Finite island model for organelle and nuclear genes in plants

R. J. PETIT*, A. KREMER & D. B. WAGNER†

*INRA, Laboratoire de Génétique et d'Amélioration des Arbres Forestiers, BP 45, 33611 Gazinet Cedex, France and
†Department of Forestry, University of Kentucky, Lexington, KY 40546-0073, U.S.A.*

Recurrence equations for genetic diversities and differentiation were developed for hermaphrodite plant species in an island model of population structure. This was made possible by the definitions of diversities at all hierarchical levels from gamete to total population and by the definition of migration rates specific to plants for both nuclear and cytoplasmic genomes. Mating system was also incorporated. Numerical computations were used to compare equilibrium values of differentiation obtained with our equations with those predicted by classical formulas. We show that the differences (sometimes high) result from the interpretations of the definition of gene diversity in a population of finite size. We interpret it as the probability that two genes sampled with replacement are different alleles (instead of without replacement). The effects of several parameters (ploidy level, mode of inheritance, outcrossing rate, population size) on genetic subdivision were evaluated. Contrary to the situation in animals, plant migration is intrinsically asymmetrical because a gene transmitted to the next generation through the male gamete may migrate in the pollen grain and in the seed, whereas a gene transmitted through the female gamete can migrate only in the seed. As a consequence, mode of inheritance (in the case of cytoplasmic genes) and outcrossing rate have strong impacts on subdivision, especially when pollen migration is larger than seed migration (a likely situation in many plant species). Parameters estimated in a survey of oak populations (*Quercus robur* L.) were used to examine whether our understanding of a real situation could be improved by the model. In particular, the rate of return to equilibrium was studied after a perturbation, i.e. a temporary decrease of population sizes (a bottle-neck).

Keywords: bottle-neck, differentiation, diversity, mating system, polyploidy, *Quercus robur*.

Introduction

An accurate description of the genetic structure of plant species, and especially of the geographic component of this structure, is desirable for conservation and breeding purposes. We have shown (Kremer *et al.*, 1991; Petit *et al.*, 1993) that the level of subdivision of organelle gene diversity among oak populations is very high, despite the fact that the species studied were characterized by low nuclear subdivision. A larger differentiation for cytoplasmic than for nuclear genes was reported for animal species (Desalle *et al.*, 1987; Hale & Singh, 1987) and was interpreted with the help of a finite island model for individuals with separate sexes (Birky *et al.*, 1989). Our objectives here are: (i) to extend the model to the case of hermaphrodite plants

with varying ploidy levels and outcrossing rates; (ii) to quantify the impact of several variables on the magnitude of genetic differentiation; and (iii) to apply the model to a real case, namely the above mentioned study of oak populations.

The theoretical approach directly derives from the classical Wright–Fisher model for nuclear genes at a single locus with neutral alleles (Wright, 1943). Genes are sampled at a discrete generation n to make the next, non-overlapping generation, $n + 1$. Drift, mutation and migration can be included. However, if mutation is much smaller than migration, it will have little effect on the level of subdivision (Gst) (Crow & Aoki, 1984; Birky *et al.*, 1989). As our focus here is on *genetic subdivision* (or *differentiation among populations*), we will not consider the effect of mutation. The parameter of genetic subdivision used (G_{st}), an extended version of Wright's *Fst* in the case of multiple alleles (Nei, 1973), was shown to *reach equilibrium very*

*Correspondence.

quickly (Crow & Aoki, 1984). This is of particular interest as indirect methods to estimate gene flow from values of differentiation rely on the assumption of equilibrium (Slatkin, 1985). Nevertheless, some populations of plants are subjected to frequent disturbances and it is seldom clear whether the equilibrium assumption is valid. This is especially problematic in the study of many forest trees, where climatic shifts that have modified the ranges of taxa during the Quaternary have occurred at a rapid pace relative to mean generation times. There is at present no straightforward method to test that equilibrium conditions have indeed been reached. We show here that dynamic study of both nuclear and cytoplasmic differentiation of a species may cast some new light on this problem.

Definitions

The general outline of the study is provided by Wright's *island model* (1943) using the following notations and assumptions. There are L (sub) populations (colonies, demes or islands). The mixed mating model is used; there is self-fertilization with probability $1 - t$ and random mating within the population with probability t. The populations all have the same size (N individuals).

Effective migration rates

For nuclear genes in animals with separate sexes, the effective migration rate which has been used by Takahata & Palumbi (1985) and Birky *et al.* (1989) is the sum of the male and female migration rates where these are defined as the proportion of migrating males (females) relative to the total population size (including both males and females):

$$m_e = m_f + m_m.$$

For cytoplasmic genes, the effective migration rate used by the same authors was:

$$m_{eo} = \alpha m_f + \beta m_m ,$$

where α is the fraction of cytoplasmic genomes in a zygote which originates from the female gamete and β is the fraction which originates from the male gamete ($\alpha + \beta = 1$). This omits the transient contribution of cytoplasmic genomes which will not be transmitted to the next generation. With balanced sex ratios, m_f and m_m each have a maximum value of 0.5, and m_{eo} is then bounded by 0.5. The cytoplasmic effective migration rate may approach its maximum value of 1 only when the sex ratio is strongly skewed in the direction of the

sex transmitting the cytoplasmic genome (e.g. a large female excess when $\alpha = 1$).

In the case of plants, migration may occur through gametes (pollen) and/or seeds. To define migration rates which can be compared with each other, we define (artificially, as female gametes are not actually mobile in plants) migration rates for male and female gametes. We first define $m_{\gamma f}$ and $m_{\gamma m}$ as the proportion of gametes of each sex which migrate, relative to the total number of all gametes. Depending on the power of the investigation, the gametes which are considered may be limited to the successful gametes or may include all gametes produced. The latter definition will, however, rarely be useful. An effective gamete migration rate will be, for nuclear genes:

$$m_{\gamma e} = m_{\gamma f} + m_{\gamma m},$$

and for cytoplasmic genes:

$$m_{\gamma eo} = \alpha m_{\gamma f} + \beta m_{\gamma m}.$$

These expressions can be related to the pollen and seed migration rates. A fertilized ovule may migrate in a seed with probability m_s (i.e. m_s is defined as the proportion of seeds which migrate, relative to the total number of seeds produced). If we consider in the above definitions only the successful gametes, then $m_{\gamma f} = m_s/2$ as there will be exactly half female and half male gametes and a female gamete will migrate with probability m_s. For the same reason, $m_{\gamma m} = 1/2(m_p + m_s - m_p m_s)$, where m_p is defined as the proportion of pollen grains that migrate, relative to the total number of successful pollen grains. This is because a male gamete may migrate in three different ways.

by pollen only: $m_p(1 - m_s)$;
by seed only: $(1 - m_p)m_s$;
by both: $m_p m_s$.
Total: $m_p + m_s - m_p m_s$.

Finally, for nuclear genes:

$$m_e = 1/2 m_s + 1/2(m_s + m_p - m_p m_s) = m_s$$
$$+ 1/2(m_p - m_p m_s) \tag{1}$$

and for cytoplasmic genes:

$$m_{eo} = \alpha m_s + \beta(m_s + m_p - m_p m_s) = m_s + \beta(m_p - m_p m_s). \tag{2}$$

Because such a definition greatly simplifies the recurrence equations, we define the migration rates as the fraction exchanged each generation by each subpopulation with the *entire population*, following Maruyama (1970). A certain proportion of individuals are then expected to be returned to the population from which they originally came.

Gene diversity parameters

The gene diversity K considered by Birky *et al.* (1989) is the probability that, at a given locus, two randomly sampled genes will be genotypically different alleles. All hierarchical levels, from zygote to the total population, were considered. We will consider in addition the gametic level. Thus, the K parameters are:

(1) K_g for sampling two distinct genes (i.e. sampling without replacement) from a single gamete;

(2) K_z for sampling two distinct genes from a single zygote;

(3) K_a for sampling two distinct genes from a single adult cell;

(4) K_b for sampling two genes from two distinct adult cells in the same individual;

(5) K_c for sampling two genes from different individuals in a population; and

(6) K_d for sampling two genes from different populations within an entire set of populations.

Because two distinct nuclear genes cannot be sampled form a single haploid gamete, K_g can be used only for diploid gametes (or gametes characterized by higher ploidy levels) and for cytoplasmic genes (because, even if the cytoplasmic genes are homoplasmic, there are several copies of these genes per gamete).

Actually, Birky *et al.* (1989) considered two kinds of sampling within a given hierarchical level, i.e. sampling in two different units (K), or sampling in two different units or in the same unit (K^*). We extend here these notations to all hierarchical levels, from the single cell to the total population. As defined, the K^* parameters correspond to the original definition of diversity (Nei, 1973). Indeed, the formula given by Nei (1973) for diversity $(h = 1 - \sum_{i=1}^{m} p_i^2$, where the sum is made over the m alleles and p_i is the frequency of allele i), corresponds to sampling with replacement of the genes, in a population of finite size. If sampling is made without replacement, the marginal allel frequencies are modified after the sampling of the first gene.

The K^* parameters are:

(1) K_g^* for sampling two genes within a gamete (the genes may be distinct or not);

(2) K_z^* for sampling two genes within a zygote;

(3) K_a^* for sampling two genes within a single adult cell;

(4) K_b^* for sampling two genes from two adult cells within an individual (the cells may be distinct or not):

$$K_b^* = \{K_{a+Kb}^*(C-1)\}/C, \tag{3}$$

where C is the number of cells;

(5) K_c^* for sampling two genes within a population (the two genes may be sampled from the same individual or from distinct individuals):

$$K_c^* = \{K_b^* + K_c(N-1)\}/N, \tag{4}$$

where N is the number of individuals;

(6) $K_d^* (= K_t$ of Birky *et al.*, 1989) for sampling two genes from anywhere in the total population:

$$K_t = \{K_c^* + K_d(L-1)\}/L, \tag{5}$$

where L is the number of subpopulations.

In the following, we will not consider the case of heteroplasmy, whether for nuclear or for cytoplasmic genes. Consequently, we have $K_a^* = K_z^* = K_b^*$ but $K_a = K_z \neq K_b$. We will use only K_z^* and K_z. In this case, eqn 4 can be re-written as:

$$K_c^* = \{K_z^* + K_c(N-1)\}/N. \tag{6}$$

Subdivision

We use G_{st}, the fractional reduction of diversity due to subdivision (Nei, 1973), as a measure of differentiation. With our definitions of diversity parameters, we have:

$$G_{st}^{(1)} = (K_t - K_c^*)/K_t. \tag{7}$$

We used this definition for both cytoplasmic and nuclear genes. However, Birky *et al.* (1989) used a different definition for nuclear genes, namely:

$$G_{st}^{(2)} = (K_t^{(2)} - K_c)/K_t^{(2)}, \tag{8}$$

where $K_t^{(2)} = \{K_c + K_d(L-1)\}/L$. We will compare both definitions of G_{st}.

Results

For comparative purposes, we derived recurrence equations for hermaphrodite animals first, and then for hermaphrodite plants. Tables 1–4 can be understood with the following considerations.

Inheritance. Because we are dealing with sexually reproducing organisms, sampled genes originate from gametes. The parameters α and β define the rules of inheritance of the genes to be studied (in particular, $\alpha \cong \beta \cong 0.5$ for nuclear genes). Hence, two genes will come from an egg with probability α^2, both will come from a sperm cell with probability β^2, and one will come from an egg and the other from a sperm cell with probability $2\alpha\beta$.

Mating system. A pollen grain may have fertilized the individual which produced it, with probability $1 - t$, or it may have fertilized at random any individual in the

population (including the individual which produced it), with probability t.

Migration. As seen in the definitions of the effective migration rates, there is a 'two-step' migration process in the case of genes coming from male gametes in plants, wherein these genes can migrate by pollen before fertilization and by seed after fertilization. However, note that genes coming from male gametes migrate only by seed and not by pollen if they fertilize the plant that produced them.

Genes sampled in different individuals. Let K'_c and K'_d be the diversities K_c and K_d at generation $n + 1$. If *at least* one of two sampled alleles has migrated, then, regardless of where these alleles were sampled at generation $n + 1$ (i.e. in two different individuals of the same or of different populations), the probability that they are genotypically different is K^*_d ($= K_t$). This is because a gene may migrate into the population from which it originated (c.f. definitions), and more than one migration event does not affect this probability. This was considered differently in the work of Birky *et al.* (1989, p. 616); our treatment has the advantage of being more consistent with the definition of the migration rate. If neither of the genes has migrated, then the probability that they are different alleles will depend on where they were sampled at generation $n + 1$ (i.e. K_d if they were sampled in different populations, or K^*_c if they were sampled from different individuals in the same population (this is because the genes may have belonged to the same individual in generation n)).

Genes sampled in the same individual (Tables 1 and 2). We focus on $K^{*\prime}_z$ instead of K'_z (Birky *et al.*, 1989) for two reasons (the ($'$) again denotes values in the next generation). Firstly, the analogy of $K^{*\prime}_z$ with K'_c and K'_d is then more obvious, as the possibility that the two sampled genes originated from gametes of the same sex is retained when $K^{*\prime}_z$ is used. Secondly, K^*_z appears to be more informative than K_z. Indeed, consider the following example: the maximum value of K_z is one at a diploid locus with two different alleles or at a tetraploid locus where all four alleles are different. If we use K^*_z instead, the diversity will be higher in the tetraploid (0.75) than in the diploid (0.5), allowing the distinction of the two cases. The maximum value of one will only be approximated when a gamete has a large number of copies of a gene each characterized by a different allele. This definition of diversity corresponds to the original concept of Nei (1973) as explained above.

If both genes sampled come from a male gamete (with probability β^2) or both from a female gamete (with probability α^2), they come from the *same* gamete

because a zygote is made up of only one gamete of each sex but they may still be different alleles in the case of polyploidy or heteroplasmy. Without polyploidy or heteroplasmy, $K^*_g = 0$ (gametes are haploid for the nuclear locus considered and gametes are also effectively haploid for cytoplasmic genes because we do not consider heteroplasmy). When the two genes are sampled from one male and one female gamete, they will be different alleles with probability K^*_z if the male gamete has fertilized the same plant (with probability $1 - t$); otherwise (with probability t) they will be different with probabilities K^*_d or K^*_c, depending on the migration events, as seen above for K_c or K_d.

Genes sampled in the same gamete. Consider the sampling of nuclear genes from a zygote to make a gamete in an autopolyploid species. It must in theory lie between the two following extremes: sampling with replacement or sampling without replacement of half of the number of copies of each gene (two copies for a tetraploid species, three for an hexaploid species, etc.). For a gene located on the centromere, all the copies in the gamete will be from non-identical chromatids (i.e. sampling without replacement). However, as a consequence of pseudoreductional meiotic events (Demarly, 1963), sister chromatids can be found in the same gamete in polyploid species. Although the true situation is probably closer to the case of sampling without replacement, it will only be possible to derive an exact solution when the frequency of such meiotic events is known. We will limit ourselves here to the presentation of both extreme cases. Firstly, if we consider sampling without replacement, we have $K_g = K_z$, $K^*_z = K_z(2x - 1)/2x$ where $2x$ indicates the ploidy level (e.g. $2x = 4$ for an autotetraploid species), and $K^*_g = K_g(x - 1)/x$, hence $K^*_g = 2K^*_z(x - 1)/(2x - 1)$. In the case of sampling with replacement, $K_g = K^*_z$ instead, and $K^*_g = K^*_z(x - 1)/x$.

Table 1 Principle of the derivation of a recurrence equation for K^*_z (zygotic diversity) at nuclear genes in hermaphrodite animals. The gene diversity for the next generation depends on the K values in the preceding generation, weighted by a probability (product of the terms situated on the left). Expansion of these terms gives eqn 9

					$K^{*\prime}_z$
α^2					K^*_g
$2\alpha\beta$	t		m^2		K^*_d
$2\alpha\beta$	t	2	m	$(1 - m)$	K^*_d
$2\alpha\beta$	t			$(1 - m)^2$	K^*_c
$2\alpha\beta$	$(1 - t)$				K^*_z
β^2					K^*_g

Recurrence equations can be derived by summing the products of the terms situated horizontally with the corresponding K parameter in Table 1 (animals) and Table 3 (plants) for $K_z^{*\prime}$, and Table 2 (animals) and Table 4 (plants) for $K_c^{\prime}$ and $K_d^{\prime}$ (presented together in the Tables). For nuclear genes, $\alpha = \beta = 1/2$, and $K_g^* = 0$ for diploid organisms. For cytoplasmic genes, in the absence of heteroplasmy, we have $K_z^* = 0$ (i.e. we assume rapid segregation in the case of biparental inheritance). Except for this difference, recurrence equations for cytoplasmic and nuclear diversities are expressed in the same way. The recurrence equations calculated using Tables 1–4 greatly simplify. We obtain the following results:

(i) Hermaphrodite animals
From Table 1, we have:

$$K_z^{*\prime} = K_g^*(\alpha^2 + \beta^2) + 2\alpha\beta\{K_z^*(1-t) + K_c^*t(1-m)^2 + K_d^*tm(2-m)\}, \tag{9}$$

and therefore

$$K_z^{\prime} = K_z^*(1-t) + K_c^*t(1-m)^2 + K_d^*tm(2-m). \tag{10}$$

From Table 2:

$$K_c^{\prime} = K_c^*(1-m)^2 + K_d^*m(2-m) \text{ and} \tag{11}$$

$$K_d^{\prime} = K_d(1-m)^2 + K_d^*m(2-m). \tag{12}$$

$K_c^{*\prime}$ can be obtained from eqn 6, $K_t^{\prime}$ from eqn 5 and $G_{st}^{\prime}$ from eqn 7 or 8. Note that K_z will be identical to K_c only when $t = 1$.

(ii) Hermaphrodite plants
From Table 3, we have:

$$K_z^{*\prime} = K_g^*(\alpha^2 + \beta^2) + 2\alpha\beta[K_z^*(1-t) + K_c^*t(1-m_s)^2 \times (1-m_p) + K_d^*t\{m_s(2-m_s) + (1-m_s)^2 m_p\}] \tag{13}$$

and

$$K_z^{\prime} = K_z^*(1-t) + K_c^*t(1-m_s)^2(1-m_p) + K_d^*t\{m_s(2-m_s) + (1-m_s)^2 m_p\}. \tag{14}$$

Table 2 Principle of the derivation of recurrence equations for K_c and K_d (diversity within and among populations) in hermaphrodite animals. Expansion of these terms gives eqns 11 and 12

				$K_c^{\prime}$	$K_d^{\prime}$
		m^2		K_d^*	K_d^*
2		m	$(1-m)$	K_d^*	K_d^*
			$(1-m)^2$	K_c^*	K_d

Table 3 Principle of the derivation of a recurrence equation for K_z^* at nuclear genes in hermaphrodite plants. Expansion of these terms gives eqn 13

						$K_z^{*\prime}$
α^2						K_g^*
$2\alpha\beta$	t		m_s^2	m_p		K_d^*
$2\alpha\beta$	t		m_s^2		$(1-m_p)$	K_d^*
$2\alpha\beta$	t	2	m_s	$(1-m_s)m_p$		K_d^*
$2\alpha\beta$	t	2	m_s	$(1-m_s)$	$(1-m_p)$	K_d^*
$2\alpha\beta$	t			$(1-m_s)^2 m_p$		K_d^*
$2\alpha\beta$	t			$(1-m_s)^2$	$(1-m_p)$	K_c^*
$2\alpha\beta$	$(1-t)$					K_z^*
β^2						K_g^*

Table 4 Principle of the derivation of recurrence equations for K_c and K_d (diversity within and among populations) in hermaphrodite plants. Expansion of these terms gives eqns 15 and 16

						$K_c^{\prime}$	$K_d^{\prime}$
α^2			m_s^2			K_d^*	K_d^*
α^2		2	m_s	$(1-m_s)$		K_d^*	K_d^*
α^2				$(1-m_s)^2$		K_c^*	K_d
$2\alpha\beta$	t		m_s^2		m_p	K_d^*	K_d^*
$2\alpha\beta$	t		m_s^2		$(1-m_p)$	K_d^*	K_d^*
$2\alpha\beta$	t	2	m_s	$(1-m_s)$	m_p	K_d^*	K_d^*
$2\alpha\beta$	t	2	m_s	$(1-m_s)$	$(1-m_p)$	K_d^*	K_d^*
$2\alpha\beta$	t			$(1-m_s)^2$	m_p	K_d^*	K_d^*
$2\alpha\beta$	t			$(1-m_s)^2$	$(1-m_p)$	K_c^*	K_d
$2\alpha\beta$	$(1-t)$		m_s^2			K_d^*	K_d^*
$2\alpha\beta$	$(1-t)$	2	m_s	$(1-m_s)$		K_d^*	K_d^*
$2\alpha\beta$	$(1-t)$			$(1-m_s)^2$		K_c^*	K_d
β^2	t^2		m_s^2		m_p^2	K_d^*	K_d^*
β^2	t^2	2	m_s^2		$m_p(1-m_p)$	K_d^*	K_d^*
β^2	t^2		m_s^2		$(1-m_p)^2$	K_d^*	K_d^*
β^2	t^2	2	m_s	$(1-m_s)$	m_p^2	K_d^*	K_d^*
β^2	t^2	4	m_s	$(1-m_s)$	$m_p(1-m_p)$	K_d^*	K_d^*
β^2	t^2	2	m_s	$(1-m_s)$	$(1-m_p)^2$	K_d^*	K_d^*
β^2	t^2			$(1-m_s)^2$	m_p^s	K_d^*	K_d^*
β^2	t^2	2		$(1-m_s)^2$	$m_p(1-m_p)$	K_d^*	K_d^*
β^2	t^2			$(1-m_s)^2$	$(1-m_p)^2$	K_c^*	K_d
β^2	$2t(1-t)$		m_s^2		m_p	K_d^*	K_d^*
β^2	$2t(1-t)$		m_s^2		$(1-m_p)$	K_d^*	K_d^*
β^2	$2t(1-t)$	2	m_s	$(1-m_s)$	m_p	K_d^*	K_d^*
β^2	$2t(1-t)$	2	m_s	$(1-m_s)$	$(1-m_p)$	K_d^*	K_d^*
β^2	$2t(1-t)$			$(1-m_s)^2$	m_p	K_d^*	K_d^*
β^2	$2t(1-t)$			$(1-m_s)^2$	$(1-m_p)$	K_c^*	K_d
β^2	$(1-t)^2$		m_s^2			K_d^*	K_d^*
β^2	$(1-t)^2$	2	m_s	$(1-m_s)$		K_d^*	K_d^*
β^2	$(1-t)^2$			$(1-m_s)^2$		K_c^*	K_d

From Table 4:

$$K_c' = K_c^*(1 - m_s)^2(1 - \beta t m_p)^2 + K_d^*\{m_s(2 - m_s)$$
$$+ \beta t(1 - m_s)^2 m_p(2 - \beta t m_p)\}, \qquad (15)$$

$$K_d' = K_d(1 - m_s)^2(1 - \beta t m_p)^2 + K_d^*\{m_s(2 - m_s)$$
$$+ \beta t(1 - m_s)^2 m_p(2 - \beta t m_p)\}. \qquad (16)$$

As for hermaphrodite animals, $K_c^{*\prime}$ can be obtained from eqn 6, K_t' from eqn 5 and G_{st}' from eqn 7 or 8. Note that K_z will be different from K_c even when $t = 1$, contrary to the case for hermaphrodite animals.

Numerical analyses

We are interested in evaluating the evolution of population subdivision. We solved the recurrence equations numerically, starting from populations with extreme non-equilibrium conditions and studying the evolution of K parameters and of G_{st}. We first compared numerical results of G_{st} obtained with our recurrence equations for hermaphrodite animals and for plants with those obtained using an equilibrium formula for G_{st}. We then focused on plants only and studied the effects

of several important parameters on nuclear and cytoplasmic subdivision, i.e. level of ploidy, mode of inheritance and outcrossing rate. Finally, we considered a tree species for which we have empirical results on nuclear and cytoplasmic differentiation: *Quercus robur* L. (the pedunculate oak). We tested the effect of population size on the rate of approach to equilibrium for nuclear and cytoplasmic genes, by simulating a population bottle-neck and then allowing populations to grow back to their carrying capacity. This was intended to mimic the early phase of post-glacial recolonization of Europe by this species.

Equilibrium values of G_{st}

We consider a diploid nuclear gene, either in hermaphrodite animals or in plants, and several sets of values for migration and effective population size. We checked numerically that $G_{st}^{(1)}$ and $G_{st}^{(2)}$ (calculated using eqns 7 and 8) had reached equilibrium. After 1000 generations, we had in all cases studied $|G_{st}' - G_{st}| \leq 10^{-15}$, whichever formula for subdivision was used. We then compared the values of $G_{st}^{(1)}$ and $G_{st}^{(2)}$ with each other and with $G_{st}^{(3)} \cong 1/[1 + 2N(L/L - 1)\{1/$

Table 5 Subdivision at equilibrium for different effective migration rates (m_e) and effective population sizes (N) in hermaphrodite animals or plants. The subdivision (G_{st}) corresponds to the ratio of interpopulation diversity relative to the total diversity, where the diversities are defined as the probability that two sampled genes are different alleles, using sampling with replacement $(G_{st}^{(1)})$ or without replacement $(G_{st}^{(2)})$ of the individuals, and $G_{st}^{(3)}$ is derived using the formula of Takahata & Nei (1984) for differentiation at equilibrium. When the same effective migration rates are used for animals and plants, the results are nearly identical. In both cases, it is important to specify exactly which definition of diversity is used to compute G_{st}, especially with high migration rates

	m_s	m_p	m_e	N	$G_{st}^{(1)}$	$G_{st}^{(2)}$	$G_{st}^{(3)}$	$G_{st}^{(1)}/G_{st}^{(2)}$
Animals			0.01495	10	0.63768	0.61875	0.61570	1.03
			0.01495	100	0.14267	0.13844	0.13809	1.03
			0.01495	1000	0.01626	0.01578	0.01577	1.03
			0.145	10	0.16123	0.11786	0.11752	1.37
			0.145	100	0.01798	0.01315	0.01314	1.37
			0.145	1000	0.00182	0.00133	0.00133	1.37
			0.28	10	0.09683	0.05020	0.05010	1.93
			0.28	100	0.01012	0.00525	0.00525	1.93
			0.28	1000	0.00102	0.00053	0.00053	1.93
Plants	0.01	0.01	0.01495	10	0.63767	0.61874	0.61570	1.03
	0.01	0.01	0.01495	100	0.14267	0.13844	0.13809	1.03
	0.01	0.01	0.01495	1000	0.01626	0.01578	0.01577	1.03
	0.1	0.1	0.145	10	0.16118	0.11782	0.11752	1.37
	0.1	0.1	0.145	100	0.01798	0.01315	0.01314	1.37
	0.1	0.1	0.145	1000	0.00182	0.00133	0.00133	1.37
	0.2	0.2	0.28	10	0.09677	0.05016	0.05010	1.93
	0.2	0.2	0.28	100	0.01012	0.00525	0.00525	1.93
	0.2	0.2	0.28	1000	0.00102	0.00053	0.00053	1.93

$(1 - m_e)^2 - 1\}]$ (eqn 4 of Takahata & Nei, 1984, using our symbols and assuming no mutation). Note that their definition of the migration rate is the same as ours. The results indicate that when the same effective migration rates are used, equilibrium G_{st} for hermaphrodite animals and plants are nearly identical (Table 5). In addition, $G_{st}^{(3)}$ closely approximates $G_{st}^{(2)}$ in all cases. Actually, the formula for $G_{st}^{(3)}$ above was derived by Nei (1975) by considering the sampling of *different genes* for the computation of diversity and hence differentiation; this is equivalent, in the case of randomly mating hermaphrodite animals (where $K_z = K_c$, as discussed above) to the sampling of *different individuals*, as in the definition of $G_{st}^{(2)}$.

On the contrary, $G_{st}^{(1)}$ is always larger than $G_{st}^{(2)}$ and $G_{st}^{(3)}$, especially in the case of high migration rate (by 37 per cent in the case where $m_e = 0.145$ and $t = 1$, and by 93 per cent when $m_e = 0.28$, regardless of the value of N). This indicates the importance of the measure of diversity considered in the computation of differentiation, despite the fact that the measure of within-population diversity itself is only slightly modified (except when N is very small). Empirical studies where Nei's (1973) concept of diversity is used actually present unbiased estimates of

$$h = 1 - \sum_{i=1}^{m} p_i^2 \,,$$

which corresponds to our K^* parameters, as discussed previously. Consequently, the definition of $G_{st}^{(1)}$ (in eqn 7) appears to be more relevant to these empirical studies. It would therefore be useful to develop new equilibrium formulas for the genetic differentiation, to which numerical and empirical results may be compared, and which would in addition include the outcrossing rate t, for animals as well as for plants. In the remaining analyses, we study only $G_{st}^{(1)}$, which will be called G_{st} for simplicity.

Effect of varying parameter values on nuclear and cytoplasmic subdivision

Several parameters were kept constant throughout. L (number of populations) was fixed at 50. The product Nm_e was also kept constant $(Nm_e = 11.3)$, based on results obtained in a study of oaks (Müller-Starck & Ziehe, 1991). We used $N = 100$ (effective number of individuals) as well as identical pollen and seed migration rates $(m_p = m_s = 0.077)$ or a high pollen/seed migration ratio $(m_p/m_s = 557$, which yields $m_p = 0.225$ and $m_s = 0.0004$ in the case where $N = 100$), as derived from relative empirical values of organelle and nuclear gene differentiation (Petit, 1992). Previous esti-

mates of nuclear (Zanetto *et al.*, 1993) and cytoplasmic (Petit *et al.*, 1993) diversities for *Q. robur* were used as initial values in the recurrence equations (except in the studies of ploidy level and mode of inheritance, where both diversities were set to the same value as the goal was to compare nuclear and cytoplasmic differentiation).

Ploidy level. We have illustrated the dedifferentiation (the evolution towards a genetically less structured organization of diversity among populations, i.e. starting from a G_{st} higher than the equilibrium one) of a biparentally inherited cytoplasmic gene, for which rapid vegetative segregation is responsible for a complete absence of heteroplasmy (Fig. 1).

This case can be compared with a diploid nuclear (i.e. biparentally inherited) gene. The rates of approach to equilibrium are similar but the level of subdivision at equilibrium is higher for the cytoplasmic gene. If we now consider a tetraploid gene, where the two genes which make up the diploid gametes are sampled either with or without replacement (discussion above) from the four genes of the sporophyte, we again note a decrease of the value of subdivision at equilibrium with an
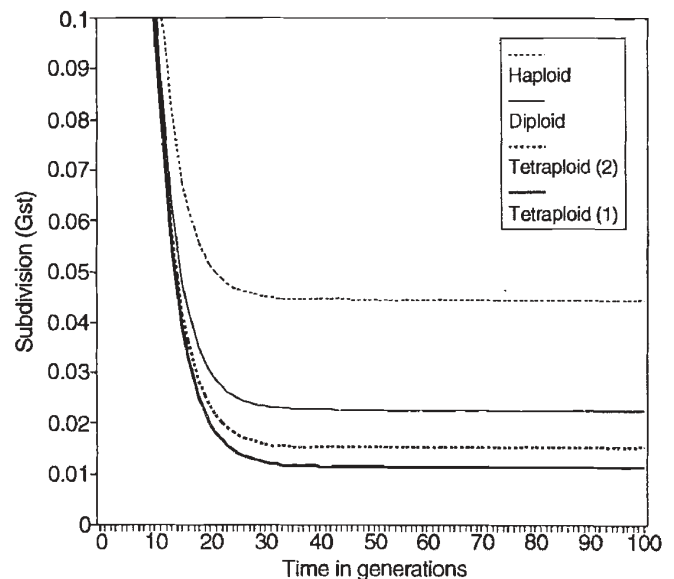

Fig. 1 Effect of the ploidy level on genetic subdivision in populations of hermaphrodite plants. G_{st} values for nuclear genes (in the case of a diploid or a tetraploid species) and for a biparentally inherited cytoplasmic gene ($\alpha = \beta = 0.5$) are plotted vs. time. Tetraploid (1): the two genes of the diploid gamete are sampled without replacement among the four genes of the sporophyte. Tetraploid (2): with replacement. Seed and pollen migration rates are equal ($m_p = m_s = 0.077$). $N = 100$, $t = 1$, initial conditions: $K_c = 0$, $K_d = 0.65$ ($G_{st} = 1$) for both nuclear and organelle genes.

increase in the ploidy level. The type of polyploid gametes that can be produced (sampling with replacement vs. sampling without replacement) also has an effect on the value of differentiation at equilibrium as sampling of genes without replacement from the sporophyte yields a lower subdivision at equilibrium than sampling with replacement.

Mode of inheritance. In plants, all possible types of cytoplasmic inheritance (i.e. maternal, biparental and paternal) have been reported. We have compared the dedifferentiation of cytoplasmic genes (with three modes of transmission, $a = 1$, 0.5 or 0) and nuclear genes ($a = 0.5$), using first identical seed and pollen migration rates (Fig. 2). Although the differences are small at equilibrium, paternally inherited genes show the lowest subdivision among cytoplasmic genes. This is clearly a result of the two-step process of paternally transmitted genes, which migrate with both pollen and seed. The fastest rate of decrease of G_{st} is also shown by paternally inherited genes. However, around the 17th generation, in this example, the nuclear G_{st} (G_{st_n}) becomes smaller than the cytoplasmic G_{st} (G_{st_c}). Biparentally inherited cytoplasmic genes show exactly the same initial behaviour as nuclear genes but again nuclear subdivision is lower at equilibrium.

Using a high m_p/m_s ratio, we see that the transmission rate has a much greater impact on the equilibrium values of cytoplasmic subdivision (Fig. 3). Although G_{st_c} is still very low at equilibrium for paternally inherited genes, it is now much higher for maternally inherited genes. If we compare strict maternal inheritance with maternally biased biparental inheritance ($a = 0.95$ or 0.99), it becomes clear that even a small paternal leakage of cytoplasmic genes has a strong impact on subdivision.

To understand the fact that G_{st_c} for paternally inherited genes may be lower than G_{st_n} during the initial non-equilibrium phase, although it is always higher at equilibrium, the formula of Crow & Aoki (1984) for the time required for G_{st_n} to go half way to equilibrium is helpful:

$$t_{1/2}^{(Nuc)} = Ln2/\{2m_e + 1/(2N_e)\}.$$

Similarly, for cytoplasmic genes, Birky *et al.* (1989) found:

$$t_{1/2}^{(Cyt)} = Ln2/(2m_{eo} + 1/N_{eo}).$$

At equilibrium, G_{st} depends on the absolute value of the product of the effective number of genes by the effective migration rate: $2N_e m_e$ for nuclear genes and $N_{eo} m_{eo}$ for cytoplasmic genes (Birky *et al.*, 1989). High values of the effective migration rate increase the rate of approach to equilibrium. For a paternally inherited
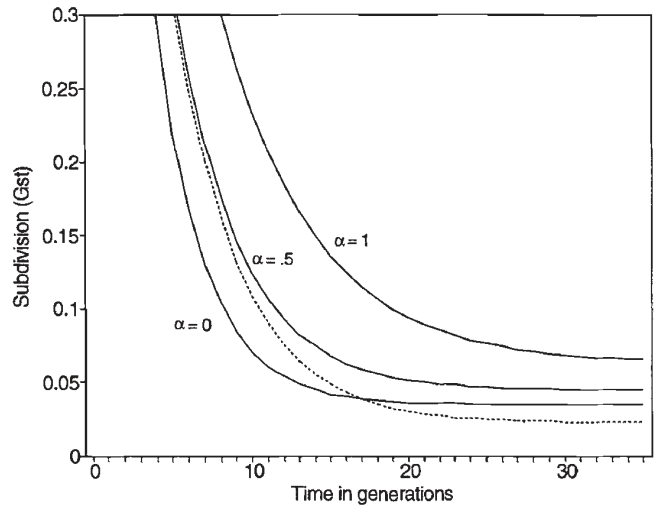

Fig. 2 Effect of the mode of inheritance on genetic subdivision in populations of hermaphrodite plants. G_{st} values for cytoplasmic genes (showing maternal, paternal or biparental inheritance) (——) and for a diploid nuclear gene (·····) are plotted vs. time. Seed and pollen migration rates are equal ($m_p = m_s = 0.077$). $N = 100$, $t = 1$, initial conditions: $K_c = 0$, $K_d = 0.65$ ($G_{st} = 1$) for nuclear and organelle genes.
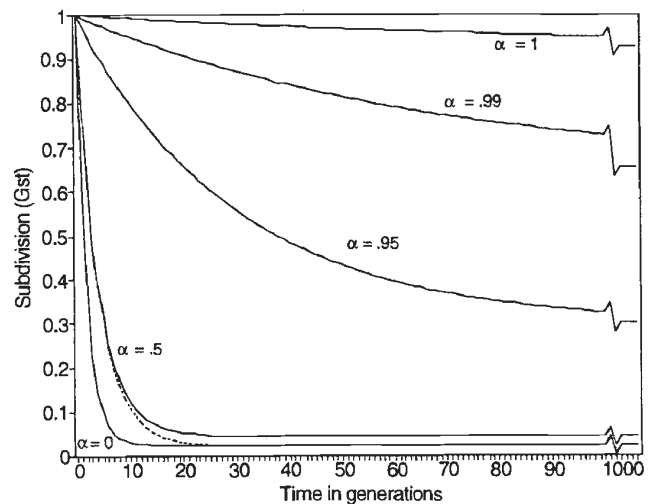

Fig. 3 Effect of the mode of inheritance on genetic subdivision in populations of hermaphrodite plants. Same as Fig. 2 except that seed and pollen migration rates are very different ($m_p = 0.225$, $m_s = 0.0004$). $N = 100$, $t = 1$, initial conditions: $K_c = 0$, $K_d = 0.65$ ($G_{st} = 1$) for nuclear and organelle genes.

gene, $m_{eo} = m_s + m_p - m_p m_s$ whereas, for nuclear genes, $m_e = m_s + 0.5(m_p - m_p m_s)$. Hence, we have $m_{eo} > m_e$ (and $t_{1/2}^{(Cyt)} < t_{1/2}^{(Nuc)}$) but $N_{eo} m_{eo} < 2N_e m_e$. For dedifferentiating populations, G_{st_c} will reach equilibrium more quickly and could therefore become

initially lower than G_{st_c}, but ultimately G_{st_n} becomes lower. For differentiating populations, on the other hand, G_{st_c} will again reach equilibrium more quickly and it will therefore remain higher than G_{st_n} from the outset. As a consequence, for conifer trees (characterized by paternal chloroplast inheritance; Chesnoy, 1987), a lower subdivision for chloroplast genes than for nuclear genes may indicate *a non-equilibrium situation where dedifferentiation is occurring.*

Outcrossing rate. For strictly maternally inherited cytoplasmic genes, t does not affect the values of gene diversities and subdivision (see eqns 13–16). However, t will affect the evolution of diversity at biparentally or paternally inherited cytoplasmic genes. The effect of t on strictly paternally inherited cytoplasmic genes is particularly important since all genes will be affected by a decrease in t (note that only those genes deriving from male gametes are affected by t in the case of biparental (e.g. nuclear) inheritance). From eqns 15 and 16, we can see that transmission and outcrossing rates are always present in the form of their product βt. Although this is not so in eqn 13, it does indicate their nearly parallel effect on migration and differentiation. For instance, a paternally inherited gene ($\alpha = 0$; $\beta = 1$) in a partial ($t = 0.05$) outcrosser was shown to display the same evolution of differentiation as a gene showing biparental ($\alpha = 0.95$; $\beta = 0.05$) inheritance in a strict outcrosser ($t = 1$) (results not shown). Hence, in a species characterized by a high pollen/seed migration ratio, a small amount of outcrossing will strongly affect differentiation, as does a small paternal leakage of cytoplasmic genes in a strict outcrosser (Fig. 3).

Effects of a bottle-neck and of population size

Our goal here is to test the effects of several possible values of population size in oaks. Postglacial population expansions of several forest tree species, including oaks, have been inferred from pollen data (Bennett, 1983). Current oak populations must have started with only a few individuals at the outset of postglacial recolonization. Using Bennett's data, we developed a *logistic population growth model* to allow directional changes in the population sizes from the beginning of postglacial recolonization. The equation we used for population growth is: $N_{t+1} = N_t R/\{1 + (R-1)N_t/K\}$, with R (reproductive rate) = 2. Indeed, sexual maturity for this oak species may be reached within 50 years (probably faster in a colonizing phase and slower during successional phases), and doubling time found by Bennett (1983) was 60–73 years in the initial phase of oak (probably *Quercus robur*) range expansion.

We assumed strict maternal inheritance of the chloroplast genome and random mating (the outcrossing rate for this species has been shown to be close to one (Bacilieri *et al.*, 1993)). We used several values of K (asymptotic value of population size) ranging from 30 to 50 000, starting with equilibrium conditions ($N = K$) and letting N drop to five individuals during the first generation of recolonization (simultaneously in all populations). The logistic model predicts about 1200 years for population size to reach 95 per cent of its value at equilibrium (K), which is consistent with the 1140 years duration of *Q. robur* expansion in Norfolk described by Bennett (1983).

Figure 4 shows how such bottle-necks may affect nuclear gene differentiation. Owing to accelerated drift, subdivision overshoots its equilibrium value during the logistic growth phase. The evolution back to equilibrium is very slow for K higher than 200. As observed nuclear subdivision is actually very low among populations of western Europe (Zanetto *et al.*, 1993), the effective population sizes of *Q. robur* may be smaller than 200 if the hypotheses underlying the model are correct.

For maternally inherited cytoplasmic genes, a bottle-neck will also increase the level of differentiation above the equilibrium value (Fig. 5). With K larger than 50, non-equilibrium conditions will prevail for cytoplasmic genes. However, as G_{st_c} was already high before the
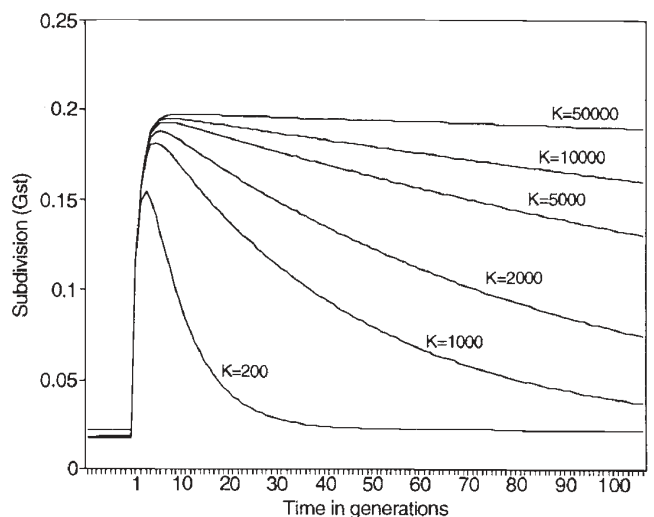
Fig. 4 Effect of population sizes at carrying capacity on the rate of approach of nuclear genetic subdivision to equilibrium in populations of hermaphrodite plants after a bottle-neck ($N = 5$). Before the bottle-neck, equilibrium conditions for subdivision were assumed. For each value of population size N at carrying capacity, the migration rates were determined assuming that $Nm_c = 11.3$ and $m_p/m_s = 557$. $t = 1$, initial conditions. $K_c = 0.373$, $K_d = 0.38$ ($G_{st} \cong 0.02$).

bottle-neck, the absolute and relative increases in G_{st_c} are limited.

Note that initial G_{st} values (before the bottle-neck) in Fig. 4 and 5 differ among the different K values. This is because G_{st} depends on population sizes (as a consequence of the dependency of K_c^* on population size; eqn 4).

Discussion

Prout (1981) introduced a 'Pollen-Seed' island model and derived a formula for the variance of nuclear gene frequencies at equilibrium. The work of Asmussen & Schnabel (1991) also included a migration model for plants in their study of the dynamic of cytonuclear structure. Here, we have defined an effective migration rate for both nuclear and cytoplasmic genes, which incorporates variable transmission rate of cytoplasmic genomes and accounts for the two-step migration of genes originating from male gametes in plants. We devised tables for the computation of recurrence equations for gene diversity, as functions of migration and the outcrossing rate, for several hierarchical levels of subdivision. These tables should provide the basis for computation of recurrence equations in even more complex situations (e.g. with dioecy or gyno-dioecy). Heteroplasmy could also be easily included, as did Birky *et al.* (1989).
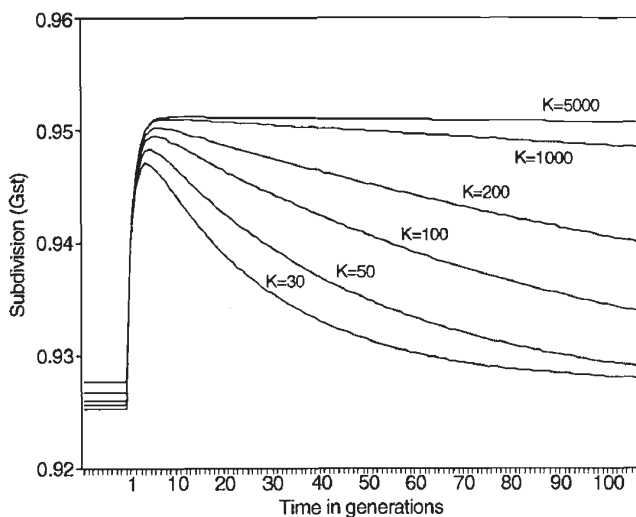

Fig. 5 Effect of population sizes on the rate of approach of cytoplasmic genetic subdivision to equilibrium in populations of hermaphrodite plants after a bottle-neck. Cytoplasmic genes are maternally inherited. Before the bottle-neck, equilibrium conditions were used. For each value of population size N at carrying capacity, the migration rates were determined assuming that $Nm_e = 11.3$ and $m_p/m_s = 557$. $t = 1$, initial conditions: $K_c = 0.0477$, $K_d = 0.65$ ($G_{st} \cong 0.92$).

To derive diversity equations, a clear definition of diversity itself is a prerequisite. We advocate the use of a gene sampling scheme with replacement, contrary to other authors who explicitly or implicitly consider sampling without replacement. In so doing, we exactly apply the definition of diversity of Nei (1973) to populations of finite size. Such a definition seems most appropriate, for example, to describe diversity in very small populations (as for instance when one considers sampling within a zygote).

We have shown that maternal and paternal cytoplasmic inheritance do not have symmetrical effects on the level of subdivision of cytoplasmic genes, even when seed and pollen migration rates are equal. When seed migration rate is much smaller than pollen migration rate, which is probably the case for most tree species (Ganeshaiah & Uma Shaanker, 1991), the equilibrium level of subdivision of maternally inherited cytoplasmic genes differs drastically from that of nuclear genes. Thus, a high pollen/seed migration ratio could account for the observed contrast between low nuclear and high cytoplasmic subdivision found in oaks (Kremer *et al.*, 1991; Petit *et al.*, 1993).

Furthermore, subdivision at equilibrium appears to be lower for nuclear genes in diploid species (and even more so in polyploid species) than for cytoplasmic genes, regardless of the mode of transmission of the cytoplasmic genes, as a consequence of the higher effective number of nuclear genes. However, we have found that a paternally inherited gene may show temporarily lower subdivision than nuclear genes in cases of *dedifferentiating* populations. This is important because a whole group of forest trees (Conifers) have paternal chloroplast inheritance (some Conifers even have paternal mitochondrial inheritance (Neale *et al.*, 1989)).

In addition, our results suggest that dedifferentiation may be more likely to occur in nature than differentiation as a temporary reduction of population size (i.e. a bottle-neck) induces a rapid increase of both nuclear and cytoplasmic subdivision above their equilibrium values. In the cases of Conifers, a test of equilibrium is possible: if accurate measures of the subdivision of paternally inherited genes and of nuclear genes are available, and if $G_{st_c} < G_{st_n}$, then a non-equilibrium situation could be inferred. Study of the range of covariation of nuclear and cytoplasmic genes at equilibrium further illustrates the possibility to detect non-equilibrium situations (Petit, 1992).

For plants with a high pollen/seed migration ratio, both paternal leakage of cytoplasmic genes (in the case of predominantly maternal cytoplasmic inheritance) and outcrossing (for predominant selfers) have a tremendous impact on cytoplasmic subdivisions.

Although this may appear intuitive, our model allows a quantitative assessment of the effect on subdivision of the transmission and outcrossing rates.

Finally, our results clearly show that the effect of a bottle-neck is to strengthen nuclear gene subdivision, regardless of the asymptotic values of population sizes (K), despite high levels of gene flow. In our oak example, for G_{st_n} to decrease from 15 to 20 per cent back to its low initial value (2 per cent) within less than 100 generations or so required relatively low population sizes (less than 200) at the demographic equilibrium, otherwise G_{st_n} would remain apparently 'frozen' at values higher than those typically observed. Another indication that these 'idealized' oak populations (Crawford, 1982) must be quite small is the existence of a heterozygote deficit (Bacilieri *et al.*, 1993) which indicates that the samples of trees studied are not panmictic. Bottle-necks could constitute a major cause of continuing departure from equilibrium for cytoplasmic genes, unless population sizes were even lower (less than 50).

On the other hand, situations have been described where the net effect of extinction and recolonization was an increase in gene flow (i.e. a decrease in differentiation) (Slatkin, 1977; McCauley, 1991). This requires that the founding groups are formed of individuals belonging to different populations and that the newly formed population grows immediately to its carrying capacity. This last hypothesis is not realistic in our example, though, as indicated by the pollen data of Bennett (1983). Moreover, the composition of the founding groups does not matter much in a situation where there is little differentiation at the outset.

To model any real situation, knowledge of the details of the population structure is necessary. The island model of population structure presents an extreme in terms of long-distance gene flow. Another extreme would be local gene flow among neighbouring populations, which would be better represented by a stepping-stone model (Malécot, 1948; Slatkin & Barton, 1989) in the case of continuously distributed species. This model of population structure should be considered in future studies of diversity and subdivision in the three major plant genomes.

References

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.

BACILIERI, R., ROUSSEL, G. AND DUCOUSSO, A. 1993. Hybridization and mating system in a mixed stand of sessile and pedunculate oak. *Ann. Sci. For.*, **50**, (Suppl. 1), 122s–128s.

BENNETT, K. D. 1983. Postglacial population expansion of forest trees in Norfolk, UK. *Nature*, **303**, 164–167.

BIRKY, (Jr.), 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.

CHESNOY, L. 1987. L'origine des organites du cytoplasme embryonnaire chez les Gymnospermes. *Bull. Soc. Bot. Fr., Actual. Bot.*, **134**, 51–56.

CRAWFORD, T. J. 1982. What is a population? In: Shorrocks, B. (ed.), *Evolutionary Ecology*, Blackwell Scientific Publications, Oxford. pp. 135–173.

CROW, J. F. AND AOKI, K. 1984. Group selection for a polygenic behavioral trait: estimating the degree of population subdivision. *Proc. Natl. Acad. Sci. USA*, **81**, 6073–6077.

DEMARLY, Y. 1963. Génétique des tétraploïdes et amélioration des plantes. *Ann. Amélior. Plantes*, **13**, 307–400.

DESALLE, R., TEMPLETON, A., MORI, I., PLETSCHER, S. AND JOHNSTON, J. S. 1987. Temporal and spatial heterogeneity of mtDNA polymorphisms in natural populations of *Drosophila mercatorum*. *Genetics*, **116**, 215–223.

GANESHAIAH, K. N. AND UMA SHAANKER, R. U. 1991. Floral sex ratios in monoecious species – Why are trees more male-biased than herbs? *Current Sci.*, **60**, 319–321.

HALE, L. R. AND SINGH, R. S. 1987. Mitochondrial DNA variation and genetic structure in populations of *Drosophila melanogaster*. *Mol. Biol. Evol.*, **4**, 622–637.

KREMER, A., PETIT, R. J., ZANETTO, A., FOUGERE, V., DUCOUSSO, A., WAGNER, D. B. AND CHAUVIN, C. 1991. Nuclear and organelle gene diversity in *Quercus robur* and *Q. petraea*. In: G. Müller-Starck and M. Ziehe (eds), *Genetic Variation in European Populations of Forest Trees*, Sauerländer's Verlag, Frankfurt am Main, pp. 141–166.

McCAULEY, D. E. 1991. Genetic consequences of local population extinction and recolonization. *Trends Ecol. Evol.*, **6**, 5–8.

MALECOT, G. 1948. *Les Mathématiques de l'Hérédité*. Masson, Paris.

MARUYAMA, T. 1970. Effective number of alleles in a subdivided population. *Theor. Popul. Biol.*, **1**, 273–306.

MÜLLER-STARCK, G. AND ZIEHE, M. 1991. Genetic variation in populations of *Fagus sylvatica* L., *Quercus robur* L., and *Q. petraea* Liebl. in Germany. In: G. Müller-Starck and M. Ziehe (eds), *Genetic Variation in European Populations of Forest Trees*, Sauerländer's Verlag, Frankfurt am Main, pp. 125–140.

NEALE, D. B., MARSHALL, K. A. AND SEDEROFF, R. R. 1989. Chloroplast and mitochondrial DNA are paternally inherited in *Sequoia sempervirens* D. Don Endl. *Proc. Natl. Acad. Sci. USA*, **86**, 9347–9349.

NEI, M. 1973. Analysis of gene diversity in subdivided populations. *Proc. Natl. Acad. Sci. USA*, **70**, 3321–3323.

NEI, M. 1975. *Molecular Population Genetics and Evolution*. North-Holland, New York.

PETIT, R. J. 1992. *Polymorphisme de l'ADN chloroplastique dans un complexe d'espèces: les chênes blancs européens*. Ph.D Thesis, University of Paris XI, Orsay.

PETIT, R. J. KREMER, A. AND WAGNER, D. B. 1993. Geographic structure of chloroplast DNA polymorphisms in European oaks. *Theor. Appl. Genet.*, **87**, (in press).

PROUT, T. 1981. A note on the island model with sex dependent migration. *Theor. Appl. Genet.*, **59**, 327–332.

SLATKIN, M. 1977. Gene flow and genetic drift in a species subject to frequent local extinction. *Theor. Popul. Biol.*, **12**, 253–262.

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.

TAKAHATA, N. AND NEI, M. 1984. F_{st} and G_{st} statistics in the finite island model. *Genetics*, **107**, 501–504.

TAKAHATA, N. AND PALUMBI, S. R. 1985. Extranuclear differentiation and gene flow in the finite island model. *Genetics*, **109**, 441–457.

WRIGHT, S. 1943. Isolation by distance. *Genetics*, **28**, 114–138.

ZANETTO, A., KREMER, A. AND LABBE, T. 1993. Differences of genetic variation based on isozymes of the primary and secondary metabolism in *Quercus petraea*. *Ann. Sci. For.*, **50** (Suppl. 1), 245s–253s.