NEWS AND COMMENTARY

New statistical insights into geneflow sampling strategies

Does flax have the answer to the GM mix-up?

MJ Wilkinson

Few things in science are as contentious or politically charged as gene flow studies on genetically modified (GM) crops. Even so, flax is very much a 'C-list' member of the GM crop hall of fame and so the prospective characterisation of pollen-mediated gene flow between a (thus far non-existent) GM flax variety and organic flax seems unlikely to set many pulses racing. However, it would be a mistake to dismiss the study by Jhala et al. (2011, in this issue) on this basis, as there is a lot more to this work than first meets the eye. Unlike the plethora of similarly labelled counterparts that pervade the literature, this study does have the capacity to influence both policy and science. Surprisingly perhaps, this is not achieved via new insights into the intricacies of flax-isolation distances, but rather in the provision of a strategy for data gathering that maximises the relevance of a negative result when searching for a rare event. True, the chosen case study is the search for a rare transgenic flax seed among a proverbial sack-full of non-GM equivalents, but the principle could be applied more broadly to gene flow and to other biological contexts.

The main thrust of the work is directed towards the perfectly valid goal of characterising gene-flow decay curves from future GM-flax fields, and the authors elected to do so by using replicated small-scale (non-GM) field trials rather than commercial fields. The deployment of a simple 'star plot' trial design to measure gene flow from a tiny central block of transgenic plants evokes 'groundhog day'-esque recollections of the many similar studies of the early 1990s (for example, McCartney and Lacey, 1991; Scheffler et al., 1993). Such experiments can have direct relevance to elements of the seed production industry, where similarly small plots are often used for seed multiplication. However, under normal circumstances the wide divergence between gene-flow rates from such small experiments and those seen between commercial fields (Timmons et al., 1995; Rieger et al., 2002)

complicates predictions of pollen-mediated transgene admixtures at the farm level and can devalue such works. This is not the case in this study, however, as the true value of this study resides in the statistical approach used to overcome the common, if not habitual, under-sampling in similar works to describe decay profiles. This is a little gem. Moreover, the authors have circumvented the implausible political question of deciding how many samples are needed to show that there are zero hybrids in a sample, a demand actually made by many in organic farming (the answer, incidentally, is simple but impossible—test all seeds!), and replaced it with the more reasonable and tractable 'how many samples are needed to be sure that there is less gene flow than a set threshold with a specified level of confidence?'

This approach has utility to regulators and enforcement authorities aiming to fix on isolation distances that allow commercial coexistence of GM and non-GM crops. This problem is not insubstantial. To illustrate, although all member countries in the European Union are legally bound to the totally arbitrary 0.9% GM admixture tolerance threshold in non-GM seeds, political expediency coupled with uncertainty over the relative merits of sometimes contradictory pollen-mediated gene-flow studies has contributed to significant variance in the isolation distances being imposed at the national level (Devos et al., 2009). However, Table 1 in this manuscript (Jhala et al., 2011) provides absolute guidance to the risk-assessment researcher over the minimum sample size to assess for gene flow below set thresholds, with a given statistical power and confidence interval for use on flax. To my mind, the approach could be readily adapted to assemble similar tables for other crops. This would provide researchers with an invaluable aid to experimental design and help regulators, policy-makers and notifiers with a simple way to compare between incongruous data sets. In reality, the effects of variable plot size still

need to be accommodated before the model has true resonance to the commercial setting. Nevertheless, the statistical approach could help address the central difficulty concerning the often inadequate sampling at longer distances of large field studies, where divergent decay functions (usually exponential or power law-based) yield ostensibly similar declines at proximal distances, but differ over larger distances where sampling is often underpowered. Clearly, the skewing of sample sizes to enable threshold limits to be maintained at increasing distances allows for far more meaningful fits to empirical data, and so provides a more resilient basis for the growing number of landscape-scale modelling efforts (for example, Shaw et al., 2006; Coleno et al., 2009; Le Bail et al., 2010).

This work provides a much-needed step toward rapid inter-study comparisons, and thereafter can only help build a scientific consensus on the effect on admixture rates relative to set thresholds. That said, one must exercise extreme care before using such works as the basis for decisions on coexistence. Moreover, seeds are typically harvested at (a minimum of) the field scale. This means that pollen-mediated gene flow between neighbouring fields will be 'averaged' across the field rather than the headline rates reported at marginal distances. This will be heavily dependent on field size and shape, and the length of shared boundaries, as well as crop density, mean farm size and the extent of shared boundaries between farms. As with so many of the issues concerning GM crops, however, agricultural reality is completely uncoupled from the main focus of scientific risk-assessment activity. As the authors point out, 'pollen is only one source of adventitious presence'. History suggests that the many sources of seed-mediated gene flow (crop volunteers, mixtures during seed multiplication, transport, planting, harvest, post-harvest transport, and handling by intermediates and end-user food industries), when considered collectively, are likely to dwarf the effects of the modest field-to-field pollen-mediated rates in almost all crops. To get a firm handle on this is a far more exacting task, largely because it involves idiosyncrasies of local supply chains and market forces. This, as they say, is a study for another day. All works on pollen-mediated gene flow run the risk that the scientific interest they foster may disappear should political interest wane. Here, however, I feel that the

908

provision of a simple strategy to determine sample sizes for set confidence intervals of falling below a named (low) threshold is something that is of far more generic appeal. I just hope the message is not simply lost in the crowd.

Conflict of interest

The author declares no conflict of interest.

Dr MJ Wilkinson is at the Institute of Biological Environmental and Rural Sciences, Aberystwyth University, Edward Llwyd Building, Aberystwyth SY23 3DA, UK.

e-mail: jjw@aber.ac.uk

Coleno FC, Angevin F, Lecroart B (2009). A model to evaluate the consequences of GM and non-GM segregation scenarios on GM crop placement in the landscape and cross-pollination risk management. *Agric Syst* **101**: 49–56.

Devos Y, Demont M, Dillen K, Reheul D, Kaiser M, Sanvido O (2009). Coexistence of genetically modified (GM) and non-GM crops in the European Union. A review. *Agronomy Sustain Dev* **29**: 11–30.

Jhala AJ, Bhatt H, Topinka K, Hall LM (2011). Pollen-mediated gene flow in flax (Linum usitatissimum L.): can genetically engineered and organic flax coexist? *Heredity* **106**: 557–566.

Le Bail M, Lecroart B, Gauffreteau A, Angevin F, Messean A (2010). Effect of the structural variables of landscapes on the risks of spatial dissemination between GM and non-GM maize. *Eur J Agronomy* **33**: 12–23.

McCartney HA, Lacey ME (1991). Wind sipersal of pollen from crops of oilseed rape (*Brassica napus* L). *J Aersol Sci* **22**: 467–477.

Rieger MA, Lamond M, Preston C, Powles SB, Roush RT (2002). Pollen-mediated movement of herbicide resistance between commercial canola fields. *Science* **296**: 2386–2388.

Scheffler JA, Parkinson R, Dale PJ (1993). Frequency and distance of pollen dispersal from transgenic oilseed rape (*Brassica napus*). *Transgenic Res* **2**: 356–364.

Shaw M, Harwood T, Wilkinson M, Elliott L (2006). Assembling spatially-explicit landscape models of pollen and spore dispersal by wind for risk assessment. *Proc Roy Soc B* **273**: 1705–1713.

Timmons AM, O'Brien ET, Charters YM, Dubbels SJ, Wilkinson MJ (1995). Assessing the risks of wind pollination from fields of genetically modified *Brassica napus* ssp. *oleifera*. *Euphytica* **85**: 417–423.

Editor's suggested reading

Ostrowski M-F, Rousselle Y, Tsitrone A, Santoni S, David J, Reboud X *et al.* (2010). Using linked markers to estimate the genetic age of a volunteer population: a theoretical and empirical approach. *Heredity* **105**: 358–369.

Ellis JR, Bentley KE, McCauley DE (2008). Detection of rare paternal chloroplast inheritance in controlled crosses of the endangered sunflower *Helianthus verticillatus*. *Heredity* **100**: 574–580.