

---

# 8

## Gene Flow in Forest Trees: Gene Migration Patterns and Landscape Modelling of Transgene Dispersal in Hybrid Poplar

GANCHO T. SLAVOV<sup>1</sup>, STEPHEN P. DIFAZIO<sup>2</sup> AND STEVEN H. STRAUSS<sup>1</sup>

<sup>1</sup>Oregon State University, Corvallis, OR 97331, USA; <sup>2</sup>Oak Ridge National Laboratory, Oak Ridge, TN 37830, USA; E-mail: Steve.Strauss@oregonstate.edu

---

### Abstract

The extent to which transgene flow from plantations can be effectively predicted, managed and monitored will be a critical factor influencing the adoption of transgenic plantations. Studies of historical and contemporary gene flow levels, via genetic structure surveys and parentage analyses, demonstrate that gene flow is generally extensive in both wind- and animal-pollinated forest tree species. Marker studies, however, give little insight into the extent and consequences of gene flow in actual deployment scenarios, because realized gene flow depends on a large number of interacting ecological and genetic factors. For example, despite the potential for extensive gene flow, there appear to be extremely low levels of introgression into wild poplar stands from plantations of hybrids. We introduce a spatial simulation model called STEVE that we have used to integrate a variety of ecological and genetic processes that affect gene flow from poplar plantations in the US Pacific Northwest. The model enables simulation of gene flow on a large landscape over long time frames (e.g. 50–100 years), and permits virtual experiments that investigate how diverse genetic, ecological and management factors might influence the magnitude and variance of gene flow. Similar approaches could be used elsewhere to help identify and prioritize research needs, suggest the strength of mitigation measures such as engineered sterility where they are warranted, and to aid in design of monitoring programmes for large-scale research and commercial applications. We present simulation results which suggest that long-distance gene flow dynamics and creation of a suitable habitat for tree establishment are of considerably more importance to predicting transgene spread than knowledge of intrapopulation mating patterns. The model also predicts that sterility mechanisms, even if incomplete, can greatly impede transgene spread.

## Introduction

Early studies of gene flow in forest trees focused on the physical dispersal potential of pollen and seeds. Direct tracking of propagule dispersal and retrospective analyses of population structure have revealed the astonishing mobility of tree pollen (Lanner, 1965; Levin and Kerster, 1974; Di-Giovanni *et al.*, 1996) and seeds (Clark *et al.*, 1998; Cain *et al.*, 2000). The last few decades have seen extraordinary growth in knowledge of gene flow patterns as a result of extensive biochemical and molecular marker studies (e.g. Loveless and Hamrick, 1984). Most of these studies have employed allozymes, which provide a low cost means for studying patterns of genetic differentiation and, by inference, gene flow. However, recent studies, particularly using highly polymorphic DNA markers, have enabled more precise and more direct inferences, particularly of within-population mating patterns and immigration (reviewed below).

Until recently, interest in gene flow was largely restricted to a modest subset of evolutionary and population biologists, and to plant breeders, for whom it was important to interpret patterns of phylogeny, adaptation and degree of contamination in seed production populations. With the advent of genetic engineering, and the biopolitical and legal controversies it has generated, gene flow has moved on to the public stage (e.g. Ellstrand, 2001). Because of the strong sentiments against genetic engineering of some sectors of the public, even small amounts of contamination appear capable of generating great concern (Thompson, 2001). The extent to which we understand, can predict and can efficiently monitor gene flow may therefore be an important biological determinant of whether transgenic crops are adopted and publicly accepted.

Because of the extensive gene flow possible from trees, and their very limited history of domestication, it has long been known that genes from intensively bred or engineered trees are highly likely to enter wild populations unless very special measures – such as use of sterile trees – are taken to avoid it (Strauss *et al.*, 1995). In the southern hemisphere, where some of the most intensive plantation forestry in the world occurs, most of the trees grown are exotics. It has been suggested that there will be less ecological and social concern over dispersal of transgenic exotics compared with that of native tree species (Burdon and Walter, 2001). However, in many of these places, the planted trees have feral, 'naturalized' forms, and in some cases these have given rise to invasive weeds that, because of their size, can have highly deleterious impacts on local ecosystems (Richardson, 1998; Ledgard, 2001). In contrast, transgenic forms of native species are expected to occupy the same niches as the native species, and thus should not displace other species (Strauss, 1999) or radically alter ecosystem processes. Dispersal of some kinds of transgenes in exotic species may therefore present more of an ecological hazard – if it promotes spread or makes control more difficult – than that from native trees.

The frequency and consequences of gene flow may therefore be globally important issues for the future of plantation forestry. Unfortunately, we

have a limited knowledge base for risk assessments. There have been very few studies of gene flow from conventional plantations to wild populations. Due to the long lifespans and generation times of trees, predictions of impacts must consider very large temporal and spatial scales, and thus embrace high levels of uncertainty about future ecological conditions (James *et al.*, 1998). New studies, and new tools, will be useful for making rational inferences about possible impacts of transgenic plantations.

We review in detail what is known about gene flow in forest trees, and then consider landscape simulation methods that might be employed for analysing, predicting and monitoring gene flow from tree plantations. As a case study, we focus on a recent analysis we undertook for poplars (genus *Populus*) in the Pacific Northwest USA. We suggest that simulation models can help to guide research and to aid in prediction and monitoring during commercial development, helping to inform environmental assessments of transgenic forestry.

## Measuring Historical Levels of Gene Flow

The advent of genetic markers revolutionized methods of measuring gene flow. Over the last three decades, rates of migration typically have been inferred from the degree of genetic differentiation among populations as measured by the fixation index  $F_{ST}$  (Wright, 1931), or its many extensions and analogues (e.g.  $G_{ST}$ , representing the interpopulation component of total gene diversity (Nei, 1973)). Other approaches for measuring historical gene flow include the 'rare allele method' (Slatkin, 1985) and coalescent methods (Beerli and Felsenstein, 1999). All such methods of measuring gene flow are 'indirect' because they apply genetic models (and underlying assumptions) to infer long-term levels of gene flow (Neigel, 1997; Sork *et al.*, 1999). Indirect measures reflect the complex interactions of all demographic parameters and evolutionary forces acting on a population, and the resulting gene flow estimates should be taken as long-term averages estimated over a large number of populations (Sork *et al.*, 1999).

Although inapplicable for estimating contemporary gene flow, indirect approaches have provided a number of valuable insights about historical forces shaping forest tree genetic structure. There have been a large number of studies of allozyme gene diversity, geographical structure and gene flow among populations of forest trees (reviewed in Govindaraju, 1989; El-Kassaby, 1991; Hamrick *et al.*, 1992; Müller-Starck *et al.*, 1992; Hamrick and Nason, 2000), and a few generalizations have emerged.

1. *Trees are characterized by higher genetic diversity and lower levels of differentiation compared with other plant groups* (Hamrick *et al.*, 1992). The interpopulation component of total gene diversity (based on  $G_{ST}$ ) of woody species rarely exceeds 10–15% (Table 8.1). This low differentiation suggests extensive gene flow among tree populations. However, some authors have hypothesized that the observed patterns of genetic variation may

**Table 8.1.** Neighbourhood sizes and per generation migration events estimated by two indirect methods for some tree genera.

Genus	$G_{ST}$	$N_e m$	$N_e m^*$	$N_a$
<i>Abies</i>	0.063	3.72	8.71	260
<i>Picea</i>	0.055	4.30	–	133–260
<i>Pseudotsuga</i>	0.074	3.13	–	–
<i>Pinus</i>	0.065	3.60	1.51–21.83	158–534
<i>Quercus</i>	0.107	2.09	2.24–6.74	–
<i>Populus</i>	0.041	5.85	–	–
<i>Eucalyptus</i>	0.169	1.23	0.76–6.51	–

$G_{ST}$ , averages from Hamrick *et al.* (1992);  $N_e m$ , per generation migration events estimated based on the  $G_{ST}$  values;  $N_e m^*$ , estimated for some species using the method of Slatkin (1985), from Govindaraju (1989);  $N_a$ , neighbourhood area in  $m^2$  (area from which the parents of some central individual may be treated as randomly drawn) estimated for some species, from Govindaraju (1988).

also be caused by other aspects of tree life history. For example, long pre-reproductive phases tend to mitigate founder effects, and long generation times delay differentiation through genetic drift (e.g. Kremer, 1994; Austerlitz *et al.*, 2000).

2. *Wind-pollinated tree species typically have interpopulation differentiation levels of less than 10%, which translates to more than two successfully established migrants per population in each generation in ideal populations.* Hamrick *et al.* (1992) reported an average interpopulation differentiation of 8% for wind-pollinated trees based on 146 data sets. In species with large and continuous ranges, interpopulation differentiation is often below 3% (e.g. *Pinus ponderosa* and *P. contorta* (El-Kassaby, 1991); *P. sylvestris*, *Picea abies*, *Quercus petraea* and *Fagus sylvatica* (Müller-Starck *et al.*, 1992)). At the opposite extreme, species with small and fragmented populations can have interpopulation differentiation in the range of 15–30% (e.g. *P. cembra*, *P. halepensis*, *P. nigra* and *Castanea sativa* (Müller-Starck *et al.*, 1992); *P. torreyana* and *P. muricata* (Hamrick *et al.*, 1992)). This observation suggests that even though long-distance pollen dispersal is possible, its effect may be insufficient for genetic homogenization of spatially isolated populations.

3. *Outcrossed animal-pollinated tree species have a detectably (but not significantly) higher degree of interpopulation genetic differentiation compared with wind-pollinated trees.* The average interpopulation differentiation for animal-pollinated trees with mixed mating systems was 10% based on 37 data sets (Hamrick *et al.*, 1992). As in wind-pollinated species, spatial distribution seemed to be a good predictor of the degree of differentiation. Moran (1992) reviewed interpopulation differentiation in Australian eucalypts and cited values in the range of 8–12% for widespread species (e.g. *Eucalyptus saligna*, *E. cloeziana* and *E. delegatensis*); 30% for the highly disjunct *E. nitens* (widespread, but with a highly discontinuous distribution); and 61% for isolated populations of *E. caesia*. However, 67 scattered and putatively isolated low

density populations of the insect-pollinated, and presumably bird-dispersed, *Sorbus torminalis* had interpopulation differentiation of only 15% (Demesure *et al.*, 2000), suggesting that gene flow rates can, in some cases, be high among spatially isolated populations (but see Austerlitz *et al.*, 2000).

## Direct Methods of Measuring Gene Flow Through Parentage Analysis

Direct observations of gene dispersal obviate the need for tenuous assumptions about historical conditions. Instead, they provide short-term 'immigration snapshots'. Generally, direct methods require the genotyping of all potential parents in a population and estimation of the proportion of progeny that could *not* have been produced by within-population mating. One approach employs simple paternity exclusion, feasible where the maternal genotype can be readily determined (Smith and Adams, 1983; Devlin and Ellstrand, 1990). However, the low variability of allozyme markers greatly limits the ability to distinguish between local and immigrant genotypes (Adams, 1992). To help overcome this problem, a number of methods employ maximum likelihood either to assign parentage (Meagher, 1986; Adams *et al.*, 1992; Smouse and Meagher, 1994; Kaufman *et al.*, 1998) or to estimate mating parameters that provide the best fit to observed progeny genotypes (Devlin *et al.*, 1988; Roeder *et al.*, 1989; Adams and Birkes, 1991).

In the early 1990s, highly variable DNA-based markers began to become affordable for parentage analyses. The high genetic resolution provided by microsatellite and amplified fragment length polymorphism (AFLP) data allowed scientists to conduct paternity analyses based on genotypic exclusion with acceptable levels of discrimination (Dow and Ashley, 1998; Streiff *et al.*, 1999; Lian *et al.*, 2001), as well as to apply maximum likelihood assignments with greater confidence (Gerber *et al.*, 2000; Kameyama *et al.*, 2000).

## Seed orchards

There have been a number of studies of genetic contamination of forestry 'seed orchards'. Seed orchards are plantations in which selected genotypes are placed to allow cross-pollination for production of seeds to be used for reforestation. Most of these orchards are within the range of native or planted populations, and distinct orchard blocks (that service distinct ecogeographical regions) are, in many cases, planted close together for management efficiency. Thus, there are often high levels of unwanted pollen immigration into the orchard blocks. This can result in substantial loss of genetic gain compared with expectations based on selection theory, and can compromise adaptability of seed orchard crops to their intended plantation environments (Adams and Burczyk, 2000).

Pollen contamination in seed orchards is usually estimated via simple paternity exclusion, adjusting the observed proportion of immigrants by the probability that an immigrant gamete will be distinguishable from the

potential orchard gamete pool (Adams and Burczyk, 2000). Other statistical procedures have also been implemented (El-Kassaby and Ritland, 1986; Plomion *et al.*, 2001). These studies have revealed great variation in pollen contamination, even when the same analytical approaches and tree species are considered (Table 8.2; Adams and Burczyk, 2000). None the less, it is clear that pollen contamination is often very large, commonly exceeding 40%, even when the closest stands of the same species are several hundred metres away.

## Commercial plantations

We know of very few studies of the effects of gene flow from forest plantations on wild populations. The best-studied cases are from interspecific hybrids in the genus *Populus*, which are often planted in proximity to wild populations. In Europe, most of the planted cottonwood hybrids include the native *Populus nigra* as a parent. These hybrids are interfertile with wild populations of *P. nigra*, whose populations are greatly reduced in extent due to destruction of riparian habitat by farming and human habitations. None the less, several studies have reported low levels of gene flow into wild *P. nigra* stands (Legionnet and Lefèvre, 1996; Benetka *et al.*, 1999, 2002; Lefèvre *et al.*, 2001; Heinze and Lickl, 2002). Likewise, in the Pacific Northwest USA, gene flow from hybrid poplar plantations into wild American black cottonwood (*Populus trichocarpa*) populations was extremely low, despite the presence of

**Table 8.2.** Estimated pollen contamination in clonal seed orchards of forest trees.

Tree species	<i>S</i>	<i>D</i> <sub>1</sub>	<i>b</i>	<i>m</i>
<i>Picea abies</i>	13.2	None	–	0.55–0.81 <sup>a</sup>
	–	–	0.10–0.17	–
<i>Picea glauca</i>	–	1000	–	0.01
<i>Pseudotsuga menziesii</i>	1.8–3.3	None	–	0.29–0.91
	20	500	–	0.11
<i>Pinus sylvestris</i>	22.9	2000	–	0.48
	3	1000	0.15	–
	6	500	0.38	–
	12.5	> 100	–	0.72
<i>Pinus taeda</i>	2	100	–	0.36
<i>Pinus pinaster</i>	–	None	0.36 <sup>b</sup>	–
<i>Pinus contorta</i>	4.4	200	0.05 <sup>c</sup>	0.08 <sup>c</sup>
<i>Pinus thunbergii</i>	0.5	> 500	0.02 <sup>d</sup>	–

*S*, area in which all potential parents have been genotyped (ha); *D*<sub>1</sub>, distance (m) to the nearest population or tree of the same species; *b*, observed proportion of immigrant pollen gametes; *m*, pollen contamination adjusted for the probability to distinguish local and migrant gametes.

References from Adams and Burczyk (2000), unless otherwise indicated: <sup>a</sup>Pakkanen *et al.* (2000); <sup>b</sup>Plomion *et al.* (2001); <sup>c</sup>Stoehr and Newton (2002); <sup>d</sup>Goto *et al.* (2002).

large male plantations in close proximity to native female trees and the large-scale clearing of wild populations for agriculture (DiFazio, 2002).

## Wild populations

Due to restrictions on marker resolution and the high cost of genotyping, spatially distinct forest stands are usually chosen for study, and parentage analysis is then applied to a sample of progeny. Typically, seeds are collected from mother trees of known genotype, and paternity estimated by comparing inferred paternal genotypes with those of all potential fathers in the analysed population (e.g. Schnabel and Hamrick, 1995; Dow and Ashley, 1998; Kaufman *et al.*, 1998). Some researchers have also sampled seedlings and/or saplings, attempting to estimate both pollen- and seed-mediated gene flow (Dow and Ashley, 1996; Isagi *et al.*, 2000; Konuma *et al.*, 2000). However, such analyses require considerably more exclusion power than paternity analysis (Marshall *et al.*, 1998).

Results from paternity analyses in wind-pollinated species generally agree with predictions from studies of genetic structure (Table 8.3). In most cases, the frequency of immigrant pollinations was over 30% (e.g. 31% in *Pinus densiflora* (Lian *et al.*, 2001); 57% in *Quercus macrocarpa* (Dow and Ashley, 1998); 65% in *Q. robur* and 69% in *Q. petraea* (Streiff *et al.*, 1999)), but appeared rather low (6.5%) in a spatially isolated population of *Pinus flexilis* (Schuster and Mitton, 2000). Remarkably, Kaufman *et al.* (1998) reported extensive pollen-mediated gene flow (a minimum of 37%) in a population of the tropical pioneer *Cecropia obtusifolia*, even though the closest population of the same species was at least 1 km away. Kaufman *et al.* (1998) suggested that successful pollen travelled as far as 10 km. Similarly, DiFazio *et al.* (unpublished results) observed extensive pollen immigration into stands of *P. trichocarpa* that were isolated by up to 16 km from the nearest ungenotyped pollen source (Table 8.4).

As with indirect methods, gene flow estimates based on parentage analysis in animal-pollinated tree species were somewhat lower, but generally similar to immigration rates in wind-pollinated trees (e.g. 17–30% in *Gleditsia triacanthos* (Schnabel and Hamrick, 1995); 20–30% in *Rhododendron metternichii* (Kameyama *et al.*, 2000); 36–68% in *Swietenia humilis* (White *et al.*, 2002); and 74% in *Magnolia obovata* (Isagi *et al.*, 2000)). These results support the notion that animals can be effective agents of long-distance pollen and seed dispersal. However, there is also considerable variance in pollen immigration between species, even when isolation distances are similar (Stacy *et al.*, 1996), cautioning against broad generalizations.

Although paternity analysis based on highly variable markers appears to be the most effective current method for measuring gene dispersal in ecological time, estimates can be greatly affected by the presence of null alleles, and the misgenotyping of complex microsatellite and AFLP phenotypes. Both of these types of errors will cause overestimates of gene flow (Marshall *et al.*, 1998). Despite high reported exclusion probabilities, it is

**Table 8.3.** Mean pollination distance and immigration estimates from parentage analyses in wind- and animal-pollinated trees.

Tree species	<i>S</i>	<i>D</i> <sub>i</sub>	<i>d</i> <sub>wp</sub>	<i>d</i> <sub>op</sub>	<i>m</i>	Reference
<b>Wind-pollinated</b>						
<i>Pinus flexilis</i>	–	> 2000	133–140	155–265	0.07	Schuster and Mitton (2000)
<i>Pinus densiflora</i>	9.1	–	68	–	0.31	Lian <i>et al.</i> (2001)
<i>Pinus pinaster</i>	0.8	–	–	–	0.15	González-Martínez <i>et al.</i> (2002)
<i>Quercus macrocarpa</i>	5	> 100	75	–	0.57	Dow and Ashley (1998)
<i>Quercus robur</i>	5.8	> 100	22–58	333	0.65	Streiff <i>et al.</i> (1999)
<i>Quercus petraea</i>	5.8	> 100	18–65	287	0.69	Streiff <i>et al.</i> (1999)
<i>Cecropia obtusifolia</i>	8.6	> 1000	–	–	0.37	Kaufman <i>et al.</i> (1998)
<b>Animal-pollinated</b>						
<i>Gleditsia triacanthos</i>	3	> 85	–	–	0.17–0.30	Schnabel and Hamrick (1995)
<i>Ficus</i> (from three different species)	–	> 1000	–	–	> 0.90	Hamrick and Nason (2000)
<i>Rhododendron metternichii</i>	1	> 50	–	–	0.20–0.30	Kameyama <i>et al.</i> (2000)
<i>Magnolia obovata</i>	69	–	131	–	0.74 <sup>a</sup>	Isagi <i>et al.</i> (2000)
<i>Neobalanocarpus heimii</i>	42	–	188–196	–	0.21–0.69 <sup>a</sup>	Konuma <i>et al.</i> (2000)
<i>Enterolobium cyclocarpum</i>	9.8	300	–	–	0.69–0.74	Apsit <i>et al.</i> (2001)
<i>Swietenia humilis</i>	–	> 1000	–	3100	0.47	White <i>et al.</i> (2002) <sup>b</sup>
<i>Eucalyptus regnans</i> <sup>c</sup>	0.5	40	16–27	–	0.49–0.51	Burczyk <i>et al.</i> (2002)

*S*, area in which all potential parents have been genotyped (ha); *D*<sub>i</sub>, distance to the nearest population (tree) of the same species (m); *d*<sub>wp</sub>, average pollination distance within the reference stand (m); *d*<sub>op</sub>, mean pollination distance from assumed dispersal curve (m); *m*, proportion of offspring with immigrant paternal gametes.

<sup>a</sup>Offspring having one or both parents located outside the reference stand.

<sup>b</sup>Estimates for population Butus/Jicarito.

<sup>c</sup>Seed orchard (values for neighbourhoods with radius 40 m are shown).

therefore important to treat many of the published estimates with caution; their accuracy will improve over the next several years as applications of molecular technology and statistical methods mature.

## Spatial Simulation Modelling

Extrapolation of short-term or historical gene flow observations to spatial and temporal scales that are relevant for management and ecological policy remains a major challenge (Levin, 1992; Turner *et al.*, 2001). Because of the time and expense required for a typical parentage analysis study, only a



**Table 8.4.** Population and gene flow statistics from three microsatellite-based studies of pollen dispersal in *Populus trichocarpa* in Oregon.

Site	$r^a$	Mothers <sup>b</sup>	Fathers <sup>c</sup>	$n^d$	$D_l^e$	$d_{wp}^e$	$d_{op}^e$	$P^f$	$M^g$	$G^h$
Willamette	0.25	5	221	255	100–300	138	847	89	55	44
Luckiamute	1	5	57	423	1000–1100	128	–	98	5	76
Vinson	10	28	54	712	2680–9760	1093	1270	273	197	34

From DiFazio *et al.* (unpublished results).

<sup>a</sup>Radius of sampled area (km).

<sup>b</sup>Number of trees from which seeds were collected.

<sup>c</sup>Number of reproductively mature male trees within sampled area.

<sup>d</sup>Number of progeny genotyped.

<sup>e</sup>Defined in Table 8.3.

<sup>f</sup>Number of seeds for which a single putative father was compatible within the sampled area.

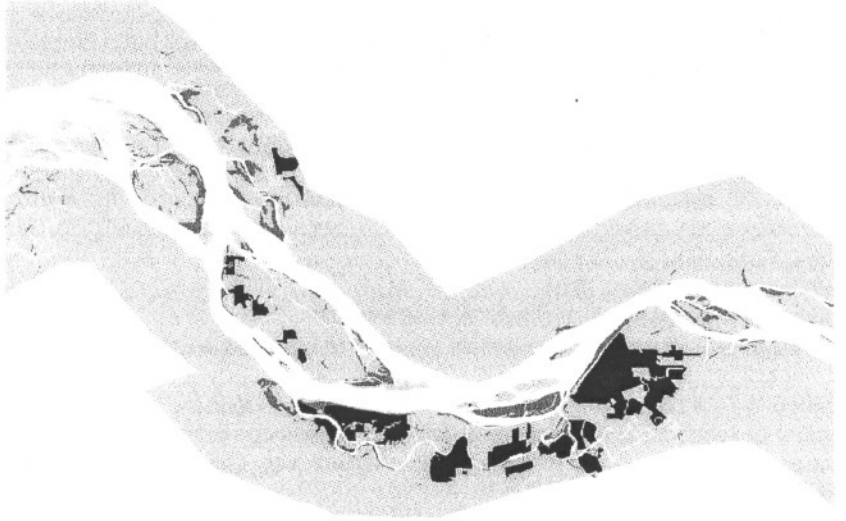
<sup>g</sup>Number of seeds for which multiple putative fathers were compatible within the sampled area.

<sup>h</sup>Percentage of seeds for which no compatible fathers were identified within the sampled area (i.e. minimum estimates of gene flow).

limited number of populations and years can be examined (Ouborg *et al.*, 1999; Cain *et al.*, 2000). However, ecologically significant levels of establishment may occur only once or twice per generation (James *et al.*, 1998), and in particular habitats. An emerging solution is the use of spatial simulation models to extrapolate results of short-term gene flow studies with knowledge of ecological processes (King, 1991; Dunning *et al.*, 1995). Such models provide an extensible framework for integrating data from disparate demographic and genetic field studies with landscape-scale analyses of ecosystem dynamics (Sork *et al.*, 1999; Higgins *et al.*, 2000). In addition, these models allow 'virtual experiments' through sensitivity analyses in which selected components of the system are manipulated to determine their importance in determining long-term outcomes (Turner *et al.*, 2001).

## Case Study: Gene Flow in Poplar

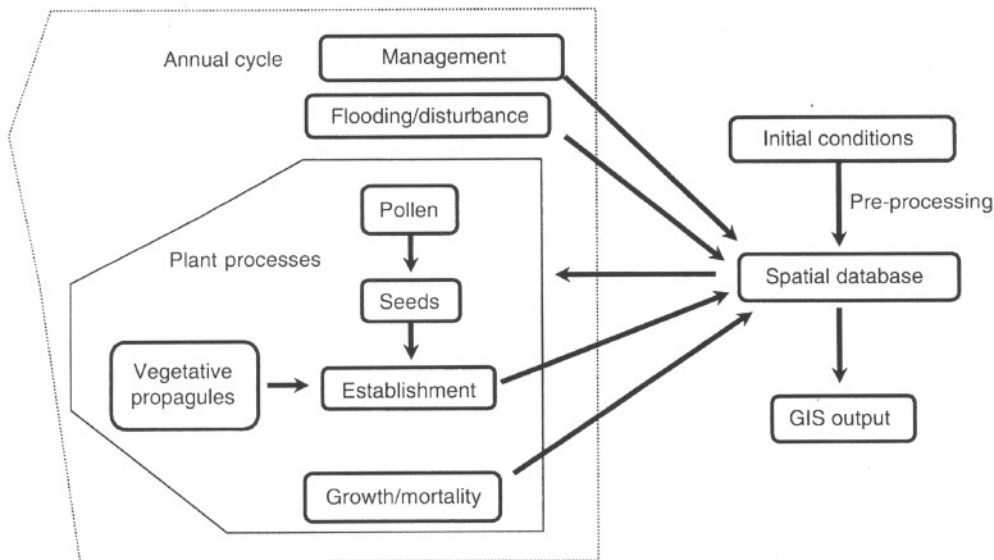
We analysed gene flow in wild black cottonwood populations, and from hybrid poplar plantations, in the northwestern USA. The primary objective of these studies was to provide data for assessing the extent of transgene dispersal that is likely to occur should transgenic hybrid poplars be cultivated in the region. We studied gene flow using parentage analysis in three wild populations with contrasting ecological characteristics (DiFazio *et al.*, unpublished results), and gathered data on seedling establishment and survival in experimental plots and in the wild (DiFazio *et al.*, 1999). We also inferred landscape-level spatial and temporal dynamics of black cottonwood establishment from a chronosequence of geographical information system (GIS) layers encompassing some of the same populations included in the field studies (Fig. 8.1).



**Fig. 8.1.** GIS representation of modelled area in northwestern Oregon (23 km × 237 km: 846 km<sup>2</sup>). White shows main channels of the Columbia River; black, areas of hybrid cottonwood plantations; dark grey, wild black cottonwood stands; light grey are non-poplar lands (mostly farms, coniferous uplands and wetlands; DiFazio, 2002).

The model, called STEVE (simulation of transgene effects in a variable environment – but also named to reflect that four Steves contributed to its development!), provides a spatially explicit representation of gene flow (DiFazio, 2002; Fig. 8.2). It operates on a landscape grid (23 km × 37 km, 100 m<sup>2</sup> cells) containing information about elevation, habitat type and poplar populations. The simulation has an annual time step, with modules to simulate creation and conversion of poplar patches, growth, reproduction, dispersal and competition within poplar cohorts. The simulations track two genotypes, transgenic and conventional. Transgenic trees originate in plantations and may spread to the wild through pollen, seed and/or vegetative propagules. The relative amounts of propagules produced in each location are proportional to basal area (i.e. trunk cross-sectional area) of each genotype, modulated by a fecundity factor.

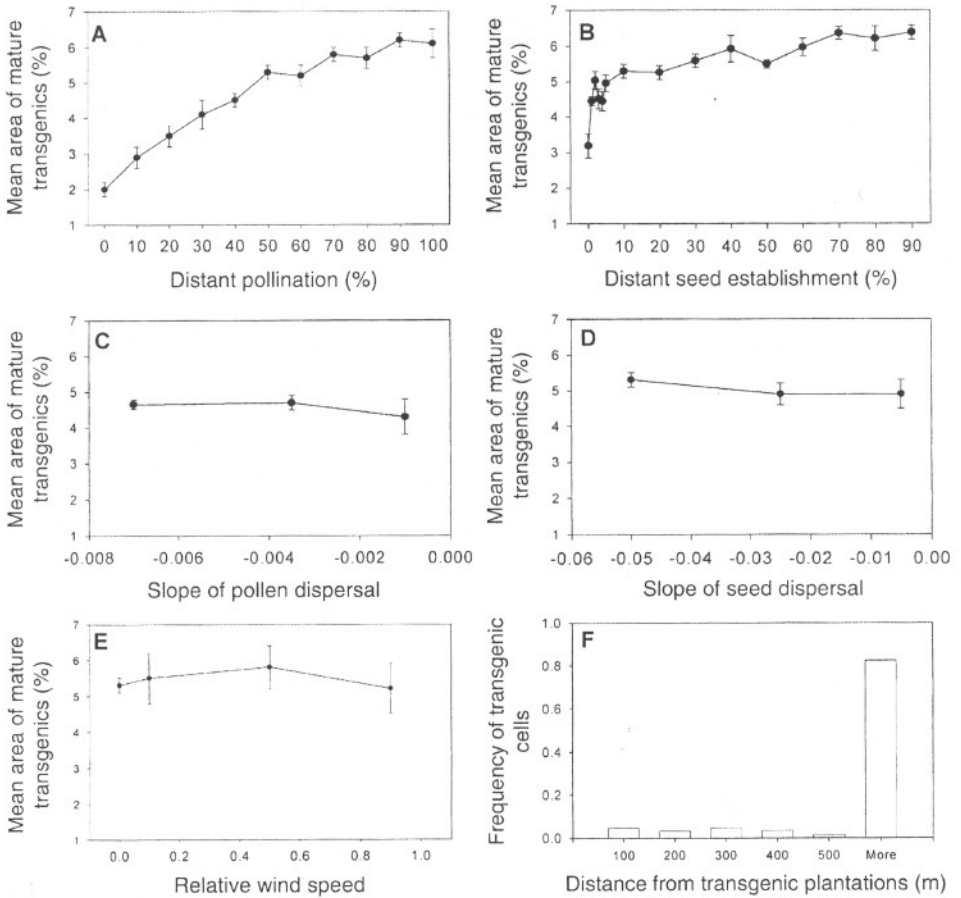
We structured and parameterized the model based on results of our field studies of gene flow. They indicated that long-distance dispersal is considerable for *Populus* (Table 8.4; DiFazio *et al.*, unpublished results), with the tail of the distribution quite 'fat' (*sensu* Kot *et al.*, 1996). We therefore chose to model gene dispersal as a two-stage process, with local dispersal modelled explicitly by a negative exponential distribution, and long-distance dispersal modelled as if a portion of the pollen and seeds were panmictic at the landscape scale. This is analogous to a mixed model approach (Clark *et al.*, 1998). Seed dispersal was modelled in the same way, though based on more limited field studies of movement, and assumed much less local and long-distance movement than for pollen. *Populus* seeds are very light and contain cotton



**Fig. 8.2.** The STEVE model. The model begins with pre-processing of GIS layers representing initial simulation conditions. Data are stored in a spatial database containing information about elevation, cover type, poplar populations, plantations and agricultural fields. Simulation begins with management activities such as plantation harvesting and herbicide spraying. Poplar establishment and mortality is simulated in the disturbance function. Seed, pollen and vegetative propagules are produced proportional to basal area of each genotype, followed by dispersal, establishment, growth and mortality. Outputs are text files and spatial data layers.

appendages that facilitate wind and water dispersal. Therefore, a portion of the seeds is expected to attain stochastic long-distance dispersal (Wright, 1952).

This method of modelling pollen and seed dispersal had major implications for gene flow from transgenic plantations. Modelled gene flow was highly sensitive to changes in the proportion of pollen and seed dispersed over long-distances (Fig. 8.3A and B), but relatively insensitive to the slope of local dispersal curves (Fig. 8.3C and D). This was primarily because poplars require very intense disturbance, abundant moisture and freedom from most competition by other plants for successful establishment. These conditions are rarely met in space and time. The majority of establishment sites therefore occurred beyond the local seed and pollen shadows of the plantations (Fig. 8.3F). Also, because long-distance dispersal was insensitive to wind in this model (pollen and seed were assumed to be panmictic at the landscape scale), wind speed had no detectable effect on gene flow from plantations (Fig. 8.3E). Long-distance dispersal ensured that a proportion of plantation-derived propagules would encounter stochastic establishment sites regardless of distance from plantations, which explains why this portion of the dispersal function was overwhelmingly important in determining gene flow. One implication of this result is that future research on gene flow in *Populus* would benefit most from better definition of the dynamics of long-distance



**Fig. 8.3.** Effects of dispersal and wind on simulated gene flow. Error bars are 1 SE from ten repetitions with each set of parameter values. (A) Effects of distant pollination on transgene flow. Distant pollination is the proportion of seeds that are fathered by trees that do not occur in the local population. This parameter has a strong effect on transgene flow, reflecting the importance of long-distance pollen dispersal. (B) Effects of distant seed establishment on transgene flow. Distant seed establishment had minor effects except at very low levels. (C and D) Effects of varying the slope of the negative exponential distributions depicting local pollen and seed dispersal, respectively. Varying this slope had little effect on gene flow. (E) Effect of relative wind speed, with wind direction set at 90°. (F) Distance of portions of transgenic cohorts (cells) from mature transgenic plantations. The local pollen and seed shadows end at 440 and 220 m, respectively.

dispersal, rather than from studies of local pollen movement and mating between trees within stands.

Sensitivity analyses allowed us to study the consequences for gene flow of many ecological conditions and transgenic deployment scenarios over a 50–100 year time frame (DiFazio, 2002). For example, we studied the consequences of:

- Transgenes that imparted herbicide resistance with respect to various scenarios of herbicide use and disturbance on the landscape;
- Transgenic trees with insect resistance, with varying levels of insect attack in wild populations;
- Reductions in fertility due to transgenic or other sources, and implications of various levels of efficiency and stability;
- The effects of transgenes with positive or negative effects under natural selection; and
- Effects of transgenic versus non-transgenic plantation area, plantation gender and rotation length (time to harvest).

Most of these simulations also included stochastic variation, so that natural environmental variances and uncertainty in parameter estimates could be reflected in model outputs. They also included a number of conservative assumptions, such that the estimates of transgene dispersal are expected to be worst-case estimates. Ideally, the model structure and parameters would be continually revised based on research results, and by data from monitoring programmes during commercial deployment. Some key results from sensitivity studies (DiFazio, 2002) were that:

- Vegetative propagation appeared to be much less important than sexual propagules in determining extent of migration under virtually all scenarios;
- An imperfect and tightly linked sterility gene could dramatically slow the spread of a gene that provided a strong selective advantage; and
- The spread of neutral genes could be greatly attenuated by sterility levels whose effectiveness (stability) was in the vicinity of 95%.

The most important contribution of spatial simulation models such as STEVE is that they provide a comprehensive, explicit logical framework for thinking about the long-term consequences of different options for deploying transgenic, as well as conventionally bred, plants. Models help to reduce the immense ecological complexity of tree gene flow to a set of specific, testable predictions that can guide further research, and inform business plans, regulatory decisions and, ultimately, public views about transgenic technology in plantation forestry.

## References

- Adams, W.T. (1992) Gene dispersal within forest tree populations. *New Forests* 6, 217–240.
- Adams, W.T. and Birkes, D.S. (1991) Estimating mating patterns in forest tree populations. In: Fineschi, S., Malvolti, M.E., Cannata, F. and Hattemer, H.H. (eds) *Biochemical Markers in the Population Genetics of Forest Trees*. S.P.B. Academic Publishing bv, The Hague, The Netherlands, pp. 157–172.
- Adams, W.T. and Burczyk, J. (2000) Magnitude and implications of gene flow in gene conservation reserves. In: Young, A., Boshier, D. and Boyle, T. (eds) *Forest*

- Conservation Genetics: Principles and Practice*. CSIRO Publishing, Collingwood, Victoria, Australia, pp. 215–224.
- Adams, W.T., Griffin, A.R. and Moran, G.F. (1992) Using paternity analysis to measure effective pollen dispersal in plant populations. *American Naturalist* 140, 762–780.
- Apsit, V.J., Hamrick, J.L. and Nason, J.D. (2001) Breeding population size of a fragmented population of a Costa Rican dry forest tree species. *Journal of Heredity* 92, 415–420.
- Austerlitz, F., Mariette, S., Machon, N., Gouyon, P.H. and Godelle, B. (2000) Effects of colonization processes on genetic diversity: differences between annual plants and tree species. *Genetics* 154, 1309–1321.
- Beerli, P. and Felsenstein, J. (1999) Maximum-likelihood estimation of migration rates and effective population numbers in two populations using a coalescent approach. *Genetics* 152, 763–773.
- Benetka, V., Mottl, J., Vacková, I., Pospíšková, M. and Dubsy, M. (1999) Estimation of the introgression level in *Populus nigra* L. populations by means of isozyme gene markers. *Silvae Genetica* 48, 218–223.
- Benetka, V., Vacková, I., Bartáková, I., Pospíšková, M. and Rasl, M. (2002) Introgression in black poplar (*Populus nigra* L. ssp. *nigra*) and its transmission. *Journal of Forest Science* 48, 115–120.
- Burczyk, J., Adams, W.T., Moran, G.F. and Griffin, A.R. (2002) Complex patterns of mating revealed in a *Eucalyptus regnans* seed orchard using allozyme markers and the neighbourhood model. *Molecular Ecology* 11, 2379–2391.
- Burdon, R.D. and Walter, C. (2001) Perspectives on risk in transgenic forest plantations in relation to conventional breeding and use of exotic pines and eucalypts. *Symposium Proceedings: Tree Biotechnology in the New Millennium*. Oregon State University, Corvallis, Oregon. Available at: <http://www.fsl.orst.edu/tgerc/iufro2001/eprocd.htm>
- Cain, M.L., Milligan, B.G. and Strand, A.E. (2000) Long-distance seed dispersal in plant populations. *American Journal of Botany* 87, 1217–1227.
- Clark, J.S., Fastie, C., Hurtt, G.S., Jackson, T., Johnson, C., King, A.G., Lewis, M., Lynch, J., Pacala, S., Prentice, C., Schupp, E.W., Webb, T. III and Wyckoff, P. (1998) Reid's paradox of rapid plant migration. *Bioscience* 48, 13–24.
- Demesure, B., Le Guerroué, B., Lucchi, G., Prat, D. and Petit, R.J. (2000) Genetic variability of a scattered temperate forest tree: *Sorbus torminalis* L. (Crantz). *Annals of Forest Science* 57, 63–71.
- Devlin, B. and Ellstrand, N.C. (1990) The development and application of a refined method for estimating gene flow from angiosperm paternity analysis. *Evolution* 44, 248–259.
- Devlin, B., Roeder, K. and Ellstrand, N.C. (1988) Fractional paternity assignment: theoretical development and comparison to other methods. *Theoretical and Applied Genetics* 76, 369–380.
- DiFazio, S.P. (2002) Measuring and Modeling Gene Flow from Hybrid Poplar Plantations: Implications for Transgenic Risk Assessment. PhD Dissertation, Oregon State University, Corvallis, Oregon. Available at: [http://www.fsl.orst.edu/tgerc/dif\\_thesis/difaz\\_thesis.pdf](http://www.fsl.orst.edu/tgerc/dif_thesis/difaz_thesis.pdf)
- DiFazio, S.P., Leonardi, S., Cheng, S. and Strauss, S.H. (1999) Assessing potential risks of transgene escape from fiber plantations. In: Lutman, P.W. (ed.) *Gene Flow and Agriculture: Relevance for Transgenic Crops*. Symposium Proceedings No. 72. British Crop Protection Council, Farnham, UK, pp. 171–176.

- Di-Giovanni, F., Kevan, P.G. and Arnold, J. (1996) Lower planetary boundary layer profiles of atmospheric conifer pollen above a seed orchard in northern Ontario, Canada. *Forest Ecology and Management* 83, 87–97.
- Dow, B.D. and Ashley, M.V. (1996) Microsatellite analysis of seed dispersal and parentage of saplings of bur oak, *Quercus macrocarpa*. *Molecular Ecology* 5, 61–627.
- Dow, B.D. and Ashley, M.V. (1998) High levels of gene flow in Bur oak revealed by paternity analysis using microsatellites. *Journal of Heredity* 89, 62–70.
- Dunning, J.B., Stewart, D.J., Danielson, B.J., Noon, B.R., Root, T.L., Lamberson, R.H. and Stevens, E.E. (1995) Spatially explicit population models: current forms and future uses. *Ecological Applications* 5, 3–11.
- El-Kassaby, Y.A. (1991) Genetic variation within and among conifer populations: review and evaluation of methods. In: Fineschi, S., Malvolti, M.E., Cannata, F. and Hattemer, H.H. (eds) *Biochemical Markers in the Population Genetics of Forest Trees*. S.P.B. Academic Publishing bv, The Hague, The Netherlands, pp. 61–76.
- El-Kassaby, Y.A. and Ritland, K. (1986) Low levels of pollen contamination in a Douglas-fir seed orchard as detected by allozyme markers. *Silvae Genetica* 35, 224–229.
- Ellstrand, N.C. (2001) When transgenes wander, should we worry? *Plant Physiology* 125, 1543–1545.
- Gerber, S., Mariette, S., Streiff, R., Bodénès, C. and Kremer, A. (2000) Comparison of microsatellites and amplified fragment length polymorphism markers for parentage analysis. *Molecular Ecology* 9, 1037–1048.
- González-Martínez, S.C., Gerber, S., Cervera, M.T., Martínez-Zapater, J.M., Gil, L. and Alia, R. (2002) Seed gene flow and fine-scale structure in a Mediterranean pine (*Pinus pinaster* Ait.) using nuclear microsatellite markers. *Theoretical and Applied Genetics* 104, 1290–1297.
- Goto, S., Miyahara, F. and Ide, Y. (2002) Monitoring male reproductive success in a Japanese black pine clonal seed orchard with RAPD markers. *Canadian Journal of Forest Research* 32, 983–988.
- Govindaraju, D.R. (1988) Life histories, neighborhood sizes, and variance structure in some North American conifers. *Biological Journal of the Linnean Society* 35, 69–78.
- Govindaraju, D.R. (1989) Estimates of gene flow in forest trees. *Biological Journal of the Linnean Society* 37, 345–357.
- Hamrick, J.L. and Nason, J.D. (2000) Gene flow in forest trees. In: Young, A., Boshier, D. and Boyle, T. (eds) *Forest Conservation Genetics: Principles and Practice*. CSIRO Publishing, Collingwood, Victoria, Australia, pp. 81–90.
- Hamrick, J.L., Godt, M.J.W. and Sherman-Broyles, S.L. (1992) Factors influencing levels of genetic diversity in woody plant species. *New Forests* 6, 95–124.
- Heinze, B. and Lickl, E. (2002) Rare, but steady, introgression in Austrian black poplar as a long-term risk? In: Van Dam, B. and Bordács, S. (eds) *Genetic Diversity in River Populations of European Black Poplar – Implications for Riparian Eco-system Management*. Proceedings of an international symposium. Csiszár Nyomda, Budapest, Hungary, Symposium held 16–20 May 2001, Szekszárd, pp. 169–175.
- Higgins, S.I., Richardson, D.M. and Cowling, R.M. (2000) Using a dynamic landscape model for planning the management of alien plant invasions. *Ecological Applications* 10, 1833–1848.
- Isagi, Y., Kanazashi, T., Suzuki, W., Tanaka, H. and Abe, T. (2000) Microsatellite analysis of the regeneration process of *Magnolia obovata* Thunb. *Heredity* 84, 143–151.

- James, R.R., DiFazio, S.P., Brunner, A.M. and Strauss, S.H. (1998) Environmental effects of genetically engineered woody biomass crops. *Biomass and Bioenergy* 14, 403–414.
- Kameyama, Y., Isagi, Y., Naito, K. and Nakagoshi, N. (2000) Microsatellite analysis of pollen flow in *Rhododendron metternichii* var. *hondoense*. *Ecological Research* 15, 263–269.
- Kaufman, S.R., Smouse, P.E. and Alvarez-Buylla, E.R. (1998) Pollen-mediated gene flow and differential male reproductive success in a tropical pioneer tree, *Cecropia obtusifolia* Bertol. (Moraceae): a paternity analysis. *Heredity* 81, 164–173.
- King, A.W. (1991) Translating models across scales in the landscape. In: Turner, M.G. and Gardner, R. (eds) *Quantitative Methods in Landscape Ecology*. Springer-Verlag, New York, pp. 479–517.
- Konuma, A., Tsumura, Y., Lee, C.T., Lee, S.L. and Okuda, T. (2000) Estimation of gene flow in the tropical-rainforest tree *Neobalanocarpus heimii* (Dipterocarpaceae), inferred from paternity analysis. *Molecular Ecology* 9, 1843–1852.
- Kot, M., Lewis, M. and van den Driessche, P. (1996) Dispersal data and the spread of invading organisms. *Ecology* 77, 2027–2042.
- Kremer, A. (1994) Diversité génétique et variabilité des caractères phénotypiques chez les arbres forestiers. *Génétique, Sélection, Évolution* 26 (Supplement 1), 105–123.
- Lanner, R.M. (1965) Needed: a new approach to the study of pollen dispersion. *Silvae Genetica* 15, 50–52.
- Ledgard, N. (2001) The spread of lodgepole pine (*Pinus contorta*, Dougl.) in New Zealand. *Forest Ecology and Management* 141, 43–57.
- Lefèvre, F., Kajba, D., Heinze, B., Rotach, P., de Vries, S.M.G. and Turok, J. (2001) Black poplar: a model for gene resource conservation in forest ecosystems. *Forestry Chronicle* 77, 239–244.
- Legionnet, A. and Lefèvre, F. (1996) Genetic variation of the riparian pioneer tree species *Populus nigra* L. I. Study of population structure based on isozymes. *Heredity* 77, 629–637.
- Levin, D.A. and Kerster, H.W. (1974) Gene flow in seed plants. In: Dobzhansky, T., Hecht, M.T. and Steere, W.T. (eds) *Evolutionary Biology* 7. Plenum Press, New York, pp. 139–220.
- Levin, S.A. (1992) The problem of pattern and scale in ecology. *Ecology* 73, 1943–1967.
- Lian, C., Miwa, M. and Hogetsu, T. (2001) Outcrossing and paternity analysis of *Pinus densiflora* (Japanese red pine) by microsatellite polymorphism. *Heredity* 87, 88–98.
- Loveless, M.D. and Hamrick, J.L. (1984) Ecological determinants of genetic structure in plant populations. *Annual Review of Ecology and Systematics* 15, 65–95.
- Marshall, T.C., Slate, J., Kruuk, L.E.B. and Pemberton, J.M. (1998) Statistical confidence for likelihood-based paternity inference in natural populations. *Molecular Ecology* 7, 639–655.
- Meagher, T.R. (1986) Analysis of paternity within a natural population of *Chamaelirium luteum*. 1. Identification of most-likely male parents. *American Naturalist* 128, 199–215.
- Moran, G.F. (1992) Patterns of genetic diversity in Australian tree species. *New Forests* 6, 49–66.
- Müller-Starck, G., Baradat, P. and Bergmann, F. (1992) Genetic variation within European tree species. *New Forests* 6, 23–47.
- Nei, M. (1973) Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences USA* 70, 3321–3323.



- Neigel, J.E. (1997) A comparison of alternative strategies for estimating gene flow from genetic markers. *Annual Review of Ecology and Systematics* 28, 105–128.
- Ouborg, N.J., Piquot, Y. and Van Groenendael, J.M. (1999) Population genetics, molecular markers, and the study of dispersal in plants. *Journal of Ecology* 87, 551–568.
- Pakkanen, A., Nikkanen, T. and Pulkkinen, P. (2000) Annual variation in pollen contamination and outcrossing in a *Picea abies* seed orchard. *Scandinavian Journal of Forest Research* 15, 399–404.
- Plomion, C., LeProvost, G., Pot, D., Vendramin, G., Gerber, S., Decroocq, S., Brach, J., Raffin, A. and Pastuszka, P. (2001) Pollen contamination in a maritime pine polycross seed orchard and certification of improved seeds using chloroplast microsatellites. *Canadian Journal of Forest Research* 31, 1816–1825.
- Richardson, D.M. (1998) Forestry trees as invasive aliens. *Conservation Biology* 12, 18–26.
- Roeder, K., Devlin, B. and Lindsay, B.G. (1989) Application of maximum likelihood methods to population genetic data for the estimation of individual fertilities. *Biometrics* 45, 363–379.
- Schnabel, A. and Hamrick, J.L. (1995) Understanding the population genetic structure of *Gleditsia triacanthos* L.: the scale and pattern of pollen gene flow. *Evolution* 49, 921–931.
- Schuster, W.S.F. and Mitton, J.B. (2000) Paternity and gene dispersal in limber pine (*Pinus flexilis* James). *Heredity* 84, 348–361.
- Slatkin, M. (1985) Rare alleles as indicators of gene flow. *Evolution* 39, 53–65.
- Smith, D.B. and Adams, W.T. (1983) Measuring pollen contamination in clonal seed orchards with the aid of genetic markers. In: *Proceedings of the 17th Southern Forest Tree Improvement Conference*, Athens, Georgia, pp. 64–73.
- Smouse, P.E. and Meagher, T.R. (1994) Genetic analysis of male reproductive contributions in *Chamaelirium luteum* (L.) Gray (Liliaceae). *Genetics* 136, 313–322.
- Sork, V.L., Nason, J., Campbell, D.R. and Fernandez, J.F. (1999) Landscape approaches to contemporary gene flow in plants. *Trends in Ecology and Evolution* 14, 219–224.
- Stacy, E.A., Hamrick, J.L., Nason, J.D., Hubbell, S.P., Foster, R.B. and Condit, R. (1996) Pollen dispersal in low-density populations of three neotropical tree species. *American Naturalist* 148, 275–298.
- Stoehr, M.U. and Newton, C.R. (2002) Evaluation of mating dynamics in a lodgepole pine seed orchard using chloroplast DNA markers. *Canadian Journal of Forest Research* 32, 469–476.
- Strauss, S.H. (1999) Report of the poplar working group. In: Traynor, P.L. and Westwood, J. (eds) *Ecological Effects of Pest Resistance Genes in Managed Ecosystems*. Virginia Polytechnic Institute, Blacksburg, Virginia, pp. 105–112.
- Strauss, S.H., Rottmann, W.H., Brunner, A.M. and Sheppard, L.A. (1995) Genetic engineering of reproductive sterility in forest trees. *Molecular Breeding* 1, 5–26.
- Streiff, R., Ducouso, A., Lexer, C., Steinkellner, H., Gloessl, J. and Kremer, A. (1999) Pollen dispersal inferred from paternity analysis in a mixed stand of *Quercus robur* L. and *Q. petraea* (Matt.) Liebl. *Molecular Ecology* 8, 831–841.
- Thompson, P.B. (2001) The ethics of molecular silviculture. In: Strauss, S.H. and Bradshaw, H.D. Jr (eds) *Proceedings of the First International Symposium on Ecological and Societal Aspects of Transgenic Plantations*. College of Forestry, Oregon State University, Corvallis, Oregon, pp. 85–91. Available at: <http://www.fsl.orst.edu/tgerc/iufro2001/eprocd.pdf>

- Turner, M.G., Gardner, R. and O'Neill, P. (2001) *Landscape Ecology in Theory and Practice: Pattern and Process*. Springer-Verlag, New York.
- White, G.M., Boshier, D.H. and Powell, W. (2002) Increased pollen flow counteracts fragmentation in a tropical dry forest: an example from *Swietenia humilis* Zuccarini. *Proceedings of the National Academy of Sciences USA* 99, 2038–2042.
- Wright, J.W. (1952) *Pollen Dispersion of Some Forest Trees*. Northeastern Forest Experiment Station Paper 46.
- Wright, S. (1931) Evolution in Mendelian populations. *Genetics* 16, 97–159.