to their small plot size, field trials for the agronomic and phenotypic characterisation of *Bt*-plants cannot be reliably used as a standalone to capture early warning signs indicating increased tolerance in the field.

- *Interaction of the GM plant with non-target organisms*: Field trials for the agronomic and phenotypic characterisation can only provide supplemental information to the results of other non-target effect studies by giving indications on the likelihood of unintended effects (e.g., through herbivore and disease associations with the GM plant)[19]. Owing to their small plot size, such field trials cannot be reliably used as a standalone to assess possible unanticipated unintended effects on non-target organisms, as they would not be suitable for trapping and counting non-target organisms, and may not necessarily have sufficient statistical power to detect unintended effects on non-target organisms.

7.1. Persistence and invasiveness of the GM plant

The risks that GM plants may pose to human and animal health and the environment are assessed as part of pre-market risk assessments. One source of potential environmental harm arising from the cultivation or import of GM plants is that they will persist in agricultural land or become invasive of non-agricultural land. To assess whether the persistence (weediness) or invasiveness potential of the GM plant is altered relative to its comparator, its general agronomic and phenotypic characteristics are measured in multi-location agronomic field trials representative of the different environments where the GM plant may be grown, and compared with those of its comparator and non-GM reference varieties[20].

From the experience gained during the evaluation of GM plant applications for cultivation and import, the EFSA GMO Panel identified the need to better define the requirements for the agronomic and phenotypic data essential for a comprehensive assessment of persistence and invasiveness. Greater clarity is needed on which agronomic and phenotypic characteristics would deliver relevant data to support such an assessment. Moreover, greater attention should be paid on how relevant agronomic and phenotypic endpoints should be selected, as this has not been specified in the EFSA (2010a) guidelines on the environmental risk assessment of GM plants.

The EFSA GMO Panel developed a decision tree, in order to inform the problem formulation for the persistence and invasiveness assessment of GM plants performed by applicants when they prepare

[19] To be in a position to assess unanticipated unintended effects, the EFSA GMO Panel requires data on interactions of the GM plant and its comparator with non-target organisms for GM plant applications for cultivation (EFSA, 2010a). Field or laboratory studies in which non-target organisms are exposed to event-specific plant material are considered suitable means to gather such data.

[20] If statistical differences in characteristics associated with persistence or invasiveness are found between the GM plant and its comparator, then the differences are evaluated further, as they may have the potential to lead to harm. Should a statistically significant difference be within the range of the non-GM reference varieties or what is known for the species, then it can be considered not biologically relevant. However, if the value of the agronomic and phenotypic characteristic in the GM plant falls outside the range of the non-GM reference varieties or the plant generally, then further evaluation of the potential for altered persistence and invasiveness would be required.

their GM plant applications. The purpose of the decision tree is to identify under what circumstances the standard agronomic and phenotypic dataset typically supplied by applicants is sufficient or not as a starting point for the assessment of persistence and invasiveness. It also indicates when and how the dataset shall be extended and/or complemented with additional measurements and/or experiments. Whether additional data or experiments are required will depend upon the plant species, trait(s), the intended use, the REs, and the conclusions drawn from the successive steps of the decision tree (see below).

The decision tree does not replace the staged approach outlined in the EFSA (2010a) guidelines to assess the persistence and invasiveness of GM plants, but complements it. It ensures that informative and case-specific data are generated in a proactive manner, in order to support the persistence and invasiveness assessment of GM plants, and that the data requirements remain proportionate to the potential risk.

This section focuses on the persistence and invasiveness of the GM plant itself, and does not consider the persistence and invasiveness of cross-compatible wild/weedy relatives that may acquire transgenes through vertical gene flow.

7.1.1. Decision tree

The EFSA GMO Panel suggests three levels of agronomic and phenotypic characterisation to inform the problem formulation for the persistence and invasiveness assessment of GM plants:

- *Level 1*: The first level of agronomic and phenotypic characterisation is based on the standard agronomic and phenotypic dataset (SD in Figure 3) that is typically supplied by applicants. This dataset consists of field trials conducted to assess the general agronomic and phenotypic characteristics of the GM plant, and studies performed under controlled conditions to assess seed and pollen characteristics of the GM plant;

- *Level 2*: The second level of agronomic and phenotypic characterisation also relies on the standard agronomic and phenotypic dataset, but is extended with additional endpoints (SD+ in Figure 3). There is also the possibility to expand the assessment of seed and pollen characteristics under controlled conditions, so as to consider additional characteristics (such as seed dormancy). These additional endpoints should deliver relevant data for the persistence and invasiveness assessment;

- *Level 3*: The last level of agronomic and phenotypic characterisation consists of the standard agronomic and phenotypic dataset extended with additional endpoints and complemented with additional experiments (SD++ in Figure 3).

The decision tree is structured around five successive steps (steps 1 to 5 in Figure 3; further details are provided in Appendix A), which are designed to determine whether, and if so, under what circumstances the standard dataset shall be extended and/or complemented with additional measurements and/or experiments. The final risk category for the GM plant strongly influences the type of agronomic and phenotypic data that is needed to usefully inform the assessment of its persistence and invasiveness according to the staged approach outlined in EFSA (2010a).

- Step 1 outlines the questions to be answered to estimate the potential persistence and invasiveness of the parental species;

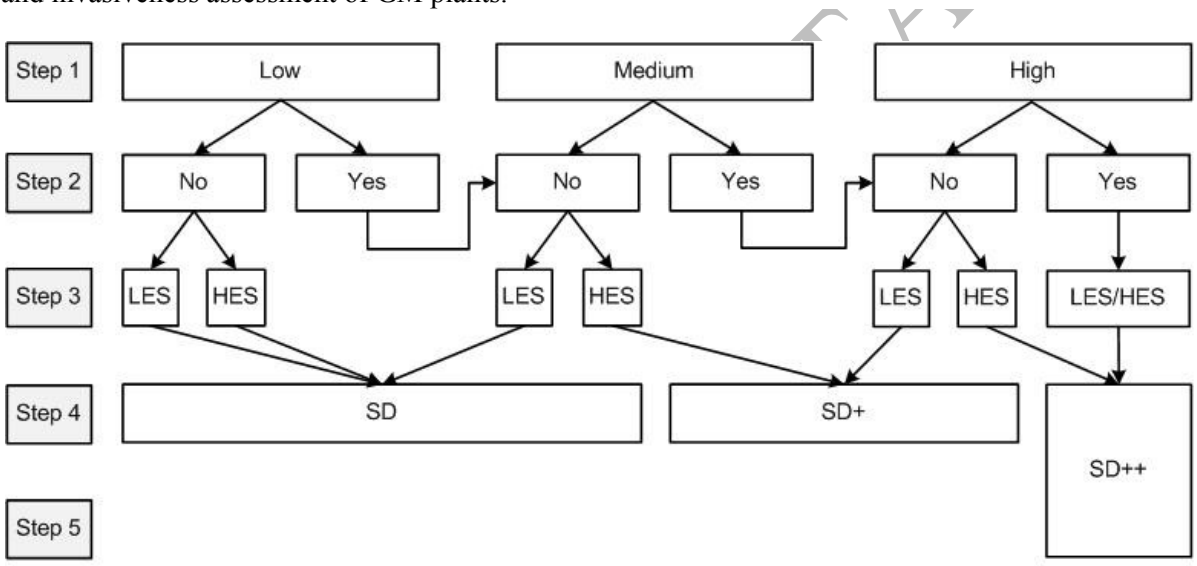
- Step 2 focuses on the critical factors limiting the persistence and invasiveness of plants, and enables to consider whether the transgene may lead to a change in persistence or invasiveness;

- Step 3 accounts for the level of exposure to the GM plant in the EU based on its intended uses (cultivation vs. import);

- Step 4 involves the selection of additional agronomic and phenotypic characteristics that are relevant for the persistence and invasiveness assessment. These characteristics are selected on a case-by-case basis, using the outcomes of the former steps of the decision tree; and

- Step 5 concerns the selection of additional experiments to complement field trials.

The outcomes from steps 1-2-3 will determine whether the plant species, trait(s) and the intended use combination under assessment has the potential to alter the persistence and invasiveness of the GM plant, and whether additional risk hypotheses should be tested via new agronomic and phenotypic measurements or experiments. Applicants are invited to use this information to explore whether the standard dataset (level 1) is sufficient for assessment at the Low end of the risk categorisation, and whether potential additional information on agronomic and phenotypic characteristics (level 2) or experiments (level 3) are needed for the Medium and High risk categories. The standard agronomic and phenotypic dataset should be generated for all GM plants. The SD+ should be generated where step 3 gives Medium-No-HES or High-No-LES. The High-No-HES and High-Yes-LES/HES category at step 3 signal that additional experiments are needed (SD++) (see Figure 3). In steps 4 and 5, additional relevant agronomic and phenotypic characteristics and experiments are to be selected or designed on a case-by-case basis, respectively, as part of the problem formulation for the persistence and invasiveness assessment of GM plants.



Legend:
Step 1 – Estimation of the persistence/invasiveness potential of the parental plant (low, medium or high category)
Step 2 – Assessment of whether the novel trait may increase persistence and invasiveness (yes or no)
Step 3 – Consideration of the level of exposure (LES or HES)
Step 4 – Selection of additional agronomic and phenotypic characteristics
Step 5 – Selection of additional experiments

Abbreviations:
LES: Low-exposure scenario (import)
HES: High-exposure scenario (cultivation)
SD: Standard AGRO/PHENO dataset
SD+: Standard AGRO/PHENO dataset with additional measurement endpoints
SD++: Standard AGRO/PHENO dataset with additional measurement endpoints and complemented with additional experiments

Figure 3: Decision tree to determine the level of agronomic and phenotypic data needed to support the persistence and invasiveness assessment of GM plants.

7.1.2. Data requirements

Based on the source of the problems and the nature of the hypothesis under test, applicants should consider the below endpoints, and provide a scientific rationale justifying their selection, as part of SD+.

- Seed dormancy potential tested under controlled conditions:
    - Seed germination testing in Petri dishes under optimal conditions, in order to establish the fraction of seeds with primary dormancy;
    - Temperature-dependent germination testing (the lower the temperature requirement, the more seeds will immediately germinate before secondary dormancy can develop); and
    - Seed germination testing of seeds with induced secondary dormancy (to examine the fraction of seeds that can acquire secondary dormancy (dark dormancy)) (Schatzki et al., 2013).
- Seed survival tested under field conditions:
    - Seed germination testing of seeds buried in the field at 10-20 cm depth for one year, or other methods to estimate seed survival in the soil (Schatzki et al., 2013); and
    - Survey of volunteers in the subsequent years in the field trials.

Experiments that can be performed as part of SD++ and from which applicants can select the most relevant ones based on the sources of the problems and the nature of the hypothesis under test, are given below. Applicants are requested to provide a scientific rationale justifying their selection of specific experiments.

- Studies to gather more accurate data on specific demographic parameters, in order to feed the demographic model and to fine-tune the population growth rate estimates:
    - Plant vigour testing under (extreme) environmental conditions;
    - Biotic and abiotic stress responses (through stress response tests performed under greenhouse conditions tests mimicking different environmental conditions);
- Survey of feral (GM) plants;
- Survey of volunteer (GM) plants in the subsequent years in the field trials, which are left unmanaged, so that they revert to natural vegetation (Raybould et al., 2012);
- Vegetation competition studies;
- Manipulative field experiments; and
- Population modelling.

CONCLUSIONS

To provide a comprehensive and harmonised approach for the agronomic and phenotypic characterisation of GM plants, and to ensure the best use of those data for the comparative analysis and food/feed and environmental risk assessment, specific guidance was developed for the agronomic and phenotypic characterisation of GM plants. Recommendations given focus on: the representativeness of the selected sites of field trials; the representativeness of the selected test materials and their quality; the experimental design of field trials; the selection and description of (ecologically) more relevant endpoints; data analysis; and risk assessment implications.

Additional conclusions will be elaborated after the public consultation.

DOCUMENTATION PROVIDED TO EFSA

1. Proposal for a self-task mandate of the EFSA GMO Panel to establish a new Working Group to supplement the guidelines on the agronomic and phenotypic characterisation of GM plants. May 2013. Submitted by the Chair of the EFSA GMO Panel.
2. Acceptance of the self-task mandate of the EFSA GMO Panel to establish a new Working Group to supplement the guidelines on the agronomic and phenotypic characterisation of GM plants. July 2013. Submitted by EFSA Executive Director.

REFERENCES

Anderson PK, Cunningham AA, Patel NG, Morales FJ, Epstein PR, Daszak P, 2004. Emerging infectious diseases of plants: pathogen pollution, climate change and agrotechnology drivers. Trends in Ecology & Evolution, 19, 535-544.

BBCH, 2001. Growth stages of mono-and dicotyledonous plants - BBCH Monograph, 2 Edition, 2001, Edited by Uwe Meier Federal Biological Research Centre for Agriculture and Forestry.

CABI, 2014. Available on line http://www.cabi.org/cpc.

Codex Alimentarius, 2007. Codex Alimentarius Commission Procedural Manual, Seventeenth edition. Codex Alimentarius Commission, Joint FAO/WHO Food Standards Programme, Rome. http://www.fao.org/docrep/010/a1472e/a1472e00.htm.

Devos Y, Aguilera J, Diveki Z, Gomes A, Liu Y, Paoletti C, du Jardin P, Herman L, Perry JN, Waigmann E, 2014. EFSA's scientific activities and achievements on the risk assessment of genetically modified organisms (GMOs) during its first decade of existence: looking back and ahead. Transgenic Research, 23, 1-25.

DS-3077, 2013. Representative Sampling – HORIZONTAL Standard. www.ds.dk. DS 3077:2013. ICS: 03. 120. 30; 13.080.05.

EC, 2009. Directive 2009/128/EC of the European Parliament and of the Council of 21 October 2009 establishing a framework for Community action to achieve the sustainable use of pesticides. Official Journal of the European Union L309, 71-86.

EFSA, 2010a. Guidance on the environmental risk assessment of GM plants. EFSA Journal, 8(11):1-111 [1879]. http://www.efsa.europa.eu/en/efsajournal/doc/1879.pdf.

EFSA, 2010b. Statistical considerations for the safety evaluation of GMOs. EFSA Journal, 8(1):1-59 [1250]. http://www.efsa.europa.eu/en/efsajournal/doc/1250.pdf.

EFSA, 2011a. Guidance for risk assessment of food and feed from GM plants. EFSA Journal, 9(5):1-37 [2150]. http://www.efsa.europa.eu/en/efsajournal/doc/2150.pdf.

EFSA, 2011b. Guidance document on selection of comparators for the risk assessment of GM plants. EFSA Journal, 9(5):1-20 [2149]. http://www.efsa.europa.eu/en/efsajournal/doc/2149.pdf.

Esbensen KH, Paoletti C, Minkkinen P, 2012a. Representative sampling of large kernel lots I. Theory of Sampling and variographic analysis. Trends in Analytical Chemistry, 32, 154-164.

Esbensen KH, Paoletti C, Minkkinen P, 2012b. Representative sampling of large kernel lots III. General considerations on sampling heterogeneous foods. Trends in Analytical Chemistry, 32, 178-184.

EU, 2013. Commission Implementing Regulation (EU) No 503/2013 of 3 April 2013 on applications for authorisation of genetically modified food and feed in accordance with Regulation (EC) No 1829/2003of the European Parliament and of the Council and amending Commission Regulations (EC) No 641/2004 and (EC) No 1981/2006. Official Journal of the European Union, L157:1-47. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2013:157:FULL:EN:PDF.

FAO/WHO, 2000. Safety aspects of genetically modified foods of plant origin. Report of a Joint FAO/WHO Expert Consultation on Foods Derived from Biotechnology, Geneva, Switzerland, 24 to 28 September 2001. FAO, Rome. http://www.who.int/foodsafety/publications/biotech/ec_sept2001/en/.

Garcia-Alonso M, 2010. Current challenges in environmental risk assessment: the assessment of unintended effects of GM crops on non-target organisms. IOBC/WPRS Bulletin, 52, 57-63.

Garcia-Alonso M, Hendley P, Bigler F, Mayeregger E, Parker R, Rubinstein C, Satorre E, Solari F, McLean MA, 2014. Transportability of confined field trial data for environmental risk assessment

of genetically engineered plants: a conceptual framework. Transgenic Research, DOI:10.1007/s11248-014-9785-0.

Giraud T, Gladieux P, Gavrilets S, 2010. Linking the emergence of fungal plant diseases with ecological speciation. Trends in Ecology & Evolution, 25, 387-395.

Horak MJ, Rosenbaum EW, Woodrum CL, Martens AB, Mery RF, Cothren JT, Burns JA, Nickson TE, Pester TA, Jiang C, Hart JL, Sammons B, 2007. Characterization of roundup ready flex cotton, 'MON 88913', for use in ecological risk assessment: Evaluation of seed germination, vegetative and reproductive growth, and ecological interactions. Crop Science, 47, 268-277.

IUSS, 2007. World Reference Base for Soil Resources 2006, first update 2007. World Soil Resources Reports No. 103. FAO, Rome.

ISTA, 2014. International Rules for Seed Testing. Seed Health Testing. ISTA, Bassersdorf, Switzerland. Available on line: http://www.ingentaconnect.com/content/ista/rules/2014/00000001/00000001.

Kok EJ and Kuiper HA, 2003. Comparative safety assessment for biotech crops. Trends in Biotechnology, 21, 439-444.

Minkkinen P, Esbensen KH, Paoletti C, 2012. Representative sampling of large kernel lots II. Application to soybean sampling for GMO control. Trends in Analytical Chemistry, 32, 165-177.

Nickson TE, 2008. Planning environmental risk assessment for genetically modified crops: Problem formulation for stress-tolerant crops. Plant Physiology, 147:2, 494-502.

OECD, 1993. Safety considerations for biotechnology: scale up of crop plants. OECD, Paris, 1-43. http://www.oecd.org/dataoecd/26/26/1958527.pdf.

OECD, 2012. OECD Seed Schemes. Guidelines for control plot tests and field inspection of seed crops. Available on line: http://www.oecd.org/tad/code/for%20website%20-%20E%20-%20%20Guidelines%20Control%20plot%20tests%20-%20100912.pdf.

OECD, 2014. OECD Seed Schemes. Schemes for the varietal certification or the control of seed moving in international trade. Available on line: http://www.oecd.org/tad/code/1_complete%20document.pdf.

Raybould A, Tuttle A, Shore S, Stone T, 2010. Environmental risk assessments for transgenic crops producing output trait enzymes. Transgenic Research, 19, 595-609.

Raybould A, Higgins LS, Horak MJ, Layton RJ, Storer NP, De La Fuente JM, Herman RA, 2012. Assessing the ecological risks from the persistence and spread of feral populations of insect-resistant transgenic maize. Transgenic Research, 21, 655-664.

Schatzki J, Allam M, Klöppel C, Nagel M, Börner A, Möllers C, 2013. Genetic variation for secondary seed dormancy and seed longevity in a set of blackseeded European winter oilseed rape cultivars. Plant Breeding, 132, 174-179.

Steven MD, Biscoe PV, Jaggard KW, Paruntu J, 1986. Foliage cover and radiation interception. Field Crops Research, 13, 75-87.

Thompson A; Taylor BN, 2008. The International System of Units (SI) (Special publication 330). Gaithersburg, MD. National Institute of Standards and Technology.

van der Voet H, Perry JN, Amzal B, Paoletti C, 2011. A statistical assessment of differences and equivalences between genetically modified and reference plant varieties. BMC Biotechnology, 11:15.

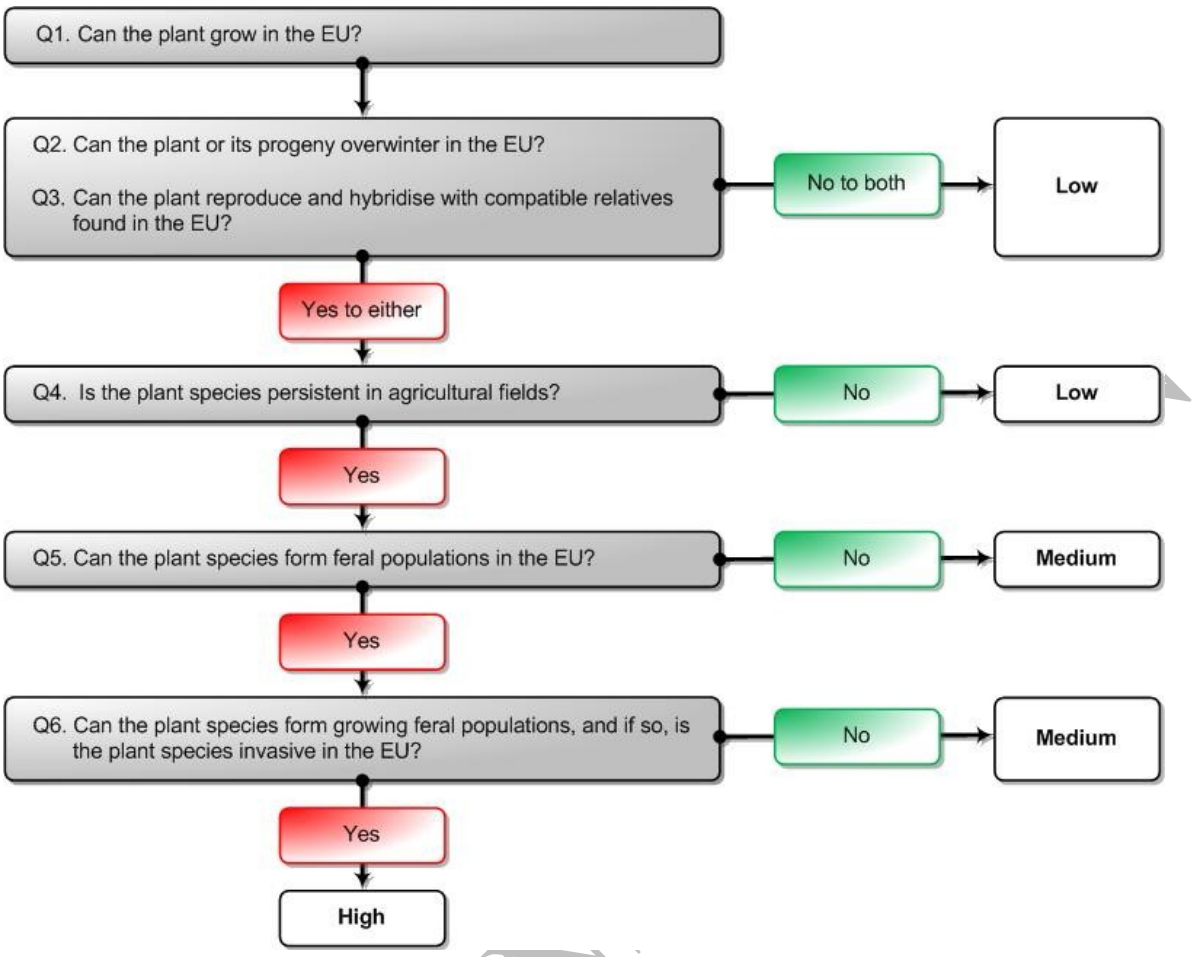
Appendix A. Decision tree to inform the problem formulation for the persistence and invasiveness assessment of GM plants

The decision tree presented in Section 7.1.1 is structured around five successive steps (steps 1 to 5 in Figure 3, above in the core text), which are designed to determine whether, and if so, under what circumstances the standard agronomic and phenotypic dataset shall be extended and/or complemented with additional measurements and/or experiments. The final risk category for the GM plant strongly influences the type of agronomic and phenotypic data that is needed to usefully inform the assessment of its persistence and invasiveness according to the staged approach outlined in EFSA (2010).

- Step 1 outlines the questions to be answered to estimate the potential persistence and invasiveness of the parental species;
- Step 2 focuses on the critical factors limiting the persistence and invasiveness of plants, and enables to consider whether the transgene may lead to a change in persistence or invasiveness;
- Step 3 accounts for the level of exposure to the GM plant in the EU based on its intended uses (cultivation vs. import);
- Step 4 involves the selection of additional agronomic and phenotypic characteristics that are relevant for the persistence and invasiveness assessment. These characteristics are selected on a case-by-case basis, using the outcomes of the former steps of the decision tree; and
- Step 5 concerns the selection of additional experiments to complement field trials.

Step 1 – Persistence and invasiveness potential of the parental species

The first step of the decision tree aims to estimate the persistence and invasiveness of the parental species of the GM plant. This is to be achieved by answering the six successive questions, which are extracted from the staged approach outlined in the EFSA (2010) guidelines on the environmental risk assessment of GM plants (see Figure A1, below). Based on the answers to the questions, plants can be categorised according to their persistence and invasiveness ability. Three risk categories are proposed: Low, Medium and High. When answering the questions, the RE in which the plant may be released should be taken into account. It therefore is suggested to refer to relevant geographical zones across the EU.



1606

1607 **Figure A1:** Questions to categorise the persistence and invasiveness potential of the parental
1608 species of the GM plant as part of step 1 of the decision tree (see Figure 3, above). The figure
1609 is based on questions outlined in Figure 4 of the EFSA (2010) guidelines for the
1610 environmental risk assessment of GM plants.

1611

1612

Examples on how experts answered the questions to complete step 1 for a number of plants are given in Table A1.

Table A1: Categorisation of plants based on their persistence and invasiveness ability, using some questions extracted from the staged approach outlined in the EFSA (2010) guidelines on the environmental risk assessment of GM plants.

Plant species	EU region (zone)	Questions extracted from EFSA (2010)						Risk category
		Q1	Q2	Q3	Q4	Q5	Q6	
Cotton	North	-	-	-	-	-	-	Low
	Central	-	-	-	-	-	-	Low
	South	+	+	-	+	+	-	Medium
Maize	North	+	-	-	-	-	-	Low
	Central	+	+/-	-	-	-	-	Low
	South	+	+	-	+/-	-	-	Low
Oilseed rape	North	+	+	+	+	+	-	Medium
	Central	+	+	+	+	+	-	Medium
	South	+	+	+	+	+	-	Medium
Potato	North	+	+	-	+	(+)	-	Medium
	Central	+	+	-	+	-	-	Medium
	South	+	+	-	+	-	-	Medium
Soybean	North	+	-	-	-	-	-	Low
	Central	+	+/-	-	-	-	-	Low
	South	+	+	-	-	-	-	Low
Sugar beet	North	+	+	+	+	+	-	Medium
	Central	+	+	+	+	+	-	Medium
	South	+	+	+	+	+	-	Medium

(+) transport of soil containing tubers (e.g. soil dumps on farms)

Step 2 – Assessment of whether the new trait changes persistence and invasiveness

In the previous step, the persistence and invasiveness potential of the parental species of the GM plant was estimated. Based on the decision tree, species were classified in three risk categories. The purpose of the next step is to identify whether the transgene may lead to a change in persistence or invasiveness of the GM plant compared to its parental species. This assessment requires the consideration of the various life cycle stages of the parental species, as this will make explicit which life cycle stage(s) limits or facilitates the persistence of a species, and whether the transgene will allow the GM plant to overcome limiting factors. Based on the analysis of a weed population model, a number of important life cycle stages and processes were identified (see below), which cover the entire life cycle of the GM plant.

Examples of biotic and abiotic factors that may limit the ability of plants to form self-sustaining populations under either cultivated or uncultivated conditions are given in Table A2. These factors may restrict or limit the niche of the plant to certain habitats, or may control its fecundity and population size. Table A2 gives an overview of intrinsic factors and the ability of the plant to deal with external factors (van Ittersum et al., 1997):

- Growth defining factors: factors, including plant characteristics, that determine the potential production level at optimum supply of all inputs;

- Growth limiting factors: include the abiotic factors such as water and nutrients that in limited supply reduce the potential production; and

- Growth reducing factors: comprise the biotic factors that reduce or hinder plant growth.

Table A2: Factors limiting the persistence and invasiveness of the parental plant species.

Life cycle stage	Intrinsic growth defining factors	Ability of plant to deal with	
		Abiotic factors	Biotic factors
Seedlings	- growth rate - growth duration - early ground cover	- moisture - temperature - salt - toxic levels of essential elements (e.g., N, P) - pH	- plant pathogens - herbivores - competing plants
Mature plants	- growth rate - growth duration - harvest index - maximum individual plant size (g/plant)	- moisture - temperature - toxic levels of essential elements - pH - soil nutrient status	- plant pathogens - herbivores - competing plants
Produced seeds	- seed size - seed number - propensity to shatter - dispersal characteristics	- temperature	- seed predators
Seedbank	- seed primary dormancy - likelihood of induction into secondary dormancy - seed longevity/survival oil/starch content - death rate in or on soil	- moisture - temperature - light level (e.g., through seed burial) - time - disturbance (e.g., flooding)	- seed predators - seed pathogens - disturbance

Applicants are requested to list the key factors limiting the population growth of the parental species, and indicate whether the transgene has the potential to alter the persistence and invasiveness of the parental species. Knowledge of the factors preventing the establishment, persistence and spread of the plant helps to assess whether the acquisition of a transgene will allow the GM plant to overcome such factors, and thus alter its persistence and invasiveness potential. Applicants shall provide a scientific rationale underpinning their conclusions. This should be achieved by describing:

- the underpinning knowledge of life cycle, growth factors and stresses of the type in Table A2 and the key factors limiting the persistence and invasiveness ability of the GM plant;

- the potentially invasive range of the plant species using climate-matching analyses under various assumptions (e.g., drought tolerance) and scenarios (e.g., irrigation, climatic change, pest pressure);

- the invasiveness potential in susceptible habitats;

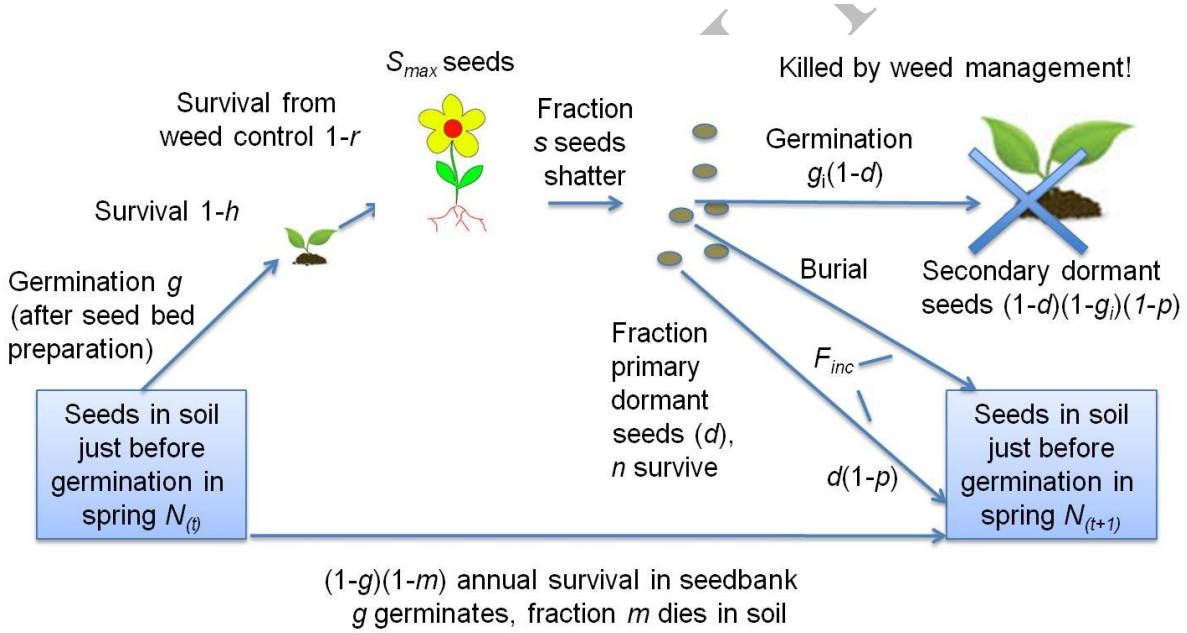
- the competitive interactions with desirable species within specific habitats; and

- potential for hybridisation with wild relatives.

If the genetic transformation alters some weedy characters of the parental species (such as the ability to make long-lived soil seedbank in the soil, small seeds, longer flowering period, shorter vegetative period before seeds are produced; reviewed by Warwick et al., 2005; Kos et al., 2011), then applicants should discuss whether, and if so, how this may affect the persistence and invasiveness potential of the GM plant compared to the parental species.

In many cases, the novel trait does not matter for survival or seed production, or may lower survival or seed production. In that case, step 2 gives a checklist of which life history parameters to consider. The final conclusion will then be that the transgene does not increase the persistence or invasiveness potential. Yet, for the species×trait combinations for which it is likely that the transgene will increase survival or seed production, or weedy characteristics based on the narrative description of the type information outlined in Table A2, the EFSA GMO Panel recommends to combine measures and quantify population growth. Demographic analysis is considered an appropriate tool to estimate/quantify population growth (Booth et al., 2010), as it gives an indication of potential persistence and invasiveness. Persistence and invasiveness are influenced by several factors at any stage of the life cycle of the plant, and traits that impact survival may therefore significantly alter population growth. For this reason, population metrics that incorporate vital rates across the whole life cycle, like the population growth rate ($\lambda=N_{t+1}/N_t$), should be considered for the evaluation of the persistence and invasiveness of specific GM plants. The different steps in population growth are outlined in Figure A2 (below). The applicant could outline which steps in Figure A2 (below) are bottlenecks, i.e., for which stages mortality is so high that population growth is below 1. Combining information on different demographic stages can also show whether growth rate is far below one or only just below.

It is important to distinguish between conditions without strong competition, like on the agricultural field or outside the field after a disturbance, and conditions with strong competition like on abandoned fields or outside the field.



$N_{(t+1)}/N_{(t)}=(1-g)(1-m)+[gS_{exp}s]F_{inc}$
with expected per capita seed production $S_{exp}=(1-r)(1-h)S_{max}$
and fraction of seeds included in the seedbank $F_{inc}=d(1-p)+(1-d)(1-g_i)(1-p)$

Figure A2: Demographic analysis to estimate/quantify population growth on the agricultural field. In the example given, seeds production occurs in autumn and successful germination, leading to surviving seedlings, occurs in spring after the farmer has prepared the seed bed, eliminating all weeds that emerged during autumn and winter (indicated by the X). Outside the agricultural field there is less or no weed control so the route through direct germination (X) should also be considered. There are three stages in the life cycle: (i) survival of seeds in the soil seedbank; (ii) production of new seeds; and (iii) incorporation of some of these seeds, the dormant ones or the seeds that acquire secondary dormancy, into the soil seedbank.

In the following, it is indicated how each life cycle stage can be quantitatively characterised. Based on all of these endpoints, parameter λ can be estimated. A value of $\lambda > 1$ indicates that the species is able to persist. If sensitive stages are identified, the next step is to verify whether the transgene is likely to influence one of the processes resembled in that specific element. For the model description we will assume that: the species produces seeds in the summer; seeds can then immediately germinate or enter the seedbank; and germination from the seedbank occurs in spring. A demographic model with probabilities for individual plants to survive and produce seeds was used. The model can be adapted to annual species by adjusting the timing of the life cycle. The model encompasses the following stages:

- *Seedbank survival*: If losses from the seedbank are low, then the number of seeds in the soil seedbank will gradually accumulate. This ability to accumulate seeds in the soil seedbank is an important feature for persistence. This is even more so for feral populations outside the agricultural field. Such populations have to bridge periods of several years between which disturbance and subsequently favourable conditions occurred to complete their life cycle.

  The most important processes responsible for seed losses in the soil seedbank are germination after preparation of the seed bed by the farmer (g). This parameter is species-dependent, yet also affected by the type of tillage (cf., depth at which the soil is prepared). In Figure A2, seed bed preparation occurs in spring, but it could also be in autumn (e.g., for oilseed rape), not long after the previous harvest. The annual mortality rate of buried seeds in the seedbank is so that of the non-germinating seeds a fraction (1-m) survives.

- *Survival from seedling to flowering plant*: The most important processes responsible for mortality of seedlings on an agricultural field are: the fraction seedlings killed through herbivores such as slugs (r), and weed control (h). Consequently, a fraction (1-r)(1-h) survives both. Outside the field, conditions will be less optimal and seedlings may die from desiccation or competition with other plants may increase mortality risk.

- *Competitive ability of individual plants*: Apart from the number of plants that will contribute to seed production, the size and competitive ability of those plants is important, as with taller and more competitive plants fewer plants are needed to replenish the soil seedbank.

- *Density-dependence of seed production*: If more plants of the same species occur closely together, then intraspecific competition occurs and this reduces the seed production potential. Likewise, seedling survival until flowering (self-thinning) can depend on density. This factor can be incorporated into the model, but it also complicates it. Given sufficient time, the weed population is no longer growing resulting in population growth rate of 1. The ability of the weed species to grow when it is still at low density (with a relatively low level of density-dependence) is therefore the best indicator for the weed species to become established.

- *Seed production (Smax)*: The ability to produce a high number of seeds is an important characteristic. Seeds can simply be counted. A rough estimate of the seed production rate (Smax) of the parental species can also be based on shoot dry matter, reproductive effort and individual seed weight (sw: g/seed).

- *Fraction of newly produced seeds incorporated in the soil seedbank (F_{inc})*: For persistence a huge fraction of the seeds should be incorporated in the seedbank. Only the fraction of the seeds that shatter (*s*) are dispersed and enter the environment. In general, a fraction (d) of the seeds is primary dormant. These seeds do not germinate even under the best conditions (like in Petri dishes). The primary dormant seeds enter the soil seedbank. Seeds that are not dormant can acquire secondary dormancy if they stay in the dark in the soil for a couple of weeks (for instance, oilseed rape). This fraction depends on environmental conditions (not allowing germination), on management regime (burial of seeds) and on genotype of the crop (Schatzki et al., 2013). In other species, seeds may not be dormant (i.e., they germinate under optimal

conditions), but conditions in the field may be unsuitable (for instance, insufficient light). Such seeds are also incorporated in the seedbank. The fraction of shattered seed that is non-dormant is called (1-d). From these seeds a fraction, g_i, germinates *immediately*, or later in the soil before secondary dormancy can be acquired. These seeds are unsuccessful in an agricultural situation because before the sowing the farmer will clear the seed bed. Outside the agricultural field, the direct germination may lead to viable plants and therefore this potential route is already indicated in Figure A2.

Table A3: Meaning of symbols used in the calculation of population growth.

Parameter	Meaning
g	Fraction of seed germination after seed bed is made
m	Annual fraction of seed mortality
r	Mortality from seedling to lowering plant through weed control
h	Other mortality from seedling to flowering plant (e.g., herbivory)
S_{max}	Seed production per plant
S_{exp}	Expected seed production per seedling after accounting for mortality (and possibly intra- and interspecific competition) $=(1-m)(1-h)S_{max}$
s	Fraction of seeds that shatters
d	Fraction of fresh seeds that is primary dormant, i.e. does not germinate under favourable conditions (wet, warm, light) in Petri dishes
g_i	Fraction of seed germination within weeks from seed shatter (and do not acquire secondary dormancy)
p	Mortality of fresh seeds that acquire secondary dormancy until the next census moment (spring in Figure A2)
F_{inc}	Fraction of fresh seeds included in the seedbank measured at moment of census $=d(1-p)+(1-d)(1-g_i)(1-p)$

In the previous paragraphs, the most important stages and transitions between stages in a plants' life cycle were discussed. Together these stages and transitions determine whether a species is likely to persist on an agricultural field. Figure A2 and Table A3 give the most important parameters. All this information can be summarised to give an estimation of λ. Values exceeding a λ-value of 1 indicate that the species is likely to persist in an agricultural field.

Both weed species, *Avena fatua* and *Chenopodium album*, species are characterised by a high λ, demonstrating their weedy character. Even when the farmer cleans the seed bed in spring, these species are hard to control. For both weed species λ-values were calculated by assuming that all produced seeds are actually incorporated in the soil seedbank. For *Brassica napus* and *B. rapa* seeds have no primary dormancy and many will germinate in late summer as soon as there is sufficient rainfall. According to the model, these seedlings are eradicated by weed management. Under this regime the weed species needs to regenerate from seedling that emerge in spring after the seed bed is cleaned. Table A4 shows that under such conditions *B rapa* could become a weed (λ=104.05), while the population growth of *B. napus* is much lower (λ=1.24), but still higher than 1. Plugging in lower success and viability as observed under feral conditions, one year after clearing a field, shows that the crop species performs much worse than its weedy relative; *B. napus* disappears where *B. rapa* can persist. It should be noted, however, that under feral conditions the model should be modified since the farmer can no longer controls seedling establishment in autumn and as a result of which seedling germination could be successful.

1768 **Table A4:** Best estimates of demographic parameters needed to calculate population growth rate.

Parameter	*Chenopodium album*[(a)]	*Avena fatua*[(a)]	*Brassica napus* – agricultural field[(b)]	*B. rapa* – agricultural field	*B. napus* – feral	*B. rapa* – feral
g	0.1	0.68	0.5	0.5	0.5	0.5
m[(c)]	0.2	0.2	0.96	0.55	0.96	0.55
r	0.9	0.9	0[(d)]	0	0	0
h	0	0	0.62	0.44	1	0.88
S_{max}	2500	300	1600	3311	1825	3311
S_{exp}	250	30	608	1854	0	397
s	1	1	0.4	0.4	0.4	0.4
d	1	1	0	0.10	0	0.10
g_i	0	0	ca 99%	ca 80%	ca 99%	ca 80%
p	0	0	0	0	0	0
F_{inc}	1	1	0.01	0.28	0.01	0.28
λ	**25.72**	**20.66**	**1.24**	**104.05**	**0.02**	**22.47**

1769 (a): Data from *C. album* and *A. fatua* from Freckleton and Watkinson (1998) and Spitters (1989), respectively.
1770 (b): Table Data from Hesse and de Jong (in preparation), Boskoop (wet habitat) for year 1 (agricultural field) and year 2
1771 (feral). Median values of 20 (*B. napus*) and 9 accessions (*B. rapa*). F_{inc} estimated as seedlings emerging in spring from
1772 seeds sown in September. This assumes no seeds are left in seedbank.
1773 (c): Data from de Jong et al. (2013).
1774 (d): We assume no control of volunteers.
1775

1776 After a first estimate of λ is made, each of the elements on which the estimate is based could be
1777 consecutively modified by 20% to determine the sensitivity of λ to the respective element. If a specific
1778 element shows to have a major impact on λ, then the next step is to verify whether the transgene is
1779 likely to influence one of the processes resembled in that specific element. As indicated before, the
1780 scheme can be used for indicating that even when a transgene has positive effects, λ is still far below 1
1781 for the species.

1782 Relevant data to feed the calculations can be retrieved from the scientific literature and that such data
1783 do not necessarily have to be generated. In this stage, it is not necessary to fully quantify all
1784 demographic parameters. The scheme may be useful for identifying bottlenecks for the crop to become
1785 feral. For instance, the exercise in Table A4 suggests some points. (i) Population growth for *C. album*
1786 and *A. fatua* is quite high so these plants are likely to be weeds. The wild *B. rapa* performs better than
1787 the crop *B. napus* in many aspects, and will be superior in the field. (iii) There is strong bottleneck for
1788 feral *B. napus* in year 2 (after soil cultivation in year 1), as seedling mortality is almost 100%. This
1789 suggests that modifications that increase seedling survival should be treated with cautions but
1790 modifications that affect other parts of the life cycle will have no effect. The scheme helps applicants
1791 to point out the bottlenecks for different species.

1792 On the agricultural field and without crop rotation, conditions are relatively constant. With crop
1793 rotation or outside the agricultural field, disturbed conditions suitable for establishment may be
1794 followed by unsuitable years. In such a case, λ-values for good years should be combined with values
1795 for bad years to estimate population growth over a longer period. In such case, populations may need
1796 to bridge a sequence of unsuitable years. A long-lived seed bank allows this, the more so if annual
1797 mortality (m) in the soil is low.

1798 On the basis of the information considered so far, step 2 decides which risk category the GM plant
1799 now falls under. A low category at step 1 could remain Low after step 2 or rise to Medium or High,
1800 while a Medium category at step 1 could remain Medium or rise to High.

1801

Step 3 – Exposure scenarios

For GM plant applications for cultivation, the principle route of exposure will be through the sowing of seeds/propagules in agricultural fields, and the consequent movement of pollen and distribution of seed or propagules to other fields and the wider environment. For GM plant applications for import, the persistence and invasiveness assessment is concerned mainly with the environmental consequences of accidental release of viable GM seeds or propagating material during import, transportation, storage, handling and processing.

When evaluating the likelihood and seriousness of harm to the environment following the cultivation of a GM plant, the environmental risk assessment assumes 100% exposure over an extended period of time (high-exposure scenarios). Exposure and potential impact are expected to be the highest under cultivation conditions (Devos et al., 2012). However, under import conditions, the context usually differs, as only few GM plants are expected to be present in the environment. Therefore, low-exposure scenarios can be expected to reduce the likelihood of one or more steps in a risk scenario, compared with a cultivation scenario (Roberts et al., 2013).

The decision tree considers two levels of exposure when determining the necessary level of agronomic and phenotypic characterisation of GM plants:

- High-exposure scenarios for GM plant applications for cultivation; and
- Low-exposure scenarios for GM plant applications for import.

Step 4 – Additional agronomic and phenotypic endpoints to the standard agronomic and phenotypic dataset (SD+)

The purpose of step 4 is to confirm the level of agronomic and phenotypic characterisation that is likely to be needed for risk assessment. If the risk after steps 1-2-3 is judged to be low, then the standard set of endpoints considered by applicants is sufficient. This list would consist of the measurements that are typically made in agronomic trials, such as plant growth stage, cover, population density, growth rate, yield and yield components (see Section 5 in the core text for an overview of mandatory and case-specific endpoints).

If the risk after steps 1-2-3 is judged to be Medium-No-HES or High, then further measurements are considered necessary. In this case, the standard agronomic dataset is augmented by measuring additional agronomic and phenotypic characteristics. The selection of these additional endpoints should be based on the limiting factors.

Based on the source of the problems and the nature of the hypothesis under test, applicants should consider the below endpoints, and provide a scientific rationale justifying their selection.

- Seed dormancy potential tested under controlled conditions:
  - Seed germination testing in Petri dishes under optimal conditions, in order to establish the fraction of seeds with primary dormancy;
  - Temperature-dependent germination testing (the lower the temperature requirement, the more seeds will immediately germinate before secondary dormancy can develop); and
  - Seed germination testing of seeds with induced secondary dormancy (to examine the fraction of seeds that can acquire secondary dormancy (dark dormancy)) (Schatzki et al., 2013).
- Seed survival tested under field conditions:

- Seed germination testing of seeds buried in the field at 10-20 cm depth for one year, or other methods to estimate seed survival in the soil (Schatzki et al., 2013); and
- Survey of volunteers in the subsequent years in the field trials.

Step 5 – Additional experiments to the standard agronomic and phenotypic dataset (SD++)

Depending on the outcomes of the previous steps in the decision tree and as part of the problem formulation, performing additional experiments may be advocated. If the risk after steps 1-2-3 is judged to be high (such as for High-No-HES and High-Yes-LES/HES), then additional experiments would be needed to generate informative data to support the persistence and invasiveness assessment of the GM plant. These experiments can take many forms depending on the hypothesis under test. Therefore, applicants shall clarify which type of additional data may be necessary to support the persistence and invasiveness assessment, and how those should be generated in the frame of the staged approach outlined in EFSA (2010).

Possible experiments, from which applicants can select the most relevant ones based on the sources of the problems and the nature of the hypothesis under test, are given below. Applicants are requested to provide a scientific rationale justifying their selection of specific experiments.

- Studies to gather more accurate data on specific demographic parameters, in order to feed the demographic model and to fine-tune the population growth rate estimates:
  - Plant vigour testing under (extreme) environmental conditions;
  - Biotic and abiotic stress responses (through stress response tests performed under greenhouse conditions tests mimicking different environmental conditions);
- Survey of feral (GM) plants;
- Survey of volunteer (GM) plants in the subsequent years in the field trials, which are left unmanaged, so that they revert to natural vegetation (Raybould et al., 2012);
- Vegetation competition studies;
- Manipulative field experiments; and
- Population modelling.

If there a sufficient data available in the scientific literature to reliably falsify the hypothesis under test, then there is no need to perform the additional experiments. However, if the level and quality of the available information is low, then the hypothesis could be tested in a tiered manner according to the staged approach outlined in EFSA (2010).

References

Booth DB, Murphy SD, Swanton CJ, 2010. Invasive plant ecology in natural and agricultural systems. CABI, Oxfordshire.

de Jong T, Isanta MT, Hesse E, 2013. Comparison of the crop species *Brassica napus* and wild *B. rapa*: characteristics relevant for building up a persistent seed bank in the soil. Seed Science Research, 23, 169–179.

Devos Y, Hails RS, Messéan A, Perry JN, Squire GR, 2012. Feral genetically modified herbicide tolerant oilseed rape from seed import spills: are concerns scientifically justified? Transgenic Research, 21, 1–21.

EFSA, 2010. Guidance on the environmental risk assessment of GM plants. EFSA Journal, 8(11):1–111 [1879]. http://www.efsa.europa.eu/en/efsajournal/doc/1879.pdf.

Freckleton RP, Watkinson AR, 1998. Predicting the determinants of weed abundance; A model for the population dynamics of *Chenopodium album* in sugar beet. Journal of Applied ecology, 35, 904–920.

Kos M, Broekgaarden C, Kabouw P, Lenferink KO, Poelman EH, Vet LEM, Dicke M, van Loon JA, 2011. Relative importance of plant-mediated bottom-up and top-down forces on herbivore abundance on *Brassica oleracea* Functional Ecology, 25, 1113–1124.

Raybould A, Higgins LS, Horak MJ, Layton RJ, Storer NP, De La Fuente JM, Herman RA, 2012. Assessing the ecological risks from the persistence and spread of feral populations of insect-resistant transgenic maize. Transgenic Research, 21, 655–664.

Roberts A, Devos Y, Raybould A, Bigelow P, Gray A, 2013. Environmental risk assessment of GE plants under low-exposure conditions. Transgenic Research, DOI:10.1007/s11248-013-9762-z.

Schatzki J, Allam M, Klöppel C, Nagel M, Börner A, Möllers C., 2013. Genetic variation for secondary seed dormancy and seed longevity in a set of blackseeded European winter oilseed rape cultivars. Plant Breeding 132, 174–179.

Spitters CJT, 1989. Weeds: population dynamics, germination and competition. In: Simulation and systems management in crop protection (p. 182-216) (Eds.: R. Rabbinge, S.A. Ward and H.H. van Laar). Pudoc, Wageningen, The Netherlands.

van Ittersum MK, Rabbinge R, 1997. Concepts in production ecology for analysis and quantification of agricultural input-output combinations. Field Crops Research, 52, 197–208.

Warwick SI and Stewart Jr. CN, 2005. "Crops come from wild plants – How domestication, transgenes, and linkage together shape ferality". In: "Crop ferality and volunteerism" edited by J. Gressel, pp. 9-30. CRC Press, Boca Raton.