



Predicting the spread of herbicide resistance in Australian canola fields

Jeanine Baker* & Christopher Preston

Department of Applied and Molecular Ecology, Cooperative Research Centre for Australian Weed Management, University of Adelaide, PMB 1, Glen Osmond, SA 5064, Australia

Received 19 March 2003; revised 1 June 2003; accepted 23 June 2003

Key words: commercial canola, gene flow, herbicide resistance, mathematical model, transgene

Abstract

A common concern expressed about the commercial release of transgenic canola into cropping systems is the risks of unwanted gene flow between varieties. Experimental data is emerging that answers some of the theoretical questions that have been posed when considering gene flow on a landscape scale. This study developed models that utilise some of this published data in an attempt to quantify the spread of transgenes in a commercial farming system. The models, which included bootstrapping the empirical data and three mathematical simulations, were compared with each other and the published data. One of the mathematical models estimated average resistance frequency by imposing a Poisson distribution around the published mean value for a single transgenic field surrounded by conventional canola fields and the other two were derived from the theory that pollen flow decreased with distance in the form of a log decay curve. The predictions of all models suggested that the average frequency of resistance occurring from pollen flow in neighbouring canola fields, even when multiple transgenic fields are adjacent to the conventional fields, are likely to be below the current internationally accepted thresholds for contamination.

Introduction

The adoption of technologies that permit the direct manipulation of the plant genome has led to a rapid increase in the availability of transgenic crop varieties (Snow & Moran-Palma, 1997). These crops offer significant production advantages such as decreased and easier herbicide use and reduced pesticide use. However, there have been calls for greater evaluation of detrimental risks to the environment and to alternate cropping systems before commercial release can occur. One of the most emotive issues surrounding proposed releases is the potential impact of genes moving from genetically modified crops into neighbouring crops, related weeds, or closely related wild relatives because of the fears that such gene flow could markedly affect farming practices, contaminate produce or irreversibly alter the ecosystem.

In the face of such controversy, the need to realistically assess the risks associated with gene flow from conventional and transgenic crops is critical. The introduction of herbicide tolerant crops, whether herbicide tolerance is the result of standard breeding practices or genetic manipulation, provides an opportunity to collect empirical data, create predictive models about the cropping system and to test the model framework.

Several models have been proposed for predicting the evolution and dynamics of intra- and inter-specific gene flow (Maxwell et al., 1990; Lavigne et al., 1996; Squire et al., 1997; Richter et al., 2000; Colbach et al., 2001a, b). With the exception of GeneSys (Colbach et al., 2001a, b) few of these models describe the demographic and genetic evolution of the plant on a landscape scale and even GeneSys relies on complicated life cycles that must be estimated from multiple parameters. All of these parameters must be estimated from the literature or field surveys and introduce increasing complexity and uncertainty to the

* Author for correspondence
E-mail: jeanine.baker@adelaide.edu.au

models. Other models describe paternity shadows for individual plants (Lavigne et al., 1998). As a result there are a lack of models based on empirical field data that predict the movement of transgenes over a landscape representative of Australian farming conditions.

The publication of a study of gene flow between commercial canola fields (Rieger et al., 2002) provided the first opportunity to create a simple model for assessing the level of pollen-mediated movement of genes between neighbouring canola fields. The Rieger et al. (2002) study quantified the level of gene flow that occurred between herbicide-resistant crops and nearby crops that did not contain the herbicide resistance genes. The study was unique in that it utilised large commercial canola fields (25–100 ha) and covered a range of environments. Only a single source field surrounded by non-herbicide resistant fields was evaluated in the Rieger et al. (2002) study. However, the data provide the opportunity to extrapolate gene flow levels when more than one field of transgenic canola is present in a farming system. The prediction of estimated levels of gene flow under such commercial conditions allows the weighing up of potential gene flow levels against the wider context of conventional breeding practices, environmental risks and the potential benefits of the transgenic variety.

Methods

Data on pollen mediated gene flow between commercial canola fields reported by Rieger et al. (2002) was used to investigate the issue of gene flow between canola fields as the spatial arrangement of conventional versus transgenic fields alters. These data consisted of the percent frequency of resistance in relation to distance from the source field and was used for four different modelling approaches. These four models estimated the average percent frequency of resistance expected in an adjacent field with no isolation barrier. After assessing the relationship between the observed frequency of resistance and distance using JMP 4.04 (SAS Institute Inc.), all models were constructed using a combination of spreadsheets (Microsoft Excel 2000 and Microsoft Visual Basic 6.0). Comparisons of the frequency values calculated by each model were compared with the observed data using matched pairs analysis (JMP 4.04).

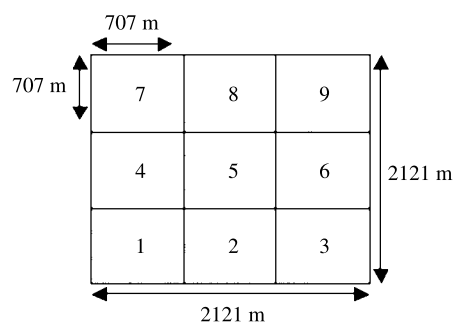


Figure 1. Spatial representation of the field structure used to estimate average percent frequency of resistance in conventional canola fields when an adjacent field (or fields) was classified as transgenic canola.

The spatial design of the fields was based on nine 50 ha fields in a square block design with no separation between the fields (Figure 1). The model parameters assumed that canola was grown for the first time in each of the fields and the user defines the fields as conventional, transgenic or non-canola for each scenario investigated. In a series of runs, the models were used independently to estimate percent frequency of resistance for a set of transects in the adjacent field. Transects were set parallel to the shared boundary of the conventional and transgenic field at 10, 50 and then at every 100 m until the furthest edge of the conventional field was reached. The overall average percent resistance in each field was then calculated to be the average across all transects. Where more than one transgenic field was included in the model, the final percent frequency of resistance in the conventional field was the sum of the contribution from each of the transgenic fields.

For each model, 1000 simulations were run for each transgenic field. These results supplied estimates of the expected average, standard deviations, minimum and maximum percent frequency of resistance within adjacent conventional canola fields. The expected percent frequency of resistance for the following quantiles: 5, 25, 75, and 95% were also calculated. Two scenarios were examined. The first scenario (Scenario 1) was a single, central transgenic field (field 5) surrounded by conventional fields. The results of this could be compared directly with the result obtained by Rieger et al. (2002). The second scenario (Scenario 2) examined the percent frequency of resistance in a conventional central field when fields 2, 4, 6, and 8 were assumed to be transgenic and represented a more extreme scenario possible in commercial cropping systems.

Model 1

The first approach estimated the average percent frequency of resistance in conventional fields *via* a simple bootstrap method which used all the data points from the Rieger et al. (2002) study for each transect. Distance effects were ignored because any point could be used to estimate the average frequency of resistance for each transect.

Model 2

The second of the models was a simple stochastic model that also ignored distance effects. This model mirrored the final observation from the Rieger et al. (2002) study that the mean frequency of resistance in a field adjacent to a single transgenic canola field would have been 0.009%. A gamma distribution function was used to impose a measure of randomness around this mean using the following equation:

$$E(F) = [0.009(\text{gammainv}(\text{rand}(), 0.8, 3))] - 0.025 \quad (1)$$

where $E(F)$ equalled the expected percent frequency of resistance, 0.009 equalled the observed average for the Rieger et al. (2002) study, any value less than zero was constrained to zero. The gammainv function returns the inverse of the gamma cumulative distribution and was used to simulate results for a variable with a skewed distribution. The form of the distribution described by Excel 2000 is described as $\text{gammainv}(\text{probability}, \alpha, \beta)$. In Equation (1), $\text{rand}()$ was the probability associated with the gamma distribution and assigned at random between 0 and 1. The parameters $\alpha(0.8)$ and $\beta(3)$ constrained the skew and tailing of the gamma distribution. For example if $\beta = 1$ then the formula would return a standard gamma distribution.

Model 3

This model was constructed from the average of the observed percent frequency of resistance at each distance for which data was collected. The assumption was that there would be a decay curve over distance for the frequency of resistance. Using a gamma distribution as described above, a level of randomness was superimposed on the final relationship determined from regression analysis undertaken with JMP 4.04. The

final equation used to estimate the percent frequency of resistance was:

$$E(F) = (0.0185 - 0.0016 \log(D)) \times \text{gammainv}(\text{rand}(), 0.35, 3) - 0.01 \quad (2)$$

where D represented the distance measured in metres from the fence line and any value less than zero was constrained to zero. The data reported by Rieger et al. (2002) had no positive values beyond 3000m, but the log form of this relationship could not be forced through zero. Hence, the results from this model were expected to overestimate the expected frequencies of resistance in each field.

Model 4

This model was similar to model 3, but data points were not averaged at each distance. Again a level of randomness was superimposed on the final relationship using a gamma distribution. It was expected the results would be very similar to those obtained using model 3, but included all data points with equal weighting. It was assumed the refining of the log decay curve might provide more realistic estimations of resistance frequencies. The final equation used to

Table 1. The estimated average frequency of resistance (%) in eight canola fields surrounding a single transgenic canola field, with no buffer zone (1000 simulations/model) compared with the observed average frequency of resistance (%) (Rieger et al., 2002)

Data source	Average (%)	Minimum (%) (1000 simulations)	Maximum (%) (1000 simulations)
Observed	0.009		
Model 1	0.010	0.000	0.046
Model 2	0.008	0.000	0.037
Model 3	0.013	0.000	0.073
Model 4	0.013	0.000	0.059

Table 2. The estimated average frequency of resistance in field specified as number 5 when surrounding fields 2, 4, 6 and 8 were designated as transgenic canola fields (1000 simulations/model)

Data source	Average (%)	Minimum (%) (1000 simulations)	Maximum (%) (1000 simulations)
Model 1	0.037	0.001	0.119
Model 2	0.030	0.003	0.078
Model 3	0.051	0.009	0.146
Model 4	0.051	0.007	0.128

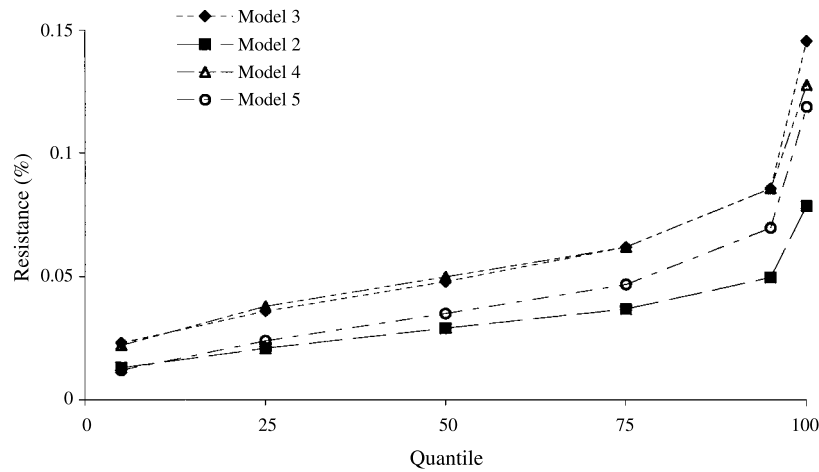


Figure 2. Simulated average frequency of resistance for models 1–4 (1000 simulations/model) in field 5 at indicated quantile when four transgenic fields surround field 5.

estimate the percent frequency of resistance was:

$$E(F) = (0.0200079 - 0.0001095 \log(D)) \times \\ \times \text{gammainv}(\text{rand}(), 0.35, 3) - 0.01 \quad (3)$$

with similar constraints to those outlined for model 3.

Results

A visual examination of the frequency data suggested that the relationship between distance and observed frequency of resistance was in the form of a decay curve (Rieger et al., 2002). Although all observations at distances greater than 3000 m were zero, at distances of less than 3000 m any assumption of a correlation between distance and frequency was not possible, with the best R^2 value for a log decay curve equalling 0.03. This led to the decision to investigate both the bootstrap and stochastic model in conjunction with the two log decay curves in order to determine the model that represented the observed data most effectively.

An assessment of the estimated range and distribution of resistance frequencies in simulated commercial canola fields was compared with the observed data. A histogram was prepared by binning the estimated frequencies of resistance in 0.01% increments and comparing all four models with the observed values (not shown). This confirmed that the gamma distributions used to impose variation around Equations (1)–(3) mirrored the observed data within reasonable limits. Positive or negative bias in the models was

examined by comparing the calculated and observed values ($n = 196$) as matched pairs and testing for any correlation, using Pearson's R , for 10 simulations of each model. The results indicated that there was no pronounced bias in any of the models.

Bootstrapping an existing data set would be expected to mirror that data set most closely because all simulations are created from the data itself. Results for model 1, the bootstrap approach, did reflect the situation as described by Rieger et al. (2002) more closely than the other models, although none of the other models were significantly different to the observed data (Table 1). Including distance as a parameter did not markedly improve model design, indicating the stochastic nature of the data. However, both theory and empirical evidence indicate that distance from the source does influence pollen dispersal (McCartney & Lacey, 1991; Scheffler et al., 1993), so models incorporating distance should be more valid in describing field situations.

The study of Reiger et al. (2002) only examined the situation for one field containing herbicide-resistant canola crop. It did not address the situation for multiple herbicide-resistant canola in the environment. One of the main advantages of predictive models is the ability to vary the parameter values or the assumptions of each simulation. In this study any of the nine fields could be set as transgenic, conventional or no canola, alone or in combination with other fields. This allows the landscape design to be altered and investigated. The following results are for a scenario where four transgenic fields would surround a central conventional field. The average frequencies of resistance

estimated in the central field were between 3.5 and 4 times that estimated for a single resistant field with models 3 and 4 returning the highest estimated average frequency of resistance (Table 2). The average frequency of resistance at the 5, 25, 50, 75, 95 and 100 quantile show a slow increase in the estimated level of resistance, with less than 5% of the simulations close to the maximum figure observed in a total of 1000 simulations (Figure 2). This suggests that there would be less than 1 in 20 occurrences where the average resistance in a field surrounded by four fields planted with transgenic canola would be over 0.1%.

Discussion

Due to differences in experimental design, the published data on gene flow for canola crops is somewhat difficult to reconcile. Studies utilising male sterile plants have found pollination at distances of up to 4 km from the source (Thompson et al., 1999), but the use of isolated male sterile plants would tend to overestimate gene flow due to lack of pollen competition. Other studies have used small scale plots to estimate the rate of outcrossing between plots (Scheffler et al., 1993, 1995; Staniland et al., 2000). Rates up to 1.5% have been measured for individual samples, but such small plots have large edge to area ratios increasing the impact of pollen exchange at the interface. Scheffler et al. (1995) reported outcrossing at 0.0038% for a distance of 400 m, however no field averages could be estimated from this study. The landscape dimensions of the Rieger et al. (2002) study was the first report providing data which could be used to develop models for predicting gene transfer under commercial growing conditions and indicated that gene flow occurs at low frequencies, but potentially over considerable distances.

Until now, simulation models to examine genetic exchange between crop populations have been constrained by a lack of empirical data and have, for the most part, been based on complicated interactions between the lifecycle of the crop and cultivation techniques, and utilised parameters derived from existing models, literature or estimated from field surveys (Squire, 1999; Colbach et al., 2001a, b). GeneSys, the most recent (Colbach et al., 2001a, b), was designed to predict the gene flow among a volunteer and cropped population of canola over time. The authors concluded that levels of outcrossing and persistence were low, although it was possible to predict canola volunteers

occurring at distances greater than 1 km. In this, the model reflected the empirical observations reported by Rieger et al. (2002), but required extensive life cycle information.

Models 1–4 presented are simple in design, based on empirical data and can rapidly evaluate and predict levels of gene flow in varying commercial situations. Some basic limitations influence each of the models in regard to the estimations of average frequency of resistance. The predictions make no allowance for point contact between fields and are based on the premise that fields have a long edge contact. It is probable that fields aligned corner to corner, even when not separated spatially, would have lower levels of gene flow. Thus, predictions based on a long edge contact are likely to overestimate the contribution to the average resistance found in each transgenic field if the spatial configuration of fields is other than shown in Figure 1, or if fields are aligned corner to corner. Another limitation is that the data used to define each of the models was also used to test the model simulations. Unfortunately, until another study of commercial fields is undertaken the models cannot be tested with other data.

Model 1, a bootstrap approach using the 196 values recorded by Rieger et al. (2002), assumes that the observed values provide a large enough sample size for re-sampling to realistically estimate values over thousands of simulations. As such this model represents the lower boundaries of what might be expected in other field situations. Model 1 was created as a means of checking the accuracy of the other models. Model 2, although it mirrors the empirical data more closely than either model 3 or 4, is limited because it is a random distribution around a known mean. There is no potential to use model 2 to investigate scenarios relative to distances between fields or how isolation barriers might influence outcomes. Both models 3 and 4 can be assumed to be overestimates, particularly as distance from the source increases. The central problem with models 3 and 4 is the relationship used to describe them (log decay curve). This curve cannot be forced through zero as the distances increase and the R^2 -values of the observed data failed to suggest that there was a true relationship between distance and frequency of resistance. However, these two models supply estimates that allow some use of distance as a parameter and provide estimates that can be regarded as the upper bounds of what may be expected in reality. As long as these limitations are borne in mind it was concluded that either model 3 or 4 would

prove useful to predict the effect of different management options for the placement of conventional and transgenic canola fields.

Using the estimates from model 4, it was predicted that the maximum value of resistance over a conventional canola field surrounded by four transgenic fields was 0.13%. As long as this level is below any set threshold level for gene flow between varieties, then combining conventional and transgenic commercial systems spatially would be possible. Systems consisting of two or three transgenic fields surrounding a conventional field would be expected to have resistance levels intermediate between one and four fields. Using the same method of estimating the frequency of resistance outlined in Methods the level of average frequency of resistance was predicted for two or three transgenic fields, selected at random around field five. The results were intermediate to the scenario with one and four fields and provide predictions of the average level of gene flow for different farming situations (Figure 3).

Other scenarios, including fields separated by non-canola fields, can be examined. For example, the maximum average gene flow from field one to field nine was calculated from 1000 simulation using model 4 and estimated to be 0.08%. In this case, it should be remembered that the further fields are separated by distance the more the likely it is that the estimated frequency of resistance would be overestimated, as a result of the limitations of model 4. Therefore, the value of 0.08% could be considered to be an overestimate, but still indicates that the level of gene flow would be low. Overall, the results concur with the empirical study undertaken by Rieger et al. (2002) and with the model predictions made by Colbach et al.

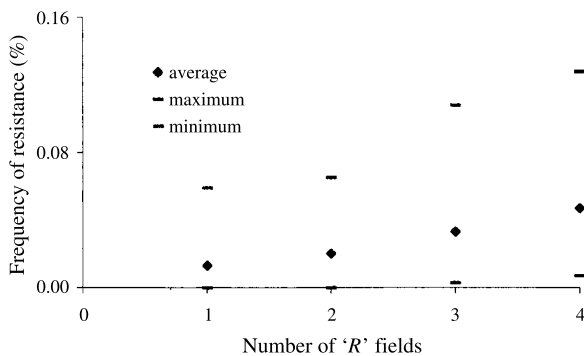


Figure 3. Average frequency of resistance in field specified as 5 (Figure 1) surrounded by transgenic canola specified as being present in fields 1, 2, 3, or 4 (model 4, 1000 simulations, transgenic fields selected at random).

(2001a, b). The current accepted level of contamination in the Japanese market is 5% and the European market is proposed at 0.9%. So, even when surrounded by four transgenic canola fields the maximum percentage of resistance estimated across the entire central canola field was 0.13%, which is far lower than the threshold levels.

Canola pollen is quite heavy and sticky, so most of it remains airborne for only a few metres. Canola pollen is subject to desiccation and even if it managed to travel a long distance, it may not be viable. The paternity shadow of a single canola plant has been shown to be limited to closely adjacent plants and selfing is the predominant method of fertilisation (Lavigne et al., 1996, 1998). Outcrossing events have been noted to occur more frequently close to the interface between two fields (Staniland et al., 2000). Edge effects have certainly been reported in several studies investigating pollen movement (Rakow & Woods, 1987; Scheffler et al., 1993, 1995; Staniland et al., 2000), although there was no strong evidence of an edge effect reported by Rieger et al. (2002). The ecology of pollen transfer and pollen competition within fields would then indicate that it is not the size of the source field that influences the amount of pollen being transferred across boundaries between neighbouring fields, but rather the length of the boundary contact. The size of the sink field, unless very small, would also have little influence due to pollen competition from within the field. This theory, along with the findings from the model scenarios, support the supposition that any average resistance observed within a sink field would be extremely low in the first year. This has implications for management recommendations to reduce unwanted outcrossing in subsequent years.

Temporal models have identified the importance of seed bank contribution to the evolution of higher rates of resistance (Squire, 1999; Colbach et al., 2001a, b), but current Australian management practices for canola rely on crop rotation to minimise the presence of volunteers. As long as genes are not stacked for herbicide resistance within the transgenic variety, thus minimising the choice of herbicides available for use, persistent volunteers can be controlled by changing herbicides. While maintaining crops at different stages of development or separating them spatially can alter crop-to-crop gene flow it appears that such complex management options would not be necessary unless seed is being grown for certification.

The consequence of intra-species gene flow needs to be assessed against the knowledge that gene flow

is a natural phenomenon and in consideration of commercially agreed threshold limits. The results from the models described above suggest that it will be possible to combine conventional and transgenic canola farming systems and meet reasonable threshold limitations. Further investigations into the temporal factors affecting the seed bank and the persistence of volunteer canola plants within managed fields and along roadsides will provide more information that can be used to refine agricultural practices that could be used to manage gene flow. This information would also assist in the development of spatial models to incorporate temporal factors without overly increasing the complexity of mathematical models used to predict the long-term effects of commercially releasing transgenic varieties of canola.

Acknowledgements

We would like to acknowledge funding support from the CRC for Australian Weed Management and the helpful comments of the anonymous reviewers. Thanks also to Dr Rieger and her co-authors for allowing the use of the raw data set used to develop the models.

References

- Colbach N, Clermont-Dauphin C et al. (2001a) GeneSys: a model of the influence of cropping system on gene escape from herbicide tolerant rapeseed crops to rape volunteers. I. Temporal evolution of a population of rapeseed volunteers. *Agric Ecosyst Environ* **83**: 235–253.
- Colbach N, Clermont-Dauphin C et al. (2001b) GeneSys: a model of the influence of cropping system on gene escape from herbicide tolerant rapeseed crops to rape volunteers. II. Genetic exchanges among volunteer and cropped populations in a small region. *Agric Ecosyst Environ* **83**: 255–270.
- Lavigne C, Godelle B et al. (1996) A method to determine the mean pollen dispersal of individual plants growing within a large pollen source. *Theor Appl Genet* **93**: 1319–1326.
- Lavigne C, Klein EK et al. (1998) A pollen-dispersal experiment with transgenic oilseed rape. Estimation of the average pollen dispersal of an individual plant within a field. *Theor Appl Genet* **96**: 886–896.
- Maxwell BD, Roush ML et al. (1990) Predicting the evolution and dynamics of herbicide resistance in weed populations. *Weed Technol* **4**: 2–13.
- McCartney HA, Lacey ME (1991) Wind dispersal of pollen from crops of oilseed rape (*Brassica napus* L.). *J Aerosol Sci* **22**: 467–477.
- Rakow G, Woods DL (1987) Outcrossing in rape and mustard under Saskatchewan (Canada) prairie conditions. *Can J Plant Sci* **67**: 147–152.
- Richter O, Boettcher U et al. (2000) Modelling of spatial spread of herbicide resistance. *Zeitschrift-fuer-Pflanzenkrankheiten-und-Pflanzenschutz*. (Special Iss. 17): 329–336.
- Rieger MA, Lamond M et al. (2002) Pollen-mediated movement of herbicide resistance between commercial canola fields. *Science* **296**: 2386–2388.
- Scheffler JA, Parkinson R et al. (1993) Frequency and distance of pollen dispersal from transgenic oilseed rape (*Brassica napus*). *Transgenic Res* **2**: 356–364.
- Scheffler JA, Parkinson R et al. (1995) Evaluating the effectiveness of isolation distances for field plots of oilseed rape (*Brassica napus*) using a herbicide-resistance transgene as a selectable marker. *Plant Breed* **114**: 317–321.
- Snow AA, Moran-Palma P (1997) Commercialization of transgenic plants: potential ecological risks. *Bioscience* **47**: 86–96.
- Squire GR (1999) Temperature and heterogeneity of emergence time in oilseed rape. *Ann Appl Biol* **135**: 439–447.
- Squire GR, Burn D et al. (1997) A model for the impact of herbicide tolerance on the performance of oilseed rape as a volunteer weed. *Ann Appl Biol* **131**: 315–338.
- Staniland BK, McVetty PBE et al. (2000). Effectiveness of border areas in confining the spread of transgenic *Brassica napus* pollen. *Can J Plant Sci* **80**: 521–526.
- Thompson CE, Squire G et al. (1999) Regional patterns of gene flow and its consequences for GM oilseed rape. Gene Flow and Agriculture. Relevance for Transgenic Crops. BCP Council, University of Keele, Staffordshire.