Supporting Information Figs S1–S5, Tables S1–S6 and Methods S1–S7

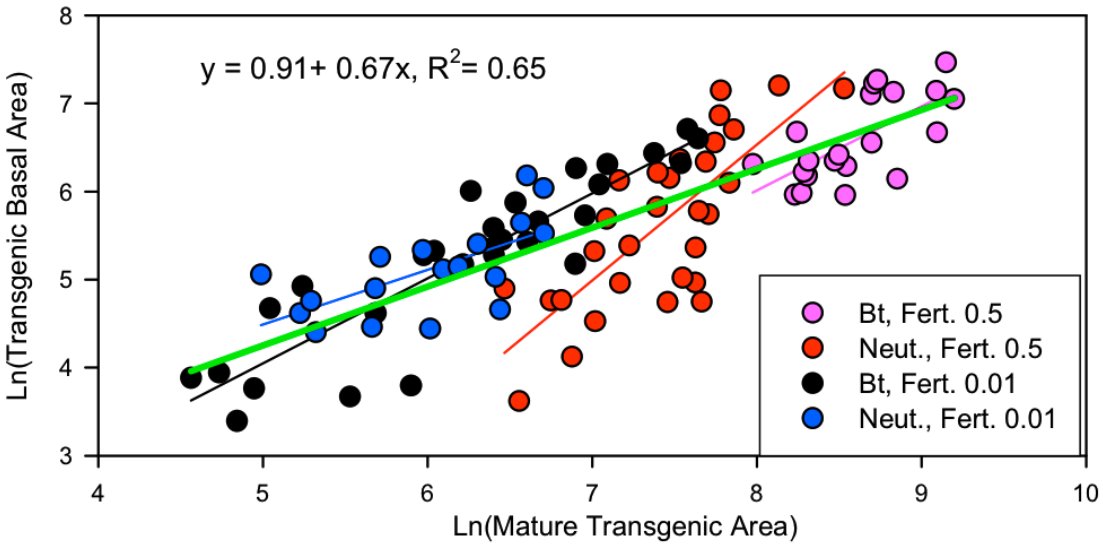

Fig. S1 Relationship between area of mature transgenics and transgenic basal area for a trait with neutral fitness (Neut.), and insect resistance (Bt) and with fertility of 0.5 and 0.1. Lines are linear regressions for each simulation; green line and equation represent a regression analysis for all data combined.

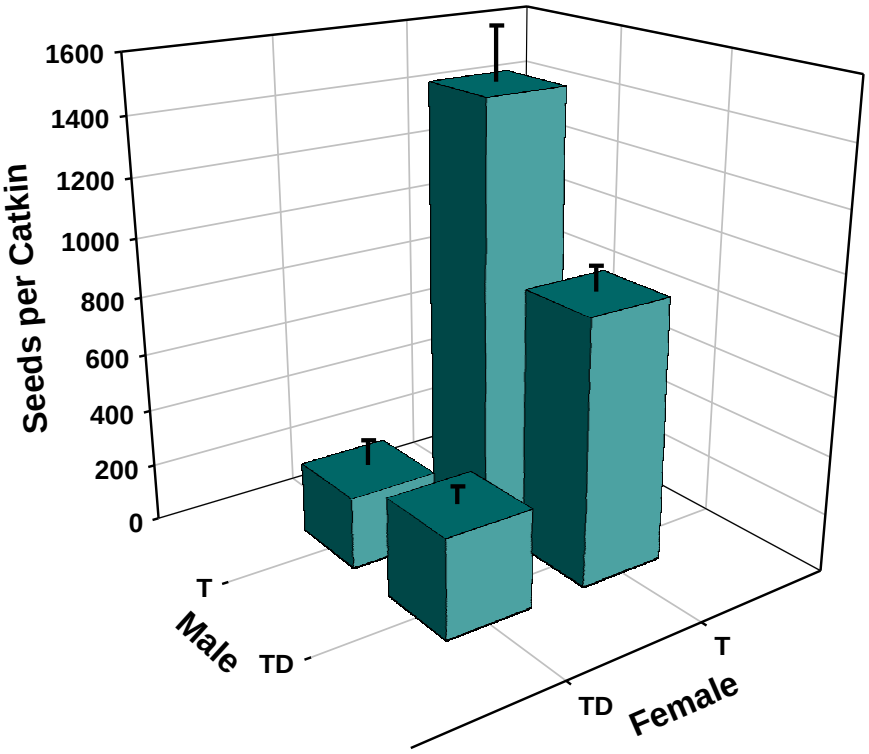

Fig. S2 Seed production per catkin from controlled crosses involving *Populus trichocarpa* × *Populus deltoides* (TD) hybrids and P. *trichocarpa* (T) wild trees. TD male, clone 47-174; TD Female, clone 49-177; T male, clone SF-41; T female, clone PTTG.

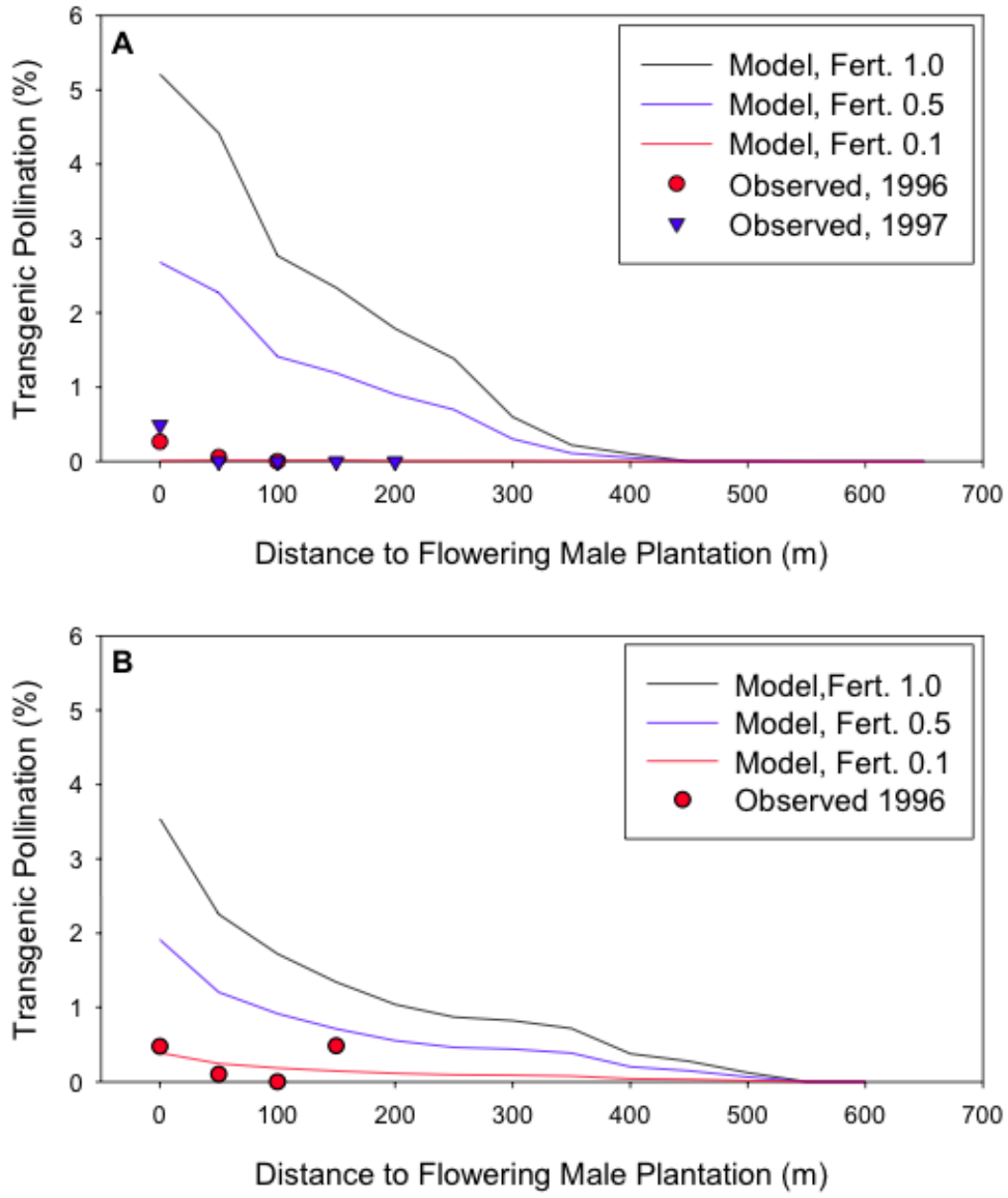

Fig. S3 Comparison of modeled versus observed pollen flow from plantations. Lines represent average percentage of seeds that were transgenic for each distance class for STEVE model simulations with landscapes mimicking landscape conditions at the time of field studies. The three lines are for different levels of transgenic fertility. Points represent observed pollination as determined from paternity analyses. **A.** Clatskanie. **B.** River Ranch.

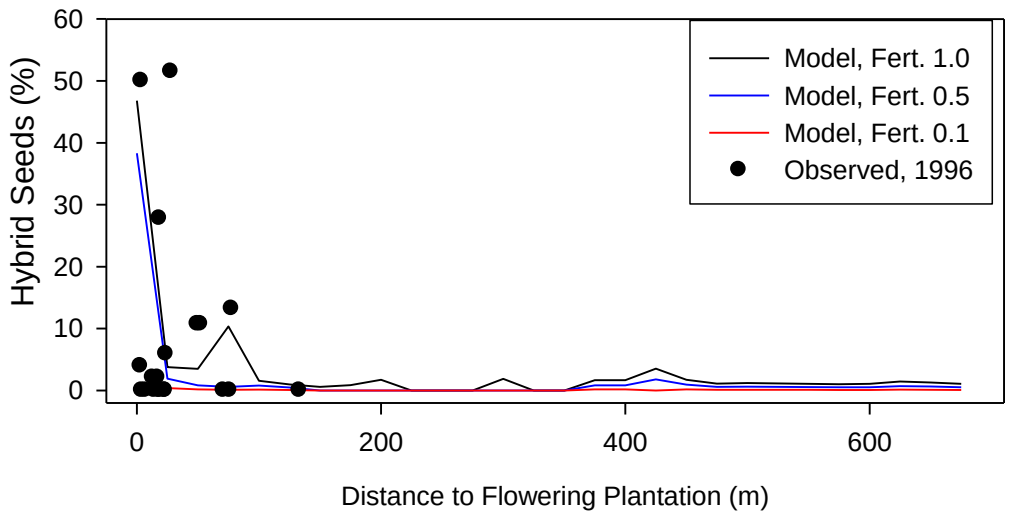

Fig. S4 Comparison of modeled versus observed seed flow from plantations. Lines represent average percentage of seeds that were transgenic for each distance class for STEVE model simulations with landscapes mimicking landscape conditions at the time of field studies. Points represent observed percentage hybrid seeds as determined from maternity analyses of seeds captured in traps near plantations at River Ranch.

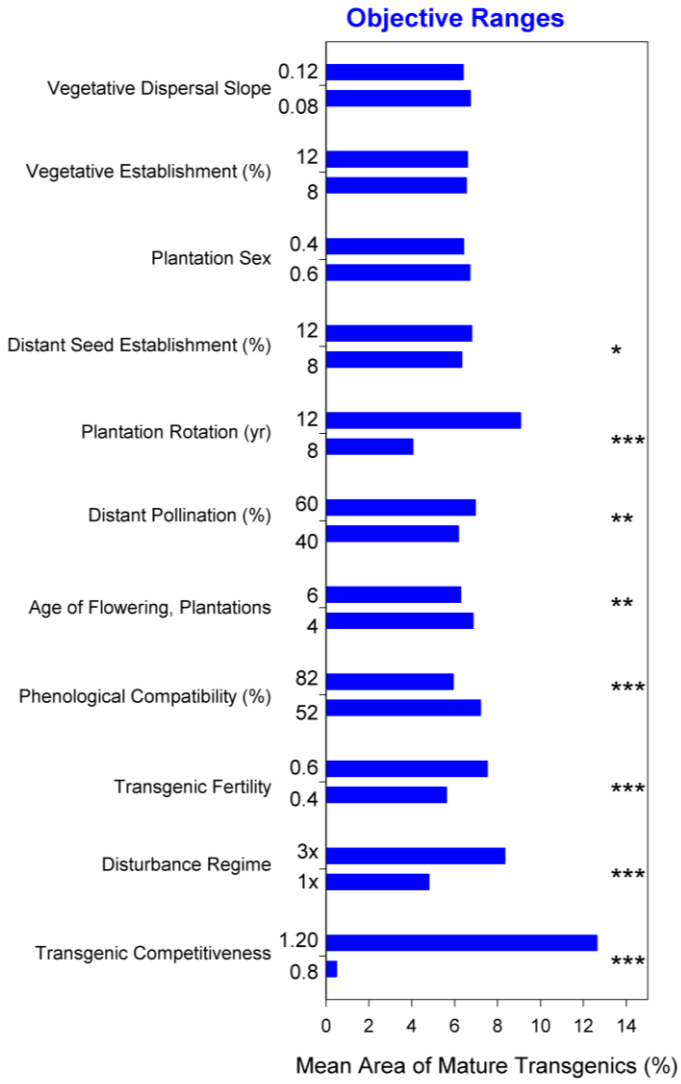

Fig. S5 Least square means from fractional factorial for lower and upper values of main effects explanatory variables in an analysis that varied each main effect by ± 10% from default values. Asterisks indicate the probability of observing larger values of F for the main effect: ***, $P < 0.0001$; **, $P < 0.01$; *, $P < 0.05$.

Supporting Methods

Methods S1 Landscape data

Habitat types were delineated as part of a joint project between Oregon State University and the U.S. Army Corps of Engineers aimed at assessing changes in wetland habitats since 1948 on the lower Columbia River (Allen, 1999). For the present study, we used data from river km 49 to 93, encompassing a strip 5 km wide on both sides of the Columbia River. This is currently an area of intensive *Populus* cultivation. Air photos were used to delineate broad habitat types in the study area (Table S1). Photos were primarily black and white and 1:48,000 in scale. Photos were laid out in adjacent flight lines, overlaid with mylar, and viewed with a stereoscope. Habitats were delineated on the mylar and subsequently zoom-transferred to United States Geological Survey 7.5' quadrangle maps to correct for variation in scale and aerial distortion. Habitats were classified in this manner for 1961, 1973, 1983, and 1991.

The data were supplemented and adapted for the current project by delineating *Populus* stands from 1991 photos only. *Populus* trees were distinguished from other hardwoods based primarily on geomorphology, crown structure, and size. *Populus* occur as linear or arcuate features, primarily within the active and historic flood plain, most often on well-drained soils on riverbanks, bars, and islands, though they can occur on upland sites if appropriate moisture conditions exist (Braatne *et al.*, 1996). In addition, crown margins have a rounded shape and a clumped foliage pattern, and *Populus* crowns are often emergent in mixed stands (DeBell, 1990).

Populus stands were divided into 'pure' (>70% cottonwood) and 'mixed' stands, and 'mature' and 'immature' size classes (< 5 m height). Generally, only stands of 0.5 ha or more were identified, though smaller patches could be delineated if they occurred in diagnostic shapes and locations (*e.g.,* linear stands on river banks). Accuracy was improved in key areas (*e.g.*, in the vicinity of *Populus* plantations) by using larger scale color photos. In addition, data were verified and corrected following multiple field visits using a hand-held Global Positioning System unit.

The mylar overlays were digitized and polygon topology was built and cleaned using Arc/Info 7.2 (ESRI Inc., Redlands, CA). This vector layer was then converted to a grid file with 10 m cells. We derived layers depicting distance and direction to the river for each cell. In addition, we obtained a 33 m digital elevation model for the study area and resampled

Table S1 Description of habitat types in original data layers (1961, 1973, 1983, 1991) (Allen, 1999; Cowardin *et al.*, 1979). *Populus* habitat types below the double line were delineated for 1991 only, and encompass several of the broader habitat types. Habitat types were subsequently merged for the simulation (STEVE codes).

Habitat Type	Description	STEVE codes
Barren Land	Sand dunes, rock lands, sandy beaches, dredge spoils, and quarries	BARR
Agriculture	Field crops, pastures, orchards	AGRI
Urban	Residential, industrial, transportation, and mining operations	URBN
Forested wetland	Wetland with > 25% persistent trees	FWET
River	River	RIVR
Lake	Lake	WATR
Slough	Slough	WATR
Reed Canary Grass	Reed canary grass	WTLD
Estuarine Intertidal Wetland	Exposed wetland flooded by tides	WTLD
Lacustrine Tidal Wetland	Shallow wetlands (< 2 m) along lake shores	WTLD
Palustrine Wetland	Non-tidal wetlands with persistent emergents, shrubs, or trees (<25%)	FWET
Riverine Intertidal Wetland	Wetland adjacent to river, subject to periodic inundation and exposure	WTLD
Riverine Tidal Wetland	Wetland adjacent to river, low gradient, water velocity influenced by tides	WTLD
Shrub/Scrub	Burned areas, recent clearcuts	SCRB
Grassland	Grassland	SCRB
Mature Coniferous	Coniferous forest, > 5 m height	MFOR
Young Coniferous	Coniferous forest, < 5 m height	IFOR
Mature Broadleaf Forest	Broadleaf forest in uplands, > 5 m height	MFOR
Young Broadleaf Forest	Broadleaf forest in uplands, < 5 m height	IFOR
Mature Mixed Forest	Mixed conifer and broadleaf forest in uplands, > 5 m height	MFOR
Young Mixed Forest	Mixed conifer and broadleaf forest in uplands, < 5 m height	IFOR
Populus Plantation	Plantation	PLNT
Young Mixed *Populus*	Up to 70% *Populus*, < 5 m height	CTWD
Mature Mixed *Populus*	Up to 70% *Populus*, > 5 m height	CTWD
Young Pure *Populus*	Greater than 70% *Populus*, < 5 m height	CTWD
Mature Pure *Populus*	Greater than 70% *Populus*, > 5 m height	CTWD

this to a 10 m grid size and derived slope and aspect using ARC/Info. All grid layers were exported as ASCII files containing a single datum for each cell of the grid. These ASCII files were subsequently converted to binary files to be used for the simulation.

Wild *Populus* Stands

Our information on *Populus* stands was approximate, having been derived principally from low-resolution air photos. We therefore created simulated populations based on localized observations as described below. Initially all trees outside plantations were nontransgenic. We assigned age randomly between 1 and 10 for 'immature' stands (those less than 5 m tall) and between 10 and 100 for mature stands.We first derived a relationship between carrying capacity and age, based in part on data from the experimental establishment plots described in the main manuscript

$$N\,max_a = 1 + 4000 * \left(\frac{e^{-\frac{a}{3}}}{1 + e^{-\frac{a}{3}}} \right)$$

where *a* is age.

For pure stands, the initial number of trees per block was randomly assigned between 70 and 100% of carrying capacity for that age. For mixed stands, initial density was randomly set up to 70% of carrying capacity. This was based on the criteria used in delineating *Populus* stands from air photos. Sex ratio was determined through a binomial sampling process. Each tree in a cohort was initially assigned sex with a 50% probability of being male or female.

Populus Plantations

Populus plantation characteristics are set from a configuration file at the time of model initiation. Plantations are divided into blocks or management units, each of which may be planted with a different clone or genotype (transgenic or conventional). Each block is assigned an age, sex, and genotype (transgenic or conventional) with the configuration file. Floral phenology is set randomly at rates determined by another input file. Values for plantations established after model initiation (*i.e.*, following harvest) are set randomly according to probabilities set by the user. Plantation density, rotation, and age of flowering can all be customized for each model run.

Plantation Management

Plantation management occurs following the general practices of forest industries that grow hybrid *Populus* in the study area. Plantations are managed in large (mean = 16 ha) single-clone blocks in an even-aged fashion, so that approximately the same area is harvested each year. Trees are harvested upon reaching rotation age, which is set at model initiation. New plantations are then established, with genotype and sex being determined with probabilities set at model initiation.

Methods S2 Disturbance

We modeled change in *Populus* stands using a chronosequence approach, a method that has been applied extensively in analyses of landscape change (*e.g.*, (Frelich *et al.*, 1993);Turner 1987). The general concept is that changes in delineated *Populus* polygons from different dates reflect establishment and mortality of *Populus* stands. We calculated transition rates between *Populus* and other habitat types by overlaying layers of adjacent years and tallying changes in habitat types of intersecting polygons. We chose to focus on two key transition periods: 1961 to 1973, and 1983 to 1991. The earlier transition represents the prevailing disturbance regime before the major flood control dams were instituted (*i.e.*, the John Day dam, 1968), and at least one large flood occurred in 1964 (Allen, 1999). The later transition period represents the post flood-control regime, though some residual effects of earlier floods might still be evident.
Annual *Populus* establishment rates were calculated as:

$$E_h = \frac{A_{ph}}{A_h t}$$

where A_{ph} is area (number of cells) of habitat type h in prior data layer that became *Populus* during the interval, A_h is total area of habitat type h in the previous data layer for which data exist in the more recent data layer, and t is number of years in the interval.
Similarly, mortality rates were calculated as:

$$M_h = \frac{A_{hp}}{A_p t}$$

where A_{hp} is area of *Populus* that became habitat type h during the interval, and Ap is total area of *Populus*. In addition to transition rates, we also determined the distribution of establishment and mortality patch sizes for each interval.

The first step in simulating creation and destruction of *Populus* populations is to select the disturbance regime. The main factor controlling disturbance in this system is flooding, and this is greatly attenuated by flood-control dams. However, large floods are still possible in this system (Fierke & Kauffman, 2005), and we assume these will occur at approximately 100-year intervals. Therefore, we instituted the pre-flood control disturbance regime (*i.e.,* 1961-1973 transition rates) with a 1% probability.

The probabilities calculated above are on an areal basis, and therefore represent the probability that a given unit of land (a cell) will be converted. However, disturbances are generally larger than one cell (100 m^2), so multiple cells must be converted in unison. Our approach was to first calculate the total area (number of cells) to be converted in a given year. For each habitat type h, area converted to *Populus* annually is:

$$C_{hp} = E_h A_h$$

Similarly, area of *Populus* to be converted to each habitat type is:

$$C_{ph} = M_h A_p .$$

Then we sampled polygons from the empirical distributions of patch sizes until the total area of the polygons exceeded the area to be converted. The probability of polygon creation at any location on the landscape then becomes:

$$P_h = \frac{N_h}{A_h}$$

where N_h is the number of polygons of habitat type h to be converted.

One further complication is that probability of polygon creation depends on spatial context of the nucleation point. For example, new *Populus* polygons are more likely to be aggregated to existing *Populus* polygons than to be created in isolation. Similarly, *Populus* mortality is more likely to occur at the edges of a stand rather than in the interior. Therefore, we forced a proportion of the new polygons to aggregate to existing *Populus* stands each year.

Agricultural fields represent a special case because they are consistently subject to anthropogenic influences. However, the edges of agricultural fields are often subject to colonization by *Populus*, especially in cases where the fields are adjacent to drainage ditches and roads. We therefore allowed colonization of field edges with a customizable probability and maximum patch size. Also, *Populus* may colonize fields containing extensively managed perennial species such as tree plantations in some areas. We therefore simulated managed lands that were not subjected to annual tilling or grazing and were thus susceptible to *Populus* establishment. Establishment within such fields was limited to a single cell (100 m^2) with establishment and mortality rates set by the user. Mortality in agricultural fields is also controlled by herbicide application, and modulated by genotype-specific herbicide tolerance.

Finally, large-scale abandonment of agricultural land has been common in this study area, and some of this land ultimately succeeds to *Populus* stands. For example, the Lewis and Clark National Wildlife refuge was established in the study area in 1971, and diked agricultural land comprised large expanses of this refuge (Allen, 1999). The dikes were eventually breached, and *Populus* stands developed by 1991. We therefore calculated the rate of conversion to *Populus* of known abandoned agricultural land, and applied this rate to fields that were abandoned at set intervals during model runs. Similarly, conversion of *Populus* stands to agriculture has been relatively common, and we included this transition as well.

Methods S3 Production and dispersal of pollen, seeds, and vegetative propagules

The primary purpose of this model is to explore the process of gene flow from transgenic *Populus* plantations. The basic simulation unit is a 100 m^2 cell, which approximately represents one fully grown tree. However, the absolute number and size of trees is not important in this model, because it operates primarily on relative proportions of transgenic and conventional genotypes. This is fortunate, because data on growth and productivity of wild *Populus* stands are largely lacking. However, it was important to derive a relationship between basal area and age, because increases in size and fecundity should decline with time. Therefore, we used the limited, anecdotal data available to us on ‘culmination’ of annual increment (DeBell, 1990) in wild populations to derive a relationship with basal area .

We related seed, pollen, and vegetative propagule production directly to basal area, because basal area is correlated with crown size, which in turn largely determines fecundity (Greene &

Johnson, 1999). We also assumed that larger crowns have a greater chance of producing the broken limbs and secondary shoots that comprise dispersed vegetative propagules. Seed and pollen production are further modulated by sex ratio and genotype-specific fecundity, which allows for simulation of transgenic sterility.

There is substantial concern about possible instability of transgenic traits, and the prospect that genetically engineered sterility could become ineffective under the relatively long rotations and varied environments encountered in tree plantations. We therefore explored the effects of instability on genetically engineered sterility by allowing genotype-specific fecundity to vary stochastically through time.

Seed and pollen production occur after trees reach maturation age, which typically occurs around 10 or 15 years for wild trees, and around 5 years for plantation trees in the study area (Stanton & Villar, 1996).

Relative pollen production is calculated for each genotype within each sexually mature cohort of trees in each *Populus* cell. Representation of pollen and seed is entirely relative, because the most important quantity is the ratio of transgenic to conventional genotypes in the propagule pools. Therefore, pollination units, P*g*, are unitless, and calculated as:

$$P_g = Ba_{ga} f_{ga} Sr_g$$

where Ba_{ga} is basal area (already defined), f_{ga} is relative fecundity of genotype *g* at age *a*, and Sr_g is sex ratio of genotype *g* within the cohort. Relative fecundity can vary annually based on a user-defined standard deviation. In addition, transgenics with reduced fecundity can have fecundity paritially restored according to a user-defined probability.

Vegetative propagule production is calculated as:

$$V_g = Ba_{ga} R$$

where *R* is a random number between 0 and 1.

Production of transgenic seeds (genotype 1) by all plantation and wild trees is calculated as:

$$S_{1a} = \left(f_{1a} Ba_{1a} \left(1 - Sr_1\right) + f_{0a} Ba_{0a} \left(1 - Sr_0\right)\right) K_1$$

where K_1 is the proportion of compatible transgenic pollen arriving at the stigma. *K* depends on pollen input from local sources, background sources, and phenological compatibility (described

in pollination section). This equation incorporates both pollination of conventional trees by transgenic pollen as well as seed production by transgenic females.

Similarly, conventional (genotype 0) seed production is calculated as:

$$S_{0a} = \left(f_{0a} Ba_{0a} \left(1 - Sr_0\right)\right) K_0 .$$

Pollen Dispersal

Pollen dispersal is potentially a key mechanism for dispersing genes from transgenic plantations. Accordingly, we expended substantial effort in characterizing factors associated with pollination success (unpublished data). In particular, we derived a relationship between pollination success and distance using paternity analysis as described in Slavov et al. (2009).

Phenology

We estimated phenological compatibility among trees by calculating the number of days of overlap in flowering time for all male and female trees within stands. We estimated flowering phenology for individual genets by repeated observations of flowers during the period of anthesis at three sites in western Oregon (two on the Columbia River within the simulated area, and one on the Willamette River near Corvallis, Oregon). We devised a scoring system that could be implemented from the ground with binoculars, and which reflects pollen shedding (for males) and receptivity (for females) (Table S2). We recorded phenology observations for up to three years, and determined the mean duration of flowering. We then extrapolated the start dates for receptivity as follows:

$$S = d_{pe} - D\frac{s_e - s_r}{s_a - s_r}$$

where d_{pe} is the earliest day on which receptive stigmata were observed, D is the average duration of receptivity or pollen shedding, S_e is the earliest receptive or shedding stage observed, S_r is the stage at which receptivity or shedding begins, and S_a is the stage at which receptivity or shedding is complete.

The end date of flowering was calculated as:

$$E = S + D.$$

We then calculated the least square mean of flowering date among all years, and calculated the pairwise overlap in flowering among all individuals in the population, with negative numbers representing the number of days separating individuals that did not overlap.

Table S2 Classes used in phenology measurements. Bold classes indicate at least some pollination is occurring.

Males
1-- Catkins not emerging
2-- Catkins emerging, but compact
3-- < 50% of catkins opening
4-- > 50% of catkins opening
5-- < 50% of catkins shedding pollen
6-- > 50% of catkins shedding pollen
7—50–90% of catkins abscissed; vegative bud break
8-- Post-pollination; all catkins abscissed
Females
1-- Catkins not emerging
2-- Catkins emerging, but compact
3-- < 50% of bracts open
4-- > 50% of bracts open
5-- < 50% of capsules distinct
6-- > 50% of capsules distinct
7-- stigmata brown or abscissed; vegetative bud-break
8-- pubescent sutures
9-- seeds shedding

Implementation in Model

For the purposes of the simulation we divided pollen dispersal into two processes: local pollination, which is sensitive to the effects of distance between potential mates, and background pollination, for which distance is not a determinant (Clark *et al.*, 1998; Higgins & Richardson, 1999). We defined the size of local pollination 'neighborhood' as the distance at which pollination success reaches background levels (*i.e.*, the point at which distance between mates is not a significant determinant of reproductive success). This value was between 400 and 500 m for both the Willamette site and eastern Oregon (Slavov *et al.*, 2009). We dispersed pollen to each *Populus* cell containing females by searching a neighborhood of 440 m radius, discounting pollen input by distance between the source and target cells as follows:

$$D_p = \beta e^{-\chi d}$$

where D_p is the scaling factor for distance between source and target cells, *d* is distance between mates, and ß and χ are means of parameters fit by nonlinear regression to two observed distributions of pollen (Slavov *et al.*, 2009). This resulted in a very large number of searches for our full landscape: 1.5 x 10^9 cells with 200,000 female *Populus* cells. In an effort to minimize the number of calculations, we created links among all *Populus* cells on the landscape and devised a search algorithm to minimize the number of cells processed. Nevertheless, the pollen dispersal algorithm accounts for the vast majority of processing time for this model.

Wind

We simulated the effects of wind speed and direction by using the scalar product between the prevailing wind vector and the vector between the source and target cells, scaled by a factor representing wind speed:

$$W = \delta - \varepsilon \cos\theta,$$

where δ and ε are parameters that vary between 0 and 1, and θ is the angle between the prevailing wind vector and the vector from source to target cells.

Phenology

Our phenology data indicated that the clones currently grown in plantations in the Pacific Northwest flowered earlier on average than sympatric native trees, but there was still extensive overlap in flowering. Others have reported substantial discontinuity between native and plantation-grown *Populus* in phenology of flowering and seed dispersal, particularly in more severe climates (U.S.Environmental Protection Agency, 1999; B. J. Thomas, University of Alberta, pers. comm.).

Therefore, we have devised a flexible method for simulating the effects of phenological discontinuities on pollination. The first step is to define the number of phenology classes (*e.g.*, early, middle, late) supported by data. The number of classes, together with an incompatibility factor, determine the degree of phenological compatibility between trees of different classes:

$$H = \frac{\left(N_p - \phi\right) - \left|p_1 - p_2\right|}{N_p - \phi}$$

where N_p is the number of phenology classes, p_i is the phenology score (range 1 to N_p) of clone *i*, and ϕ is an incompatibility parameter, set for the entire population between 0 and N_p-1. Values of *H* of 0 or less indicate complete incompatibility. This technique allows simulation of a broad gradation of intercompatibility. Furthermore, plantation and wild trees can be assigned different phenology distributions, allowing simulation of phenological discontinuities. For our simulations we examined three distributions of phenology each for plantation and wild trees: observed distributions, uniform distributions (equal number in each class) and nonoverlapping distributions (all plantation trees in early class, all wild trees in late class(es). For each set of distributions, we tested cases with two and three phenology classes equally divided over the flowering period, and Φ = 1 and 2. This resulted in a range spanning complete compatibility between plantations and wild trees, and complete incompatibility due to disjunct flowering.

Pollination

Pollination success of each genotype is a function of the relative quantity of compatible pollen arriving in the target cell. Alternatively, this can be viewed as a function of the cumulative probability of pollination by that genotype, summed over all cells on the landscape. This can be a function of distance, phenological overlap, and/or direction. Pollen input (or probability of pollination) of each genotype *g* from each cell *i* in the neighborhood of the target cell is calculated as

$$P_{gi} = P_g D_p W H.$$

The proportion of seeds of each genotype is determined from the proportional representation in the pollen arriving in the cell (see seed production equation above). A certain proportion of the pollinations (γ) derive from local trees in the neighborhood, and the remainder from distant trees (the 'background pollen cloud', estimated as 0.5 from our gene flow studies). Therefore, for *i* cells in the neighborhood, and a landscape with *j Populus* cells, the proportion of compatible pollen of genotype g arriving at the cell is:

$$K_g = \gamma \frac{\sum_i P_{gi}}{\sum_g \sum_i P_{gi}} + (1-\gamma) \frac{\sum_j P_{gj}}{\sum_g \sum_j P_{gj}}.$$

Seed Dispersal

We estimated seed dispersal both by direct measurements of seed movement, and by using maternity analysis on seedlings and seeds captured in the vicinity of plantations.
We performed direct measurements by setting seed traps at various distances from isolated wild *Populus* trees, and from the edge of isolated blocks of hybrid plantations. These traps consisted of 0.25 m^2 wire mesh coated with "Stikem Special" adhesive (Seabright Laboratories, Emeryville, CA), and mounted on 1 m wooden posts. We counted and removed seeds once every 7 days for two weeks.

We also collected seeds in the vicinity of plantations using mesh bags suspended on a wooden frame. We emptied traps twice weekly, germinated seeds, and identified hybrids based on leaf morphology, and Random Amplified Polymorphic DNA (Welsh & McClelland, 1990). We also used microsatellites (Morgante & Olivieri, 1993) to identify specific mother trees, and calculated dispersal distances.

Vegetative Dispersal

We had little direct data on vegetative dispersal for this study, so we chose to rely on a retrospective assessment of clone size as an integrated index of successful vegetative dispersal distances. We estimated genetic distances among ramets by the total difference in estimated microsatellite allele size for 10 or more loci. We analyzed these data using the Unweighted Pair Group Method with Arithmetic mean (UPGMA; Sneath & Sokal, 1973) to identify putative clones, accounting for somatic mutations and scoring errors within a clone (Tuskan *et al.*, 1996). We confirmed these identifications with field observations of phenology, sex, and morphology. We then calculated the minimum and mean distance between individual ramets and all other ramets of the clone.

In the model, seeds are dispersed explicitly only in the direct vicinity of establishment sites. We fit a negative exponential curve to the frequency distribution of dispersal distances, and seed

and vegetative dispersal are implemented much like pollen dispersal, except the default neighborhood is 220 m rather than 440 m. Local seed dispersal is also subject to the influence of wind, so that seed input from each source cell, i, is:

$$S_{gi} = S_g D_s W$$

where S_g is seed production for genotype g (defined above), D_s is the relationship between seed dispersal and distance, and W is the effect of wind (same as for pollen).
In addition, total production of seeds of each genotype is tallied on the landscape to allow for a 'background' seed cloud.

Input of vegetative propagules from each source cell, i, only depends on distance:

$$V_{gi} = V_g D_v .$$

where D_v is the relationship between vegetative propagule dispersal and distance.

Methods S4 Establishment and competition

Density and Extent

Data on *Populus* establishment come primarily from experimental plots we established in the vicinity of *Populus* plantations at two sites, Columbia and Willamette (DiFazio *et al.* 1999). We cleared 1 m^2 plots in the vicinity of competing vegetation and monitored input of *Populus* seeds, and emergence and growth of seedlings over two years. We repeated the experiment for two years at the Columbia site and three years at the Willamette site. However, despite weekly supplemental watering, we only observed substantial establishment in the first year of the study, 1996, which was the year of a large flood and elevated water tables. Such episodic establishment is to be expected for this species (Braatne *et al.*, 1996; Auble & Scott, 1998; Bradley & Smith, 1986; Scott *et al.*, 1997; Stromberg, 1997).

Propagule Type

Analogous to pollination, establishment of each genotype is determined primarily by proportional representation in propagules dispersed to the cell. However, the situation is more complicated for establishment because there are several propagule types: locally

Table S3 Frequency and extent of cottonwood clones in the wild for a variety of studies.

Site	Ramets	Genets	Clonality (%)[1]	Distance Between Ramets (m)	Long Distance (%)[2]	Source
Lower Nisqually	62	61	2	-	-	McKay 1996
Lower Cowlitz	35	29	17	-	-	McKay 1996
Upper Nisqually	55	46	16	-	-	McKay 1996
Upper Cowlitz	52	43	17	0-77	-	McKay 1996
Lethbridge	194	29	85	1.2-9.2	0	Gom and Rood 1999
Yakima 1(lower)	51	46	10	<10	0	Reed 1995
Yakima 2(lower)	48	48	0	<10	0	Reed 1995
Yakima 3(upper)	18	32	36	<10	0	Reed 1995
Yakima 4(upper)	25	27	48	<10	0	Reed 1995
Beaugency (nigra)	118	114	3	<5	0	Legionnet et. al. 1997
Oldman River	57	43	75	-	~1	Rood *et al.* 1994
Fraser R., Similkameen R.	Many	?	~5%	-	<1	Galloway and Worrall 1979
Davidson	104	57	45.2	-	-	Unpublished data
Willamette	287	221	31.4	0.6-76	6.6	Present study
Clatskanie	94	57	45.7	2.8-98	20	Present study
River Ranch	54	45	17	-	-	Present study

[1]Clonality is the percentage of ramets in the stand that had at least one other identical ramet in the stand.
[2]Potential long-distance dispersal. Includes ramets separated by 10 m or more (Nisqually, Cowlitz, Willamette, Clatskanie), or ramets resulting from small propagules (Oldman,Fraser/Similkameen), which presumably could have resulted from long-distance dispersal.

produced seeds, background seeds, and vegetative propagules. Unfortunately, data on the relative success of these different propagules are generally lacking. However, we were able to derive estimates of vegetative versus seedling establishment based on inferences from existing clone structure, and we can infer likely ranges of long-distance seed dispersal based on propagule characteristics and expectations for physical dispersal.

We estimated rates of vegetative versus seedling establishment at 4 sites as

$$V = 1 - \frac{G}{R}$$

where G is the number of genets and R is the total number of ramets.
We also derived similar estimates for a variety of published studies on *P. trichocarpa* and the ecologically similar species *P. nigra* (Table S3).

From the perspective of this simulation model, vegetative establishment directly adjacent to the parent tree is treated as vegetative growth, because this occurs within a cell, which is the unit of simulation. Therefore, we differentiated long-distance (> 10 m) from local vegetative dispersal in calculating rates of vegetative establishment (Table S3).

There have been few direct studies of long-distance seed dispersal by wind in trees (i.e., beyond 500 m) (Clark *et al.*, 1998; Higgins & Richardson, 1999), and none in *Populus*. Furthermore, the genetic data generated for this study were intended primarily for studying pollen dispersal and for differentiating plantation from wild seed sources, and they are largely inadequate for tracking long-distance seed flow. *Populus* seeds are tiny (~0.4 mg dry weight, personal observation) with plumed appendages that facilitate dispersal by both wind and water (Braatne *et al.*, 1996). However, our observations of local seed dispersal indicated that seeds do not move as far as pollen, and we can therefore expect that background seed input will be less than background pollen input.

Establishment of each genotype in a neighborhood of *i* cells on a landscape with *j* cells is proportional to its representation in each of the propagule pools, as follows:

$$L_g = RL_{\max}\left(\kappa\frac{\sum_i S_{gi}}{\sum_g\sum_i S_{gi}} + \lambda\frac{\sum_j S_{gj}}{\sum_g\sum_j S_{gj}} + \mu\frac{\sum_i V_{gi}}{\sum_g\sum_i V_{gi}}\right)$$

where R is a uniform random variate from 0 to 1 (the same for each genotype), L_{max} is the maximum density of establishment (2000 seedlings/100 m^2 by default), and κ, λ, and μ are

the proportions of local seed, background seed, and vegetative propagules in the established cohort. These proportions sum to one.

The random variate accounts for absolute differences in propagule input, variation in interspecific competition, and differences in habitat suitability.

Competition and Mortality

Data on density-dependent mortality came primarily from the same experimental plots that provided data on establishment. We followed the fate of seedlings over two to three year periods in 29 plots at two sites (DiFazio *et al.* 1999). In addition, we drew upon data from a total of 18 field studies in which growth and survival data were gathered for transgenic trees and controls. These trials were intended to assess the effects of traits such as resistance to the herbicides glyphosate and glufosinate, leaf beetle resistance (cry3A), floral sterility, expression of a putative broad-spectrum disease resistance gene (bacterio-opsin), and expression of a variety of selectable and visible marker genes (Strauss *et al.*, 2001b; Meilan *et al.*, 2000a; Meilan *et al.*, 2000b; Meilan *et al.*, 1999). These studies provided growth and survival data under selectively neutral conditions (with respect to the transgenic trait), and, where appropriate, under a selective regime that should favor the transgenics (*e.g.*, with herbicide spraying for herbicide resistant trees, and with insect attack for insect-resistant trees).

The relative competitiveness of transgenics controls the rate at which transgenic and nontransgenic trees die during density-dependent mortality (*i.e.*, self-thinning). Competitive effects of transgenes are simulated through effects on size (see basal area calculation), and density-dependent mortality. This is similar to the Lotka-Volterra equation for two-species interactions (*e.g.*, (Shugart, 1998; MacArthur & Levins, 1967), except the competitive differential of one genotype is the exact opposite of that of the alternate genotype. In addition to having strong direct effects on competition and mortality, the competitiveness parameter indirectly affects seed and pollen production, which depend on basal area.

Mortality of conventional trees is

$$M_0 = N_0\left(\left(\frac{N_0 + (1+\alpha)N_1}{N\,max_a}\right) - 1\right)$$

and mortality of transgenics is

$$M_1 = N_1\left(\left(\frac{N_0(1-\alpha) + N_1}{N\,max_a}\right) - 1\right)$$

where α is a competitive differential, Ng is the number of trees of genotype g that are present in the cohort, and $Nmax_a$ is the carrying capacity of a cell for age a (previously defined). α is the relative difference in growth or resource acquisition of transgenic trees relative to average trees (or half the difference between transgenic and conventional trees). This value can be fixed at initiation, and/or altered in response to selective pressures such as insect herbivory (see below).

There is no theoretical maximum value for α, but effective maxima and minima are determined by the rate at which one of the genotypes declines to 0 in the cohort. Density-dependent mortality occurs until a single tree occupies the cell, and that tree persists until eliminated by stochastic disturbance (see disturbance section). If transgenic and conventional trees decline to 0 in the same year, we randomly select a genotype to occupy the cell.

Sex ratio is recalculated each year for each genotype, and equal numbers of male and female trees die initially. When there are 100 or fewer trees, we determine the number of males of each genotype that will die by sampling from a binomial distribution in which each dying tree has a probability of 0.5 of being male. This allows the chance development of skewed sex ratios. ovule production, pollen viability, pollen tube growth rates, and embryo development. Also, density-dependent mortality following germination is quite intense in the early years of the cohort, and competitive effects can be almost immediately apparent. Therefore, differences in establishment ability can be incorporated into the density-dependent mortality competition coefficient.

Methods S5 Sensitivity analysis

One of the key features of the model is it allows us to perform sensitivity analyses, which consist of 'virtual experiments' in which we vary individual parameters, holding all else constant at reasonable values, and assess changes in model outcomes (Haefner, 1996). Sensitivity analyses require a very large number of model runs, and therefore substantial computing power. To speed processing and allow more model runs, we devised a test landscape which contained many of the features of the real landscape, but on a much smaller scale (full landscape = 46,000 ha, test landscape = 2,500 ha). Important characteristics such as habitat types and *Populus* edge:interior ratios were similar between the test landscape and the full landscape. For the bulk of the sensitivity analyses, we explored a scenario with a very high component of *Populus*

plantations on the landscape, 50% of which were transgenic. In addition, we assessed transgene flow on a landscape containing a single small plantation (19 ha), which simulates a large, isolated field trial.

Using the test landscapes, we analyzed the effects of 30 variables and a variety of interactions, analyses that required more than 8000 model runs of 50 years each. For the sensitivity analyses we set all parameter values at a default baseline condition (Table S4). These conditions are hypotheses, and the sensitivity analyses depict the consequences of deviating from this condition. We varied selected parameters individually and assessed changes in model outcomes for 10 repetitions of each scenario. Each repetition was initialized with a different random number, so stochastic processes such as establishment and mortality varied in space and time for each repetition.

The response for the sensitivity analyses was the 'mean area of mature transgenics', which is the percentage of area of mature *Populus* occupied by transgenic trees outside of plantations (*i.e.*, wild *Populus* populations). This response was averaged over the final 25 years of the run to dampen interannual variability and simplify presentation of results. Time trends for individual scenarios were generally concordant with trends of the 25 year means. We chose this response value on the assumption that the area of mature transgenics best represents ecologically significant gene flow.

Table S4 Baseline conditions for sensitivity analyses.

Parameter	**Value**
Landscape Area	2500 ha (500 x 500 pixels)
Plantation Area	485 ha (19.4% of landscape)
Transgenic Plantations	240 ha (9.6% of landscape)
Plantation Rotation	12 years, even aged
Plantation Density	1500 trees/ha
Plantion Sex Ratio	50% female, by area
Transgenic Fertility	1 (fully fertile)
Transgenic Competitive Advantage	0
Initiation of flowering, plantations	5 years
Initiation of flowering, wild	10 years
Pollen Dispersal	Neighborhood, 440m,50%; slope, -0.007; intercept, 0.67;
Seed Dispersal	Neighborhood, 220m,90%; slope, -0.05; intercept, 0.9;
Vegetative Dispersal	Neighborhood, 220m,100%; slope, -0.1; intercept, 0.6;
Phenology Classes	1 (all trees fully compatible)
Wind	No influence of wind on dispersal
Maximum Establishment Density	20 seedlings/m^2
Density-Dependent Mortality	Slope, -0.33
Basal Area Increase	Slope, 0.04
Establishment Rates	Empirical
Maximum Establishment, Agricultural Edges	0.05 ha

Young transgenic trees may be destined to be eliminated by competition before they attain a stature that would allow significant ecosystem effects. For most model runs, this response reached an apparently stable equilibrium within the 50 year runs, indicating that it is a good indicator of long-term levels of transgene flow under modeled conditions. It is also important to note that the grain of the STEVE model, 100 m^2 cells, is rather coarse in that multiple trees may occupy a single cell until the age of 25. Therefore, the area of mature transgenics can be misleading because a cell is considered 'occupied' even if transgenics represent a minority of the trees present in the cell. In fact, in a typical scenario with neutral transgenes, transgenic trees constituted a minority of trees in most cells, and representation declined with age. Most of the established transgenics derived from background seed flow (91%), as demonstrated by the long distances between transgenic cohorts and plantations. These background transgenics largely disappeared from the landscape as they were eliminated by density-dependent mortality, and older transgenics tended to be closer to plantations. Therefore, the influence of transgenics would likely be minor for much of the area we are considering to be 'occupied' by transgenics with our transgene flow estimate. Basal area is probably a more accurate measure of the potential influence of transgenics in the wild, because this integrates age (size) and density. However, basal area was linearly related to mean area of mature transgenics across a broad range of values. Therefore, our measure of transgene flow reveals the same general trends as basal area.

Variation of individual parameters and selected combinations provides a great deal of information about the factors controlling transgene flow in our model. However, it is likely that many of these parameters interact, such that changing the value of one parameter alters the effects of other parameters. We therefore sought to explore the effects of the most important parameters in a fractional factorial experiment. Fractional factorial experiments allow exploration of main effects and selected levels of interactions by using higher level interactions as aliases for the level of interest, indicated by the resolution of the design. This allows exploration of interactions of a large number of factors with a modest number of scenarios. We performed a resolution V fractional factorial with 11 factors at two levels per factor. A resolution V design allows discrimination of all main effects and two-factor interactions. This required 128 scenarios, with 5 repetitions per scenario (Box *et al.* 1978). We performed this analysis for a landscape with commercial-scale transgenic plantations and a transgenic field trial. We chose factors that had a substantial effect on gene flow in single-factor analyses (competitiveness,

disturbance, fertility, phenology, plantation reproductive maturity, and distant seed establishment), or for which interactions with important parameters were deemed likely (plantation sex, vegetative establishment, vegetative dispersal). We chose two levels for each parameter, representing extremes of a reasonable expected distribution of values (Table S5). Where possible, we chose the lowest value of a parameter that caused a substantial response in single-factor analyses (*i.e.*, where the response began to reach an asymptote). Significance was assessed based on standard F-statistics ($P < 0.05$ of observing a larger F-value).

The fractional factorial analysis was potentially susceptible to experimenter bias because only two levels were examined for each factor, and the levels were chosen somewhat subjectively to reflect reasonable values that could have a strong influence on gene flow. We explored the robustness of the results by repeating the experiment with factors varied +/- 20% from our best estimates of parameter values (Table S5).

Table S5 Parameter values used in fractional factorial analyses. 'Biological Range' was selected based on biologically reasonable upper and lower estimates for parameters. Factors were varied +/- 20% from initial parameter estimates for 'Objective Range'.

Parameter	Abbrev.	Biological Range		Objective Range	
		Lower	Upper	Lower	Upper
Fertility	FER	0.01	1	0.4	0.6
Transgenic Competitiveness	COMP	0.99	1.2	0.8	1.2
Plantation Maturity	PLANT	5	8	4	6
Phenology	PHEN	25% compat.[a]	65% compat.	52% compat.	82% compat.
Disturbance	DIST	1x(-15%)[b]	3x (+15%)	1x(-15%)	3x (+15%)
Distant Pollination[c]	POLCLD	10%	50%	40%	60%
Distant Seed Establishment[d]	SDCLD	1%	10%	8%	12%
Vegetative Establishment	VEG	1%	40%	8%	12%
Vegetative Dispersal Slope	VGD	-0.05	-0.1	-0.08	-0.12
Plantation Sex[e]	SEX	0	0.5	0.4	0.6
Rotation	ROT	8	12	8	12

[a] Phenology is expressed as compatibility with wild trees relative to a case with one phenology class (complete overlap in flowering among all trees)
[b] Disturbance rates were selected based on changes in wild *Populus* populations. Empirical disturbance rates resulted in a 15% reduction in *Populus* populations over a 50 year simulation. Enhancing establishment 3-fold resulted in a 15% increase in wild *Populus* over 50 years.
[c] Distant pollination is proportion of seeds that are fathered by nonlocal males (determined by total proportion of pollen produced on landscape)
[d] Distant seed establishment is proportion of seedlings derived from nonlocal seeds
[e] Ratio of male to female plantation blocks

Methods S6 Risk assessment simulations

We explored probable levels of gene flow under a range of scenarios using a landscape that represents a large area of poplar cultivation on the lower Columbia River (36.8 km x 23.0 km). We also simulated a landscape with more extensive potential poplar habitat in upland sites, as would be found in northern Washington and British Columbia. Simulations were initiated with our best estimates for parameter values, but with substantial stochasticity incorporated to reflect natural variation and uncertainty (Table S6). We allowed stochastic variation in fertility, competitiveness, pollen flow, seed flow, vegetative establishment, and disturbance, all of which were identified as important in the sensitivity analyses. Little information was available on variation in these parameters, so we generally used a standard deviation equivalent to 50% of the mean value, as described below.

We simulated cultivation of transgenics with neutral fitness effects in the wild and various levels of fertility. In addition, we explicitly simulated fitness effects of two of the most prominent transgenic traits in forestry: insect resistance and herbicide resistance.

We performed 30 repetitions of most risk assessment scenarios and calculated the mean and 99% confidence interval of the response (transgene flow) through time.

Stochasticity in Fertility

Fertility was varied in two ways: base fertility was varied among new poplar cells and plantation blocks to reflect differences among genotypes. In addition, fertility was varied annually to reflect environmental influences (*e.g.*, weather, interspecific competition). Annual variation can be cumulative (*i.e.*, changes persist from year to year), or independent (fecundity is reset to base value annually). Annual variation is not allowed for completely sterile trees, but fecundity can be restored with a probability and rate determined by the user, at which point annual variation may ensue.

Table S6 Baseline conditions for risk assessment analyses.

Parameter	**Value**
Landscape Area	46,000 ha (2287 x 3681 pixels)
Plantation Area	2,348 ha (19.4% of landscape)
Transgenic Plantations	1,200 ha ((9.6% of landscape)
Plantation Rotation	12 years, even aged
Plantation Density	1,100 trees/ha
Plantation Sex Ratio	50% female, by area
Transgenic Fertility	0.5
Transgenic Competitive Advantage	0
Initiation of flowering, plantations	5 years
Initiation of flowering, wild	10 years
Pollen Dispersal	Neighborhood, 440m,50%; slope, -0.007; intercept, 0.67;
Seed Dispersal	Neighborhood, 220m,90%; slope, -0.05; intercept, 0.9;
Vegetative Dispersal	Neighborhood, 220m,100%; slope, -0.1; intercept, 0.6;
Phenology Classes	3 (72% compatible)
Wind	No influence of wind on dispersal
Maximum Establishment Density	20 seedlings/m^2
Density-Dependent Mortality	Slope, -0.33
Basal Area Increase	Slope, 0.04
Establishment Rates	Empirical
Maximum Establishment, Agricultural Edges	0.05 ha

Stochasticity in Competitiveness

Variation in the transgenic competitive differential occurs simultaneously in space and time to reflect influences of factors such as weather, microsite, and interspecific interactions. The competitive deviate is sampled from a normal distribution according to a user-defined standard deviation.

Stochasticity in Disturbance

We varied rates of disturbance annually by sampling transition probabilities from a negative exponential distribution, with a mean determined by the empirically determined transition rates, as described above. This mimics the natural process of poplar establishment, which consists primarily of rare bursts of establishment in response to large-scale disturbances (primarily due to flooding) (*e.g.*, (Braatne *et al.*, 1996).

Pollination and Establishment

We expect that gene flow by pollen, seeds, and vegetative propagules will vary spatially due to effects of local populations, weather, topography, and other factors not explicitly considered in the model. We therefore allowed random variation in rates of background pollination for each female. Similarly, we allowed variation in proportion of establishment from local seed, background seed, and vegetative propagules for each establishment site. All parameters were sampled from normal distributions with user-defined standard deviations.

Methods S7 Insect resistance scenario

Poplar trees are susceptible to attack by a wide variety of herbivores, including leaf-eating chrysomelid beetles (Dickmann & Stuart, 1983). One of the scenarios that we examined was the cultivation of insect-resistant transgenics (*e.g.*, trees containing the Bt endotoxin gene, which has been used to create beetle-resistant poplar trees: Meilan *et al.* 2000). We sought to examine how mean growth enhancements determined in field trials would translate to actual fitness advantages and transgenic gene flow, given variation in insect pressure and growth enhancement in the wild.

We obtained data on insect resistance of transgenic hybrid poplars containing a gene encoding a modified version of the Cry3a endotoxin from Bacillus thuringiensis (Bt). The growth advantage was approximately 13% for transgenic versus conventional clones of the same

genotype (Meilan *et al.*, 2000b). These studies were carried out under nearly ideal moisture and nutrient conditions, but with high insect pressure. These trials therefore provide idealized estimates of the short-term advantage that this transgene might confer on trees growing in the wild (Strauss *et al.*, 2001a).

We simulated insect attack in the wild as a stochastic process that affected patches of trees up to 0.5 ha in size. Creation of an insect attack patch occurred with a predetermined probability, and was independent between years (*i.e.*, insects disappeared and reappeared annually at random locations). This resulted in a range of insect pressures on the landscape. The attack was manifested as the relative growth advantage of a transgenic tree at that site of attack, and this value was determined by sampling from a normal distribution of transgenic insect resistance. Random variation within patches represents some combination of local spatial variation in intensity of insect attack, variation in other biotic (*e.g.*, fungal pathogens, other herbivores) and abiotic stresses (*e.g.*, desiccation, shading, inundation), that might limit response to protection from insects, and variation in innate resistance of conventional trees (James & Newcombe, 2000). We simulated scenarios with a range of insect pressures and mean transgenic advantages, assuming a standard deviation in transgenic advantage equivalent to 50% of the mean advantage.

References

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

Auble GT, Scott ML. **1998.** Fluvial disturbance patches and cottonwood recruitment along the upper Missouri River, Montana. *Wetlands* **18:** 546-556.

Braatne JH, Rood SB, Heilman PE. **1996.** Life history, ecology, and reproduction of riparian cottonwoods in North America. In: Stettler RF, Bradshaw HD, Jr., Heilman PE, Hinckley TM, eds. *Biology of Populus and its implications for management and conservation*. Ottawa, Canada: NRC Research Press, 57-85.

Bradley CE, Smith DG. **1986.** Plains cottonwood recruitment and survival on a prairie meandering river floodplain, Milk River, southern Alberta and northern Montana. *Canadian Journal of Botany* **64:** 1433-1442.

Clark JS, Fastie C, Hurtt G, Jackson ST, Johnson C, King GA, Lewis M, Lynch J, Pacala S, Prentice C, Schupp EW, Webb TI, Wyckoff P. **1998.** Reid's paradox of rapid plant migration: dispersal theory and interpretation of palaeoecological records. *Bioscience* **48:** 13-24.

Cowardin LM, Carter V, Golet FC, LaRoe ET. **1979.** *Classification of wetlands and deepwater habitats of the United States*. Washington, D.C.: Office of Biological Services, U.S.D.I. Fish and Wildlife Service.

DeBell DS. **1990.** *Populus trichocarpa* Torr. & Gray, Black Cottonwood. In: Burns RM, Honkala BH, eds. *Silvics of North America Vol. 2. Hardwoods. USDA For. Serv. Agric. Handbook 654*. Washington D.C: USDA Forest Service, 570-576.

Dickmann D, Stuart K. **1983.** *The Culture of Poplars in Eastern North America*. East Lansing, MI: Michigan State University.

Fierke MK, Kauffman JB. **2005.** Structural dynamics of riparian forests along a black cottonwood successional gradient. *Forest Ecology and Management* **215:** 149-162.

Frelich LE, Calcote RR, Davis MB. **1993.** Patch Formation and Maintenance in an Old-Growth Hemlock-Hardwood Forest. *Ecology* **74:** 513-527.

Greene DF, Johnson EA. **1999.** Modelling recruitment of Populus tremuloides, Pinus banksiana, and Picea mariana following fire in the mixedwood boreal forest. *Canadian Journal of Forest Research* **29:** 462-473.

Haefner JW. **1996.** *Modeling Biological Systems: Principles and Applications*. New York: Chapman and Hall.

Higgins SI, Richardson DM. **1999.** Predicting plant migration rates in a changing world: The role of long-distance dispersal. *American Naturalist* **153:** 464-475.

James RR, Newcombe G. **2000.** Defoliation patterns and genetics of insect resistance in cottonwoods. *Canadian Journal of Forest Research* **30:** 85-90.

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

Meilan R, Han KH, Ma C, James RR, Eaton JA, Stanton BJ, Hoien E, Crockett RP, Strauss SH. **2000a.** Development of glyphosate-tolerant hybrid cottonwoods. *Tappi Journal* **83:** 164-166.

Meilan R, Ma C, Eaton JA, Miller LK, Crockett RP, DiFazio SP, Strauss SH. **2000b.** High levels of Roundup and leaf-beetle resistance in genetically engineered hybrid cottonwoods. In: Blatner KA, Johnson JD, D.M.Baumgartner, eds. *Hybrid Poplars in the Pacific Northwest: Culture, Commerce, and Capability*. Pullman, WA: Washington State University, 29-38.

Meilan R, Ma C, Strauss SH. **1999.** Stability of herbicide resistance and GUS expression in transgenic hybrid poplars during several years of field trials and vegetative propagation. *Hortscience* **34:** 557.

Morgante M, Olivieri AM. **1993.** PCR-amplified microsatellites as markers in plant genetics. *Plant Journal* **3:** 175-182.

Scott ML, Auble GT, Friedman JM. **1997.** Flood dependency of cottonwood establishment along the Missouri River, Montana, USA. *Ecological Applications* **7:** 677-690.

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

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.

Sneath PHA, Sokal RR. **1973.** *Numerical taxonomy; the principles and practice of numerical classification*. San Francisco, W. H. Freeman.

Stanton BJ, Villar M. **1996.** Controlled reproduction of Populus. In: Stettler RF, Bradshaw HD, Jr., Heilman PE, Hinckley TM, eds. *Biology of Populus and its implications for management and conservation*. Ottawa, Canada: NRC Research Press, 113-138.

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

Strauss SH, Meilan R, DiFazio SP, Brunner AM, Carson J. **2001b.** *Tree Genetic Engineering Research Cooperative (TGERC) Annual Report:2000-2001*. Corvallis, Oregon: Forest Research Laboratory, Oregon State University.

Stromberg JC. **1997.** Growth and survivorship of fremont cottonwood, Goodding willow, and salt cedar seedlings, after large floods in central Arizona. *Great Basin Naturalist* **57:** 198-208.

Tuskan GA, Francis KE, Russ SL, Romme WH, Turner MG. **1996.** RAPD markers reveal diversity within and among clonal and seedling stands of aspen in Yellowstone National Park, U.S.A. *Canadian Journal of Forest Research* **26:** 2088-2098.

U.S.Environmental Protection Agency. 1999. *Biological Aspectsof Hybrid Poplar Cultivation on Floodplains in Western North America- A Review.* Seattle, WA, USA: Environmental Protection Agency. EPA Document No. 91-R-99-002.
Ref Type: Serial (Book,Monograph)

Welsh J, McClelland M. **1990.** Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Research* **18:** 7213-7218.