SCIENTIFIC OPINION

Scientific Opinion on the assessment of potential impacts of genetically modified plants on non-target organisms[1]

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

European Food Safety Authority (EFSA), Parma, Italy

ABSTRACT

The European Food Safety Authority (EFSA) asked the Panel on Genetically Modified Organisms to establish a self-tasking Working Group with the aim of (1) producing a scientific review of the current guidance of the GMO Panel for Environmental Risk Assessment (ERA), focusing on the potential impacts of GM plants on Non-Target Organisms (NTOs), (2) proposing criteria for NTOs selection, and (3) providing advise on standardized testing methodology. This initiative was undertaken in response to a need and request from a wide range of stakeholders, including the European Commission and Member States. In first instance, the self-tasking Working Group on Non-Target Organisms (EFSA NTO WG) mainly considered impacts of GM plants on invertebrate species, but also took account of ecosystem functions that could be altered. The EFSA NTO WG considered the necessity for clear and objective protection goals, for which assessment and measurement endpoints shall be developed; the need to initiate the scientific risk assessment by setting testable hypotheses; criteria for appropriate selection of test species and ecological functional groups; appropriate laboratory and field studies to collect relevant NTO data; and the use of statistical techniques that should be an integral part of experimental design. The EFSA NTO WG considered the range of approaches and methodologies of ERA of NTOs as described in the current literature and proposed risk assessment approaches based on selection of functional groups and individual species within a tiered approach. The present scientific opinion provides guidance to risk assessors for assessing potential effects of GM plants on NTOs, together with rationale for data requirements in order to complete a comprehensive ERA for NTOs. In this respect, guidance to applicants as outlined in the present opinion has been inserted in the updated Guidance Document of the EFSA GMO Panel for the ERA of GM plants.

1 On request from EFSA, Question No EFSA-Q-2008-089, adopted on 21 October 2010.

2 Panel members: Hans Christer Andersson, Salvatore Arpaia, Detlef Bartsch, Josep Casacuberta, Howard Davies, Patrick du Jardin, Gerhard Flachowsky, Lieve Herman, Huw Jones, Sirpa Kärenlampi, Jozsef Kiss, Gijs Kleter, Harry Kuiper, Antoine Messéan, Kaare Magne Nielsen, Joe Perry, Annette Poeting, Jeremy Sweet, Christoph Tebbe, Atte Johannes von Wright, and Jean-Michel Wal. Correspondence: GMO@efsa.europa.eu

3 Acknowledgement: The Panel wishes to thank the members of the self-tasking Working Group on Non-Target Organisms: Salvatore Arpaia, Detlef Bartsch, Marc Delos, Achim Gathmann, Rosie Hails, Paul Henning Krogh, Jozsef Kiss, Barbara Manachini, Joe Perry, Jeremy Sweet, Claudia Zwahlen for the preparatory work on this scientific opinion, and EFSA staff: Sylvie Mestdagh for the support provided to this scientific opinion. In addition the Panel wishes to thank the three external referees, including Sue Hartley and Gabor Lövei, for reviewing an early draft of the document during its development phase, as well as Toby Bruce and Joop van Loon for providing technical input.

Suggested citation: EFSA Panel on Genetically Modified Organisms (GMO); Scientific Opinion on the assessment of potential impacts of genetically modified plants on non-target organisms. EFSA Journal 2010; 8(11):1877. [72 pp.] doi:10.2903/j.efsa.2010.1877. Available online: www.efsa.europa.eu/efsajournal.htm

KEY WORDS

Ecosystems services, environmental risk assessment (ERA), focal species, genetically modified (GM) plants, non-target organisms (NTOs), protection goals, species selection, unintended effects, tiered approach.

SUMMARY

The European Food Safety Authority (EFSA) asked the Panel on Genetically Modified Organisms to establish a self-tasking Working Group with the aim of (1) producing a scientific review of the current guidance of the GMO Panel for Environmental Risk Assessment (ERA), focusing on the potential impacts of GM plants on Non-Target Organisms (NTOs), (2) proposing criteria for NTOs selection, and (3) providing advise on standardized testing methodology. This initiative was undertaken in response to a need and request from a wide range of stakeholders, including the European Commission and Member States.

In first instance, the self-tasking Working Group on Non-Target Organisms (EFSA NTO WG) mainly considered impacts of GM plants on invertebrate species, but also took account of ecosystem functions that could be altered.

The EFSA NTO WG considered the necessity for clear and objective protection goals, for which assessment and measurement endpoints shall be developed; the need to initiate the scientific risk assessment by setting testable hypotheses; criteria for appropriate selection of test species and ecological functional groups; appropriate laboratory and field studies to collect relevant NTO data; and the use of statistical techniques that shall be an integral part of experimental design. The EFSA NTO WG considered the range of approaches and methodologies of ERA of NTOs as described in the current literature and proposed risk assessment approaches based on selection of functional groups and individual species within a tiered approach.

The EFSA GMO Panel has recently updated the Guidance Document for the ERA of GM plants (EFSA, 2010b), including guidance for assessing potential effects of GM plants on NTOs. This opinion further describes the data requirements and gives the scientific rationale in order to complete a comprehensive ERA for NTOs. In addition, it includes examples of methodologies and stepwise approaches.

TABLE OF CONTENTS

BACKGROUND

In the European Union (EU), any genetically modified organism (GMO) and its derived products are subject to a risk assessment before they can be placed on the EU market. According to Directive 2001/18/EC that regulates the deliberate release into the environment of GMOs (EC, 2001), GMOs should only be authorised for placing on the market after a scientific assessment of any risks which they might present for human and animal health and for the environment. The general principles and methodology to be followed in this environmental risk assessment (ERA) are defined in Annex II of Directive 2001/18/EC. Both documents require that, in the context of the ERA, potential interactions between a genetically modified (GM) plant and non-target organisms (NTOs) are considered, including direct and indirect, as well as immediate and delayed effects. In this scientific opinion, potential NTOs are defined as all those species directly and/or indirectly exposed to the GM plants, and which are not targets of the newly expressed metabolite(s) in these plants.

To assist and guide applicants in the preparation and presentation of their GM plant market authorisation applications/dossiers, the scientific Panel on Genetically Modified Organisms of the European Food Safety Authority (EFSA GMO Panel) established a guidance document on the risk assessment of GM plants and derived food and feed (EFSA, 2006a). In line with Directive 2001/18/EC, the guidance document provides recommendations on how to address possible immediate and/or delayed environmental impacts resulting from direct and indirect interactions between the GM plant and NTOs. In Section 9.5 of the EFSA GMO Panel guidance document (EFSA, 2006a), general risk assessment principles to be followed in the frame of NTOs testing are defined, while the use of a tiered approach – where information collected in lower tiers directs the extent and nature of experiments conducted in higher tiers – is recommended.

However, the debate on NTO testing approaches and views is ongoing in the scientific literature. To stimulate discussion on the different approaches and views reflected in the scientific literature and to further develop scientific approaches on the ERA of GM plants, EFSA organised a scientific colloquium that was attended by different stakeholders (EFSA, 2008). Based on the discussions held at this colloquium, participants made a list of recommendations that included the use of conceptual models in NTO testing, the definition of clear and objective protection goals, as well as the use of prospective power analysis (i.e. a power analysis that is conducted before experiments are started, and which is used to inform levels of replication).

Following the discussions held and recommendations made on NTO testing at the EFSA scientific colloquium and acknowledging the different NTO testing approaches debated in the scientific literature, EFSA established a self-tasking working group on NTO (NTO WG)[4] in March 2008. Subsequently, a 2-year mandate of the European Commission (EC) reinforced the activities of the NTO WG.

NTO WG activities focused mostly on non-target arthropods and included an analysis of relevant scientific literature for assessing different NTO testing approaches, and for developing more detailed guidance in this area. Based on the work performed by the NTO WG and in line with the requirements of the EC mandate, the EFSA GMO Panel produced this scientific opinion that aims at guiding applicants through the sequential steps of an ERA for NTO-related issues, and at explaining the rationale behind the criteria suggested for NTO testing by the EFSA GMO Panel.

An updated Guidance Document for the ERA of GM plants is being prepared by the EFSA GMO Panel. That document will contain in a condensed format the guidance to applicants on the assessment of potential impacts of GM plants on NTOs. This scientific opinion can then be considered as a

[4] http://registerofquestions.efsa.europa.eu/roqFrontend/questionsList.jsf

detailed background document to support the specific section of NTOs of the updated Guidance Document for the ERA of GM plants.

This scientific opinion considers impacts of GM plants on ecosystem functions and biodiversity but does not consider impacts of GM microorganisms, deliberate releases of GM plants into the environment for experimental purposes, or the deliberate release of GM animals.

Three scientific referees, with expertise in risk assessment of NTOs, were invited to review and comment on the opinion during its development phase. In addition, the opinion was presented to stakeholders (e.g. biotech companies, Member States, environmental non-governmental organisations) during a three-day consultation in June 2009. The draft opinion was subsequently submitted for comments by the public during a two-month consultation period. Finally, EFSA and its GMO Panel met the stakeholders to further discuss the comments they had sent through the public consultation, respectively in June and September 2010.

TERMS OF REFERENCE

Recognising the importance and complexity of assessing possible environmental impacts of GM plants on NTOs, the EFSA GMO Panel decided to update the NTO-specific sections of its guidance document on GM plants. Therefore, in March 2008, EFSA established a self-tasking working group on NTO (NTO WG)[5] with the aim of (1) producing a scientific review of the current guidance document of the EFSA GMO Panel for Environmental Risk Assessment (ERA), focusing on the potential impacts of GM plants on NTOs; (2) proposing criteria for NTOs selection; and (3) advising on standardised testing methodologies.

Subsequently, the EFSA GMO Panel received a 24-month mandate from the European Commission, including a public consultation, to further develop and update its guidelines by covering the following points:

a) ERA of potential effects of GM plants on NTOs through

   i. the development of criteria for the selection of NTOs and representative species thereof, focusing on arthropods and other invertebrates, and also considering other relevant NTOs in different trophic levels;

   ii. the selection and recommendation of appropriate methods to study the potential effects of GM plants on these non-target organisms;

b) development of criteria for field trials to assess the potential ecological effects of the GM plants in receiving environments (including experimental design and analysis to ensure sufficient statistical power);

c) identification of the EU geographic regions where the GM plants (combinations crop + trait) may be released and the selection of representative receiving environment(s) which reflect the appropriate meteorological, ecological and agricultural conditions;

d) selection of appropriate techniques to assess potential long-term effects of GM plants including experimental and theoretical methodologies, and recommendations for establishing relevant baseline information.

[5] http://registerofquestions.efsa.europa.eu/roqFrontend/questionsList.jsf

Given the complexity of the topic and the large number of public comments expected, the duration of the mandate to deliver a scientific opinion on NTOs as well as related recommendations to be included into the updated Guidance Document for the ERA of GM plants was extended till November 10, 2010.

ASSESSMENT

Despite considerable variation among ERA frameworks for GM plants worldwide (Hill, 2005), risk assessment generally comprises **several sequential steps**: (step 1) problem formulation, a critical first step, including hazard identification; (step 2) hazard characterisation that examines potential hazards and their magnitude; (step 3) exposure characterisation that estimates levels and likelihood of exposure; and (step 4) integrative risk characterisation in which the magnitude of consequences and the likelihood of occurrence are integrated.

Risk characterisation (step 4) may identify risks that require management measures. Therefore, risk management strategies should be described and/or proposed by applicants. Finally an evaluation of the overall risk of the GM plant(s) (step 5) shall be made taking into account the results of the ERA and associated levels of uncertainty, the weight of evidence and the risk management strategies proposed in the receiving environment(s).

The ERA is conducted starting with step 1 and moving to step 6; step 2 and 3 can however be carried out in parallel (see Figure 1). The successive steps comprising the ERA of GM plants are discussed in the following sections focusing on the interactions between GM plants and NTOs. Further details can also be found in the updated Guidance Document of the EFSA GMO Panel for the ERA of GM plants (EFSA, 2010b).

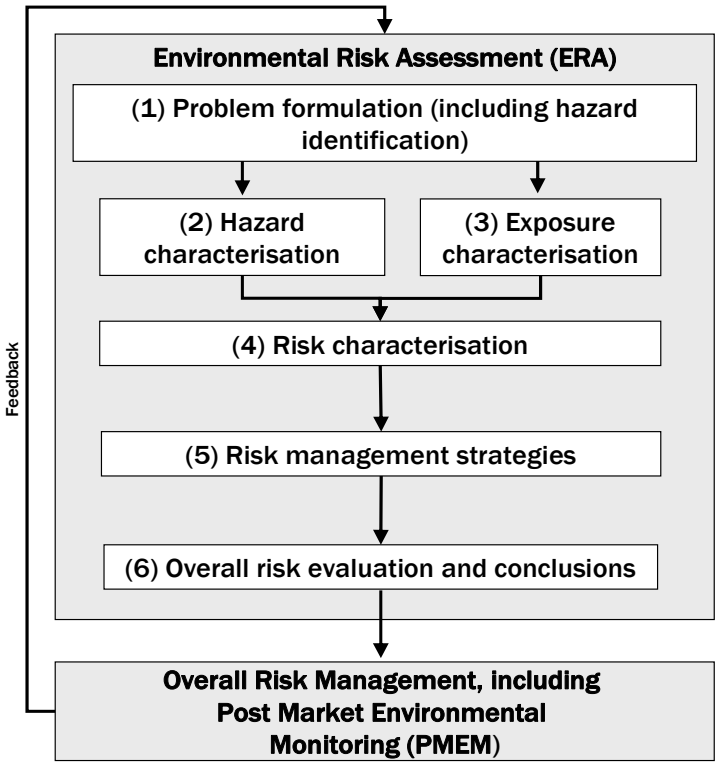

Figure 1: Six steps of the environmental risk assessment (ERA) and the relationship to overall risk management, including post-market environmental monitoring, according to Directive 2001/18/EC and Regulation (EC) No. 1829/2003.

1. Problem Formulation

1.1. Introduction

Through the identification and formulation of the problem, a broadly-stated problem shall be transformed into a manageable analysis that will be relevant for regulatory decision-making. In this respect, the most important questions to be addressed (= testable hypotheses) are to be identified by applicants (see Section 2.2.1 of EFSA, 2010b).

The GM plant itself is a potential stressor in the environment, in addition to the transgenes and its products. Environmental impacts can be a consequence of changes to the GM plant, the effects of the introduced traits and changes in management.

Problem formulation starts with the identification of potential hazards through a comparison of the GM plant with its conventional counterpart. Any differences identified are initially assessed theoretically in the problem formulation process in order to establish their potential environmental consequences. While some differences may be deemed irrelevant to the assessment, others will need to be practically evaluated for their potential to cause harm. Environmental harm is determined in relation to environmental features or goals which are considered important in order to sustain healthy and viable ecosystems. Thus emphasis is placed on functional biodiversity and ecosystem services as well as conservation of biodiversity.

1.2. Environmental protection goals

EU legislation aims at protecting the environment to ensure a high quality of life for current and future generations and to conserve global biodiversity. Some of these legally defined environmental protection goals are relevant in the context of NTO testing, as GM plants introduced into the EU may interact with several NTOs in various receiving environments (for background information, see Table 1). While Directive 2001/18/EC specifically applies to GMOs, other EU legal and strategic documents, as listed in Table 1, relate to environmental protection goals and should therefore be considered by the applicants. To scientifically assess these potential interactions, it is thus necessary to test hypotheses and identify clear assessment endpoints in the context of protection goals for biodiversity and ecosystem services (EFSA, 2010a). Ecosystems services include all services provided by ecosystems, e.g. production of food, fuel, fibre and medicines, regulation of water, air and climate, maintenance of soil fertility, cycling of nutrients. Ecosystems services are characterised by the fact that humans benefit directly from these natural assets and processes. Therefore, problem formulation starts with the definition of explicit, unambiguous and representative targets for protection in order to establish assessment endpoints (Suter, 2000) that are extracted from public policy environmental protection goals (see Table 1).

Specifically when considering NTOs, the receiving environment consists of: the managed terrestrial ecosystem (e.g. agro-ecosystem) including the GM cultivated fields, orchards and plantations and their margins and the wider environment (e.g. other adjacent GM or non-GM cultivations and non-cultivated habitats) and, where relevant, aquatic ecosystems. Therefore, protection goals need to be selected taking into consideration these defined environments.

The conservation and protection of biodiversity in the EU is of great importance. In line with the approach used by ACRE (2001), the EFSA GMO Panel applies a broad approach for the definition of environmental protection goals that include the wider biodiversity. Some environmental protection goals, such as those for wild species, are defined in Council Directive 92/43/EEC of 21 May 1992 on the conservation of natural habitats and of wild fauna and flora (EC, 1992). Directive 92/43/EEC aims to protect biodiversity through the conservation of natural habitats, wild fauna and flora in the European territory of Member States to which the Treaty applies. This Directive describes the

protection and conservation aims for natural and semi-natural habitats, conservation and conservation status. However, the approach followed by the EFSA GMO Panel does not only comprise the protection of species and habitats, but also that of ecosystem functions, such as pollination, biological control, soil functions, and quality of water and air. For example, carabid species that are well known generalist predators in agro-ecosystems might be one typical NTO group specifically in coleopteran resistant crops. The role of carabids as natural pest control agents in reducing pest populations has been reviewed and demonstrated in various crop stands (Kromp, 1999).

In a human-managed context, sustainable land use (e.g. for agriculture and forestry) is considered a primary environmental protection goal. For the benefit of sustainable production, the scope is to maintain a certain level of biodiversity, providing essential ecosystem services, including biological control of pests and diseases, nutrient fixing and cycling, decomposition of plant materials, maintenance of soil quality and fertility, and structural stability. Therefore, the criterion of functional biodiversity is deemed important in this context, since preserving the functional biodiversity may guarantee the quality of production systems (e.g. agro-ecosystems) and ensure their sustainability. Applicants shall consider whether a GM plant and its use are directly and/or indirectly (e.g. through food web interactions, scale of adoption) potentially harmful to species guilds involved in ecosystem functions.

For instance, in agro-ecosystems, soil functioning is a primary ecosystem function to be preserved. Soil is a physical and chemical matrix supporting plant growth, a source of nutrients and a habitat for species. This biodiverse environment contains several trophic levels and numerous varieties of flora and fauna. FAO (2008) indicates that soil biodiversity can be assessed, managed and conserved, showing examples of successful and unsuccessful practices which have been historically adopted in various regions of the world to manage soil biodiversity. Moreover, the Conference of the Parties to the Convention on Biological Diversity (CBD) identified soil biodiversity as an area requiring particular attention. General knowledge of the functions within soil provides risk assessors with necessary background information for the ERA. The close interaction between cultivation and soil processes leads to contacts between soil organisms (directly and indirectly) and the GM traits expressed by GM plants. In any such assessment, the significance of impacts on biodiversity of the soil system in terms of its functionality and ecosystem services needs to be addressed. Thus factors such as nutrient cycling and decomposition as well as impacts on beneficial and pathogenic associations should be considered.

Table 1: Examples of public policy environmental protection goals related to NTOs and their EU legal bases. Directive 2001/18/EC[(a)] specifically applies to GM plants; other legislations as listed below should be considered by the applicants, even though GM plants may not be specifically mentioned.

Protection goals		Legal basis		NTO function/ecosystem services[(q)]
Areas of protection		**Background**	**Scope**	
Biodiversity conservation	Species of conservation or cultural value; red list species // Protected habitats; landscapes	Directive 2004/35/EC[(b)]	Environmental liability	Sustainable agriculture
		Directive 92/43/EEC[(c)]	Conservation of natural habitats and of wild fauna and flora	Maintaining viable populations & pollination, herbivory, predation, parasitism
		Directive 2009/147/EC[(d)]	Conservation of wild birds	
		Regulation 338/97[(e)]	Protection of endangered wild fauna and flora	
		Action plan for biodiversity[(f)]	Conservation of biodiversity	
		Biodiversity strategy[(g)]	Conservation of biodiversity	
		Biodiversity action plan for the conservation of natural resources[(h)]	Conservation of natural resources	Breeding resource
		Biodiversity action plan for agriculture[(i)]	Conservation of biodiversity	
		Bern convention[(j)]	Conservation of European wildlife and natural habitats	
		Convention on biological diversity[(k)]	Conservation of biological diversity	
Agro-ecological functions	Soil	Directive 2004/35/EC	Environmental liability	Functional aspects (biological activity)
		Thematic strategy for soil protection[(l)]	Preservation of soil functions	
	Water	Directive 2000/60/EC[(m)]	Water protection	
	Production systems; plant health	Regulation 1107/2009[(n)]	Marketing of Plant Protection Products	
		Directive 2009/128/EC[(o)]	Sustainable use of PPP	Integrated Pest Management
		Biodiversity strategy	Sustainable use of biodiversity	Natural regulating mechanisms (biocontrol)
		Thematic strategy on the sustainable use of natural resources[(p)]	Sustainable use of natural resources	

(a): Directive 2001/18/EC of the European Parliament and of the Council of 12 March 2001 on the deliberate release into the environment of genetically modified organisms and repealing Council Directive 90/220/EEC
(b): Directive 2004/35/CE of the European Parliament and of the Council of 21 April 2004 on environmental liability with regard to the prevention and remedying of environmental damage
(c): Council Directive 92/43/EEC of 21 May 1992 on the conservation of natural habitats and of wild fauna and flora
(d): Council Directive 2009/147/EC of 30 November 2009 on the conservation of wild birds
(e): Council Regulation (EC) No 338/97 of 9 December 1996 on the protection of species of wild fauna and flora by regulating trade therein
(f): Commission Communication of 22 May 2006 "Halting the loss of biodiversity by 2010 - and beyond - Sustaining ecosystem services for human well-being" COM(2006) 216
(g): Communication from the Commission to the Council and the European Parliament of 4 February 1998 on a European Community biodiversity strategy COM(1998) 42
(h): Commission Communication of 27 March 2001 to the Council and the European Parliament: Biodiversity Action Plan for the Conservation of Natural Resources (Volume II) COM(2001) 162
(i): Commission Communication of 27 March 2001 to the Council and the European Parliament: Biodiversity Action Plan for Agriculture (Volume III) COM(2001) 162
(j): Council Decision 82/72/EEC of 3 December 1981 concerning the conclusion of the Convention on the conservation of European wildlife and natural habitats (Bern Convention)
(k): Council Decision 93/626/EEC of 25 October 1993 concerning the conclusion of the Convention on Biological Diversity
(l): Commission Communication of 22 September 2006 entitled "Thematic strategy for soil protection" COM(2006) 231
(m): Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy
(n): Regulation (EC) No 1107/2009 of the European Parliament and of Council of 21 October 2009 concerning the placing of plant protection products and repealing Council Directives 79/117/EEC and 91/414/EEC.
(o): 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.
(p): Communication from the Commission of 21 December 2005 - Thematic Strategy on the sustainable use of natural resources COM(2005) 670
(q): Ecosystem services are linked to human activities, whilst ecosystem functions are used in a broader context.

Mankind benefits from a multitude of resources and processes that are supplied by natural and managed ecosystems. An ecosystem can be defined at the most basic level as a natural unit of living beings (animals, plants and microorganisms) interacting with their physical environment. Ecosystem services are defined as services provided by the natural environment that benefit people. The Millennium Ecosystem Assessment[6] identifies four broad categories of ecosystem services (MEA, 2005):

- provisioning services (food & feed, freshwater, wood and fiber, fuel),
- regulating services (climate-, disease-, flood- and water regulation),
- cultural services (aesthetic, spiritual, educational recreational),
- supporting services (pollination of crops, nutrient cycling, soil formation, primary production).

European agro-ecosystems exist in a matrix of land used for other purposes, and impacts of agricultural practices can cross boundaries. Agriculture has a record of affecting biodiversity and its functioning at several levels. There is a concern that GM plant use may exacerbate negative impacts on biodiversity in agro-ecosystems. Therefore, the wider biodiversity in itself is to be considered in the selection of environmental protection goals. In this context, biodiversity is interpreted broadly and covers both species richness and agro-eco-functions providing ecosystem services. Ecosystem functions depend on the number of species, their abundances and assemblages. In a particular assemblage, the abundance of any species naturally fluctuates and the decline of a certain population might be compensated by another species within the same guild without adversely affecting functionality (Finke and Snyder, 2008). In other cases, species diversity might be of importance for a conservation purpose. In the context of general impact assessments on (wider) biodiversity, the EU has put in place assessment procedures to identify projects of high risk to biodiversity, in line with Article 14 of the Convention on Biological Diversity (CBD[7]).

When performing ERAs, applicants are asked to address environmental protection goals legally as outlined in EU legislation. Examples provided in Table 1 should thereby be used as an environmental protection goal checklist. At landscape level, for instance, a broad range of diverse habitats (e.g. agricultural areas, field crops but also natural habitats) are present. Therefore, in the ERA of interactions between GM plants and NTOs, attention shall be paid to consider conservation and protection objectives. In line with the EFSA GMO Panel opinion on Post-Market Environmental Monitoring (PMEM) (EFSA, 2006b), applicants are already requested to consider environmental protection goals associated with the cultivation of GM plants.

More detailed guidance on problem formulation procedures and how to apply problem formulation to specific areas of risk addressed in an ERA is provided in Section 2.2.1 of the updated guidance document for the ERA of GM plants (EFSA, 2010b).

[6] http://www.millenniumassessment.org
[7] http://www.biodiv.org

1.3. Receiving environments

1.3.1. Background

The EFSA GMO Panel guidance document on GM plants (EFSA, 2006a) gives special emphasis on the receiving environment and an assessment of the potential impact of GM plant use on wider biodiversity in the agro-ecosystem and in adjacent non-crop habitats is mentioned. Under Section 9 entitled 'Potential changes in the interactions of the GM plant with the biotic environment resulting from the genetic modification', the receiving environment is mentioned as follows: "*Data should be provided from field experiments in areas representative of those geographical regions where the GM plant will be grown commercially in order to reflect relevant meteorological, soil and agronomic conditions. Where data from field studies on other continents are supplied, the applicant should submit a reasoned argument that the data is applicable to European conditions*". In addition, subsection 9.5 entitled 'Interactions of the GM plant with non-target organisms' specifically states that '*An assessment is required of the possible immediate and/or delayed environmental impact resulting from direct and indirect interactions of the GM plant with non-target organisms (also taking into account organisms which interact with target organisms), including impact on population levels of competitors, herbivores, symbionts (where applicable), predators, parasites and pathogens*'.

A baseline of the receiving environment, including its organisms, their interactions and their known variations, is determined before any (harmful) characteristics of the GM plant can be identified. The potential interaction of GM plants with environmental diversity in terms of flora and fauna, climatic conditions, habitat composition, ecosystem functions and human interventions shall consider the range of environments potentially exposed to GM plants in the EU. However, in practice, the tested scenarios may only cover a subset of the multiplicity of these factors. Field trials shall be designed in order to provide relevant information in the range of receiving environments.

1.3.2. Receiving environments – Principles

The receiving environment(s) is the environment into which the GM plant(s) will be released and into which the transgene(s) may spread. The receiving environment(s) is characterized by three components (see Figure 2):

- The GM plant (e.g. plant species, genetic modification(s) and intended uses(s));
- The Geographical Zones (e.g. the climate, altitude, soil, water, flora, fauna, habitats);
- The Management Systems (e.g. land use and production systems, other cultivated GM plants, cultivation practices, integrated pest management, non-production activities and nature conservation activities).

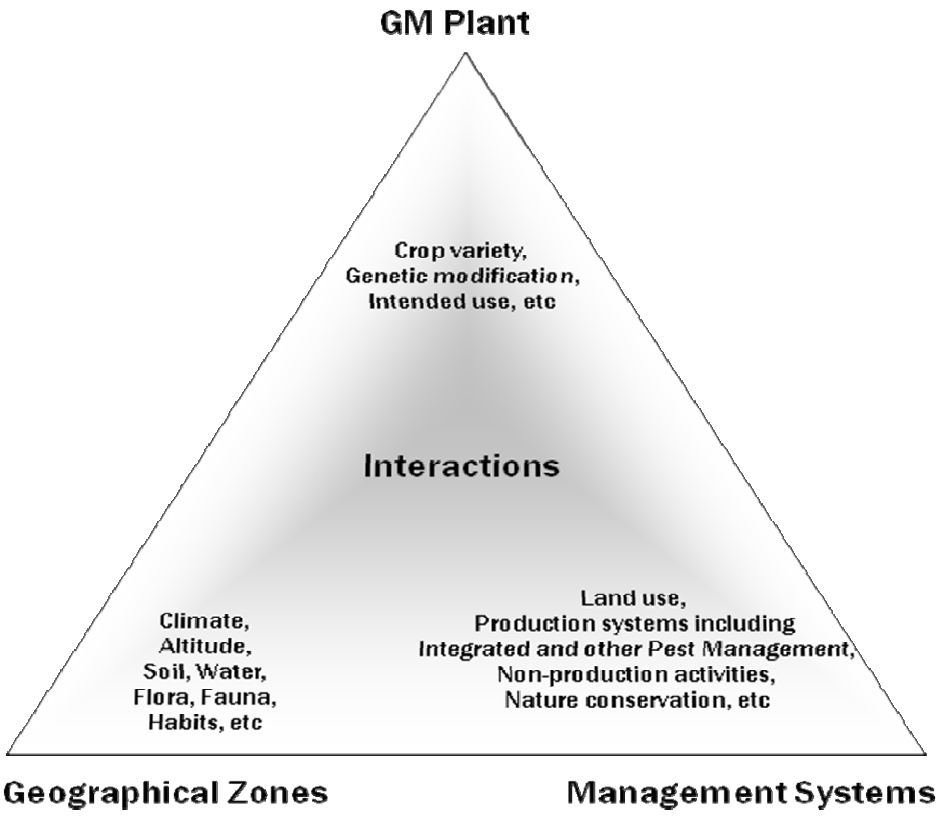

Figure 2: The receiving environment(s) is characterised by (A) the GM plant including its intended use(s), (B) the Geographical Zones, and (C) the Management Systems. Examples of attributes of (A), (B), and (C) that could interact are provided in the triangle.

In the component "Management Systems", land use and production systems shall be considered as these systems can differ significantly within and between geographical regions (e.g. irrigated maize versus non-irrigated cultivation). Moreover, in a specific region, cultivation of GM plants for different purposes may have specific risk assessment implications (e.g. green maize for biogas or silage with early harvest compared to grain maize).

The three components listed above result in biotic and abiotic interactions that shall be considered by the applicants when establishing representative scenarios considering receiving environments for carrying out the ERA of a GM plant (Figure 2 and Table 3). A broad range of environments in terms of fauna and flora, climatic conditions, habitat composition and ecosystem functions and human interventions occurs in EU. Accordingly, GM plants will potentially interact with those differing environments.

The ERA shall be carried out on a case-by-case basis, meaning that the required information varies depending on the types of the GM plants and trait(s) concerned, their intended use(s) and the potential receiving environment(s). There may be a broad range of environmental characteristics (regional-specific) to be taken into account. To support a case-by-case assessment, it may be useful to classify regional data, reflecting aspects of the receiving environment(s) relevant to the GM plant (e.g. botanical data on the occurrence of wild relatives of GM plants in different agricultural or (semi)natural habitats of Europe, or effects of production systems on the interactions between the GM plant and the environment)).

Applicants shall take into account the potential risk implications of the presence of any other GM plants that have been placed on the market in the same receiving environments, including interactions between the specific cultivation characteristics (e.g. use of plant protection products) associated with the different GM plants. In addition, applicants shall consider likely and/or predicted trends and changes to receiving environments, and how these might interact with the GM plants. For example, the spread of invasive maize pest (Western Corn Rootworm, *Diabrotica virgifera virgifera* LeConte) in Europe and its population build up in maize cultivation areas will result in changed cultivation practice, insecticide use and may increase the likelihood of adoption of rootworm resistant maize hybrids.

Relevant baseline(s) of the receiving environment(s), including production systems, indigenous biota and their interactions, should be established to identify any potentially (harmful) characteristics of the GM plant. Relevant baselines refer to current production systems for which generally published literature is available. These baseline(s) serve as a point of reference against which future changes can be compared. The baseline(s) will depend to a considerable extent on the receiving environment(s), including biotic and abiotic factors (for example, natural preserved habitats, agricultural farmland or contaminated land).

1.3.3. Potential cultivation areas of GM plants in the EU receiving environments

The receiving environment will generally include the environment where the specific plant (species) has already been cultivated, but may also include areas where the new traits will allow cultivation outside of former cultivation areas (e.g. for GM plants with tolerance to abiotic and biotic environmental stresses or providing new economic benefits). Thus both the plant and the transgenic trait(s) determine where the GM plant will most likely be grown. Some GM plants (e.g. cotton, rice) can realistically be cultivated in some geographical zones only, while others, like maize, may be cultivated more widely in Europe. Transgenic traits such as biotic (e.g. pest resistance) and abiotic (e.g. drought and salt) stress tolerance will also determine where GM plants are likely to be grown. Therefore, all these elements shall be taken into account when defining the receiving environment(s) (e.g. considering geographical zones) for the ERA of each GM plant.

In addition, in relation to NTOs, the composition of species assemblages in different receiving environments of the GM plant needs to be considered. Also farming and cultivation practices within receiving environments might have to be considered: these practices can differ significantly between regions (e.g. irrigated maize versus non-irrigated cultivation elsewhere) and influence associated non-target biota.

There are many climatic, ecological, agricultural and political ways of defining geographical regions or zones in Europe and examples of existing definitions are provided in Section 1.3.4. The variety of the methods and criteria, used to define these zones, reflects the diversity and multivariate nature of the characteristics of potential receiving environments of a GM plant. In some cases, such methods may assist applicants to select study sites. However, applicants should also consider selecting sites, where the exposure and impacts are expected to be highest, and where it is anticipated that if effects exist they will be detected. Applicants shall explain why the results of their studies in certain receiving environments are considered representative for other receiving environment(s).

According to the EFSA GMO Panel guidance document (EFSA, 2006a), '*environmental risk assessments should be carried out for each of the different environmental compartments that are likely to be exposed to the GM plant'*. In the case of arthropods, the receiving environment consists of (1) the managed terrestrial ecosystem (e.g. agro-ecosystem) including the GM cultivated fields, orchards, plantations and their margins, and (2) the wider environment (e.g. other adjacent GM or non-GM cultivations and non-cultivated habitats) and, where relevant, aquatic ecosystems.

An example: while arthropod assemblages including NTOs are generally similar over various EU maize cultivation areas due to available food source (maize, weeds, herbivore preys and hosts) and perform similar ecological functions, some important differences occur. For example, acariphagous coccinellid *Stethorus punctillum* and herbivorous leaf beetles (*Phyllotreta* species) are present or abundant in warmer regions in Europe while absent in others. Therefore, selection of representative cultivation areas shall consider these differences in distribution of possible focal species. Arthropod assemblages, inhabiting areas outside agricultural fields, potentially interacting with GM plants might be more region- or zone-specific. Applicants shall therefore consider these assemblages and/or relevant species of these assemblages when selecting field sites. The case-by-case approach would cover the heterogeneity of zones outside the field.

1.3.4. Geographical zoning concepts

There are different zoning concepts in EU defined for various purposes, some of which could be considered. However no single zoning concept could be used for all purposes and instead, a flexible and case-by-case approach is advised (see in Table 2 and Appendix I-Road map). The following zoning concepts might be considered in the framework of ERA of GM plants:

a) Plant protection product registration-based zoning;

b) Phytogeographic zoning;

c) Natura 2000;

d) SEAMLESS zoning approach;

e) LANMAP.

a) According to the new Regulation on Plant Protection Products (EC, 2009a), concerning the **placing of plant protection products** (PPP) on the market (and repealing Council Directives 79/117/EEC and 91/414/EEC), approvals for these products would be granted by geographical zones in the EU. Zones are defined in this Regulation as areas where agricultural, plant health and environmental (including climatic) conditions are comparable. For this purpose, 3 geographical zones have been defined to cover Europe:

- Zone A = North: Denmark, Estonia, Latvia, Lithuania, Finland and Sweden;
- Zone B = Centre: Belgium, Czech Republic, Germany, Ireland, Luxembourg, Hungary, the Netherlands, Austria, Poland, Romania, Slovenia, Slovakia and the United Kingdom;
- Zone C = South: Bulgaria, Greece, Spain, France, Italy, Cyprus, Malta and Portugal.

The zoning for placing plant protection products on the market offers valuable information for applicants on registered pesticides in 3 zones in Europe that could be considered when ERA of GM plants are put in cultivation, production practice context.

b) Phytogeographic zoning subdivides the Circumboreal region (Eurasia and North America) into a number of floristic provinces (Takhtajan, 1986). Three provinces (Atlantic, Central European, Illyrian) cover vast majority of the area of Member States within the European Union. However, some parts of several Member States belong to more than one province (such as Arctic, Euxinian, Eastern European, Northern European provinces). This zoning defines provinces based on distribution of plant species.

However, this zoning focuses on floristic characteristics but does not cover and reflect differences arising from agricultural activities or cultivation.

c) Natura 2000 is a network of special protected areas within the European Union. It covers the special areas of conservation under the Habitats Directive and special protection areas under the Birds Directive. Natura 2000 areas are areas of importance to the Community that have been designated by Member States of the EU. The Natura 2000 concept proposes 9 biogeographical regions (across 27 Member States) for covering the European ecological diversity. These biogeographical units are differentiated into the following regions: Alpine; Atlantic; Black Sea; Boreal; Continental; Macaronesian; Mediterranean; Pannonian; and Steppic Region. However, this zoning focuses on habitat and species protection, but does not cover and reflect differences arising from agricultural activities or cultivation.

The Indicative Map of European Biogeographical Regions (Natura 2000 or its modified version), was developed with the purpose of defining in practice the biogeographical regions mentioned in Art.1 c) (iii) of Council Directive 92/43/EEC of 21 May 1992 on the conservation of natural habitats and wild fauna and flora (EC, 1992; Evans, 2006). Consequently this map and the information on 9 biogeographical regions may offer specific information for the ERA (e.g. on natural habitats, wild fauna and flora, endangered, vulnerable, rare and endemic species) but are not related to plants being cultivated and associated NTO arthropod assemblages in production systems in Europe.

d) SEAMLESS zoning approach:

Since GM plant deployment will primarily (but not exclusively) be linked to agricultural areas and within these to arable fields and any nearby semi-natural habitats, an agriculture land use oriented zoning that also considers the differing regional quality for crop plant cultivation would reflect a more practical approach for GM plant risk assessments. Therefore, a zonation like that recently proposed in the SEAMLESS research project (van Ittersum et al., 2008) could be more relevant for the ERA of GM plant for cultivation purpose. Undoubtedly, geographical zones based on a scientific rationale, taking into account realistic agricultural situations, will be more valuable for GM risk assessment.

In the agriculture-environment oriented zonation generated by the SEAMLESS project, the current EU territory is – based on upscaling of farm-scale and physical data by a statistical approach – differentiated into 12 environmental zones. According to this approach, EU territory is differentiated into the following Zones: Boreal; Nemoral; northern Atlantic; central Atlantic; Lusitanian; Continental; Pannonian; northern Alpine; southern Alpine; mountainous Mediterranea; northern Mediterranea; and southern Mediterranea.

The selected regions according to this zonation are characterised not just in relation to climatic factors, but also to agricultural aspects such as preferred farming types.

Each of the above examples for various zonings reflects the main goal (e.g. flora, habitat and species conservation, economically relevant pesticide approval, more integrated agricultural and environmental modelling system) of the given zonation but can not be automatically transferred to the risk assessment of GM plants. For instance, the continental zone – also according to the SEAMLESS zonation – covers about 20 % of the EU surface, but even within these biogeographical (or agro-ecological) regions, there will be significant differences in NTO species, habitat characteristics, flora and fauna of these habitats.

e) LANMAP: Mücher et al. (2010) have developed a new hierarchical European Landscape Classification that can be used as a framework for, e.g., indicator reporting and environmental sampling. Landscapes are ecological meaningful units where many processes and components

interact. And as such, landscapes themselves have resulted from long-term interactions of natural abiotic, biotic and anthropogenic processes. The authors argue that a good understanding of landscapes is essential for its assessment, protection, management and planning. An internationally consistent approach is proposed to be obligatory and the production of landscape classifications and associated maps should be an important tool in this context. Although intuitive maps are available there are no consistent quantitative maps of European landscapes. In Mücher et al. (2010), landscapes are regarded as forming recognizable parts of the earth's surface and as showing a characteristic ordering of elements. It is argued that the complex nature of the underlying scientific concepts, which sometimes overlap and conflict, requires an objective and consistent methodology. As there are many regional differences in landscape properties, it is crucial to strike the right balance between reducing the inherent complexity and maintaining an adequate level of detail. Against this background, a European Landscape Map (LANMAP) has been produced, making use of available segmentation and classification techniques on high-resolution spatial data sets. LANMAP is a landscape classification of Pan-Europe with four hierarchical levels; using digital data on climate, altitude, parent (geologic) material and land use as determinant factors; and has as many as 350 landscape types at the most detailed level. According to Mücher et al. (2010), LANMAP is thus far limited to a biophysical approach, since there is a lack of consistent and European-wide data on cultural–historical factors.

The SEAMLESS zoning approach (12 zones) and the LANMAP (European Landscape Map) classification are more detailed and provide more complex maps (the latter uses as many as 350 landscape types). Both could be relevant information sources in specific cases of ERA of GM plants or in long-term monitoring programs.

Each of the above examples for various zonings reflects the main goal (e.g. habitat and species conservation, economically relevant pesticide approval, more integrated agricultural and environmental modelling system) of the given zonation, but can not be automatically transferred to the risk assessment of GM plants. For instance, the continental zone – also according to the SEAMLESS zonation – covers about 20% of the EU surface, but even within these biogeographical (or agro-ecological) regions there will be significant differences in NTO species, habitat characteristics, flora and fauna of these habitats.

Nevertheless, the zoning systems can provide useful supplementary information on the relatedness of meteorological and agricultural conditions and therefore the types of flora and fauna that may be associated with GM plants in these regions, assisting in identification of areas for conducting field trials.

1.3.5. Conclusion and guidance to applicants

The ERA shall consider interactions between the GM plant, its trait(s) and its receiving environment(s) which are identified in more details in Table 2 hereunder. Cultivation areas may cover one or more regions or zones in EU.

Table 2: Schematic steps for the selection process of relevant areas for field tests for effects on NTOs

Step 1 → Plant	Consider the present cultivation areas of the plant
Step 2 → Plant x Trait	Revise the present cultivation areas according to the nature of the trait: 1. add potential future cultivation areas, 2. according to the nature of the trait, concentrate on those areas where the plant is most likely to be grown.
Step 3 → Plant x Trait x NTO	Identify focal NTO guilds from all relevant functional groups in the production system
Step 4 → Plant x Trait x NTO x Region/zone[8]	Revise areas selected in Step 2 according to Step 3 considering levels of exposure and likely success of field studies so that proper RA can be conducted: 1. consider NTOs in adjacent habitats, 2. consider consequences of gene flow for potential secondary exposure, 3. consider management practices (including crop rotation and crop protection).
Step 5 → Final decision	Decide for area(s) for field tests according to requirements outlined in Section 1.8.

Applicants should initially consider representative scenarios, including a worst-case scenario where the exposure and impact are expected to be the highest. For the set of selected receiving environment(s), applicants shall describe:

- The characteristics of these receiving environments where the plant is likely to be distributed;
- The representative management systems (use of the plant, crop rotation, other GM plants, cultivation techniques);
- The range of relevant biotic and abiotic interactions (e.g. interactions between plants and NTOs, NTOs species assemblages) likely to occur in the receiving environment(s) taking into consideration the range of natural environmental conditions, protection goals (including those related to species differences across Europe), and production systems. Where appropriate, the presence of cross-compatible wild/weedy relatives nearby, the ability of the GM plant to form feral populations and hence the potential impacts on the receiving environments should be considered.

Based on the criteria listed above, applicants shall provide evidence that data generated are representative of the range of receiving environment(s) where the crop will be grown in the EU, e.g. for the selection of field trial sites according to Section 1.8.5.

1.4. Assessment endpoints

Because protection goals are general concepts, they need to be translated into measurable assessment endpoints. Thus the assessment endpoint is an explicit expression of the environmental value that is to be protected. This necessitates defining (a) *species* and (b) *ecosystem functions* that could be adversely affected by the GM plant, and that require protection from harm. To allow regulatory decision-making, assessment endpoints should be defined by applicants as far as possible using measurable criteria relevant to the case under study, so that change in these endpoints can be identified. These endpoints are operationally defined by an ecological entity (e.g. a natural enemy species, a pollinator species, a species of conservation concern, a soil function) and its attributes of that entity (both ecological and socio-economic e.g. regulation of arthropod pest populations, pollination of plants, of conservation concern) that could potentially be impacted by the GM plant use and that require protection from harm (Suter, 2000). For example, some ladybirds species are predators of crop pests (e.g. aphids) providing ecological services and reproduction and age class structure are some of their important attributes. In this example, parameters such as mortality, reproduction and age class structure together may form an assessment endpoint.

From a practical point of view the species assemblage in a non GM production system shall be considered, specifically describing the functional groups active in these agro-ecosystems (see Appendix I-Road map). Among these functional groups, assessment endpoints need to be defined specifically considering protection goals. The relationship between assessment endpoints and protection goals shall be specified.

1.5. Limits of concern

Applicants should clearly place the work on NTO testing in the context of environmental damage. This would mean that damage should be measurable, the significance of any damage be defined, and its representativeness for the receiving environment be covered. Hence, once assessment endpoints have been set, the 'environmental' quality to be preserved is to be defined (limits/threshold of concern, trigger values, decision criteria), as it enables defining and identifying the level of difference between the GM plant and its conventional counterparts that may lead to harm and trigger regulatory concern (see Section 1.8 for further details). It is thus important to identify if potential impacts on NTOs result in environmental harm and are different from those related to the conventional counterpart.

In this scientific opinion, environmental damage is defined as a measurable adverse change in a natural resource (e.g. a protected species, ecosystem service or other environmental entity of conservational relevance), or as a measurable impairment of a natural resource service which may occur directly or indirectly. The following three principles merit attention in the context of assessment of impact on NTOs:

- 'Damage or harm' means a measurable adverse change. This definition has implications for the ERA in respect of practicality since there is a need to quantify the effects on NTOs (EC, 2004);

- The significance of any damage needs to be evaluated on a case-by-case basis particularly in relation to the population (size) and potential for recovery of the affected NTO. This definition has implications for the assessment of the magnitude of observed changes;

- Consequences for the receiving environments, protection goals and ecosystem services.

This process includes defining the magnitude and both the spatial and the temporal scales relevant for the entity and the attribute to be preserved. The magnitude should describe to what extent the 'environmental' quality should be preserved (or above what threshold a change would be considered a

disturbance in 'environmental' quality). The spatial and temporal scales are the habitats in which the 'environmental' quality and the period during which the 'environmental' quality should be preserved, respectively (Storkey et al., 2008; Sanvido et al., 2009).

The issue of selecting an 'appropriate' or 'acceptable' baseline level of biodiversity for any agro-ecosystem is widely debated. Logically, an 'acceptable' level of biodiversity needs to be defined in terms of a 'minimum' biodiversity level for the efficient and sustainable functioning of the particular agro-ecosystem (i.e., providing essential 'ecosystem services', including biological control of pests and diseases, nutrient fixing and cycling, decomposing plant materials, maintenance of soil quality and fertility and structural stability). Once this level of biodiversity has been set for a particular agro-ecosystem, it should then be possible to design studies to determine whether a GM plant will maintain this required or desired level of biodiversity. The required level of biodiversity in a particular agricultural system is often subjective and a cultural response in a human-managed habitat, rather than a basic and definitive biological measure. Since agro-ecosystems are heavily human-modified environments, it is logical to expect biodiversity levels to depend upon how that agro-ecosystem is managed. They will therefore vary from region to region, from Member State to Member State, and from season to season, depending upon many parameters (e.g. nature of the particular environment, farming system, weed pressure, soil type and climatic conditions). Agro-ecosystems comprise crop areas, field margins and other semi-natural habitats that may be utilized by NTOs in several ways. It is therefore important that the ERA takes into account the possible threats to biodiversity within the agro-ecosystems and in the surrounding habitats, particularly considering the possible implications for protected areas and natural habitats that might be in proximity of cropping areas.

The applicants shall select assessment endpoints and define limits of concern based on a well-defined problem formulation that facilitates a scientifically sound study in relation to environmental protection goals. These limits of concern are specific for each assessment endpoint and need to be related to significant effects at the population level.

1.6. Conceptual model

Once assessment endpoints have been extracted from environmental protection goals (e.g. ecological functions and the representative non-target species for those functions have been selected), a methodology is required that will enable risk characterisation and the production of relevant information for regulatory decision-making. This is generally done on the basis of a conceptual model and an analysis plan (EPA US, 1998; Hill and Sendashonga, 2003; Wilkinson et al., 2003; Hilbeck et al. 2004; Raybould, 2006, 2007a,b; Nelson and Banker, 2007; Andow et al., 2008; Nickson, 2008; Romeis et al., 2008; Storkey et al., 2008; Wolt et al., 2010).

The conceptual model shall describe all relevant exposure scenarios of how harm to the assessment endpoint may arise from the GM plant in a way that allows for a characterisation of risks. Therefore, key interactions between the GM plant, the assessment endpoints and pathways of exposure through which the GM plant may affect them either directly or indirectly (= exposure profile), and potential impact of the GM plant to the environment (e.g. Wolt et al., 2010) need to be described.

1.6.1. Exposure profiles

The conceptual model shall include the available information on the nature of the stressor, its intended use(s) (including the intended scale of cultivation), exposure routes or profiles (see Section 2.3), and potential responses of the assessment endpoint as a result of exposure. A GM plant introduces additional potential stressors into the environment: the transgene in an organismal context, its products and the GM plant itself. In this respect, the scope of GM plant market authorisation applications is

relevant as it can cover different intended uses such as import, processing, food, feed and cultivation. The level of exposure of NTOs to the GM plant will depend on the intended uses of a GM plant:

- In cases where the application does not include cultivation in the EU, direct environmental exposure of NTOs to the GM plant is via the accidental release into the environment of seeds or propagules during transportation and processing. This may result in sporadic occurrence of feral GM plants and therefore exposure of NTO populations is likely to be negligible. The ERA will then focus on indirect exposure to products of the GM plant (e.g. through manure and faeces from animals fed the GM plant, and other by-products of industrial processes).

- In cases where the application includes cultivation in the EU, the level of environmental exposure is estimated on a case-by-case basis depending upon several factors. These include the biological and ecological characteristics of the GM plant and its transgene(s), the range of expected scales and frequencies of GM plant use, the receiving environment(s) where the GM plant is likely to be cultivated, and the interactions among these factors (e.g. Andow and Zwahlen, 2006; EFSA, 2006a; Garcia-Alonso et al., 2006).

1.6.2. Hypotheses

A well-structured conceptual model should allow the identification and formulation of relevant and testable hypotheses that arise from the consideration of potentially significant risks. These hypotheses are necessary to make assumptions and predictions about how a stressor could affect an assessment endpoint. Within the analysis phase of the ERA, several hypotheses may be plausible, then testable hypotheses are translated into one or more rigorous statistical hypotheses which are amenable to testing and corroboration.

The information considered to formulate testable hypotheses can take many forms starting from the data generated during product development by applicants, and including published scientific literature, scientific and expert opinions, stakeholders' deliberations and experience gained from other similar GMOs. This information should summarise existing knowledge of the system (plant-plant products-environment-hazard-exposure) under study.

If possible, appropriate risk formulations to be considered in the risk characterisation should be established in the analysis plan by describing the way the exposure measurement relates to the hazard measurement (Wolt et al., 2010).

1.7. Analysis plan

The last step of the problem formulation comprises an analysis plan in which decisions should be made about the most appropriate ways to measure the response of each assessment endpoint of the GM plant. In this planning phase, data needed and the approach to be taken for data acquisition and synthesis are delineated in order to test hypotheses formulated in the conceptual model.

Realistic scenarios should be placed in the context of an analysis plan by describing and selecting (1) the various measures to be used in the assessment and subsequent risk characterisation; and through the description of (2) methods and criteria of measurement.

The selection and prioritizing of both measures to be used and testing needed should help to focus on data relevant for risk characterisation (Raybould, 2006) and help the allocation of human and financial resources in a proper way (Qi et al., 2008). A properly constructed analysis plan, based on a conceptual model that is clearly linked to assessment endpoints, will guide the collection of data that are relevant to demonstrate the safety of GM plant use. The EFSA GMO Panel considers that this

approach makes the risk assessment process comprehensive and transparent by explicitly stating significant assumptions underlying the ERA.

1.7.1. Species selection

In any ecosystem, there is usually a high number of NTO species that may be exposed to GM plants. Considering that not each of these species can be tested, a representative subset of NTO species (referred to as 'focal species') shall be selected, on a case-by-case basis, for consideration in the risk assessment of each GM plant. This selection of species is generally based on several criteria (e.g. Birch et al., 2004; Todd et al., 2008), including the ecological relevance of the species, sensitivity to known or potential stressors, anthropocentric value, testability, exposure pathways (e.g. predators and parasitoids through preys and hosts).

The number and type of species to be tested will depend upon the hypotheses generated in the conceptual model. Therefore, NTO testing shall start with a clear problem formulation to enable the development of decision trees for species selection.

1.7.1.1. NTO species selection approach

There are several criteria suggested for species selection to conduct ERA for GM plants in the scientific literature.

For example, selecting species that are representative of their genera and/or of particular functional groups (including herbivores, pollinators of cultivated and wild plants, predators and parasitoids of pest organisms and decomposers of plant material) and that can be tested under laboratory and/or field conditions (Romeis et al., 2008) . These species are considered 'surrogate species'. Surrogate species are selected on the basis of their exposure to the environmental stressor. For instance, if the Bt protein is expressed in pollen, honeybees (*Apis mellifera*) would be considered a useful surrogate for pollinator taxa and related bee species. They are themselves present and economically important in many different crops and regions and thus are a useful combination of surrogate and key species (see below). More specific, crop-associated species may be selected that represent an important genus (e.g. *Orius* spp.), and other taxa may be selected that are broadly representative of whole families (e.g. *Aphidius* spp.) that are known to be important for ecosystem services. The pest species that are screened for their sensitivity to the insecticidal protein during product development can also serve as surrogates for NTO's. The familiarity with the species to be selected as a laboratory organism is deemed important in this approach. In addition, the problem formulation may consider species of anthropocentric significance, including those with special aesthetic or cultural value (e.g. the peacock butterfly, *Inachis io*) or species classified as threatened or endangered. The concept of using surrogates is widely applied in regulatory toxicity testing, in monitoring effects of environmental pollutants and in conservation biology to indicate the extent of anthropogenic influences, to monitor population changes of other species and to locate areas of high biodiversity. Species selection would normally prioritize the functional role of these taxa, so that conclusions from the risk assessment address important processes and are broadly applicable.

The concept of selecting species that belong to certain functional groups and/or are of anthropocentric value was also an important component of the concepts presented by Andow et al. (2006b). A 'key species' selection process was developed in a stepwise[9] 'ecological approach' to address effects on NTOs and biological diversity. The essential components of the "ecological approach" include the following steps: a) a risk endpoint selection process, b) a process relying on hypotheses to guide the characterisation of exposure, adverse effects and risk, and c) a transparent prioritization of the selected

[9] Originally the word 'tiering' was used.

species based on ecological characteristics for the specific system. Significant properties of the methodology include: a) it uses all available scientific information, including locally available expertise; b) it relies first on qualitative information and methods and proceeds to quantitative approaches only as necessary, c) it is structured in a way to overcome the lack of information, specifically addressing uncertainties and d) it considers the special needs of regions with high biodiversity. This stepwise selection is used to filter out approximately 6 to 10 key species for GM plant/trait combinations (which may include species of aesthetic or cultural value as well), which should then be examined in more detail.

The EFSA GMO Panel recognises that in the various existing approaches there are useful suggestions for NTO testing procedures (see EFSA, 2008). However in elaborating this document, some new elements are proposed, for instance the document does not use the terms 'surrogate' or 'key' species but instead introduces the term 'focal species' for the selection of test species (see section 1.7.1.2).

1.7.1.2. Guidance for selection of test species ('focal species')

The EFSA GMO Panel recommends a species selection process for the applicants as outlined below (see also Appendix I-Road map). This process includes four steps that lead applicants to a decision on the focal NTO species to be used:

Step 1 - Identification of functional groups:

As a first step in species selection, it is necessary to identify the ecosystem functions and services (including maintenance of herbivores as part of food web, pollination, regulation of arthropod pest populations by natural enemies and decomposition of plant material) provided by the production system (e.g. agro-ecosystem) and the functional groups of species involved, in the environment(s) where the GM plant is likely to be grown.

Step 2 - Categorisation of NTO species from identified functional groups:

In the second step, the main species linked to the functional groups identified in the previous step should be listed, considering the GM plant and the organisms associated to its receiving environment(s) (Birch et al., 2004; Hilbeck et al. 2006). An indicative list detailing the ecological role for common invertebrates in agro-ecosystems is provided in Table 3. Some taxonomically related species and/or life stages of the same species may have different ecological roles (e.g. different feeding habits) and this aspect should be considered.

Table 3: Examples of functional groups (exposure through trophic interactions)

Functional group		Examples of taxonomic groups
Herbivores		Phloem-feeders: aphids (*Hemiptera*: *Aphididae*), leafhoppers (e.g. *Hemiptera*: *Cicadellidae*), certain *Heteroptera*
		Cell-content feeders: thrips (*Thysanoptera*: *Thripidae*), spider mites (*Acarina*) and *Nematoda* (*Tylenchida*: *Meloidogynidae*)
		Chewing: leaf beetles (*Coleoptera*: *Chrysomelidae*), *Lepidoptera* larvae, *Diptera* larvae, grasshoppers (*Orthoptera Ensifera*), gastropods (*Mollusca*, *Gastropoda*)
Natural enemies	Predators	Beetles: *Coleoptera* (e.g. *Coccinellidae*, *Carabidae*, *Staphilinidae*)
		Predatory bugs: *Heteroptera* (e.g. *Nabidae*, *Anthocoridae*)
		Predatory flies: *Diptera* (e.g. *Syrphidae*)
		Lacewings: *Neuroptera* (e.g. *Chrysopidae*, *Hemerobidae*)
		Thrips: *Thysanoptera* (e.g. Aeolothrips)
		Spiders & harvestmen: *Araneae* and *Opiliones*
		Mites: Acarina (e.g. *Phytoseiidae*)
		Nematoda (e.g. *Mononchus* sp)
	Parasitoids	Hymenoptera (e.g. *Ichneumonidae*, *Braconidae*, *Aphelinidae*)
	Parasites & Pathogens	Bacteria, fungi, viruses
	Entomopathogenic organisms	*Nematoda* (e.g. *Heterorhabditidae*, *Steinernematidae*), pathogenic microorganisms
Pollinators		Solitary and social bees (*Hymenoptera*: *Apidae*), hover flies (*Diptera*: *Syrphidae*); *Coleoptera* (e.g. *Melyridae*, *Curculionidae*, *Scarabaeidae*)
Decomposers		*Diptera* larvae (e.g. *Phoridae*, *Sciaridae*), *Nematoda* (e.g. *Rhabditidae*, *Dorylaimidae*), springtails (*Collembola*), mites (*Acarina*), earthworms (*Haplotaxida*: *Lumbricidae*), *Isopoda*, microorganisms
Plant symbionts		Rhizobacteria, mycorrhiza

In the categorisation of relevant NTO species, additional species of economic, aesthetic or cultural value, or species of conservational importance considered as threatened or endangered may also need to be included.

Step 3 - Ranking species based on the ecological criteria:

From the list built in step 2 of species selection, applicants shall prioritize NTO species from each relevant functional group (Birch et al., 2004; Hilbeck et al. 2006).

The main criteria to be considered in this prioritization process are:

- Species exposure to the GM plant under field conditions, specifically considering life stages present during the period of exposure;
- Known sensitivity of the species to the product(s) expressed in the GM plant;
- Linkage to the production system (e.g. agro-ecosystem), and presence of alternative food source;
- Abundance;

- Interactions with target species (trophic and plant-mediated);
- Species vulnerability (i.e. are certain populations already threatened and thus more vulnerable to additional pressures?);
- Relevance to adjacent habitats, including natural and semi-natural habitats.

Step 4 - Final selection of focal species:

Based on the considerations addressed in the previous steps of species selection, a restricted number of focal species needs to be selected from each functional group. A theoretical framework for focal species selection is presented in Figure 3. At this stage, some practical criteria may be considered in the final selection of focal species. It may be that, among the prioritized species, some can be tested more effectively under laboratory conditions, or are more likely to be available in sufficient numbers in the field to give statistically meaningful results (Gathmann et al., 2006; Todd et al., 2008). Legal constraints may limit testing of certain NTOs (e.g. protected species), so this aspect may also influence the final choice of focal species.

It is expected that, at the end of the selection process, the applicants have selected at least one focal species from each relevant functional group identified in the problem formulation for further consideration in the ERA. Different possible sources of exposure for each focal species (in the most relevant developmental stages) to be tested should be considered in the focal species selection process.

The sensitivity of non-target herbivore populations exposed to the GM plant can indicate selectivity of the GM trait. In addition any changes in their populations can influence populations of other species (e.g. natural enemies). This information is useful in assessing consequences of the GM plant for the pest status of herbivores as well as for a wider range of species that may not be directly exposed to the GM plant and its target pests. For these reasons, herbivores should be considered when selecting possible focal species.

For field trials, estimation of ecosystem functions and services could complement or replace data on focal species. Ecological functions (such as pollination, biological control, soil functions[10]) depend on the number of species, their abundances and different types of assemblages. In a particular assemblage, the abundance of any species naturally fluctuates and the decline of a certain population might be compensated by another species within the same guild without adversely affecting functionality (Naranjo, 2005a,b). For example, the overall predation rate of a guild of predatory species could be selected as an assessment endpoint in field trials (Arpaia et al., 2009). Likewise, evaluating the earthworm community as a whole might provide data that are more ecologically relevant than measuring the effects on a single (focal) earthworm species.

Since soil is also a complex environment, a range of laboratory bioassays is used to produce reliable predictions of what could happen to the soil ecosystem. Such studies may be approached by using species sensitivity distributions and by requiring specific tests for the main functions and life-forms. Biological indicators in particular are often very dynamic and sensitive to changes in soil conditions. Consequently, they are often used as indicators of short-term changes in soil quality. Biological indicators include populations of micro-, meso- and macro-organisms, or the study of its structure using multiple endpoints that address both diversity and processes. In addition to using biological indicators, it is necessary to include effects on overall soil functions and the effects of functional groups as a whole. Measurements of ecosystem processes such as soil respiration, biomass decomposition or nutrient dynamics could be used to assess the effect of a GM plant on important soil functions. Soil respiration and other important indicators could be used as measurement endpoints for

[10] E.g. soil respiration, biomass decomposition, and nutrient dynamics.

soil microbial activity. Other measurement endpoints could include important soil properties, such as the changes in soil organic matter (SOM) and soil texture.

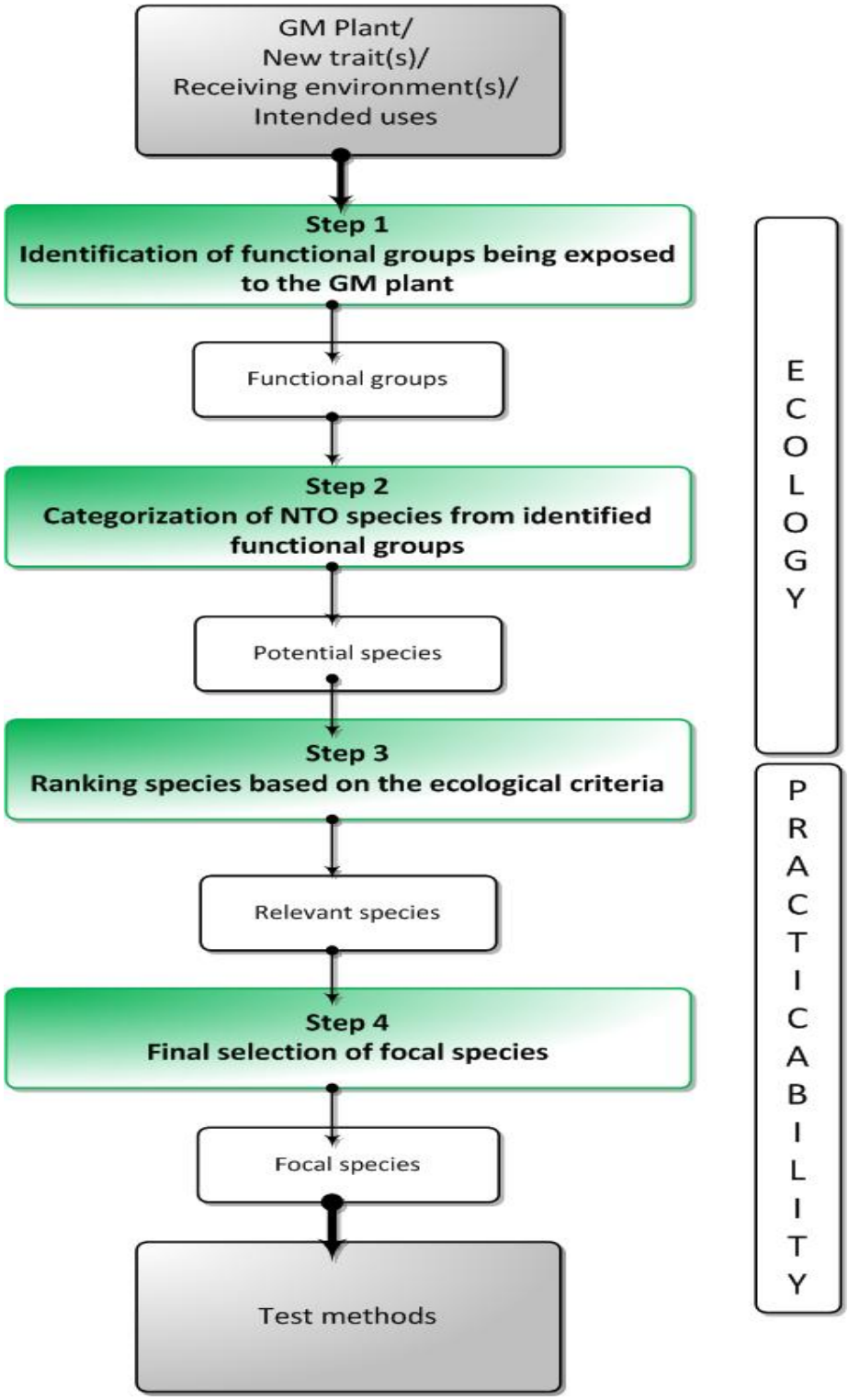

Figure 3: Four steps for selecting focal NTO species to be tested (modified after Hilbeck et al., 2008).

1.7.2. Definition of measurement endpoints

Through the formulated hypotheses, assessment endpoints are made operational into quantitatively measurable endpoints, termed measurement endpoints. Indicators of change, that will be recorded as part of the comparative risk assessment, need to be defined and established by applicants through measurement endpoints. These measurement endpoints should constitute measures to characterize both exposure and/or hazard, and shall be selected when there is an univocal interpretation of the biological data, *i.e.* how to relate the results to the assessment endpoint.

Both lethal and sublethal effects are relevant in the assessment of a possible hazard for a given NTO species. Testing for sublethal effects is important since it can give indications of possible long-term effects. An appropriate measurement endpoint for NTO testing is relative fitness (or some component of relative fitness), which is the relative lifetime survival and reproduction of the exposed versus unexposed non-target species (Birch et al., 2004). It is therefore important that NTO tests consider toxic effects (short-term mortality, longevity). In addition reproduction parameters (e.g. number and size of offspring, percentage of eggs hatching, sex ratio of progeny, age of sexual maturity), growth pattern, development rate and, when appropriate, behavioural characteristics (e.g. searching efficiency, predation rates, food choice) shall be considered. An alteration in plant metabolism could substantially affect components of the life history of organisms associated with these plants and consequently alter the growth of NTO populations (Charleston and Dicke, 2008).

The abundance and species richness of certain groups of NTOs at a relevant life-stage within a landscape or region are typical measurement endpoints. The choice of specific measurement endpoints shall be done according to the problem formulation on a case-by-case basis.

Long-term effects on NTOs populations or functional guilds are a substantial element of the ERA, meaning that, in the context of NTO testing, reproduction parameters and testing over multiple generations could be considered as appropriate endpoints. In addition modelling and/or post-market environmental monitoring can also be suitable methods for addressing potential long-term effects (see Section 2.3.4 of EFSA, 2010b).

Measures of hazard: Measures of hazard represent the measurable change of the measurement endpoint(s) in response to the GM plant and/or its products to which it is exposed (Storkey et al., 2008). Measures of hazard may be an acute lethal concentration resulting in the death of, e.g. 50% of the organisms tested or the effective response concentration for chronic effects measured or altered reproduction (e.g. fecundity), growth, development and behaviour in a receptor population (Wolt et al., 2010). These measurements can be expressed as effective concentration affecting a x percentage of individuals (EC_x). In addition, it is necessary to consider reproduction parameters (e.g. number and size of offspring, percentage of eggs hatching, age of sexual maturity), growth pattern, development rate and behavioural characteristics (e.g. searching efficiency, predation rates, food choice) may also be appropriate measures of hazard for long-term effects. At population level, an important predictor is the intrinsic rate of increase (r_m) that integrates measures of survivorship and fecundity (e.g. Romanow et al., 1991; Stark and Wennengren, 1995). Moreover, the calculation of the instantaneous rate of increase (r_i) allows a good estimate of r_m for the study of insect populations at lower tiers (Walthall and Stark, 1997).

Measures of exposure: Measures of exposure shall describe the contact or co-occurrence of the GM plant with the valued entity, and can be expressed as predicted (or estimated) environmental concentrations (PEC or EEC). The description of the novel attribute of the GM plant (e.g. transgenic protein) in terms of the route, frequency, duration, and intensity of exposure for the change relative to the valued entity is considered relevant information (Wolt et al., 2010). Both plant and NTO features assume an important role here, for instance overlapping of the NTO biology (e.g. life cycle stages) with the spatio-temporal concentration of the transgene are to be considered to quantify exposure. If a

non-target species is not directly exposed to the transgene from the plant but indirectly via other target or non-target species, these pathways of exposure need to be evaluated.

1.7.3. Hypotheses testing & Tiered approach

Any type of genetic modification of plants results in intended effects but may also result in unintended effects. The ERA is focused on the identification and characterisation of both effects with respect to possible adverse impacts on human and animal health and the environment. Effects can be direct and indirect, immediate and delayed, including cumulative long-term effects.

Intended effects are those that are designed to occur and which fulfil the original objectives of the genetic modification. Alterations in the phenotype may be identified through a comparative analysis of growth performance, yield, pest and disease resistance, *etc*. Intended alterations in the composition of a GM plant compared to its conventional counterpart, may be identified by measurements of single compounds. These effects may also inadvertently impact NTOs.

Unintended effects of the genetic transformation are considered to be consistent (non-transient) differences between the GM plant and its conventional counterpart, which go beyond the primary intended effect(s) of introducing the transgene(s). Since these unintended effects are event specific, applicants must supply data on the specific event. Sources of data that may reveal such effects are: Molecular characterisation, Compositional analysis, Agronomic and phenotypic characterisation and GM plant-environment interactions (see Section 1.7.3.2 below).

A case study approach describing how the GM plant may adversely affect NTOs or their ecological functions is proposed as outlined in Table 4. Based on plant-trait-NTO interactions, five possible cases can be foreseen. On one hand, GM plants may express new proteins/metabolites that have (Ia) toxic properties; (Ib) non-toxic properties; or (Ic) unknown toxicity. On the other hand, GM plants may have an intentionally altered composition, in which metabolic pathways known to affect NTO-plant relationships (e.g. glucosinolates in *Brassicaceae*, alkaloids in *Solanaceae*, lignin in trees) are altered (IIa), or not altered (IIb).

In all of those five cases, the metabolism and/or the composition of the GM plants may in addition be unintentionally altered as a consequence of the genetic modification in a way that could affect NTO-plant relationships (‘unintended effects’).The presence of unintended effects in GM plants can be due to different reasons (e.g. pleiotropic effects) and it is well documented in the scientific literature (BEETLE project, 2009)

Only in some of the five identified cases (i.e. Ia, Ic and IIa), can a specific hypothesis be formulated to assess plausible intended effects (e.g. a GM plant intentionally altered to produce biologically active compounds may produce the same effects on non-target species).

To test these hypotheses and thus assess possible adverse effects on NTOs, relevant data need to be supplied and considered by the applicants (see also Appendix I-Road map).

For the two remaining classes of GM plants, only the absence of possible unintended effects on NTOs needs to be demonstrated according to the principle described below.

Table 4: Identified cases and hypotheses testing

	GM plants expressing new proteins/metabolites with:			**GM plants with intentionally altered composition**	
	Toxic properties	Non-toxic properties	Unknown toxicity	Alteration of metabolic pathways known to affect NTO-plant relationships	No alteration of metabolic pathways known to affect NTO-plant relationships
	Ia	**Ib**	**Ic**	**IIa**	**IIb**
Possible effects of the transformation process	Intended and unintended	Unintended	Intended and unintended	Intended and unintended	Unintended
Could specific hypotheses be defined?	yes	No, see Section 1.7.3.2 below	yes	yes	No, see Section 1.7.3.2 below

1.7.3.1. Specific hypothesis-driven investigation

For case studies Ia, Ic, and IIa, specific hypotheses can be formulated and assessed (e.g. the new metabolite can be toxic to some non-target species, or the change in the metabolic pathway will possibly influence the plant's interactions with other organisms on various trophic levels) according to the flow chart illustrated in Figure 4.

Based on specific hypotheses, NTO risk assessment can be performed in a tiered manner; whereby, hazards are evaluated within different tiers that progress from worst-case scenario conditions framed in highly controlled laboratory environments to more realistic conditions in the field. Three main tiers can be used, which comprise experimental tests under controlled conditions (e.g. laboratory tests under tier 1a and 1b and semi-field tests under tier 2), and field tests (tier 3). Within a tier, all relevant data shall be gathered to assess whether there is sufficient information to conclude on the risk at that tier. In case no reliable risk conclusions can be drawn, further data might be needed. Decision of moving between tiers needs to be driven by trigger values. These values shall be set for the species under consideration taking into account the intrinsic toxicity (e.g. estimated by effective concentration (ECx) of the newly expressed products and the expected concentration in the plant), and the sensitivity of the NTO developmental stages (examples of trigger values for NTOs are provided in EPPO guidelines[11]).

Based on the experience with Cry toxins, tier 1 tests generally seem to represent useful predictors for results at higher tier tests (Duan et al., 2010) provided that designs include all ecologically relevant ways of exposure. When laboratory studies are performed, both *in vitro* and *in planta* tests (tiers 1a and 1b) should be done to reach a reliable risk conclusion after tier 1. Tier 1a testing is of crucial importance for the ERA if no or little data on the metabolites expressed by similar GM traits are available (e.g. Table 4: case Ic). Tier 1a tests require purified metabolites in the same form as expressed in the GM plant. Tier 1b complements the results obtained with purified metabolites as they give indications on possible interactions between plant compounds and reflect realistic exposure conditions through bioavailability. In fact, Duan et al. (2008) demonstrated that laboratory studies incorporating tri-trophic interactions of Cry1-expressing plants, herbivores and parasitoids were better

[11] http://archives.eppo.org/EPPOStandards/era.htm

correlated with the decreased field abundance of parasitoids than were direct exposure assays. Where purified metabolites are not available, only tier 1b studies shall be conducted using GM plant material that guarantees exposure to both transgene products and the plant. Likewise, it is possible that for some NTO focal species no reliable protocols for performing such experiments exist, in this case the applicants may perform this type of test on some focal species only. In all justified cases where testing on a lower tier is not appropriate (e.g. test organisms cannot be reared in the laboratory), applicants can perform tests at the next tier.

Some impacts on multi-trophic interactions and ecosystem functions may not be observed in tier 1 tests. Higher tier testing may therefore be needed on a case-by-case basis before decisions on the level of risks can be made. In particular, field testing is essential to investigate trait versus environment interactions when laboratory tests give reason to assume a possible adverse effect.

The NTO testing phase can be finalized when sufficient information is compiled to reject the tested hypotheses. Applicants, who conclude that further tests are not required, based on available information, are required to explain the rationale for this conclusion. If at any tier adverse effects are detected, a hazard characterisation is required to determine the biological relevance of these effects. Also, the use of more NTO species in the same functional group might help to clarify how common these adverse effects might be for the specific agro-ecosystem. In some cases it might necessary to go back to the problem formulation phase, to redefine a hypothesis and to design additional experiments to generate the data needed.

For stacked events not expressing biocidal compounds, if scientific knowledge does not indicate possibility of synergistic or antagonistic interactions between these compounds that may affect NTOs, then no specific testing is necessary.

Stacked events expressing at least one biocidal compound may have different adverse effects on NTOs than the single events due to synergistic, additive or antagonistic effects. Applicants shall perform studies (or providing existing data) with combined administration of proteins when the genetic modification results in the expression of two or more proteins in the GM plant. *In planta* tests with the stacked event shall be included in tier 1 studies.

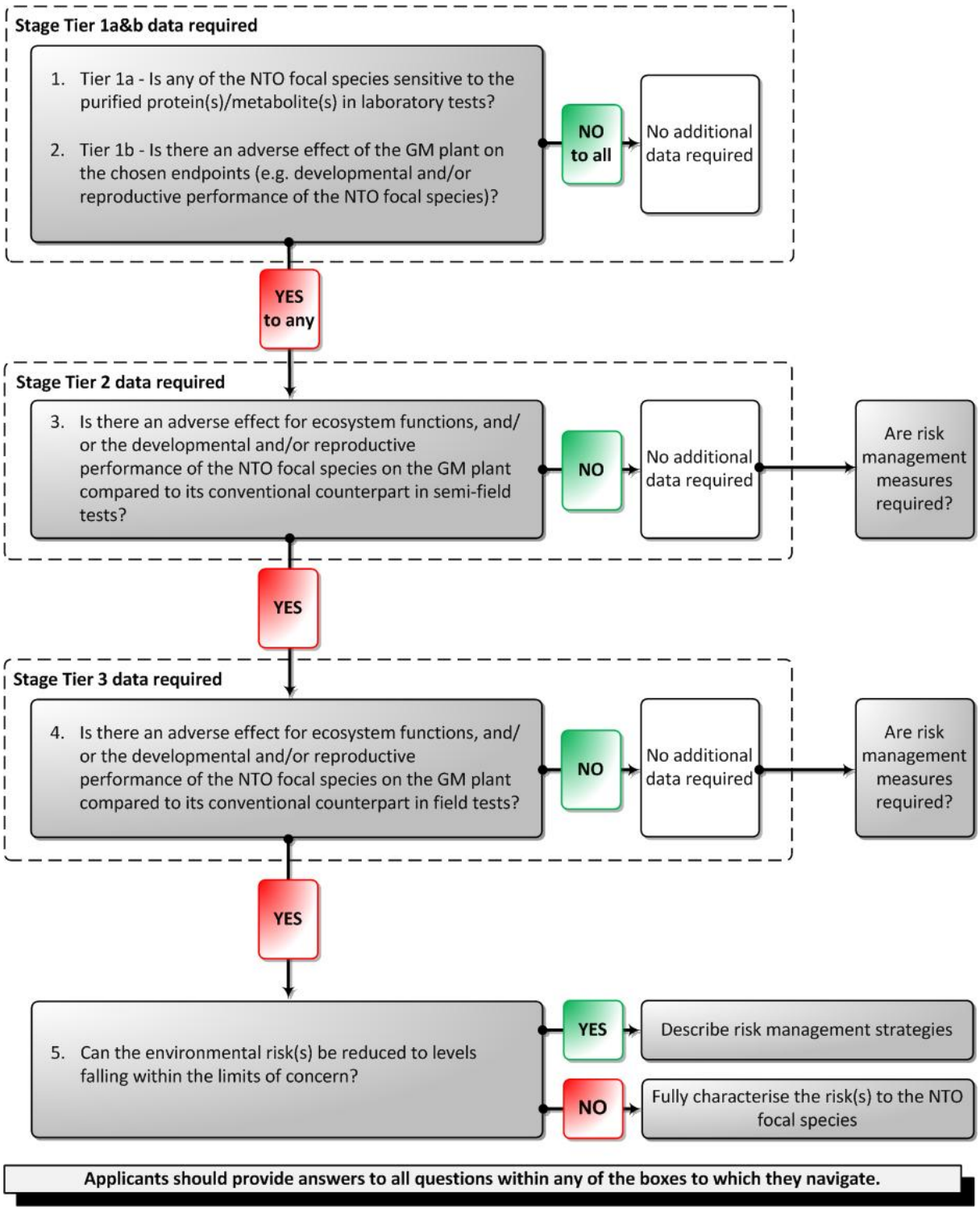

Figure 4: Decision tree for carrying out a specific-hypothesis driven investigation. Applicants shall provide answers to all questions within any of the boxes to which they navigate. The questions are divided into three stages (Tiers 1→2→3). Only if all the questions of a stage are answered negatively (answer: NO), are no additional data required. If at least one question of a stage is answered positively (answer: YES), applicants shall move to the next stage and address all the questions of that stage.

1.7.3.2. Data requirement for the evaluation of possible unintended effects

GM plants may have unintended adverse effects on biodiversity through interactions with populations of other species associated or sympatric with the GM plant. It is important that species richness and ecological functions, especially considering guilds that provide ecosystem services, are not disrupted to the extent that populations decline and/or vital functions are impaired. Unintended impacts of GM plants on species richness and ecological functions shall be considered in the ERA.

Problem formulation thus seeks to collect all available information to decrease uncertainty of unintended effects to an acceptable level. The evidence to exclude the likelihood of unintended effects on NTOs can come from numerous sources including data already collected for other parts of the risk assessment, collating all the appropriate information from these data sources to provide a weight-of-evidence approach. Data sources relative to plant-environment interactions (see point n° 4 hereunder) are always necessary to support the possible exclusion of unintended effects.

The possible sources of data are indicated below:

1. **Molecular characterisation** providing information on both the insert (e.g. promoters, insertion site including flanking regions) and any alterations to the genetic profile of the recipient plant. Data from the molecular characterisation can indicate whether there are general differences between the GM plant and its conventional counterpart. However, these data only provide valuable information about the occurrence of unintended differences between the GM plant and its conventional counterpart which may have an effect on GM plant-NTO interactions. In addition, data on the expression levels of transgenic products in different plant parts and stages that are exposed to NTOs are required (EPA, 1999; Nguyen & Jehle, 2007).

2. **Compositional analysis** is an important component of the comparative food/feed safety assessment that enables the identification of potential unintended effects on food/feed safety. An additional source of information in the frame of the NTOs assessment can be data from an **extended compositional analysis** which focuses on plant parts (e.g. pollen, nectar, leaves, stem, roots) that are consumed by NTOs, and which are not always considered under the food/feed safety assessment. Such an extended analysis can help to identify the likelihood of occurrence of unintended effects of GM plants that could affect NTO guilds and their functionality. An extended compositional analysis should in particular consider key components influencing the nutritional value for NTOs and secondary plant compounds relevant to plant defence mechanisms. As an example, glucosinolates and glycoalkaloids are important in regulating arthropod species assemblages on oilseed rape and potatoes and therefore a compositional analysis on leaves, stem and tubers may provide an indication of possible unintended effects on non-target herbivores. These crops are routinely analysed for these compounds and standardized protocols are available for testing the plants. It is considered important that identified compositional differences in the GM plant are discussed in an environmental context, specifically considering (a)biotic stresses affecting plant responses. However, if the extended compositional analysis indicates substantial differences and/or non-equivalence, further studies that are based on specifically formulated hypotheses are required to determine the biological relevance of these effects.

3. **Agronomic/phenotypic characterisation** carried out in specifically designed field trials may enable the detection of potential unintended differences between GM plants and their conventional counterparts e.g. differences in pest and disease occurrence (EFSA, 2009d). Notwithstanding the limited value of the agronomic/phenotypic field trials, mostly due to their plot size, for the assessment of possible unintended effects on NTOs, applicants may however use data from these field trials, which provide indirect indications deriving from general plant

characteristics, in order to inform conclusions as to the likelihood of unintended effects, e.g. about herbivore and disease associations with the GM plant.

4. **Plant-environment interactions** can be studied starting from tests carried out with *in planta* material in lab and field tests. In this respect, field-generated data, e.g. data related to NTO guilds and their functionality are a fundamental source of information in the majority of cases. Several aspects will need to be taken into consideration when obtaining these types of data from manipulative tests, field trials and/or field surveys (see Section 1.8 for further details).

The applicants are requested to consider all the information available from these different data sources and to ensure that some field generated data are included. The use of field-generated data from outside the EU may be informative in this context, but applicants must justify why these data are relevant to the ecological functionality of receiving environments in the EU where the GM plant will be grown. Since unintended effects are to a large extent event specific, data from other events or from similar events in other plant species will carry little weight in supporting an application.

Unintended impacts of the specific management and cultivation of GM plants are considered in Section 3.5 of the updated guidance document for the ERA of GM plants (EFSA, 2010b).

1.7.4. Design of protocols – Laboratory and field trials

Once specific measurement endpoints are chosen and given a priority, appropriate methods and criteria of measurement shall be selected and described in the analysis plan. This includes information on studies to be conducted, the appropriate tier for analysis, the design of experimental protocols with the definition of the appropriate statistical power (Marvier, 2002; Lovei and Arpaia, 2005; Romeis et al., 2008; Storkey et al., 2008; Perry et al., 2010).

1.7.4.1. Laboratory studies

Two kinds of methodologies are relevant for laboratory studies. First, existing conventional eco-toxicology methodologies (e.g. OECD, ISO, EPPO, IOBC standardized methods) can be used and adapted in order to assess the sensitivity of the NTO to different levels of exposure to the GM plant-produced proteins. The methodologies must be adapted to fulfil the measurement endpoint requirements. Secondly, an *in planta* experimental protocol is required in which the GM plant-NTO interactions are evaluated at exposure levels likely to occur in the field. For *in planta* studies, the testing scheme should ensure that the food used is ecologically relevant for the chosen NTO life stage to be tested (e.g. mimicking the trophic interactions existing in nature), and that specimens are exposed to the expected concentration throughout the study duration.

In addition to the above examples, several first tier studies that have been published in scientific literature can be considered by applicants.

All laboratory tests shall satisfy the following requirements:

- The endpoint and species are unequivocally identified;
- The rationale for the selection of the species and endpoint is given;
- Variability is sufficiently low for precise effect level estimation;
- Exposure to known quantities of testing material is maintained throughout the study;

- The experiment is conducted for a time span adequate to reliably estimate measurement endpoints.

When reproduction is an endpoint, the following requirements shall also be fulfilled:

- The processes of the reproductive biology must be included in the testing phase;
- The life-history must be known: age at maturation, duration of egg development, and instars subjected to exposure;
- Optimization of conditions for growth and reproduction must be provided by the test substrate and food supply.

Applicants can develop their own protocols for particular NTO species that are considered in the ERA. In this case, it is requested that, among others, the following aspects of the experimental protocols are correctly addressed:

- Organisms used during tests should be healthy and of similar age;
- The biological performance of organisms used as controls should be within acceptable limits (control mortality less than e.g. 20% depending on the testing system and organism);
- Environmental conditions in growth chambers, mesocosms and greenhouses should be described explicitly and justified;
- Plant material should be checked for transgene expression;
- Direct and indirect exposure pathways should be clearly identified in the experimental setup.

When designing experiments with natural enemies, the following additional requirements shall be considered:

- The suitability of artificial diet or surrogate host/prey species vs. natural food (e.g. some species do not grow well or do not reproduce when reared on artificial diet);
- Host/prey herbivores have to be properly exposed (possibly from hatching) to the right treatments;
- A uniform supply of prey/host quality, age, etc;
- The availability of additional food sources for species with mixed feeding habits (e.g. availability of pollen, honey or sugar solution, possibility for sucking from plants, etc.);
- The availability of an appropriate oviposition surface for predators;
- The provision of particular microhabitats (e.g. providing additional sources of water-saturated surfaces).

For tier 1a it is assumed that the test substance can be dosed and conventional testing approaches of chemicals can be followed. The sensitivity of the endpoint must be presented as EC10 and EC50 with confidence intervals. Laboratory practices (e.g. environmental conditions, specimen handling) should be according to standardized and published testing procedures. Limitations of some laboratory protocols should be considered (Lovei and Arpaia, 2005) when designing tests and concluding test results. When novel or non-standardised testing procedures are used, it shall be demonstrated that the method is appropriate, reproducible, reliable and of correct sensitivity.

BOX 1 Examples of standardized toxicity tests

OECD 220 Enchytraeid reproduction
OECD 222 Earthworm reproduction
OECD 232 Collembolan reproduction
OECD 226 Predatory mite reproduction
OECD guidance no. 56 Organic matter decomposition
OECD 216 Soil Microorganisms: Nitrogen Transformation Test
OECD 217 Soil Microorganisms: Carbon Transformation Test
OECD 213 Honeybees, Acute Oral Toxicity Test
OECD 214 Honeybees, Acute Contact Toxicity Test
ISO 15685 Potential ammonium oxidation
ISO 17155 Microbial soil respiration
ASTM E 2172. Nematode Caenorhabditis elegans
ISO 23753-1 Dehydrogenase activity
ISO 15952 Soil quality - Effects of pollutants on juvenile land snails (Helicidae)
IOBC recommended tests (Candolfi et al., 2000)
EPPO recommended tests[12]

In addition to above examples, several first tier studies have been conducted and are published that might be considered by applicants:

NTOs	Tier 1 studies/Publications
PREDATORS	
Coccinella septempunctata (Coleoptera: Coccinellidae)	Stacey et al. (2006)
Poecilus chalcites (Coleoptera: Carabidae)	Duan et al. (2006)
Chrysoperla carnea (Neuroptera: Chrysopidae)	Lawo & Romeis (2008); Li et al. (2010)
Chrysoperla carnea (Neuroptera: Chrysopidae)	Duan et al. (2010)
Orius insidiosus (Heteroptera: Anthocoridae)	
POLLINATORS	
Apis mellifera (Hymenoptera: Apidae)	Davidson et al. (1977); Brødsgaard et al. (1998); Rose et al. (2007)
Bombus spp. (Hymenopetra: Apidae)	Morandin & Winston (2003); Babendreier et al. (2008)
PARASITOIDS	
Aphidius ervi	Kramarz and Stark (2003)
Cotesia marginiventris	Ramirez-Romero et al. (2007)
Cotesia plutellae	Schuler et al. (2003)
NON-TARGET HERBIVORES	
NT herbivores in Bt cotton	Thi Thu Cuc et al. (2008)
Helix aspersa aspersa Müller (snail)	Kramarz et al. (2007)

The OECD tests for soil organisms cover several functional groups: e.g. predators, fungivores, detrivores, primary and secondary decomposers. However, for the above listed tests, assessing the impacts of GM plants might require significant modifications of the accepted procedures. For instance, OECD tests with honey bees were standardized considering the acute effects of a chemical product. Therefore, the test of acute contact toxicity is clearly not applicable for GM plants and even the acute

[12] http://archives.eppo.org/EPPOStandards/era.htm

oral toxicity test (OECD 213) may not be considered adequate due to the limited time span of the protocol.

The *in planta* testing required for tier 1b needs particular consideration of modifications of the standard procedures to allow for exposure to plant material. NTOs in tier 1b tests could be exposed to plant material through whole plants, plant parts (e.g. leaves, pollen) or ground plant material in diets or soil.

For *in planta* tests where feeding is an important route of exposure, it will not normally be possible to produce doses of the GM product that exceed the concentrations in plant tissues. Thus the normal level will act as the maximal exposure concentration in a test. Doses lower than the maximal dose can be made by dilution with a near-isogenic non-GM variety and EC10 and EC50 effect levels may be obtained. Different levels of exposure can also be achieved by mixing levels of GM plant material into the test substrates, e.g. soil, and a true dose-response relationship can be established delivering EC10 and EC50 effect levels. Appropriate controls for the effects of these diet regimes can be made by making similar mixtures with near isogenic non-GM materials.

In order to provide optimal nutrition in ecotoxicological tests particularly for some soil organisms, a food source may be added. The amount of additional food source may need to be adjusted in order to ensure worst-case exposure.

When the aim is to demonstrate equivalence of the GM plant to the appropriate comparator, the standard tests should include the appropriate comparator as a negative control at an exposure level identical to the GM plant, as well as a positive chemical control to prove the functionality of the experimental setup, as advised in the pesticide test guideline.

1.7.4.2. Field trials

Experimental complexity and variability increases from tier 1 (e.g. toxicological studies), to bi- and tritrophic studies with plant parts, bi- and tritrophic studies with whole plants, to field assemblage studies. Laboratory testing provides the best way to control and manipulate experimental conditions (environmental factors, set-up) and to limit complexity and variability. In contrast, field tests allow the evaluation of trait x environment interactions, but they exhibit the highest experimental complexity and provide the lowest ability to control experimental conditions due to large natural variability.

The objectives of field trials are:

- To identify and study exposure routes (including trophic relationships) and confirm observed effects in lower tier experiments;
- To provide feedback for further testing hypotheses;
- To study food chain and indirect effects;
- To determine effects of scale on NTO populations, including effects on generations and other spatio/temporal interactions;
- To study effects of interactions between several NTOs species in natural environment(s).

Field testing for NTOs is of special importance for certain species that cannot be tested in laboratory (e.g. rearing methods and experiences are not available). Field testing provides a broader range of arthropods in terms of species number, life stages, exposure to abiotic and biotic stress, complexity of trophic interactions, etc. that cannot be reproduced in the laboratory. Moreover, field experiments offer the opportunity to assess the functioning of ecosystem services in conditions of cultivation (e.g.

Naranjo, 2005a,b). On the other hand field experiments are characterised by a multiplicity of factors increasing the variability of the test system, which might reduce the sensitivity of the detection system. In consequence only large effects on the environment will be detected by field experiments. Hence, attention should be paid to the trade-off between standardised laboratory tests in lower tiers and the testing of NTO species in field experiments.

Due to the lack of well-defined standards, the number of NTO tests on decomposers is very limited. In relation to this, some biogeochemical processes cannot be mirrored in an artificial environment (such as pot experiments). Therefore, field tests may produce valuable results and important conclusions under more realistic conditions.

Design and analysis of field trials for NTOs should be performed according to the criteria explained in Section 1.8.

Closely linked to the analysis of field trials, the EFSA GMO Panel recognizes the importance of sampling methods. In this respect, the EFSA GMO Panel provides the applicants with examples of sampling methods that can be of use in the NTOs ERA:

- ISO 23611-1:2006 Soil quality -- Sampling of soil invertebrates -- Part 1: Hand-sorting and formalin extraction of earthworms
- ISO 23611-2:2006 Soil quality -- Sampling of soil invertebrates -- Part 2: Sampling and extraction of micro-arthropods (Collembola and Acarina)
- ISO 23611-3:2007 Soil quality -- Sampling of soil invertebrates -- Part 3: Sampling and soil extraction of enchytraeids
- ISO 23611-4:2007 Soil quality -- Sampling of soil invertebrates -- Part 4: Sampling, extraction and identification of soil-inhabiting nematodes
- ISO 10381-6:2009 Soil quality -- Sampling -- Part 6: Guidance on the collection, handling and storage of soil under aerobic conditions for the assessment of microbiological processes, biomass and diversity in the laboratory
- ISO/TS 10832:2009 Soil quality -- Effects of pollutants on mycorrhizal fungi -- Spore germination test
- ISO 7828-1985 Water quality -- Methods of biological sampling -- Guidance on handnet sampling of aquatic benthic macro-invertebrates

The importance of high quality field trials in the ERA of GM plants is widely accepted (Scientific Colloquium[13], EFSA (2008)). One crucial aspect is the increase in "ecological realism" that can be achieved as tests move from lab, to semi-field, and to field. NTOs will be in contact with GM plants in a multitrophic context and therefore the estimated impact on ecological functioning will be improved with the increasing scale of the experimental setup. On the other hand, it is well known that a lack of the control that is afforded by field trials and there is difficulty in establishing causal relationships in the field. This topic has clear implications with the concept of the receiving environment (see Section 1.3.) in which field trials should be tailored. As stated in the preamble section, the approach to ERA in the present document is to retain some flexibility. However, it is clear that the completion of studies

[13] http://www.efsa.europa.eu/EFSA/ScientificOpinionPublicationReport/EFSAScientificColloquiumReports/efsa_locale-1178620753812_EnvironmentalRiskAssessmentofGeneticallyModifiedPlants.htm

on NTOs at different scales guarantees the best description of the GM agro-ecosystems. Applicants should explain the rationale of their approach to risk assessment for non-target species according to the steps outlined in Section 1.7.1. For field trials conducted within the overall ERA framework, conceptual models should be developed which allow measurement endpoints to be identified.

➨Single species vs. multiple species analysis

Herbaceous agro-ecosystems host rather complex food webs, normally encompassing up to six trophic levels. Trophic relationships between arthropods are rather variable between seasons and include numerous species of herbivores, carnivores and omnivores (species with mixed diet regimes). Most field studies with GM plants were conducted with the aim of sampling one or more non-target species over time, to comparatively estimate their abundances. In most of these studies, results were obtained with typical univariate analyses for each taxon investigated, while the evaluation of multiple endpoints is more recent (e.g. Whitehouse et al., 2005 in cotton; Dively, 2005 in maize; Arpaia et al., 2007 in eggplant). The main goal of using multiple endpoints is the necessity of analysing simultaneously several variables considered to be somewhat related and potentially all important. It is quite possible that non significant results obtained from independent tests may become significant when data are analysed together, and this might occur because each single taxon contributes only for a little amount to the whole variability. Another main reason for trying to record data simultaneously is due to the fact that usually we have incomplete a priori knowledge of the trophic relationships between taxa in the field (see Section 1.8.).

A species-assemblage approach is aimed at evaluating overall ecological functions (e.g. natural pest control), rather than specifically concentrating on a given beneficial species. This approach may have benefits considering a) the lack of a recognized indicator species for evaluating the effects of Cry proteins expressed in plants (e.g. Arpaia et al., 2007), and b) the possibility of revealing unintended effects of GM plants on higher trophic levels in field studies where a large number of arthropod species are present at the third and fourth trophic levels. Using this approach, Naranjo (2005a) indicated that sentinel eggs and pupae of *Pectinophora gossypiella* (Saunders) in GM cotton fields experienced the same rates of predation in both Bt and non-Bt plots in a 5-years field study in spite of the fact that reductions in density of several predator taxa in Bt cotton were observed in a companion study (Naranjo, 2005b). Of particular concern is the fact that certain important ecological/functional groups are not tractable model species commonly used in laboratory tests (e.g. soil processes/species are not easily tractable but are ecologically essential).

➨Appropriate comparators and baselines

The choice of appropriate comparators is discussed in several documents (e.g. EFSA, 2007; EFSA, 2009c). This document will only deal with the specific issue of alternative pest control methods to be used when comparing results of field studies aimed at assessing non-target effects of GM plants. In a specific IOBC working group meeting, participants suggested that the "current agricultural practices" should be the base for comparative testing and monitoring of GM crops (Arpaia, 2004). Romeis et al. (2006) suggest that insecticide-treatments should be the main basis for comparison of risk, "unless other comparisons are of practical relevance". It should be emphasised that other comparisons may well acquire relevance in some contexts. Factors that may influence this determination include the policy goals of the regulatory authority or the potential users of the technology. For example, strategic goals for the adoption of certain pest management regimes, (e.g. Integrated Pest Management, biological control etc.), could be considered to be the appropriate basis for risk comparisons. In Europe for instance, starting from year 2014 Integrated Pest Management will become the only permitted approach to pest management in order to ensure the sustainability of pest control practices (EC, 2009b). Alternatively, the information on the potential adopters may determine the appropriate conventional comparison. For instance, in the case of cultivation of maize, there are very different production methods and so appropriate comparators should reflect the pest management at different

sites, areas or regions. In some situations, it would be wrong to assume that Bt crops would replace insecticide use entirely, as they may be used as complementary strategies. Indeed, in some Bt cotton and maize cropping systems, additional insecticides are used at low levels (e.g. Pemsl et al., 2005; Yang et al., 2005). The choice of the appropriate risk comparisons and these options need to be considered during the problem formulation phase at the beginning of the risk assessment, on a case-specific basis. The EFSA GMO Panel therefore recommends that comparative assessments be made with "current agricultural practices" for the conventional crop in the area of study, with special emphasis on pest control techniques adopted. Applicants must therefore specify and justify the choice of comparators deemed most relevant for the dossier (see Section 1.8.).

1.8. General statistical principles

This section applies to data collected from experiments in which specific hypotheses are tested. When such experiments are conducted in the field they are termed 'trials' throughout this section. This section does not apply to data obtained from surveys or observational data.

For ERA, the applicants shall list explicitly *in words* all the questions that each study, be it a field trial, a trial in semi-field conditions or a laboratory study, was designed to address. In addition, each of these questions shall be re-stated *in formal terms*, in the form of the precise null hypothesis that was tested to answer the question. This shall apply equally to those studies that seek confirmatory data on unintended effects when some evidence already exists, as to those that take an ecotoxicological approach with a specific null hypothesis. For field trials, the applicants shall provide a clear and explicit statement concerning the minimum levels of abundance[14] acceptable for each taxa sampled, below which results would lack credibility (for an example, see Heard *et al*. (2003), section 2F). Applicants shall supply justification for the values chosen. In mathematical modelling for the assessment of long-term or large-scale effects, the applicants shall state explicitly all assumptions made and provide justifications for each. The principles underlying the statistical tests of difference and equivalence (EFSA, 2009c) described below are to provide information with quantified uncertainty that may be used by biologists in risk characterisation of those endpoints for which differences or lack of equivalence are found. In order to place differences or lack of equivalence into context, allowance must be made for the distinction between statistical and biological significance. The two approaches are complementary: statistically significant differences may point to biological changes caused by the genetic modification, but these may or may not be relevant on safety grounds (see limits of concern, below). For risk assessment it is not the function of statistical analysis to provide results that lead automatically to a particular decision; instead, the case-by-case approach shall remain paramount.

The ERA is often hampered by the difficulty of conducting experiments with sufficient statistical power (see below). The use of meta-analysis (Marvier et al., 2007) is an option for applicants to consider, but is not mandatory. It may be useful to quantify studies that may not all have the power to be individually significant, in the statistical sense, and also to provide an overview of broad patterns when individual studies appear to contradict one other.

The comparative analysis referred to above shall involve two approaches: (i) a proof of difference, to verify whether the GM plant is different from its conventional counterpart(s) and might therefore be considered a potential risk depending on the type of the identified difference, extent and pattern of exposure; and (ii) a proof of equivalence to verify whether the GM plant is equivalent or not to its

[14] The following is an example of an appropriate statement concerning abundance: All species with mean abundance per site per occasion greater than 0.5 were analysed. Species within the genus *Pterostichus* that individually failed to meet this criterion were pooled and since the pooled mean abundance per site per occasion was greater than 0.5 these aggregate data were analysed in a category denoted as 'Other *Pterostichus*'. Very few individuals from the genera *Amara* and *Brachinus* were caught and these were therefore not analysed.

conventional counterpart(s) (Perry et al., 2009) within bounds defined by so-called 'limits of concern' (see below). For each measurement endpoint, the level of environmental protection to be preserved is expressed, directly or indirectly, through the setting of 'limits of concern' which may take one of two forms. For lower-tier studies (see Section 1.7) the limits of concern will usually be trigger values which, if exceeded, will usually lead to further studies at higher tiers. Then the relationship of the limits of concern to environmental protection goals is indirect. For higher-tier studies, especially field studies, the limits of concern shall reflect more directly the minimum ecological effects (in positive and negative directions) that are deemed biologically relevant. For field studies, at least one of the limits of concern shall represent the minimum effect that is considered by applicants potentially to lead to environmental harm (see also Section 2.3.3 of EFSA, 2010b). If this limit is exceeded then detailed quantitative modelling of exposure may be required to scale up adverse effects at the field level both temporally (to seasons, generations, rotations) and spatially (to farms, landscapes, regions and ecosystems) (EFSA, 2008). Baseline data can be used to define the limits of concern. Purely as a guide, for laboratory studies, a multiplicative effect size of 20% is often taken as the trigger value for further, higher-tier studies. Similarly, for semi-field testing, a trigger value of 30% has been used previously. For field studies, several studies, both in the USA and in the EU (Heard et al., 2003), have adopted 50% as a limit of concern, which is a reasonable level. By contrast, the effect size threshold for classification set by IUCN for butterflies is a reduction in population size of at least 30% over three generations (but here 'population' is defined at a larger than field scale). Note that, unless there is explicit justification, limits of concern for lower-tier studies shall usually be less than those for higher-tier studies, since it makes no sense for the results from laboratory studies to exclude from further study effects that might be manifest in the field. Whatever are the limits of concern adopted, applicants shall state their value and justify the choice explicitly, for each measurement endpoint. For field studies, it will usually be the lower limit, which might correspond for example to a decrease in the abundance of a particular species in the presence of the GM plant relative to that for the conventional counterpart, that will be defined as the threshold effect deemed to be of just sufficient magnitude to cause environmental harm. Notwithstanding this general approach, it is acknowledged that the multiplicity and diversity of questions that might be posed in an ERA may demand alternative statistical approaches, on a case-by-case basis.

All test materials, the GM plant and conventional counterpart(s), whether in the field, in semi-field conditions or in the laboratory, shall be fully randomised to the experimental units. Other aspects of experimental design are addressed below.

Whether analysis is of field, semi-field or laboratory data, results shall be presented in a clear format, using standardised scientific units. Applicants shall provide the raw data and the programming code used for the statistical analysis in an editable form. Other aspects of reporting and analysis are addressed below.

1.8.1. Testing for difference and equivalence

In testing for a difference the null hypothesis is that there is no difference between the GM plant and its conventional counterpart, against the alternative hypothesis that a difference exists. In testing for equivalence the null hypothesis is that there is lack of equivalence, in the sense that the difference between the GM plant and its conventional counterpart is at least as great as a specified minimum size, against the alternative hypothesis that there is no difference or a smaller difference than the specified minimum between the GM plant and its conventional counterpart. Rejection of the null hypothesis (*i.e.* a finding that the difference is no greater than this minimum size) is required in order to conclude that the GM plant and the conventional counterpart are unambiguously equivalent for the measurement endpoint considered. The two approaches are complementary: statistically significant differences may point to biological changes caused by the genetic modification, but these may or may not be relevant from the viewpoint of environmental harm. For studies that use extra comparators, the analysis shall encompass separate difference tests (between the GM plant and each of its different comparators) and

separate equivalence tests (between the GM plant and each of its different comparators), and these shall be reported similarly. Further discussion of the principles of equivalence testing, with practical examples, is given in EFSA (EFSA, 2009c).

1.8.2. Specification of the effect size and the limits of concern

Major parts of the risk assessment dossier are problem formulation and risk characterisation. Notwithstanding the well-known distinction between biological relevance and statistical significance (Perry, 1986) risk characterisation cannot be done without relating effects to potential harm. Therefore it is essential to specify for each effect variable a minimum effect size which is considered to potentially have a relevant impact on the receiving environment(s). Based on such effect sizes, power analyses aid transparency and may engender public confidence that risk to the consumer is well-defined and low (Marvier, 2002); these require specification of the magnitude of the effect size that the study is designed to detect. Good scientific studies are planned carefully enough for the experimenters to have a reasonable idea of the size of effect that the study is capable of detecting. For all these reasons, for each study, whether in the field, in semi-field conditions or in the laboratory, applicants shall state explicitly the size of the effect that it is desired to detect in the study, for each measured endpoint. The effect size may be asymmetric, and in particular may be set as zero in one direction to yield a non-inferiority form of the equivalence test (Laster and Johnson, 2003). The magnitude of the effect size that the study is designed to detect will generally be greater for trials designed to provide confirmatory field data for the assessment of unintended effects on non-target organisms than for specific hypotheses (see Section 1.7.3). The effect size will often be placed on the multiplicative scale; however, the natural scale or some other scales are admissible alternatives, on a case-by-case basis. In principle, where more than one comparator is used different effect sizes may be specified for the different comparators; however, this is unlikely to be necessary in practice. Applicants shall provide a full justification for all effect sizes chosen.

Applicants shall state explicitly how the chosen effect size(s) relates to the limits of concern through the minimum relevant ecological effect that is deemed biologically relevant. Usually, these quantities will be identical; applicants shall justify cases where this is not so. Applicants shall state explicitly the limits of concern that were used for each equivalence test. If justified appropriately, more than one pair of limits of concern may be set for each measurement endpoint; an equivalence test shall then be performed for each pair of limits.

1.8.3. Power analysis

For each study, be it a field trial, a trial in semi-field conditions or a laboratory study, applicants shall ensure that the design is such that the difference test has sufficient statistical power to provide reasonable evidence (Perry et al., 2009). Statistical power is the probability of detecting an effect of a given size, when such a real effect exists. In medical science, a level of 80% is usually considered to be an acceptable level for statistical power, but it is recognised that for ecological GM field trials the restriction on the land available for experimentation combined with unavoidable environmental heterogeneity usually necessitates some compromise between the replication required for high power and the experimental resources available (Perry et al., 2003). Notwithstanding, optimal experimental design shall be directed to attain power as high as possible.

For each study, applicants shall provide an analysis that estimates the power for each difference test on each measurement endpoint, based on the stated effect size and assuming a 5% type I error rate. The analysis shall be done at the planning stage of the study. The power analysis shall use only information verifiable as available prior to the study; under no circumstances shall data from the study itself be used. For field trials, since each field trial at a site on a particular occasion shall have sufficient replication to be able to yield a stand-alone analysis if required (see below), this power

analysis shall relate to a single site, not to the entire set of trials. For situations where many species are sampled such as in field trials, the power analysis is required only for those species of prime importance and those expected to be the most abundant.

1.8.4. Experimental environment

The first decision in conducting a study is whether the questions asked are best answered by data produced in the laboratory, mesocosm, semi-field, field or region.

Laboratory studies are used particularly in tier 1 studies (e.g. see Sections 1.7.3 and 1.7.4).

As is clear from Section 1.7.3.2, the effect of plant-environment interactions can be studied starting from studies that encompass a range of environmental scales. For this, hazards are evaluated within environments that progress from worst-case scenario conditions with laboratory experiments, up to ecological field trials with relatively large plots.

The laboratory environment is favoured for studies where it is important to control and define closely the conditions for tested organisms. Since environmental variability and interfering factors which can mask potential effects are minimised, laboratory studies yield results of relatively high precision. The laboratory environment is used particularly for the identification of acute and direct impacts of GM products and metabolites on individuals. In particular dose-response relationships may be well described. It also provides the possibility to study indirect and multi-trophic effects at small scales. Trait-environment interactions may be studied in the laboratory, but only to a limited extent. The laboratory is often used as an initial environment in the tiered approach, particularly for tier 1 studies (see Section 1.7.3). In a laboratory study, decisions must be made whether test materials should be of synthetic or *in-planta* form (see Sections 1.7.3 and 1.7.4).

Semi-field trials are manipulative test systems that are designed to control the inherent variability of the environment. They usually incorporate some form of protected environment or containment, such as field cages or screen houses, designed both to isolate the organisms under test and exclude unwanted biotic (e.g. predators) or non-biotic (e.g. rainfall) factors. Semi-field trials allow exposure to ambient weather and light conditions. The larger cages may result in more natural behavioural interactions between the organisms and plants tested. The semi-field environment is not subject to large variations in the ecology of habitats, and any variability due to different receiving environments is suppressed. Semi-field trials may have greater sensitivity than less-controlled open field trials and it may be that lower levels of statistically significant differences may therefore be detected. Examples include studies on possible indirect effects on non-target pollinators using bees in screen house trials. Mesocosms are experimental ecosystems that can be used to perform tests under realistic semi-field conditions. Examples include studies of biogeochemical cycles using residue decomposition (see Section 3.6 of EFSA, 2010b), although litterbag experiments within field trials provide a more realistic alternative.

Field trials allow the study of indirect and multi-trophic effects at larger scales, including at some cases the population level. Trait-environment interactions may be tested validly. Although they must, by definition, suffer from less ability to control environmental conditions and therefore yield results subject to greater environmental variability, they provide the only way in which lower-tiered results may be validated under natural conditions. They allow experimental tests of parameters of importance in ecosystem functioning (such as the predation and/or parasitism rate of a species, the decomposition rate of plant residues, etc.) and the estimation of overall ecosystem functions (such as pollination, natural pest control, etc). Another advantage of field trials is that genotype x environment interactions may be studied in the receiving environment(s).

Field surveys are scientifically designed studies without a hypothesis and where there is no experimental imposition of treatments. However, data are collected in the receiving environment(s). For example, these may provide appropriate data relevant to the identification of unintended effects on non-target organisms and to changes in plant fitness (see Section 3.1 of EFSA, 2010b).

The importance of field trials in the ERA of GM plants is widely accepted (EFSA, 2008). One crucial aspect is the increase in ecological realism that can be achieved as the scale of tests move up from laboratory through mesocosm to semi-field, field and region. For example, when any organism is in contact with a GM plant within a multi-trophic context, identification of the impacts on ecological functioning is facilitated by an increase of scale of the experimental arena.

Field testing for environmental effects of GM plants is of special importance because there are organisms for which particular ecological or behavioural tests in the laboratory fail to encompasss realistic conditions (for example in some studies of species that are highly mobile, such as adult butterflies or bees; or species for which rearing methods are inadequate; see Section 1.7.4). Field testing allows a wide range of arthropod characteristics to be assessed (such as species number, life stages, exposure to abiotic and biotic stress, complexity of trophic interactions) that cannot easily be reproduced in laboratory settings. Conversely, laboratory studies may incorporate controlled conditions that are impossible to reproduce in the field, which may prevent the identification of causal relationships. Attention shall therefore be paid to the differences in inferences that may be drawn between standardised tests and field testing.

Due to the lack of well-defined standards, the number of tests on necrotrophic decomposers is very limited and, in particular, some biogeochemical processes cannot be investigated in artificial environments, such as pot experiments. Therefore, field trials may be essential to produce results for such functional groups.

The importance of high quality field trials in the ERA of GM plants is widely accepted (EFSA, 2008). One crucial aspect is the increase in ecological realism that can be achieved as the scale of tests move up from laboratory through mesocosm to semi-field, field and region. For example, when any organism is in contact with a GM plant within a multi-trophic context, identification of the impacts on ecological functioning is facilitated by an increase of scale of the experimental arena.

1.8.5. Experimental design

Experimental designs for laboratory experiments shall conform to accepted international standards and protocols such as those published, for example, by OECD or similar organisations specialising in ecotoxicology.

For field trials, the principle shall be followed that each field trial at a site on a particular occasion shall have sufficient replication to be able to yield a stand-alone analysis if required, although the main analysis shall derive inferences from averages over the complete set of field trials at all sites and years. The level of within-site replication shall be informed by the power analysis referred to above. Notwithstanding this, it is most unlikely that less than three replicates per site would provide an adequate design. A completely randomized or randomized block experimental design is usually appropriate; appropriate extensions to these designs are discussed by (Perry et al., 2009). Applicants shall justify explicitly why the different sites selected for the trials are considered to be representative of the range of receiving environments where the crop will be grown, reflecting relevant meteorological, ecological, soil and agronomic conditions. The choice of plant varieties shall be appropriate for the chosen sites and shall also be justified explicitly. Within each site the GM plant and its conventional counterpart(s) and any additional test material, where appropriate, shall be identical for all replicates. Environmental variation is manifest at two scales: site-to-site and year-to-year. The primary concern is not environmental variation *per se*, but whether potential differences

between the test materials vary across environmental conditions (*i.e.* statistical interactions between test material and environmental factors, often referred to as genotype by environment (GxE) interactions). Hence, in addition to within-field replication there is a need to replicate over sites and years to achieve representativeness across geography and climate. Unless explicit appropriate justification is given by the applicants, each field trial shall be replicated over at least two years, within each of which there shall be replication over at least three sites. In the case that sites cover a very restricted geographic range, further replication of trials, over more than two years, may be required. The use of data from different continents may be informative, but applicants must justify explicitly why the sites within these continents are representative of the range of receiving environments where the GM plant will be grown, reflecting relevant meteorological, ecological, soil and agronomic conditions. In particular, applicants must provide explicit reasons when data from field trials in EU Member States are not available.

However, these explicit requirements above for replication to achieve representativeness do not apply to confirmatory field data for the assessment of unintended effects on non-target organisms when some evidence already exists (see below and Section 1.7), or to the great variety of field trials designed to provide data for a wide range of purposes, to assess aspects of potential persistence and invasiveness (see Section 3.1 of EFSA, 2010b). Many experimental designs used for research purposes are available in the literature as a guide for the very specific requirements for such trials. Data concerning phenotypic and agronomic characteristics of plants is often derived from the same trials designed to supply data for compositional analysis; statistical guidance (EFSA, 2009c,d) has already been prepared for compositional trials and the requirements above do not apply to them. However, for some non-food, non-feed applications for cultivation, such as potatoes modified to enhance the content of the amylopectin component of starch, compositional trials may not be conducted. Then, the experimental design of phenotypic and agronomic trials shall follow the guidance in this section.

For non-target organisms, plant performance and data on environmental measurement endpoints (e.g. agronomic characteristics, including herbivore interactions with the plant, responses to specific environmental exposure) may provide indications concerning the likelihood or otherwise of unintended effects (see Section 1.7.3). This may, for example, include evidence for unchanged ecosystem functions. Under the weight-of-evidence approach (see Section 1.7.3.2), data from field trials may be used to provide such confirmatory data to underpin conclusions that unintended effects are unlikely. While the requirement for statistical power for these field trials shall be carried out as outlined in Section 1.8.3, the requirements for representativeness may be relaxed. Hence, as long as there is explicit justification, under these circumstances, there is no requirement for a minimum number of sites and/or years.

Experimental units (field plots) that are of the spatial scale of a whole or half-field are probably of most use for post-commercialisation studies, for monitoring or mitigation. For pre-commercialisation experimentation, smaller plots, where variation may be controlled and defined treatments imposed more easily, are more appropriate for experimental units (Perry et al., 2009). It is recommended to separate plots within sites, often by strips of bare soil of specified width, and to sample towards the centre of plots to avoid edge-effects. Unless the experiment is set up specifically to study residual effects from one season to the next or to study long-term effects, it is recommended not to utilise exactly the same plots over more than one year at a particular site (Perry et al., 2009).

When it is desirable to assess several different GM plants for one crop species (e.g. *Zea mays*) the generation of data for the comparative assessment of these different GM varieties may be produced simultaneously, at the same site and within the same field trial, by the placing of the different GM plants and their appropriate conventional counterparts in the same randomized block. This is subject to two conditions which shall be strictly met: (i) each of the appropriate counterpart(s) shall always occur together with its particular GM plant in the same block; (ii) all the different GM plants and their

counterpart(s) shall be fully randomized within each block. For further details, and for the use of partially balanced incomplete block designs see EFSA (2009c).

In general, it is easier to impose controlled conditions in semi-field trials, and these are not subject to environmental variability to the same extent as are field trials. However, if semi-field trials do not control conditions then the need to test in different environments (at different sites and/or in different years) shall be considered.

For some GM perennial plants (e.g. trees), the plants themselves may be more appropriate experimental units than are field plots (Petersen, 1994). Care should be taken to choose an experimental design that does not suffer unduly from loss of plants during the trial. Whilst it is largely unnecessary to control for positional variation, plant-to-plant variability should be minimised when selecting experimental material.

1.8.6. Analysis and reporting

It is recommended that applicants prepare an experimental design protocol and a statistical analysis protocol for each study (Perry et al., 2009 for a suggested checklist). It is recommended that the experimental design protocol comprises full information on the study, and includes but is not restricted to: (i) a list of the measurement endpoints, and why they were included; (ii) a description of and justification for of the experimental design; (iii) a description of the experimental units including dimensions; (iv) the blocking structure of the experimental units, in terms of the factors that represent it, their levels and whether the factors are nested or crossed; (v) the sampling regime, within and between experimental units, and through time; (vi) any repeated measurements made in the study; (vii) the test materials and the justification for their inclusion; (viii) the treatment structure of the study, in terms of the factors that represent it and their levels; (ix) a list of the interactions, if any, that are of interest, and why they are; and (x) a description of how the treatment factors listed will be randomized to the experimental units specified in the blocking structure above.

It is recommended that the statistical analysis protocol comprises full information on the analysis, and includes but is not restricted to: (i) a description of the generic form of the analysis and why it was chosen; (ii) the criteria for identifying outliers; (iii) a description of the likely transformations planned, with reasons; (iv) justification for any distributional assumptions; (v) the scale on which the effects in the experiment are assumed to be additive; and (vi) justification for any other assumptions made in the analysis.

For field trials, the protocols shall also include: (i) details of the management of the fields before sowing including the cropping system and rotation; (ii) the dates of sowing; (iii) the soil types; (iv) insecticide and herbicide use and use of any other plant protection products or techniques; (v) climatic and other cultivation/environmental conditions during growth, and where appropriate during harvest; (vi) relevant details of the field margins and neighbouring fields; (vii) brief descriptions of pest and disease infestations.

When many measurement endpoints have been included in a study (e.g. where the endpoints represent several NTO species), the results of all endpoints for which sufficient records have been obtained shall be reported, not just those deemed to be of particular biological or statistical interest. Data transformation may be necessary to ensure normality and to provide an appropriate scale on which statistical effects are additive. As is routine in ecological applications, for many measurement endpoint response variables, a logarithmic transformation (or a generalized linear model with a logarithmic link function) may be appropriate. In such cases, any difference between the GM plant and any other test material is interpreted as a ratio on the natural scale. However, for other measurement endpoints the logarithmic transformation may not be optimal and the natural scale or another scale may be more suitable.

Allowance must be made for possible correlations between repeated measurements from the same experimental units. This is especially important (i) where sampling is repeated over several occasions during a season; and (ii) where the GM plant is a perennial.

Analyses will involve a test for difference and a test for equivalence. Specifically, for a particular measurement endpoint, the mean difference(s) between the GM plant and its conventional counterpart(s) is computed and a 90% confidence interval constructed around it, as in (Perry et al., 2009). This mean(s), these confidence limits and all equivalence limits shall be displayed on a graph(s) similar to Figure 1 of (EFSA, 2009c), but where values are plotted relative to a zero baseline defined by the mean of the GM plant test materials (see Figure 3 of Perry et al., 2009) and example therein). The line of zero difference on the logarithmic scale corresponds to a multiplicative factor of unity on the natural scale. The horizontal axis shall be labelled with values that specify the change on the natural scale. In the case of logarithmic transformation, changes of 2x and ½x will appear equally spaced on either side of the line of zero difference.

Both the difference test and the equivalence test may be implemented using the well-known correspondence between hypothesis testing and the construction of confidence intervals. In the case of equivalence testing the approach used shall follow the two one-sided tests (TOST) methodology (e.g. Schuirmann, 1987) by rejecting the null hypothesis when the entire confidence interval falls between the equivalence limits. The choice of the 90% confidence interval corresponds to the customary 95% level for statistical testing of equivalence. Since the confidence interval graph is used also for the test of difference, each difference test will have a 90% confidence level. Although 1 in 10 of these tests is expected to yield a significant result by chance alone, applicants shall report and discuss all significant differences observed between the GM plant, its conventional counterpart and, where applicable, any other test material, focussing on their biological relevance within the context of risk characterisation. Regarding the simultaneous tests of difference and equivalence, each outcome from the graph shall be categorised and the respective appropriate conclusion shall be drawn, exactly as described in EFSA (2009c).

1.8.7. Statistical analysis of field trials

The main analysis shall address all field trials simultaneously and shall be based on the full dataset from all sites. Accordingly, the form of the equivalence test shall be that termed ‘average equivalence’ in the drug testing literature (Wellek, 2002). The use of a statistical mixed model is an important feature of analysis for food-feed assessments because of the need to estimate the natural variation of the commercial varieties. However, as stated in section 1.8.1 above, for ERA it is recommended that equivalence limits are set explicitly. Therefore the use of commercial varieties for this purpose is not necessary, although it might be appropriate for other biological reasons. Hence it is not recommended that statistical mixed models be required forms of analysis, as they are for food-feed assessments (Perry et al., 2009). Indeed, it is recommended to use simple statistical models; effects due to environmental factors such as seasons and sites may be represented by fixed factors if desired. Applicants shall ensure that each analysis has the potential to identify any interactions between sites and years and the test materials. For each measurement endpoint studied, applicants shall make an explicit statement concerning the presence or absence of any such interactions. If interactions are found, the possible reasons for their existence and the implications for the inferences drawn from the trials shall be discussed. Applicants shall also provide a table or graph giving, for each site and year and for each (transformed) measurement endpoint, the means and standard errors of means of the GM plant and its conventional counterpart(s), and any other test material, where applicable.

Diversity indices are not recommended for general risk assessment in pre-commercialisation studies, because it is most unlikely that studies will yield sufficient samples of individuals to characterise indices adequately or that a sufficient degree of ecological background information will exist to give confidence that biodiversity can be represented adequately as a single number. By contrast,

multivariate approaches may be useful, especially for summarising data and for analysing principal response curves (Perry et al., 2009).

Further discussions and motivations underpinning the above statistical guidance may be found in Perry et al. (2009).

1.9. Uncertainties

Directive 2001/18/EC and the Guidance Notes supplementing Annex II to Directive 2001/18/EC define risk as the product of the magnitude of the consequences of the hazard and the likelihood of the adverse effect. Both the effect and the likelihood are measured with uncertainty.

ERA has to take into account uncertainty at various levels. Uncertainties may arise from problem formulation: limitations in the data (e.g. limited exposure data), gaps in the effect database, model choice, the limitation of the test systems and measurement endpoints selected, inadequacy of study designs and the uncertainties in extrapolating between species (EFSA, 2009a). Scientific uncertainty may also arise from differing interpretations of existing data, publication bias or lack of some relevant data. Uncertainty may relate to qualitative or quantitative elements of the analysis. The level of knowledge or data for a baseline is reflected by the level of uncertainty, which shall be discussed by the applicants. Applicants shall in addition assess the degree of uncertainty within the ERA in comparison with the current uncertainties displayed in the scientific literature.

Although it may be impossible to identify all the uncertainties, the assessment shall include a description of the types of uncertainties encountered and considered during the different risk assessment steps. Their relative importance and their influence on the assessment outcome shall be described (EFSA, 2009a). Any uncertainties inherent in the different steps of the ERA (steps 1 to 5) shall be highlighted and quantified as far as possible; this might be done by adapting the methodology outlined by (Risbey and Kandlikar, 2007). Distinction shall be made between uncertainties that reflect natural variations in ecological and biological parameters (including variations in sensitivity in populations or varieties), and possible differences in responses between species. Estimation of uncertainties in experimental data shall be handled by proper statistical analysis, while quantification of uncertainties in assumptions (e.g. extrapolation from environmental laboratory studies to complex ecosystems) may be more difficult, but shall be discussed fully (Morgan and Henrion, 1990; Finley, 1994) . The absence of data essential for the environmental risk assessment shall be indicated and the quality of existing data shall be discussed. It should be clear from the discussion how this body of information has been taken into account when the final risk characterisation is determined. Risk characterisation may be qualitative and, if possible, quantitative depending on the issue to be addressed and the available data. The terms for the expression of risks and associated uncertainties shall be as precise as possible. For instance, expressions like '*no/negligible/acceptable/significant risk*' need, where possible, further numerical quantification in terms of probability of exposure and/or occurrence of adverse effects (see also Section 2.2.1 of EFSA, 2010b).

It is recognised that an ERA is only as good as our state of scientific knowledge at the time it was conducted. Thus, under current EU legislation, ERAs are required to identify areas of uncertainty or risk which relate to areas outside current knowledge and the limited scope of the ERA. These include such factors as the impact of the large-scale exposure of different environments when GM plants are commercialised, the impact of exposure over long periods of time and cumulative long-term effects. When uncertainty factors (EFSA, 2009a) are used, an explanation of their basis and a justification of their appropriateness need to be provided, or a reference to documents where that information may be found shall be included. When point estimates are used for uncertain quantities, justification for the values chosen and assessment of their influence on the assessment shall be included (EFSA, 2009a).

Predicting impacts of GM plants on complex ecosystems which are continually in flux is difficult and largely based on experiences with other introductions and an understanding of the robustness of ecosystems. It is recognised that an environmental risk assessment is limited by the nature, scale and location of experimental releases, which biospheres have been studied and the length of time the studies were conducted. Probabilistic methods could be used to determine ranges of plausible values rather than single values or point estimates, which are subsequently combined in order to quantify the uncertainty in the end result. These methods could provide a powerful tool to quantify uncertainties associated with any steps in the environmental risk assessment. When such probabilistic approaches are used, the outcome of the environmental risk assessment should be characterised by reporting a distribution of the risk estimates[15]. However, the use of quantitative methods does not remove the need for a qualitative evaluation of the remaining uncertainties (EFSA, 2009a).

Scientific knowledge from the literature and experience gained from growing GM plants encompassed in PMEM following past applications and approvals may also inform the risk assessment process. Notwithstanding the requirement to fully assess all possible risks based on reliable data, this is but one example of the responsibility on applicants continually to update ERA in the light of new knowledge.

2. Risk characterisation

2.1. Introduction

Risk characterisation is defined as: 'The quantitative or semi-quantitative estimate including attendant uncertainties, of the probability of occurrence and severity of adverse effect(s)/event(s) in a given population under defined conditions based on hazard identification, hazard characterisation and exposure assessment' (SSC, 2000). This section describes how the risk characterisation step should be carried out and gives examples of issues to be addressed. Risk characterisation for NTOs involves generating, collecting and assessing information on a GM plant and its products in order to determine its impact on the environment relative to existing baselines, and thus its relative safety. The final risk characterisation should result in informed qualitative, and if possible quantitative, guidance to risk managers. It should explain clearly what assumptions have been made during the risk assessment, and what is the nature and magnitude of uncertainties associated with establishing these risks.

2.2. Hazard characterisation

The objective of this step is to characterize possible adverse effects due to the exposure of NTOs to the GM plant and is addressed in the problem formulation and the body of information informing the problem formulation (see Section 1). The principles of performing the hazard characterisation are described in section 1.7.3 for the envisaged types of plant genome modifications. For example Cry proteins are known to be toxic to some insect species after ingestion, the hypothesis to be tested is therefore if a particular focal species might suffer upon ingestion of plant parts expressing such toxins. The adverse effects will need to be quantified using the relevant lethal and sublethal endpoint measurements, as indicated above (see Section 1.4). Hazard assessment should consider possible effects at different ecological scales (e.g. organismal level, population levels).

2.3. Exposure characterisation

An exposure assessment is conducted to determine whether and to what degree the focal species come into contact with the transgene product. This assessment requires information on the phenotypic pattern of transgene expression in the various parts of the plant over the growing season. This

[15] Examples of probabilistic approaches applied for ERA of pesticides may be found at http://www.eufram.com.

exposure can be bitrophic via exposure to the GM plant (or plant parts) or can occur through higher trophic level exposure (see Figure 5). Moreover, exposure may happen after the transgene moved via gene flow (pollen, seed, horizontal) to other plants that may then cause exposure (e.g. pollen deposited on leaves of wild host plants for non-target *Lepidoptera*). The diet regime for each NTO (in the most relevant biological instar) is paramount in considering potential exposure. The overlap of the life cycle and developmental stages of the focal species and the phenology of the GM plants needs to be evaluated. Exposure may also happen after the transgene has moved via dispersal of pollen and grain/seed in and away from the cultivation site of the GM plant (e.g. pollen deposited on leaves of host plants for non-target *Lepidoptera*, *Coleoptera*). In addition gene flow via outcrossing may result in gene expression in related species and result in different levels of exposure to other NTO species. The exposure of NTOs and their life stages requires a knowledge of their temporal and spatial distribution in relation to the distribution of the GM plant. Based on the specific biological characteristics, the likelihood of exposure needs to be estimated.

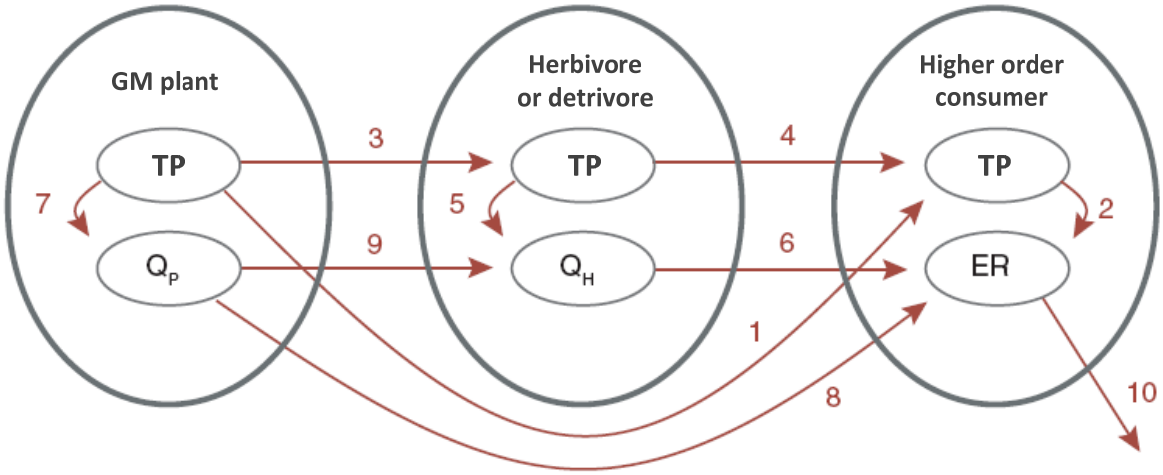

Figure 5: Tritrophic interactions showing (1) GM plant, (2) herbivore or detrivore, and (3) higher order consumer (such as natural enemy), illustrating the first part of hypotheses useful for ERA. TP is transgenic product, Q_P is plant quality, Q_H is herbivore quality and ER is the ecological response (e.g. death, delayed development) by the natural enemies. The remaining parts of the hypotheses that are not depicted here would show how the response of the higher order consumer (such as natural enemy) might lead to an adverse environmental effect (link 10). There are five hypotheses or pathways by which an ER could occur: (Direct 1) 1→2; (Direct 2) 3→4→2; (Indirect 1) 3→5→6; (Indirect 2) 7→9→6; (Indirect 3) 7→8. Figure adapted from Andow et al. (2006a).

In assessing the possible direct or indirect exposure of NTOs to GM plants and its products, worst-case scenario should be used considering expected concentration of the product along the food web.

2.4. The result of risk characterisation

Based on the conclusions of the hazard and exposure characterisation, applicants shall estimate each identified risk that a GM plant will cause to NTOs considering the magnitude of the effects detected and the likelihood of their occurrence. Applicants shall summarize the outcomes of the ERA considering intended and unintended effects as outlined in Section 1.7.3. Unintended effects shall be excluded by applicants by resorting to the weight-of-evidence approach. Hence applicants shall conclude on risk for NTOs taking into account focal species as well as the overall functionality of the

agro-ecosystem. Applicants shall provide an assessment of the range of effects likely to occur in different receiving environments based on the collected data and other relevant information.

Considering receiving environment-plant-trait combinations, applicants are also required to characterize the risk (a) in the production site of the GM plant and (b) outside the production site in different habitats (e.g. adjacent crops and other non-crop habitats) where relevant exposure of sensitive NTOs may occur. Quantification of risks and its relative uncertainties shall be provided in relation to each selected assessment endpoint and upscaling of data from lab, semi-field and field trials to landscapes considering the expected adoption rate of GM plants. The conclusions of risk characterisation shall assess the consequences of each identified risk to NTOs and applicants shall propose appropriate risk management measures where levels of risk exceed threshold levels.

3. Risk management strategies

In situations where risk due to the GM plant and/or its product(s) on NTOs and related ecosystem services has been identified and characterized, applicants should propose appropriate risk management strategies. These strategies should be designed, under assumptions of high exposure scenarios, to reduce the risk to a level considered acceptable (criteria defining this acceptability should be explicitly discussed) and their ability to reduce risks should be documented. The implementation of measures should fit to common principles e.g. the principles of good agricultural practice and Integrated Pest Management that are being introduced by Member States under the Framework Directive on the sustainable use of pesticides in the EU (EC, 2009b).

These mitigation measures may include measures to reduce exposure in order to reduce risk to NTOs and ecosystem services. Examples might be the planting of non-Bt plants as border rows (EFSA, 2009b) or, where feasible, detasseling of GM maize plants in border rows in order to limit Bt maize pollen dispersal outside of the maize field. Also, the establishment and maintenance of habitats (ecological compensation areas) that provide refugia, feeding source, etc. for NTO populations over larger area and time might also be considered (Boller et al., 2004).

Applicants should also consider the implications of the introduction of the GM plant on present cultivation and farming practices. Applicants should describe how the GM plant will be introduced into Integrated Pest Management and farming systems so that present pest management strategies and practices contribute to sustainability of pest management. These practices that should be in line with general IPM principles (EC, 2009b) may cover rotation of crops and crop varieties, use of pesticides with different modes of action in order to maintain and support natural regulating mechanisms, including beneficial NTOs.

These mitigation measures and strategies should be devised in the light of a long-term management and maintenance of NTOs and ecosystem services in rural landscapes.

4. Post-market environmental monitoring

The EFSA GMO Panel refers to its opinion on PMEM for further advice to applicants (EFSA, 2006b) as well as to Section 4 of the updated ERA guidance document (EFSA, 2010b).

CONCLUSIONS AND GUIDANCE

Applicants shall conclude on the risk of intended and unintended effects on NTOs taking into account focal species considering all relevant ecosystem services. Applicants shall provide an assessment of the range of effects likely to occur in relevant EU receiving environments based on the collected data and other relevant information. Applicants are also required to characterize the risk (a) in the

production site of the GM plant and (b) outside the production site in different habitats considering relevant exposure routes. Quantification of risks and its relative uncertainties shall be provided in relation to each selected assessment endpoint in comparison to relevant baselines. The consequences of these risks for all relevant protection goals, including the overall functionality of the ecosystems, integrated pest management and the sustainability of production systems, shall be considered.

The conclusions of risk characterisation shall assess the consequences of each identified risk to NTOs and applicants shall propose appropriate risk management measures where levels of risk exceed acceptable threshold levels.

REFERENCES

ACRE, 2001. Guidance on the assessment of the impact on wider biodiversity from proposed cultivation of GM crops. http://webarchive.nationalarchives.gov.uk/20081107165902/http://www.defra.gov.uk/environment/acre/biodiversity/guidance/index.htm

Andow DA, Zwahlen C, 2006. Assessing environmental risks of transgenic plants. Ecology Letters 9(2), 196-214.

Andow DA, Lövei GL, Arpaia S, 2006a. Ecological risk assessment for Bt crops. Nature Biotechnology 24, 749-751.

Andow DA, Birch NAE, Dusi AN, Fontes EMG, Hilbeck A, Lang A, Lövei GL, Pires CSS, Sujii ER, Underwood E, Wheatley RE, 2006b. Non-target and biodiversity ecological risk assessment for GM crops. Proceedings of 9th International Symposium on the Biosafety of Genetically Modified Organisms (ISBR), September 24-29 2006, Jeju Island, Korea, pp 68-73.

Andow DA, Hilbeck A & Nguyen Van Tuat, 2008. Environmental Risk Assessment of Genetically Modified Organisms, Volume 4, Challenges and Opportunities with Bt Cotton in Vietnam. CABI Publishing, Wallingford, UK. http://www.gmoera.umn.edu/public/publications/books.html

Arpaia S, 2004. Workshop report – Monitoring/bioindicators. IOBC/Bulletin OILB/SROP 27(3), 205-208.

Arpaia S, Di Leo GM, Fiore MC, Schmidt JEU, Scardi M, 2007. Composition of Arthropod Species Assemblages in Bt-expressing and Near Isogenic Eggplants in Experimental Fields. Environmental Entomology 36(1), 213-227.

Arpaia S, Schmidt JEU, Di Leo GM, Fiore MC, 2009. Oviposition of the Colorado potato beetle (*Leptinotarsa decemlineata* - Say) and natural predation on its egg masses in Bt-expressing fields. Biocontrol Science and Technology 19(9), 971-984.

ASTM, 2008. Standard Guide for Conducting Laboratory Soil Toxicity Tests with the Nematode Caenorhabditis elegans. ASTM E2172 – 01.

Babendreier D, Reichhart B, Romeis J, and Bigler F, 2008. Impact of insecticidal proteins expressed in transgenic plants on bumblebee microcolonies. Entomologia Experimentalis et Applicata 126, 148–157.

BEETLE_report, 2009. Long-term effects of genetically modified (GM) crops on health and the environment (including biodiversity): prioritization of potential risks and delimitation of uncertainties. German Federal Office of Consumer Protection and Food Safety, BLaU-Umweltstudien and Genius GmbH.

Birch NAE, Wheatley RE, Hilbeck A, Arpaia S, Lovei GL, Songa J, Sétamou M, Sithanantham S, Ngi-Song A, Anyango B, Muyekho F, Ochieno D, Lelmen E, Kalama P, Degaga E, Soares Melo I,

Pitelli R, Sujii ER, Capalbo D, Fontes EMG, Wan F, Ogwang J, Smith J, Van Son N, Quang Tan T, 2004. Biodiversity and Non-Target Impacts: a Case Study of Bt Maize in Kenya. In: Hilbeck A & Andow D [eds.]. Environmental Risk Assessment of Genetically Modified Organisms, Volume 1: A Case Study of Bt Maize in Kenya, 117-185. CABI Publishing, Wallingford, UK.

Blondel J, 2003. Guilds or functional groups: does it matter? Oikos, 100, 223-231.

Boller EF, Häni F, Poehling HM, 2004. Ecological infrastructures: ideabook on functional biodiversity at the farm level, Landwirtschaftliche Beratungszentrale Lindau, Lindau, Suisse.

Brodsgaard CJ, Ritter W, Hansen H, 1998. Response of in vitro reared honey bee larvae to various doses of *Paenibacillus larvae larvae* spores. Apidologie 29, 569–578.

Candolfi MP, Blümel S, Forster R, Bakker FM, Grimm C, Hassan SA, Heimbach U, Mead-Briggs MA, Reber B, Schuck R, Vogt H, 2000. Guidelines to evaluate side-effects of plant protection products to non-target arthropods – IOBC, BART and EPPO Joint Initiative - IOBC/WPRS, Gent Belgium.

Charleston DS, Dicke M, 2008. Designing experimental protocols to investigate the impact of GM crops on non-target arthropods. COGEM onderzoeksrapport CGM 2008-01. 61 pp.

Crawley MJ, Hails RS, Rees M, Kohn D and Buxton J, 1993. Ecology of transgenic oilseed rape in natural habitats. Nature 363**,** 620 – 623.

Cummins KW, 1974. Structure and function of stream ecosystems. Bioscience 24, 631-641.

Davidson EW, Morton HL, Moffett JO, Singer S, 1977. Effect of *Bacillus sphaericus* strain SSII-1 on honey bees, *Apis mellifera*. Journal of Invertebrate Pathology 29, 344–346.

Dively GP, 2005. Impact of Transgenic VIP3A * Cry1Ab Lepidopteran-resistant Field Corn on the Nontarget Arthropod Community. Environmental Entomology,34, 1267-1291.

Duan JJ, Paradise MS, Lundgren JG, Bookout JT, Jiang C and Wiedenmann RN, 2006. Assessing nontarget impacts of *Bt* corn resistant to corn rootworms: tier-1 testing with larvae of *Poecilus chalcites* (Coleoptera: Carabidae). Environmental Entomology 35**,** 135–142.

Duan JJ, Marvier M, Huesing J, Dively G, Huang ZY, 2008. A Meta-Analysis of Effects of Bt Crops on Honey Bees (Hymenoptera: Apidae). PLoS ONE 3(1): e1415. doi:10.1371/journal.pone.0001415

Duan JJ, Lundgren JG, Naranjo S, Marvier M, 2010. Extrapolating non-target risk of *Bt* crops from laboratory to field. Biology letters 6(1), 74-77.

EC, 1992. Directive 92/43/EEC Council Directive 92/43/EEC of 21 May 1992 on the conservation of natural habitats and of wild fauna and flora. Official Journal of the European Communities L206, 7-50. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31992L0043:EN:html

EC, 2001. Directive 2001/18/EC of the European Parliament and of the Council of 12 March 2001 on the deliberate release into the environment of genetically modified organisms and repealing Council Directive 90/220/EEC. Official Journal of the European Communities L106, 1-39. http://europa.eu.int/eur-lex/pri/en/oj/dat/2001/l_106/l_10620010417en00010038.pdf

EC, 2002. Council Decision of 3 October 2002 establishing guidance notes supplementing Annex VII to Directive 2001/18/EC of the European Parliament and of the Council on the deliberate release into the environment of genetically modified organisms and repealing Council Directive 90/220/EEC. Official Journal of the European Communities L280**,** 27-36. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2002:280:0027:0036:EN:PDF

EC, 2004. Directive 2004/35/CE of the European Parliament and of the Council of 21 April 2004 on environmental liability with regard to the prevention and remedying of environmental damage.

Official Journal of the European Communities L143, 56-75. http://europa.eu.int/eurlex/pri/en/oj/dat/2004/l_143/l_14320040430en00560075.pdf

EC, 2009a. Regulation (EC) No 1107/2009 of the European Parliament and of Council of 21 October 2009 concerning the placing of plant protection products on the market and repealing Council Directives 79/117/EEC and 91/414/EEC. Official Journal of the European Union L309, 1-50. http://www.furs.si/law/EU/ffs/eng/NovaEUzakonodaja/32009R1107.pdf

EC, 2009b. 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, 71-85. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2009:309:0071:0086:EN:PDF

EFSA, 2006a. Guidance document of the Scientific Panel on Genetically Modified Organisms for the Risk Assessment of Genetically Modified Plants and Derived Food and Feed. The EFSA Journal 374, 1-115.

EFSA, 2006b. Opinion of the Scientific Panel on Genetically Modified Organisms on the Post Market Environmental Monitoring (PMEM) of genetically modified plants, The EFSA Journal 319, 1-27.

EFSA 2007, Guidance Document of the Scientific Panel on Genetically Modified Organisms for the risk assessment of genetically modified plants containing stacked transformation events The EFSA Journal 512, 1-5.

EFSA, 2008. Environmental risk assessment of genetically modified plants - challenges and approaches. IN colloquium Es (Ed.) EFSA scientific colloquium - summary report.

EFSA, 2009a. Guidance of the Scientific Committee on Transparency in the Scientific Aspects of Risk Assessments carried out by EFSA. Part 2: General Principles. The EFSA Journal 1051**,** 1-22.

EFSA, 2009b. Scientific Opinion of the Panel on Genetically Modified Organisms on applications (EFSA-GMO-RX-MON810) for the renewal of authorisation for the continued marketing of (1) existing food and food ingredients produced from genetically modified insect resistant maize MON810; (2) feed consisting of and/or containing maize MON810, and maize MON810 for feed use (including cultivation); and of (3) food additives and feed materials produced from maize MON810, all under Regulation (EC) No 1829/2003 from Monsanto. The EFSA Journal 1149**,** 1-85.

EFSA, 2009c. Scientific opinion on Statistical considerations for the safety evaluation of GMOs, on request of EFSA. The EFSA journal 1250**,** 1-62.

EFSA, 2009d. Updated Guidance Document for the risk assessment of genetically modified plants and derived food and feed. Annex 1 to the Scientific report of EFSA prepared by the GMO Unit on Public Consultation on the Updated Guidance Document of the Scientific Panel on Genetically Modified Organisms (GMO) for the risk assessment of genetically modified plants and derived food and feed. The EFSA Scientific Report, 293, 1-18.

EFSA, 2010a. Scientific opinion on the development of specific protection goal options for environmental risk assessment of pesticides, in particular in relation to the revision of the Guidance Documents on Aquatic and Terrestrial Ecotoxicology. The EFSA Journal 1821**,** 1-55.

EFSA, 2010b. Scientific opinion on guidance for the Environmental Risk Assessment of genetically modified plants. The EFSA Journal 8, 1879.

EPA US, 1998. Guidelines for Ecological Risk Assessment. IN EPA (Ed.). Environmental Protection Agency. http://www.epa.gov/superfund/programs/nrd/era.htm

EPA, 1999: Biopesticide fact sheet: *Bacillus thuringiensis* Cry1A(b) delta-endotoxin and the genetic material necessary for its production in corn. Office of Pesticide Programs, US Environmental Protection Agency
www.epa.gov/pesticides/facsheets/fs006430t.htm

Evans D, 2006. The habitats of the European Union Habitats Directive. The Royal Irish Academy 106(3), 167-173.
Biology & Environment: Proceedings of the Royal Irish Academy

FAO, 2008. Land and Water development division; Soil biodiversity Portal: http://www.fao.org/ag/agl/agll/soilbiod/rootalk.stm

Finke DL, Snyder WE, 2008. Niche partitioning increases resource exploitation by diverse communities. Science 321(5895), 1488 – 1490.

Finley BP, D, 1994. The Benefits of Probabilistic Exposure Assessment: Three Case Studies Involving Contaminated Air, Water, and Soil. Risk Analysis 14, 53-73.

Garcia-Alonso M, Jacobs E, Raybould A, Nickson TE, Sowig P, Willekens H, Van Der Kouwe P, Layton R, Amijee F, Fuentes AM, Tencalla F, 2006. A tiered system for assessing the risk of genetically modified plants to non-target organisms. Environmental Biosafety Research 5, 57–65.

Gathmann A, Wirooks L, Hothhorn L, Bartsch D, Schuphan I, 2006. Impact of Bt-maize pollen (MON810) on lepidopteran larvae living on accompanying weeds. Molecular Ecology 15, 2677–2685.

International Organization for Standardisation (ISO), 1985. Water quality -- Methods of biological sampling -- Guidance on handnet sampling of aquatic benthic macro-invertebrates. ISO 7828.

International Organization for Standardisation (ISO), 2005. Soil quality -- Determination of dehydrogenase activity in soils -- Part 1: Method using triphenyltetrazolium chloride (TTC), ISO 23753-1.

International Organization for Standardisation (ISO), 2006. Soil quality -- Effects of pollutants on juvenile land snails (Helicidae) -- Determination of the effects on growth by soil contamination, ISO 15952.

International Organization for Standardisation (ISO), 2006. Soil quality -- Sampling of soil invertebrates -- Part 1: Hand-sorting and formalin extraction of earthworms. ISO 23611-1.

International Organization for Standardisation (ISO), 2006. Soil quality -- Sampling of soil invertebrates -- Part 2: Sampling and extraction of micro-arthropods (Collembola and Acarina). ISO 23611-2.

International Organization for Standardisation (ISO), 2007. Soil quality -- Sampling of soil invertebrates -- Part 3: Sampling and soil extraction of enchytraeids. ISO 23611-3.

International Organization for Standardisation (ISO), 2007. Soil quality -- Sampling of soil invertebrates -- Part 4: Sampling, extraction and identification of soil-inhabiting nematodes. ISO 23611-4.

International Organization for Standardisation (ISO), 2009. Soil quality -- Sampling -- Part 6: Guidance on the collection, handling and storage of soil under aerobic conditions for the assessment of microbiological processes, biomass and diversity in the laboratory. ISO 10381-6.

International Organization for Standardisation (ISO), 2009. Soil quality -- Effects of pollutants on mycorrhizal fungi -- Spore germination test. ISO/TS 10832.

International Organization for Standardisation (ISO), (under development). Soil quality -- Determination of potential nitrification and inhibition of nitrification -- Rapid test by ammonium oxidation. ISO/CD 15685.

International Organization for Standardisation (ISO), (under development). Soil quality -- Determination of abundance and activity of soil microflora using respiration curves. ISO/CD 17155.

Heard MS, Hawes C, Champion GT, Clark SJ, Firbank LG, Haughton AJ, Parish AM, Perry JN, Rothery P, Scott RJ, Skellern MP, Squire GR & Hill MO, 2003. Weeds in fields with contrasting conventional and genetically modified herbicide-tolerant crops. 1. Effects on abundance and diversity. Philosophical Transactions of the Royal Society of London Biological Sciences 358, 1819-1832.

Hilbeck A, Andow D, Birch ANE, Bong BB, Capalbo DMF, Fitt GP, Fontes EMG, Heong KL, Johnston J, Osir EO, Snow A, Songa J, Underwood E and Wan FH, 2004. The GMO Guidelines Project: Development of International Scientific Environmental Biosafety Testing Guidelines for Transgenic Plants. In: L.E. Ehler, R. Sforza and T. Mateille (eds) Genetics, Evolution and Biological Control. CABI Publishing, Wallingford, UK, pp 207-218.

Hilbeck A, Andow DA, Arpaia S, Birch ANE, Fontes EMG, Lövei GL, Sujii E, Wheatley RE and Underwood E, 2006. Methodology to support non-target and biodiversity risk assessment, pp. 108-132 CABI Publishing, Wallingford, UK.

Hilbeck A, Jänsch S, Meier M, Römbke J, 2008. Analysis and validation of present ecotoxicological test methods and strategies for the risk assessment of genetically modified plants. BfN – Skripten 236, 287 pp.

Hill RA, 2005. Conceptualizing risk assessment methodology for genetically modified organisms. Environmental Biosafety Research 4, 67–70.

Hill R, Sendashonga C, 2003. Conservation Biology, Genetically Modified Organisms, and the Biosafety Protocol. Conservation Biology 20(6), 1620-1625.

Kramarz PE, De Vaufleury A, Carey M, 2007. Studying the effect of exposure of the snail *Helix aspera* to the purified Bt toxin, Cry1Ab. Applied soil ecology 37, 169-172.

Kramarz P, Stark JD, 2003. Population level effects of cadmium and the insecticide imidacloprid to the parasitoid, *Aphidius ervi* after exposure through its host, the pea aphid, *Acyrthosiphon pisum* (Harris). Biological Control 27, 310–314.

Kromp B, 1999. Carabid beetles in sustainable agriculture: a review on pest control efficacy, cultivation impacts and enhancement. Agriculture, Ecosystems and Environment 74, 187-228.

Laster LL, Johnson MF, 2003. Non-inferiority trials: the 'at least as good as' criterion. Statistics in Medicine 22**,** 187-200.

Lawo NC, Romeis J, 2008. Assessing the utilization of a carbohydrate food source and the impact of insecticidal proteins on larvae of the green lacewing, *Chrysoperla carnea.* Biological Control 44, 389–398.

Li Y, Meissle M, Romeis J, 2010. Use of maize pollen by adult *Chrysoperla carnea* (Neuroptera: Chrysopidae) and fate of Cry proteins in *Bt*-transgenic varieties. Journal of Insect Physiology 56, 157-164.

Lovei GL, Arpaia S, 2005. The impact of transgenic plants on natural enemies a critical review of laboratory studies. Entomologia Experimentalis et Applicata 114, 1–14.

Marvier M, 2002. Improving risk assessment for non-target safety of transgenic crops. Ecological Applications 12, 1119-1124.

Marvier M, McCreedy C, Regetz J and Kareiva P, 2007. A meta-analysis of effects of Bt cotton and maize on nontarget invertebrates. Science 316**,** 1475-1477.

Millennium Ecosystem Assessment (MEA), 2005. Ecosystems and Human Well-being: Synthesis. Island Press, Washington, DC, 160 pp.

Morandin LA, Winston ML, 2003. Effects of novel pesticides on bumble bee (*Hymenoptera*: *Apidae*) colony health and foraging ability. Community and Ecosystem Ecology 32, 555–563.

Morgan MG and Henrion M, 1990. Uncertainty: a guide to dealing with uncertainty in quantitative risk and policy analysis, Cambridge University Press.

Mücher CA, Klijn JA, Wascher DM, Schaminée JHJ, 2010. A new European Landscape Classification (LANMAP): A transparent, flexible and user-oriented methodology to distinguish landscapes. Ecological Indicators 10, 87–103.

Naranjo SE, 2005a. Long-term assessment of the effects of transgenic *Bt* cotton on the function of the natural enemy community. Environmental Entomology 34, 1211-1223.

Naranjo SE, 2005b. Long-Term Assessment of the Effects of Transgenic Bt Cotton on the Abundance of Non-target Arthropod Natural Enemies. Environmental Entomology 34, 1193-1210.

Nelson KC, Banker MJ, 2007. Problem formulation and options assessment handbook: A guide to the PFOA process and how to integrate it into environmental risk assessment (ERA) of genetically modified organisms (GMOs). GMO-ERA Project.
Available at: www.gmoera.umn.edu.

Nguyen HT, Jehle JA, 2007. Quantitative analysis of the seasonal and tissue-specific expression of Cry1Ab in transgenic maize MON810. Journal of Plant Diseases and Protection 114(2), 82-87.

Nickson TE, 2008. Planning Environmental Risk Assessment for Genetically Modified Crops: Problem Formulation for Stress-Tolerant Crops. Plant Physiology 147, 494–502.

OECD (Organisation for Economic Co-Operation and Development) (1998a): Guideline for Testing of Chemicals No. 213. Honeybees, Acute Oral Toxicity Test. Paris, France.

OECD (Organisation for Economic Co-Operation and Development) (1998b): Guideline for Testing of Chemicals No. 214. Honeybees, Acute Contact Toxicity Test. Paris, France.

OECD (Organisation for Economic Co-Operation and Development) (2000a): Guideline for Testing of Chemicals No. 216. Soil microorganisms, Nitrogen Transformation Test. Paris, France.

OECD (Organisation for Economic Co-Operation and Development) (2000b): Guideline for Testing of Chemicals No. 217. Soil microorganisms, Carbon Transformation Test. Paris, France.

OECD (Organisation for Economic Development) (2003a): Guideline for testing of chemicals No. 220: Enchytraeidae reproduction test. Paris, France.

OECD (Organisation for Economic Development) (2004): Guideline for Testing Chemicals No. 222: Earthworm reproduction test. Paris, France.

OECD (Organisation for Economic Development) (2006): Guidance Document on the breakdown of organic matter in litter bags. OECD series on testing and assessment No. 56. Paris, France.

OECD (Organisation for Economic Development) (2008): Guideline for Testing Chemicals No. 226: Predatory mite (*Hypoaspis* (*Geolaelaps*) *aculeifer*) reproduction test in soil. Paris, France.

OECD (Organisation for Economic Development) (2009): Guideline for Testing Chemicals No. 232: Collembolan reproduction test in soil. Paris, France.

Pemsl D, Waibel H, Gutierrez AP, 2005. Why do some Bt-cotton farmers in China continue to use high levels of pesticides? International Journal of Agricultural Sustainability 3, 44–56.

Perry JN, 1986. Multiple comparison procedures: a dissenting view. Journal of Economic Entomology 79, 1149-1155.

Perry JN, Rothery P, Clark SJ, Heard MS, Hawes C, 2003. Design, analysis and statistical power of the Farm-Scale Evaluations of genetically modified herbicide-tolerant crops. Journal of Applied Ecology 40**,** 17-31.

Perry JN, Ter Braak CJF, Dixon PM, Duan JJ, Hails RS, Huesken A, Lavielle A, Marvier M, Scardi M, Schmidt K, Tothmeresz B, Schaarschmidt F, Van der Voet H, 2009. Statistical aspects of environmental risk assessment of GM plants for effects on non-target organisms. Environmental Biosafety Research 8**,** 65-78.

Perry JN, Devos Y, Arpaia S, Bartsch D, Gathmann A, Hails RS, Kiss J, Lheureux K, Manachini B, Mestdagh S, Neemann G, Ortego F, Schiemann J, and Sweet JB, 2010. A mathematical model of exposure of non-target Lepidoptera to *Bt*-maize pollen expressing Cry1Ab within Europe. Proceedings of the Royal Society B: Biological Sciences, 2010-01-10, 1-9.

Petersen RG, 1994. Agricultural field experiments: design and analysis. Mercel Dekker Inc. New York. 409 pp.

Qi A, Perry JN, Pidgeon JD, Haylock LA, and Brooks DR, 2008. Cost-efficacy in measuring farmland biodiversity—lessons from the Farm Scale Evaluations of genetically modified herbicide-tolerant crops. Annual Applied Biology 152, 93–101.

Ramirez-Romero R, Bernalb JS, Chaufaux J, Kaiser L, 2007. Impact assessment of Bt-maize on a moth parasitoid, *Cotesia marginiventris* (*Hymenoptera*: Braconidae), via host exposure to purified Cry1Ab protein or Bt-plants. Crop Protection 26, 953–962.

Raybould A, 2006. Problem formulation and hypothesis testing for environmental risk assessments of genetically modified crops. Environmental Biosafety Research 5 (3), 119-125.

Raybould A, 2007a. Environmental Risk Assessment of Genetically Modified Crops: General Principles and Risks to Non-target Organisms. BioAssay 2, 8.

Raybould A, 2007b. Ecological versus ecotoxicological methods for assessing the environmental risks of transgenic crops. Plant Science 173, 589-602.

Risbey JS and Kandlikar M, 2007. Expressions of likelihood and confidence in the IPCC uncertainty assessment process. Climatic Change 85, 19-31.

Romanow LR, De Ponti OMB, Mollema C, 1991. Resistance in tomato to the greenhouse whitefly: analysis of population dynamics. Entomology Experimentalis et Applicata 60, 247-259.

Romeis J, Meissle M, Bigler F, 2006. Transgenic crops expressing *Bacillus thuringiensis* toxins and biological control. Nature Biotechnology 24, 63-71.

Romeis J, Bartsch D, Bigler F, Candolfi MP, Gielkens MMC, Hartley SE, Hellmich RL, Huesing JE, Jepson PC, Layton R, Quemada H, Raybould A, Rose RI, Schiemann J, Sears MK, Shelton AM, Sweet J, Vaituzis Z, Wolt JD, 2008. Assessment of risk of insect-resistant transgenic crops to non-target arthropods. Nature Biotechnology 26, 203-208.

Rose R, Dively GP, Pettis J, 2007. Effects of Bt corn pollen on honey bees: Emphasis on protocol development. Apidologie 38, 368–377.

Sanvido O, Romeis J, Bigler F, 2009. An approach for post-market monitoring of potential environmental effects of Bt-maize expressing Cry1Ab on natural enemies. Journal of Applied Entomology 133, 236-248.

Schuirmann DJ, 1987. A Comparison of the 2 one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability. Journal of Pharmacokinetics and Biopharmaceutics 15**,** 657-680.

Schuler TH, Denholm I, Clark SJ, Stewart CN, Poppy GM, 2003. Effects of Bt plants on the development and survival of the parasitoid *Cotesia plutellae* (*Hymenoptera*: *Braconidae*) in

susceptible and Bt-resistant larvae of the diamondback moth, *Plutella xylostella* (*Lepidoptera*: *Plutellidae*). Journal of Insect Physiology 50, 435–443.

Smith TM, Shugart HH, Woodward FI (eds.),. 1997. Plant Functional Types: Their Relevance to Ecosystem Properties and Global Change. Cambridge: Cambridge University Press.

SSC, 2000. First Report of the Scientific Steering Committee on the Harmonisation of Risk Assessment Procedures. European Commission, Health & Consumer Protection Directorate-General. http://europa.eu.int/comm/food/fs/sc/ssc/out83_en.pdf
Appendices: http://europa.eu.int/comm/food/fs/sc/ssc/out84_en.pdf

Stacey D, Graser G, Mead-Briggs M, Raybould A, 2006. Testing the impact on non-target organisms of insecticidal proteins expressed in transgenic crops. IOBC/WPRS Bulletin 29, 165–172.

Stark JD, Wennergren U, 1995. Can population effects of pesticides be predicted from demographic toxicological studies? Journal of Economical Entomology 88, 1089-1096.

Steneck RS, 2001. Functional groups. – In: Levin, S. A. (ed.), Encyclopedia of biodiversity. Vol. 3. Academic Press, pp. 121–139.

Storkey J, Bohan DA, Haughton AJ, Champion GT, Perry JN, Poppy GM, Woiwod IP, 2008. Providing the evidence base for environmental risk assessments of novel farm management practices. Environmental Science and Policy 11 (7), 579-587.

Suter GW, 2000. Generic assessment endpoints are needed for ecological risk assessment. Risk analysis 20, 173-178.

Takhtajan, A., 1986. Floristic Regions of the World. (translated by T.J. Crovello & ed. by A. Cronquist). University of California Press, Berkeley. 522 pp.

Thi Thu Cuc N, Sujii ER, Wilson LJ, Underwood E, Andow DA, Van Hao M, Zhai B, Van Chién H, 2008. Potential effect of transgenic cotton on non-target herbivores in Vietnam. In: Environmental Risk Assessment of Genetically Modified Organisms: Vol. 4. Challenges and opportunities for Bt Cotton in Vietnam. Andow DA, Hilbeck A & Van Tuàt N (eds.), CABI Publishing, Wallingford, UK. pp 138-175.

Todd JH, Ramankutty P, Barraclough EI, Malone LA, 2008. A screening method for prioritizing non-target invertebrates for improved biosafety testing of transgenic crops. Environmental Biosafety Research 7, 35–56.

Van Ittersum MK, Ewert F, Heckelei T, Wery J, Alkan Olsson J, Andersen E, Bezlepkina I, Brogaard S, Donatelli M, Flichman G, Olsson L, Rizzoli A, van der Wal T, Wien JE, Wolf J, 2008. Integrated assessment of agricultural systems - A component-based framework for the European Union (SEAMLESS). Agricultural Systems 96, 150-165.

Walthall WK, Stark JD, 1997. Comparison of two population-level ecotoxicological endpoints: The intrinsic (rm) and instantaneous (ri) rates of increase. Environmental Toxicology and Chemistry, 16(5) 1068-1073.

Wilkinson M, Tepfer M, 2009. Fitness and beyond: preparing for the arrival of GM crops with ecologically important novel characters. Environmental Biosafety Research 8, 1-14.

Wellek S, 2002. Testing Statistical Hypotheses of Equivalence. CRC Press, ISBN: 1584881607. 284 pp.

Whitehouse MEA, Wilson LJ, Fitt GP, 2005. A Comparison of Arthropod Communities in Transgenic Bt and Conventional Cotton in Australia. Environmental Entomology 34, 1224-1241.

Wilkinson MJ, Sweet J, Poppy GM, 2003. Risk assessment of GM plants: avoiding gridlock? Trends in Plant Science Volume 8, Issue 5, 208-212. http://www.sciencedirect.com/science/article/B6TD1-48FSWX5-1/2/e8ed52fdfccd08ea9398638a8511c0e6

Wolt JD, Keese P, Raybould A, Fitzpatrick JW, Burachik M, Gray A, Olin SS, Schiemann S, Sears M and Wu F, 2010. Problem formulation in the environmental risk assessment for genetically modified plants. Transgenic Research 19(3), 425-436.

Yang PY, Iles M, Yan S, Jolliffe F, 2005. Farmers' knowledge, perceptions and practices in transgenic Bt cotton in small producer systems in Northern China. Crop Protection 24, 229–239.

1

APPENDIX I:

ROAD MAP ON HOW TO COLLECT SUFFICIENT DATA FOR ENVIRONMENTAL RISK ASSESSMENT ON NTOS (EXAMPLES ARE BASED ON A GENERIC CASE OF CRY PROTEIN-EXPRESSING GM MAIZE)[16]

STEP 1: ANALYZING RELEVANT PROTECTION GOALS

Applicants should consider (among others) the following targets, where appropriate:

1.a. In agro-ecosystems

- Natural regulatory mechanisms controlling pest populations
- Pollination
- Soil biodiversity and ecosystem services
- Healthy plant stands
- Biodiversity in ecological infrastructure (e.g. field margins)
- Sustainability of pest management practices

1.b. In adjacent (non-managed) habitats

- Protected and endangered species in protected areas (possibly exposed via pollen dust on their host plants)
- Water bodies (possibly exposed through accidental presence of plant parts in the water)
- Breeding resource (e.g. for birds)
- Pollination

[16] The examples listed in this Road map are only indicative to illustrate the logical process underlying the proposed Environmental Risk Assessment for NTOs and should not be used as a standard set of experiments for similar cases.

STEP 2: SELECTING RECEIVING ENVIRONMENT(S) FOR TESTING NTOS

Steps for the selection	Criteria to meet	Remarks	Examples
2.1. Consider production area of the plant	Potential receiving environments for field tests should be **representative** for the broad range of climatic (temperature, rainfall, etc.) conditions in EU	Present and likely near future areas should also be considered	Maize can be grown from Mediterranean regions up to Denmark
2.2. Plant x trait	Potential receiving environments should cover **typical production areas** where the trait will be relevant	Certain traits are not likely to be relevant for all EU areas (significance of target pest, of the trait itself, etc.)	Certain North-Western regions could be excluded if e.g. Western corn rootworm is the target pest
3. Plant x trait x NTO (in-field)	Potential receiving environments should **be characterized in relation to NTO species, their guilds** need to be identified in advance	There are differences in NTO species assemblages under different climatic conditions	
3a	**Herbivore** species be **representative** for European receiving environments and be in **suitable density** for sampling and analysis	Some NTOs prefer humid climate while others do dry and/or warm climate.	aphids humid climate, spider mites dry warm climate,
3b	**Predators, parasitoids** should be **representative** for European receiving environments and be in **suitable density** for sampling and analysis	Predator and parasitoid density depends among other on prey/host availability which should be taken into account	populations of acariphagous cocinellids are supported by spider mite availability, while aphidiphagous coccinellids and lacewings are supported by aphids
3c	**Pollinators, pollen feeders** (where relevant) to be considered	Plants maintain/attract different species and at different extent. Some predators are mixed feeders.	where relevant (*Orius* sp., honey bees, Coccinellids)
3d	**Decomposers** to be considered	Many decomposers are present at higher density in humid areas or season. Other in dry and warm	selected *Diptera* and/or springtails Scarabeids
3e	Species of **conservation or cultural** value	Only in-field exposure (i.e. if nettles are present in the field)	*Inachis io and Vanessa atalanta*
3f	Consider general statistical requirement	In general, 2 years and 3 sites	
4. Plant x trait x NTO (adjacent habitats)	Potential receiving environments should allow field tests, observations on **NTO species in adjacent habitats** (e.g. field margins) based on **prior exposure**	Composition and structure of adjacent habitats and therefore, species maintained by these	*Inachis io and Vanessa atalanta*

	analysis	habitats show high variability across regions.	
5. Production system context	Along the selection process, consider sites with **typical current agronomic practices**	Present production practice should be used as comparator.	irrigated or non-irrigated maize; different pest management practices,
6. Spatial and temporal context	Potential receiving environments should contribute to risk conclusions to be placed into a **spatial and temporal** context.		expected adoption rate of maize, rotation practice
7. Risk management implications	Potential receiving environments should allow **conclusions and suggestion valid for other EU receiving environments**.		

STEP 3: ASSESSMENT ENDPOINTS AND LIMITS OF CONCERNS

According to the ecological function and guild of species, examined above, applicants shall set assessment endpoints and limits of concern (see examples on aphids and coccinellids 1.4. and protected *Lepidoptera* 1.5).

STEP 4: SELECTION OF FOCAL SPECIES

4.1. Construction of a faunal list. An example is shown in Table 1.
4.2. Prioritization of species according to ecological criteria. In Table 1 an exercise to rank species based on the proposed criteria is suggested, and it is based on an hypothetical case of a Cry3-expressing maize.

Functional group	Taxon	Environment	Feeding habits	Exposure	Sensitivity	Abundance	Linkage to the production system	Vulnerability	Trophic interactions with target species	Relevance to semi-natural habitats nearby	Rating
Predators	**Coleoptera: Coccinellidae**	Plant canopy	Pollen and insects	Directly from plant tissues + indirectly through prey	Likely	Low to high	Medium	No	No	High	A
Predators	**Neuroptera: Chrysopidae**	Plant canopy	Pollen and insects	Direct+indirect	Less likely	Low	Medium	Unlikely	No	Medium	B
Predators	**Coleoptera: Carabidae**	Above and below ground	Wide range of foods including insects	Mostly indirect	Likely	High	Medium	Unlikely	Yes	High	A
Predators	**Soil mites**	Below ground	Wide range of insects including eggs	Likely	Unlikely	Medium	Low	Unlikely	Yes (by eggs)	Low	C
Predators	**Spiders**	Above and below ground	Generalist predator	Mostly indirect	Unlikely	High	Low	Unlikely	Yes	Medium to high	B/C
Herbivores	**Homoptera: Aphidiidae**	Plant canopy	Sap-feeders	Unlikely	Unlikely	Low to high	High	Unlikely	No	Low	B
Herbivores	**Coleoptera: Elateridae**	soil	Root feeders, tissue chewers	Direct	Likely	Low to medium	Medium	Unlikely	No	High	A
Herbivores	**Homoptera: Cicadellidae**	Plant canopy	Sap feeders	Unlikely	Unlikely	Medium to high	Medium	Unlikely	No	Medium	B

Herbivores	**Acarinae**	Plant canopy	Cell –content feeders	Likely	Unlikely	Medium to high	Low	Unlikely	No	High	A
Herbivores	**Thysanoptera**	Plant canopy	Cell –content feeders	Likely	Unlikely	Medium to high	Low	Unlikely	No	High	A
Herbivores	**Nematodes**	soil	Cell –content feeders	Likely	unlikely	Medium to high	Medium	Unlikely	Indirect	Low	A
Herbivores	**Lepidoptera**	Plant canopy and soil	Several plant parts, tissue chewers	Likely	unlikely	Low to high	Medium to high	Unlikely	Indirect	Medium to high	A
Herbivores	**Coleoptera: Chrysomelidae**	Plant canopy	All plant parts, tissue chewers	Direct	Likely	Low to high	High	Unlikely	Indirect	Medium to high	A
Entomopathogenic organisms	**Entomopathogenic nematods**	soil	Insects as host	Indirect	Unlikely	Medium	High	Unlikely	Yes	Low	A
Parasitoids	**Hymenoptera: Braconidae**	Above ground	Insects as larvae Nectar and other possible sugar sources as adults	Unlikely	Unlikely	Medium (linked to host populations)	Low	Unlikely	Only for some species	Medium to high	B
Parasitoids	**Hymenoptera: Aphidiinae**	Above ground	Insects as larvae Nectar and other possible sugar sources as adults	Unlikely	Unlikely	Medium (linked to host populations)	Low	Unlikely	No	Medium to high	B

Parasitoids	**Hymenoptera: Trichogramatidae**	Above ground	Insects eggs as larvae Nectar and other possible sugar sources as adults	Indirect	Unlikely	Medium (linked to host populations)	Medium to high	Unlikely	No	Low	B
Parasitoids	**Diptera:Tachinidae**	Above ground	Insects as larvae Nectar, pollen and other possible sugar sources as adults	Direct and indirect	Unlikely	medium	Medium to high	unlikely	Only for some species	medium	A/B
Pollinators / Pollen feeders	**Hymenoptera: Apiidae**	Plant canopy	Pollen and nectar	Direct	Unlikely	Medium	Low	Likely	No	High	A
Decomposers	**Collembola**	Soil	Decaying organic matter, including bacteria and fungi	Direct and indirect	Unlikely	High	Low	Unlikely	Possible	Unlikely	A
Decomposers	**Nematods**	Soil	Decaying organic matter, including bacteria and fungi	Direct and indirect	Unlikely	High	Low	Unlikely	Possible	Unlikely	A
Decomposers	**Enchytraeids**	Soil	Decaying organic matter, including bacteria and fungi	Direct and indirect	Unlikely	High	Low	Unlikely	Unlikely	Unlikely	B

Decomposers	**earthworms**	Soil	Decaying organic matter, including bacteria and fungi	Direct and indirect	Unlikely	High	Low	Unlikely	Unlikely	Unlikely	B
Decomposers	**mites**	Soil	Decaying organic matter, including bacteria and fungi	Direct and indirect	Unlikely	High	Low	Unlikely	Unlikely	Unlikely	B

Conclusion from the prioritization process:

- Based on the above ranking the ladybird *Coccinella septempunctata* and the guild of ground beetles are selected for specific studies in the tiered approach.
- NT herbivores in this species selection process are not deemed important apart from aphids that however will be studied during tritrophic experiments for the impact on parasitoids.
  Specific hypotheses-driven experiments shall be performed according to the tiered approach illustrated in Figure 6 in the text.

STEP 5: DEFINITION OF MEASUREMENT ENDPOINTS AND TESTING PROTOCOLS (as described in the opinion)

STEP 6: INFORMATION NEEDED FOR DETECTION OF UNINTENDED EFFECTS

These type of data are relevant in detecting potential unintended effects and demonstrate through a weight-of-evidence approach the absence of unintended effects that might have adverse impact on NTO focal species (see Step 4). The following types of data are useful to support the weight of evidence of lack of unintended effects. However, *in planta* generated data shall always be included in data set provided.

6.1. Molecular Characterisation data

- promoters,
- insertion site,
- flanking regions,
- stability of the inserts,
- similarity newly expressed proteins with their natural form (e.g. changes in some amino-acids),
- protein expression data for leaves, stalk, grains, roots and pollen,
- ranges of expression of new proteins (EPA, 1999*),*
- variations over growing season for plant tissues.

6.2. Extended compositional analysis

To focus on plant parts which express the transgenic product(s). In case of maize: seeds, leaves, flowers/pollen (depending on the promoter being used) and roots could be analyzed. Since in maize there is not a specific class of compounds for which direct quantitative relationships with the biology of some NTO species are known, these analyses could be focussed on compounds already being analyzed for food/feed purposes.
Protein content in pollen is considered the best proxy to estimate pollen quality.

6.3. Agronomic and phenotypic data

- Plant height, flowering time,
- Seeds and pollen morphology,
- Susceptibility to pests and diseases (the field should be of an adequate size to effectively monitor very mobile pests).

6.3. Plant-environment interactions

In planta studies for unintended effects:

- "semi-field" conditions will be accessible to parasitoids. Artificially infested maize plants will attract *Aphydiinae* and the overall parasitisation rate can be easily scrutinized by counting mummies;
- A confined study with *Collembola* (i.e. litter bag) is feasible and useful for studying biological parameters of this taxon.

GLOSSARY

Note: The definitions provided in this glossary are to be considered in the context of the EFSA guidance on ERA of GM plants and the scientific opinion on the assessment of potential impacts of GM plants on NTOs.

Assessment endpoint: is defined as a natural resource or natural resource service that needs protection. It is the valued attribute of a natural resource worth of protection (Suter, 2000).

Baseline: is defined as a point of reference against which future changes can be compared (EC, 2002).

Biogeographical region or zone: is defined as spatial scale of Earth's surface containing related biotic (e.g. fauna and flora) and abiotic (e.g. climate, soil, or elevation) conditions.

Case-by-case: is defined as the approach by which the required information may vary depending on the type of the GMOs concerned, their intended use and potential receiving environment, taking into account i.a. GMOs already in the environment (EC, 2001).

Deliberate release: is defined as any intentional introduction into the environment of a GMO or a combination of GMOs for which no specific containment measures are used to limit their contact with and to provide a high level of safety for the general population and the environment (EC, 2001).

Desk study: is defined as an investigation of relevant available information, often before starting practical study of a problem.

Ecosystem services: include all services provided by ecosystems, e.g. production of food, fuel, fibre and medicines, regulation of water, air and climate, maintenance of soil fertility, cycling of nutrients. Ecosystems services are distinct from ecosystem functions by virtue of the fact that humans, rather than other species, benefit directly from these natural assets and processes (MEA, 2005).

Effects:

Adverse effects: are defined as a harmful and undesired effects consisting of measurable changes of protected entities (*e.g.* change in a natural resource or measurable impairment of a natural resource service) beyond accepted ranges.

Unintended effects: are defined as consistent differences between the GM plant and its conventional counterpart, which go beyond the primary intended effect(s) introducing the target gene(s).

Direct effects: are defined as primary effects on human health or the environment which are a result of the GMO itself and which do not occur through a causal chain of events (EC, 2001).

Indirect effects: are defined as to effects on human health or the environment occurring through a causal chain of events, through mechanisms such as interactions with other organisms, transfer of genetic material, or changes in use or management (EC, 2001).

Immediate effects: are defined as effects on human health or the environment which are observed during the period of the release of the GMO. Immediate effects may be direct or indirect (EC, 2001).

Delayed effects: are defined as effects on human health or the environment which may not be observed during the period of the release of the GMO, but become apparent as a direct or indirect effects either at a later stage or after termination of the release (EC, 2001).

Cumulative long-term effects: are defined as the accumulated effects of consents on human health and the environment, including flora and fauna, soil fertility, soil degradation of organic material, the feed/food chain, biological diversity, animal health and resistance problems in relation to antibiotics (EC, 2001).

Environmental harm: is defined as a measurable adverse change in a natural resource or measurable impairment of a natural resource service which may occur directly or indirectly (EC, 2004).

Environmental risk assessment: is defined as the evaluation of risks to human health and the environment, whether direct or indirect, immediate or delayed, which the deliberate release or the placing on the market of GMOs may pose and carried out in accordance with Annex II (EC, 2001).

Fitness: is defined as the number of seeds (or propagules) produced per seed sown, and includes the whole life cycle of the plant (Crawley et al., 1993). Enhanced fitness can be defined as a characteristic of an individual or subpopulation of individuals that consistently contribute more offspring to the subsequent generation (Wilkinson and Tepfer, 2009).

Functional groups: are defined as non-phylogenetic, aggregated units of species sharing an important ecological characteristic and playing an equivalent role in the community (Cummins, 1974; Smith et al., 1997; Steneck, 2001; Blondel, 2003).

Genetically modified organism (GMO): is defined as an organism, with the exception of human beings, in which the genetic material has been altered in a way that does not occur naturally by mating and/or natural recombination (EC, 2001).

Hazard (harmful characteristics)**:** is defined as the potential of an organism to cause harm to or adverse effects on human health and/or the environment (EC, 2002).

Limits of concern: are defined as the minimum ecological effects that are deemed biologically relevant and that are deemed of sufficient magnitude to cause harm. These limits of concern are set for each assessment endpoint in the problem formulation.

Measurement endpoint: is defined as a quantifiable indicator of change in the assessment endpoint, and constitute measures of hazard and exposure (e.g. fitness, growth, behaviour, development).

Problem formulation: is defined as the process including the identification of characteristics of the GM plant capable of causing potential adverse effects to the environment (hazards) of the nature of these effects, and of pathways of exposure through which the GM plant may adversely affect the environment (hazard identification). It also includes defining the assessment endpoints and setting of specific hypothesis to guide the generation and evaluation of data in the next risk assessment steps (hazard and exposure characterisation).

Production system: is defined as the specific use of the GM plant, the context in which the GM plant is grown, its cultivation (including crop rotation), harvesting and management, and the genetic background in which the transgenic trait has been introduced.

Protection goals: are defined as natural resources (e.g. arthropod natural enemies, bees) or natural resource services (e.g. regulation of arthropod pest populations, pollination) that are to be protected as set out by EU legislations.

Risk: is defined as the combination of the magnitude of the consequences of a hazard, if it occurs, and the likelihood that the consequences occur (EC, 2002).

Receiving environment: is defined as the environment into which the GM plant(s) will be released and into which the transgene(s) may spread.

Stacked events: are GM plants in which two or more single events have been combined by conventional crossing.

Step-by-step approach: is used in this ERA GD to describe the six assessment steps (1. Problem formulation; 2. Hazard characterisation; 3. Exposure characterisation; 4. Risk characterisation; 5. Risk management strategies and 6. Overall risk evaluation and conclusions) for the ERA. This assessment approach is different from the Stepwise approach defined hereunder.

Stepwise Approach: is defined as all the steps (used in the sense of 'containment-level') beginning with experiments in the contained use system through temporarily and spatially restricted deliberate release up to placing on the market, where data should be collected stepwise as early as possible during the procedure (EC, 2002).

Stressor: the GM plant itself, the transgene(s) in this organismal context and its products, are all considered as potential stressor.

Weight-of-evidence approach: is defined as the use of scientific evidence from various data sources to support assessment conclusions.