

How wind can shape genetic variation and adaptability in metapopulations of wind-pollinated forest trees through its effect on pollen dispersal

Elizabeth M. Gillet*

Abstract

In wind-pollinated tree species, wind conditions during pollination are a primary determinant of mating. The oft patchy distribution of trees over spatially heterogeneous environments suggests a metapopulation approach to the study of wind effects on pollen contributions to seeds. A mathematical model of distance-dependent pollen dispersal among five populations on a transect yields distributions of pollen sources, and thus their potentially adaptive genetic characteristics, in the seeds of each population. For scenarios of symmetric and directional dispersal over variable mean dispersal distance, structural characteristics of the seeds are quantified by effective number of pollen sources within populations and compositional differentiation of pollen contributions among populations. These measures avoid the problems of the fixation index F_{ST} , which confounds diversity and difference. Effects of wind depend on location. The population with the highest effective number of pollen sources in its seeds need not be the most representative nor even central. Symmetric dispersal over long distances can be statistically indistinguishable from random dispersal, but directional dispersal is not. Consequences for adaptation and practical implications for forest management and seed harvesting are discussed.

Keywords: *Pollen dispersal, wind-pollination, forest tree, adaptation, metapopulation model, diversity, effective number, compositional differentiation*

Zusammenfassung

Wie Wind genetische Variation und Anpassungsfähigkeit in Metapopulationen windbestäubter Waldbäume durch Auswirkung auf die Pollenverbreitung beeinflussen kann

Bei windbestäubten Baumarten sind Windverhältnisse eine wesentliche Determinante von Paarungen. Die oft fragmentierte Verteilung von Bäumen über räumlich heterogene Umweltbedingungen legt einen Metapopulationsansatz nahe, um die Auswirkungen des Windes auf die Pollenbeiträge zu Samen zu untersuchen. Ein mathematisches Modell der entfernungsabhängigen Pollenverbreitung zwischen fünf Populationen an einem Transekt ergibt die Häufigkeitsverteilung von Pollenquellen mit ihrer potentiell adaptiven genetischen Information in den Samen jeder Population. Für Szenarien von symmetrischer bzw. gerichteter Verbreitung über variable mittlere Distanz werden strukturelle Eigenschaften der Samen durch eine effektive Anzahl von Pollenquellen in Populationen und der kompositionellen Differenzierung zwischen Populationen quantifiziert. Diese Maße sind frei von der Problematik des Fixierungindexes F_{ST} , welcher Vielfalt mit Differenz vermischt. Wind wirkt sich ortsabhängig aus. Die Population mit der höchsten effektiven Anzahl von Pollenquellen muss weder die repräsentativste sein noch im Zentrum liegen. Symmetrische Verbreitung über weite Distanzen kann von zufälliger Verbreitung statistisch ununterscheidbar sein, gerichtete Verbreitung jedoch nicht. Konsequenzen für Anpassung und praktische Implikationen für die Waldbewirtschaftung und die Saatgutgewinnung werden diskutiert.

Schlüsselwörter: *Pollenverbreitung, Windbestäubung, Waldbäume, Anpassung, Metapopulationsmodell, Diversität, effektive Anzahl, kompositionelle Differenzierung*

* Georg-August-Universität Göttingen, Buisen-Institut – Forstgenetik und Forstpflanzenzüchtung, Buisenweg 2, 37077 Göttingen, Germany

1 Introduction

In wind-pollinated plants, meteorological conditions during pollination are primary determinants of mating, in that they influence how many pollen from any one plant are available to fertilize the ovules of any other plant. Depending on wind direction, horizontal wind velocity, and turbulence but also on temperature and humidity (including washout by rain) after pollen emission (Prentice, 1985; Niklas, 1985; Jackson and Lyford, 1999; Helbig et al., 2004), wind-borne pollen can settle near source or be carried long distances. Tree pollen has been shown to be transported over distances of tens to hundreds of kilometers without losing viability (Williams, 2010; Buschbom et al., 2011; Kremer et al., 2012).

In spatially heterogeneous environments, the genetic variants transferred with the migrating pollen may therefore encounter different environmental conditions than those at its source. The adaptive genetic information (i.e., alleles at gene loci involved in adaptation) that it contributes to the seeds of a tree in a different environment may be detrimental or perhaps beneficial for survival of the offspring that descend from it. Between pollination seasons, variation in wind velocity and direction creates differences in the sources of pollen and thus the adaptive genetic information represented in each year's seeds. Wind thus shapes the spatial and the temporal distribution of adaptive genetic information in seeds, giving it considerable influence on the adaptability of forests to environmental change. This is especially true in temperate latitudes, where most tree species are wind-pollinated.

Since conspecific trees often occur in environmentally more or less uniform patches, and since these patches may give rise to the formation of subpopulations, mating can be characterized as the pattern of effective pollen dispersal from (sub)population to (sub)population. This calls for adaptive interpretation at the level of a metapopulation, which is defined as a set of partially isolated populations connected by gene flow (Levins, 1969). Populations of any size can form a metapopulation, even single trees. Extending metapopulation analysis beyond the basic consideration of colonization/extinction of habitats, the focus here is on the relative contributions of pollen from different sources to the seeds produced by each population and by the metapopulation as a whole. This metapopulation concept is reminiscent of that frequently conceived in community ecology as a set of local communities of potentially interacting species (here pollen of different sources) that are linked by species dispersal (here pollen dispersal) (see Liebold et al., 2004).

Each population in a metapopulation harbors a unique genetic structure (alleles, genotypes) at each gene locus that is involved in adaptation. Its alleles at each such locus contribute via pollen dispersal to the seeds of other populations, just as other populations contribute their alleles via pollen dispersal to the seeds of this population. Populations may or may not share alleles at a locus, so that pollen dispersal may change the frequency of the constituent alleles or even introduce new alleles. Since it is practically impossible to trace the dispersal of alleles at all of the presumably large number of adaptive gene loci among the populations, it is meaningful in

metapopulation analysis to label each pollen with its source population as a synopsis of the complex adaptive genetic characteristics that have allowed the trees in this population to survive to maturity. Labelling of pollen by source population specifies a genetic trait within the set of all pollen produced in the metapopulation, where the trait state stands for the specific genetic structure of the source population.

The adaptive potential in a population's seeds is the range of environments in which a sufficient number of offspring can survive to support a viable population (Gregorius, 1977; Bergmann et al., 1990). This is primarily determined by the spectrum of environments in the source populations of the pollen contributions in interaction with the local environment of the population's ovule contributions. In changing environments, a large adaptive potential might be a safeguard for survival (adaptive reserve), but under constant environmental conditions, a high proportion of pollen from environments that differ from a population's local environment endangers survival (genetic load, migration load). Whether a sufficient number of offspring will survive to maintain the population depends on the distribution of environments around it, the degree of differences among the environments, and the direction and distance of pollen dispersal, all of which may vary over space and time.

In population genetics, the paradigm model of mating is panmixis. The panmixis hypothesis rests on three assumptions: (a) the same trait is viewed in both gametic sexes, (b) the frequency distribution of trait states is the same in both sexes (sexual symmetry), and (c) gametes fuse at random. If the trait state is the genotype at a diploid gene locus, panmixis yields Hardy-Weinberg proportions in the offspring. In the metapopulation framework of mating among populations, panmixis amounts to (a) consideration of the source population as a trait (state) for pollen as well as for ovules, (b) assumption that the distribution of source populations among the pollen produced within the metapopulation is the same as among the ovules, and (c) stochastic independence of pollen and ovule sources in the zygotes of the metapopulation. Panmixis is thus realized on the metapopulation level only if the relative contribution of each population through pollen equals its relative contribution through ovules to all zygotes, and in the absence of zygotic selection, to all seeds formed in the metapopulation. Even though it is frequently addressed as random dispersal, in a more strict sense this situation refers to equal mating success for pollen sources within populations. True random dispersal would, however, require special mechanisms of distance-independent pollen dispersal that are physically difficult to construe. It may, however, be possible that certain patterns of pollen dispersal approximate random dispersal so closely that statistical testing based on limited samples of the seeds cannot discriminate between them (Spasojevic et al., 2014). This may explain the frequent failure to detect deviation from the hypothesis of panmixis using real sampling data. One question is whether long distance dispersal can realize true random dispersal. Model characteristics that allow for such situations will be given special attention in order to allow for realistic interpretation of these observations.

The vagaries of wind defy general analytical treatment of mating for wind-pollination. By simulating wind conditions in models, possible effects of wind on the pattern of pollen dispersal among populations, and ultimately on adaptive potentials, can be demonstrated. To this end, a probabilistic model is presented that accepts wind direction and mean dispersal distance as the input variables that shape the dispersal of pollen beyond their source populations to other populations. By applying the model to scenarios of populations lined along a transect, a setting that is representative of a wide range of real situations (e.g. roadside, stream, lakeshore, valley), pollen dispersal need be considered in only one dimension. This makes it easier to separate the effects of wind direction and pollen dispersal distance from other factors that play a role in two-dimensional systems. Effects of wind direction and wind intensity are demonstrated for scenarios of symmetric (non-directional) and asymmetric (directional) pollen dispersal paired with increasing dispersal distance along the transect. Random mating on the metapopulation level provides a reference scenario. Under simple assumptions on gamete production and fertilization, the output of the model is the pollen source distribution in the seeds produced within each population and in the metapopulation as a whole.

The outcome of pollen dispersal is commonly analyzed in terms of amounts of gene flow among populations and the potential implication of preferential matings within or between these populations, as referenced above. Numerous experimental studies apply paternity analysis to the seeds of single mother trees in order to estimate the relative contributions of pollen donors (backward migration) or the distribution of pollen dispersal distances, including wind direction (Smouse et al., 2001; Austerlitz et al., 2004; Burczyk et al., 2004; Vranckx et al., 2014). In a wider evolutionary context, what ultimately counts are the effects of the mating system on providing potentials for adaptation, including general limits set by the available genetic variation.

The present paper therefore focuses on outlining the structural effects of different modes of wind-mediated pollen dispersal on the genetic variation within populations on a transect. It is shown that quantification of the distribution of pollen sources among populations by a different set of measures than those commonly used helps us understand how wind affects pollen source variation of populations along a transect. These measures set quantitative limits to the adaptive potential of populations when the metapopulation is spread over heterogeneous environments, showing how wind can shape adaptability to different types of environmental variation, whether static, gradual, or unstructured (chaotic).

Basically, the diversity, or effective number, of pollen sources in the seeds of a population represents the spectrum of pollen sources within the metapopulation. The compositional differentiation among populations reflects differences, or specializations, in the distribution of pollen sources in seeds between populations. The degree to which differentiation is absent corresponds to the homogeneity of the populations, allowing comparison with results for

metacommunity models showing that dispersal has a homogenizing effect on species composition (Loreau et al., 2003; Mouquet and Loreau, 2003; Münkemüller et al., 2011). The commonly used fixation index F_{ST} or G_{ST} (Wright, 1978; Nei, 1973) is included to show how it confounds the effects of variation (polymorphism) within populations with differences between populations, making it less useful for prediction of adaptive potential.

2 Methods

2.1 The model

Consider $n = 5$ (sub)populations P_1, \dots, P_5 lined along a transect far from other trees of the same species. The populations are of equal length $L = 100$ by some arbitrary measure (e.g. meter, decameter) and separated by treeless gaps of the same length to simulate fragmentation. The reason for division of the transect into equal stretches of populations and gaps is its suitability for analysis of the effects of partial isolation and pollen loss on the one hand and connectivity between populations via pollen dispersal on the other hand on within-population diversity. All five populations together are regarded as a metapopulation connected by pollen dispersal between populations. Seeds are heavy and do not disperse.

Model inputs are wind direction and mean dispersal distance (a composite of horizontal wind speed, turbulence, and humidity) during a single pollination season. Pollen produced within a population is carried by wind rightwards along the transect (e.g. towards populations of higher number) with probability p_r or leftwards with probability $p_l = 1 - p_r$.

Dispersal distance is governed by a dispersal kernel, i.e., the probability density function (pdf) of the distance $x \geq 0$ that pollen travels in either direction before settling. Equal dispersal distance in both directions is a simplification that may be realized, for example, by daily changes in convection up and down mountains or valleys.

A number of different dispersal kernels have been proposed and tested using data from paternity analysis on seeds or by studying the physical movement of small particles (Clark, 1998; Austerlitz et al., 2004; Katul et al., 2005; Kuparinen et al., 2007). Common to all is the property that the probability of reaching a certain distance is a decreasing function of this distance. This decrease is initially steep for small dispersal distance x and becomes more gradual as x increases, asymptotically approaching 0. These smooth functions neglect the possibility of occasional bulk dispersal of pollen from the same source to a spatially confined area (Lanner, 1966) and they may basically differ from dispersal kernels for non-wind-pollinated species, especially for insect pollination (see e.g. Gregorius et al., 2011). The frequent claim that they are limited to mean dispersal distances of not more than a few hundred meters may partially result from the infeasibility of identifying more distant pollen donors within the radially expanding pool of potential donors for a sufficient numbers of seeds.

The (negative) exponential distribution in its one-dimensional form

$$\theta_\gamma(x) = \frac{1}{\gamma} \exp\left(\frac{-x}{\gamma}\right)$$

is applied in this model, because it specifies the standard dispersal kernel against which other kernels are interpreted as deviations. The distance parameter γ in the same unit of length as x equals the mean (or expected) dispersal distance and defines the rate of decrease as $-1/\gamma$. The pdf θ_γ is unique in that it is the only distribution with a constant rate of decrease over distance (memorylessness).

Given the direction parameters p_r and $p_l = 1 - p_r$ and the kernel θ_γ with its mean distance parameter, and assuming that pollen production is uniformly distributed along the length of each population, the probability m_{ij} that pollen produced in P_i settles in P_j (forward migration rate) equals

where $q_{ij} = |i - j| \cdot 2L$ (see Figure 1).

Since pollen can settle outside of the bounds of the populations, the proportion $1 - \sum_j m_{ij}$ of the pollen produced by population P_i is lost, i.e., excluded from mating a priori. Dispersal is well-behaved in the sense that the amount of pollen distributed from one population to another decreases with the distance between them.

In order to model the relative contributions of the source populations through pollen to the seeds of a given population, additionally assume that all populations produce the same number of pollen and the same number of ovules. Then the post-dispersal frequency of pollen from source P_i in P_j equals $p_i(j) = m_{ij} / (\sum_k m_{kj})$ (backward migration rate). Assuming further that pollen production is sufficient to ensure the (non-selective) fertilization of all ovules (with high probability), the frequency of seeds produced by P_j whose pollen was contributed by source P_i also equals $p_i(j)$. The outcome of the model is the pollen source distribution

$$m_{i,j} = p_l \cdot \frac{\gamma}{L} \cdot \left(\exp\left(\frac{-q_{ij} + L}{\gamma}\right) - 2 \exp\left(\frac{-q_{ij}}{\gamma}\right) + \exp\left(\frac{-q_{ij} - L}{\gamma}\right) \right), \text{ if } j < i$$

$$m_{i,j} = \frac{\gamma}{L} \cdot \left(\exp\left(\frac{-q_{ij} + L}{\gamma}\right) - 2 \exp\left(\frac{-q_{ij}}{\gamma}\right) + \exp\left(\frac{-q_{ij} - L}{\gamma}\right) \right), \text{ if } j = i$$

$$m_{i,j} = p_r \cdot \frac{\gamma}{L} \cdot \left(\exp\left(\frac{-q_{ij} + L}{\gamma}\right) - 2 \exp\left(\frac{-q_{ij}}{\gamma}\right) + \exp\left(\frac{-q_{ij} - L}{\gamma}\right) \right), \text{ if } j > i$$

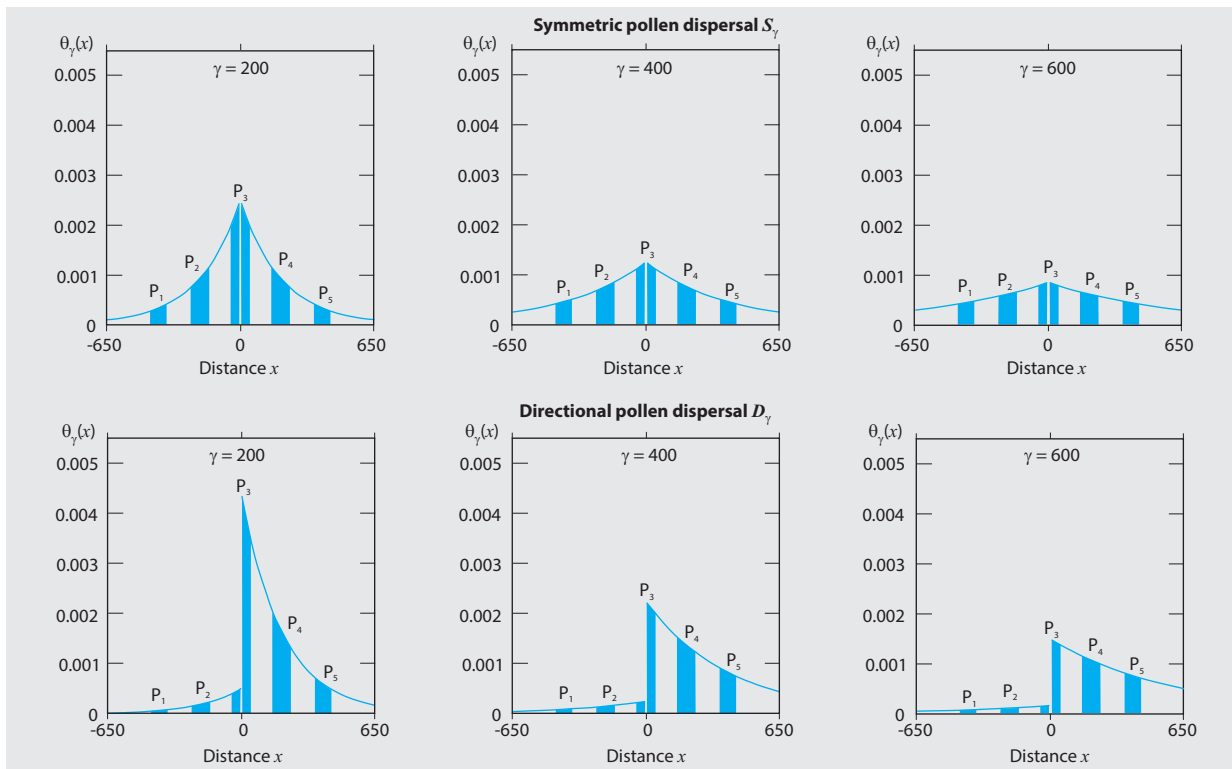


Figure 1

Dispersal of a single pollen: Probability density function $\theta_\gamma(x)$ for pollen that is produced at the center of P_3 to settle at a distance x in either direction under Scenarios S_γ and D_γ . The total area below the curve, including the two tails that extend to infinity, equals 1. The area of the shaded region below P_j equals the probability that such pollen settles in P_j . The area of a white region is proportional to the probability that the pollen settles in the gap between the two neighboring populations.

$(p_1(j), \dots, p_5(j))$ in the seeds produced by each of the five populations.

The forward migration rates $m_{i,j}$ are invariant to changes in the unit of length or to a proportional change in γ and L that leaves the ratio γ/L constant (scale invariance). Increasing L but not γ would reduce the $m_{i,j}$, i.e., the dispersal among populations. By appropriate choice of the dispersal distance γ and population length L and the unit of length in which they are measured, the model can simulate metapopulations at any spatial scale for which the underlying dispersal kernel and the assumptions can be considered meaningful.

2.2 Scenarios

Symmetric dispersal S_γ Pollen is given equal chances $p_r = p_l = 0.5$ of dispersing in either direction along the transect over distances following the pdf θ_γ . For selected mean dispersal distances $\gamma \in \{200,400,600\}$, (forward) dispersal probabilities $m_{i,j}$ and the (backward) pollen source distributions $(p_1(j), \dots, p_5(j))$ within the seeds produced by each of the five populations are listed in Table 1 and plotted as horizontal percent bar charts in Figure 2c. Interestingly, the symmetry of dispersal yields $m_{i,j} = p_j(i)$, that is, the probability that pollen of P_i disperses to P_j equals the probability that a seed of P_j has its pollen contribution from P_i .

Table 1

Pollen dispersal probabilities, pollen loss, and pollen source distributions in seeds for the random dispersal scenario R and the symmetric dispersal scenarios S_γ for $\gamma \in \{200,400,600\}$. The pollen dispersal probability (forward dispersal, left side) is shown as the standardized probability $m_{i,j} / \sum_k m_{i,k}$ that pollen produced in P_i (row head) and not lost will settle in P_j (column head) for $i,j = 1, \dots, 5$ (see Figure 2). Loss equals the probability $1 - \sum_k m_{i,k}$ that it does not settle within the bounds of any population. Assuming equal pollen and equal ovule production over populations and non-selective fertilization of all ovules, the pollen source frequencies $p_j(i)$ (backward dispersal, right side) in the seeds of population of P_i (row head) equal $m_{i,j} / \sum_k m_{i,k}$ for each source population P_j (column head). Note the equality of forward and backward dispersal.

R	Probability of pollen dispersal from P_i to P_j					Loss	Frequency of pollen from P_j in seeds of P_i				
	$j = 1$	2	3	4	5		$j = 1$	2	3	4	5
$i = 1$	0.200	0.200	0.200	0.200	0.200	0.000	0.200	0.200	0.200	0.200	0.200
2	0.200	0.200	0.200	0.200	0.200	0.000	0.200	0.200	0.200	0.200	0.200
3	0.200	0.200	0.200	0.200	0.200	0.000	0.200	0.200	0.200	0.200	0.200
4	0.200	0.200	0.200	0.200	0.200	0.000	0.200	0.200	0.200	0.200	0.200
5	0.200	0.200	0.200	0.200	0.200	0.000	0.200	0.200	0.200	0.200	0.200
S_{200}	$j = 1$	2	3	4	5		$j = 1$	2	3	4	5
$i = 1$	0.594	0.262	0.096	0.035	0.013	0.641	0.594	0.262	0.096	0.035	0.013
2	0.210	0.476	0.210	0.077	0.028	0.552	0.210	0.476	0.210	0.077	0.028
3	0.073	0.200	0.454	0.200	0.073	0.530	0.073	0.200	0.454	0.200	0.073
4	0.028	0.077	0.210	0.476	0.210	0.552	0.028	0.077	0.210	0.476	0.210
5	0.013	0.035	0.096	0.262	0.594	0.641	0.013	0.035	0.096	0.262	0.594
S_{400}	$j = 1$	2	3	4	5		$j = 1$	2	3	4	5
$i = 1$	0.408	0.270	0.163	0.099	0.060	0.717	0.408	0.270	0.163	0.099	0.060
2	0.223	0.337	0.223	0.135	0.082	0.658	0.223	0.337	0.223	0.135	0.082
3	0.128	0.212	0.320	0.212	0.128	0.640	0.128	0.212	0.320	0.212	0.128
4	0.082	0.135	0.223	0.337	0.223	0.658	0.082	0.135	0.223	0.337	0.223
5	0.060	0.099	0.163	0.270	0.408	0.717	0.060	0.099	0.163	0.270	0.408
S_{600}	$j = 1$	2	3	4	5		$j = 1$	2	3	4	5
$i = 1$	0.337	0.255	0.183	0.131	0.094	0.766	0.337	0.255	0.183	0.131	0.094
2	0.220	0.290	0.220	0.158	0.113	0.728	0.220	0.290	0.220	0.158	0.113
3	0.151	0.210	0.278	0.210	0.151	0.716	0.151	0.210	0.278	0.210	0.151
4	0.113	0.158	0.220	0.290	0.220	0.728	0.113	0.158	0.220	0.290	0.220
5	0.094	0.131	0.183	0.255	0.337	0.766	0.094	0.131	0.183	0.255	0.337

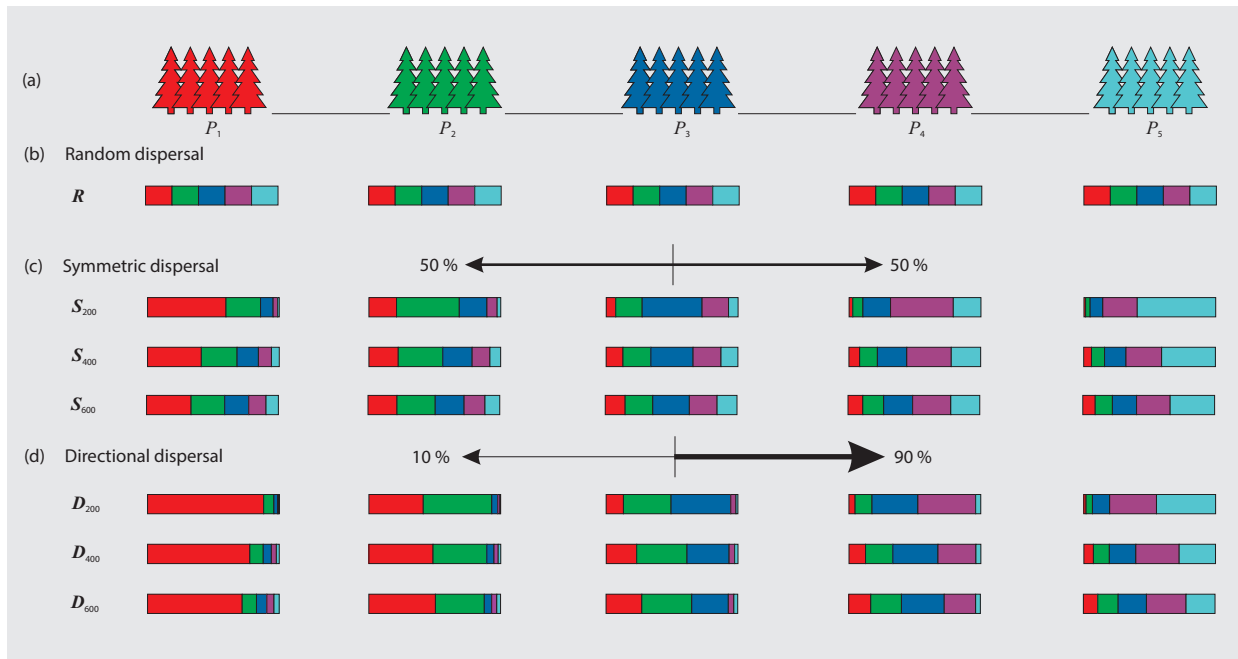


Figure 2
 Distribution of pollen sources in seeds: (a) Diagram of the transect showing five populations of equal length $L = 100$ separated by gaps of the same length. (b to d) For each scenario and each population P_i ($j = 1, \dots, 5$), the colored horizontal histogram of total length 1 shows the relative contribution $m_{j,i}$ of each of the five populations P_j as pollen sources to the seeds of P_i (e.g. the length of the red segment is proportional to the relative frequency of seeds with pollen from the “red” population P_1). (b) Random dispersal scenario R . (c) Symmetric dispersal scenario S_γ (50 % leftwards: 50 % rightwards) for mean dispersal distances $\gamma \in \{200, 400, 600\}$. (d) Directional dispersal scenario D_γ (10 % leftwards: 90 % rightwards) for $\gamma \in \{200, 400, 600\}$.

Table 2

Pollen dispersal probabilities, pollen loss, and pollen source distributions in seeds for the directional dispersal scenarios D_γ for $\gamma \in \{200, 400, 600\}$ (see Table 1 for explanation). Note the inequality of forward and backward dispersal.

	Probability of pollen dispersal from P_i to P_j					Loss	Frequency of pollen from P_j in seeds of P_i				
	$j = 1$	2	3	4	5		$j = 1$	2	3	4	5
D_{200}											
$i = 1$	0.448	0.355	0.131	0.048	0.018	0.524	0.880	0.078	0.028	0.010	0.004
2	0.039	0.438	0.348	0.128	0.047	0.514	0.412	0.519	0.046	0.017	0.006
3	0.015	0.040	0.453	0.360	0.132	0.530	0.132	0.360	0.453	0.040	0.015
4	0.006	0.017	0.046	0.519	0.412	0.590	0.047	0.128	0.348	0.438	0.039
5	0.004	0.010	0.028	0.078	0.880	0.758	0.018	0.048	0.131	0.355	0.448
D_{400}											
$i = 1$	0.276	0.329	0.200	0.121	0.073	0.583	0.775	0.102	0.062	0.038	0.023
2	0.038	0.287	0.342	0.207	0.126	0.599	0.486	0.408	0.054	0.033	0.020
3	0.026	0.042	0.320	0.381	0.231	0.640	0.231	0.381	0.320	0.042	0.026
4	0.020	0.033	0.054	0.408	0.486	0.718	0.126	0.207	0.342	0.287	0.038
5	0.023	0.038	0.062	0.102	0.775	0.851	0.073	0.121	0.200	0.329	0.276
D_{600}											
$i = 1$	0.220	0.300	0.215	0.154	0.110	0.641	0.717	0.109	0.078	0.055	0.040
2	0.036	0.238	0.325	0.233	0.167	0.669	0.505	0.370	0.056	0.040	0.029
3	0.030	0.042	0.277	0.379	0.271	0.716	0.271	0.379	0.277	0.042	0.030
4	0.029	0.040	0.056	0.370	0.505	0.787	0.167	0.233	0.325	0.238	0.036
5	0.040	0.055	0.078	0.109	0.717	0.890	0.110	0.154	0.215	0.300	0.220

Directional dispersal D_r : Prevailing winds in the rightward direction (i.e., towards populations with higher numbers) are modelled by giving pollen the probability $p_r = 0.9$ of dispersing rightwards and $p_l = 0.1$ of dispersing leftwards. The dispersal distance in either direction follows θ_r . Table 2 shows that m_{ij} and $p_j(i)$ are no longer equal. Figure 2d shows the horizontal percent bar charts of $(p_1(j), \dots, p_5(j))$.

Random dispersal R : In spatially less explicit population genetic models, pollination is assumed to occur at random, contributing equal proportions of pollen from all pollen sources to the seeds of all populations, i.e., $p_i(j) = \frac{1}{5}$ for all populations P_i and P_j . Random dispersal is included here as a reference scenario, since it is still a baseline hypothesis against which other models are tested based on actual data.

2.3 Measures of variation and their meaning

Structural characteristics of the distribution of pollen sources in the seeds produced within the populations of a metapopulation are quantified using the following measures (see classification of Gregorius and Gillet, 2015). Each seed is characterized by its “population” (i.e., the population of its seed parent), and its pollen source (i.e., the population of its pollen parent) or “type”.

2.3.1 Effective number of pollen sources (types)

The adaptedness of the seeds produced by a population to a constant environment and its adaptive potential to changing environments are dependent on the spectrum of adaptive variation that it is contributed through pollen by the populations. The amount of pollen source variation (i.e., pollen source diversity) within the seeds of a population is measured here as the effective number of populations that contributed pollen according to Rényi diversity $v(j) = (\sum_i p_i(j))^2$ of order 2 (or Hill number 2D, where $p_i(j)$ is the frequency of pollen from P_i in the seeds of P_j , $v(j)$ ranges from its minimum of 1 for monomorphism (i.e., fixation of the seeds to a single pollen source) to its maximum of five for equal contributions $p_i(j) = \frac{1}{5}$ of all five source populations P_i . The Rényi-diversity of order 2 equals its own effective number of pollen sources and also the effective number by the Simpson diversity $(1 - \sum_i p_i(j)^2)$ (Gregorius, 1991).

The effective number of pollen sources in the seeds of the metapopulation, i.e., the aggregate of the seeds produced by the five populations, is expressed by the Rényi diversity $v_T = (\sum_i p_i^2)^{-1}$ (T stands for total), where the frequency p_i of pollen of source P_i in the metapopulation's seeds equals $\sum_j c_j p_i(j)$ and c_j is the relative frequency of seeds carrying pollen of P_j . Here $c_j = \frac{1}{5}$ holds due to the equality of the numbers of ovules and their complete fertilization in each population.

2.3.2 Compositional differentiation of pollen sources

Wind-dispersal of pollen creates differences between populations in the distribution of pollen sources in their seed production and thus in the spectrum of adaptive variants. Compositional differentiation measures the tendency for seeds produced by different populations to have pollen contributions from different sources or, inversely, for seeds with pollen contributions from the same source population to have been produced in the same population (Gregorius, 2010; Gregorius, 2014; Gregorius and Gillet, 2015). Differentiation is complete if the pollen sources among the seeds produced by the different populations are disjoint, and it is absent if all populations have the same distribution of pollen sources in their seeds (Gregorius and Roberds, 1986). In its “pure” form, compositional differentiation should be independent of levels of diversity (i.e., effective number) within the populations (Gregorius and Gillet, 2015). A measure that considers only the frequency differences between the pollen sources in the seeds, and not their diversities, is the complementary compositional differentiation

$$\delta_{SD} = \sum_{j=1}^n c_j \delta_{SD}(j)$$

(Gregorius and Roberds, 1986), where

$$\delta_{SD}(j) = \frac{1}{2} \sum_{i=1}^n |p_i(j) - \bar{p}_i(j)|$$

is the absolute difference between the pollen source distribution $(p_1(j), \dots, p_n(j))$ in the seeds of population P_j and the pollen source distribution $(\bar{p}_1(j), \dots, \bar{p}_n(j))$ in this population's complement, i.e., the aggregate of the seeds produced by all other populations. The frequency of pollen from P_i in the complement of the seeds of P_j equals $\bar{p}_i(j) = \sum_{k,k \neq j} c_k p_i(k) / (1 - c_j)$. Thus δ_{SD} equals the mean difference between the pollen source distributions in the seeds of each population and in their complements. The range of δ_{SD} stretches from its minimum of 0 for identical pollen distributions in all populations to its maximum of 1 when no two populations share pollen from the same population. In other words, $\delta_{SD}(j)$ measures the uniqueness of the seed production of P_j and $1 - \delta_{SD}(j)$ measures its representation for the others. δ_{SD} measures the mean uniqueness or disjunction of the seeds produced among the populations and $1 - \delta_{SD}$ the homogeneity.

2.3.3 Fixation index F_{ST} or G_{ST}

Because of its wide application in the analysis of (sub)population structure and as an overview of results that are obtainable for indices of fixation versus indices of compositional differentiation, a brief account of this index is provided.

The general term “differentiation” as it is commonly used in the literature is ambiguous. It is often interpreted to mean the difference in the distributions of types (here, pollen sources) among populations, but the most commonly used measures of “differentiation” actually quantify the degree of monomorphism of types within populations. This family of measures comprises various standardizations of the increase in Simpson diversity $H_T = 1 - \sum_i p_i^2$ of types in the metapopulation compared to the mean $H_S = 1 - \sum_j c_j p_j(j)^2$ of the Simpson diversities within the populations. The oldest member of this family measures this increase relative to H_T as

$$F_{ST} = G_{ST} = \frac{H_T - H_S}{H_T}$$

(Wright, 1978; Nei, 1973). Its minimum $F_{ST} = 0$ is assumed when the pollen source distributions in the seeds of all populations are identical, which suggests that it might measure “difference”, but at the other extreme, its maximum $F_{ST} = 1$ is reached only when the seeds of all populations are monomorphic for pollen from only one source (though not from the same source in all populations, in which case F_{ST} is undefined). As a consequence, F_{ST} is maximal even if all but one of a large number of populations are fixed for the same type, making them identical, and only the remaining population is fixed for a second type. As pointed out by Gregorius and Roberds (1986), this contradicts the common interpretation of F_{ST} as a measure of the “difference” among populations. This was recognized by Sewall Wright, who introduced F_{ST} in the context of his Island Model of population genetics as a *fixation index*, that is, a measure of “differentiation within the

total array in the sense of the extent to which the process of fixation has gone towards completion” (Wright, 1978, p. 82). F_{ST} is “thus not a measure of degree of differentiation in the sense implied in the extreme case by absence of any common allele”; i.e., it does not measure compositional differentiation. Moreover, Wright’s interpretation as a fixation index assumes fulfillment of the highly idealistic assumptions of the Island Model. As opposed to δ_{SD} , F_{ST} depends heavily on the diversities within populations, decreasing rapidly towards 0 as the degree of polymorphism increases within populations (Hedrick, 1999; Gregorius et al., 2007).

The fact that F_{ST} approaches zero for high polymorphism within populations despite high differentiation between them as well as for low differentiation marks a salient point in interpretation of this index, and this will be given special emphasis in the present paper. Beyond this, other standardizations of $(H_T - H_S)$ may not solve this problem. G'_{ST} (Hedrick, 2005), for example, assumes its maximum for complete fixation, for complete compositional differentiation, or both, meaning that it does not discriminate between tendencies towards monomorphism and disjunction (Gregorius et al., 2007). Jost’s D is maximal only for complete disjunction, but its intermediate values are sensitive to the within-population diversity (Jost, 2008; Gillet, 2013).

2.3.4 Comparison of compositional differentiation and fixation

Both F_{ST} and δ_{SD} assume their minimum of 0 when the types are identically distributed in all populations, independently of the number of types in relation to the number of

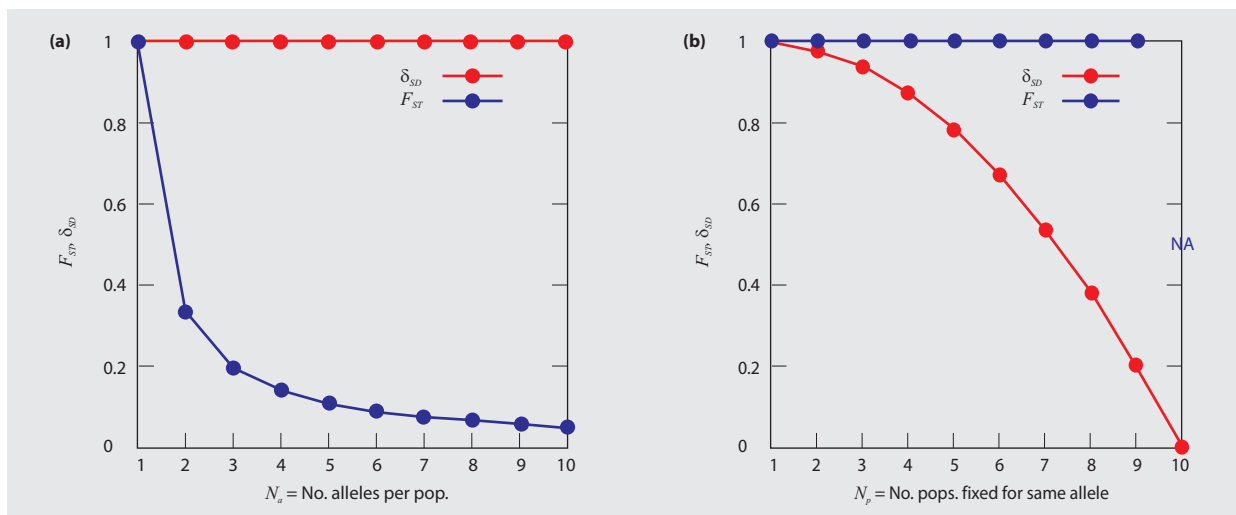


Figure 3

Examples of deviant characterizations of structural variation in metapopulations by two measures of variation among populations, the compositional differentiation δ_{SD} (red) and the fixation index F_{ST} (blue). (a) Two populations each possess N_a equally frequent alleles that are not shared between populations. For $N_a = 1$, both measures equal their maximum of 1. As the populations become more polymorphic with rising diversity $v_2 = N_a F_{ST}$, F_{ST} approaches its minimum of 0, while the non-sharing of alleles leaves the diversity-independent δ_{SD} at its maximum. (b) Ten populations are fixed for one allele, N_p of the populations for the same allele, the others for different alleles. For $N_p = 1$, both measures are maximal. As N_p increases, the decreasing difference of each population to its complement reduces δ_{SD} to 0 for $N_p = 10$, while the monomorphism of all populations leaves F_{ST} maximal for all N_p except 10, for which F_{ST} is undefined (“NA”).

populations. Complete fixation ($F_{ST} = 1$) and complete compositional differentiation ($\delta_{SD} = 1$), in contrast, occur simultaneously only when all populations are fixed for a different type, implying that the number of populations equals the number of types (here, pollen sources). Here it happens in the absence of dispersal ($\gamma = 0$), where all seeds contain local pollen only.

In general, F_{ST} and δ_{SD} show quite different behavior, demonstrating that fixation (commonly called “differentiation”) and compositional differentiation are very different concepts. At the extremes, fixation cannot be complete if there are more types than populations, and compositional differentiation cannot be complete if there are more populations than types. If, for example, each population is polymorphic for N equally frequent alleles that are not shared between populations, the diversity-dependence of F_{ST} pushes it towards 0 for increasing polymorphism, while δ_{SD} remains at its maximum of 1 (Figure 3a). On the other hand, if all populations are monomorphic and if an increasing number of populations are fixed for one particular allele, the decreasing disjunction between populations pushes δ_{SD} from 1 towards 0, while F_{ST} remains maximal but is undefined if all are fixed for the same allele (Figure 3b). For more rigorous treatment of the connection between fixation (or “apportionment”) and compositional differentiation, see the papers of Gregorius et al., 2007; Gregorius, 2010; 2014; 2016.

2.3.5 Adaptive potentials and migration load

Adaptedness to a constant environment: A high proportion $p_j(j)$ of local pollen ensures adaptedness in the current environment, since offspring containing local pollen contributions have a better chance of surviving than offspring containing pollen from most other environments. Pollen contributions from the other populations may confer poorer adaptation in a population’s environment, reducing the number of offspring that can survive. The potential migration load as the genetic load contributed via pollen dispersal to the offspring of P_i (i.e., the seed parent is in P_j) is measured as the proportion

$$p_{imm}(j) = 1 - p_j(j)$$

of immigrant pollen. For high potential migration load, the effective number of sources of immigrant pollen becomes important. A small effective number of sources bears the danger of flooding the populations with poorly adapted pollen from a few sources, while a higher effective number signifies a larger number of adaptive variants and thus a higher capacity for adaptation.

Adaptability to environmental change: In a changed environment, offspring descending from local pollen may have poor survival, while offspring containing pollen that immigrated from a population well adapted to the new environment would survive. The proportion $p_{imm}(j)$ of immigrant pollen is now a measure of the adaptive potential in the offspring. A high proportion is necessary for adaptability, though it is not sufficient to guarantee that pollen from

adapted populations will be among them. Nevertheless, a large effective number

$$v_{imm}(j) = \sum_{i,i \neq j} \left(\frac{p_i(j)}{\sum_{k,k \neq j} p_k(j)} \right)^2$$

of sources of immigrant pollen increases the chances that some of the pollen will come from populations that are adapted to the new environment, giving offspring the ability to survive. The adaptive potential of the offspring of P_i in the face of environmental change is expressed as a combination of $p_{imm}(j)$ and $v_{imm}(j)$.

3 Results

Application of the model to the scenarios yields the following characterizations of the pollen source distributions along the transect, both in the seeds of the populations and in the aggregate seeds of the metapopulation. Since Scenario R is the idealized situation to which the other scenarios are compared, its characterization is presented first.

3.1 Random dispersal R

Under the hypothesis of random dispersal, pollen from each population has equal probability of settling in each of the five populations, i.e., $m_{ij} = \frac{1}{5}$ for all i, j . Under the further assumptions of the model, i.e., equal pollen production and non-selective ovule fertilization, the frequency $p_i(j)$ of every pollen source P_i in the seeds produced by each population P_j equals $\frac{1}{5}$. The pollen source distributions are illustrated by the five evenly striped horizontal percent bar charts in Figure 2b. Metapopulation structure is characterized for random dispersal as follows:

- *Effective number:* The equal proportions of all pollen sources in the seeds of each population yield equal proportions among the seeds produced by the metapopulation, so that the effective number of pollen sources is the actual number of populations ($v_T = v(j) = 5$ for all P_j).
- *Compositional differentiation:* Compositional differentiation is absent ($\delta_{SD} = 0$), yielding complete homogeneity of the populations for the distribution of pollen sources in their seeds. The seeds of every population are completely representative of the seeds of the metapopulation ($\delta_{SD}(j) = 0$).
- *Fixation index:* F_{ST} equals 0 due to the complete homogeneity of seeds among populations.
- *Adaptive potential:* The proportion $p_{imm}(j) = \frac{4}{5}$ of immigrant pollen in each population is large, among which each of the other four populations is equally represented ($v_{imm}(j) = 4$).

3.2 Symmetric dispersal S_γ

The pollen source distributions among the seeds of the populations are shown for selected mean dispersal distances

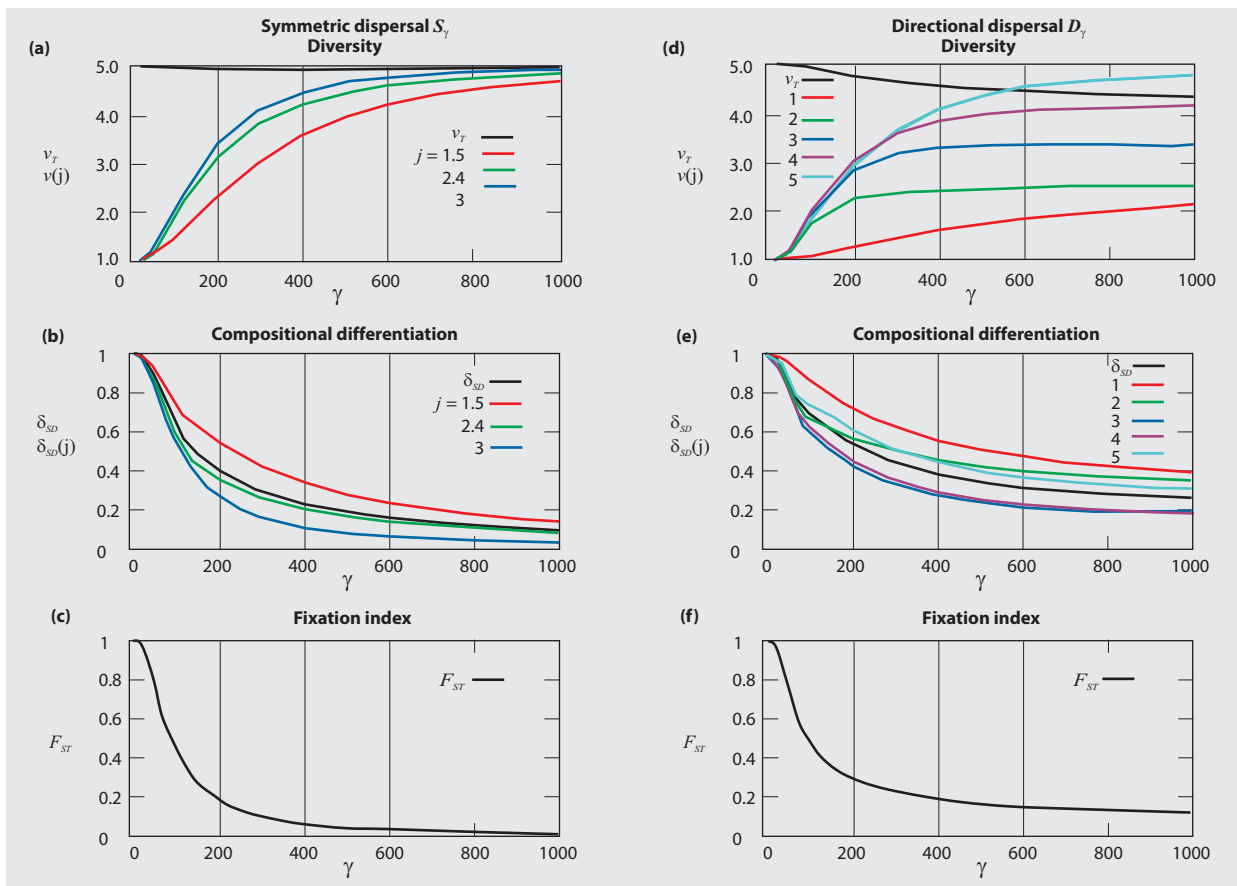


Figure 4
 Measures of variation of the pollen source distributions in the seeds for scenarios S_γ and D_γ for mean dispersal distance γ up to 1000, i.e., 10 times the length L of the populations and the gaps. Dotted lines mark the scenarios in Tables 1 and 2; the numbers in the legends refer to the populations P_j .

Table 3

Variation of pollen sources in the seeds of the populations and the metapopulation for the scenarios in Tables 1 and 2.

Scenario	Effective number of pollen sources in seeds						Pollen loss
	$v(1)$	$v(2)$	$v(3)$	$v(4)$	$v(5)$	v_T	
R	5.000	5.000	5.000	5.000	5.000	5.000	0.000
S_{200}	2.317	3.118	3.375	3.118	2.317	4.935	0.583
S_{400}	3.584	4.202	4.446	4.202	3.584	4.949	0.678
S_{600}	4.201	4.582	4.738	4.582	4.201	4.967	0.741
D_{200}	1.280	2.263	2.823	3.000	2.884	4.771	0.583
D_{400}	1.620	2.458	3.295	3.853	4.085	4.573	0.678
D_{600}	1.862	2.516	3.370	4.062	4.529	4.468	0.741
Scenario	Compositional differentiation among populations pollen contributions to seeds of variation					δ_{SD}	Partitioning F_{ST}
	$\delta_{SD}(1)$	$\delta_{SD}(2)$	$\delta_{SD}(3)$	$\delta_{SD}(4)$	$\delta_{SD}(5)$		
R	0.000	0.000	0.000	0.000	0.000	0.000	0.000
S_{200}	0.566	0.379	0.294	0.379	0.566	0.424	0.198
S_{400}	0.351	0.218	0.124	0.218	0.351	0.244	0.062
S_{600}	0.246	0.152	0.076	0.152	0.246	0.169	0.029
D_{200}	0.737	0.585	0.457	0.474	0.635	0.560	0.306
D_{400}	0.569	0.467	0.294	0.306	0.456	0.395	0.189
D_{600}	0.482	0.411	0.233	0.242	0.375	0.325	0.149

$\gamma \in \{200,400,600\}$ as bar charts in Figure 2c. Pollen source variation is characterized as follows (see Table 3 and Figures 2a to c and 4a):

- **Effective number:** Pollen source variation is unevenly distributed over the populations (Figure 2c). Within populations, the effective number of pollen sources in the seeds increases differentially from 1 (monomorphism) for $\gamma = 0$ towards the maximum of five with increasing distance of symmetric dispersal. The increase is steepest for short-distance dispersal and flattens out quickly as dispersal distance increases. The increase becomes faster when moving from the periphery to the center, giving the central population P_3 the highest effective number and both peripheral populations P_1 and P_5 the lowest. In the seeds of the metapopulation, the effective number v_T of pollen sources is only slightly lower than the maximum of five for all dispersal distances (Figure 4a).
- **Compositional differentiation:** Starting from complete differentiation ($\delta_{SD} = 1$) for $\gamma = 0$, pollen sources become more homogeneous as dispersal distance increases (Figure 4b, “snails” in Figure 5a). Differentiation seems to be approaching $\delta_{SD} = 0$ for longer dispersal distances ($\delta_{SD} = 0.053$ for $\gamma = 2000$). The pollen source distribution of the central population best represents the seeds of
- the metapopulation, i.e., $\delta_{SD}(3)$ is the smallest $\delta_{SD}(j)$ and is even lower than the average differentiation δ_{SD} among all populations.
- **Fixation index:** When there is no wind ($\gamma = 0$), all populations are initially monomorphic for pollen source, yielding $F_{ST} = 1$. As dispersal distance γ increases, F_{ST} approaches 0 (i.e., identical pollen source distributions in all populations) more rapidly than δ_{SD} (Figure 4c). The faster decrease of F_{ST} towards 0 is due to the rapid increase in diversity within the populations. The dependence of F_{ST} on within-population diversity implies that F_{ST} would start out being much smaller than 1 for $\gamma = 0$ if allelic variation within the pollen contributions to the seeds were taken into consideration instead of only the pollen source (see Figure 1).
- **Adaptive potential:** The proportion $p_{imm}(j)$ of pollen immigrating into the populations is initially absent for $\gamma = 0$ and increases toward $p_{imm}(j) = 0.8$, the value for random dispersal, as dispersal distance increases (Figure 6a). The peripheral populations have a slightly lower proportion of immigrant pollen than the others. The effective number $v_{imm}(j)$ of sources of immigrant pollen is undefined (non-existent) for $\gamma = 0$ and increases from an initial positive value of 1 for the peripheral populations P_1

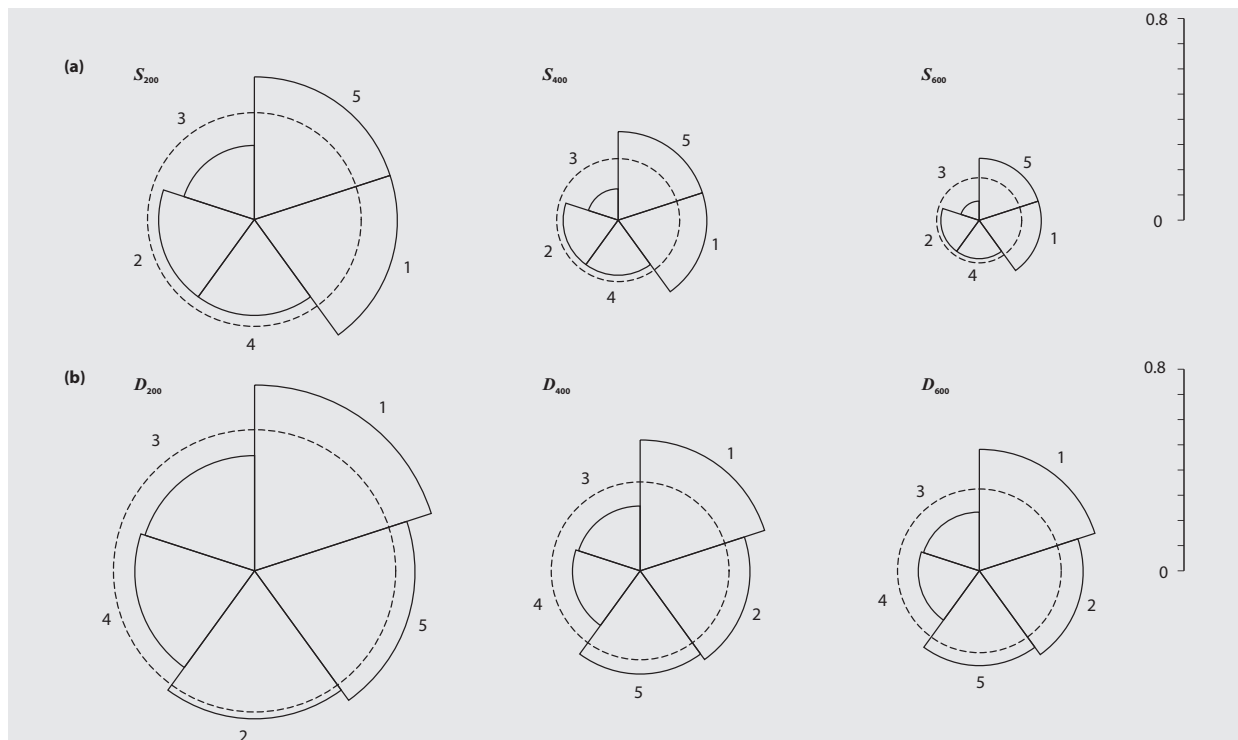


Figure 5

Snail diagrams of compositional differentiation δ_{SD} among the pollen source distributions of the seeds for three symmetric and three directional scenarios in Table 1 and 2 (after Gregorius and Roberds, 1986). The radius of the dotted circle is proportional to δ_{SD} . Population P_j is represented as a sector of radius proportional to $\delta_{SD}(j)$ and central angle $c_j = 2\pi/5$ marked by the number j . Populations are arranged in ascending order, starting with the population of smallest difference $\delta_{SD}(j)$ from its complement placed to the left of the top vertical axis and proceeding counterclockwise to the population of largest $\delta_{SD}(j)$. The smaller the snail, the more homogeneous are the pollen source distributions among populations. The population with the lowest $\delta_{SD}(j)$ has the most representative pollen source distribution, that with the highest $\delta_{SD}(j)$ has the most unique distribution. Since $\delta_{SD}(j) = 0$ holds for the random dispersal scenario R , its snail is a single point (not shown).

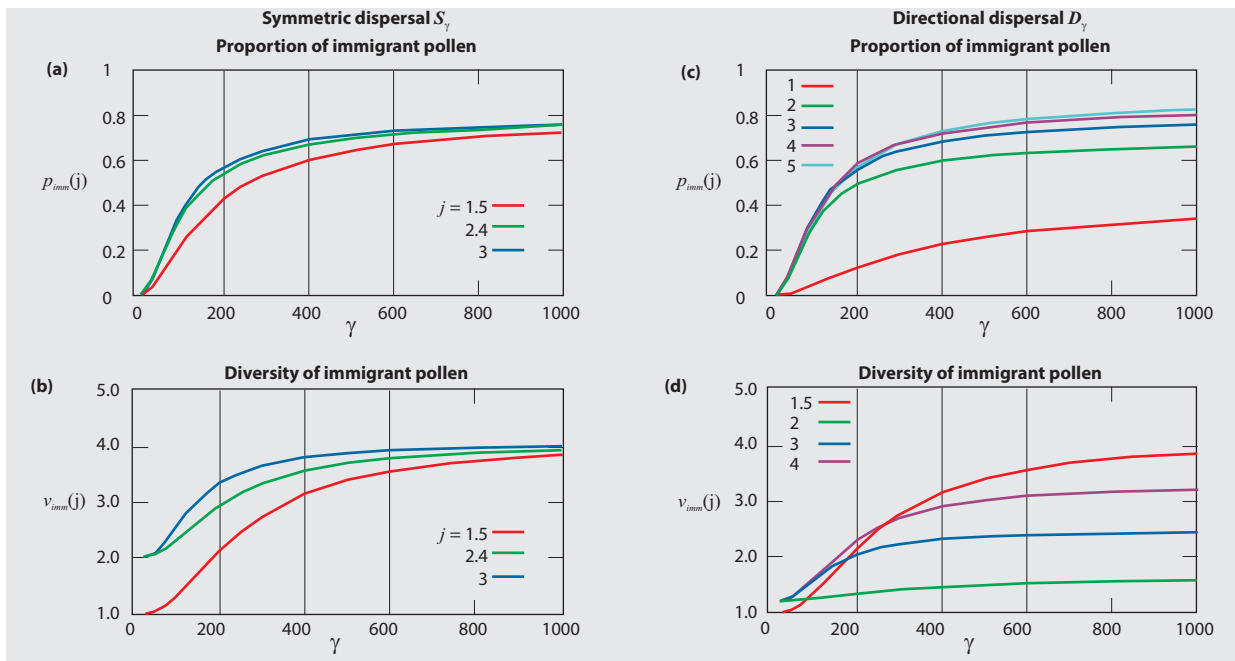


Figure 6

Adaptive potential: Proportion of immigrant pollen $p_{imm}(j)$ and its diversity $v_{imm}(j)$ in the seeds of each population P_j for scenarios S_γ and D_γ and mean dispersal distance γ up to 1000. The numbers in the legends refer to the populations P_j .

and P_5 and from 2 for the three interior populations as dispersal distance increases (Figure 6b). It remains highest in the central population and lowest in the peripheral populations. As dispersal distance increases, the effective number of immigrant pollen sources approaches 4 for all populations, as for random dispersal. The seeds of the peripheral populations P_1 and P_5 have the lowest potential migration load, since the proportion of immigrant pollen is smaller. The effective number of sources of immigrant pollen is also lower than for the interior populations. Both the migration load and the effective number of sources of immigrant pollen increase as dispersal distance increases.

3.3 Directional dispersal D_γ

The nine-fold chance (9:1) of dispersing rightwards, or downwind, along the transect yields a high level of asymmetry for pollen source distributions around the central population, as shown for $\gamma \in \{200, 400, 600\}$ in Figure 4d. Pollen source variation is characterized as follows (see Table 3 and Figures 2d to f and 4b):

- **Effective number:** As for symmetric dispersal, the pollen source variation is unevenly distributed over the populations (Figure 2d). The effective number v_j of pollen sources within populations increases in prevailing wind direction from its lowest level in the leftmost, or upwind, population P_1 to its highest level in the rightmost, or downwind, population P_5 (Figure 4d). Only the downwind population's effective number approaches the maximum of 5, and even exceeds the effective number v_γ of pollen sources in the metapopulation. In the

metapopulation, the effective number v_γ of pollen sources is lower than under symmetric dispersal and actually decreases with increasing dispersal distance (Figure 4d), deviating from the value of five for random dispersal.

- **Compositional differentiation:** Among the populations, the distributions of pollen sources in the seeds initially become more homogeneous as dispersal distance increases, but their compositional differentiation levels off well above 0 even for long-distance dispersal ($\delta_{SD} = 0.234$ for $\gamma = 2000$), which is over 4 times that for symmetric dispersal (Figure 4e, "snails" in Figure 5b). The pollen source distribution in the seeds of each population is less representative of the metapopulation's seeds than for symmetric winds (smaller $\delta_{SD}(j)$), and the most representative is now not only the central population but also its downwind neighbor.
- **Fixation index:** F_{ST} for pollen sources in the seeds decreases from its initial value of 1 (no wind for $\gamma = 0$) as dispersal distance γ increases. F_{ST} again falls faster than δ_{SD} , but as opposed to symmetric dispersal, F_{ST} remains well above the value 0 that corresponds to identical pollen source distributions. This rules out random mating, which is a special case of identical distributions. The reason is the persistence of differences in pollen source distributions among the populations, as is confirmed by the positive compositional differentiation δ_{SD} (Figure 4f).
- **Adaptive potential:** The proportion $p_{imm}(j)$ of pollen immigrating into the populations increases from its initial absence for $\gamma = 0$ as dispersal distance increases, but only the three populations farthest downwind tend toward the proportion of 0.8 for random dispersal

(Figure 6c). The upwind population has by far the lowest immigration for all positive dispersal distances, and its neighbor also remains well below the value for random dispersal. The effective number $v_{imm}(j)$ of sources of immigrant pollen is undefined for $\gamma = 0$, then increases from 1 as dispersal distance increases (Figure 6d). It does not exceed the corresponding number for symmetric dispersal for any population. It remains small for population P_2 . Populations P_3 and P_4 have successively larger effective numbers that level out far below 4, the value for random dispersal. The two peripheral populations P_1 and P_5 have the same effective number as under symmetric dispersal, but here it is the highest among all populations. As dispersal distance increases, the adaptive potential of the seeds of all populations increases, since more pollen immigrates and the effective number of its sources increases. For distances over ca. 300, the downwind population P_5 has the highest adaptive potential, since it has the most immigrant pollen $p_{imm}(j)$ and the highest effective number of sources $v_{imm}(j)$. Though the upwind population P_1 has the same effective number as P_5 , the proportion of immigrant pollen upon which the effective number is based is extremely small and thus of little import for adaptation. P_2 has more immigration than P_1 but a lower effective number of sources.

4 Discussion

4.1 Effects of wind on structural characteristics in seeds

Variation of the effective number of pollen sources in seeds along the transect: In comparison to random dispersal, where the seeds of every population and the metapopulation itself have the maximum effective number of pollen sources (here five), the scenarios demonstrate that wind conditions can have a pronounced effect on the effective number of pollen sources, i.e., diversity. For symmetric dispersal, the seeds of the metapopulation are a comprehensive adaptive reserve due to the near maximality of the effective number of pollen sources, but for directional dispersal the metapopulation becomes less diverse as dispersal distance increases (black curves in Figures 4a, d). Within populations, the pollen source diversity varies along the transect. For symmetric dispersal of all dispersal distances, the central population P_3 has the highest and the peripheral populations P_1 and P_5 the lowest diversity. For directional dispersal, in contrast, the population of highest diversity moves from P_4 to the peripheral population P_5 as γ passes from 200 to 300, in contradiction to the central-peripheral hypothesis (Lira-Noriega and Manthey, 2014), while P_1 has very low diversity for all γ . For $\gamma \geq 600$, the diversity in the downwind population P_5 is even higher than in the metapopulation, making the seeds of this population a more comprehensive adaptive reserve than the entire seed production of the metapopulation.

Differences in the distribution of pollen sources in seeds along the transect: The partial connectedness among populations that is typical for metapopulations has an adaptational

advantage over single unconnected populations ($\delta_{SD} = 1$) as well as over one big population ($\delta_{SD} = 0$) in that intermediate differentiation between the constituent populations is a factor of additional genetic variation. For populations on a transect, it would seem adaptationally advantageous for the central population to be the least differentiated, since its short distance to all of the other populations and best representation of the metapopulation's adaptive variants would allow it to replenish lost variants in the others. This is true for all symmetric wind conditions, where the central population P_3 is the least differentiated. It also holds for directional dispersal up to $\gamma = 800$, but it is perhaps surprising that for stronger winds the neighboring downwind population P_4 becomes the least differentiated. Even for γ as high as 2000, P_4 still maintains the least differentiation.

As opposed to directional dispersal, symmetric dispersal over longer dispersal distance can have a homogenizing effect on pollen source distributions in seeds, ultimately causing them to become statistically indistinguishable from random dispersal for samples of the sizes commonly encountered in experimental studies. To see this, consider two additional scenarios S_{1000} and D_{1000} of symmetric and directional dispersal, respectively, of large mean dispersal distance $\gamma = 1000$ and a sample of 100 seeds from each population that closely matches the pollen source distribution in the model (data not shown). Under the hypothesis of random dispersal, the pollen source distributions would be nearly identical, i.e., homogeneous, among samples. For S_{1000} , the likelihood-ratio test of homogeneity yields a non-significant P-value of 0.183, leading to acceptance of the hypothesis of homogeneity. For D_{1000} , in contrast, the test yields a highly significant P-value of 0.000, leading to rejection of homogeneity. Thus it may be difficult to detect deviation from random dispersal in seeds formed by symmetric long-distance dispersal, while the deviation from random dispersal should be obvious in seeds formed by directional dispersal. Yet, distinguishing between symmetric and directional dispersal as the causes for an observed significant deviation from random dispersal requires an analysis of the local differentiation patterns or explicit paternity analysis in order to avoid erroneous conclusions about mating.

Relationship between diversity and differentiation: The scenarios show that wind dispersal refutes the commonly held thought that the population with the highest diversity should be the most representative for the metapopulation, that is, the least differentiated. While it does hold for symmetric dispersal with respect to the central population P_3 for all dispersal distances (Figures 4a, b), it is not generally valid for directional dispersal (Figure 4d). For example, in the directional scenarios, P_5 has the highest pollen source diversity for $\gamma \geq 300$, but P_4 is the least differentiated for $\gamma \leq 800$ (Figures 4d, e).

4.2 Effects of wind on adaptability to environmental variation

Survival of a metapopulation and its member populations depends on the relationship between the amount and the

effective number of sources of immigrant pollen and the type of environmental variation to which the populations are subjected. In the following, hypothetical implications of the simulations for responses of individual populations and the metapopulation as a whole for potential adaptive challenges are pointed out. Such challenges and their responses differ depending on the degree and kind of environmental variability, whether locally stable, locally gradual, or unstructured.

Stable environment: If the populations inhabit different but stable environments, short-distance dispersal maintains adaptedness of offspring by ensuring a high proportion of local pollen. As dispersal distance increases, increasing pollen immigration introduces maladapted pollen into the seeds as migration load that may reduce offspring survival. Symmetric dispersal gives all seeds nearly the same proportion of immigrant pollen, which becomes very large for longer dispersal distance, endangering survival. Directional dispersal gives all but the upwind population about the same amount of immigration as symmetric dispersal. Only the upwind population maintains a low migration load over all dispersal distances, ensuring its continued adaptedness for longer dispersal distances.

Gradual environmental change: If the transect follows an environmental gradient and if environmental change is gradual, environmental conditions may just shift in one direction along the transect. In such a situation, a neighboring population may have already experienced the environmental conditions that are now facing a population. The more pollen this neighbor contributes to the population's seeds, carrying with it adaptive genetic information that was beneficial in the neighbor's old environment, the better would be the chances of survival in the changed environment. Symmetric dispersal contributes more pollen from adjacent populations than from more distant ones (Figure 4c). The three interior populations have two direct neighbors and thus good chances of surviving gradual environmental change in either direction, while the one peripheral population whose single direct neighbor has not experienced the new environment is vulnerable. Thus the offspring of the three interior populations and one of the peripheral populations should be able to survive under symmetric dispersal. For directional dispersal, the upwind neighbor contributes considerably more pollen than for symmetric dispersal (Figure 4d). For $\gamma \geq 400$, even more pollen is contributed by the upwind neighbor than by the local population itself (Table 2, Scenarios D_{400} and D_{600}). Thus directional dispersal benefits survival for all but the upwind population if environmental change follows the prevailing wind direction. If change proceeds against the prevailing wind direction, all populations are endangered.

Environmental change by disturbance: Environmental disturbances are sudden, often local, and temporary. Populations that are hit by a disturbance may have reduced survival, causing them to contribute little pollen to other populations. In addition, immigrant pollen finds few ovules to fertilize in the disturbed populations. Disturbances thus reduce the degree of connectedness among populations and, due to the decrease in the amount of pollen from the disturbed

populations, the effective number of pollen sources in the metapopulation's seed production. A population's recovery from disturbance may depend on the immigration of pollen from other populations that had received the disturbed population's local adaptive genetic variants in previous generations.

In a highly connected and thus nearly undifferentiated metapopulation that arises from symmetric, longer-distance dispersal, all populations accumulate adaptive variants from the entire metapopulation in undisturbed times. In the presence of pronounced environmental differences among the populations, this high diversity would present a large migration load within the populations as long as they are undisturbed. Nevertheless, their ability to provide disturbed populations with these populations' own locally adapted variants through pollen after cessation of the disturbance could ensure persistence of the metapopulation. In a more loosely connected and thus well-differentiated metapopulation that arises from symmetric dispersal over intermediate distances, populations accumulate the adaptive genetic variants of the nearby populations and can return them after the disturbance. In this case, the effective number of pollen sources is smaller, but so is the migration load. Thus symmetric dispersal of intermediate distance is better than longer-distance symmetric dispersal when metapopulations are subject to disturbance. In contrast, prevailing directional winds spread the locally adapted variants to populations farther downwind and prevent these variants from returning when the disturbance ends.

In summary, short-distance dispersal or an upwind location maintain initial adaptedness in a constant environment. In a gradually changing environmental gradient, prospective adaptive potential depends on the direction of environmental change: the three interior populations and one of the peripheral populations have the highest adaptive potential for symmetric dispersal, while directional dispersal either ensures survival of all but the upwind population or endangers the survival of all populations. For the recovery of populations subject to disturbance, symmetric dispersal of intermediate distance is advantageous.

4.3 Practical implications

The model findings have practical implications for forest management, especially in the face of climate change, and in particular for interpretation of the results of common garden trials. The planting of trees originating from a wide range of different environments in each of a limited number of environments is used to predict which provenance would be best adapted to which environment for the purpose of reforestation. The genetic variation needed for adaptation to a changing environment is often judged based on phenotypic variability shown in the trial. For example, based on data from such trials, Yeaman and Jarvis (2006) found that variation in growth response is correlated with environmental heterogeneity in the region surrounding the population of origin in lodgepole pine (*Pinus contorta* Dougl. ex Loud). It would be interesting to see whether consideration of wind

direction during pollination in the measure of environmental heterogeneity, in addition to dispersal distance, would affect this result. In another study, Kapeller et al. (2016) detected higher variability for growth parameters for trees originating from higher alpine elevations in Norway spruce (*Picea abies* [L.] Karst.). Perhaps the greater genetic diversity at higher elevations is due to the transport of pollen from lower elevations by the morning updrafts on mountainsides (“valley winds”), during the time of day when pollen shed is heaviest. Directional winds may be a natural form of “assisted gene flow”, which is increasingly being discussed as a management measure to augment the dispersal of pollen or seeds from populations that are already under the same environmental conditions that the recipient population is expected to experience (Aitken and Bemmels, 2016).

The results also have implications for seed harvesting in natural stands. The common objective is to obtain a representative sample of the genetic variants in the stand by harvesting seeds from a sufficient number of trees that are preferably spread over the interior of the stand. The expectation is that pollen contributions to the seeds are a random sample from the entire stand. The model results contradict this expectation by demonstrating that the genetic diversity in the seeds of a tree depends on its location relative to the prevailing wind direction within the stand. This is confirmed, for example, by findings that the amount of genetic diversity contained in seeds from single trees increases from the upwind edge of a stand of European beech (*Fagus sylvatica* L.) towards the downwind edge (Gillet and Ziehe, 2012).

4.4 Outlook

There are several ways to extend this relatively simplistic model to allow for more variability. One is to consider additional forms or families of probability density functions (Austerlitz et al., 2004). The exponential function θ_γ applied here is only one member of the two-parameter exponential power family with distance parameter γ and shape parameter $c = 1$ (Clark, 1998; Hardy, 2009). In some studies, fat-tailed kernels with their increased probability of long-distance dispersal ($c < 1$) have been found to fit well with data (Clark, 1998; Austerlitz et al., 2004; Katul et al., 2005; Robledo-Arnuncio and Gil, 2005; Goto et al., 2006; Piotti et al., 2012). Their effects on metapopulation structure should be quantitatively stronger but not qualitatively different than the effects demonstrated here. Another extension would be to drop the assumption that all ovules are fertilized. By limiting the ratio of pollen-to-ovule production within the populations, the effect of fertilization probabilities less than 1 could be examined, as proposed by Gregorius (1983). Populations that receive less pollen would produce fewer seeds, skewing the distribution not only of pollen sources and but also ovule sources in the seed production of the metapopulation. Thirdly, the model framework can be applied to examine the meaningfulness of new and emerging measures of variation that describe different kinds of effective number and untangle the long-standing confusion between differentiation and apportionment (Gregorius,

2010, 2014, 2016; Gregorius and Gillet, 2015) in the framework of pollen dispersal and adaptation. Ultimately, the subsequent spatial distribution of seeds and its implications could be added.

Acknowledgements

The author thanks Claire G. Williams for motivating the application of a more general model of mine to pollen dispersal, Martin Ziehe and Hans-Rolf Gregorius for endless discussions on how to measure variation, and very especially two anonymous reviewers for valuable suggestions on all parts of the paper. This work was funded by grant Zi 662/7-1 of the Deutsche Forschungsgemeinschaft.

References

- Aitken SN, Bemmels JB (2016) Time to get moving: assisted gene flow of forest trees. *Evol Appl* 9:271–290
- Austerlitz F, Dick CW, Dutech C, Klein EK, Oddou-Muratorio S, Smouse PE, Sork VL (2004) Using genetic markers to estimate the pollen dispersal curve. *Mol Ecol* 13:937–954
- Bergmann F, Gregorius H-R, Larsen JB (1990) Levels of genetic variation in European silver fir (*Abies alba*): are they related to the species' decline? *Genetica* 82:1–10
- Burczyk J, Lewandowski A, Chalupka W (2004) Local pollen dispersal and distant gene flow in Norway spruce (*Picea abies* [L.] Karst.). *For Ecol Manage* 197:39–48
- Buschbom J, Yanbaev Y, Degen B (2011) Efficient long-distance gene flow into an isolated relict oak stand. *J Hered* 102:44–472
- Clark JS (1998) Why trees migrate so fast: confronting theory with dispersal biology and the paleorecord. *Am Nat* 152:204–224
- Gillet EM (2013) DifferInt: compositional differentiation among populations at three levels of genetic integration. *Mol Ecol Resour* 13:95–964
- Gillet EM, Ziehe M (2012) Auswirkungen von Genfluss zwischen Beständen auf großräumige genetische Variationsmuster, Metapopulationsstudien am Beispiel des Sollings. *Forstarchiv* 83:4–11
- Goto S, Shimatani K, Yoshimaru H, Takahashi Y (2006) Fat-tailed gene flow in the dioecious canopy tree species *Fraxinus mandshurica* var. *japonica* revealed by microsatellites. *Mol Ecol* 15:2985–2996
- Gregorius H-R (1977) Some fundamental relationships between genetic and genotypic multiplicity in diploid populations. *Math Biosci* 34:267–277
- Gregorius H-R (1983) Efficiency of random pollination and optimal sex ratio. *Math Biosci* 66:263–271
- Gregorius H-R (1991) On the concept of effective number. *Theor Popul Biol* 40:269–283
- Gregorius H-R (2010) Linking diversity and differentiation. *Diversity* 2:370–394
- Gregorius H-R (2014) Partitioning of trait variation among communities: measures of apportionment and differentiation based on binary sampling. *Theor Ecol* 7:313–324
- Gregorius H-R (2016) Effective numbers in the partitioning of biological diversity. *J Theor Biol* 409:133–147
- Gregorius H-R, Degen B, König A (2007) Problems in the analysis of genetic differentiation among populations: a case study in *Quercus robur*. *Silvae Genet* 56(3–4):190–199
- Gregorius H-R, Gillet EM (2015) Classifying measures of biological variation. *PLOS One* 10:e0115312
- Gregorius H-R, Kownatzki D, Höltnen AM (2011) Spatial patterns of mating relations in wild cherry (*Prunus avium* L.). *Perspect Plant Ecol Evol System* 13:36–44
- Gregorius H-R, Roberds JH (1986) Measurement of genetical differentiation among populations. *Theor Appl Genet* 71:826–834
- Hardy OJ (2009) How fat is the tail? *Heredity* (Edinb) 103:437–438
- Hedrick PW (1999) Perspective: highly variable loci and their interpretation in evolution and conservation. *Evolution* 53:313–318

- Hedrick PW (2005) A standardized genetic differentiation measure. *Evolution* 59:1633-1638
- Helbig N, Vogel B, Vogel H, Fiedler F (2004) Numerical modelling of pollen dispersion on the regional scale. *Aerobiologia* 3:3-19
- Jackson ST, Lyford ME (1999) Pollen dispersal models in quaternary plant ecology: assumptions, parameters, and prescriptions. *Bot Rev* 65:39-75
- Jost L (2008) G_{ST} and its relatives do not measure differentiation. *Mol Ecol* 17:4015-4026
- Kapeller S, Dieckmann U, Schueler S (2016) Varying selection differential throughout the climatic range of Norway spruce in Central Europe. *Evol Appl* 10(1):25-38, doi:10.1111/eva.12413
- Katul GG, Porporato A, Nathan R, Siqueira M, Soons MB, Poggi D, Horn HS, Levin SA (2005) Mechanistic analytical models for long-distance seed dispersal by wind. *Am Nat* 166:368-381
- Kremer A, Ronce O, Robledo-Arnuncio JJ, Guillaume F, Bohrer G, Nathan R, Bridle JR, Gomulkiewicz R, Klein EK, Ritland K, Kuparinen A, Gerber S, Schueler S (2012) Long-distance gene flow and adaptation of forest trees to rapid climate change. *Ecol Lett* 15:378-392
- Kuparinen A, Markkanen T, Riikonen H, Vesala T (2007) Modeling air-mediated dispersal of spores, pollen and seeds in forested areas. *Ecol Model* 208:177-188
- Lanner RM (1966) Needed: a new approach to the study of pollen dispersion. *Silvae Genet* 15:50-52
- Levins R (1969) Some demographic and genetic consequences of environmental heterogeneity for biological control. *Bull Entomol Soc Am* 15:237-240
- Liebold MA, Holyoak M, Mouquet N, Amarasekare P, Chase JM, Hoopes MF, Holt RD, Suring JB, Law R, Tilman D, Loreau M, Gonzalez A (2004) The metacommunity concept: a framework for multi-scale community ecology. *Ecol Lett* 7:601-613
- Lira-Noriega A, Manthey JD (2014) Relationship of genetic diversity and niche centrality: a survey and analysis. *Evolution* 68:1082-1093
- Loreau M, Mouquet N, Holt RD (2003) Meta-ecosystems: a theoretical framework for a spatial ecosystem ecology. *Ecol Lett* 6:673-679
- Mouquet N, Loreau M (2003) Community patterns in source-sink metacommunities. *Am Nat* 162:544-557
- Münkemüller T, Travis JMJ, Burton O, Schiffrers K, Johst K (2011) Density regulated population dynamics and conditional dispersal alter the fate of mutations occurring at the front of an expanding population. *Heredity* 106:678-689
- Nei M (1973) Analysis of gene diversity in subdivided populations. *Proc Nat Acad Sci USA* 70:3321-332
- Niklas KJ (1985) The aerodynamics of wind pollination. *Bot Rev* 51:328-386
- Piotti A, Leonardi S, Buiteveld J, Geburek T, Gerber S, Kramer K, Vettori C, Vendramin GG (2012) Comparison of pollen gene flow among four European beech (*Fagus sylvatica* L.) populations characterized by different management regimes. *Heredity (Edinb)* 108:322-331
- Prentice C (1985) Pollen representation, source area, and basin size: toward a unified theory of pollen analysis. *Quaternary Res* 23:76-86
- Robledo-Arnuncio JJ, Gil L (2005) Patterns of pollen dispersal in a small population of *Pinus sylvestris* L. revealed by total-exclusion paternity analysis. *Heredity* 94:13-22
- Smouse PE, Dyer RJ, Westfall RD, Sork VL (2001) Two-generation analysis of pollen flow across a landscape: I. Male gamete heterogeneity among females. *Evolution* 55:260-271
- Spasojevic MJ, Copeland S, Suding KN (2014) Using functional diversity patterns to explore metacommunity dynamics: a framework for understanding local and regional influences on community structure. *ECO* 37:939-949
- Vranckx G, Mergeay J, Cox K, Muys B, Jacquemyn J, Honnay O (2014) Tree density and population size affect pollen flow and mating patterns in small fragmented forest stands of pedunculate oak (*Quercus robur* L.). *For Ecol Manage* 328:254-261
- Williams CG (2010) Long-distance pine pollen still germinates after meso-scale dispersal. *Am J Bot* 97:1-10
- Wright S (1978) Variability within and among natural populations. Chicago: Univ Chicago Pr, 580 p, Evolution and the genetics of populations 4
- Yeaman S, Jarvis A (2006) Regional heterogeneity and gene flow maintain variance in a quantitative trait within populations of lodgepole pine. *Proc R Soc Lond B* 273:1587-1593

Glossary

Seeds of a population:

The set of all seeds formed by fertilization of the population's ovules during the pollination season.

Seeds of the metapopulation:

Aggregate of the seeds produced by all populations $P_1 - P_5$

Complement of population P_j :

Aggregate of the seeds produced by all populations except P_j

Pollen source distribution of population P_j :

Vector of relative frequencies $p_i(j)$ of seeds generated through fertilization of its ovules by pollen from the source populations P_i

Effective number of pollen sources in population P_j :

Rényi-diversity $v(j)$ of the pollen source distribution in the seeds of P_j

Representation of population P_j :

Difference $\delta_{SD}(j)$ between the distribution of pollen sources in its own seeds and in the complement, equaling 1 for complete disjunction, 0 for complete identity

Homogeneity among populations:

Degree of similarity $1 - \delta_{SD}$, i.e., of absence of compositional differentiation, among the pollen source distributions in the seeds of the populations, equaling 1 for identity of all distributions, 0 for complete disjunction

Proportion of local pollen for population P_j :

Relative frequency $p_j(j)$ of seeds fertilized by its own pollen

Proportion of immigrant pollen for population:

Relative frequency $p_{imm}(j) = \sum_{k,k \neq j} (1 - p_k(j))$ of seeds fertilized by pollen from all other populations, presenting potential migration load under constant environmental conditions and adaptive potential under changing environmental conditions.

Effective number of immigrant pollen sources in population P_j :

Diversity $v_{imm}(j)$ of the pollen contributions of all other populations, equaling 4 for equal representation of all pollen sources, 1 for monomorphism

Central population:

Population located at the center of the transect (P_3)

Peripheral population:

Population at the edge of the transect (P_1, P_5) as opposed to interior (P_2, P_3, P_4)

Prevailing wind direction:

For directional winds, direction in which the majority of pollen disperses, i.e., rightwards towards the higher-numbered populations

Upwind/downwind population:

Population located at the upwind/downwind end of the transect with respect to the prevailing wind direction (P_1 resp. P_5)