Roundtable

Fitness and beyond: Preparing for the arrival of GM crops with ecologically important novel characters

Mike WILKINSON1* and Mark TEPFER2

1 Institute of Biological Environmental and Rural Sciences, Edward Llwyd Building, Wales University, Aberystwyth, Ceredigion SY23 3DA, UK
2 Plant Virology Group, ICGEB Biosafety Outstation, Via Piovega 23, 31056 Ca’ Tron di Roncade, Italy

The seemingly inexorable expansion of global human population size, significant increases in the use of biofuel crops and the growing pressures of multifunctional land-use have intensified the need to improve crop productivity. The widespread cultivation of high-yielding genetically modified (GM) crops could help to address these problems, although in doing so, steps must also be taken to ensure that any gene flow from these crops to wild or weedy recipients does not cause significant ecological harm. It is partly for this reason that new GM cultivars are invariably subjected to strict regulatory evaluation in order to assess the risks that each may pose to the environment. Regulatory bodies vary in their approach to decision-making, although all require access to large quantities of detailed information. Such an exhaustive case-by-case approach has been made tractable by the comparative simplicity of the portfolio of GM crops currently on the market, with four crops and two classes of traits accounting for almost all of the area under cultivation of GM crops. This simplified situation will change shortly, and will seriously complicate and potentially slow the evaluation process. Nowhere will the increased diversity of GM crops cause more difficulty to regulators than in those cases where there is a need to assess whether the transgene(s) will enhance fitness in a non-transgenic relative and thereafter cause ecological harm. Current practice to test this risk hypothesis focuses on attempting to detect increased fitness in the recipient. In this paper we explore the merits and shortcomings of this strategy, and investigate the scope for developing new approaches to streamline decision-making processes for transgenes that could cause unwanted ecological change.

Keywords: environmental risk assessment / fitness / fitness parameters / GM crops / invasiveness / weediness

THE CONTEXT OF ENVIRONMENTAL RISK ASSESSMENT

Throughout the world, wherever transgenic crops are grown outside the confines of the laboratory, there is a regulatory system for governing and monitoring the conditions under which experimental field trials and/or commercial cultivation is sanctioned. There is considerable variation between individual countries in the legislative frameworks that dictate how decisions to permit, delay or halt the release of a particular transgenic event are made (Capalbo et al., 2003; Cardwell and Kerr, 2008; Guehlstorff and Hallstrom, 2005; Jepson et al., 2005; Kalaitzandonakes et al., 2007; Nasiruddin and Nasim, 2007; Ramjoue, 2007; Salleh, 2006; Zafar, 2007). Nevertheless, there appears to be a growing harmonization of the underlying process that governs decision-making, with all systems adopting a case-by-case approach in which large bodies of evidence are first presented and then evaluated in a reiterative fashion.

The documents made in support of individual submissions are usually rather weighty and data-rich; they invariably include substantial information relating to the biology and agronomy of the crop, the biology and nutritional qualities of the specific event itself, details of the transgenes it contains, their site of insertion and expression characteristics, as well as data on the biology of the crop’s wild relatives and of any associated flora and fauna (e.g. Rowland 2002; http://www.aphis.usda.gov/biotechnology/brs_main.shtml; http://www.efsa.europa.eu/EFSA/ScientificPanels/efsa_locale-1178620753812-GMO.htm). The onus of responsibility for generating the large volumes of data in support of an individual event invariably falls on the notifier (usually the company wishing to commercialize the GM crop). This requirement,
coupled with associated risk management costs (e.g. Wilson et al., 2008) may have a significant role in restricting commercial releases of GM crops to those generated by large, multi-national firms, and so may be argued to restrict broader application of the technology. Naturally, it is also far easier to reconfigure a submission made for one regulatory system and use it for a new country than it is to create a submission de novo, especially given the effective levels of communication and co-operation between regulatory bodies as required by the Cartagena Protocol, and dictated by the mutual interests of international trade. To what extent these factors have acted in concert to favor a restricted list of ‘first wave’ commercial transgenic events is open to question. However, what is clear is that to date, the extremely rapid growth in commercially-grown transgenic crops has almost entirely taken place on the back of four crops (soybean, cotton, maize and rapeseed) and four transgene types (insect resistance conferred by Bacillus thuringiensis (Bt) cry1 and cry3, and herbicide tolerance conferred by bar or pat (glufosinate resistance) and modified EPSPS (glyphosate resistance)).

We can confidently expect this comparative simplicity to change in the relatively near future, as the numbers of countries growing GM crops expands further, as the number of traits controlled by transgenes diversifies, and as the number of transgenes contained within each transgenic cultivar increases. It can be reasoned that GM crop range expansion poses fewer difficulties than transgene proliferation. Most early commercial GM crops contain only the well-established trait-crop categories (e.g. James, 2005), and GM crop expansion into new countries was almost invariably based on a cultivar for which there was already a portfolio of submission documents for other countries and an established safety record. This pattern is unlikely to change in the near future. Less certain is the nature of further changes in transgenic cultivars in countries with a large and established production of GM crops. Some transgenic crops are certainly now approaching saturation of the available market, and there is only limited scope for further growth. One option then is to stack transgenes through the cross-pollination of complementary events that have already received approval and to combine these by crossing with novel transgenes as they are developed. Some indication of an increase in the number and complexity of stacked events can be taken from the abundance of field releases of maize events containing ten or more transgene-mediated traits in the US (http://www.isb.vt.edu/cfdocs/fieldtests1.cfm). It is difficult to determine the precise number of transgenes in any event granted field approval in the US because the identity of many genes is withheld because it is deemed to be Confidential Business Information (CBI). It is nevertheless possible to infer a minimal number of transgenes from the number of different source organisms identified for the transgenes contained in an event. Using this approach, between 1995 and 2000, there was just one stacked event containing more than ten registered traits, compared with 11 between 2000 and 2005, and 14 so far from 2005–2008.

Interestingly, one maize event (permit 07-257-108RM) contained 65 different source organisms for transgenes whose identity was withheld under CBI. Added to this complexity is the enormous diversification of transgenic traits being incorporated into GM plants. This is aptly indicated by the high diversity of traits of transgenic lines currently being trialed in the US alone. It is likely that many of these events will not ultimately progress to unrestricted release on a commercial scale, although it seems inevitable that significant numbers will, and that this will require new sets of information to be provided by the notifier.

Transgenes that enhance product quality or those that promote various forms of abiotic stress tolerance (e.g. drought, cold, salinity and generic abiotic stress) predominate among the new traits being developed by the large multinational companies (http://www.isb.vt.edu/cfdocs/fieldtests1.cfm). It is therefore on these traits that regulators and risk assessment research scientists now need to focus greater attention. As the proportion of transgenic cultivars approaches saturation for existing GM crops, so there is increasing commercial impetus to explore the value of introducing transgenes into other crops of economic significance. This process may well accelerate as the diversity of transgenes broadens.

It should also be pointed out that the traits introduced into GM crops so far are mediated by proteins that are not expected to have significant direct or indirect effects on the physiology and metabolism of the plant. For instance, Bt proteins are not expected to have any physiological effects on the host plant, and the effects of herbicide-tolerance genes should also be negligible, since they encode proteins that either inactivate the herbicide (pat and bar genes), or encode enzymes that are tolerant to it (e.g. modified EPSPS). In contrast, the various strategies for creating resistance/tolerance to abiotic stresses are all based on significant changes in the plant’s metabolism, and thus may require even more complex data sets for risk assessment.

Particularly for certain new crop/trait combinations, the amount of information required for risk assessment is far greater than for a crop with a successful track history of authorized GM cultivars. Some of these data will relate to the biology of the crop itself, and so should be relatively easy to assemble. Other data, such as those relating to possible ecological harm (e.g. effects mediated by cross-compatibility of wild relatives or exposure of non-target species), may be more problematic.
Thus, the anticipated increase in the complexity and diversity of GM crops over the coming years means that notifiers will be required to compile a growing list of information for each submission. The same trend will drive regulators to develop simplified systems to identify the key issues required for decision-making. Pragmatism dictates that this may well involve a reduction in the overall information requested and a refocusing on data that aids decision-making, or more rigorous application of the tiered approach (see Box 1).

**Box 1. The tiered approach to environmental risk assessment.**

Within the GM crop risk assessment community, there is now a consensus that a tiered approach is an important concept for efficient and effective risk assessment (Wilkinson et al., 2003b). Briefly, Tier 1 entails first testing the risk hypothesis under highly controlled (lab or greenhouse) worst-case conditions. If it can be shown in a sufficiently robust manner in Tier 1 studies that either harm or exposure is negligible, then it is possible to conclude that risk is negligible, and higher tier studies are not necessary. If Tier 1 results leave reasons for concern about risk, then it is appropriate to carry out Tier 2 studies, in which the issue is tested under more realistic conditions, e.g. in small-scale confined field trials. Here also, if either harm or exposure can be shown to be negligible, then higher tier (larger-scale) studies would not be required. The non-trivial issue in this strategy is designing and carrying out lower tier studies robust enough to be predictive of effects that would be observed at higher ones.

The quest for pertinent information is most challenging when the aim is to avoid ecological harm, particularly that pertaining to natural and semi-natural communities found outside the farmed environment. Quite apart from the multitude of possible scenarios that could cause change, difficulty lies in defining what actually constitutes an unacceptable change (i.e. harm) and what level of change is deemed acceptable. This boundary is known as an assessment endpoint, and can be used to provide ‘trigger values’, such as the decline of a named species to a defined population size. It is here that the crux of the decision-making procedure lies: lack of confidence that the unacceptable/acceptable boundary will not be crossed favors delay (whilst additional evidence is assembled) or rejection of the submission. However, it is not necessary to have an accurate prediction of (in this case) the population size of the species under consideration in order to make a decision, merely confidence that it will not traverse the acceptable/unacceptable boundary.

This then leaves us with two problems. First, we must try to identify all plausible changes that could constitute harm (an unacceptable change). Second, we must set trigger values (boundaries), and assess whether the GM crop under consideration is likely to cause sufficient change to cross that boundary. In both agricultural and ecological settings, these changes are most readily recognized and described when they relate to an organism whose population size we are concerned may expand (e.g. a weed, pest or invasive), or contract, in the case of an organism that we wish to protect (e.g. a beneficial non-target organism, a species of cultural importance, or an endangered species). It is probably for this reason that much of the research in this area has focused on predicting changes to the fitness of the crop, of a wild relative, or of an associated species, that could lead either to population expansion or contraction. In this sense, any change in fitness per se has been used as a proxy for the consequences that may arise from the changed fitness. However, as we outline below, it does not necessarily follow that changed fitness will lead to a change in population size, that a change in population size will necessarily lead to harm, that harm cannot occur without a change in population size, or that the context in which we measure fitness necessarily relates to the acceptable/unacceptable boundary being considered. The primary objectives of this article are therefore to explore how we can define and anticipate the consequences of changed fitness to an organism of concern within the context of an ecological risk assessment.

**MEASURING CHANGED FITNESS AND ITS CONSEQUENCES**

**Measuring fitness and/or other pertinent parameters**

To date there has been little consistency in the meaning of the phrase ‘enhanced fitness’ when applied to risk studies associated with the release of transgenic crops. This is partly because there is some confusion between the concepts of ‘enhanced fitness’, which we define here as a characteristic of an individual or sub-population of individuals that consistently contribute more offspring to the subsequent generation, and ‘fitness parameters’, which we define here as one or more features of an individual or subpopulation of individuals that could contribute towards enhanced fitness. This distinction is more than mere semantics, since it is entirely possible for an individual to exhibit features such as increased seed set (a commonly measured fitness parameter), and yet not display enhanced fitness, because of confounding factors such as density-dependent seedling predation. Alternatively, it is also entirely possible for an individual with reduced seed set to show enhanced fitness, perhaps through an enhanced ability to compete, or through...
the predominant importance of vegetative reproduction. It is for this reason that a growing number of authors have elected to use demographic modeling approaches to estimate changes in the innate growth rate of a population (defined as lambda ($\lambda$)). This is an important step, since an increase in population size within cultivated fields (i.e. weediness) or in non-farm environments (i.e. invasiveness) has often been considered a potential harm that could be caused by GM crops.

Modeling strategies provide extremely powerful tools for predictive ecology, primarily because they identify the key stages in life history progression that determine population growth. Their utility for risk assessment rests on the fact that they provide empirical evidence for whether transgenic plants are outperforming non-transgenic equivalents in field tests. The validity of such predictions rests heavily on how closely the growing conditions in which experiments are performed mimics those found in the field. This is seemingly easiest to achieve when the risk assessment scenario being considered relates to predicting the risk of enhanced weediness in an agricultural setting of a volunteer transgenic crop or of an introgressed wild relative. Under these circumstances, it is relatively easy to simulate the soil conditions and surrounding flora (predominantly the crop), and so to measure the performance of the transgenic sub-population under conditions broadly similar to those found in the field. But even under this simplified system, there can be difficulty in confidently predicting performance given the diversity of growing conditions and agronomic practices used across the range of the crop in the nation under consideration. To some extent these problems can be mitigated through replicated trials, although it may be problematic to accommodate for scale. The problems are far more pronounced when the risk scenario relates to a wild relative growing outside the agricultural environment.

Ultimately, however, it should be remembered that these studies are not meant to provide accurate predictions of population size increase; they are merely intended to impart sufficient information to assist the decision-making process. In this context, an increase in lambda is generally taken as the proxy for assessment endpoint trigger values. Under certain scenarios, this approach is seemingly valid although in others, an increase in lambda may be an inappropriate tool to assist the decision-making process.

Increased population growth can sometimes provide a useful trigger in its own right in a specific as well as a generic sense. For example, if the concern centers on the possibility that a transgenic crop or an introgressed wild relative will become more pernicious as a weed, then any consistent increase in lambda over multiple sites and years would be sufficient to elicit caution. Conversely, consistent equality or decline in lambda amongst the transgenic subpopulation would indicate a lack of concern relating to this hazard. On the other hand, it is sometimes inappropriate or insufficient to focus on fitness of the transgenic subpopulation as the aid for decision-making. For example, if the species of concern (protection goal) is a herbivore (endpoint species), and is found by Tier 1 experiments to be entirely sensitive to the transgene product, then activity must center on estimating whether exposure of that species to the transgene is sufficient to cause significant decline in population size below acceptable levels. The Monarch butterfly story is exemplary in this regard (Box 2).

**Box 2. The Monarch butterfly controversy seen in retrospect.**

The publication by Losey et al. (1999) showing that Monarch larvae are sensitive to Bt maize pollen, which can be regarded as a typical Tier 1 study, was at the origin of a major controversy that could probably have been avoided. That Monarch larvae could be harmed by Bt maize pollen was not surprising, since Cry1Ab is toxic to many Lepidoptera, but in the absence in 1999 of exposure data from the field, it was impossible at the time to determine the actual risk to the populations of Monarchs. Once the exposure studies had been carried out in exemplary fashion (Sears et al., 2001), the controversy gradually disappeared. The important point is that either the risk scenario had not been identified in the initial stages of the risk assessment, or if it had been identified, it was presumed that risk was negligible, on the assumption that conditions in the over-wintering sites in Central America are more likely to determine the size of the Monarch population. Unfortunately, this did not consider the possibility that the risk of any decline in this iconic species would be universally regarded as unacceptable. That the scenario was not adequately considered in the original assessment is the primary cause for concern, and this issue is made more acute given that this was an assessment on an established crop with an established transgene. It is important that we learn from this experience and seek generic approaches that minimize the chances of similar omissions.

**The need for improved problem formulation**

In order to avoid further Monarch-type controversies, it is the problem formulation stage at the very start of the risk assessment process where we should focus our attention. Problem formulation considers the characteristics of the transgene, the crop, its wild relatives and any associated species and seeks to construct realistic scenarios describing how the presence of a transgene in a certain setting
may result in some form of harm. This process inevitably relies heavily on the experience, knowledge and creativity of the various stakeholders involved, and thus far has been made comparatively simple by the predominance of just two types of traits (herbicide tolerance and insect resistance) and four crops (usually considered outside their centers of origins). As the number and diversity of GM crops increases, along with the number of traits conferred by the transgenes they contain, so does the potential complexity of plausible scenarios. The challenge then rests on identifying generic approaches that minimize the need for creative inspiration. One attractive option is to use ‘enhanced fitness’ as a proxy to cover a broad raft of scenarios leading to harm. The thinking is that if there is no fitness advantage/disadvantage, then surely there is negligible scope for harm. There are two important considerations that must be made when using such a strategy:

• First, it is vital that one identifies in which organism fitness advantage or disadvantage will determine whether or not harm will occur.
• Second, it is important that the tests for significant changes to fitness are both appropriate and resource-effective.

In order to address the first point, we should remember that there are essentially four categories of organism in which a change in fitness could lead to environmental harm, viz:

• the crop itself,
• (weed) flora and on-farm fauna associated with the crop,
• cross-compatible wild relatives of the crop,
• flora and fauna associated with the crop or crop relative in non-farm environments.

Effects of enhanced fitness

Enhanced fitness in the crop will lead to a series of scenarios relating to the capacity of the GM crop to become a more pernicious problem as a volunteer weed in other crops or as a potential invasive species of communities outside the farmed environment. The former is by far the most direct and simple to test with existing approaches. In part, this is because we usually already know much about the propensity of a particular crop to become a substantial volunteer problem. This provides a useful comparator for the relatively simple question of whether the transgenic lines are likely to exacerbate this to such an extent that volunteer populations become unacceptably high. It is relatively easy to anticipate the sorts of traits that may influence fitness in an on-farm environment and also to mimic the farmed environment sufficiently closely in order to test these hypotheses.

It is rather more difficult to anticipate whether a transgene will confer enhanced fitness outside the farmed environment. Here again, however, there is usually an established literature describing the communities in which feral crops can be found, and in many cases, how persistent non-transgenic populations are in these settings. This information provides a useful baseline upon which plausible scenarios can be built. However, in cases where the crop or its wild relative is known to appear in reasonable numbers outside the agricultural field, it may not be evident whether or not these populations are truly self-sustaining. This point can be exemplified by a number of case studies. For instance, feral populations of cultivated rapeseed frequently occupy disturbed communities in the margins of agricultural fields, on roadside verges or on building sites (Timmons et al., 1995). These populations can be large, but can also vary in size significantly from one year to the next (Crawley and Brown, 1995). Direct observations of these populations between years represents an easy and direct way to assess their stability, and in the case of rapeseed, has generally tended to imply a trend towards extinction (Crawley and Brown, 1995; Pesson et al., 2001; Wilkinson et al., 1995). Such studies all suggest that plant abundance is attributable to a balance between the frequency of fresh spillages and the speed of decay to local extinction. However, the reality is almost certainly more complex, and also includes the compounding factors such as the frequency of disturbance needed for recruitment, and the extent to which secondary dormancy extends the longevity of the seed bank. For instance, Pesson et al. (2001) used genetic analyses to reveal that some members of roadside rapeseed populations apparently originated from cultivars that had not been marketed for at least eight years, and inferred that more emphasis should be placed on understanding the mechanisms underpinning dynamics of these populations.

Deployment of life-history modeling approaches has proved a popular and informative strategy to address this need, and has generally confirmed a tendency towards extinction and emphasized the importance of seedbank dynamics and fresh spillage (Aono et al., 2006; Garnier et al., 2006; Pivard et al., 2008a, 2008b; Saji et al., 2005; von der Lippe and Kowarik, 2007a, 2007b; Yoshimura et al., 2006). These data are useful because they provide a valuable baseline comparator for any transgenic crop. Care must nevertheless be exercised when interpreting the results of such works. There seems little doubt that regulators would enforce restrictions in any situation where the transgene increases fitness and there is associated population expansion ($\lambda > 1$). Decisions are less predictable if populations are in decline. For example, if transgenic feral rapeseed populations were predicted to decline to extinction on the basis of field trials,
but to do so more slowly than conventional rapeseed (i.e. $\lambda$ conventional $< \lambda$ transgenic $< 1$), or else to be more prone to spillage ($\lambda$ conventional = $\lambda$ transgenic $\leq 1$, but transgenic recruitment increased because of changed handling), paradoxically the effect would still be a net increase in overall feral population size. In this case the increased incidence of population establishment through seed spillage and/or the reduced rate of decline of these populations to extinction would inevitably lead to an increase in the number of feral rapeseed plants when considered across the landscape as a whole even though each individual population is still destined for extinction. Whether either scenario constitutes a problem depends on the adopted risk-assessment philosophy. For instance, if the regulators define the unacceptable boundary as strictly expanding feral populations of transgenic plants (i.e. $\lambda > 1$) then these scenarios would not be a problem. Conversely, if a comparative approach is adopted (no worse than the existing situation), then an increase in overall abundance of feral plants could be considered to constitute a problem whether it is caused by increased spillage or by reduced extinction rates. This difference may also influence the appropriate choice of experiment to aid decision-making. For the former, one needs to determine whether the feral populations are on a trajectory towards local extinction, in which case an in situ life history study may be favored. For the latter, a simple comparison of transgenic and non-transgenic plants for key life history traits may suffice. This case illustrates the key difficulty facing regulators and other decision-makers: the need for clarity in defining the acceptable/unacceptable boundary. Some might argue that an increase in feral rapeseed would constitute a problem whatever the cause, although the natural extension of this logic is that steps should also be taken to regulate other plausible causes of population expansion such as transportation and harvest (affecting spillage rates) and construction/land management (affecting soil disturbance and so plant recruitment) and even the agricultural popularity of the crop. In reality, the nub of the argument actually focuses around the consequences of an expanded population size rather than population growth per se, with the lack of enhanced fitness being taken as a proxy for the absence of negative environmental consequence(s) (i.e. harm) to aid decision-making. The reasoning being that ‘no enhanced fitness’ also means ‘no consequence of enhanced fitness and so no harm’. Taken in this context, the need for a comparator is paramount and the basis of the evaluation should be ‘the transgenic event shows no more fitness than the non-transgenic equivalent’.

When taking this stance, it is important to give careful consideration to scenarios where there is potential for an uncoupling of the relationship; situations where there is no apparent enhanced fitness (at least in its current setting) but harm is likely, or where enhanced fitness is likely but it is unlikely to result in harm. Clearly, the former is generally of greater concern, and has greatest resonance where the basis of concern is that the transgenic plants will invade into new communities and thereafter cause harm. These situations are well known from alien species introductions (e.g. Bailey and Wissskirchen, 2006; Bailey et al., 2007), but are particularly difficult to predict, because the factors that currently constrain a wild relative or crop to a particular community (or communities) is invariably unknown. It is at this point that it is useful to consider the traits conferred by the transgene(s). For many traits, and this is particularly the case for the traits in the presently released GM crops, it is difficult to envisage scenarios where the addition of a single additional trait will lead to the successful invasion by the crop species of a community in which the crop was previously unknown. This was first examined in the series of seminal works by Crawley and colleagues, in which the capacity of transgenic herbicide-tolerant rapeseed plants to invade new plant communities was assessed directly in a series of medium/long-term transplant experiments in the presence and absence of the herbicide (Crawley et al., 1993). Careful analysis of the demography of each population clearly indicated that most populations rapidly declined towards extinction, and that the transgenic lines were no more likely to survive than their conventional counterparts. Thus, the populations failed to elicit concern on the basis of either projected population growth or relative performance compared with non-GM rapeseed. These findings have been supported retrospectively by the total absence of invasive transgenic herbicide-tolerant rapeseed in any of the natural plant communities considered in the study. With the huge benefit of hindsight, it may be tempting to dismiss these findings as entirely predictable or even unnecessary within a modern regulatory framework. Moreover, the scenario falls at the question ‘would herbicide tolerance confer any advantage in the absence of the herbicide?’ given that there are no wild communities that are regularly sprayed with broad-spectrum herbicides. This viewpoint would grossly underplay the importance of these works, however, since they do provide a tractable methodological approach for evaluating the invasive capacity of transgenic crops or of transgenic introgressant wild relatives. The need for a reliable predictor of invasiveness is becoming increasingly germane as the diversity of transgenes in GM crops increases. Specifically, the next wave of transgenic cultivars will include some that possess traits conferring abiotic-stress tolerance, such as drought or salt tolerance and the capability to withstand biotic stresses, such as disease resistance or herbivory. Given that these characteristics have been associated with selection in natural habitats (e.g. Eveno et al., 2008; Franks et al., 2008; Kane
and Rieseberg, 2007; Roelofs et al., 2008), it seems entirely plausible that these plants could exhibit advantageous characteristics across a range of settings outside the farmed environment, giving rise to ecological release. In such cases, the approach devised by Crawley and co-workers could provide a plausible third Tier evaluation for increased invasiveness, although the strategy appears rather protracted and demanding of space to have practical utility for early-stage screening. There consequently remains a pressing need for early-Tier experiments to test for enhanced invasive potential.

An attractive alternative strategy may be to center attention on identifying those features that currently constrain a crop or wild relative to a particular community profile. In some instances, this may be possible from the established literature or knowledge of the crop and its wild relatives. For example, it would be entirely reasonable to speculate that populations of the sea-cliff dwelling wild *Brassica oleracea* (a relative of rapeseed) or of the coastal *Beta maritima* (a relative of sugar beet), although somewhat salt tolerant, may be constrained from invading some of their neighboring maritime communities by a comparative sensitivity to high salt concentrations. At the same time, the present degree of salt tolerance of these species is very unlikely to be important for the invasion of more inland, ‘low-salt’ communities. Less clear is whether the introduction of a transgene providing raised salt tolerance into the entirely salt-sensitive Riverside *Brassica rapa* would enable this species to invade the kind of coastal habitats currently occupied by *B. oleracea*. Likewise, it is difficult to predict whether a transgenic drought-tolerant sunflower, soy, wheat, cotton or their relatives would be likely to invade natural arid communities, even if introduction of the trait allows cultivation of these crops (and associated weeds) to be extended into these areas. Lack of knowledge of the biological feature(s) that currently constrain a crop or wild relative to a particular ecological profile represents the primary factor confounding the decision-making process in such instances. Whilst it is intuitively clear that traits such as these would probably improve the performance of the transgenic individuals when placed in the appropriate setting, other factors such as disease resistance, absence of seed dormancy, susceptibility to herbivory may nevertheless be of overriding importance and so prevent establishment. On the other hand, there are many cases where there is probably already sufficient information to make a reasoned judgment for transgenes conferring some novel traits. For instance, it is difficult to envisage a particularly plausible scenario where the introduction of transgenes that enhance the vitamin content of seeds (Yusuf and Sarin, 2007), product quality (Lu and Kang, 2008), processing characteristics (Brummell et al., 2002) or even herbicide tolerance (Tan et al., 2006) could free any wild relative or crop from ecological constraint and enable it to invade new communities. There is consequently adequate *a priori* evidence upon which to base the decision concerning which events should be guided to tiered evaluation and which do not merit further examination.

Having eliminated the implausible scenarios, Tier 1 assessments can be designed on the basis of crude exposure. In the case of salt tolerance, simple exposure tests could be designed so that the progression to Tier 2 tests could be based on the ability of transgenic plants to survive salt concentrations set some level below that observed in natural saline-dominated communities. Similarly, Tier 1 tests for drought-tolerant GM plants would assess the ability to survive maximum water availability in the arid-dominated habitats under consideration. The underlying reasoning being: if the transgenic plants cannot survive the levels of abiotic stress experienced by native plants in these communities, then they will be unable to invade them. However, if the GM plants are able to survive in these conditions of average stress, it would be important to also assess their ability to survive under conditions that would correspond to exceptional conditions (e.g. extreme drought), since these may be particularly important for determining long-term survival in the target environment. Tier 2 experiments are rather more challenging, given the complexity and heterogeneity of natural communities. Here again, however, it must be remembered that the primary goal of the exercise is to aid decision-making rather than to perform a truly predictive ecological study. Arguably the simplest approach would be to conduct a life history comparison between transgenic and conventional equivalents under growing conditions that simulate those of the natural community under consideration. If the transgenic plants perform no better than the conventional plants, then they are unlikely to be more invasive. Similarly, if the transgenic plants are predicted to progress towards extinction, then they are also unlikely to invade, provided that numbers of founder populations (spillage rates) are not expected to be raised and population decay rates are not projected to be slow (λ is only slightly below 1). Progression beyond Tier 2 would necessitate a simulation of the natural community or transplant experiments of the type developed by Crawley and colleagues. Thus, for the majority of cases the tiered approach to risk assessment continues to offer practical solutions to help regulators and decision-makers address the immediate problem of handling the diversification of transgenes and crops expected in the near future.

**THE ROLE OF PUBLIC-SECTOR RESEARCH**

If the established tiered approach to risk assessment has the capacity to evaluate the new wave of transgene constructs when they arrive, then it is important to establish
the most effective role for public sector research. On the one hand, it can be reasonably argued that much of the work published to date under the broad guise of risk assessment has subsequently proved to have little or no value to the decision-making process and so only of passing relevance to the regulators (Raybould, 2007). On the other, the general public perceives that the public sector has greater independence than industry. It also has the luxury of time to generate data that has generic value that, if framed appropriately, could help stake-holders focus attention on the higher risk scenarios. This is probably the most important role for the public sector.

The first area where public sector research could help regulation revolves around the need for a procedural framework that addresses the key residual issues associated with problem formulation. Moreover, whilst the tiered approach to risk assessment has a strong record for guiding GM regulatory decision-making, and is sufficiently flexible to address a wide range of risk scenarios, it is totally dependent upon a comprehensive coverage of risk scenarios at the problem formulation stage. Omission of a plausible risk scenario arguably provides the greatest scope for error in the risk assessment process and therefore requires careful attention.

As outlined above, enhanced or reduced fitness provides a fitting proxy for harm in a high proportion of cases, and can be applied to entire rafts of risk scenarios that share a common denominator. For example, there is a multitude of risk scenarios associated with population expansion of a crop wild relative growing within its existing communities, including decline in any of its associated flora or fauna. In this context, the establishment by tiered assessment that the transgenic wild relative exhibits no greater fitness than the non-transgenic equivalents effectively discounts all of these scenarios unless there is excessive migration by spillage into the target communities. Difficulty lies in identifying the communities that contain the wild relatives or else that could potentially contain transgenic wild relatives after release from ecological constraint.

Lists of plant communities containing crop wild relatives are already available for some geographic locations, and can be usefully deployed to guide problem formulation. Such information is currently lacking for associated fauna, and this requires fresh data gathering. To illustrate, Rodwell (1991) compiled comprehensive lists of the major and minor component species in different natural plant communities found in the United Kingdom. Cross-referencing of these lists can provide a useful assemblage of known associated species for cross-compatible crop wild relatives. Ford et al. (2006) exploited this resource, and added their own field surveying for associated insects and vertebrates, and then combined these data with species listed under IUCN threat categories (Cheffings and Farrell, 2005) and other listings of species afforded conservational protection. They used these data to identify those plant and animal associates of wild *Brassica rapa* and *B. oleracea* (both cross-compatible relatives of rapeseed) for which legislative protection already exists.

The utility of such lists relates directly to the problem formulation stage and by reference to associated literature, helps define which interactions are likely to warrant greatest attention. In the above study, Ford et al. (2006) identified the one provisionally red-data-listed micro-moth species *Selania leplastriana* (Kent County Council, 2000) that uses *B. oleracea* as a larval food source. It follows that this associate becomes a natural candidate to consider should a Lepidoptera-specific cry1A Bt transgene become widespread within these populations.

Given the innate variability of the natural environment, thought should be given to the considerable difficulty in evaluating for possible enhanced fitness when selection for the trait could be stochastic in strength, distribution and frequency, and where, as in the case of *B. oleracea*, the crop relative is relatively long-lived. Under these circumstances, occasional exceptional years may have disproportionate importance to the overall fitness conferred by the transgene when considered over several generations. For this kind of trait/recipient combination, design of the tiered experiments requires very careful thought if unforeseen spread of a transgene is to be avoided. The most direct approach would be to set conditions in the early-tier experiments to exceed levels that are expected in exceptional years. This *caveat* applies particularly to transgenes conferring resistance to diseases that are prone to epidemics, to abiotic stresses such as exceptional drought that are only occasionally limiting within a population, and to resistance to herbivores that undergo occasional population explosions. In recognition of such difficulties, the generation of data of value for risk management and mitigation measures seems likely to become a key part of the decision-making process. Such data would be particularly valuable if it throws light on the plausibility of prevention or correction measures. A comprehensive knowledge and understanding of known associates of the crop wild relatives would also assist the process of prioritizing efforts to screen for likely consequences and where necessary, to trigger the instigation of corrective measures.

**Identifying environments most likely to be affected**

The outstanding issues to be resolved are the identification of those transgenes most likely to elicit a non-obvious or indirect release from ecological constraint, and to identify which communities are most under threat of invasion. Reference to the associates of wild relatives and feral crops in the natural and semi-natural...
Beyond fitness in environmental risk assessment

Table 1. Empirical studies of airborne pollen dispersal from isolated plots of oilseed rape (*Brassica napus*).

<table>
<thead>
<tr>
<th>Study</th>
<th>Plot size of oilseed rape</th>
<th>% pollen density detected/inferred at 400 m</th>
</tr>
</thead>
<tbody>
<tr>
<td>McCartney and Lacey (1991)</td>
<td>20 m²</td>
<td>0</td>
</tr>
<tr>
<td>Scheffler et al. (1993)</td>
<td>64 m²</td>
<td>&lt; 0.00033%</td>
</tr>
<tr>
<td>Timmons et al. (1996)</td>
<td>10 ha</td>
<td>10–11%</td>
</tr>
<tr>
<td>Cai et al. (2008)</td>
<td>4.5 m²</td>
<td>&lt; 0.015%</td>
</tr>
</tbody>
</table>

communities that they currently inhabit may provide a useful start point for a preliminary tier screen to identify which communities are under threat of invasion. The reasoning is thus: communities that share no species in common with those that currently contain the wild relative in the same geographic region are *a priori* less likely to be realistic candidates for invasion following transgene recruitment. Moreover, if no extant species can span the ecological requirements of both communities, it is rather improbable that the genetic modification of a single member (*i.e.* the crop wild relative) will allow it to do so. At the other extreme, it seems far more plausible that a transgenic wild relative found in one community may be capable of invading another that shares many of the same species in common, especially if the introduced trait appears likely to result in constraint release. In this way, simple cross-references of communities containing feral crops or their wild relatives should yield a reasonable list of candidate communities that may be most vulnerable to invasion by transgenic genotypes. Having established such a list, one obvious option would be to adopt tiered experimentation as outlined above, culminating, only where necessary, in life history transplantation approach as pioneered by Crawley et al. (1993). However, it may not always be possible to perform transplantation experiments, particularly if the candidate receiving community is afforded strict legislative protection because of conservational and/or cultural importance. In these instances it may be necessary to select a similar community type as a reasonable approximation. The problem of conceiving workable early-tier experiments for some transgene/relative combinations remains, however, particularly when the relative is long-lived and the relationship between gene function and enhanced fitness is obscure. In the short term, the most probable outcome is that industry and the regulatory machinery will focus on transgenic events that can be evaluated readily using adaptations of the existing tiered strategy; atypical cases will probably be withdrawn, deferred or else delayed by requests for substantial quantities of additional information.

**Modeling pollen dispersal and gene flow**

Clearly, one area in which the public sector has invested a substantial effort in the general field of risk assessment is in the generation of predictive models to anticipate pollen dispersal and gene flow. Although less interesting from the point of view of fundamental research, and hence less studied, GM seed dispersal can also be a significant means of transgene movement (Aono et al., 2006; Saji et al., 2005). Early predictive models intended for this use focused largely on the seemingly tractable problem of anticipating airborne dispersal of pollen from GM fields. This is nominally with a view to establishing isolation distances for seed multiplication, to anticipate scope for inter-field admixture by wind-mediated gene flow, and to determine scope for long-range gene flow to allopatric wild relatives. Initial attempts were highly empirical, and typically used small plots of transgenic lines as pollen donors, with radiating stands of non-transgenic individuals as pollen recipients and/or transects of volumetric pollen traps. However, variance between such studies was often wide, as aptly illustrated by the disparate estimates for dispersal over 400 m predicted for rapeseed (Tab. 1). Consensus among mathematical models attempting to describe the shape of pollen decay curves also proved elusive, due largely to the low numbers of records at the more distant sites, whether exponential (Lavigne et al., 1998; Manasse, 1992) or the fatter-tailed dispersal curve functions (Devaux et al., 2005; Shaw et al., 2006) were used to describe decay rates. Not surprisingly, regulators made little direct use of these findings, except for rapidly making the presumption of gene flow in cases where gene flow was known to be possible. Models nevertheless continued to improve in their predictive power, especially once the importance of plot size became recognized as a key factor, so that there have been several works capable of predicting pollen dispersal characteristics over a landscape scale (*e.g.* Lavigne et al., 2008; Shaw et al., 2006).

Justification of such models has subsequently evolved, and now either seeks to contribute to holistic models to describe the management issue of admixture between GM and non-GM crops on a landscape scale (*e.g.* Angevin et al., 2008), or more pertinently to assessing the prospects for ecological harm, to develop more predictive temporally- and spatially-explicit models that describe the location, pattern and speed of hybridization and transgene spread. The first of these described the pattern and location of hybrids between rapeseed and...
B. rapa across the United Kingdom (Wilkinson et al., 2003a). From the perspective of GM regulation, the usefulness of information generated by this and more sophisticated models as they arise lay largely in the medium-term, as the quality and quantity of parameterization data improves, along with the predictive power of the model (e.g. see Lavigne et al., 2008). One appropriate example is that a working knowledge of the distribution and abundance of hybrids forms a firm basis upon which to design first-Tier efficacy tests for genetics-based biocontainment measures. Considering the case of GM rapeseed in the UK, less than a thousand F_1 hybrids are expected annually between the crop and B. oleracea (Ford et al., 2006), whereas F_1 hybrids between the same crop and B. rapa may sometimes number over 50,000 (Wilkinson et al., 2003a). Viewed in this context, a biocontainment measure that fails one time in 10,000 would be effective for the former species but not the latter. It also follows that a first-Tier test for biocontainment efficacy would need to be of sufficient size to provide statistical confidence of system compromise in less than 1 in 50,000 attempted cross-pollinations. The next generation of new models will incorporate population dynamics and demographic modeling, together with spatially-explicit models describing population density and distribution and gene movement. These models will have greater predictive capacity, and will help frame worse case scenarios for submissions that pass through to Tier 3. Most importantly, they can be used to inform post-release monitoring efforts, target which aspects of the population or community to observe, and thereby identify significant divergence from expectations to allow for remedial action at an early stage. The same models will also help anticipate the efficacy of mitigation measures and guide regulators in targeting where to perform mitigation and to monitor its effectiveness.

Evaluating the potential for ecological release

There is some scope for optimism that the more problematic potential submissions can be addressed, and that the provision of generic data sets may obviate the need for some of the tiered evaluation. It is here, in the creation of new assessment protocols and through the development of predictive understanding that public sector science has potential to add value to the risk assessment process and thereby to enrich the diversity of low-risk transgenic cultivars that reach the market.

One attractive proposition is to gather data sets that indicate the identity of genetic features and traits that are currently limiting population size or the ecological range of a specified wild relative, so that transgenes conferring traits of concern can be readily identified prior to their inception. The science of predictive ecology is still very much in its infancy, but will undoubtedly improve as we gradually enhance our capacity to exploit genomic, transcriptomic, proteomic and metabolomic datasets from the model and crop species within the context of a crop relative growing in a natural or semi-natural setting. One area of particular interest centers on increasing our understanding of how the expression profiles of genes in stress-response pathways of plants relate to the nature and severity of the biotic or abiotic factors eliciting the response. Moreover, should we progress to a state of knowledge where particular stress-gene expression profiles are associated with different stressors, then we may become more able to identify which factor is causing the highest levels of stress amongst the presumably weaker individuals occupying the margins of the populations or sited on the species’ ecological boundary.

There is now an increasing understanding of the underlying molecular basis of control for the diversity of stress response systems deployed by plants. Remarkably, there is evidence for shared elements in the signaling pathways in the hypersensitive response to pathogens and those in response to drought or saline exposure (de Torres-Zabala et al., 2007; Mur et al., 2006; Pandey et al., 2004). A predictive understanding of these systems in the models and crops is probably still some way off, but has the potential to be extremely useful for risk assessment efforts concerning the stress responses of crop relatives in natural settings.

Perhaps more tractable in the shorter term is the prospect of using association genetics approaches to identify regions of the genome associated with domestication or ecotype differentiation, or when used in combination with summary statistics (e.g. Beaumont et al., 2002; Tallmon et al., 2004), to identify regions of the genome or even genes under selection. Again, both strategies could help throw new light on the kinds of gene regions and traits that are under selection as a plant reaches the margins of its ecological boundary. The provision of such knowledge would identify those genome regions, genes and traits that carry most and least concern in facilitating ecological range expansion of a crop or wild relative.

A similar intriguing prospect is the possibility of combining experimental and modeling strategies that were originally developed for use in other scientific disciplines. For example, Verhoeven et al. (2008) combined the conventional ecological strategy of reciprocal transplantation with quantitative trait loci (QTL) approaches, more usually associated with crop genetics, to identify regions of the genome responsible for ecotype differentiation. Using this dual strategy, the authors performed reciprocal transplantation experiments using mapping populations created between ecotypes of wild barley and found that a QTL associated with flowering-time variation is
Beyond fitness in environmental risk assessment

an important determinant of fitness in each population’s native habitat. Bratteler et al. (2006) also adopted a map-based approach to identify QTLs responsible for ecotype differentiation in Silene vulgaris. Gardner and Latta (2006) used an extension of this strategy that is perhaps more relevant to a risk assessment perspective. These authors grew recombinant inbred lines of Avena barbata under glasshouse conditions to simulate water stress in the natural environment and identified a QTL that conferred different traits in the alternate environments. An intriguing further extension of these strategies would be to combine life history modeling to identify life history stages responsible for population growth, as exemplified by Stokes et al. (2004) on Ulex species and by Bullock et al. (2008) on Rhinanthus minor populations, with map-based transplantation methods such as those used by Verhoeven et al. (2008). Such a line of attack would not only provide information on which traits and QTLs are responsible for ecotype differentiation, but also help identify those associated with transition through each life stage, and identify which life stage transition is of greatest importance in determining ecotype transition. In this way, regulators would be provided with the identity of traits that are most likely to allow range expansion, but crucially, also those that are unlikely to be important. Such data are based on understanding, and so would effectively circumvent the need for some categories of empirical testing. Indeed, there is now a wide range of new genomics-based strategies that are being deployed to identify genes and traits under selection (for review see Kannenberg and Widmer, 2008). The application of these methodologies increase the prospect of developing tractable strategies for identifying ecologically-limiting traits, QTLs and/or genes that would have direct value for the regulatory community. Collectively, the hope is that these protocols should allow future risk assessors to clearly define which new traits or transgenes are most likely to cause concern and in which recipient communities.

Once the harm is identified (e.g. decline of a named beneficial insect species in a named community/location), trigger values can be established to define the acceptable/unacceptable boundary (assessment endpoint). In this case, this might be a particular threshold population size that defines ‘unacceptable’. It is almost always unnecessary for the risk assessor to accurately estimate the size of the endpoint species in order to determine whether it falls above or below the critical threshold; rapid assessment based on Tier 1 experiments will usually suffice. Of greater concern is the possibility that the risk scenario is so poorly defined as to be untestable. This is most likely to apply to concerns relating to ecosystem function and community structure. Here, regulators and legislators must work hard to define more clearly what elements of each would constitute unacceptable harm. Only then can procedures be established to assess and measure risk.

CONCLUSIONS AND RECOMMENDATIONS

As described in the introduction, the GMO-regulatory system will soon be facing a greater diversity of GM crop species, including ones expressing transgenes that will directly affect important ecological parameters such as resistance to major biotic and abiotic stresses. It is clear that this will place extraordinary stress on the regulatory system, making it particularly important to clarify, and wherever possible to simplify, how the environmental impact of GM crops is assessed. If this is not done in a timely fashion, this could have two highly negative consequences. First, this could lead to refusal to deploy GM crops that are clearly of great practical use, and second, if the crops in question are of over-riding importance for the agricultural or broader economy, the inability of the current system to deal with risk assessment may have the effect of leading to wholesale dismantlement of the present regulatory oversight.

There are many features that could contribute to attaining the required increase in efficiency and realism in GMO risk assessment; we will address two here, in which the public sector should be expected to play an essential role.

• First, as mentioned above, there must be better understanding of the proper role for public-sector research. There is no need for the public sector – except when it is the notifier – to engage in routine characterization of a specific GM crop event. However, there is truly a great need for the public sector to engage in research that is generic, particularly since this is often labor- and time-intensive. This includes studies such as those described in the previous section in the area of the impact of modified fitness of crops and crop relatives, but generic research can be equally important in other areas to GMO risk assessment. For instance, it took nearly 10 years to come to a point of completion regarding the potential impact of recombination in virus-resistant transgenic plants (Aaziz and Tepfer, 1999; de Wispelaere et al., 2005; Turturo et al., 2008).

• Second, in order to avoid future incidents like the “Monarch controversy”, which greatly discredited the existing regulatory oversight in the minds of the general public, greater clarity in formulation and prioritization of risk hypotheses must be achieved, and in a manner that takes into account the public’s management goals. As mentioned above, this improved hypothesis formulation and prioritization will also be essential to prevent unjustified regulatory bottlenecks.

Environ. Biosafety Res. 8, 1 (2009)
The key here is to proceed in two steps: first an exhaustive catalog of risk hypotheses, which must be made accessible to the general and scientific public; second, evaluation of each risk hypothesis. It may be argued that this could open a veritable Pandora’s box of endless quibbling about minor potential risks. However, retrospectively, if this procedure had been used with the present GM crops, this would not have significantly increased the burden on either the notifier or the regulators, since the vast majority of risk hypotheses could have been shown to be unimportant from already available knowledge, and would have served to focus attention on a very small number of potential issues.

In closing, it is important to mention that if the GMO regulatory process is to meet the challenge of more numerous GM crops expressing more ecologically complex traits, this will require particular attention to the time frame. Early and complete listing and prioritization of risk hypotheses must take place early, i.e. shortly after reaching proof of principle, in order to provide adequate time for any additional research that may be required. This is of particular importance with traits of increased complexity, which may raise issues that will be more difficult to resolve quickly. The Editors of *Environmental Biosafety Research* (http://www.ebr-journal.org) would be happy for the journal to serve as a forum for “Case Studies” that would support development and prioritization of catalogs of risk hypotheses for novel GMOs.

**ACKNOWLEDGEMENTS**

This paper is based on discussions at the Second Biosafenet Seminar, entitled “Beyond fitness: how do we predict the consequences of enhanced fitness following gene flow from GM crops”, which was held at the ICGEB Biosafety Outstation in Ca’ Tron di Roncade (Italy) on 23–26 January 2008. We thank the participants for their helpful comments. The Biosafenet project is supported by the European Commission (contract # 043025). We thank Wendy Craig for helpful comments and discussion.

Received November 4, 2008; accepted January 9, 2009.

**REFERENCES**


Beyond fitness in environmental risk assessment


Environ. Biosafety Res. 8, 1 (2009)


