Gene flow and simulation of transgene dispersal from hybrid poplar plantations

Stephen P. DiFazio[1], Stefano Leonardi[2], Gancho T. Slavov[1,3,4], Steven L. Garman[5,6], W. Thomas Adams[6] and Steven H. Strauss[6]

[1]Department of Biology, West Virginia University, Morgantown, WV 26506-6057, USA; [2]Dipartimento di Scienze Ambientali, Università di Parma, 43100 Parma, Italy; [3]Department of Dendrology, University of Forestry, Sofia 1756, Bulgaria; [4]Institute of Biological, Environmental and Rural Sciences, Aberystwyth University, Aberystwyth SY23 3EB, UK; [5]National Park Service, PO Box 848, Moab, UT 84532, USA; [6]Department of Forest Ecosystems and Society, Oregon State University, 3180 SW Jefferson Way, Corvallis, OR 97331, USA

Summary

Author for correspondence:
Steven H. Strauss
Tel: +1 541 737 6578
Email: steve.strauss@oregonstate.edu

- Gene flow is a primary determinant of potential ecological impacts of transgenic trees. However, gene flow is a complex process that must be assessed in the context of realistic genetic, management, and environmental conditions.
- We measured gene flow from hybrid poplar plantations using morphological and genetic markers, and developed a spatially explicit landscape model to simulate pollination, dispersal, establishment, and mortality in the context of historical and projected disturbance and land-use regimes.
- Most pollination and seed establishment occurred within 450 m of the source, with a very long tail. Modeled transgene flow was highly context-dependent, strongly influenced by the competitive effects of transgenes, transgenic fertility, plantation rotation length, disturbance regime, and spatial and temporal variation in selection. The use of linked infertility genes even if imperfect, substantially reduced transgene flow in a wide range of modeled scenarios. The significance of seed and vegetative dispersal was highly dependent on plantation size.
- Our empirical and modeling studies suggest that transgene spread can be spatially extensive. However, the amount of spread is highly dependent on ecological and management context, and can be greatly limited or prevented by management or mitigation genes such as those that cause sexual infertility.

Introduction

There are substantial concerns about the spread of transgenic plants into wild and feral plant communities. These concerns are heightened for perennial species such as trees and grasses that have undergone little domestication and that provide extensive ecological services (James *et al.*, 1998; Ellison *et al.*, 2005; Hoenicka & Fladung, 2006; Reichman *et al.*, 2006). A first step in assessing the potential ecological effects of transgenes beyond areas of cultivation is the estimation of potential transgene dispersal (Wilkinson *et al.*, 2003b; Hails & Morley, 2005; Chandler & Dunwell, 2008). However, such studies present substantial technical and theoretical challenges because forest trees are at an early stage of domestication, and have extensively outcrossing mating systems, long life cycles, and long-distance dispersal of pollen and/or seeds (James *et al.*, 1998; Williams & Davis, 2005; Williams, 2010). Transgene spread must therefore be considered over very large spatial scales and long timeframes, precluding conventional experimental approaches in contained environments (James *et al.*, 1998; Kuparinen & Schurr, 2008).

Furthermore, regulatory restrictions would prevent direct study of transgene dispersal in most species of trees and environments before their full deregulation, further complicating accurate assessment of transgenic risks (Strauss *et al.*, 2010).

Various methods have been proposed to constrain transgene flow, but an integrated framework is needed to assess the potential utility of these approaches. For example, the release of pollen and/or seeds can be prevented through incorporation of a sterility transgene in the transformation construct (Brunner *et al.*, 2007; Kwit *et al.*, 2011). Alternatively, in cases where flowering is needed for commercial or ecological purposes, it has been proposed that incorporation of other types of 'mitigation transgenes' could be an effective containment strategy. For example, genes that produce reduced stature and therefore provide high productivity in a plantation setting but low competitiveness in the wild have been proposed as mitigation transgenes (Gressel, 1999; Al Ahmad *et al.*, 2006). Another approach where flowering is permitted is to use recombinases to excise transgenes from genomes during gamete development, but before pollen or seed release (Moon *et al.*, 2009). Chloroplast transformation would

effectively prevent pollen flow in most angiosperms, but allow spread by seed (Daniell, 2002). However, this approach has come under criticism as providing incomplete containment in some scenarios due to transmission by seed (Stewart & Prakash, 1998; Allainguillaume *et al.*, 2009) or incomplete maternal inheritance (Svab & Maliga, 2007). Finally, the most common way of controlling transgene spread is through management practices, including harvesting on short rotations before flowering begins, and/or the use of buffer zones where pollination and seed establishment are prevented, thereby mitigating spread outside the confines of plantations. Such strategies have been called into question because of the incomplete containment they provide in empirical (Watrud *et al.*, 2004; Reichman *et al.*, 2006) and simulation studies (Kuparinen & Schurr, 2008).

Modeling has played a key role in assessments of the efficacy of containment strategies and, more generally, in attempts to assess potential impacts of transgene dispersal at a landscape or regional scale. For example, potential gene flow from *Brassica napus* fields has been assessed at a regional scale in the United Kingdom using simulation models that incorporate mechanistic representations of pollen and seed dispersal, informed by empirical data on hybridization and establishment and field surveys of wild relatives (Wilkinson *et al.*, 2003a). Similarly, potential dispersal of transgenic loblolly pine seeds has been inferred from mechanistic modeling of seed dispersal based on seed terminal velocity and turbulent flow (Williams *et al.*, 2006), and a deterministic population genetics model parameterized with demographic and silvicultural data (Williams & Davis, 2005). Alternatively, transgene introgression has been modeled as a metapopulation following an island model, with underlying competitive dynamics represented by a standard selection model (Meirmans *et al.*, 2009). While valuable insights have been gained from such studies, inferences are limited due to the lack of spatial and ecological context (Sears *et al.*, 2001; Wilkinson *et al.*, 2003a; Allainguillaume *et al.*, 2009).

A spatially explicit modeling approach that incorporates landscape-scale analyses of ecosystem dynamics, empirical measures of gene flow, land management regimes, and specific fitness effects of transgenes enables 'virtual experiments' to help identify the critical factors determining levels of potential transgene spread. For example, the AMELIE model combines mechanistic models of seed and pollen dispersal with phenomenological models of plant establishment and competition in a spatially explicit context (Kuparinen & Schurr, 2007). However, even this model lacks essential details about management practices and landscape composition that are necessary for evaluating realistic scenarios of transgenic crop cultivation. Here we report a model that fulfills most of the above criteria, enabling the testing of hypotheses important to ecological assessments and regulatory decisions about transgenic crops. Our specific goals in this work were to: directly measure levels of gene flow and establishment from poplar plantations into adjacent wild populations; assess the relative importance of biological and management factors in determining the extent of transgene spread in a simulation modeling framework; and assess how management might mitigate or promote spread, and whether linked infertility genes can be effective in attenuating spread.

We report on gene flow from hybrid poplar plantations into wild populations of black cottonwood (*Populus trichocarpa*) in the Pacific Northwest region of the USA. Hybrid poplar plantations have been cultivated in the study region for > 20 yr in close proximity to stands of wild *P. trichocarpa* (Fig. 1), enabling us to gather empirical data from conventional plantations as a baseline for calibrating our models of transgene dispersal from possible future transgenic plantations. We found that potential transgene spread can be spatially extensive. However, the large majority of dispersal is local and can be substantially attenuated or prevented by management practices such as early harvest, coppicing, or the use of linked sterility or other mitigation genes.

Materials and Methods

Choice of study species

We selected poplar as a model for studying potential transgene spread because it is the preeminent model species for forest biotechnology (Brunner *et al.*, 2004), it is a potential biofuel crop (Perlack *et al.*, 2005; Hinchee *et al.*, 2009), and its long-standing cultivation in the Pacific Northwest USA allowed us to study a realistic model ecosystem in which transgene deployment might occur. Transgenic poplars have been used in numerous regulated field trials in the USA and elsewhere (Strauss *et al.*, 2004), and are currently being commercially cultivated in China (Sedjo, 2005). There is a large potential for additional transgenic applications because the poplar genome has been sequenced (Tuskan *et al.*, 2006), many genotypes are amendable to genetic transformation, and transformation appears capable of improving its high value for bioremediation (Doty *et al.* 2007) as well as a number of other traits (Boerjan, 2005). Furthermore, poplar is an excellent model for study of large-scale transgene flow because several of its biological characteristics make extensive gene flow possible (Vanden Broeck *et al.*, 2004, 2006; Meirmans *et al.*, 2010). Poplars are dioecious and wind-pollinated, and produce abundant, small seeds with cotton-like appendages that facilitate long-distance dispersal by wind and water (DeBell, 1990). Finally, wild relatives are often interfertile with cultivated clones, and extensive wild populations commonly occur in the vicinity of commercial plantations.

Study sites

Studies of plantation gene flow were performed in the vicinity of *P. trichocarpa* × *P. deltoides* (TD) hybrid poplar plantations growing in the Pacific Northwest USA. Two of the sites were located within large-scale, commercial hybrid poplar farms on the lower Columbia River (Fig. 1), which also formed the basis for the simulation modeling described below. The River Ranch site was located near the confluence of the Westport Slough and the Columbia River, and contained *c.* 110 ha of flowering hybrid poplar clones (46°8′6″N, 123°20′20″W). The Clatskanie site was located near the confluence of the Clatskanie and Columbia Rivers (46°7′35″N, 123°13′1″W) (Fig. 1).

The Willamette site was located north of Corvallis, OR (44°35′12″N, 123°11′30″W) and contained a 2.5 ha plantation
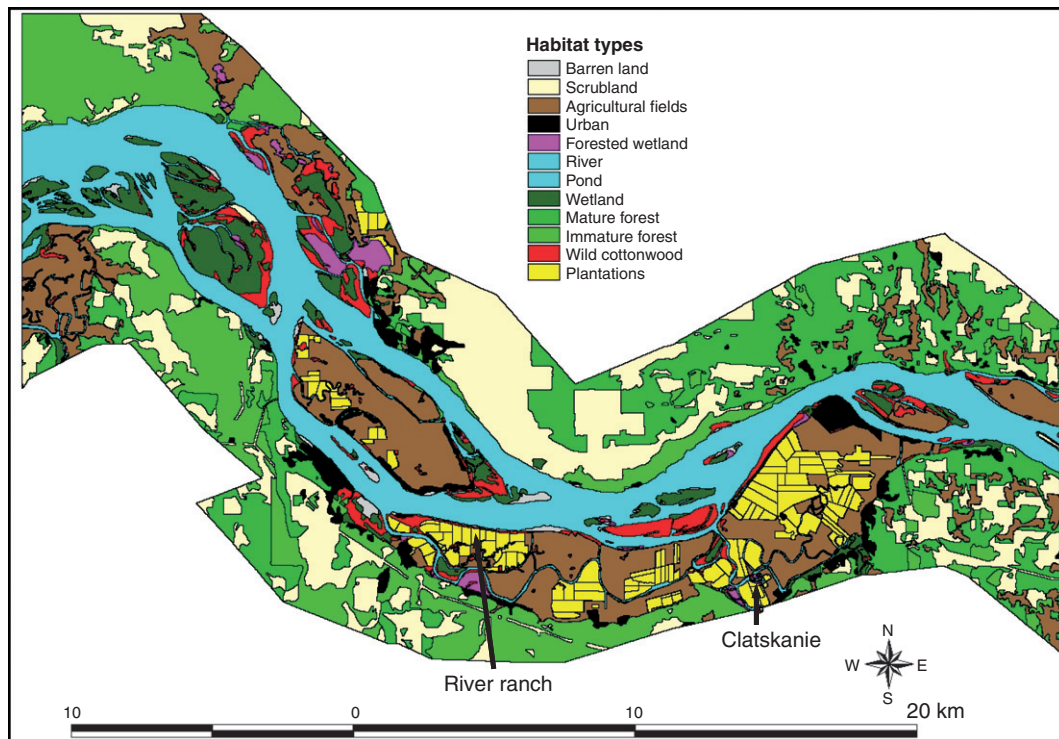
Fig. 1 Habitat types for the modeled landscape. Habitat types were delineated based on 1991 air photos and converted to a GIS layer, which was subsequently converted to grid format with 10 × 10 m cells, and then converted to a binary input file for the simulation model. Descriptions of habitat types can be found in Table S1.

that was established in 1990 to test growth of 27 hybrid poplar clones, principally diploid *P. trichocarpa* × *P. deltoides*. Adjacent to the plantation was a riparian population of *P. trichocarpa* consisting of large, mature trees, and an abandoned gravel pit with smaller trees that had apparently become established within the previous 20 yr (Slavov *et al.*, 2010).

Studies of natural establishment of hybrid seedlings were performed in additional locations where wild black cottonwood (*Populus trichocarpa* Torr. & Gray) seedlings had recently become established in close proximity to large flowering plantations of TD hybrid trees at three sites: along the Columbia River, between the Clatskanie River and the Westport Slough (46°8′48″N, 123°13′30″W) with 2350 ha of plantations (Fig. 1); on the Skagit River in Washington (48°29′41″N, 122°9′15″W) with 44.4 ha of plantations; and on the Fraser River in British Columbia (49°13′16″N, 121°55′33″W) with 190 ha of plantations.

Detection of hybrid gene flow

Establishment, survival, and growth rates of seedlings were assessed using experimental 'establishment' plots that were established in areas near plantations, where vegetation was cleared and trees were watered during dry periods. These permissive conditions were intended to emulate the rare opportunities for recruitment for a small seeded, highly shade-intolerant species in a region with long summer droughts. These establishment plots were paired with adjacent seed traps constructed from wood frames and mosquito netting. In total, 45 plots and traps were

established along four transects at each site located parallel to the plantation edges at distances ranging from 10 to 100 m. Seeds were collected weekly from traps and germinated in the lab to enable identification of hybrids. Naturally established seedlings were tagged and measured weekly within the establishment plots.

Seeds and seedlings derived from plantation hybrids were readily identified using morphological and molecular markers that were specific for *P. deltoides*, which is not native to the study area. Over 9900 seeds collected from 56 wild *P. trichocarpa* trees at the three sites and 4300 seeds from seed traps were germinated in a glasshouse, then grown in a common garden at 30 cm spacing with irrigation and weed control for 1 yr. Seedlings were morphologically analyzed for evidence of TD hybrid parentage using one of several foliage characteristics that distinguish *P. deltoides* from *P. trichocarpa*, including petiole shape, abaxial leaf color, leaf margin dentation, and leaf blade shape (Eckenwalder, 1984).

All putative hybrids plus a sample of 1650 randomly selected seedlings were subjected to molecular analysis. We first used six unlinked dominant RAPD markers (OPA2-640, OPA2-475, UBC105-570, UBC406-700, UBC413-310, UBC417-1900) and one microsatellite marker (*win3*) (Heinze, 1997) that were highly species-specific. These markers were present in all 33 *P. trichocarpa* × *P. deltoides* F_1 hybrids that were growing at the study sites, and absent in a random sample of *P. trichocarpa* trees from the same study sites ($n = 178$). These markers were presumably heterozygous in the F_1 hybrids, so the probability of correctly identifying hybrid progeny with this method was 0.99 based on independent Mendelian segregation of the markers. The presence of any single marker was sufficient for declaration

of putative TD hybrid parentage. Morphological and molecular assessments were found to be in agreement 99.9% of the time. We next used 10 microsatellite markers with an average of 17.5 alleles per locus and an average expected heterozygosity of 0.77 (Slavov *et al.*, 2009). Using these markers, we confirmed putative hybrid seedlings using parentage analysis based on modified genotypic exclusion (Slavov *et al.*, 2009).

Paternity analysis

Pollen flow was estimated for the Marchel site based on paternity analysis for five *P. trichocarpa* mother trees and 45–48 seedlings per tree. Mothers, progeny, and all male trees within the sampled area were genotyped using the 10 microsatellite loci described above. We assigned paternity using a genotypic exclusion method that accounts for cryptic gene flow using Monte Carlo simulations and allows up to three mismatching loci between offspring and putative fathers to compensate for genotyping error, mutations, and null alleles (Slavov *et al.*, 2005, 2009). We used maximum likelihood to model pollination distances by exponential, exponential power, and Weibull distributions, but these provided a poor fit to observed data, severely under-predicting long-distance dispersal. We therefore modeled pollination using a mixture model, with local pollination described by an exponential distribution, and long-distance dispersal modeled as a uniform distribution, which provided an excellent fit to the observed pollination data (Slavov *et al.*, 2009).

Simulation model

We developed a simulation model to explore the implications of transgene dispersal at the large spatial scales and long timeframes required for ecological assessment. The model, called 'Simulation of Transgene Effects in a Variable Environment' (STEVE), simulated most aspects of the gene flow process, including pollination, seed dispersal, stand establishment, and density-dependent mortality (Fig. 2). The model and underlying data are described in
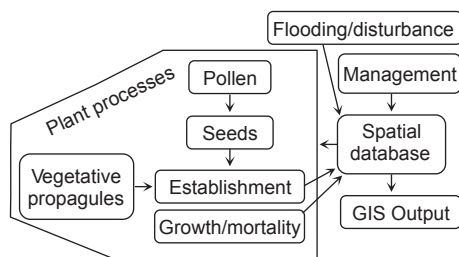

Fig. 2 The STEVE model. Model begins with preprocessing of GIS layers representing initial simulation conditions. Data are stored in a spatial database containing information about elevation, cover type, poplar populations, plantations, and agricultural fields. Simulation begins with management activities 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.

detail in Supporting Information Methods S1–S7. The model was coded in C (source code available upon request from first author) with modules representing each of the plant and management processes described below. Each iteration, representing an annual cycle, begins with reading data from a spatial database containing information about each pixel on the landscape. Variables contained in these data structures are altered by plant and management processes and written to the spatial database at the end of each cycle. Pixel data are parsed with C utilities and perl scripts and converted to GIS layers using ARC/INFO tools for visualization.

Our target landscape was a 46 631 ha area along the Columbia River, where hybrid poplar plantations are a significant component of the landscape (Fig. 1). The model started with a specified landscape containing various habitat types, including plantations of transgenic trees (Table S1). Because *P. trichocarpa* in this area has specialized ecological requirements for regeneration (Braatne *et al.*, 2007), the rate of establishment of propagules was determined by prevailing disturbance and land-use regimes. The simulations therefore included alterations of landscape composition due to management activities, ecological succession, and natural disturbance events (e.g. flooding). Transition rates were inferred from a chronosequence of air photos. Briefly, we created a GIS representation of habitat types in the study area by interpretation of air photos from the Army Corps of Engineers (Allen, 1999). This included delineation of *P. trichocarpa* populations and hybrid poplar plantations. We compared GIS layers from the years 1961, 1973, 1983, and 1991 to estimate rates of transition among habitat types and therefore rates and patterns of establishment of *P. trichocarpa* in different habitat types. Model structure was also based on information on poplar demography and silviculture from the literature, and on consultations with resource management professionals (Methods S2). The simulation model employed a spatially explicit landscape with 100 m^2 pixels where poplar seedling cohorts became established and then changed through time as competitive exclusion occurred, until a single tree occupied the cell or the habitat was altered by disturbance. This pixel size approximates the crown-width of a mature *P. trichocarpa* tree.

Poplar mating, dispersal, regeneration, growth, and competition occurred on an annual basis for 50 yr time periods. Empirical and theoretical values of model parameters, and their variances, were specified to study how they singly, or in combination, influenced transgene spread. The primary measure of transgene spread was the percentage of all pixels on the landscape containing mature (usually > 5 yr old) wild poplars that had at least one transgenic tree in the cohort (referred to as % area of mature transgenics). This measure is directly related to the total basal area of transgenics (Fig. S1), and is therefore a reasonable index of the extent of transgenic plants on the landscape.

In most runs the model assumed that one-half of the pollen was distributed locally according to an exponential pollen dispersal kernel and one-half was the result of landscape-wide pollen immigration where transgene frequencies were calculated as a proportion of all pollen produced on the landscape. This assumption was consistent with the mixture model parameters we

estimated using paternity analysis data and maximum likelihood (Slavov *et al.*, 2009). Probability of pollination was modulated by phenological overlap (Table S2) and wind direction. Seed movement was calculated similarly to pollen, except that only 1–10% of seed movement was considered to be non-local based on results from seed trap studies. Vegetative dispersal was primarily local and based on an exponential function parameterized by observations of interclonal distances in wild stands (Methods S3; Table S3).

Competitive effects of transgenes were simulated through effects on size using the equation:

$$\text{Ba}_{ga} = \alpha \frac{N_g t (1 - e^{-0.04a})}{N \max_a} \qquad \text{Eqn 1}$$

where Ba_{ga} is the basal area of genotype g (transgenic or conventional) at age a, α is the relative difference in growth of transgenic trees relative to average trees (or half the difference between transgenic and conventional trees), N_g is the number of trees of each genotype, $N\max_a$ is the carrying capacity at age a, and t is time (Methods S4).

Transgene effects were also simulated through differential effects on mortality using an approach similar to the Lotka–Volterra equation for two-species interactions (MacArthur & Levins, 1967; Shugart, 1998), except the competitive differential of one genotype is the exact opposite of that of the alternate genotype. Mortality of conventional trees was:

$$M_0 = N_0 \left(\left(\frac{N_0 + (1 + \alpha) N_1}{N \max_a} \right) - 1 \right) \qquad \text{Eqn 2}$$

and mortality of transgenics was:

$$M_1 = N_1 \left(\left(\frac{N_0 (1 - \alpha) + N_1}{N \max_a} \right) - 1 \right) \qquad \text{Eqn 3}$$

For insect resistance simulations, the value of α depended on the amount of insect herbivory that occurred in a pixel (Methods S4).

We tested the functioning of the STEVE model by parameterizing runs to represent the same sites where we assessed pollen and seed dispersal as described above. We created GIS layers representing these sites and assessed pollen and seed flow at locations approximating the distances and directions to the mother trees, seed traps, and establishment plots described above. We studied the influence of all major model parameters on transgene spread using sensitivity analyses (Methods S5; Table S4). For computational efficiency in sensitivity analyses, we created a smaller artificial landscape (2500 ha) containing the same relative proportions of the different habitat types as the full landscape. This portion of the landscape was based primarily on the River Ranch site used to generate much of the empirical data described above. We simulated two levels of cultivation of transgenic trees: a large field trial (19 ha) in an agricultural setting, and large-scale cultivation (480 ha, 50% transgenic). For factors with substantial effects in isolation, we performed a fractional factorial analysis, where each factor was investigated at two levels in a systematic sample of all factorial combinations to reveal main effects and two-factor interactions. This was a resolution V fractional factorial experiment (Box *et al.*, 1978), in which 11 main effects (Tables 1, S5) and all two-factor interactions were examined. This required 128 factor combinations, selected using the Factex procedure of SAS.

Results

Empirical observations of gene flow from plantations

Parentage analyses revealed that cultivated TD hybrid males successfully pollinated female *P. trichocarpa* trees growing in close proximity to plantations, though hybrid offspring accounted for < 0.5% of seeds analyzed (Fig. 3). Hybrid seedlings (i.e. those with at least one cultivated TD hybrid parent) grew at least as well as *P. trichocarpa* seedlings in establishment experiments adjacent to plantations, and in a common garden (Fig. 4).

Spatially explicit simulations of transgene dispersal and establishment

We assessed the functioning of the STEVE model by generating model predictions for pollen and seed flow for three different levels of transgenic fertility, 0.1-, 0.5-, and 1.0-fold compared to

Table 1 Parameters varied in fractional factorial analyses

| Parameter	Description
Vegetative dispersal	Rate parameter of the exponential distribution used to model the spread of vegetative propagules.
Vegetative establishment	Percentage of established individuals in a new cohort that are derived from vegetative propagules.
Plantation sex	Ratio of male to female plantation blocks.
Distant seed establishment	Percentage of seedlings derived from nonlocal seeds (i.e. from a seed cloud that is panmictic at the landscape scale).
Rotation	Time between plantation establishment and harvest.
Distant pollination	Percentage of seeds that are sired by nonlocal males (i.e. from a pollen cloud that is panmictic at the landscape scale).
Age of flowering plantations	Age at which plantations trees become reproductively mature.
Phenological compatibility	Percent overlap in floral phenology between transgenic and wild trees.
Transgenic fertility	Seed or pollen production of transgenic trees relative to wild trees.
Disturbance regime	Rate of disturbance. 'Obs.' is the empirical regime that results in a 15% reduction of poplar populations over 50 yr, as observed for the study area. 3× is a regime that causes a 15% increase in wild poplar over the same period.
Transgenic competitiveness	A competitive differential for transgenic trees. Reflects differences in size (basal area) between transgenic and nontransgenic trees and rates at which transgenic and nontransgenic trees die during density-dependent mortality.
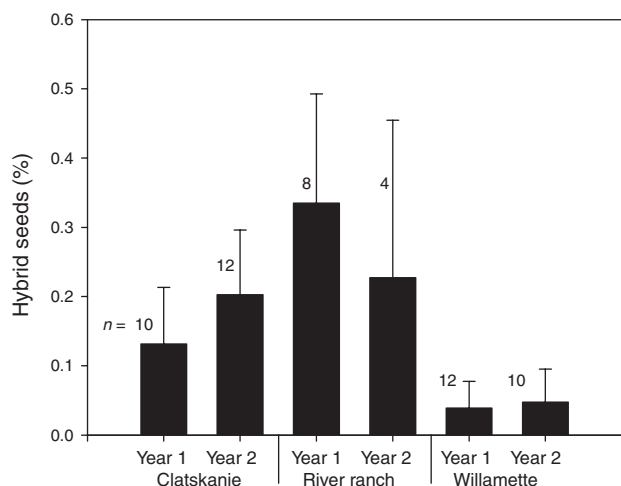
Fig. 3 Proportion of hybrid seeds from wild *Populus trichocarpa* trees near plantations. Error bars represent SE among mothers, and numbers to the left of error bars are number of mother trees sampled.
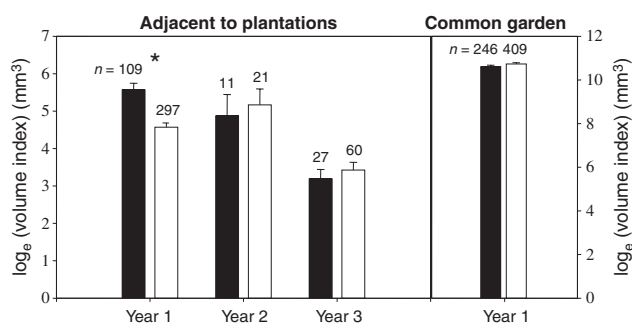

Fig. 4 Volume index (the product of height and the square of basal diameter) for 'hybrid' (TD hybrid parentage; closed bars) and 'wild' (open bars) *Populus trichocarpa* seedlings in experimentally disturbed plots within 50 m of plantations, and for seedlings derived from traps adjacent to the plots and grown in a common garden in Corvallis, OR, USA. *n* is the number of seedlings measured. Error bars represent SE among plots. *, $P < 0.05$.

non-transgenic trees (i.e. at rates that bracket levels of fertility observed in controlled crosses, Fig. S2) for the same sites where pollen and seed flow were empirically determined. These predictions were closest to fitting observed levels of pollen and seed flow at the lowest fertility levels tested, though the model showed a strong tendency to over-predict transgene flow (Figs S3, S4).

Using sensitivity analyses, we tested a wide variety of biological and managerial factors for their potential influence on transgene dispersal (Table 1). A fractional factorial sensitivity analysis revealed that fertility, competitiveness, plantation rotation length, and the disturbance regime had the most important and consistent effects on transgene spread (Fig. 5). Because the fractional factorial allows simultaneous testing of multiple parameters at a range of levels, factors that are significant in this analysis can be considered to be the most important determinants of transgene spread. Fertility was by far the single largest factor affecting transgene spread. Incomplete sterility (fertility of 0.01 relative to wild-type) gave a dramatic reduction in the extent of spread, as did various forms of unstable sterility (Brunner *et al.*, 2007). The

importance of fertility was driven in part by the wide range of parameter values tested, ranging from 1 to 100%. However, fertility was still an important factor in a fractional factorial in which all parameters were varied ± 10% from their default values, though parameters such as competitiveness, rotation length, and disturbance regime had stronger effects on transgene spread (Fig. S5). Therefore, the range of parameters tested does determine the relative importance of each variable in the sensitivity analysis. However, fertility was still the most important variable when all other factors were varied in biologically reasonable ranges.

We also compared fractional factorial sensitivity analyses for two different scenarios of transgenic poplar cultivation: an isolated 19 ha 'field trial' and a 480 ha 'commercial cultivation' scenario, where transgenic plantations occupied nearly 10% of the landscape. In our simulations, the relative effects of some variables differed markedly between these scenarios, suggesting scale-dependent effects on factors controlling transgene dispersal (Fig. 5). Transgene spread was significantly affected by parameters related to the relative degree and distance of spread of vegetative propagules, and to the extent of long-distance seed movement only for the commercial cultivation scenario, but not the field trial scenario. By contrast, increasing the proportion of pollinations that were due to long-distance rather than local pollen dispersal resulted in significantly higher transgene spread in the field trial scenario but not the commercial scenario (Fig. 5). In all cases, however, spread from the commercial planting was similar to or < 25-fold higher than from the field trial, similar to the ratio of transgenic plantation sizes between the two scenarios (480/19).

Simulations of transgene fitness and partial sterility

For simulations with stochastic variation on the full landscape, we first considered the behavior of transgenes with no effect on competitiveness (e.g. marker transgenes) by simulating their spread over the entire study landscape (46 631 ha) over a 50 yr period (Methods S6; Table S6). Stochastic variation was incorporated for all key parameters identified from the sensitivity analyses. The maximum level of abundance of neutral transgenes was *c.* 1% when relative fertility levels were equal to that of fully fertile hybrid trees (i.e. relative fertility of 0.5; Fig. 6). With the introduction of a sterility gene (relative fertility of 0.001), the abundance was *c.* 0.4% (Fig. 6), and was mostly accomplished by localized dispersal through vegetative propagules (Brunner *et al.*, 2007).

In all scenarios with neutral transgenes, there was a strong tendency for establishment sites to be near plantation sources (Fig. 7). Transgenes occurred in a diversity of sites throughout the riparian areas of the landscape due to the combined effects of long distance gene flow via pollen and seeds, and the sporadic creation of suitable habitat due to flooding and other disturbance. However, transgenic trees had mostly disappeared from these sites by age 20 due to density-dependent mortality, and mature transgenic trees were very rare at sites distant from plantations (Fig. 7).
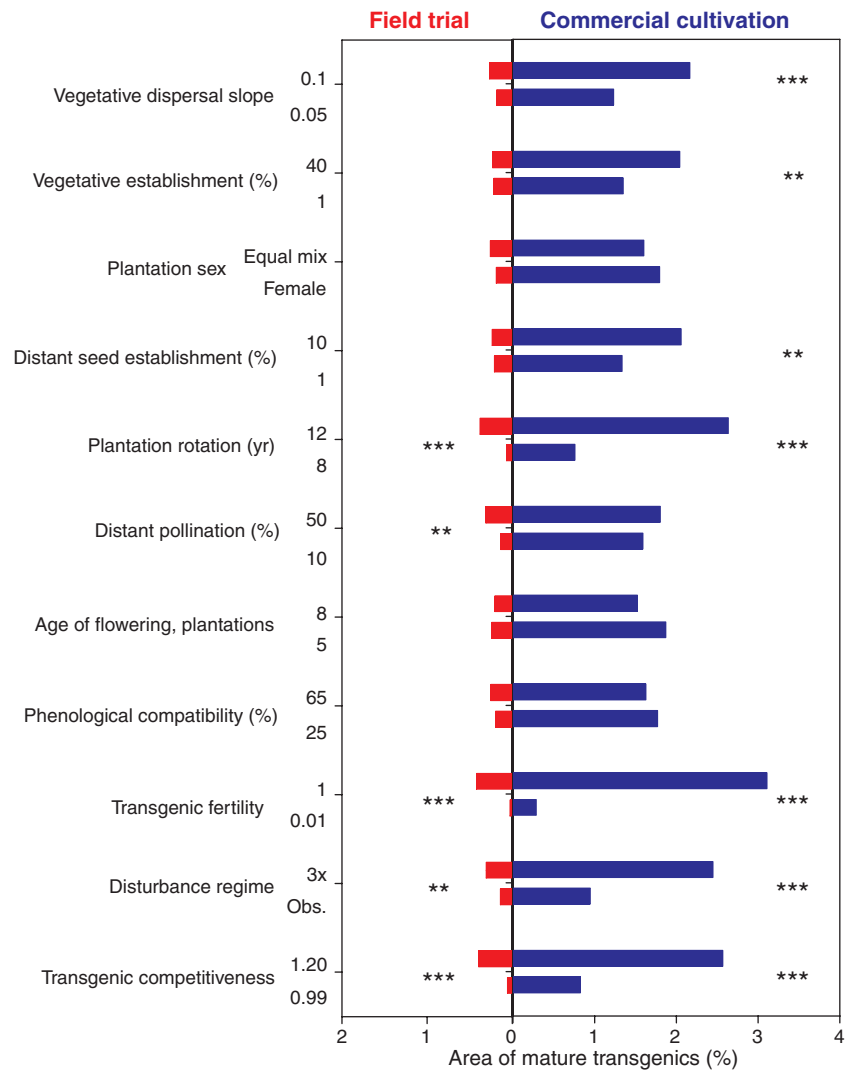
Fig. 5 Least square means for main effects from a resolution V fractional factorial experiment. Response was the percentage of all pixels with mature wild poplars that had at least one transgenic tree in the cohort. Results are for a field trial scenario and commercial cultivation over a 50 yr simulation. Variables are described in Table 1 and the Supporting Information Methods, section. **, $P < 0.01$; ***, $P < 0.001$.

We modeled transgene spread in relation to various combinations of insect herbivory (% of leaf area affected) and associated fitness benefit from having a resistance transgene when insects are active (Methods S7). Because the fitness landscape for most traits is expected to be highly environment-dependent (Raybould, 2007; Warwick et al., 2009; Beckie et al., 2010), we considered the effects of different distributions and intensities of detrimental insect activity. Substantial levels of pest pressure (> 27% of area) and transgenic competitive differential (> 30%) were required for transgenic trees to exceed 5% areal abundance at the end of the simulation period (Fig. 8). Surprisingly, when both a strong resistance transgene (competitive differential of 30%, 50% of area affected) and a linked but imperfect transgene for fertility reduction were employed (relative fertility of 0.001), the level of transgene abundance after 50 yr was virtually identical to that of a neutral gene (Fig. 6).

Discussion

We described a series of empirical and simulation studies that provided insights into the factors likely to govern the rate and extent of transgene dispersal, establishment, and introgression from hybrid poplar plantations into wild sympatric populations of *P. trichocarpa* in the Pacific Northwest USA. A full risk assessment requires the identification of endpoints of concern, enumeration of hazards contributing toward the risk of those endpoints, assessment of potential exposure to the hazards, and a delineation of options for mitigating or minimizing undesired outcomes (Graham et al., 1991; Hails & Morley, 2005; Raybould, 2007; Chandler & Dunwell, 2008). The tools and data for determining gene dispersal that we describe here – because they address the magnitude of 'exposure' and the distribution of impact – will have application to all of these phases of risk assessment.

Potential for long-distance transgene dispersal

Not surprisingly, our empirical research suggested that some level of transgene spread from fertile hybrid poplar plantations was highly likely. In some settings, transgenic plantations would likely be phenologically compatible with extensive, co-occurring wild populations. Furthermore, results from controlled crosses
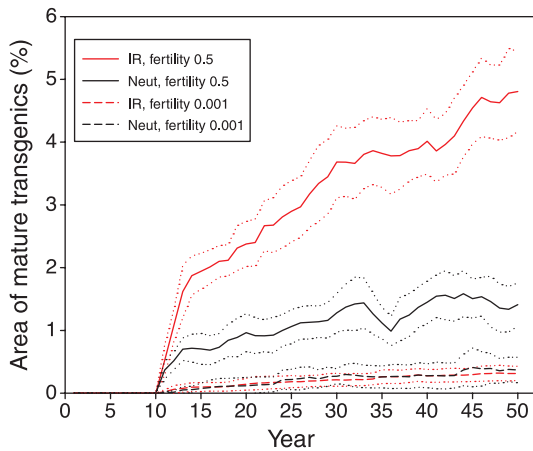
Fig. 6 Results from simulations of neutral transgene (Neut) and insect resistance (IR) transgene with a 30% fitness advantage under two levels of fertility over 50 yr. Lines represent average transgene spread by year on full landscape, with upper and lower 99% confidence intervals (dotted lines) based on 30 runs of the model.

and field experiments suggested that hybrid parentage would not inhibit establishment and competitiveness of seedlings, although these results are based on controlled conditions and may not hold in the wild. Nevertheless, transgenes have the potential to introgress in a manner similar to native genes once they become established via backcrosses (Stewart *et al.*, 2003).

Long-distance dispersal is likely to be a major determinant of the rate of transgene spread in natural populations (Clark *et al.*, 1998; Nathan *et al.*, 2002; Higgins *et al.*, 2003; Smouse *et al.*, 2007). We have also recently demonstrated that gene flow covers great distances in *P. trichocarpa*, with effective pollination distances possibly averaging as much as 7.6 km (Slavov *et al.*, 2009). Therefore, there appears to be substantial potential for long-distance transgene dispersal and rapid introgression from poplar plantations into wild populations when there is a large
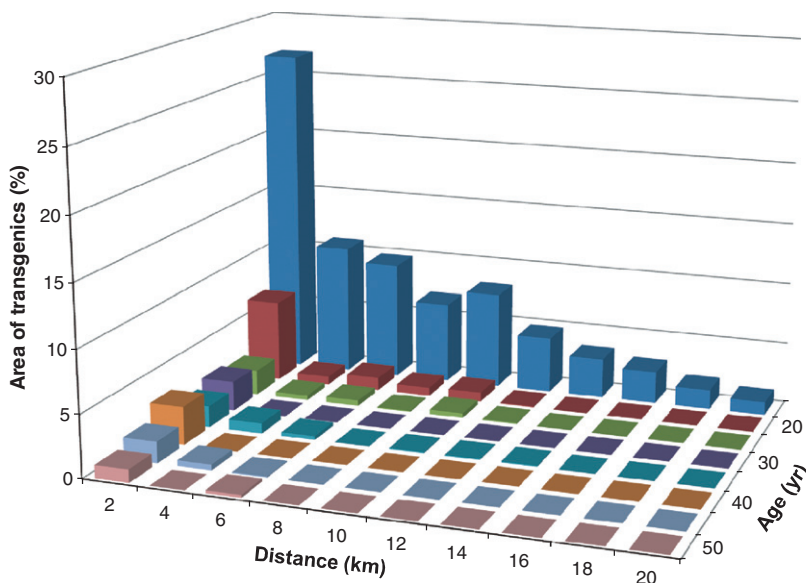
excess of transgenic vs wild pollen, or a strong selective advantage imparted. This finding is consistent with studies of dispersal and introgression of alleles from exotic cultivated poplar trees in Canada (Meirmans *et al.*, 2010; Thompson *et al.*, 2010) and Europe (Benetka *et al.*, 1999; Pošpíšková & Šálková, 2006; Smulders *et al.*, 2008). However, the level of introgression from plantations may be extremely low (e.g. far below 1% beyond only 10 m from plantations), even after many decades of plantation culture (Vanden Broeck *et al.*, 2005; Csencsics *et al.*, 2009). The actual level of introgression is likely to depend on the relative proportions of propagules, hybrid fertility, species compatibilities, and fitness of hybrid derivatives in the wild.

Importance of ecological context

Our modeling studies provided additional insight into the factors that would likely control the rate and extent of transgene spread. Our results suggest that some of the previous predictions of extensive transgene spread (e.g. Williams *et al.*, 2006) were not supported. One of our most surprising findings was that the presence of sympatric, compatible wild populations did not necessarily lead to a rapid, monotonic increase of transgene frequencies, contrary to predictions of other simulations of transgene spread (Kuparinen & Schurr, 2007; Meirmans *et al.*, 2009). For transgenes not providing a strong selective advantage, this was mostly due to chance exclusion of transgenics from cohorts of seedlings that were dominated by non-transgenic trees, particularly at sites distant from plantations. This effect was accentuated by the long juvenile period of plantation trees during which no seed or pollen is produced, the large and continuous influxes of non-transgenic pollen and seed into wild stands, and by the episodic occurrence of establishment opportunities in the vicinity of plantations. This latter effect is not captured by models lacking spatially explicit landscapes and empirically determined disturbance regimes. For all types of genes, the ability to greatly diminish both male and female fertility via hybrid breeding and genetic engineering had a

Fig. 7 Frequency of pixels containing transgenic trees as a function of tree age and minimum distance from transgenic plantations that were mature at the time of establishment. Simulations (30 iterations) were run for the same scenario as in Fig. 5, with a neutral transgene and stochastic variation in fitness. As a result of competitive exclusion, transgenic trees occurred at frequencies > 10% only by age 20 and within 2 km from their source plantation.
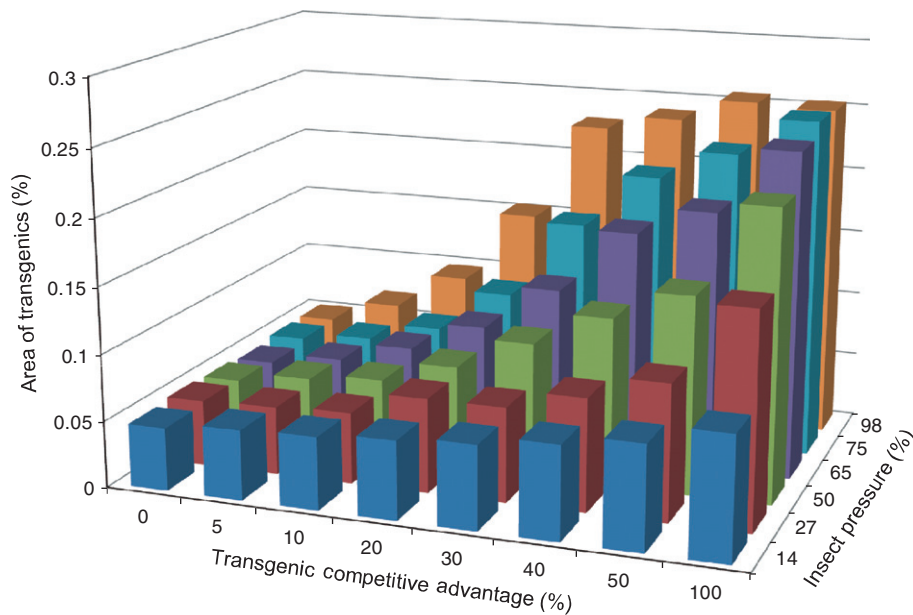
Fig. 8 Transgene flow under a variety of insect pressures (expressed as % of trees affected) and mean levels of transgenic competitiveness (expressed as the % advantage relative to conventional trees). High insect pressure in the wild (> 27% of trees affected) and transgenic competitiveness (≥ 50%) was required before transgenic advantage due to an insect resistance transgene resulted in transgene spread > 10%.

strong mitigating effect (Strauss *et al.*, 1997). Therefore, the proximity and interfertility of wild and non-transgenic populations do not necessarily provide a conduit for unmitigated transgene spread, as has often been assumed (Hancock, 2003; Chapman & Burke, 2006).

Ecological context also makes it difficult to predict transgenic fitness differentials over the long term and in wild settings due to spatial and temporal variation in selection pressures (Gould, 1998; Mason *et al.*, 2003; Warwick *et al.*, 2009). Nonetheless, there is substantial concern about potential environmental impacts of transgenes that confer a fitness benefit in the wild, especially for pest or stress tolerant transgenic trees with feral or wild relatives (van Frankenhuyzen & Beardmore, 2004; Beckie *et al.*, 2010). Because insect resistant poplars have already been produced and planted in precommercial trials (Sedjo, 2005), we explored a range of scenarios related to deployment of this trait. Rates of spread were highly dependent on the context in which the transgenics were deployed. In some scenarios, transgene spread increased monotonically throughout the duration of the simulation, suggesting that substantial establishment in wild populations could occur with high insect pressure, a strong transgenic advantage, and no mitigation transgenes. However, the level of herbivory in wild populations was nearly as important for determining spread as the direct fitness effects of the transgene in the presence of insects, and insect pressure is known to vary stochastically over time and space in natural populations, such as tent caterpillar outbreaks in aspen populations in boreal regions (Roland, 1993; Huang *et al.*, 2008). This highlights the importance of an integrated risk assessment program that combines results of lab studies and small-scale field trials with surveys of insect pressure in wild populations (Chandler & Dunwell, 2008).

Model verification

Our model consistently over-predicted transgene dispersal and establishment of transgenic organisms in the wild compared with that observed in our empirical studies. This was most likely due to assumptions built into the model that were purposely biased toward enhanced transgenic establishment. Two of the more important assumptions were a lack of pollen and seed limitation (i.e. all females had full seed production regardless of the absolute amount of pollen input, and all establishment sites received excess seeds), and panmixia for the long-distance dispersal component that occurred at the scale of the entire landscape. The low rates of observed gene flow from hybrid plantations are likely to be the result of several factors, including: a low proportion of planted compared with wild trees on the landscape, even where there has been substantial reduction of wild poplars due to agriculture, as is typical of the study area (Fig. 1); the limited number of years in which planted trees flower due to the delayed reproduction of trees (beginning at *c.* 5 yr of age in a harvest cycle of 6–12 yr); and the reduced fertility of hybrids compared with wild trees (Fig. S2). Therefore, transgenic trees with neutral or slightly deleterious combinations of transgene loci are unlikely to establish a conspicuous presence except very near plantations. In the direct vicinity of plantations, 25% of cohorts contained transgenics in the early years of establishment, before declining to *c.* 5% by 20 yr of age as a result of density-dependent mortality (Fig. 7).

Verification of the model predictions requires long-term field plantings and monitoring over many years. However, another surprising finding from our work was that small-scale field trials were not very effective for identifying factors that are important for gene flow from commercial-scale plantations (Fig. 5). This

discrepancy can be explained by the rarity of establishment sites near plantations in our system, and the small size of the field trial relative to natural populations. Both of these factors effectively precluded establishment at sites beyond the local seed and vegetative dispersal neighborhoods. By contrast, long-distance pollen dispersal was disproportionately important in the field trial scenario because these were the primary means by which transgenes were dispersed to distant establishment sites. However, the absolute amount of predicted transgene flow was consistent between scenarios and proportional to the area of plantations. From this standpoint, small-scale trials will be valuable for predicting some dimensions of transgene flow.

The modeling results suggest that gene flow would best be estimated from commercial-scale plantations rather than extrapolated from small-scale field trials, as has been observed in other systems (Beckie et al., 2010). However, the simulations also suggest that there are a wide variety of transgenes that could be safely employed in an 'adaptive management' (i.e. learn by doing) strategy where gene flow is predicted to be low due to the use of linked sterility-imparting genes. This may allow monitoring of transgenic fitness and spread in ecologically realistic settings and at appropriate scales with a tolerable level of risk (Kareiva et al., 1996; Beckie et al., 2010).

Factors controlling transgene spread

Our model allowed for quantitative evaluation of mitigating genes that might inhibit transgene spread. Some transgenes are themselves likely to have negative effects on fitness of trees in the wild, and thus retard their own spread. This may include genes that promote bioremediation, encode new industrial products such as biopolymers, reduce stature, or modify wood quality to make it more susceptible to chemical or biological breakdown during pulping or liquid fuel production (Strauss et al., 2001; Strauss, 2003). Reductions of lignin content often lead to fitness-reducing traits, such as reduced growth and survival, sensitivity to stress, and lodging (Pedersen et al., 2005). Alternatively, transgenes may be closely linked or flanked by mitigation genes (Gressel, 1999). Scenarios with low but not complete infertility (e.g. relative fertility of 0.001; Fig. 6), or with unstable sterility (Brunner et al., 2007), can be viewed as forms of mitigation. Our simulations showed that such genes, even if their effects are modest, incomplete, or unstable, can be effective mitigating agents for transgene flow.

Our modeling studies provided insights into the relative importance of diverse biological and management factors in determining rates of transgene spread. Transgenic competitiveness, fertility, and rates and patterns of dispersal have received substantial attention in recent analyses of transgene dispersal (Haygood et al., 2004; Hails & Morley, 2005; Chapman & Burke, 2006; Meirmans et al., 2009). However, we have shown that levels of transgene spread are strongly affected by ecological and management factors that affect habitat creation and the abundance of mature transgenic trees on the landscape. Recent analyses based on simple, context-free population genetics models have reached conclusions that clearly do not hold in our system. For example, Chapman & Burke (2006) used a relatively simple stepping-stone model of gene flow to infer that gene flow is relatively unimportant compared to the selective effects of transgenes in determining rates of spread. Similarly, Haygood et al. (2004) and Meirmans et al. (2009) simulated transgene spread using non-dimensional population genetic models with no explicit ecological or management components. In contrast to our findings, these studies concluded that recurrent gene flow from transgenic crops could overwhelm wild populations over time, even in the presence of partial sterility of the transgenic crop (Haygood et al., 2004) or when cultivation of the transgenic crop is halted after only 15 yr of flowering (Meirmans et al., 2009). One difference in our philosophical approaches is these latter studies generally consider outcomes on evolutionary time-scales of thousands to tens of thousands of years, while we confined our simulations to decadal scales. Our choice was driven by the desire to perform our simulations in a time frame that is relevant to management and policy timeframes, and because of the vast, unknowable climatic, land use, and biotic uncertainties associated with long-term projections. Nonetheless, our results demonstrate that for near- to mid-term assessments of potential transgene spread, it is essential to explicitly consider the ecological and agronomic contexts of transgene deployment for risk assessment (Snow et al., 2005).

The implications of the work presented here extend well beyond our specific model system. Our simulations suggest that for transgenes with a wide variety of effects, a tree with extensive pollen and seed flow and large populations of interfertile wild relatives will not necessarily result in a substantial fraction of transgenic trees in wild populations. The key element, which has been missing from other studies thus far, is the competitive interactions between wild and transgenic trees in a highly dynamic and heterogeneous ecological and management setting. Although our range of direct inference is necessarily limited to poplar trees in the Columbia River Gorge where the simulations occurred, the lessons appear general, and the model can be readily extended to other organisms for which large-scale disturbance is a prerequisite for establishment and for which habitat requirements and transition rates are well-defined. Adapting the model to other scenarios only requires the development of GIS layers from different habitats and landscapes, and altering the model parameters (controlled by simple text files) to fit the biology of the species under study. The approach we propose here therefore provides a flexible way to assess the complex interacting factors that control gene spread, whether from transgenic, exotic, or wild species.

References

Al Ahmad H, Dwyer J, Moloney M, Gressel J. 2006. Mitigation of establishment of *Brassica napus* transgenes in volunteers using a tandem construct containing a selectively unfit gene. *Plant Biotechnology Journal* **4**: 7–21.

Allainguillaume J, Harwood T, Ford CS, Cuccato G, Norris C, Allender CJ, Welters R, King GJ, Wilkinson MJ. 2009. Rapeseed cytoplasm gives advantage in wild relatives and complicates genetically modified crop biocontainment. *New Phytologist* **183**: 1201–1211.

Allen TH. 1999. *Areal distribution, change, and restoration potential of wetlands within the lower Columbia River riparian zone, 1948–1991*. PhD thesis, Oregon State University, Corvallis, OR, USA.

Beckie HJ, Hall LM, Simard MJ, Leeson JY, Willenborg CJ. 2010. A framework for post-release environmental monitoring of second-generation crops with novel traits. *Crop Science* **50**: 1587–1604.

Benetka V, Mottl J, Vackova K, Pošpíšková M, Dubsky M. 1999. Estimation of the introgression level in *Populus* nigra L. populations by means of isozyme gene markers. *Silvae Genetica* **48**: 218–223.

Boerjan W. 2005. Biotechnology and the domestication of forest trees. *Current Opinion in Biotechnology* **16**: 159–166.

Box GP, Hunter WG, Hunter JS. 1978. *Statistics for experiments*. New York, NY, USA: John Wiley and Sons.

Braatne JH, Jamieson R, Gill KM, Rood SB. 2007. Instream flows and the decline of riparian cottonwoods along the Yakima River, Washington, USA. *River Research and Applications* **23**: 247–267.

Brunner AM, Busov VB, Strauss SH. 2004. Poplar genome sequence: functional genomics in an ecologically dominant plant species. *Trends in Plant Science* **9**: 49–56.

Brunner AM, Li JY, DiFazio SP, Shevchenko O, Montgomery BE, Mohamed R, Wei H, Ma C, Elias AA, VanWormer K *et al.* 2007. Genetic containment of forest plantations. *Tree Genetics & Genomes* **3**: 75–100.

Chandler S, Dunwell JM. 2008. Gene flow, risk assessment and the environmental release of transgenic plants. *Critical Reviews in Plant Sciences* **27**: 25–49.

Chapman MA, Burke JM. 2006. Letting the gene out of the bottle: the population genetics of genetically modified crops. *New Phytologist* **170**: 429–443.

Clark JS, Fastie C, Hurtt G, Jackson ST, Johnson C, King GA, Lewis M, Lynch J, Pacala S, Prentice C *et al.* 1998. Reid's paradox of rapid plant migration: dispersal theory and interpretation of palaeoecological records. *BioScience* **48**: 13–24.

Csencsics D, Angelone S, Paniga M, Rotach P, Rudow A, Sabiote E, Schwab P, Wohlhauser P, Holderegger R. 2009. A large scale survey of *Populus nigra* presence and genetic introgression from non-native poplars in Switzerland based on molecular identification. *Journal for Nature Conservation* **17**: 142–149.

Daniell H. 2002. Molecular strategies for gene containment in transgenic crops. *Nature Biotechnology* **20**: 581–586.

DeBell DS. 1990. *Populus trichocarpa* Torr. & Gray, Black Cottonwood. In: Burns RM, Honkala BH, eds. *Silvics of North America Vol. 2. Hardwoods. USDA Forest Service Agriculture Handbook 654*. Washington, DC, USA: USDA Forest Service, 570–576.

Doty SL, James CA, Moore AL, Vajzovic A, Singleton GL, Ma C, Khan Z, Xin G, Kang JW, Park AY *et al.* 2007. Enhanced phytoremediation of volatile environmental pollutants with transgenic trees. *Proceedings of the National Academy of Sciences, USA* **104**: 16816–16821.

Eckenwalder JE. 1984. Natural intersectional hybridization between North American species of *Populus* (Salicaceae) in sections *Aigeiros* and *Tacamahaca*. I. Population studies of *P. × parryi*. *Canadian Journal of Botany* **62**: 317–324.

Ellison AM, Bank MS, Clinton BD, Colburn EA, Elliott K, Ford CR, Foster DR, Kloeppel BD, Knoepp JD, Lovett GM *et al.* 2005. Loss of foundation species: consequences for the structure and dynamics of forested ecosystems. *Frontiers in Ecology and the Environment* **3**: 479–486.

van Frankenhuyzen K, Beardmore T. 2004. Current status and environmental impact of transgenic forest trees. *Canadian Journal of Forest Research* **34**: 1163–1180.

Gould F. 1998. Sustainability of transgenic insecticidal cultivars: integrating pest genetics and ecology. *Annual Review of Entomology* **43**: 701–726.

Graham RL, Hunsaker CT, O'Neill RV, Jackson BL. 1991. Ecological risk assessment at the regional scale. *Ecological Applications* **1**: 196–206.

Gressel J. 1999. Tandem constructs: preventing the rise of superweeds. *Trends in Biotechnology* **17**: 361–366.

Hails RS, Morley K. 2005. Genes invading new populations: a risk assessment perspective. *Trends in Ecology & Evolution* **20**: 245–252.

Hancock JF. 2003. A framework for assessing the risk of transgenic crops. *BioScience* **53**: 512–519.

Haygood R, Ives AR, Andow DA. 2004. Population genetics of transgene containment. *Ecology Letters* **7**: 213–220.

Heinze B. 1997. A PCR marker for a *Populus deltoides* allele and its use in studying introgression with native European *Populus nigra*. *Belgian Journal of Botany* **129**: 123–130.

Higgins SI, Clark JS, Nathan R, Hovestadt T, Schurr F, Fragoso JMV, Aguiar MR, Ribbens E, Lavorel S. 2003. Forecasting plant migration rates: managing uncertainty for risk assessment. *Journal of Ecology* **91**: 341–347.

Hinchee M, Rottmann W, Mullinax L, Zhang CS, Chang SJ, Cunningham M, Pearson L, Nehra N. 2009. Short-rotation woody crops for bioenergy and biofuels applications. *In Vitro Cellular & Developmental Biology-Plant* **45**: 619–629.

Hoenicka H, Fladung M. 2006. Biosafety in *Populus spp.* and other forest trees: from non-native species to taxa derived from traditional breeding and genetic engineering. *Trees – Structure and Function* **20**: 131–144.

Huang JG, Tardif J, Denneler B, Bergeron Y, Berninger F. 2008. Tree-ring evidence extends the historic northern range limit of severe defoliation by insects in the aspen stands of western Quebec, Canada. *Canadian Journal of Forest Research* **38**: 2535–2544.

James RR, DiFazio SP, Brunner AM, Strauss SH. 1998. Environmental effects of genetically engineered woody biomass crops. *Biomass & Bioenergy* **14**: 403–414.

Kareiva P, Parker IM, Pascual M. 1996. Can we use experiments and models in predicting the invasiveness of genetically engineered organisms? *Ecology* **77**: 1670–1675.

Kuparinen A, Schurr FM. 2007. A flexible modelling framework linking the spatio-temporal dynamics of plant genotypes and populations: application to gene flow from transgenic forests. *Ecological Modelling* **202**: 476–486.

Kuparinen A, Schurr FM. 2008. Assessing the risk of gene flow from genetically modified trees carrying mitigation transgenes. *Biological Invasions* **10**: 281–290.

Kwit C, Moon HS, Warwick SI, Stewart CN. 2011. Transgene introgression in crop relatives: molecular evidence and mitigation strategies. *Trends in Biotechnology* **29**: 284–293.

MacArthur RH, Levins R. 1967. The limiting similarity, convergence, and divergence of coexisting species. *American Naturalist* **101**: 377–385.

Mason P, Braun L, Warwick SI, Zhu B, Stewart CN Jr. 2003. Transgenic Bt-producing *Brassica napus*: *Plutella xylostella* selection pressure and fitness of weedy relatives. *Environmental Biosafety Research* **2**: 263–276.

Meirmans PG, Bousquet J, Isabel N. 2009. A metapopulation model for the introgression from genetically modified plants into their wild relatives. *Evolutionary Applications* **2**: 160–171.

Meirmans PG, Lamothe M, Gros-Louis MC, Khasa D, Perinet P, Bousquet J, Isabel N. 2010. Complex patterns of hybridization between exotic and native North American poplar species. *American Journal of Botany* **97**: 1688–1697.

Moon HS, Li Y, Stewart CN. 2009. Keeping the genie in the bottle: transgene biocontainment by excision in pollen. *Trends in Biotechnology* **28**: 3–8.

Nathan R, Katul GG, Horn HS, Thomas SM, Oren R, Avissar R, Pacala SW, Levin SA. 2002. Mechanisms of long-distance dispersal of seeds by wind. *Nature* **418**: 409–413.

Pedersen JF, Vogel KP, Funnell DL. 2005. Impact of reduced lignin on plant fitness. *Crop Science* **45**: 812–819.

Perlack RD, Wright LL, Turhollow AF, Graham RL, Stokes BJ, Erbach DC. 2005. *Biomass as feedstock for a bioenergy and bioproducts industry: the technical feasibility of a billion-ton annual supply*. Oak Ridge, TN, USA: Oak Ridge National Laboratory.

Pošpíšková M, Šálková I. 2006. Population structure and parentage analysis of black poplar along the Morava River. *Canadian Journal of Forest Research* **36**: 1067–1076.

Raybould A. 2007. Ecological versus ecotoxicological methods for assessing the environmental risks of transgenic crops. *Plant Science* **173**: 589–602.

Reichman JR, Watrud LS, Lee EH, Burdick CA, Bollman MA, Storm MJ, King GA, Mallory-Smith C. 2006. Establishment of transgenic herbicide-resistant creeping bentgrass (Agrostis stolonifera L.) in nonagronomic habitats. *Molecular Ecology* **15**: 4243–4255.

Roland J. 1993. Large-scale forest fragmentation increases the duration of tent caterpillar outbreak. *Oecologia* **93**: 25–30.

Sears MK, Hellmich RL, Stanley-Horn DE, Oberhauser KS, Pleasants JM, Mattila HR, Siegfried BD, Dively GP. 2001. Impact of Bt corn pollen on monarch butterfly populations: a risk assessment. *Proceedings of the National Academy of Sciences, USA* **98**: 11937–11942.

Sedjo R. 2005. Will developing countries be the early adapters of genetically engineered forests? *AbBioForum* **8**: 205–212.

Shugart HH. 1998. *Terrestrial Ecosystems in Changing Environments*. New York, NY, USA: Cambridge University Press.

Slavov GT, Howe GT, Gyaourova AV, Birkes DS, Adams WT. 2005. Estimating pollen flow using SSR markers and paternity exclusion: accounting for mistyping. *Molecular Ecology* **14**: 3109–3121.

Slavov GT, Leonardi S, Adams WT, Strauss SH, DiFazio SP. 2010. Population substructure in continuous and fragmented stands of *Populus trichocarpa*. *Heredity* **105**: 348–357.

Slavov GT, Leonardi S, Burczyk J, Adams WT, Strauss SH, DiFazio SP. 2009. Extensive pollen flow in two ecologically contrasting populations of *Populus trichocarpa*. *Molecular Ecology* **18**: 357–373.

Smouse PE, Robledo-Arnuncio JJ, Gonzalez-Martinez SC. 2007. Implications of natural propagule flow for containment of genetically modified forest trees. *Tree Genetics & Genomes* **3**: 141–152.

Smulders MJM, Beringen R, Volosyanchuk R, Broeck AV, Van Der Schoot J, Arens P, Vosman B. 2008. Natural hybridisation between *Populus nigra* L. and *P. x canadensis* Moench. hybrid offspring competes for niches along the Rhine River in the Netherlands. *Tree Genetics & Genomes* **4**: 663–675.

Snow AA, Andow DA, Gepts P, Hallerman EM, Power A, Tiedje JM, Wolfenbarger LL. 2005. Genetically engineered organisms and the environment: current status and recommendations. *Ecological Applications* **15**: 377–404.

Stewart CN, Halfhill MD, Warwick SI. 2003. Transgene introgression from genetically modified crops to their wild relatives. *Nature Reviews Genetics* **4**: 806–817.

Stewart CN, Prakash CS. 1998. Chloroplast-transgenic plants are not a gene flow panacea. *Nature Biotechnology* **16**: 401.

Strauss SH. 2003. Genetic technologies – Genomics, genetic engineering, and domestication of crops. *Science* **300**: 61–62.

Strauss SH, Brunner AM, Busov VB, Ma CP, Meilan R. 2004. Ten lessons from 15 years of transgenic *Populus* research. *Forestry* **77**: 455–465.

Strauss SH, DiFazio SP, Meilan R. 2001. Genetically modified poplars in context. *Forestry Chronicle* **77**: 271–279.

Strauss SH, Kershen DL, Bouton JH, Redick TP, Tan HM, Sedjo RA. 2010. Far-reaching deleterious impacts of regulations on research and environmental studies of recombinant DNA-modified perennial biofuel crops in the United States. *BioScience* **60**: 729–741.

Strauss SH, Knowe SA, Jenkins J. 1997. Benefits and risks of transgenic Roundup Ready Cottonwoods. *Journal of Forestry* **95**: 12–19.

Svab Z, Maliga P. 2007. Exceptional transmission of plastids and mitochondria from the transplastomic pollen parent and its impact on transgene containment. *Proceedings of the National Academy of Sciences, USA* **104**: 7003–7008.

Thompson SL, Lamothe M, Meirmans PG, Perinet P, Isabel N. 2010. Repeated unidirectional introgression towards *Populus balsamifera* in contact zones of exotic and native poplars. *Molecular Ecology* **19**: 132–145.

Tuskan GA, DiFazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S, Rombauts S, Salamov A *et al.* 2006. The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* **313**: 1596–1604.

Vanden Broeck A, Cottrell J, Quataert P, Breyne P, Storme V, Boerjan W, Van Slycken J. 2006. Paternity analysis of *Populus nigra* L. offspring in a

Belgian plantation of native and exotic poplars. *Annals of Forest Science* **63**: 783–790.

Vanden Broeck A, Storme V, Cottrell JE, Boerjan W, Van Bockstaele E, Quataert P, Van Slycken J. 2004. Gene flow between cultivated poplars and native black poplar (*Populus nigra* L.): a case study along the river Meuse on the Dutch–Belgian border. *Forest Ecology and Management* **197**: 307–310.

Vanden Broeck A, Villar M, Van Bockstaele E, Van Slycken J. 2005. Natural hybridization between cultivated poplars and their wild relatives: evidence and consequences for native poplar populations. *Annals of Forest Science* **62**: 601–613.

Warwick SI, Beckie HJ, Hall LM. 2009. Gene flow, invasiveness, and ecological impact of genetically modified crops. In: Mousseau TA, Schlichting CD, eds. *Year in evolutionary biology 2009*. Oxford, UK: Blackwell Publishing, 72–99.

Watrud LS, Lee EH, Fairbrother A, Burdick C, Reichman JR, Bollman M, Storm M, King G, Van de Water PK. 2004. Evidence for landscape-level, pollen-mediated gene flow from genetically modified creeping bentgrass with CP4 EPSPS as a marker. *Proceedings of the National Academy of Sciences, USA* **101**: 14533–14538.

Wilkinson MJ, Elliott LJ, Allainguillaume J, Shaw MW, Norris C, Welters R, Alexander M, Sweet J, Mason DC. 2003a. Hybridization between *Brassica napus* and *B. rapa* on a national scale in the United Kingdom. *Science* **302**: 457–459.

Wilkinson MJ, Sweet J, Poppy GM. 2003b. Risk assessment of GM plants: avoiding gridlock? *Trends in Plant Science* **8**: 208–212.

Williams CG. 2010. Long-distance pine pollen still germinates after meso-scale dispersal. *American Journal of Botany* **97**: 846–855.

Williams CG, Davis BH. 2005. Rate of transgene spread via long-distance seed dispersal in *Pinus taeda*. *Forest Ecology and Management* **217**: 95–102.

Williams CG, Ladeau SL, Oren R, Katul GG. 2006. Modeling seed dispersal distances: implications for transgenic *Pinus taeda*. *Ecological Applications* **16**: 117–124.

Supporting Information

Additional supporting information may be found in the online version of this article.

Fig. S1 Relationship between area of mature transgenics and transgenic basal area for a trait with neutral fitness (Neut.), and insect resistance (Bt) and with fertility of 0.5 and 0.1.

Fig. S2 Seed production per catkin from controlled crosses involving *P. trichocarpa* × *P. deltoides* (TD) hybrids and *P. trichocarpa* (T) wild trees.

Fig. S3 Comparison of modeled versus observed pollen flow from plantations. Points represent observed pollination as determined from paternity analyses.

Fig. S4 Comparison of modeled versus observed seed flow from plantations. Points represent observed percentage hybrid seeds as determined from maternity analyses of seeds captured in traps near plantations at River Ranch.

Fig. S5 Least square means from fractional factorial for lower and upper values of main effects explanatory variables in an analysis that varied each main effect by ± 10% from default values.

Table S1 Description of habitat types in original data layers

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.