Integrated Environmental Assessment and Management — Volume 6, Number 2—pp. 287–300

General Recommendations for Soil Ecotoxicological Tests Suitable for The Environmental Risk Assessment of Genetically Modified Plants

Jörg Römbke,† Stephan Jänsch,† Matthias Meier,‡ Angelika Hilbeck,‡ Hanka Teichmann,§ and Beatrix Tappeser§*

†ECT Oekotoxikologie, Böttgerstrasse 2-14, D-65439 Flörsheim, Germany
‡EcoStrat, Hottingerstrasse 32, Zürich, Switzerland
§German Federal Agency for Nature Conservation (BfN), Konstantinstrasse 110, Bonn, Germany

(Submitted 30 March 2009; Returned for Revision 13 July 2009; Accepted 24 July 2009)

ABSTRACT

Before a genetically modified plant (GMP) can be placed on the market in the European Union (EU), an environmental risk assessment has to be conducted according to EU-Directive 2001/18/EC or Regulation (EC) No. 1829/2003 of the European Parliament and of the Council. However, no harmonized concept for ecotoxicological testing is available today that considers the characteristics of GMPs as a whole. In fact, to date, mainly ecotoxicological tests originally developed and standardized for pesticides are used for this purpose. Frequently in these tests, not the whole GMP is tested but only specific transgene products (mainly toxins). In this contribution, ecotoxicological methods developed for the testing of pesticides are evaluated for whether they are suitable for risk assessment of GMPs as well. In total, 105 test methods covering a wide range of terrestrial invertebrates, microbes, and plants (laboratory, semifield, and field levels) were assessed. Only 7 of them had already been used with GMPs, and in about 20 studies the existing tests methods were modified, mostly in a way such that nonstandard species were used. In the laboratory, few earthworm and nontarget arthropod (NTA) species as well as collembolans and isopods were tested, and, in the field, only the litter-bag test was used. Clearly, more species than these few standard organisms currently in use have to be selected for testing purposes. A more detailed analysis of GMP tests with soil invertebrates published in the literature revealed that some of the relevant GMP exposure routes, such as via bulk soil, soil porewater, and litter from GMPs, are well covered. However, studies addressing either consumption of GMPs themselves or secondary exposure after GMPs have been taken up by invertebrates that feed on living or dead GMPs are underrepresented. Integr Environ Assess Manag 2010;6:287–300. © 2009 SETAC

Keywords: Test methods Invertebrates Microbes Selection criteria Test strategy

Legal and Policy Analysis

INTRODUCTION

Before a genetically modified plant (GMP) can be placed on the market in the European Union (EU), an environmental risk assessment (ERA) has to be conducted according to EU-Directive 2001/18/EC (on the deliberate release into the environment of genetically modified organisms; EC 2001) or Regulation (EC) No. 1829/2003 of the European Parliament and of the Council (EC 2003). Ecotoxicological tests investigating adverse effects of a GMP on the biotic environment are an important element of the ERA. To date, testing of adverse effects of GMPs on nontarget organisms (NTO) relies on ecotoxicological tests originally developed and standardized for pesticides (Candolfi et al. 2000; Romeis et al. 2008). Although this ecotoxicological concept is widely used in the application dossiers of GMPs for regulatory approval, it does not fulfill the requirements of Directive 2001/18/EC. This document explicitly demands an ERA on a case-by-case basis, which considerably broadens the requirements used for pesticide registration. A case is described in Annex II of Directive 2001/18/EC as a combination of the crop plant (its biology, ecology, and agronomy), the novel trait relating to its intended effect and phenotypic characteristics of the GM crop plant (the GMP), and the receiving environment related to the intended use of the GMP. The emphasis on the receiving environment is an outstanding difference between pesticide and GMP ERA. Equally important is the fact that even persistent pesticides, e.g., with a DT90 (time to 90% degradation) greater than 100 or even 365 d, remain stable for only a limited time in the environment. Genetically modified plants, in contrast, are capable of self-reproduction, meaning that they might increase, potentially spread, and exist for an unlimited time in the environment, at least by human standards. Then, they might be impossible to purge if detrimental effects on the environment are detected after their release. For this reason, the Directive 2001/18/EC requires that the testing is carried out ''step-by-step'' when introducing a GMP into the environment. This means that the scale of release may only be gradually increased if no unacceptable risk for human health and the environment has been identified in the previous step. This is a crucial difference from the risk assessment of pesticides, for which a safe use may already be identified at the laboratory level. The extrapolation to the

* To whom correspondence may be addressed:
j-roembke@ect.de
Published on the Web 8/20/2009.
DOI: 10.1897/IEAM.2009-043.1

field level is then accomplished by applying so-called safety or assessment factors (see below for details). This is not possible with GMPs when considering the whole plant and not only the concentration of an isolated novel toxin.

Thus, the following mismatches of the existing approach for the ERA of GMPs with the requirements of Directive 2001/18/EC can be identified.

- Frequently, only specific transgene products (i.e., a toxin) are assessed and not the whole GMP.
- Testing is performed with a small set of standard test species but not those that are ecologically relevant in the receiving environment where the specific GMP is used.
- The primary focus is on the laboratory level, and semifield or field studies are not necessarily asked for, which contradicts the step-by-step principle of a gradual release into the environment.

A detailed critique of the currently applied concept can be found in Hilbeck et al. (submitted).

Hence, no harmonized concept for the ecotoxicological testing of GMPs is available today that fully includes the requirements of Directive 2001/18/EC based on the characteristics of whole GMPs. For the ecotoxicological testing of GMPs, test organisms and methods should be selected according to the case definition given above. We must test whether the GMP, its use, or the transgene product can affect structural (i.e., related to individual species) or functional (i.e., related to ecosystem services provided by the whole community) endpoints.

This paper seeks to learn from the experience in the field of ecotoxicology of pesticides and to evaluate the suitability of pesticide testing methods for GMPs. First, the current ERA model of pesticides in the EU and the resulting ecotoxicological testing requirements and, second, general criteria relevant for the selection of test species and ecotoxicological test methods and for the compilation of test strategies are summarized. Third, existing test methods developed for the evaluation of effects of pesticides on terrestrial organisms (excluding birds and mammals) are compiled. Afterward, a detailed review of the tests already performed with soil invertebrates and GMPs published in the scientific literature is presented. Finally, the suitability of these methods for the toxicity assessment of GMPs is discussed, using general selection criteria (test species, methods, and strategy) and more specific criteria for GMPs. Recommendations are given that existing test methods be modified for GMP testing and that new tests should be developed.

TESTING REQUIREMENTS FOR THE ENVIRONMENTAL RISK ASSESSMENT OF PESTICIDES IN THE EUROPEAN UNION

Already in 1991, the EU published harmonized requirements for placing on the market of plant protection products (EU 1991). Later, details of the testing requirements and the risk assessment procedure were elaborated further (e.g., EU 1997; EC 2002). However, to address an increasing number of requests regarding how these requirements should be interpreted, technical advice was given in semiofficial documents that were often developed as a result of discussions among stakeholders (i.e., agencies, industry, and academia; EPPO 2003; Römbke et al. 2003). In practice, the ERA has to be carried out separately for each environmental compartment (water, sediment, soil, and air). In a first step, the concentration expected in the environment (predicted environmental concentration [PEC]) and the concentration with a specific effect in the environment (no observed effect concentration [NOEC], 50% lethal concentration [LC50], or 50% effective concentration [EC50]) are determined (Leeuwen and Hermens 2001). The toxicity of a pesticide (e.g., mortality or effects on growth or reproduction) is measured in laboratory tests with individual species. To quantify the risk of a pesticide, the quotient between the toxicity value (NOEC, LC50, or EC50) and the exposure value (PEC) is calculated (toxicity exposure ratio [TER]). With the assumption that the pesticide is used according to the principles of good agricultural practice, including manufacturer's intended use, and depending on the comparison of the TER and certain safety or assessment factors (see the Definition of Safety/Assessment Factors below), it is decided whether the use of the pesticide can be considered safe. If such safe use cannot be assumed, the authorities can require safety measures (e.g., buffer zones between treated areas and surface waters, lower application rates). If risks cannot be avoided by safety measures, the pesticide may not be given authorization.

Ecotoxicological test methods for pesticides were developed within the last 30 y in order to identify possible problems occurring when these substances enter the environment. In the mid-1970s of the last century, the Chemicals Testing Program was initiated by the Organisation for Economic Co-operation and Development (OECD) in order to reduce trade barriers. One of the main tasks of this program was to harmonize and to standardize the various test methods that had been developed in several industrial countries (Forbes and Forbes 1994). In the early 1980s, OECD published a set of guidelines for testing the fate and

Definition of Safety/Assessment Factors (EU TGD, 2003, Chapter 3.3, p. 103):

When assessing the environmental risk of chemicals several assumptions have to be made which allow, even though uncertain, an extrapolation from single-species short-term toxicity data to ecosystem effects. In particular it is assumed that ecosystem sensitivity depends on the most sensitive species and that protecting ecosystem structure protects community function. Since for most substances the pool of data is very limited, it is recognized that, while having no strong scientific validity, empirically derived factors must be used. In applying such factors, the intention is to predict a concentration below which an unacceptable effect will most likely not occur. In establishing the size of these factors, a number of uncertainties must be addressed to extrapolate from single-species laboratory data to a multispecies ecosystem:

- intra- and interlaboratory variation of toxicity data
- intra- and interspecies variations (biology variance)
- short-term to long-term toxicity extrapolation
- laboratory data to field impact extrapolation.

In general, the size of an assessment factor, usually covering a range between 1000 and 1, decreases with increasing information on the effects of a substance, i.e. in terms of the number of species tested or trophic levels covered.

effects of chemicals, mainly pesticides (OECD 1984a), the number of which has now increased to 55. Many countries have adopted the OECD guidelines in their national legislation.

The use of standardized test guidelines follows a hierarchical (tiered) order (Fig. 1). The first tier consists of simple, short-term, and low-cost single-species tests. Tests are performed under assumed worst-case conditions and are designed to determine the effects of 1 or repeated applications of the pesticide over a wide range of concentrations (e.g., LC50 value). Depending on the ERA results on this first tier, higher-tier tests, often performed on the semifield or field level, may follow (Cairns 1981; Bradbury et al. 2004). The results of complex field studies, e.g., with earthworms (ISO 1999a,b), clearly are of greater ecological relevance. However, at the same time, these studies are often difficult to interpret because test conditions, such as weather, are rarely reproducible and test results are highly variable. The final aim of the tiered test strategy is to ensure that only substances with a high potential for causing an adverse effect in the environment are tested in complex and resource-demanding tests (in terms of time and costs). However, a tiered strategy is adequate to ensure the protection of the environment only if the criteria for when to stop testing (or to proceed to a higher tier) are scientifically sound (Kareiva et al. 1996; Chapman 2002). In this context, a wrong decision not to proceed with testing (type II error) must clearly be avoided; otherwise, serious consequences for the environment could occur (Forbes and Forbes 1994).

In standardized test methods, usually only a few test species are used. Most of them are easy to cultivate, genetically uniform, and of medium sensitivity to a wide range of pesticides. An ecological relevance of the test organisms is desirable but often plays a secondary role because of practicability considerations (e.g., the main representative for soil organisms is the compost worm *Eisenia fetida* (OECD 1984b) that rarely if ever occurs in soils of arable fields). At the first tier, a minimum set of test species is required (earthworms, microorganisms, plants [6 species], bees, and 2 species of nontarget arthropods [NTAs; parasitic wasps and predatory mites]). On higher tiers, more species can be tested: in soil, collembolans and predatory mites, or, as further representatives of nontarget arthropods, beetles, spiders, or green lacewings (however, the latter have not been standardized by OECD so far [EC 2002]).

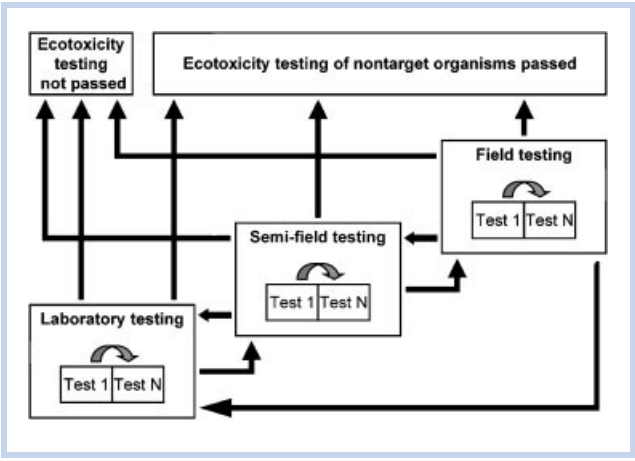

Figure 1. Testing strategy for the registration of pesticides in the European Union. The arrows show the possible lines of consequence depending on testing results. Testing follows a tiered scheme (laboratory, semifield, field), but shortcuts and feedback are possible. The decision on whether ecotoxicity testing is passed or not can occur at all tiers.

Within the ERA for the terrestrial compartment, the risk of an individual pesticide is assessed for each organism group (bees, NTAs, earthworms, microbes, other macrofauna, and plants) independently. It should be noted that birds and mammals are not seen as part of the terrestrial compartment (exception: evaluation of secondary poisoning), so these organisms are not considered in this contribution. Depending on the tiers covered, the number of tests submitted or required for the registration of 1 pesticide active ingredient and its (potentially various) formulations can vary considerably. As an example, the testing efforts for an insecticide that is currently approved by the EU are briefly presented here. This pyrethroid insecticide (active ingredient) has already been produced for about 2 decades and is used globally in many crops. In total, the applicant provided 23 standard tests with soil organisms and 17 tests with bees and nontarget arthropods. In addition, 54 more data sets from specific regions, performed with nonstandard species, were presented. The data package provided is not an exceptionally large case. Rather, the number of tests reflects the broad mode of action of this pesticide and the high number of crops and geographical regions where it is applied.

GENERAL CRITERIA FOR THE SELECTION OF TEST SPECIES, TEST METHODS, AND TEST STRATEGIES

Before addressing the tests used or proposed for the testing of pesticides, some general criteria relevant for all ecotoxicological test methods are briefly summarized (Keddy et al. 1995; Römbke et al. 1996; Løkke et al. 2002). All these criteria have to be considered when new testing methods for GMPs are to be developed (or existing testing methods are going to be modified). They are divided into 3 categories: 1) criteria for test species, 2) test methods, and 3) test strategies (the latter group will not be discussed in detail in this paper).

Criteria for test species

To obtain reproducible and justifiable test results with the identified test methods, the species used have to fulfill the following criteria.

- Ecological relevance: Is the species dominant in terms of abundance or biomass, and does the species play a key role in food webs or act as an ecosystem engineer? Remark: For practicability reasons, often a different representative of the group to which the most suitable test species belongs is chosen.
- Keeping and breeding: Is the species easy to keep and breed in the laboratory? Does it have a quick succession of generations, and is mass breeding possible throughout the year? Remark: It is assumed that no catches in the field are necessary.
- Exposure: Does the species live in close contact with the soil, plants, or plant residues?
- Sensitivity: Is the species moderately sensitive to a wide range of anthropogenic stress factors? Remark: A broad sensitivity spectrum is more important than a high sensitivity to individual stress factors.

- Ecological tolerance: Does the species have a low sensitivity to fluctuations in environmental conditions such as soil properties, temperature, or moisture?
- Distribution: Is the species widely distributed in the environment?

Criteria for test methods

The following criteria should be used for the evaluation of test methods.

- Exposure verification (qualitative and quantitative): Are the most important exposure pathways (e.g., porewater, food) covered? Is the exposure of the test organisms known (e.g., in the case of pesticides, via residue analysis)?
- Standardization: Is the method published as a validated (international) guideline?
- Practicability: Is much effort required to perform the test? Is the test application limited by high demands on laboratory personnel (e.g., taxonomic knowledge of certain animal species)?
- Rejection standards: Can criteria be defined for the validity of the test method (e.g., accepted mortality or effect level in the negative controls and/or the regular testing of reference substances)?
- Documentation and/or experience: Is enough data known from a particular test method to evaluate new results?
- Endpoints: Are the number of endpoints and their sensitivity high enough to react to different modes of action?

Criteria for a test strategy

Four criteria should be used when compiling a test strategy.

- Taxonomy and physiology: Are various taxonomic and physiological groups (e.g., arthropods, oligochaetes) covered?
- Trophic levels: Are various trophic levels (e.g., saprophagous, predatory, etc.) covered?
- Exposure: Are various exposure pathways (including bioaccumulation) covered?
- Endpoints: Are various structural and functional endpoints (i.e., biodiversity and ecosystem services) covered?

Although the first 3 criteria are generally regarded as being relevant for the selection of test methods, the study of effects on ecosystem services such as organic matter decomposition requires an explanation. Often, such effects can be traced back to an effect on key species or ecosystem engineers (Lavelle et al. 1997), but these species do not fulfill the criteria for a laboratory test species (e.g., they cannot be kept in the laboratory). In other cases, small, often undetectable effects on various single species or their interactions may trigger greater effects on a higher level of biological organization (i.e., communities or the whole ecosystem), meaning that field studies such as the earthworm test (ISO 1999a) or the litter-bag test (OECD 2006) can be required.

COMPILATION OF EXISTING TEST METHODS WITH TERRESTRIAL ORGANISMS USED FOR THE ECOTOXICITY TESTING OF PESTICIDES

To evaluate the existing tests for their potential suitability for risk assessment of GMPs, 105 individual test methods used for ecotoxicity testing of pesticides with soil invertebrates, plants, and microorganisms were compiled and described in a uniformly structured way (for details, see Hilbeck et al. 2008). It should be noted that it was not the aim to present a complete list of all test methods for pesticides ever recommended. Only those methods were included in the compilation that have been used in ecotoxicological studies and have been standardized or are in the process of being standardized. Each method was classified according to its investigation level (laboratory, semifield, or field) and within each group according to the respective taxonomic group. The criteria used for the evaluation included information on the standardization status, the trophic level of the test organism (Groot and Dicke 2002), the test design, the substrate used, the endpoints measured, the duration, the experimental conditions, and any other information that was deemed of importance for the risk assessment of GMPs. In particular, measures to ensure the reproducibility of the test results (e.g., validity criteria and the use of reference test substances) were pointed out. Finally, it was considered whether the test was already used for the assessment of GMPs (for details, see Hilbeck et al. 2008). These test methods were evaluated according to 5 different criteria (Table 1): 1) Testing level (i.e., level of ecological complexity), 2) taxonomic group, 3) ecological group, 4) status of standardization, and 5) whether the method has already been used for GMPs.

Testing level

By far, most of these tests are laboratory tests (86), whereas 13 are semifield methods, and only 6 were developed for field studies.

Taxonomic group

The test species used belong to a wide range of taxonomic groups (Table 1): microbes, plants, nematodes, oligochaetes, insects, isopods, arachnids, and others. Eight of the 105 methods were classified as multispecies test systems. Because the groups listed here are inconsistent in their taxonomic representation, the resulting numbers are of limited value. For example, although all plants are summarized in 1 category, animals are divided into 6 groups. As mentioned above, the number of tests is governed mainly by the legal requirements for the registration of pesticides. For example, many oligochaete tests were developed not only because large earthworms are widely distributed and ecologically relevant but also because they were accepted by the public and in the scientific community as the most important soil invertebrates (Lavelle et al. 1997). Together with their easy handling and testing, they seemed to be the perfect surrogate for the vast number of soil invertebrate species. In contrast, the development of the many insect tests was driven by the implementation of integrative plant protection measures. To identify those pesticides that had side effects on beneficial arthropods (mainly predators or parasitoids of pest species), there was a need for appropriate methods.

Ecological group

Test species were classified according to the following ecological groups (Table 1): primary producers, decomposers, consumers first-order (herbivores), consumer second-order

Table 1. Results of a literature review: Classification of terrestrial test methods (excluding those for birds and mammals) according to 5 criteria (for details see Hilbeck et al. 2008)

	Level of testing		
	Laboratory	**Semifield**	**Field**
All tests	86	13	6
Taxonomic group[a]			
Microbes	11	0	0
Plants	13	0	0
Nematodes	7	0	0
Oligochaetes	15	0	1
Insects	26	6	2
Isopods	3	1	0
Arachnids	9	0	1
Others	3	0	0
Multispecies	0	6	2
Ecological group[a]			
Primary producers	13	0	0
Decomposers	40	2	2
Consumers, first-order (herbivores)	10	1	0
Consumers, second-order (predators)	28	4	3
Consumer third-order (parasitoids)	2	0	0
Pollinators	3	2	2
Not assignable	0	5	0
Standardization[a]			
OECD	10	0	1
ISO	18	0	1
IOBC	10	1	1
EPPO	2	1	1
USA (ASTM, etc.)	7	0	0
Environ. Canada	6	0	0
Other nations	6	1	2
Literature proposals	33	10	0

[a]Total number is higher than 105 because in some methods more than 1 taxonomic (e.g., 2-species test) or ecological (1 species can be part of more than one trophic level) group is covered, and sometimes 1 method is standardized by more than 1 organization.

(predators), consumer third-order (parasitoids), pollinators, and not assignable. The resulting numbers clearly reveal a bias for decomposers (mainly because of the many oligochaete tests), whereas the pollinators and primary producers are underrepresented. Also, the number of tests with herbivores is lower than one would expect given their ecological importance. Within the group of first-order consumers, only few species are truly living on green plant material; many are in fact bacterial feeders, such as the nematodes included here. Also, more than 2 tests with parasitoids have been proposed in the literature, but none of them is sufficiently standardized to be included in this survey. Finally, the tests that could not be assigned to a group cover whole communities (e.g., multispecies tests).

Status of standardization

Standardization of tests is carried out by international organizations (OECD, the International Organization for Standardization [ISO], the International Organization for Biological Control [IOBC], European Plant Protection Organization [EPPO]) and also by national agencies (the American Society for Testing and Materials [ASTM], US Environmental Protection Agency [USEPA], US Food and Drug Administration [USFDA], Environment Canada, German Biologische Bundesanstalt, etc.; Table 1). In addition, several test methods have been proposed in the literature, often with the aim of becoming standardized after some experience has been collected. Among the 105 terrestrial test methods described, more than one third (43) are proposals from the open literature, which usually are acceptable only as additional information for ERA of pesticides. Experiences gained in some of the research projects described in the open literature have been incorporated into test guidelines. In addition, it is possible to use a method from the literature for higher tier testing, i.e., in cases when specific situations do occur.

TESTS ALREADY PERFORMED WITH GMPs

Tests performed according to standardized guidelines

A comparative literature review including the scientific literature as well as application dossiers and unpublished reports revealed that the following methods originally developed for pesticides have already been used for the assessment of GMPs.

- Ammonium oxidation—rapid test (ISO 2002a)
- Microbial soil respiration (ISO 2002b)
- Earthworm acute mortality, reproduction, and avoidance behavior tests with *Eisenia fetida* (OECD 1984b; ISO 1999a, 2007)
- Collembolan reproduction test with *Folsomia candida* (ISO 1999b)
- Breakdown of organic matter in soil (litter-bag test; OECD-Guidance Document 56)

These 7 test methods reflect the legal requirements of the ERA for pesticides or were developed for this purpose.

Review of soil invertebrate laboratory tests with GMPs

Here, those laboratory studies published in the scientific literature are reviewed that were carried out in the context of effect determination of GMPs using true soil-inhabiting invertebrate species and actual GMP material (i.e., the whole-plant approach). This review includes those tests that were not performed according to standard guidelines but were modified to suit the specific requirements of GMPs. Only few laboratory tests have been described that assess the effects of GMPs on single species of soil invertebrates.

Earthworms. Ahl Goy et al. (1995) exposed *Eisenia fetida* to leaf extracts of ECB (European corn borer)-tolerant maize, corresponding to 0.35 mg of the delta-endotoxin CryIA(b)/kg soil, in artificial soil for 14 d, hence probably following the standard OECD guideline 207 (OECD 1984b). They assumed the concentration as being 785 times higher than the expected concentration in the soil, when maize plants will be incorporated into the soil after harvest. No effects on survival or weight gain were observed in comparison with nonspecified control maize. No details are given concerning the extraction method of the *Bacillus thuringiensis* (*Bt*) toxin from the leaves, and no estimation can be made for how the toxin might be comparable to a situation in which *Bt* toxin enters the soil from decaying leaf material in the field.

Saxena and Stotzky (2001) performed laboratory tests with commercially purchased *Lumbricus terrestris* in a natural field soil planted with *Bt* (NK4640Bt) or isogenic non-*Bt* maize for 40 d and in soil amended with ground, air-dried biomass (leaves, stems, and roots) of *Bt* or non-*Bt* maize (1% plant material in 500 g soil) for 45 d. No significant differences in mortality and earthworm weight were observed. The presence of the *Bt* toxin in the soil from the earthworms' guts was verified at test end through immunological assays and bioassays. However, the amount of the *Bt* toxin was not quantified.

Zwahlen et al. (2003) fed adult field-collected *L. terrestris* with N4640Bt and isogenic maize leaf litter in field soil for 200 d. The plant material was not incorporated into the soil but placed on the soil surface according to the feeding habits of *L. terrestris*. The initial Cry1Ab toxin concentration in the transgenic *Bt* leaves was 15.5 μg/g dry weight leaf. The toxin concentration decreased to 1.2 μg/g dry weight leaf during the first 40 d of the trial but remained at a level of 0.2 to 0.7 μg/g dry weight leaf until the end of the trial. No lethal effects were observed. No statistically significant differences in relative weights were observed during the first 160 d of the trial, but, after 200 d, adult *L. terrestris* had a statistically significant weight loss of 18% of their initial weight when fed *Bt* maize litter compared with a weight gain of 4% in non-*Bt* maize fed earthworms.

Vercesi et al. (2006) performed various tests with field-collected *Aporrectodea caliginosa* in natural soil. Finely ground leaves of MEB307 *Bt* and near-isogenic maize were incorporated into the soil up to concentrations of 5 g/kg dry weight soil. Cow dung was supplied as an additional food source. The content of the *Bt* toxin Cry1Ab was determined to be 9.6 μg/g in MEB307 (dry leaves). Adult and juvenile earthworms were exposed for 28 d and 14 weeks, respectively. No effects on survival, growth, development, or cocoon production were observed in the *Bt* maize treatments. However, a slight but statistically significant effect on cocoon hatchability was observed, with an NOEC of 3 g dry mass/kg and an EC10 of 4.2 g dry mass/kg soil. Growth of juvenile *A. caliginosa* was unaffected when the earthworms were kept in pots with a growing *Bt* maize plant for 28 d. The fungicide benomyl served as a positive control but did not always show statistically significant effects. Hence the sensitivity of the test system and the exposure to the test item could not always be demonstrated. Thus, the validity of some of the studies' results remains unclear.

Collembolans. Yu et al. (1997) fed leaf discs or milled leaves of transgenic cotton lines 81 and 249 (control: parent variety Coker 312) and transgenic and nontransgenic potato leaves to *Folsomia candida* on a field soil for 7 to 8 weeks. For the transgenic cotton lines, expression rates were up to 0.1% soluble Cry1Ab protein or 10 to 25 μg protein/g fresh weight plant tissue. Transgenic potato leaves had an expected expression of 0.1% soluble Cry3A protein or 10 to 20 μg protein/g fresh weight of plant. Bean leaves soaked with

cadmium nitrate served as a positive control. No effects on body length or reproduction parameters of *F. candida* were observed with the *Bt* treatments.

Romeis et al. (2003) fed dried root material of Greina and Golin KP4 (killer protein) transgenic and nontransgenic (isolines) wheat varieties to *F. candida* on plaster of Paris and activated charcoal. The animals were exposed individually and in groups of 10 until after the third oviposition and for 8 weeks, respectively. Food material was provided ad libitum in a 1:10 mixture with baker's yeast on small pieces of filter paper and renewed every week. No effects on life-history parameters mortality, oviposition, cluster size of oviposition bouts, skipping of oviposition bouts, insect weight after third egg laying, or egg viability were observed.

Heckmann et al. (2006) investigated the effects of feeding dried ground root tissue of 2 *Bt* maize varieties (Cascade and MEB307) and their isogenic varieties (Rivaldo and Monumental) to laboratory-cultured *Protaphorura armata* on plaster of Paris and activated charcoal for 4 weeks. The amount of Cry1Ab expressed in the root tissue was determined to be 1.37 and 1.01 μg/g for varieties Cascade and MEB307, respectively. No effects on mortality and body surface area were observed.

Nematodes. Saxena and Stotzky (2001) observed no effect on the number of nematodes in a natural field soil planted with *Bt* (NK4640Bt) or isogenic non-*Bt* maize for 40 d and in soil amended with ground, air-dried biomass (leaves, stems, and roots) of *Bt* or non-*Bt* maize (1% plant material in 500 g soil) for 45 d. The experiments were conducted in parallel with tests performed with the earthworm *Lumbricus terrestris* (see above).

Isopods. Escher et al. (2000) fed pre-decomposed X4335-EPR *Bt* maize and isogenic maize leaves to field-collected *Porcellio scaber* on plaster of Paris and activated charcoal for 8 d in food-choice experiments and for 7 months in a reproduction trial. No differences in consumption of either *Bt* or non-*Bt* maize were found. There was also no difference in the number of juveniles per female. Differences in juvenile mortality and adult and juvenile weight gain were found to be related to higher food quality of *Bt* maize resulting from a slightly lower C/N ratio, a lower lignin content, and a higher content of soluble carbohydrates. No measurements on the level of toxin expression were performed.

Wandeler et al. (2002) performed 20-d feeding experiments with field-collected *P. scaber*, 2 *Bt* maize varieties (Max88 and N4640Bt), and 6 conventional varieties (N4640 being isogenic to N4640Bt) on plaster of Paris. Analysis of the 2 *Bt* maize varieties by enzyme-linked immunosorbent assay (ELISA) indicated an initial Cry1Ab toxin concentration of 19.7 and 2.9 μg/g dw in N4640Bt and Max88, respectively. After 20 d, the toxin concentration decreased to 15.5 and 1.1 μg/g, respectively. The presence of the *Bt* toxin in the isopods' gut was verified after the experiment by ELISA. *P. scaber* fed statistically significant less from N4640Bt leaves than from its control N4640. Max88 was consumed statistically significantly more than N4640Bt, but there was no statistically significant difference from N4640. Within the 6 nontransgenic maize varieties, a wide range of consumption was detected. The transgenic maize variety N4640Bt equaled the poorly consumed varieties, whereas Max88 was one of the most often consumed varieties.

Oribatid mites. Yu et al. (1997) fed leaf discs or milled leaves of transgenic cotton lines 81 and 249 (control: parent variety Coker 312) leaves to *Oppia nitens* on a field soil for 7 weeks. For the transgenic cotton lines, expression rates were up to 0.1% soluble Cry1Ab protein or 10 to 25 μg protein/g fresh weight plant tissue. No effects on population growth rates of *O. nitens* were observed with the *Bt* treatments.

DISCUSSION OF THE SUITABILITY OF THESE METHODS FOR THE TOXICITY ASSESSMENT OF GMPs

Here, soil invertebrate species and methods that have already been used for the toxicity assessment of GMPs are discussed. This includes whether they fulfill the above-mentioned general selection criteria for test species and methods as well as their specific suitability for the evaluation of GMPs. The selection criteria for the compilation of a test strategy are not considered here, because the methods reviewed are not part of a coordinated test strategy but isolated tests originating from the scientific community.

Test species

The results for assessment of whether the species tested so far fulfill the general criteria for the selection of test species are given in Table 2. Here these species are discussed in more detail, in particular concerning ecological relevance and keeping and breeding efforts.

As the most commonly used test organisms for ERA of pesticides, *Eisenia fetida* and *Folsomia candida* are an obvious first choice for conducting ecotoxicological tests with GMPs. However, these are standard test organisms that were selected primarily because of their amenability to laboratory culturing and sensitivity to a wide range of pesticides or heavy metals, but they do not usually occur in agricultural habitats and, hence, are not necessarily ecologically relevant (Jänsch et al. 2005).

Lumbricus terrestris is a deep-burrowing (anecic) species and ecologically a highly important ecosystem engineer in central European agricultural soils. It feeds on soil surface plant litter. It has repeatedly been used in ecotoxicological assays but is difficult to handle because of its size and long life cycle, and it cannot easily be cultured in the laboratory on a mass scale. For this reason, it has not been included as a standard test species in the laboratory assessments of pesticides in Europe, but its occurrence and the effects on this species are important parameters for the performance and evaluation of the standardized earthworm field trial (ISO 1999a). The same is true for *Aporrectodea caliginosa*, a horizontal-burrowing inhabitant of the upper mineral soil (endogeic).

Terrestrial isopods (e.g., *Porcellio scaber*) belong to the soil macrofauna and live mainly close to the soil surface or even in the litter layer. Although they usually have a minor role in central or northern European regions, their importance is clearly higher in the Mediterranean area. Especially at sites with often dry soils, they are important decomposers, which, together with millipedes or ants (and termites in the tropics), can take over the role of earthworms more or less completely (Garcia 2004). Some species, such as *P. scaber*, can be kept and bred in the laboratory quite well. They have been increasingly used when investigating the importance of

Table 2. Fulfillment of general selection criteria for ecotoxicological test species

Test species already used for GMP	Ecological relevance	Keeping/ breeding	Exposure	Sensitivity	Ecological tolerance	Distribution	References
Eisenia fetida	Low	Easy	Given	Broad	High	Low	Ahl Goy et al. 1995
Lumbricus terrestris	High	Difficult	Given	Broad	Medium	Wide	Saxena and Stotzky 2001; Zwahlen et al. 2003
Aporrectodea caliginosa	Medium	Difficult	Given	Broad	High	Wide	Vercesi et al. 2006
Folsomia candida	Low	Easy	Given	Broad	High	Wide	Romeis et al. 2003; Yu et al. 1997
Protaphorura armata	Unknown	Easy	Given	Broad	High	Wide	Heckmann et al. 2006
Nematoda	Low to medium	Easy to difficult	Given	Broad	Unknown	Wide	Saxena and Stotzky 2001
Porcellio scaber	Medium	Easy	Given	Broad	Medium	Wide	Escher et al. 2000; Wandeler et al. 2002
Oppia nitens	Unknown	Easy	Given	Unknown	Unknown	Wide	Yu et al. 1997

exposure pathways via food as well as in studies looking at bioaccumulation of pesticides.

The ecological relevance of other soil invertebrate species is less well investigated. For example, the influence of nematodes and oribatid mites on soil processes and functions is certainly important when considering their extremely high numbers in many soils (Petersen and Luxton 1982). However, it is very problematic if not impossible to identify individual species responsible for these activities because of their high taxonomic diversity (at 1 site, easily more than 100 species and subspecies can occur). In addition, only parasitic (and, thus, economically relevant) nematode species have been investigated in laboratory or field tests studying the consequences of anthropogenic stress. In the case of oribatid mites, the situation is even worse: because of their disputed taxonomy, very few data sets concerning their ecological relevance or their reaction to anthropogenic stress are available.

A limiting factor in the use of ecologically relevant species for the assessment of GMPs will be difficulties in breeding and handling of a certain species because of its specific biology and ecological requirements. Species that will not easily reproduce in the laboratory on a reasonable time scale will obviously not be a suitable subject for reproduction or life-cycle assays. In some cases, it may be acceptable to collect animals from the field, but this has the obvious disadvantage that availability of test animals might be limited and vary strongly between seasons and ecological regions. Also, the quality (e.g., age, individual fitness) of the test animals will strongly vary in space and time, affecting the comparability and reproducibility of studies. Selection and quality criteria for field collection of test species would have to be very well defined and strictly followed. In the studies reviewed here, some species were field collected (*L. terrestris*, *A. caliginosa*, *P. scaber*), whereas others originated from laboratory cultures (*E. fetida*, *F. candida*, *Protaphorura armata*, *Oppia nitens*).

All species included in this review live within or in close contact with the soil environment compartment. Hence, it can be assumed that all of them will potentially be exposed to GMP material through the soil pathway. Previous experience (excluding O. *nitens*) from the assessment of pesticides suggests broad sensitivity of these species to a wide range of toxicants. Their ecological tolerance can be classified as medium to high (unknown for nematodes and O. *nitens*). All of these species are widely distributed throughout central Europe, except *E. fetida*, which is commonly found only in anthropogenic habitats of rich organic matter, such as compost heaps.

Test methods

Exposure. Test parameters such as duration, temperature, light regime, moisture, and substrate should be chosen with respect to the specific GMP, the test organism, the intended exposure route, and the observed endpoints. Ideally, a situation should be created that resembles realistic field conditions as closely as possible while retaining the intended advantages of laboratory trials such as short duration, practicability, controllability, low variability, and repeatability. This means that, for example, natural soils should be favored, although in certain cases the use of artificial substrates such as OECD artificial soil or plaster of Paris might be more appropriate. In the studies listed in Table 2, artificial soil was used in 1 study (Ahl Goy et al. 1995), and 4 studies were carried out with natural soils (Yu et al. 1997; Saxena and Stotzky 2001a; Zwahlen et al. 2003; Vercesi et al. 2006). The remaining 4 studies, all with arthropod test organisms, used plaster of Paris and activated charcoal (Escher et al. 2000; Wandeler et al. 2002; Romeis et al. 2003; Heckmann et al. 2006). In most cases, when choosing natural soils as testing substrates, one must again consider the demands of the test species as well as the potential receiving environment of the GMP. In this respect, soil classification concepts such as BBSK (Soil Biological Site Classification; Römbke and Breure 2005) and Euro-Soils (Römbke and Amorim 2004) could provide useful assistance.

In the tests for soil invertebrates published so far, mainly exposure through direct feeding on dead GMP material has been assessed. Generally, this is a reasonable approach for an initial assessment of toxin-expressing GMPs. In the case of

earthworms, 2 studies assessed exposure through direct feeding by mixing GMP material into the soil (Saxena and Stotzky 2001; Vercesi et al. 2006). The same 2 studies also tried to address exposure to root exudates.

Studies using GMP material have often estimated exposure by quantifying the amount of *Bt* toxin present in the plant material. A much more accurate approach to determine exposure is by measuring the presence of the toxin in the test animal. This was done in only 1 study (Saxena and Stotzky 2001). Care must be taken when feeding animals with plant material that would not normally be their preferred food source in the GMP-receiving environment (see, e.g., Yu et al. 1997; Romeis et al. 2003). The influence of food quality has to be considered, because food components may mask detrimental effects of the GMP or may produce false-positive results.

One study was performed using *Bt* toxin extracted from the GMP and incorporated into the test soil (Ahl Goy et al. 1995). The methodology of extraction may affect the structure of the toxin and lead to an altered exposure situation compared with that to be expected under field conditions. Using *Bt* toxins extracted from transgenic *Bt* plant material is yet a better approximation of a realistic exposure than using microbial surrogate toxin, but a validated methodology for extraction must still be developed, including verification of the bioactivity.

Tests investigating the effects of secondary poisoning of soil organisms (i.e., tests with organisms of higher trophic levels such as predatory mites) are missing to date. After all, the plausibility and realism of the exposure route are critical, although laboratory tests are always simplified approximations of realistic exposure scenarios compared with the field. A battery of test organisms for the assessment of GMPs should cover all relevant exposure routes.

Standardization. Generally, any test system proposed for an ERA of GMPs should potentially be able to meet the requirements of ISO and OECD standardization and quality standards of GLP (good laboratory practice) to allow for a transparent, repeatable, and justifiable evaluation of GMP. For example, GMPs used in tests have to be characterized in a standardized way. Among the studies reviewed here, only 2 were performed according to existing guidelines or literature proposals (Ahl Goy et al. 1995; Vercesi et al. 2006). All other methods were newly developed or so strongly modified that they do not fulfill many standardization criteria.

Practicability. The duration of the studies evaluated varied between 8 d and 7 months (Escher et al. 2000). Most studies lasted for 4 to 8 weeks, which can be considered a reasonable time frame for a laboratory assessment of GMP. From the usually rather brief descriptions in the scientific literature, the actual effort in performing the studies is difficult to estimate. However, it can be assumed that those studies using species easy to keep and breed and with a short duration and simple test setup require relatively little testing effort (e.g., Ahl Goy et al. 1995; Yu et al. 1997; Romeis et al. 2003; Heckmann et al. 2006).

Those tests performed with field-collected animals and with a long duration and rather complex test setup require much more training and effort (e.g., Escher et al. 2000; Zwahlen et al. 2003; Vercesi et al. 2006).

Rejection standards. The choice of an appropriate control is crucial. Hence, when assessing a certain GMP, the control material should originate from the isogenic variety of the GMP and be treated exactly like the GMP material. This was usually the case in the above-mentioned studies. For the reviewed studies, it is unknown whether validity criteria were previously defined. At least 1 study used a positive control to confirm the exposure and sensitivity of the test system (Vercesi et al. 2006). However, details on how positive controls could look in GMP testing have not been defined.

Documentation and/or experience. As can be derived from the standardization status of the methods described in the reviewed studies, very little to no previous experience exists for a comparison with new test results. In probably all cases, it is the first effort to assess GMP material.

Endpoints. When considering possible endpoints of ecotoxicological testing of GMPs, ideally one would want to cover all relevant life-cycle parameters of a certain test species. Realistically, laboratory testing will have to concentrate on those endpoints that are most likely to be sensitive to an expected impact of the GMP. These are often sublethal endpoints such as feeding behavior, reproduction, or growth. Acute lethal effects should of course be investigated but should not be the focus. Functional parameters such as organic matter decomposition may partially be covered in laboratory experiments, e.g., through feeding trials, but will most likely be better investigated in the field, e.g., by performing litter-bag studies (Römbke et al. 2003) such as those carried out by Cortet et al. (2006). In the laboratory studies reviewed, the above-mentioned endpoints have been covered, mostly concentrating on sublethal parameters but not involving entire life-cycle studies.

RECOMMENDATIONS FOR THE APPLICATION OF EXISTING AND THE DEVELOPMENT OF NEW TEST METHODS FOR GMPs

For risk assessment of GMPs, it is recommended to select test species from organism groups that are ecologically relevant for the receiving environments and cover different exposure routes as well as different taxonomic and physiological groups. The test species to be selected has to additionally fulfill the above-mentioned practicability criteria to be testable in the laboratory. The main challenge will be to expand the spectrum of species from these groups beyond the standard ecotoxicological test species in order to account for differences in the receiving environments and exposure pathways (mainly via feeding) more adequately.

A careful analysis of all possible exposure routes of a nontarget organism to GMPs is essential for selecting the proper test species (Hilbeck et al. 2008). Exposure is highly dependent on the characteristics of the novel trait of the GMP (e.g., toxin expression, herbicide resistance) and its expression patterns. Current ecotoxicity testing of GMPs has been focused on *Bt* plants, for which potential exposure routes are relatively easy to predict. In the future, cultivation of GMPs may well lead to additional exposure routes not considered so far. Nontarget organisms may be exposed to GMP material through the following routes.

- Direct feeding on living (e.g., roots, tubers) or dead (e.g., plant litter, roots) GMP material (on the soil surface or after incorporation by, for example, ploughing)

- Exposure to novel proteins through soil particles or pore water after degradation of GMP material
- Exposure to root exudates through soil particles or pore water
- Secondary exposure to the novel proteins through feeding on other organisms that have incorporated GMP material (including decomposing GMP residues in soils)

In comparing these exposure scenarios with those scenarios already covered in the reviewed standardized guidelines (contaminated or spiked [field] soil, spiked dung [veterinary pharmaceuticals], spray application on plants, spray application on glass plates, spiked food or oral dose, direct contact), it can be concluded that the following exposure scenarios may be applicable to the assessment of GMPs, requiring little or no modification of test methods.

- Field soil, e.g., from GMP cultivation (or spiked in the laboratory)
- Dung from animals fed with GMP material
- Direct contact with GMPs
- Feeding with GMP material (in part)

However, some exposure routes either have not been addressed at all so far or at least have to be extended: direct feeding of living GMP material, exposure through root exudates, and feeding of organisms that have incorporated GMP material or their toxins (''secondary poisoning''). On the other hand, some exposure routes currently used in pesticide tests are not helpful in GMP risk assessment, e.g. spray application on glass plates or other artificial substrates (e.g., in NTA tests). In this context, it has to be stated that acute tests (including those focusing on mortality as the endpoint) cannot be recommended for the ERA of GMPs, because strong short-term effects are unlikely to occur in the environment based on current experience. The same is true when looking at summary endpoints such as microbial respiration in soil. In contrast, the use of methods addressing microbial diversity should be encouraged (so far, no such test has been validated or standardized internationally, but ISO Working Group TC 190 SC4 has started to work on this issue).

In Tables 3 and 4, the experiences discussed above are summarized in order to clarify which current test methods have to be rejected, which can be modified, and which are

Table 3. Standardized terrestrial test originally developed for pesticides or other chemicals not suitable for testing of whole GMPs as part of their ERA

Organism group	Guideline	Reason for rejection
Microorganisms		
C-transformation	OECD 217	Not sensitive according to pesticide testing experience
Biomass, SIR	ISO 14240-1	Not sensitive according to pesticide testing experience
Biomass, fumigation	ISO 1120-2	Not sensitive according to pesticide testing experience
Microbial respiration	ISO 16072	Not sensitive according to pesticide testing experience
Respiration curves	ISO 17155	Not sensitive according to pesticide testing experience
Plants		
Root elongation	ISO 11269-1	Not sensitive according to pesticide testing experience
Vegetative vigor	OECD 227	Exposure scenario (spraying on plant leaves) not GMP-relevant
Screening acute	ISO 17126	Not sensitive according to pesticide testing experience
Soil invertebrates		
Earthworm acute	OECD 207	Not sensitive according to pesticide testing experience
Earthworm acute	ISO 11268-1	Not sensitive according to pesticide testing experience
Beetle larvae	ISO 20963	Not sensitive according to pesticide testing experience
Bees and NTAs		
Bee, acute oral	OECD 213	Exposure scenario not GMP-relevant
Bee, acute contact	OECD 214	Exposure scenario not GMP-relevant
Wasp, glass plate	IOBC	Exposure scenario not GMP-relevant
Mite, glass plate	IOBC	Exposure scenario not GMP-relevant
Ladybird	IOBC	Exposure scenario not GMP-relevant
Egg parasitoid	IOBC	Exposure scenario not GMP-relevant

Note that the evaluation focused on the test method, not the test species. IOBC = guidelines given in Candolfi et al. (2000).

Table 4. Standardized terrestrial test methods originally developed for pesticides or other chemicals suitable for testing of whole GMPs as part of their ERA

Organism group	Guideline	Recommended modification
Microorganisms		
N-transformation	OECD 216[a]	Field soil with GMP litter or from GMP cultivation
N-transformation	ISO 14238[a]	Field soil with GMP litter or from GMP cultivation
Rapid pot. N-transf	ISO 15685	Field soil with GMP litter or from GMP cultivation
Enzyme kits	ISO 22939[a]	Field soil with GMP litter or from GMP cultivation
Dehydrogenase, TTC	ISO 23735-1[a]	Field soil with GMP litter or from GMP cultivation
Dehydrogenase, INT	ISO 23735-1[a]	Field soil with GMP litter or from GMP cultivation
Arthrobacter	ISO 10871[a]	Field soil with GMP litter or from GMP cultivation
Plants		
Seedling emergence	OECD 208	Field soil with GMP litter or from GMP cultivation
Seedling emergence	ISO 11269-2	Field soil with GMP litter or from GMP cultivation
Chronic plant	ISO 22030	Field soil with GMP litter or from GMP cultivation
Soil invertebrates		
Enchytraeid reproduction	OECD 220	Field soil with GMP litter or from GMP cultivation
Enchytraeid reproduction	ISO 16387	Field soil with GMP litter or from GMP cultivation
Earthworm reproduction	OECD 222	Field soil with GMP litter or from GMP cultivation
Earthworm reproduction	ISO 11268-2	Field soil with GMP litter or from GMP cultivation
Earthworm avoidance	ISO 17512-1[a]	Field soil with GMP litter or from GMP cultivation
Earthworm (field)	ISO 11268-3	Evaluation of sublethal endpoints
Nematoda reproduction	ISO 10872[a]	Field soil with GMP litter or from GMP cultivation
Collembola reproduction	ISO 11267	Field soil with GMP litter or from GMP cultivation
Collembola avoidance	ISO 17512-2[a]	Field soil with GMP litter or from GMP cultivation
Predatory mite	OECD Draft	Field soil with GMP litter or from GMP cultivation
Snail growth	ISO 15952	Field soil with GMP litter or from GMP cultivation
Litter-bag (field)	OECD 56	Substrate: GMP material (e.g., straw)
Dung flies	OECD Draft	Dung from GMP-fed animals
Dung beetles	OECD Draft	Dung from GMP-fed animals; evaluation of sublethal endpoints
Bees and NTAs		
Bee (cage)	EPPO 1/170	Exposure to GMP pollen
Bee (field)	EPPO 1/170	None necessary
Lacewing	IOBC	Secondary exposure via prey
Predatory bug	IOBC	Secondary exposure via prey
Staphylinid reproduction	IOBC	Secondary exposure via prey
Carabid reproduction	IOBC	Secondary exposure via prey
Carabid semifield	IOBC	Secondary exposure via prey
Spider	IOBC	Secondary exposure via prey
Mites (field)	IOBC	Secondary exposure via prey

The evaluation focused on the test method, not the test species. IOBC = guidelines given in Candolfi et al. (2000).
[a]Little experience from pesticide testing concerning range of sensitivity.

acceptable in their present form for use with GMPs. It should be noted that Tables 3 and 4 include only those tests that have been internationally standardized by OECD, ISO, EPPO, and IOBC (including draft methods). For these, the most experience from pesticide and other chemical ERAs exists, they fulfill the highest standardization criteria, and they are the best accepted in a regulatory context. For the question of which tests should be developed anew, it can be concluded that, particularly for NTAs, new methods are needed, but existing experience (e.g., cultivation methods) can be used for their development. Any new test methods will have to take the variable properties of GMPs and their receiving environments into account, e.g., by providing flexibility in the choice of the exposure scenario and the actual test species.

CONCLUSIONS AND OUTLOOK

The experiences with the environmental risk assessment of pesticides in the EU can be summarized as follows. A clear and comprehensible legal basis is available, but technical details are explained in Guidance Documents, which are regularly updated. The number of tests required is manageable; depending on the effect profile, 7 to 35 terrestrial tests have to be performed with an active ingredient of a pesticide and its formulations in order to obtain an approval. In cases of concern (i.e., if the outcome of the risk assessment indicates that safe use is not possible for the planned amount and frequency of applications in a certain crop), the number, complexity, effort, and ecological relevance of tests increases on higher tiers of the testing strategy. The selection and the design of these tests are often discussed between industry and agencies, using information provided by ad hoc working groups consisting of representatives of universities, industry, and agencies. With these experiences as a starting point, the following recommendations concerning the ERA of GMPs can be given.

- Legal testing requirements must be clear. To reach this goal, a testing program similar to that for pesticides should be demanded by the European regulatory authorities for GMPs as well. However, this does not mean just copying the requirements; modifications are necessary in order to address the specific properties of GMPs (such as self-reproduction) and their ERA (such as the importance of the receiving environment).
- When selecting tests or test strategies, open discussions between all interested parties are very helpful. Therefore, technical guidance papers should be written under the auspices of neutral organizations such as the Society for Environmental Toxicology and Chemistry (SETAC) or the OECD Test Guidelines Program.
- For the registration of pesticides, ecologically relevant, higher tier tests are accepted. Consequently, for the risk assessment of GMPs, such tests should be implemented too. In other words, the sole performance of the simple tier 1 tests is not sufficient for GMPs (Figure 2), which is an important difference from the ERA of pesticides (Figure 1). In addition, the number of tests required for a specific GMP should reflect the number of geographical regions where it is grown.
- For the registration of an active ingredient of a pesticide in the EU, at least 1 representative formulation of this pesticide has to be tested as well (on the national level, the potential risk of those formulations to be used in that country have to be assessed as well), i.e., the final test items are pesticide formulations. These additional tests are performed in order to assess potential synergistic effects caused by the interaction of several active ingredients belonging to the same formulation or between 1 (or more) active ingredients and those formulation additives added in order to improve application efficiency. Consequently, and by analogy with these complex pesticide formulations, the whole GMP should be the ultimate test item (Hilbeck et al., submitted).

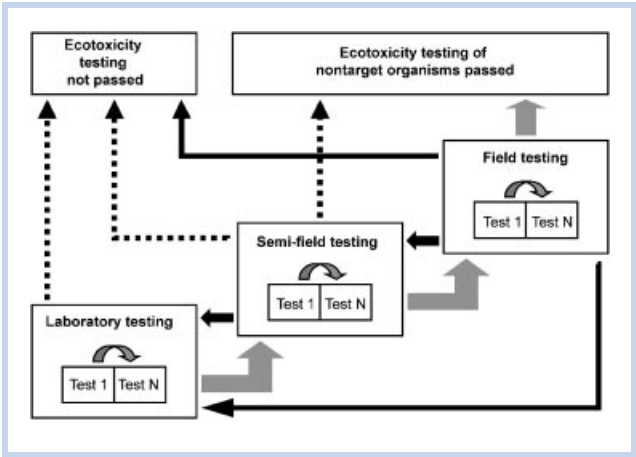

Figure 2. Proposed testing strategy for the risk assessment of GMPs in the European Union. The arrows show the possible lines of consequence depending on testing results. Dotted arrows represent lines of consequences, the possibilities for which will still have to be specified by regulatory decision makers. Testing follows a tiered scheme (laboratory, semifield, field), but shortcuts and feedback are possible (black arrows). The decision that ecotoxicity testing is passed can occur only after safe use has been demonstrated at all tiers (gray arrows).

Based on the evaluation of 105 test methods with soil organisms developed for pesticide testing and published by standardization organizations or in the scientific literature, the following recommendations concerning the selection of appropriate test methods for the ERA of GMPs are possible. In addition, the experiences gained in the relatively few studies with soil organisms and GMPs as well as proposals from the literature are included here.

- Testable (i.e., cultivable and practical) species that are ecologically relevant for the receiving environments should be selected. They should cover different exposure routes such as soil porewater and in particular the uptake via food, e.g., litter from GMPs. Also, different taxonomic and physiological groups should be included.
- Existing standard test methods for, e.g., earthworms, collembolans, or predatory mites, should be modified in order to save resources.
- New tests, e.g., with isopods, have to be developed to cover exposure routes that are particularly important for GMPs.
- Chronic endpoints (e.g., feeding behavior, biomass, or reproduction) are preferred over acute or lethal ones.

- In addition to structural also functional endpoints such as organic matter decomposition have to be tested because of their high ecological relevance, but more research is needed to adapt existing methods to the GMP requirements (e.g., different environments).
- Soil microorganisms should be included in the test battery, but probably structural endpoints (e.g., genetical) will be more appropriate than activity parameters.

The final aim of the test strategy is the selection of a case-specific battery of standardized tests relevant for the ERA of GMPs. Preferably, the same test method can be used for different species of the same organism group, e.g., representing different regions or environments within Europe. The main challenge will be to increase the number of species from these groups beyond the existing standard ecotoxicological test species in order to account more adequately for differences in the receiving environments, behavioral types, and exposure pathways (mainly via feeding).

REFERENCES

Ahl Goy P, Warren G, White J, Privalle L, Fearing P, Vlachos D. 1995. Interaction of an insect tolerant maize with organisms in the ecosystem. In: Landsmann J, Casper R, editors. Key biosafety aspects of genetically modified organisms. Workshop, 10-11 April 1995 in Braunschweig. Germany. Berlin (DE), Blackwell Wissenschafts-Verlag.p 50–53.

Bradbury SP, Feijtel TCJ, van Leeuwen CJ. 2004. Meeting the scientific needs of ecological risk assessment in a regulatory context. *Environ Sci Technol* 1:463–470.

Cairns J Jr. 1981. Sequential versus simultaneous testing for evaluating the hazards of chemicals to aquatic life. *Mar Envionr Res* 4:165–166.

Candolfi M, Blümel S, Forster R, Bakker FM, Grimm C, Hassan SA, Heimbach U, Mead-Briggs M, Reber B, Schmuck R, Vogt H. editors. 2000. Guidelines to evaluate side-effects of plant protection products to non-target arthropods. Gent (BE): International Organization for Biological and Integrated Control of Noxious Animals and Plants, West Paleartic Regional Section (IOBC/WPRS). 158 p.

Chapman PM. 2002. Integrating toxicology and ecology: Putting the ''eco'' into ecotoxicology. *Marine Poll Bull* 44:7–15.

Cortet J, Andersen MN, Caul S, Griffiths B, Joffre R, Lacroix B, Sausse C, Thompson J, Krogh PH. 2006. Decomposition processes under *Bt* (*Bacillus thuringiensis*) maize: Results of a multi-site experiment. *Soil Biol Biochem* 38:195–199.

[EC] European Commission. 2001. EU Directive 2001/18/EC of 12 March 2001 on the deliberate release into the environment of genetically modified organisms and repealing Council Directive 90/220/EC. Official Journal of the European Communities L 106/1-38 (17.4.2001).

[EC] European Commission. 2002. Guidance document on terrestrial ecotoxicology. Draft Working Document Sanco/10329/2002 rev. 2 final, 17 October 2002, Luxembourg (LU).

[EC] European Commission. 2003. Regulation No. 1829/2003 of the European Parliament and of the Council of 22 September 2003 on genetically modified food and feed. Luxembourg (LU).

[EPPO] European Plant Protection Organization. 2003. Environmental risk assessment scheme for plant protection products. *Bull OEPP/EPPO Bull* 33:99–244.

Escher N, Kach B, Nentwig W. 2000. Decomposition of transgenic *Bacillus thuringiensis* maize by microorganisms and woodlice *Porcellio scaber* (Crustacea: Isopoda). *Basic Appl Ecol* 1:161–169.

[EU] European Union. 1991. Council Directive Concerning the Placing of Plant Protection Products on the Market. No. 91/414/EEC. Brussels (BE).

[EU] European Union. 1997. Council Directive Establishing Annex VI to Directive 91/414/EEC Concerning the Placing of Plant Protection Products on the Market. No. 97/57/EC, Brussels (BE).

Forbes VE, Forbes TL. 1994. Ecotoxicology in theory and practice. London UK. Chapman & Hall, 264 p.

Garcia M. 2004. Effects of pesticides on soil fauna: Development of ecotoxicological test methods for tropical regions. Göttingen (DE). Cuvillier Verlag, 281 p.

Groot AT, Dicke M. 2002. Insect-resistant transgenic plants in a multi-trophic context. *Plant J* 31:387–406.

Heckmann L-H, Griffiths B, Caul S, Thompson J, Pusztai-Carey M, Moar WJ, Andersen MN, Krogh PH. 2006. Consequences for *Protaphorura armata* (Collembola: Onychiuridae) following exposure to genetically modified *Bacillus thuringiensis* (*Bt*) maize and non-*Bt* maize. *Environ Pollut* 142:212–216.

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. Federal Agency for Nature Conservation, Bonn-Bad Godesberg (DE): *BfN-Skripten* 236:287 pp.

Hilbeck A, Römbke J, Meier M, Jänsch S, Teichmann T, Tappeser B. (submitted). Environmental risk assessment of genetically modified plants—concepts and controversies. *Perspect Plant Ecol Evol Systemat*.

[ISO] International Organization for Standardization. 1998. Soil quality—Effects of pollutants on earthworms (*Eisenia fetida*). Part 2: Determination of effects on reproduction. ISO 11268-2. Geneva (CH).

[ISO] International Organization for Standardization. 1999a. Soil quality—Effects of pollutants on earthworms—Part 3: Guidance on the determination of effects in field situations. ISO No. 11268-3. Geneva (CH).

[ISO] International Organization for Standardization. 1999b. Soil quality—Inhibition of reproduction of collembola (*Folsomia candida*) by soil pollutants. ISO No. 11267. Geneva (CH).

[ISO] International Organization for Standardization. 2002a. Soil quality—Determination of potential nitrification—Rapid test by ammonium oxidation. ISO 15685. Geneva (CH).

[ISO] International Organization for Standardization. 2002b. Soil quality—Determination of the activity of the soil microflora using respiration curves. ISO 17155. Geneva (CH).

[ISO] International Organization for Standardization. 2007. Soil quality—Avoidance test for evaluating the quality of soils and the toxicity of chemicals. Test with earthworms (*Eisenia fetida/andrei*). ISO 17512-1. Geneva (CH).

Jänsch S, Amorim MJ, Römbke J. 2005. Identification of the ecological requirements of important terrestrial ecotoxicological test species. *Environ Rev* 13:51–83.

Kareiva P, Stark J, Wennergren U. 1996. Using demographic theory, community ecology and spatial models to illuminate ecotoxicology. In Baird DJ, Maltby L, Greig-Smith PW, Douben PET, editors. ECOtoxicology: Ecological Dimensions. London (UK). Chapman & Hall, p 13–23.

Keddy CJ, Greene JC, Bonnell MA. 1995. Review of whole-organism bioassays: Soil, freshwater sediment, and freshwater assessment in Canada. *Ecotoxicol Environ Saf* 30:221–251.

Lavelle P, Bignell D, Lepage M, Wolters V, Roger P, Ineson P, Heal OW, Dhillion S. 1997. Soil function in a changing world: the role of invertebrate ecosystem engineers. *Eur J Soil Biol* 33:159–193.

Leeuwen CJ, Hermens JLM. 2001. Risk assessment of chemicals: An introduction. Dordrecht (NL). Kluwer Academic, 374 p.

Løkke H, Janssen CR, Lanno RP, Römbke J, Rundgren S, Van Straalen NM. 2002. Soil toxicity tests—invertebrates. In: Fairbrother A, Glazebook PW, van Straalen N, Tarazona JV, editors. Test methods to determine hazards of sparingly soluble metal compounds in soils. Pensacola (FL). SETAC, p 34–57.

[OECD] Organization for Economic Co-Operation and Development. 1984a. OECD Guidelines for testing of chemicals. Paris (FR).

[OECD] Organization for Economic Co-Operation and Development. 1984b. OECD-guideline for testing of chemicals No. 207. Earthworm acute toxicity test. Paris (FR).

[OECD] Organization for Economic Co-Operation and Development. 2006. OECD Series on Testing and Assessment No. 56. Guidance document on the breakdown of organic matter in litterbags. ENV/JM/MONO 23. Paris (FR).

[OECD] Organization for Economic Co-Operation and Development. 2008. Guideline for the testing of chemicals. Draft: Determination of developmental toxicity of a

test chemical to Dipteran dung flies (*Scathophaga stercoraria* L. [Scathophagidae] and *Musca autumnalis* De Geer [Muscidae]). Paris (FR).

Petersen H, Luxton M. 1982. A comparative analysis of soil fauna populations and their role in decomposition processes. *Oikos* 39:287–388.

Römbke J, Amorim MJ. 2004. Tackling the heterogeneity of soils in ecotoxicological testing: A EURO-soil based approach. *JSS J Soils Sediments* 4:276–281.

Römbke J, Bauer C, Marschner A. 1996. Hazard assessment of chemicals in soil. Proposed ecotoxicological test strategy. *ESPR Environ Sci Poll Res* 3:78–82.

Römbke J, Breure AM. editors. 2005. Ecological soil quality: Classification and assessment. *Ecotoxicol Environ Saf* 62:185–308.

Römbke J, Heimbach F, Hoy S, Kula C, Scott-Fordsmand J, Sousa P, Stephenson G, Weeks J, editors. 2003. Guidance document: Effects of plant protection products on functional endpoints in soil (EPFES). Pensacola (FL). SETAC, 92 p.

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 R, Schiemann J, Sears MK, Shelton AM, Sweet J, Vaituzis Z, Wolt JD. 2008. Assessment of risk of insect-resistant transgenic crops to nontarget arthropods. *Nat Biotechnol* 26:203–208.

Romeis J, Battini M, Bigler F. 2003. Transgenic wheat with enhanced fungal resistance causes no effects on *Folsomia candida* (Collembola: Isotomidae). *Pedobiologia* 47:141–147.

Saxena D, Stotzky G. 2001. *Bacillus thuringiensis* (*Bt*) toxin released from root exudates and biomass of Bt corn has no apparent effect on earthworms, nematodes, protozoa, bacteria and fungi in soil. *Soil Biol Biochem* 33:1225–1230.

[TGD]. 2003. Technical Guidance Document on Risk Assessment in support of Commission Directive 93/67/EEC on Risk Assessment for new notified substances, Commission Regulation (EC) No 1488/94 on Risk Assessment for existing substances, Directive 98/8/EC of the European Parliament and of the Council concerning the placing of biocidal products on the market. Brussels (BE).

Vercesi ML, Krogh PH, Holmstrup M. 2006. Can *Bacillus thuringiensis* (*Bt*) corn residues and *Bt*-corn plants affect life-history traits in the earthworm *Aporrectodea caliginosa*? *Appl Soil Ecol* 32:180–187.

Wandeler H, Bahylova J, Nentwig W. 2002. Consumption of two *Bt* and six non-*Bt* corn varieties by the woodlouse *Porcellio scaber*. *Basic Appl Ecol* 3:357–365.

Yu L, Berry RE, Croft BA. 1997. Effects of *Bacillus thuringiensis* toxins in transgenic cotton and potato on *Folsomia candida* (Collembola: Isotomidae) and *Oppia nitens* (Acari: Orbatidae). *J Econ Entomol* 90:113–118.

Zwahlen C, Hilbeck A, Howald R, Nentwig W. 2003. Effects of transgenic *Bt* corn litter on the earthworm *Lumbricus terrestris*. *Mol Ecol* 12:1077–1086.