

Estimating the relative rates of pollen and seed migration among plant populations

R. A. ENNOS

*Institute of Ecology and Resource Management, University of Edinburgh, Darwin Building, King's Buildings,
Mayfield Rd, Edinburgh EH9 3JU, UK*

Interpopulation gene flow in plants is mediated by a combination of pollen and seed dispersal. The effectiveness of pollen and seeds in bringing about gene flow depends upon the mode of inheritance of the genetic marker. For nuclear and paternally inherited markers, gene flow occurs in both pollen and seed. For maternally inherited markers, genes are only dispersed in seeds. As a result, levels of population differentiation under drift-migration equilibrium are expected to differ for markers with contrasting modes of inheritance, and the extent of such differences should be related to the relative levels of pollen and seed migration among populations. A model is developed that relates expected values of population differentiation (F_{st}) for nuclear, paternally and maternally inherited markers, to pollen and seed migration rates. The model is used to estimate the relative rates of seed and pollen flow in six species of plants where F_{st} values are available for both nuclear and maternally inherited markers. Estimates of (pollen flow/seed flow) range from four in wild barley to 200 in oaks, and this pattern of variation is consistent with the reproductive characteristics of the species concerned.

Keywords: cpDNA, mtDNA, gene flow, pollen flow, seed flow.

Introduction

Gene flow among plant populations can take place in two ways. The first involves dispersal of pollen to a different population, successful fertilization of an ovule by this pollen, and finally establishment of the resulting seed within the site. Gene flow can also occur by dispersal of seed, and the successful establishment of dispersed seed within a new population. A complete description of interpopulation gene flow in plants must, therefore, include an assessment of the relative importance of pollen and seeds as agents of gene flow.

Attempts to measure gene flow in plants centre on two approaches. The indirect approach (Slatkin, 1985) makes use of data on population genetic structure (F_{st} , G_{st} or distribution of rare alleles) gathered from isozyme surveys. By fitting such data to models of population differentiation under a balance between migration and genetic drift, quantitative estimates of Nm , the product of effective population size and migration rate, can be calculated (Slatkin & Barton, 1989). Nm is a useful measure of the effectiveness of gene flow in preventing genetic divergence of populations through drift (Govindaraju, 1988, 1989). However, it provides no absolute measure of gene flow, nor

does it enable gene flow via pollen to be distinguished from that via seed.

Direct approaches to measuring interpopulation gene flow in plants have concentrated almost exclusively on gene flow via pollen (Ellstrand, 1992). They rely on the fact that if appropriate genetic markers are available, paternity exclusion analysis can be used to classify successful pollen as either local (from the same population) or foreign. Pollen migration rates are then calculated as the proportion of foreign pollen among all successful pollen (Ellstrand & Marshall, 1985). This approach has been widely applied for measuring pollen flow into both artificial and natural populations (Friedman & Adams, 1985; Nagasaka & Szmidt, 1985; Ellstrand *et al.*, 1989; Adams & Birkes, 1990). These studies have been instrumental in stimulating a re-assessment of the importance of interpopulation gene flow in plants. However, they are unable to provide any information on the levels of seed flow that occur among plant populations (Levin, 1984).

In recent years, advances in molecular genetics have provided techniques that allow population differentiating to be measured not only for nuclear genetic markers, but also for markers on the organelle genomes (Wagner *et al.*, 1987; Neale *et al.*, 1988;

Kremer *et al.*, 1991; Dong & Wagner, 1993; Strauss *et al.*, 1993). In contrast to biparentally inherited nuclear markers, organelle markers show predominantly uniparental inheritance (Birky, 1988; Clegg, 1989). Chloroplast genomes are generally maternally inherited in angiosperms, and paternally inherited in gymnosperms (Conde *et al.*, 1979; Neale & Sederoff, 1989); mitochondrial genomes are maternally inherited in most plants, with the exception of some gymnosperms (Neale *et al.*, 1991; Wagner *et al.*, 1991; Palmer, 1992).

As a consequence of these different modes of inheritance, we anticipate that the extent of gene flow among populations will differ for biparentally, maternally and paternally inherited markers (Birky *et al.*, 1983, 1989; Takahata & Palumbi, 1985). For instance, we anticipate that gene flow for maternally inherited markers will be lower than for nuclear markers, since only nuclear markers will be carried from one population to the next by pollen. Therefore, the extent of population differentiation is expected to vary among nuclear, maternally and paternally inherited markers for the same set of populations (Birky *et al.*, 1989). Moreover, in plant populations, the extent of these differences is likely to be a function of the relative amounts of interpopulation pollen and seed flow. Thus, it may be possible to make use of simultaneous measurements on population differentiation for markers with different modes of inheritance to obtain insights into pollen and seed flow among populations.

In this paper I relate levels of gene migration for nuclear, maternally inherited, and paternally inherited markers to levels of interpopulation pollen and seed flow. I then derive expressions for the expected levels of genetic differentiation among populations for these three types of markers under a genetic drift-migration balance. These expressions are couched in terms of the rates of interpopulation pollen and seed flow. I then show how approximate estimates of the levels of (pollen flow/seed flow) can be calculated if data on genetic differentiation, for markers with different modes of inheritance, are available for the same set of plant populations. Finally I use these approximations to compare (pollen flow/seed flow) for six species of plants for which appropriate data are available.

Rates of gene flow for nuclear, paternally inherited and maternally inherited markers

The analysis is based on a model which assumes that the plant species under study is both diploid and hermaphrodite. The species is distributed as an infinite number of island populations, each of size N , linked by a constant level of gene flow. For simplicity the popula-

tions are initially assumed to be random mating. Consider a single population within this metapopulation. The population is constant in size, and therefore the number of seeds giving rise to the next generation in each population is N .

Let the proportion of pollen, derived from individuals within the reference population, that successfully fertilizes ovules within this population be x . The proportion of pollen originating from individuals in other populations, and successfully fertilizing ovules within the reference population, is therefore $(1-x)$. This will be referred to as the rate of interpopulation pollen flow.

Let the proportion of seeds contributing to the next generation, that are derived from plants within the reference population, be y . The proportion of successfully establishing seeds coming from other populations is therefore $(1-y)$. This will be referred to as the rate of interpopulation seed flow.

With these assumptions, we can now derive an expression for the per generation interpopulation migration rates of biparentally inherited nuclear genes, in terms of the per generation interpopulation migration rates of seed and pollen. In order to facilitate this calculation, it is convenient to deal separately with the non-dispersed and dispersed seed components contributing to the next generation.

For the non-dispersed seed component, the total number of seeds establishing that come from the reference population is Ny . The number of migrant genes establishing within the population as a consequence of pollen flow is $Ny(1-x)$. The number of non-migrant genes establishing in the population in the non-dispersed seed component is Ny (female gametes) + Nyx (male gametes) = $Ny(1+x)$. As a check we note that the total number of genes establishing in the non-dispersed seed component is $Ny[(1-x) + (1+x)] = 2Ny$.

For the dispersed seed component, the total number of dispersed seeds that establish in the population is $N(1-y)$. The number of migrant genes establishing as a consequence of this seed flow is $2N(1-y)$. The number of non-migrant genes establishing as a consequence of this seed flow is 0. Summing the non-dispersed and dispersed seed components gives the total number of nuclear genes establishing as $2Ny + 2N(1-y) = 2N$.

Using these results we can express the migration rate for nuclear genes, m_b (the proportion of the total nuclear genes establishing each generation that are migrants) in terms of the amounts of interpopulation seed and pollen flow. Thus,

$$m_b = [Ny(1-x) + 2N(1-y)]/2N = 1 - [y(1+x)]/2.$$

Let us now consider the interpopulation migration rate for paternally inherited markers. Again we will deal with the non-dispersed and dispersed seed components separately. If there is no heteroplasmy for the paternally inherited genome, the total number of paternally inherited genes establishing in the non-dispersed seed component will be Ny . The number of migrant paternally inherited genes establishing in this non-dispersed seed as a consequence of pollen flow is $Ny(1-x)$. The number of non-migrant paternally inherited genes establishing in the non-dispersed seed component is Nyx .

Within the dispersed seed component the total number of paternally inherited genes is $N(1-y)$. The number of migrant paternally inherited genes establishing from these dispersed seeds is $N(1-y)$. The number of non-migrant paternally inherited genes establishing from dispersed seed is 0. As a check, the total number of paternally inherited genes establishing from non-dispersed plus dispersed seed is therefore $Ny + N(1-y) = N$.

Thus, the migration rate for paternally inherited genes, m_p , the proportion of migrant genes among the total in terms of interpopulation seed and pollen migration is:

$$m_p = [Ny(1-x) + N(1-y)]/N = 1 - yx.$$

Similar arguments can be used to obtain the expression for the migration rate of maternally inherited genetic markers, m_m , as:

$$m_m = N(1-y)/N = (1-y).$$

The results of these calculations of migration rate for genetic markers with the three types of inheritance patterns are summarized in Table 1. Having established

these expressions for rates of gene migration among populations, we can use them to derive expressions for the extent of genetic differentiation among populations for nuclear, paternally inherited, and maternally inherited markers under an equilibrium between genetic drift and selection.

Population differentiation under migration – drift equilibrium

Nuclear markers

Following Wright (1951), the extent of genetic differentiation is measured in terms of the inbreeding coefficient F_{st} . For an island model, with an infinite number of ideal, random mating, diploid populations of size N , and a migration rate among populations of m_b per generation, the amount of genetic differentiation among populations for selectively neutral, nuclear markers, $F_{st(b)}$, is given by:

$$F_{st(b)} = 1/(4Nm_b + 1). \quad (\text{Wright, 1951})$$

Substituting in the expression for migration rate of nuclear genes in terms of pollen and seed migration, m_b (eqn 1), gives:

$$F_{st(b)} = 1/[4N(1-y(1+x)/2) + 1]. \quad (1)$$

Rearranging this expression gives:

$$N = (1/F_{st(b)} - 1)/[4 - 2y(1+x)]. \quad (1a)$$

Paternally inherited markers

For paternally inherited genes, assuming no heteroplasmy, only one copy of each gene is found per in-

Table 1 Relationship between gene migration rates for biparentally (m_b), paternally (m_p) and maternally (m_m) inherited genetic markers and the levels of pollen ($1-x$) and seed ($1-y$) flow among populations of size N

Seed component	Category	Nuclear markers	Paternal markers	Maternal markers
Non-dispersed	Migrant genes	$Ny(1-x)$	$Ny(1-x)$	0
	Non-migrant genes	Nyx	Nyx	Ny
	Total genes	$2Ny$	Ny	Ny
Dispersed	Migrant genes	$2N(1-y)$	$N(1-y)$	$N(1-y)$
	Non-migrant genes	0	0	0
	Total genes	$2N(1-y)$	$N(1-y)$	$N(1-y)$
Total	Migrant genes	$Ny(1-x) + 2N(1-y)$	$Ny(1-x) + N(1-y)$	$N(1-y)$
	Non-migrant genes	Nyx	Nyx	Ny
	Total genes	$2N$	N	N
Gene migration rate		$m_b = [1 - y(1+x)/2]$	$m_p = (1-yx)$	$m_m = (1-y)$

dividual. The behaviour of organelle genes under drift will then be equivalent to the behaviour of nuclear genes in haploids (Birky *et al.*, 1983, 1989; Takahata & Palumbi, 1985). For haploid populations of size N , linked by migration at a rate m per generation, the amount of genetic differentiation, F_{st} , at drift-migration equilibrium is given by:

$$F_{st} = 1/(2Nm + 1).$$

Substituting in the expression for migration rate of paternally inherited genes in terms of pollen and seed migration (eqn 2) gives the equilibrium level of genetic differentiation for paternally inherited markers, $F_{st(p)}$, as:

$$F_{st(p)} = 1/[2N(1 - xy) + 1]. \quad (2)$$

Rearranging this expression gives:

$$N = (1/F_{st(p)} - 1)/[2(1 - xy)]. \quad (2a)$$

Maternally inherited markers

For maternally inherited genes, with one copy per individual and migration rate m_m , genetic differentiation at migration-drift equilibrium, $F_{st(m)}$, is given by:

$$F_{st(m)} = 1/(2Nm_m + 1).$$

Substituting in the expression for migration rate of maternally inherited genes in terms of pollen and seed migration (eqn 3) gives:

$$F_{st(m)} = 1/[2N(1 - y) + 1]. \quad (3)$$

Rearranging this expression gives:

$$N = (1/F_{st(m)} - 1)/[2(1 - y)]. \quad (3a)$$

Relationship between population differentiation for nuclear, paternally inherited, and maternally inherited genes

The extent of genetic differentiation at equilibrium between migration and drift is a function of effective population size and migration rates. In the absence of inbreeding within populations, effective population size is identical for the three types of markers. Moreover the seed and pollen migration rates that influence population differentiation for nuclear genes, also affect the levels of population differentiation for paternally inherited and maternally inherited genes. For these reasons we anticipate that the three values of F_{st} , for markers with contrasting modes of inheritance, will not be independent. It is therefore important to establish the relationship between these three measures of F_{st} . In order to do this we equate eqns 1, 2 and 3 for population size, N , given above. Before proceeding it is help-

ful to make the following substitutions.

$$\text{Let } (1/F_{st(b)} - 1) = A,$$

$$(1/F_{st(p)} - 1) = B, \text{ and}$$

$$(1/F_{st(m)} - 1) = C.$$

This gives us the three equations:

$$N = A/[4 - 2y(1 + x)], \quad (1b)$$

$$N = B/[2(1 - xy)], \quad (2b)$$

$$N = C/[2(1 - y)]. \quad (3b)$$

Equating 2b and 3b gives:

$$B[2(1 - y)] = C[2(1 - xy)],$$

$$Cxy = C - B + By,$$

$$x = (C - B + By)/Cy.$$

Substituting in (1b) and equating with (3b) gives:

$$A/[4 - 2y(1 + (C - B + By)/Cy)] = C/[2(1 - y)],$$

$$A(1 - y) = C(2 - y - (Cy + By - By^2)/Cy),$$

$$A(1 - y) = C(1 - y + B/C - By/C),$$

$$A(1 - y) = C(1 - y)(1 + B/C),$$

$$A = B + C.$$

Thus, the population differentiation F_{st} statistics, for genes with different modes of inheritance, are not independent. They are related according to the formula:

$$(1/F_{st(b)} - 1) = (1/F_{st(p)} - 1) + (1/F_{st(m)} - 1). \quad (4)$$

Comparison of population differentiation among nuclear, paternally inherited and maternally inherited markers

Using eqns 1, 2 and 3 derived above, it is possible to compare the levels of genetic differentiation for markers with contrasting modes of inheritance when there is an equilibrium between drift and migration. It is also possible to determine the conditions under which inequalities exist between the levels of genetic differentiation for these different markers.

Comparison of $F_{st(m)}$ and $F_{st(b)}$

Nuclear markers migrate both by pollen and seed, whereas maternally inherited markers only migrate via seed. We would therefore expect $F_{st(m)}$ to exceed $F_{st(b)}$ at migration-drift equilibrium. According to eqns 1 and 3, population differentiation for maternally inherited genes $F_{st(m)}$ will be greater than that for nuclear genes $F_{st(b)}$ when:

$$1/[2N(1-y)+1] > 1/[4N(1-y(1+x)/2+1)],$$

$$4N - 2Ny - 2Nxy + 1 > 2N - 2Ny + 1,$$

$$2N(1-xy) > 0,$$

$$1 > xy.$$

Values of x and y can vary between 0 and 1. Therefore, population differentiation will be greater for maternally inherited than for nuclear genes when the values of either x or y , or both x and y , are less than 1.0. These conditions are met when there is either pollen migration, seed migration, or both occurring among populations.

Comparison of $F_{st(m)}$ and $F_{st(p)}$

Paternally inherited markers are expected to show less population differentiation than maternally inherited markers since they are dispersed in pollen as well as in seed. Making use of eqns 2 and 3 the conditions for $F_{st(m)}$ to be greater than $F_{st(p)}$ are:

$$1/[2N(1-y)+1] > 1/[2N(1-xy)+1],$$

$$2N(1-xy)+1 > 2N(1-y)+1,$$

$$1 > x.$$

Thus, so long as pollen migration takes place among populations, maternally inherited markers will always show greater differentiation than paternally inherited markers.

Comparison of $F_{st(p)}$ and $F_{st(b)}$

Using eqns 1 and 2, population differentiation for paternally inherited genes, $F_{st(p)}$, will be greater than for nuclear genes, $F_{st(b)}$, when:

$$1/[2N(1-xy)+1] > 1/[4N(1-y(1+x)/2+1)],$$

$$4N - 2Ny - 2Nxy + 1 > 2N - 2Nxy + 1,$$

$$2N - 2Ny > 0,$$

$$1 > y.$$

So, as long as seed migration takes place among populations, population differentiation for paternally inherited markers will be greater than that for nuclear markers.

Estimating relative levels of seed and pollen migration by comparisons of F_{st} values for markers with contrasting modes of inheritance

Comparison of F_{st} for nuclear and maternally inherited markers

If measurements of population differentiation are available for both nuclear and maternally inherited markers, eqns 1b and 3b can be used to derive an approximate estimate of the relative amounts of seed and pollen flow among populations. Equating the two formulae gives:

$$A/[4-2y(1+x)] = C/[2(1-y)],$$

$$A - Ay = 2C - Cy - Cxy,$$

$$(A - 2C)(1-y) = Cy(1-x),$$

$$(1-x)/(1-y) = (A - 2C)/Cy.$$

When seed flow among populations is small, the value of y is close to 1.0. In this situation the ratio of pollen flow into the population, $(1-x)$, relative to seed flow into the population, $(1-y)$, is approximated by the expression:

$$(\text{Pollen flow/Seed flow}) = (A - 2C)/C. \quad (5)$$

Comparison of F_{st} for paternally and maternally inherited genes

Where simultaneous measurements of population differentiation for paternally and maternally inherited markers are available, a second approximation for the relative rates of pollen and seed flow can be made. Equating eqns 2b and 3b gives:

$$B/[2(1-xy)] = C/[2(1-y)],$$

$$B - By = C - Cxy,$$

$$y(B - Cx) = B - C,$$

$$y = (B - C)/(B - Cx),$$

$$(1-y) = (B - Cx - B + C)/(B - Cx),$$

$$(1-y) = C(1-x)/(B - Cx),$$

$$(1-x)/(1-y) = (B - Cx)/C.$$

For small values of pollen flow, when x lies close to 1.0, a reasonable approximation for the relative rates of pollen and seed flow is:

$$(\text{Pollen flow/Seed flow}) = (B - C)/C. \quad (6)$$

Modification of model for inbreeding populations

For simplicity, the model relating genetic differentiation and gene flow used above assumes that random mating occurs within populations. In other words, the extent of inbreeding, F_{is} , for each population is taken to be zero. In these circumstances, the variance effective size of each of these ideal populations is N , and the effective number of biparentally inherited nuclear alleles at each locus in each population is $2N$. The effective number of uniparentally inherited organelle alleles within each population, in contrast, is N .

In many plant populations the assumption of random mating is violated, particularly as a result of self-fertilization and its attendant inbreeding. Values of F_{is} can range from 0 to 1 depending upon the level of self-fertilization in the population. At inbreeding equilibrium the value of F_{is} is given by $s/(2-s)$, where s is the rate of selfing. Where this type of inbreeding occurs in a population of size N , its variance effective population size is reduced to $N/(1+F_{is})$ (Caballero & Hill, 1992). The effective number of biparentally inherited nuclear alleles at a locus in a population of size N is therefore reduced to $2N/(1+F_{is})$. Inbreeding does not, however, affect the variance effective number of uniparental organelle alleles in the subpopulation, which remains at N .

In order to account for the effects of systematic inbreeding, arising from self-fertilization, on the variance effective population size for nuclear markers, the value N is replaced by $N/(1+F_{is})$ in eqns 1 and 1a. However, because inbreeding does not alter the variance effective number of alleles for uniparentally inherited organelle markers, eqns 2, 2a, 3 and 3a remain unchanged. Taking note of these conclusions, more general formulae relating $F_{st(b)}$, $F_{st(p)}$ and $F_{st(m)}$ values under any level of inbreeding, F_{is} , in the population, can be given. With inbreeding, the relationship between the measures of population differentiation for markers with different modes of inheritance becomes:

$$(1/F_{st(b)} - 1)(1 + F_{is}) = (1/F_{st(p)} - 1) + (1/F_{st(m)} - 1). \quad (4a)$$

Inbreeding also affects the value of population differentiation for nuclear markers, $F_{st(b)}$, relative to the values of $F_{st(p)}$ and $F_{st(m)}$. The condition for greater population differentiation for maternally inherited than for nuclear markers becomes:

$$F_{st(m)} > F_{st(b)}, \quad \text{when } (1 - F_{is}) > y(x - F_{is}),$$

and the condition for greater population differentiation for paternally inherited than for nuclear markers becomes:

$$F_{st(p)} > F_{st(b)}, \quad \text{when } (1 - F_{is}) > y(1 - xF_{is}).$$

Finally, the more general formula for obtaining an approximate estimate of interpopulation pollen flow relative to seed flow, from measurements on genetic differentiation of nuclear and maternally inherited markers, becomes:

$$\begin{aligned} (\text{Pollen flow})/(\text{Seed flow}) &= (1 - x)/(1 - y) \\ &= [A(1 + F_{is}) - 2C]/C. \quad (5a) \end{aligned}$$

Note that all the above formulae reduce to the original form when F_{is} is equal to zero.

Application of theory for estimating relative rates of interpopulation pollen and seed flow

A number of sets of data have recently become available in which measurements of population differentiation for both biparentally inherited nuclear markers and maternally inherited organelle markers are reported (see Table 2 for references). This provides us with the opportunity to apply the theory outlined above to estimate relative rates of pollen and seed flow among populations of plants with diverse reproductive biology. Estimates of $F_{st(b)}$ are all from surveys of isozyme variation. Data for differentiation of maternally inherited markers comes from RFLP chloroplast DNA markers in the angiosperms *Quercus robur/petraea* and *Hordeum spontaneum*, and from RFLP mitochondrial DNA markers in the case of the gymnosperms *Pinus contorta*, *P. radiata*, *P. attenuata* and *P. muricata*. For the sake of these calculations the inbreeding coefficients, F_{is} , for the outbreeding *Quercus* (oak) and *Pinus* (pine) species are both assumed to be 0, while F_{is} for the highly inbreeding *H. spontaneum* (wild barley) is assumed to be 1.0 (Brown *et al.*, 1978). Estimates of pollen migration rate relative to seed migration rate range from nearly 200, for the oak species, to four, for wild barley (Table 2). Figures for the pine species are intermediate between these extremes.

Discussion

This paper has demonstrated that it is possible to estimate the relative rates of pollen and seed migration among plant populations from a simple comparison of F_{st} values for nuclear and maternally inherited organelle genetic markers. The approach can be contrasted with other recent theoretical developments aimed at measuring seed and pollen components of gene flow in plants. These more sophisticated analyses also employ data collected jointly from nuclear and organelle genomes, but exploit the fact that, under appropriate conditions, pollen and seed flow between genetically differentiated populations will generate dis-

Table 2 Estimates of the ratio of pollen to seed flow for six plant species making use of data on the inbreeding coefficient F_{is} and measures of population differentiation for nuclear, $F_{st(b)}$, and maternally inherited, $F_{st(m)}$, markers

Species	F_{is}	$F_{st(b)}$	$F_{st(m)}$	(Pollen flow/Seed flow)	References
<i>Quercus petraea</i> / <i>Q. robur</i> complex	0	0.037	0.884	196	Kremer <i>et al.</i> (1991)
<i>Pinus contorta</i> (mean over subspecies)	0	0.0608	0.66	28	Wheeler & Guries (1982) Dong & Wagner (1993)
<i>P. contorta</i> ssp. <i>latifolia</i>			0.82	68	
<i>P. contorta</i> ssp. <i>contorta</i>			0.56	18	
<i>P. contorta</i> ssp. <i>murrayana</i>			0.59	20	
<i>P. radiata</i>	0	0.13	0.833	31	Millar <i>et al.</i> (1988) Strauss <i>et al.</i> (1993)
<i>P. attenuata</i>	0	0.12	0.863	44	
<i>P. muricata</i>	0	0.22	0.882	24	
<i>Hordeum</i> <i>spontaneum</i>	1.0	0.49	0.735	4	Nevo <i>et al.</i> (1979) Brown (1979) Neale <i>et al.</i> (1988)

equilibria between nuclear and organelle markers (Asmussen *et al.*, 1987). Through measurement and analysis of these 'cytonuclear disequilibria' it should be possible to estimate rates of pollen and seed flow between populations (Asmussen *et al.*, 1987; Asmussen & Schnabel, 1991; Schnabel & Asmussen, 1992). Application of these techniques to real populations awaits appropriate data collection (possibly from hybrid zones where the necessary genetic differentiation between populations is present) and the development of statistical techniques for estimating pollen and seed flow rates from cytonuclear frequencies (Schnabel & Asmussen, 1992).

In comparison with the methods outlined in this paper, the analysis of cytonuclear disequilibria is more restricted with respect to the range of situations that can be analysed. Significant population differentiation must be present before measurable disequilibria are likely to be detected. Such analysis requires that nuclear and organelle genotypes be determined jointly within the same individuals, a condition that is not required when F_{st} statistics for nuclear and organelle genomes are being estimated. Finally, it is likely that rather larger sample sizes per population will be required to measure disequilibrium statistics with the accuracy required for estimation of pollen and seed flow rates than would be the case if a simple comparison of F_{st} statistics were being undertaken (Asmussen & Schnabel, 1991; Schnabel & Asmussen,

1992). However, experience from more extensive practical application of these two approaches is required before their advantages and disadvantages can be fully assessed.

Returning to the results of the present study, a number of critical questions needs to be addressed. The first is whether the analysis generates results that are biologically reasonable. For all six species investigated, estimated rates of pollen migration are greater than rates of seed migration, as might be expected. However, differences among species are substantial. The greatest contrast between pollen and seed migration rates is found for the oak species complex, where interpopulation pollen flow is estimated to be 200 times greater than interpopulation seed flow. This result may be accounted for by the reproductive system of the species. Oaks are predisposed to high rates of interpopulation pollen dispersal since they are outbreeding, wind-pollinated, and release pollen from a substantial height. On the other hand, dispersal of acorns by birds and rodents is likely to be relatively restricted, and opportunities for establishment of dispersed seed may be rare (Sork, 1984). Thus, it is not inconceivable that pollen migration rates are as much as 200 times higher than seed migration rates.

In outbreeding pine species, interpopulation pollen dispersal is again expected to be substantial. Indeed, direct measures of pollen flow into seed orchards and along natural clines show significant pollen flow over

many km (Millar, 1983; Nagasaka & Szmids, 1985). However, the discrepancy between pollen and seed flow may not be as large as in oaks, because pine seeds are also adapted to long range wind dispersal and have been shown to establish successfully over distances as great as 70 km (Cwynar & MacDonald, 1987). Estimates of pollen/seed migration rates ranging from 20 to 60 are not inconsistent with these facts.

In contrast to the two figures for outbreeding forest tree species, the rate of interpopulation pollen flow for the annual inbreeding species, wild barley, is estimated to be only four times that of seed flow. Opportunities for interpopulation pollen dispersal in such a highly selfing species are expected to be rare and it is not surprising that pollen and seed flow should be of the same order of magnitude for this species (Brown *et al.*, 1978). We can conclude from this rather sparse set of results that the figures on the relative rates of pollen and seed migration, derived from the analysis developed in this paper, are compatible with the biological characteristics of the species concerned.

Although these quantitative estimates of seed to pollen migration rates seem plausible on the basis of the species' biology, it is important to note that many assumptions underlie the analysis, assumptions that may limit the accuracy and validity of the results. Let us begin by considering the nature of the genetic variants that are being employed in the analysis. The model assumes that the rates of mutation at the loci concerned (in both nuclear and organelle genomes) are very low compared with the rates of interpopulation gene flow for these markers. Mutation does not, therefore, affect the equilibrium levels of F_{st} , nor does it lead to heteroplasmy of organelle genomes (Birky *et al.*, 1983, 1989; Takahata & Palumbi, 1985). This assumption is likely to hold for isozyme loci in the nuclear genome, and restriction site variants in the chloroplast genomes of plants (Birky, 1988). More information is currently needed on the rate at which structural rearrangements take place in the mitochondrial genome of plants to ensure that they occur at a sufficiently low rate to be used in this analysis. Such structural variants were used to measure $F_{st(m)}$ in the pine studies analysed here (Dong & Wagner, 1993; Strauss *et al.*, 1993). The high rate of differentiation for mitochondrial structural rearrangements among these, and other plant populations, suggests that the rate of mtDNA structural rearrangement within species is sufficiently low for such variants to be used (Saumitou-Laprade *et al.*, 1993).

A second crucial assumption made in the model is that uniparental inheritance of the organelle genomes is complete, and that heteroplasmy, arising from biparental inheritance or a very high mutation rate, is

absent. Evidence on chloroplast and mitochondrial genomes so far suggests that, while biparental inheritance and heteroplasmy have been detected, they are relatively rare phenomena for the species analysed in Table 2, and individual plants generally contain only one form of chloroplast and one form of mitochondrial genome (Kremer *et al.*, 1991; Wagner *et al.*, 1991; Strauss *et al.*, 1993).

Apart from questioning the appropriateness of the experimental data, it is necessary to assess the realism of the genetic model employed in this analysis, before accepting the results. The island model of population differentiation that forms the framework for the analysis makes a large number of assumptions that are not upheld in nature. These include constant and equal population sizes, and constant and equal migration rates among populations. Estimates of the relative amounts of pollen and seed migration that emerge when such a model is applied must necessarily be average figures, which gloss over the heterogeneity in interpopulation migration rates that is likely to be found in nature (Slatkin & Barton, 1989; Ellstrand, 1992). Thus, they should be regarded not as absolute, but as 'realistic' figures that will be most useful in a comparative context. The model also assumes that populations are at drift-migration equilibrium. Although it is known that drift-migration equilibrium levels of F_{st} are reached quite rapidly (Birky *et al.*, 1989), caution needs to be exercised in using data from species with long generation times, especially where disturbance of populations is known to have occurred in the recent past. For such populations, migration-drift equilibrium levels of F_{st} may not have been achieved after the disturbance.

In addition to these theoretical limitations, the methods described above have a number of practical shortcomings when used to compare pollen and seed migration rates. The first is that replicate estimates of the parameters $F_{st(p)}$ and $F_{st(m)}$ cannot be obtained. This is due to the fact that neither the chloroplast nor the mitochondrial genomes recombine, and are in complete linkage disequilibrium. Each organelle genome is effectively one gene. Estimates of F_{st} for different marker loci on these genomes are not independent. This contrasts with the situation for the nuclear genome in a random mating species, where independent estimates of $F_{st(b)}$ can be obtained from loci that are in linkage equilibrium.

Another limitation is that estimates of pollen/seed migration are only valid when there are low absolute levels of seed migration (comparison of $F_{st(b)}$ and $F_{st(m)}$) or low absolute values of pollen migration (comparison of $F_{st(p)}$ and $F_{st(m)}$). Since pollen migration rates are generally believed to be greater than seed migration

rates, the former assumption is likely to be the more realistic. Thus, comparison of $F_{st(b)}$ and $F_{st(m)}$ are likely to give more reliable values for pollen/seed migration rates than comparisons of $F_{st(p)}$ and $F_{st(m)}$. This point is perhaps rather academic, because there is only a limited number of species, chiefly gymnosperms, where paternally inherited markers are available.

Although the purpose of developing the model was to allow quantitative estimates of pollen/seed migration to be calculated, it also generates a number of more general and qualitative predictions about the levels of population differentiation for selectively neutral markers with different modes of inheritance. The first is that the expected levels of genetic differentiation for nuclear, paternally and maternally inherited markers are simply related according to eqn 4a. This prediction could be tested with data from gymnosperm populations in which nuclear, paternally inherited (chloroplast) and maternally inherited (mitochondrial) genomes occur (Neale & Sederoff, 1989).

The second set of predictions is that for any degree of inbreeding in the population, genetic differentiation for nuclear and paternally inherited markers will be greater than for maternally inherited markers, provided that some seed establishes in the home population ($y > 0$) and some pollen flow occurs among populations [$(1-x) > 0$]. When there is no inbreeding, differentiation for paternally inherited markers will be greater than for nuclear markers so long as some seed flow occurs [$(1-y) > 0$]. However, the inequality is reversed when the population is completely inbreeding ($F_{is} = 1$), provided that some seed establishes in the home population ($y > 0$), and some pollen flow occurs [$(1-x) > 0$]. These results are amenable to experimental test. They also emphasize the fact that the degree of inbreeding within plant populations, as well as the levels of gene flow among populations, has a significant effect on levels of population differentiation for genetic markers with contrasting modes of inheritance.

The analyses outlined above would not have been possible without the extra information on organelle genome diversity that has been provided by the application of molecular biological techniques. A major weakness of the analysis, however, is that it only makes use of qualitative information about organelle genome variability, i.e. whether organelle genomes are the same or different for particular genetic markers. The challenge for the future will be to develop techniques that will incorporate information on the degree of differences among organelle variants, so that more sophisticated insights into gene flow in plant populations can be obtained.

Acknowledgements

This work was initiated during the tenure of a CIBA-GEIGY Senior Research Fellowship at the Swedish Agricultural University, Umeå, Sweden. I am very grateful to Professor D. Lindgren, Dr A. Szmidt and Dr X.-R. Wang for their stimulation and hospitality during my stay. Completion of the work in Edinburgh has been supported by Research Grant GR3/8606 from the Natural Environment Research Council.

References

- ADAMS, W. T. AND BIRKES, D. S. 1990. Estimating mating patterns in forest tree populations. In: Hattemer, H. H. and Fineschi, S. (eds) *Biochemical Markers in the Population Genetics of Forest Trees*, pp. 157–172. S.P.B. Academic Publishing, The Hague.
- ASMUSSEN, M. A. AND SCHNABEL, A. 1991. Comparative effects of pollen and seed migration on the cytonuclear structure of plant populations. I. Maternal cytoplasmic inheritance. *Genetics*, **128**, 639–654.
- ASMUSSEN, M. A., ARNOLD, J. AND AVISE, J. C. 1987. Definition and properties of disequilibrium statistics for associations between nuclear and cytoplasmic genotypes. *Genetics*, **115**, 755–768.
- BIRKY, C. W. 1988. Evolution and variation in plant chloroplast and mitochondrial genomes. In: Gottlieb, L. D. and Jain, S. K. (eds) *Plant Evolutionary Biology*, pp. 23–55. Chapman and Hall, London.
- BIRKY, C. W., MARUYAMA, T. AND FUERST, P. 1983. An approach to population and evolutionary genetic theory for genes in mitochondria and chloroplasts, and some results. *Genetics*, **103**, 513–527.
- BIRKY, C. W., FUERST, P. AND MARUYAMA, T. 1989. Organelle gene diversity under migration, mutation and drift: equilibrium expectations, approach to equilibrium, effects of heteroplasmic cells, and comparison to nuclear genes. *Genetics*, **121**, 613–627.
- BROWN, A. H. D. 1979. Enzyme polymorphism in plant populations. *Theor. Pop. Biol.*, **15**, 1–42.
- BROWN, A. H. D., ZOHARY, D. AND NEVO, E. 1978. Outcrossing rates and heterozygosity in natural populations of *Hordeum spontaneum* Koch. in Israel. *Heredity*, **41**, 49–62.
- CABALLERO, A. AND HILL, W. G. 1992. Effective size of non-random mating populations. *Genetics*, **130**, 909–916.
- CLEGG, M. T. 1989. Molecular diversity in plant populations. In: Brown, A. H. D., Clegg, M. T., Kahler, A. L. and Weir, B. S. (eds) *Plant Population Genetics, Breeding and Genetic Resources*, pp. 98–115. Sinauer, Sunderland, MA.
- CONDE, M. F., PRING, D. R. AND LEVINGS, C. S. 1979. Maternal inheritance of organelle DNAs in *Zea mays*–*Zea perennis* reciprocal crosses. *J. Hered.*, **70**, 2–4.
- CWYNAR, L. C. AND MACDONALD, G. M. 1987. Geographic variation of lodgepole pine in relation to population history. *Am. Nat.*, **129**, 463–469.

- DONG, J. AND WAGNER, D. B. 1993. Taxonomic and population differentiation of mitochondrial diversity in *Pinus banksiana* and *Pinus contorta*. *Theor. Appl. Genet.*, **86**, 573-578.
- ELLSTRAND, N. C. 1992. Gene flow among seed plant populations. *New Forests*, **6**, 241-256.
- ELLSTRAND, N. C. AND MARSHALL, D. L. 1985. Interpopulation gene flow by pollen in wild radish, *Raphanus sativus*. *Am. Nat.*, **126**, 606-616.
- ELLSTRAND, N. C., DEVLIN, B. AND MARSHALL, D. L. 1989. Gene flow by pollen into small populations: data from experimental and natural stands of wild radish. *Proc. Natl. Acad. Sci. U.S.A.*, **86**, 9044-9047.
- FRIEDMAN, S. T. AND ADAMS, W. T. 1985. Estimation of gene flow into two seed orchards of loblolly pine (*Pinus taeda* L.). *Theor. Appl. Genet.*, **69**, 609-615.
- GOVINDARAJU, D. R. 1988. Relationship between dispersal ability and levels of gene flow in plants. *Oikos*, **52**, 31-35.
- GOVINDARAJU, D. R. 1989. Estimates of gene flow in forest trees. *Biol. J. Linn. Soc.*, **37**, 345-357.
- KREMER, A., PETIT, R., ZANETTO, A., FOUGERE, V., DUCOUSSO, A., WAGREN, D. AND CHAUVIN, C. 1991. Nuclear and organelle gene diversity in *Quercus robur* and *Q. petraea*. In: Muller-Starck, G. and Ziehe, M. (eds) *Genetic Variation in European Populations of Forest Trees*, pp. 141-166. Sauerlander's Verlag, Frankfurt.
- LEVIN, D. A. 1984. Immigration in plants: an exercise in the subjunctive. In: Dirzo, R. and Sarukhan, J. (eds) *Perspectives on Plant Population Ecology*, pp. 242-260. Sinauer, Sunderland, MA.
- MILLAR, C. I. 1983. A steep cline in *Pinus muricata*. *Evolution*, **37**, 311-319.
- MILLAR, C. I., STRAUSS, S. H., CONKLE, M. T. AND WESTFALL, R. D. 1988. Allozyme differentiation and biosystematics of the California close-cone pines (*Pinus* subsect. *Oocarpae*). *Syst. Bot.*, **13**, 351-370.
- NAGASAKA, K. AND SZMIDT, A. E. 1985. Multilocus analysis of external pollen contamination of a Scots pine (*Pinus sylvestris* L.) seed orchard. In: Gregorius, H.-R. (ed.) *Population Genetics in Forestry, Lecture Notes in Biomathematics*, **60**, pp. 134-138. Springer, Berlin.
- NEALE, D. B. AND SEDEROFF, R. R. 1989. Paternal inheritance of chloroplast DNA and maternal inheritance of mitochondrial DNA in loblolly pine. *Theor. Appl. Genet.*, **77**, 212-216.
- NEALE, D. B., SAGHAI-MAROOF, M. A., ALLARD, R. W., ZHANG, Q. AND JORGENSEN, R. A. 1988. Chloroplast DNA diversity in populations of wild and cultivated barley. *Genetics*, **120**, 1105-1110.
- NEALE, D. B., MARSHALL, K. A. AND HARRY, D. E. 1991. Inheritance of chloroplast and mitochondrial DNA in incense cedar (*Calocedrus decurrens*). *Can. J. For. Res.*, **21**, 717-720.
- NEVO, E., ZOHARY, D., BROWN, A. D. H. AND HABER, M. 1979. Genetic diversity and environmental associations of wild barley, *Hordeum spontaneum* in Israel. *Evolution*, **33**, 815-833.
- PALMER, J. P. 1992. Mitochondrial DNA in plant systematics: applications and limitations. In: Soltis, P. S., Soltis, D. E. and Doyle, J. J. (eds) *Molecular Systematics of Plants*, pp. 36-49. Chapman and Hall, New York.
- SAUMITOU-LAPRADE, P., ROUWENDAL, G. J. A., CUGEN, J., KRENS, F. A. AND MICHAELIS, G. 1993. Different CMS sources found in *Beta vulgaris* ssp. *maritima*: mitochondrial variability in wild populations revealed by a rapid screening procedure. *Theor. Appl. Genet.*, **85**, 529-535.
- SCHNABEL, A. AND ASMUSSEN, M. A. 1992. Comparative effects of pollen and seed migration on the cytonuclear structure of plant populations. II. Paternal cytoplasmic inheritance. *Genetics*, **132**, 253-267.
- SLATKIN, M. 1985. Gene flow in natural populations. *Ann. Rev. Ecol. Syst.*, **16**, 393-430.
- SLATKIN, M. AND BARTON, N. H. 1989. A comparison of three indirect methods for estimating average levels of gene flow. *Evolution*, **43**, 1349-1368.
- SORK, V. L. 1984. Examination of seed dispersal and survival in red oak, *Quercus rubra* (Fagaceae) using metal tagged acorns. *Ecology*, **65**, 1020-1022.
- STRAUSS, S. H., HONG, Y.-P. AND HIPKINS, V. D. 1993. High levels of population differentiation for mitochondrial DNA haplotypes in *Pinus radiata*, *muricata* and *attenuata*. *Theor. Appl. Genet.*, **86**, 605-611.
- TAKAHATA, N. AND PALUMBI, S. R. 1985. Extranuclear differentiation and gene flow in the finite island model. *Genetics*, **109**, 441-457.
- WAGNER, D. B., FURNIER, G. R., SAGHAI-MAROOF, M. A., WILLIAMS, S. M., DANCIK, B. P. AND ALLARD, R. W. 1987. Chloroplast DNA polymorphisms in lodgepole and jack pines and their hybrids. *Proc. Natl. Acad. Sci. U.S.A.*, **84**, 2097-2100.
- WAGNER, D. B., DONG, J., CARLSON, M. R. AND YANCHUK, A. D. 1991. Paternal leakage of mitochondrial DNA in *Pinus*. *Theor. Appl. Genet.*, **82**, 510-514.
- WHEELER, N. C. AND GURIES, R. P. 1982. Population structure, genetic diversity and morphological variation in *Pinus contorta* Dougl. *Can. J. For. Res.*, **12**, 595-606.
- WRIGHT, S. 1951. The genetical structure of populations. *Ann. Eugen.*, **15**, 323-354.