

# Effects of Colonization Processes on Genetic Diversity: Differences Between Annual Plants and Tree Species

Frédéric Austerlitz,\* Stéphanie Mariette,<sup>†</sup> Nathalie Machon,<sup>‡</sup> Pierre-Henri Gouyon\*  
and Bernard Godelle\*<sup>§</sup>

\*Laboratoire Evolution et Systématique, Université Paris-Sud, F-91405 Orsay, France, <sup>†</sup>Laboratoire de Génétique et d'Amélioration des Arbres Forestiers, INRA, F-33611 Gazinet, France, <sup>‡</sup>Conservatoire Botanique du Bassin Parisien, Muséum National d'Histoire Naturelle, F-75005 Paris, France and <sup>§</sup>Institut National Agronomique Paris-Grignon, F-75005 Paris, France

Manuscript received May 15, 1999  
Accepted for publication November 11, 1999

## ABSTRACT

Tree species are striking for their high within-population diversity and low among-population differentiation for nuclear genes. In contrast, annual plants show much more differentiation for nuclear genes but much less diversity than trees. The usual explanation for this difference is that pollen flow, and therefore gene flow, is much higher for trees. This explanation is problematic because it relies on equilibrium hypotheses. Because trees have very recently recolonized temperate areas, they have experienced many foundation events, which usually reduce within-population diversity and increase differentiation. Only extremely high levels of gene flow could counterbalance these successive founder effects. We develop a model to study the impact of life cycle of forest trees, in particular of the length of their juvenile phase, on genetic diversity and differentiation during the glacial period and the following colonization period. We show that both a reasonably high level of pollen flow and the life-cycle characteristics of trees are needed to explain the observed structure of genetic diversity. We also show that gene flow and life cycle both have an impact on maternally inherited cytoplasmic genes, which are characterized both in trees and annual species by much less diversity and much more differentiation than nuclear genes.

THE present distribution of genetic variability in forest tree species is peculiar. Isozyme data, collected by Hamrick *et al.* (1992) and Hamrick and Godt (1996) on a large number of species, indicate that trees maintain a significantly higher level of genetic diversity within species (0.177 on average) and within populations (0.148 on average) than annual plants (respectively 0.154 and 0.101 on average) for nuclear genes. Forest trees also show a lower level of genetic differentiation among populations, measured with  $G_{ST}$  (Nei 1973), a coefficient of gene differentiation equivalent to Wright's (1951)  $F_{ST}$ : 0.084 on average for woody long-lived perennial species and 0.355 on average for annual plants. The low differentiation observed with isozymes has recently been confirmed by molecular data for several species (for instance, the oak *Quercus petraea*; Le Corre *et al.* 1997a).

Analysis with cytoplasmic markers of chloroplast DNA, which is maternally inherited in most angiosperms, gives completely different results: a clear geographic structure is observed. For two European species of oaks, the estimations of  $G_{ST}$  are 0.905 for *Q. robur* and 0.925 for *Q. petraea* (Petit *et al.* 1993a). Comparable values are given for various tree species in Ennos (1994). Genetic

structure is clearly much higher for maternally inherited cytoplasmic markers than for nuclear ones. Data on the structure of cytoplasmic genetic diversity in annual or short-lived plants are very rare. For *Silene alba*, a short-lived perennial,  $F_{ST}$  is 0.613 for chloroplast genes and 0.128 for nuclear genes on a regional scale in a recently colonized region (McCaughey 1994, 1997). The values are, respectively, 0.624 and 0.222 for *S. vulgaris* (McCaughey 1998). There are also few data for paternally inherited chloroplast genes in gymnosperms. They usually show much less differentiation than maternally inherited mitochondrial genes (Dong and Wagner 1994; Latta and Mitton 1997).

The high within-population diversity and low differentiation of trees' nuclear genes is unexpected. During the last glacial period of ~100,000 years (Andersen and Borns 1994), tree species from the Northern Hemisphere were confined to a few southern refuges completely isolated from one another. About 15,000 years ago, they began to recolonize all temperate areas and some species have only recently reached their modern range limits (Huntley and Birks 1983; Huntley 1990). As pointed out by Kremer (1994), because of trees' long life span, there have been only 100–1000 generations (depending on the species) since the end of the last glacial period. Our previous results (Austerlitz *et al.* 1997) showed that the successive foundation events that occur during colonization yield a strong genetic differentiation and low within-population diver-

Corresponding author: Frédéric Austerlitz, Rm. 152, ENRS Bldg., 14 College Farm Rd., Rutgers University, New Brunswick, NJ 08901-8551. E-mail: austerlitz@aesop.rutgers.edu

sity, especially in populations far from refuges. This is the classically denoted founder effect: the loss of diversity in newly founded populations due to a small number of founders.

Various authors have experimentally investigated the impact of recent colonization events on genetic diversity, both on annual plants and tree species, on a more local scale. They show that colonization events usually yield strong founder effects in plant species (Westerbergh and Saura 1994; Raybould *et al.* 1996), but it is not the case for several tree species (Mariette *et al.* 1997; Raspé and Jacquemart 1998).

The usual explanation for the limited differentiation of nuclear genes in trees and for their high differentiation for maternally inherited cytoplasmic genes is that there is much more migration through pollen than through seeds. Ennos (1994) showed that the difference between nuclear and cytoplasmic gene differentiation in oaks could be explained by a 200:1 ratio between gene flow through pollen and seeds, respectively. Nevertheless, his results are calculated for an island model at equilibrium. Equilibrium hypotheses seem unrealistic because recolonization is recent for forest tree species, and an island model on the scale of most temperate areas is doubtful. To be more realistic, Le Corre (1997) simulated the process of European recolonization with a two-dimensional (2D) stepping-stone model with non-overlapping generations. To explain the observed  $F_{ST}$  values, she had to take an even higher ratio of pollen flow *vs.* seed flow.

None of these studies take into account an important feature of tree species. Unlike short-lived annual and perennial plant species, trees have a long life span and grow vegetatively for a long period before reproduction. Birches, alders, elms, and poplars reproduce at  $\sim 10$  years; maples and chestnuts between 20 and 30 years; and oaks and beeches at  $\sim 50$  years. Great longevity, several age or size classes, overlapping generations, and a long juvenile phase are the main characteristics of the life cycle of trees.

The first aim of this article is to show the impact of this life-cycle difference between trees and short-lived plants on genetic diversity and population structure during colonization. Our hypothesis is that due to the duration of the juvenile phase, a newly founded tree population will grow for many years only through the arrival of new migrants, increasing the number of founders of a population and therefore decreasing the founder effect. The second aim is to study the impact of the persistence of several refuges, isolated from one another during the glacial period.

We develop a Markovian approach that follows the evolution of genetic structure for nuclear and maternally inherited cytoplasmic genes of plants with a tree life cycle during and after a colonization period, following long isolation. Our colonization starts from one or several refuges and is one- or two-dimensional. In

previous models of tree populations, only long generation time was taken into account, trees being thus modeled as semelparous long-lived plants with no overlap between generations (bamboo-like models; Petit *et al.* 1993b; Ennos 1994; Le Corre *et al.* 1997b). Here trees are also characterized by a juvenile phase and overlapping generations. They are compared with plants having no juvenile phase or nonoverlapping generations, with either a 1-year generation (annual plants) or the same generation time as trees. In the latter case, all plants in a cohort reproduce and die simultaneously. This corresponds to "trees" as modeled in previous studies, but not to any existing plant, and it allows us to characterize the effect of life cycle independently of generation length. To also separate the effects of life cycle from other effects, all the comparisons are made with the same effective size, the same amount of gene flow through seeds and pollen, and the same population growth rate.

## MATERIALS AND METHODS

**Demographic model within each population:** We simulate a tree life cycle of several size classes and include a density-dependence factor. In classical models of population dynamics, the transition from one population state to another is based on a matrix that gives annual survival and reproduction rates. These models were originally used by Lewis (1942) and Leslie (1945) for human demography and dynamics of animal populations that were age structured.

They were adapted to populations of forest trees, either natural (for example, Bosch 1971) or managed (Usher 1966, 1969). Following these models, we structured populations according to the size of the individuals rather than the age, as in Lefkovich (1965). If  $\mathbf{A}$  is the annual transition matrix and  $\mathbf{N}(t)$  and  $\mathbf{N}(t+1)$  are the vectors of the numbers of trees in each size class, respectively, at years  $t$  and  $t+1$ , we write  $\mathbf{N}(t+1) = \mathbf{A} \cdot \mathbf{N}(t)$ .

$$\mathbf{A} = \begin{pmatrix} P_{11} & 0 & 0 & \dots & f \\ P_{12} & P_{22} & 0 & \dots & 0 \\ 0 & P_{23} & P_{33} & & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & P_{k-1,k} & P_{kk} \end{pmatrix},$$

where  $P_{ii}$  is the proportion of trees that stay alive and remain in class  $i$  from  $t$  to  $t+1$ ;  $P_{i+1,i}$  is the proportion of trees of class  $i$  that stay alive and move from the class of size  $i$  to  $i+1$  during the time interval  $t$  to  $t+1$ ; and  $f$  is the number of offspring per individual of the last class at each generation. Thus, we have  $k-1$  classes of juveniles and only one class of adult trees that can reproduce. The proportion of individuals that die in class  $i$  in one time unit is  $1 - P_{ii} - P_{i+1,i}$ .

However, models based on the Leslie matrix yield exponential growth or decrease of the size of each class, which is not realistic enough if long-term predictions are required. Therefore we include a density-dependent growth regulation function, adapted from Buongiorno *et al.* (1995). This regulation acts as follows: when density increases, the probabilities of moving to the next class and fecundity decrease, whereas mortality increases. This is performed by assigning to the individuals of each class a given stand basal area (which is a mea-

sure of the space occupied by an individual). Then, at each time  $t$ , all  $P_{i+1}$  and  $f$  are multiplied by a reduction coefficient  $\alpha(t)$  ( $0 \leq \alpha(t) \leq 1$ ), which is computed so that the total stand basal area of the whole population grows logistically to an equilibrium value (see details in appendix a).

The demographic data available (see, for instance, Buongiorno and Michie 1980; Harcombe 1987) subdivide tree populations into a small number of classes and are very specific for a given species in a given area. Because we were interested in features that are common to all forest tree species, we chose parameters that were representative of the main characteristics of forest trees. The number of classes ( $k$ ) is 25, so trees can never reproduce before the age of 25, and fecundity is high in the absence of density dependence ( $f = 250$ ). The other parameters of the projection matrix  $\mathbf{A}$  are given in appendix a, and they were chosen so that at equilibrium an individual reaches the adult class on average at 50 years and the average age of the individuals in the adult class is 100 years. An important point of this model is that, for a newly founded population, due to low density, trees reaching the adult class are much younger. With the chosen parameter, tree populations are filled to carrying capacity within a few hundred years (*i.e.*, a few generations).

The results obtained for trees are compared with those obtained for plants with no juvenile phase and nonoverlapping generations. For these plants, the population size at equilibrium is set at the effective size of tree populations at equilibrium ( $N_e$ ), which is calculated using the method of Orive (1993). These plants are modeled with a single class ( $k = 1$ ). All individuals are born, reproduce, and die within a single generation (that is,  $P_{11} = 0$ ). The population grows logistically toward its equilibrium size. The initial growth rate  $r$  is directly related to the fecundity  $f$ :  $f = 1 + r$ . The generation time can be either 1 year (annual plants) or 100 years, which is approximately the generation time of trees in our model. Various values of the fecundity were tested:  $f = 1.01, 1.1, 1.5$ , or 3. It is difficult to compare precisely the growth of trees and plants with nonoverlapping generations. The value of 1.1 is the fecundity necessary for the population of annual plants to grow approximately at the same rate as trees, as defined in our model. The value of 3 is the corresponding value for the plants with nonoverlapping generations of 100 years. These values were determined by empirical comparison of the growth of stand basal area in both cases.

**Colonization models:** In each case, we have a total of  $d$  populations connected by seed and pollen flow in a one- or two-dimensional stepping-stone model (see Figure 1).

We wanted to first study the interaction between the colonization process and tree life cycle from a theoretical point of view. We thus modeled colonization for a one-dimensional stepping-stone metapopulation (in which seeds and pollen migrate unidirectionally to the neighboring populations; see Figure 1A). At the beginning of a simulation, all sites are empty, except for the refuge, within which there is a population at demographic and genetic equilibrium (this is obtained by letting this population evolve alone until it reaches equilibrium). The total number of sites ( $d$ ) is 15. Colonization is then one-dimensional and unidirectional. Every 100 years a new site is opened to colonization, as a consequence of progressive warming, making the total colonization period last 1400 years. Subsequently, we let the populations evolve after the colonization period until time  $t = 5000$  years.

The second model of colonization is a two-dimensional stepping-stone model (see Figure 1B). Because the process is highly computer intensive, the total number of populations was set at only 102 (6 along the  $x$ -axis times 17 along the  $y$ -axis). We started with three refuge areas. In each refuge there were 4 populations. Several initial conditions for genetic diver-

sity and differentiation were tested; in most cases we started with the expected equilibrium value for the 102 populations, connected by gene flow, as indicated. Before starting the colonization, we let the 12 populations evolve during 80,000 years (the average length of glacial periods, see Hays *et al.* 1976; Pailard 1997), with no gene flow through seeds or pollen among them. After the isolation period, the colonization starts, with 1 new area of populations being open for colonization every 500 years. This new area consists of 6 populations on a line (see Figure 1B). This yields a colonization period of 7000 years. Then the populations evolve, remaining at their equilibrium size, for a postcolonization period of 8000 years. The entire process lasts 95,000 years. The length of the different periods is realistic (Huntley and Birks 1983; Kremer 1994), but due to the limitations in the number of populations, the interval between two colonization events had to be long.

In each case the migration parameters are  $m_s$  and  $m_p$ , the proportion of seeds and pollen produced by the adults of a population  $p$  that exit the population. Migrants arrive with equal probability to each of the populations to which  $p$  is connected. Regulation by density dependence acts on migrant seeds in the population in which they arrive in the same way that it acts on local seeds.

**Computation of genetic parameters:** The mutation rate is set at  $10^{-6}$  for nuclear and cytoplasmic genes. We use a Markovian approach to calculate the probabilities of identity by descent (IBD) iteratively from one generation to the next, assuming an infinite allele model. Time is expressed in units of years, except for the plants with long nonoverlapping generations, where it is expressed in units of generations. Individuals are diploid for their nuclear genes and haploid for their cytoplasmic ones, which are strictly maternally inherited. We assume a panmictic reproduction among all the individuals of the adult class. The probabilities  $f_{p,i,q,j}^c$  and  $f_{p,i,q,j}^n$  of identity by descent at time  $t$  for two cytoplasmic or nuclear genes belonging, respectively, to individuals of population  $p$  and class  $i$  and of population  $q$  and class  $j$ , are calculated iteratively every unit of time, after migration and reproduction occur, for all populations  $p$  and  $q$  that are not empty at time  $t$  and all classes  $i$  and  $j$  between 1 and  $k$  (see details in appendix b). The following computations are the same for nuclear or cytoplasmic genes, and  $f_{p,i,q,j}$  denotes both  $f_{p,i,q,j}^c$  or  $f_{p,i,q,j}^n$  in this subsection.

The expected diversity  $H_p$  in each population  $p$  is then calculated as follows:  $H_p = 1 - f_{p,p}$  where  $f_{p,p}$  is the average probability of identity by descent within population  $p$ ,

$$f_{p,p} = \frac{2}{N_p(N_p - 1)} \left( \sum_{1 \leq i \leq k} \frac{N_{p,i}(N_{p,i} - 1)}{2} f_{p,i,p,i} + \sum_{1 \leq i \neq j \leq k} N_{p,i} N_{p,j} f_{p,i,p,j} \right),$$

where  $N_{p,i}$  is the number of individuals in each population  $p$  and each class  $i$ , and  $N_p$  is the total number of individuals in population  $p$ . This allows us to calculate the average within-population diversity  $H_S$  by averaging  $H_p$  over every population in  $\Omega$ , the set of populations that are not empty at time  $t$ . In the same way, we can calculate the average probability of identity by descent  $f_{p,q}$  for two genes chosen respectively in populations  $p$  and  $q$ ,  $p \neq q$ ,

$$f_{p,q} = \frac{1}{N_p N_q} \left( \sum_{1 \leq i \neq j \leq k} N_{p,i} N_{q,j} f_{p,i,q,j} \right),$$

The average identity by descent  $f_T$  for two genes chosen in any population is then calculated

$$f_T = \frac{1}{d_\Omega} \sum_{p,q \in \Omega} f_{p,q}$$

where  $d_\Omega$  denotes the number of populations in  $\Omega$ . The ex-

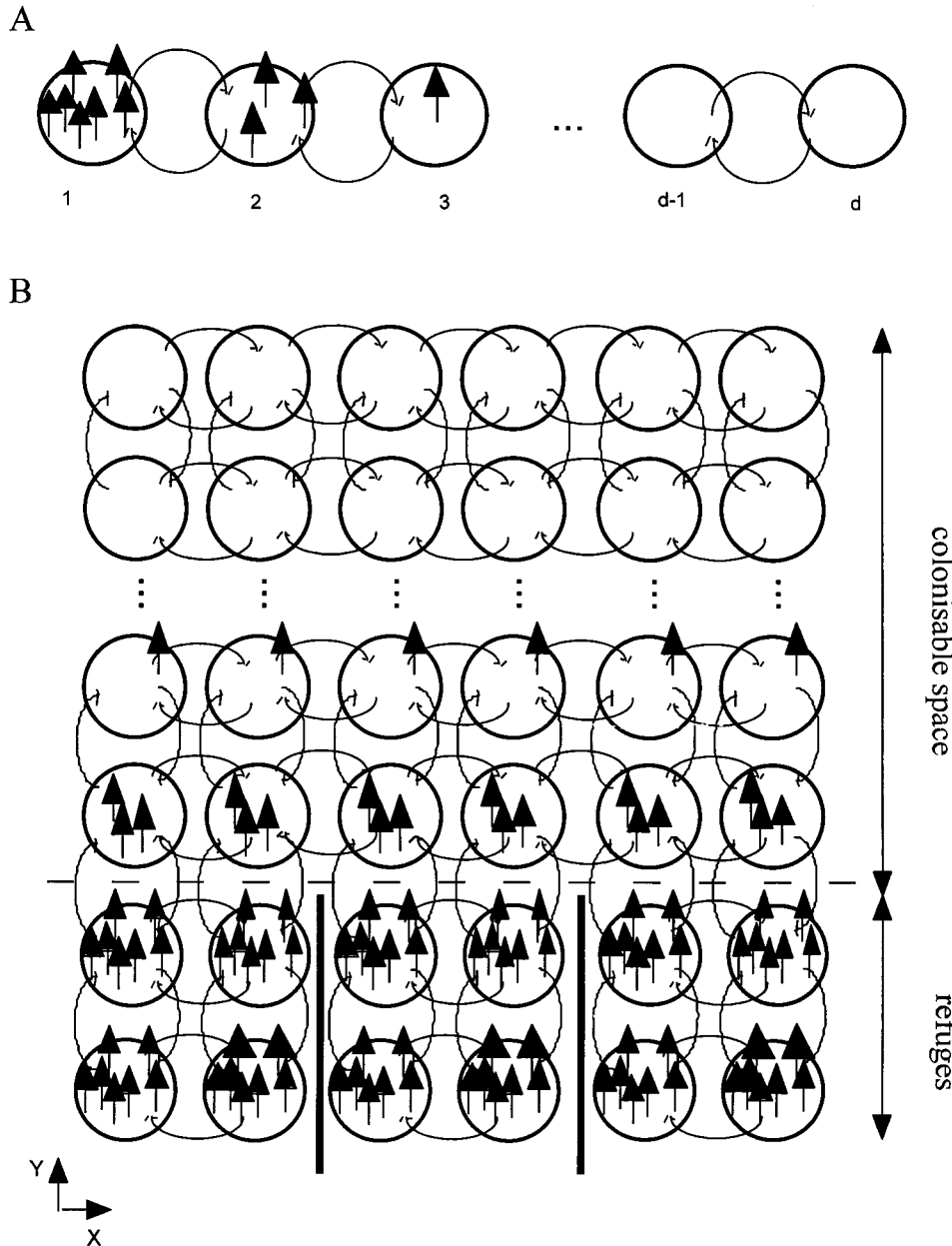


Figure 1.—Schematic representation of the 1D and 2D colonization processes. In both cases, the arrows indicate the populations that are connected by gene flow through seeds and pollen. In the 1D colonization (A), only the first population is full and at demographic and genetic equilibrium; all other populations are empty. The simulation begins with the colonization of the second population, and one population is colonized subsequently every 100 years. In the 2D colonization (B), starting from a given value of differentiation, the three groups of four refuge populations drift independently, one from another, over 80,000 years (isolation period). Then the whole space is colonized in the  $y$  direction, one line of populations every 500 years (colonization period, which lasts 7000 years). Then, we let the system evolve until time  $t = 95,000$  years is reached (postcolonization period).

pected global diversity is  $H_T = 1 - f_T$  and the expected  $F_{ST}$  can be deduced using  $F_{ST} = 1 - H_S/H_T$ .

## RESULTS

**Evolution of within-population diversity in the one-dimensional colonization process:** For annual plants (Figure 2A), there was very strong founder effect: diversity in each population decreased substantially during the first generations after each colonization event. A great proportion of diversity was lost due to founder effect but was then recovered relatively quickly. Figure 2A gives the case of no pollen flow. For the case of high pollen flow, the founder effect was almost of the same magnitude, but the recovery of diversity was much faster (data not shown).

For trees, the situation was very different. Even with

no pollen flow, the foundation event had a very limited effect: genetic diversity decreases only slightly in the newly founded populations (Figure 2B). Afterward, it increased only very slowly. Nevertheless, trees reached the same equilibrium value as annual plants after several hundred thousand years (data not shown). There was almost no founder effect and each founded population retained almost all the diversity of its ancestral population. Increasing pollen flows reduced the founder effect even further (data not shown).

**Evolution of among-population differentiation ( $F_{ST}$ ) in the one-dimensional colonization process:** For annual plants,  $F_{ST}$  strongly increased during the colonization period (Figure 3, A–C), followed by a gradual decrease and again a slow increase toward its equilibrium value. With  $N_e = 1000$  and  $m_s = 0.0002$ , even in the case of high pollen flow ( $m_p = 0.01$ ),  $F_{ST}$  for nuclear genes

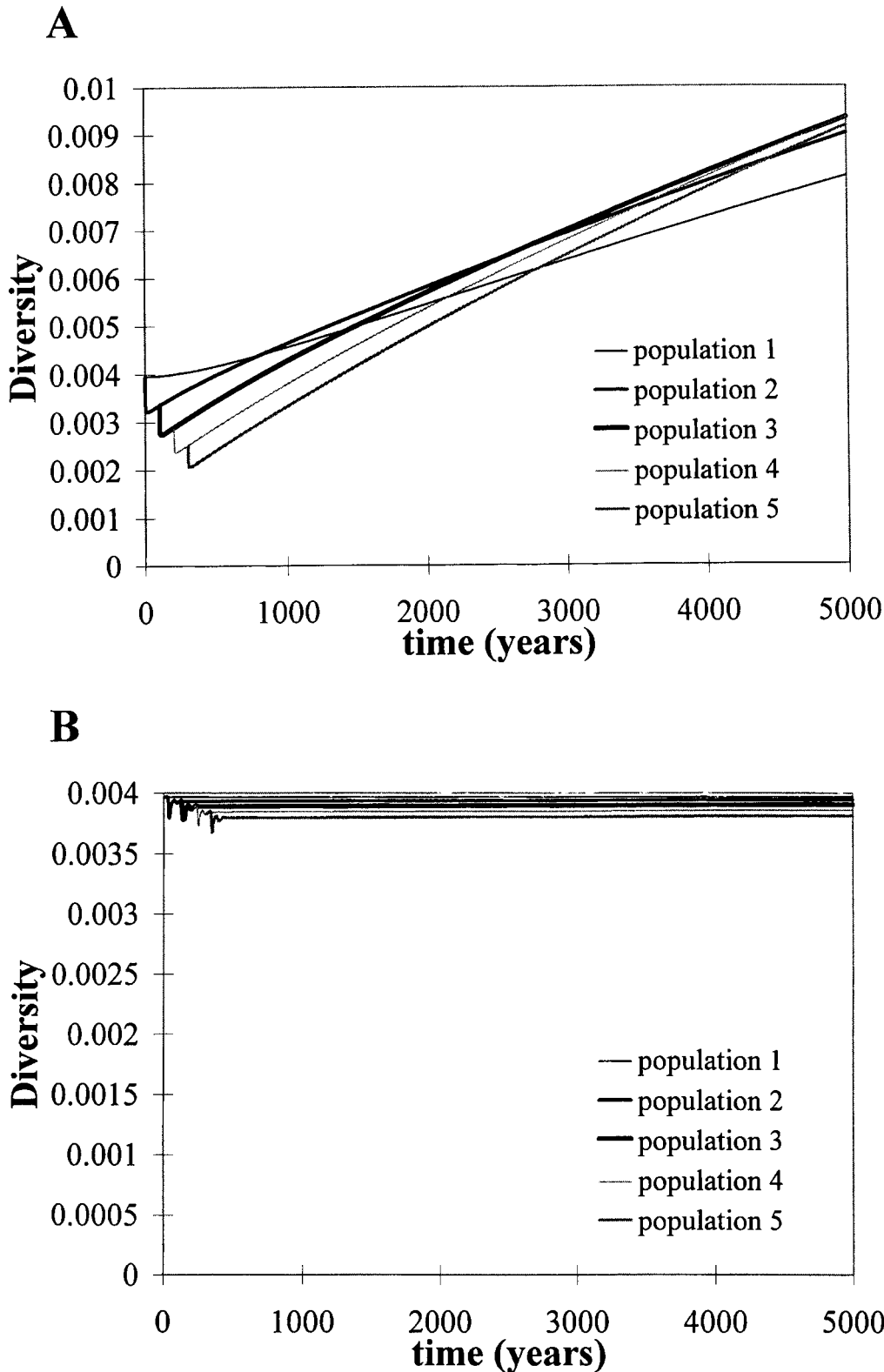


Figure 2.—Within-population diversity of annual plants (A) and trees (B) for the first five populations of the 1D colonization process. The founder effect is much greater for annual plants than for trees. Parameters are as follows:  $m_s = 0.0002$ ,  $m_p = 0$ ,  $N_e = 1000$ ,  $d = 15$ .

reached a maximum value of 0.13 (Figure 3A). This maximum value was higher when pollen flow decreased: 0.28 when  $m_p = 0$  (Figure 3B). The value was 0.36 for cytoplasmic genes (Figure 3C).

For trees,  $F_{ST}$  for nuclear genes increased only very slowly during the colonization period. It then increased

even slower toward its equilibrium value, which was the same for trees and annual plants. Even with no pollen flow, the  $F_{ST}$  value at the end of the colonization period was only 0.036, much lower than the final equilibrium value. For cytoplasmic genes, the increase in differentiation was slightly larger in the colonization period, yield-

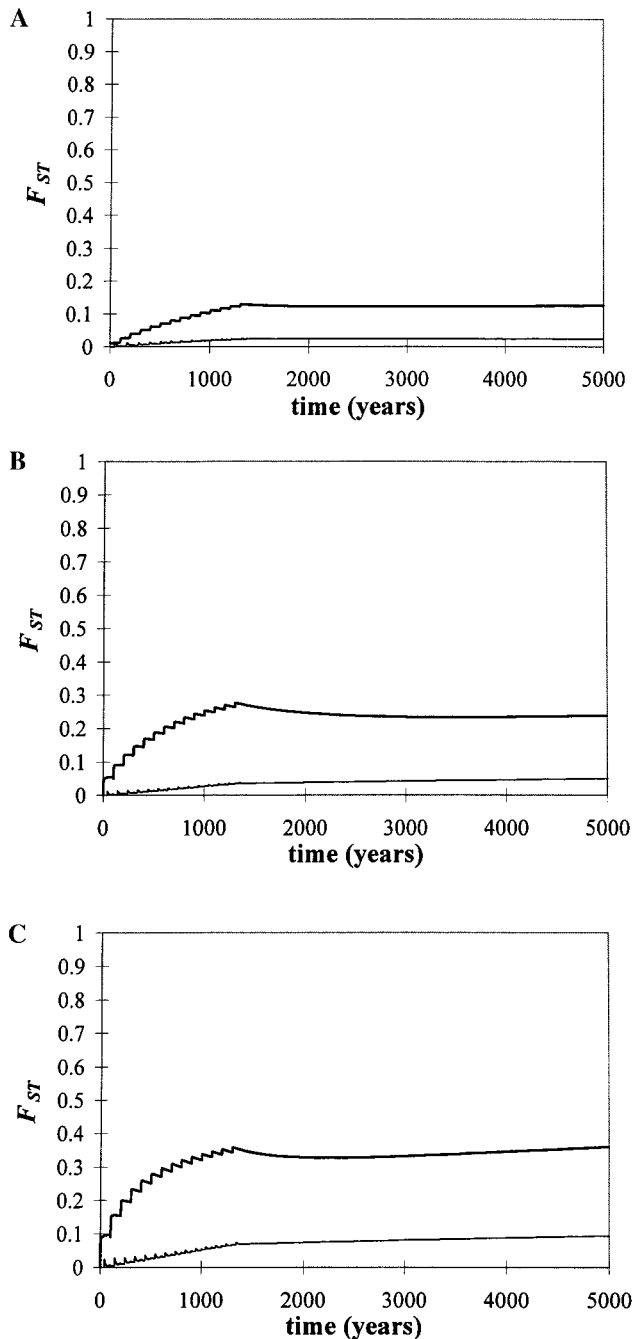


Figure 3.— $F_{ST}$  plotted against time since the beginning of the colonization period in the 1D colonization process, for trees (thin lines) and annual plants (thick lines), for nuclear genes (A and B), or maternally inherited cytoplasmic genes (C). In all cases,  $d = 15$ ,  $m_s = 0.0002$ ,  $N_e = 1000$ . For nuclear genes, the pollen flow  $m_p$  was, respectively, 0.01 (A) and 0.00 (B).

ing an  $F_{ST}$  of 0.07 at the end of the period. Again, this value was much lower than the equilibrium value. In all cases the rate of increase of  $F_{ST}$  decreased during colonization. The founder effect was decreased for increasing fecundity of annual plants, but it remained higher than the one to which trees were subjected (result not shown).

**Evolution of among-population differentiation ( $F_{ST}$ ) in the two-dimensional colonization process:** Several features appeared when two-dimensional processes were taken into account (Figure 4).

*Nuclear genes:* First,  $F_{ST}$  increased during the isolation period of 80,000 years. The increase was much stronger for annual plants in all cases. For example, in Figure 4A, with  $N_e = 1000$ ,  $m_s = 0.0005$ , and  $m_p = 0.1$ , starting from equilibrium value,  $F_{ST}$  of nuclear genes only increased from 0.021 to 0.093 for trees during the isolation period, whereas for annual plants it increased from 0.021 to 0.92, having almost reached equilibrium.

For trees, the rate of increase during this isolation period depended slightly on pollen flow: if we started from 0.021, it increased up to 0.093 for  $m_p = 0.1$  during the isolation period (Figure 4A) and up to 0.12 for  $m_p = 0.01$  (Figure 4B). The rate of increase of  $F_{ST}$  during the isolation period depended much more on population effective size ( $N_e$ ), when the comparisons were made with the same number of migrants (*i.e.*, the same values of  $N_e m_s$  and  $N_e m_p$ ). For the nuclear genes of trees, with  $N_e m_s = 0.5$  and  $N_e m_p = 10$ , starting from the initial equilibrium value of  $F_{ST}$  of 0.17,  $F_{ST}$  increased only to 0.23 in 80,000 years when  $N_e = 1000$  (Figure 4C), whereas when  $N_e = 100$  it increased much more, to 0.63 (Figure 4D).

Second,  $F_{ST}$  for annual plants decreased during the colonization period that followed. For trees, it decreased with high pollen flow (Figure 4A) or low effective population size (Figure 4D); otherwise it increased slightly (Figure 4, B and C). Third, during the postcolonization period, there was a decrease of  $F_{ST}$  that was very sharp in the case of annual plants but smooth for trees. As a consequence, the  $F_{ST}$  for annual plants fell quickly below that of trees. The rate of decrease also depended on the effective population size, a smaller  $N_e$  making  $F_{ST}$  decrease much faster. Therefore,  $F_{ST}$  of annual plants declined almost immediately below that of trees in the case of  $N_e = 100$  (Figure 4D).

*Cytoplasmic genes:* Concerning cytoplasmic genes (Figure 4, E and F), with an  $N_e m_s$  of 0.5, the equilibrium value of  $F_{ST}$  was already very high,  $>0.8$ . If we started at  $t = 0$  from this value (Figure 4E), the isolation period brought almost no change for trees, whereas it yielded an increase for annual plants. Then, as for nuclear genes,  $F_{ST}$  for annual plants fell rather quickly below that of trees.

However, if we started with an  $F_{ST}$  of 0.17 (Figure 4F), as for nuclear genes (Figure 4C), even with an  $N_e$  value of 1000,  $F_{ST}$  of trees for cytoplasmic genes increased much more than for nuclear genes during the isolation period. It increased even more sharply during the colonization period and then slowly approached the equilibrium.

**Comparison with long-living plants with nonoverlapping generations in the two-dimensional colonization process:** For a plant with nonoverlapping generations and a generation time similar to that of trees, the isola-

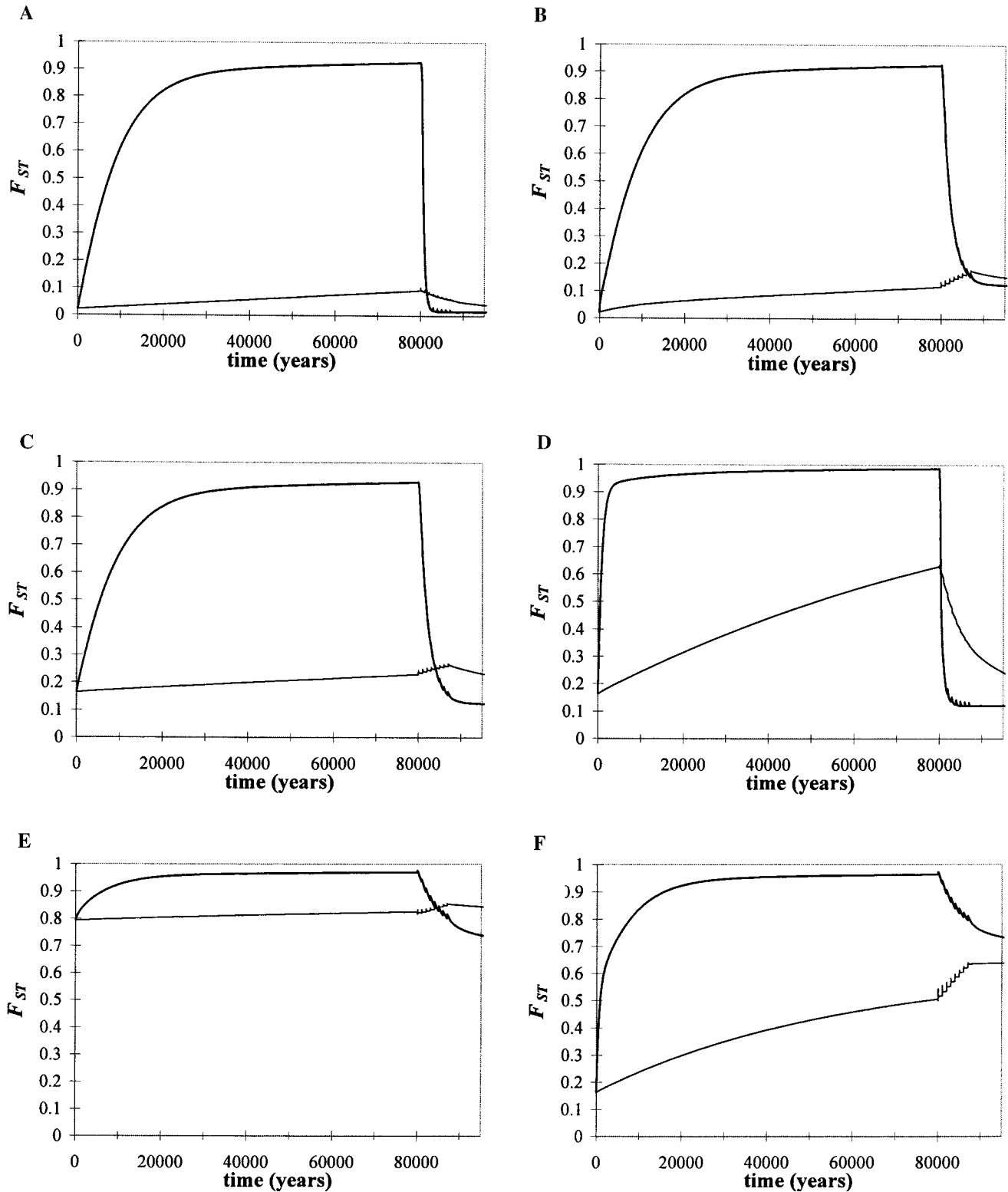


Figure 4.— $F_{ST}$  plotted against time for the 2D colonization process, with  $6 \times 17$  populations (see details in materials and methods), for trees (thin lines), and annual plants (thick lines). A, B, C, and D represent cases of nuclear genes with different values of the initial differentiation  $F_{ST0}$  (which is in some cases the equilibrium value) and the others parameters. (A)  $m_s = 0.0005$ ,  $m_p = 0.1$ ,  $N_e = 1000$ ,  $F_{ST0} = 0.02$  (equilibrium value); (B)  $m_s = 0.0005$ ,  $m_p = 0.01$ ,  $N_e = 1000$ ,  $F_{ST0} = 0.02$ ; (C)  $m_s = 0.0005$ ,  $m_p = 0.01$ ,  $N_e = 1000$ ,  $F_{ST0} = 0.17$  (equilibrium value); (D)  $m_s = 0.005$ ,  $m_p = 0.1$ ,  $N_e = 100$ ,  $F_{ST0} = 0.17$  (equilibrium value). E and F correspond to cytoplasmic genes: (E)  $m_s = 0.0005$ ,  $N_e = 1000$ ,  $F_{ST0} = 0.82$  (equilibrium value); (F)  $m_s = 0.0005$ ,  $N_e = 1000$ ,  $F_{ST0} = 0.17$ .

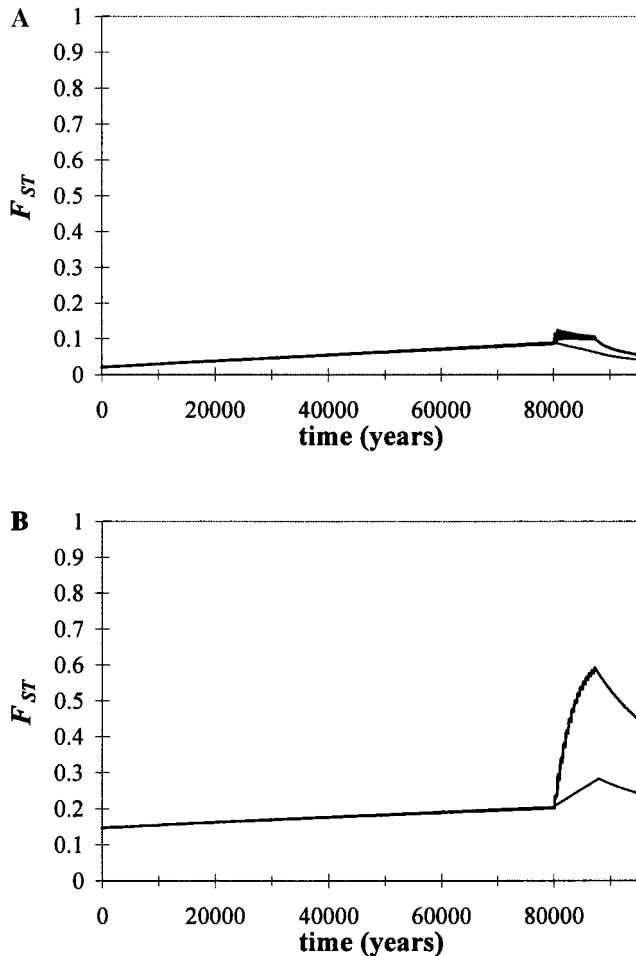


Figure 5.— $F_{ST}$  plotted against time for the 2D colonization process with  $6 \times 17$  populations for trees (thin lines) and plants with nonoverlapping generations and the same generation length as trees (thick lines), with a fecundity  $f$  of 3. The given values correspond to nuclear genes with, for A,  $m_s = 0.0005$ ,  $m_p = 0.1$ ,  $N_e = 1000$ ,  $F_{ST0} = 0.02$  (equilibrium value); and for B,  $m_s = 0.0005$ ,  $m_p = 0.01$ ,  $N_e = 1000$ ,  $F_{ST0} = 0.17$  (equilibrium value).

tion period had the same impact as on forest trees for nuclear (Figure 5) or cytoplasmic genes (data not shown). The colonization period yielded a stronger increase in  $F_{ST}$  of nuclear genes for the plant with nonoverlapping generations than that for the tree species. This difference was very low when pollen flow was very high ( $m_p = 0.1$ , Figure 5A). It was much higher and remained so long after the colonization period was over, when pollen flow had a lower value ( $m_p = 0.01$ , which still makes  $N_{em}m_p = 10$ , Figure 5B). Again, the founder effect decreased when increasing the fecundity of these plants, but it remained higher than that to which trees are subjected (results not shown).

#### DISCUSSION

Our results, which are summarized in Table 1, are useful for understanding the impact of various demo-

graphic parameters that have influenced the present distribution of genetic diversity in forest trees and annual plants. We show that the founder effect can be dramatically reduced when the life cycle of trees is taken into account, even with limited pollen flow. Therefore we did not need to introduce very high levels of pollen flow to explain the contrast between low differentiation of nuclear genes and the high differentiation of maternally inherited cytoplasmic genes.

Thanks to the one-dimensional colonization process, we showed that the founder effect seems to be much more limited for trees than for annual plants. This can be explained as follows: when a tree population is founded, growth for the first several years is because of new juvenile migrants (seed flow) and not because of reproduction. Therefore, when the first trees reach reproductive age in the newly founded population, a non-negligible part of the space is already occupied by juveniles from seeds that arrived years before.

On the contrary, as we had shown in a previous work (Austerlitz *et al.* 1997), the first annual plants that arrive in an empty site can reproduce the next year, and therefore the offspring of these first occupants have the opportunity to colonize the whole space. Therefore, the subsequent loss of diversity is greater for annual plants than for trees, and the differentiation among populations is also greater.

We emphasize that the key factor in avoiding the founder effect of trees species is not overlapping generations but delayed reproduction, which allows a large increase in the number of initial founders of a given population before reproduction begins. To check this more directly we made simulations with plants having overlapping generations but no juvenile phase, with the same growth rate and equilibrium effective population size. For these plants  $F_{ST}$  increased as for annual plants during colonization (results not shown). Therefore, even if some life-history traits, like the occurrence of a seed bank, can have an impact on annual plants similar to that of overlapping generations (Templeton and Levin 1979), they cannot reduce the founder effect for those plants.

This process can be observed in well-described examples of recent colonization events. In a study on *Prunus avium*, Mariette *et al.* (1997) showed that a very young population ( $\sim 120$  years old) showed almost the same level of diversity as a long-established one, and neither population was differentiated ( $G_{ST} = 0.014$ ). This is especially interesting because this species is insect-pollinated, so that a high level of gene flow through pollen is not expected. Seeds are dispersed at random by animals, so the number of seeds per year that fall and are able to germinate on a given site is presumably low. In spite of limited recruitment, the delay in reproduction allows genetic diversity to accumulate. A similar pattern can be observed in *Sorbus aucuparia*, another entomophilous tree species, where a recent colonization of a pla-



**TABLE 1**  
**Main results obtained with the different models for diversity within and among populations**

Model	Life cycle	Main results
1D colonization process	Trees	Low founder effect, low differentiation regardless of the pollen flow for nuclear and cytoplasmic genes.
1D colonization process	Annual plants	High founder effect, high differentiation, except with very high pollen flow for nuclear genes.
2D colonization process	Trees	Provided pollen flow and effective population sizes are high enough, cytoplasmic genes differentiate among refuges during the glacial period but not nuclear genes. The colonization period does not have much effect. Therefore, after colonization, the situation is very different between nuclear and cytoplasmic genes. The homogenization is then slow.
2D colonization process	Annual plants	Both nuclear and cytoplasmic genes differentiate strongly during the glacial period, regardless of the level of pollen flow. Because $F_{ST}$ is already very high, it falls rapidly during the colonization period and continues to fall quickly during the homogenization period.
2D colonization process	Plants with nonoverlapping generations and long generation time	Same results as for trees during the isolation period. Strong founder effect during colonization, very slow homogenization afterward; $F_{ST}$ much higher than for trees at the end, except for very high pollen flow.

teau in Belgium yielded only a small loss of genetic diversity (Raspé and Jacquemart 1998).

By contrast, two species of *Silene* in Hawaii provide a good example of the impact of a recent colonization on genetic diversity and differentiation of perennial species with no juvenile phase (Westerbergh and Saura 1994). The older populations of these species clearly show the highest level of internal diversity, and founder effect is extremely high for populations derived from recent colonization of lava beds. Also, as expected, these recent colonizations yield very high levels of  $F_{ST}$  (0.677). *S. alba*, which colonized Virginia ~200 years ago, does not show much differentiation for nuclear genes ( $F_{ST} = 0.13$ ; see McCauley 1994, 1997). This illustrates how subsequent homogenization can counterbalance initial founder events relatively quickly, the discrepancy with cytoplasmic genes ( $F_{ST} = 0.67$ ) coming presumably from the differences between seed and pollen flow.

Our model also helps us to understand the general pattern of genetic diversity observed on a continental scale for temperate forest tree species: both the differences with annual plants as well as the opposite patterns of diversity of nuclear and cytoplasmic genes. On this last point, the two-dimensional colonization process shows that the differences between these two kinds of differentiation reflect not only differential gene flow during and after the colonization period, but also the fact that we have much differentiation between refuges for cytoplasmic genes and very little differentiation for nuclear genes during the isolation period.

In this two-dimensional process, pollen flow within refuges slows down differentiation by drift among the populations. Differentiation for nuclear genes only oc-

curs among the refuges. For cytoplasmic genes, on the other hand, populations within refuges also diverge. This process is accelerated because the effective population size for cytoplasmic genes is half that of nuclear genes. Annual plants, because of their short generation time, have the time to diverge from one refuge to another for both nuclear and cytoplasmic genes. Effective population size is a very important parameter; tree refuges have the time to diverge genetically from one another during the glacial period if effective population size is ~100 but not if it is ~1000. This result is consistent with results (Slatkin 1991) on the island model, for which the rate of convergence toward equilibrium is approximately governed by the smaller of  $1/N_e$  and  $m/d$ , where  $N_e$  is the effective size of the demes,  $m$  the migration rate, and  $d$  the number of demes.

Provided that tree populations were already only slightly differentiated for nuclear genes when they entered the refuges and that the effective population sizes in the refuges were large enough, the tree populations might have maintained this low level of differentiation throughout the glacial period. Because recolonization does not yield a strong increase in differentiation for trees,  $F_{ST}$  would remain relatively low, without invoking a high level of pollen flow (between 1% and 10% of pollen coming from the neighbor populations, that is, between 20 and 200 times more than seed flow).

One of the most surprising results is the decrease of  $F_{ST}$  during the colonization period for the two-dimensional stepping-stone model for annual plants and also in some cases for trees. The explanation is probably that because  $F_{ST}$  had reached a very high value at the end of the isolation period, the restoration of gene flow between the refuges in the colonization period

overwhelmed the effects of foundation events. Le Corre *et al.* (1997b) found a similar result. The number of founders of each new population was probably also important; other results (Le Corre and Kremer 1998) have shown that  $F_{ST}$  can decrease during a colonization process if the number of founders is high enough.

The results obtained for trees are not just a consequence of their generation time. Plants with nonoverlapping generations and the same generation time as trees resembled trees in the change in their population structure during the isolation periods. However, they differentiated more during the colonization period (Figure 5) and did not return to equilibrium within the subsequent 8000 years, except in the case where pollen flow was very high.

The two-dimensional pattern of colonization also played a major role, increasing the number of source populations for each newly founded population and thus strongly reducing the founder effect. Also, when gene flow was high enough ( $Nm > 1$ ), the rate of convergence was increased in a two-dimensional, compared to a one-dimensional, stepping-stone model (Maruyama 1971; Slatkin 1991).

Another important factor is pollen flow. A key indicator of its role is that paternally inherited chloroplasts show much less differentiation than maternally inherited mitochondria in pines (Latta and Mitton 1997). Nevertheless, Latta and Mitton show that while mitochondrial DNA is very differentiated between colonization lines coming from different refuges ( $F_{ST} = 0.679$ ), it is almost not differentiated along a colonization line ( $F_{ST} = 0.005$ ), well in agreement with the results of our one-dimensional colonization model.

Moreover, explanations based only on high pollen flow (Ennos 1994) cannot account for the observation that the  $F_{ST}$  value for tree species does not depend on their pollen dispersal ability, because animal-pollinated species do not show more differentiation than wind-pollinated species (Hamrick *et al.* 1992). On the other hand, low differentiation can be explained by life cycle, which does not depend on the level of pollen flow.

Our model reveals the complementarity of all processes. Figure 5 shows that if the tree life cycle was not taken into account, a huge level of pollen flow would be necessary to explain observed differentiation. Nevertheless, with the same level of seed and pollen flow, nuclear genes'  $F_{ST}$  for annual plants falls below that of trees at the end of the process because of the rapidity of homogenization of annual plant populations (see Figure 4). Therefore, higher levels of pollen flow or larger effective population sizes for trees than for annual plants are also necessary to explain the observed differences. Additionally, the two-dimensional process appears to be a key explanation for the lack of genetic differentiation.

Because the simulation is time intensive, we used a smaller number of populations than would be realistic

in temperate areas. Increasing the number of populations may have increased the effect of colonization on differentiation, but probably not more for trees than for annual plants. With additional populations, more time would have been required to reach equilibrium, especially in a stepping-stone model. This could have generated a situation in which annual plants would not have had the time to return to equilibrium, therefore explaining in part why they have a high  $F_{ST}$ .

Several other features could be integrated into the model. For instance, Le Corre *et al.* (1997b) have shown by a simulation study that events of long distance dispersal of seeds by birds have to be invoked to explain the extremely high speed with which forest trees have recolonized the whole temperate zone after the last glacial period. This result agrees with direct observation (Bossema 1979; Darley-Hill and Johnson 1981). For annual species, these long distance dispersal events would generate a founder effect that is much greater than that obtained by a stepping-stone colonization process, some newly founded populations remaining isolated for a long period. It will be of interest to see if tree population dynamics might reduce this high founder effect. These rare events of long distance dispersal could explain the observation of a very strong differentiation for chloroplast DNA at a local scale, as pointed out by Petit *et al.* (1997).

Our results have shown that the level of differentiation of trees is a consequence of the differentiation that existed before the last glacial period. The last million years have seen cycles of  $\sim 100,000$  years with glacial periods of 80,000 years, followed by short warm periods (Hays *et al.* 1976; Pailard 1997). There were probably successive events of isolation followed by colonization. Under our hypothesis, one might think that the particular life cycle of tree species permits them to pass through these successive events without suffering many changes in the structure of their nuclear gene diversity, while the cytoplasmic gene differentiation was more deeply shaped by the process. In some cases, the effective size of the refuges may not have been large enough, yielding much more among-population differentiation and much less within-population diversity (see, for instance, the case of *Pinus pinea* described by Fallour *et al.* 1997). The loss of diversity through fragmentation has been observed for some tree species (Billington 1991).

The method that we developed here is relatively specific to plant populations, because migration is presumed to occur through pollen and seeds. Nevertheless, the method could be adapted to the case of animals or humans, for which migration can also occur for adults or whole families. One might expect the same avoidance of founder effects during colonization for long-lived species, because even if the first migrants in a newly founded population can reproduce immediately, their descendants will then have to wait for many years before being able to reproduce, permitting the population to

be filled by the arrival of new migrants as in the case for trees, thus limiting the founder effect.

This could help explain the low genetic differentiation of humans on a worldwide scale, which is rather surprising, because the colonization of the whole world by humans is probably very recent (Cavalli-Sforza *et al.* 1994) and could have involved major founder effects. Differences between populations account for only 15% of the global diversity (Barbujani *et al.* 1997; Barton 1997), whereas the average  $F_{ST}$  of mammals is  $\sim 0.242 \pm 0.030$  (Ward *et al.* 1992).

We thank Evelyne Heyer, Bernard Jung-Muller, Mark Kirkpatrick, Antoine Kremer, Valérie Le Corre, and Jacquie Shykoff for helpful discussions and comments on the manuscript. We thank the associate editor and two anonymous reviewers for helpful comments. We also thank Liz Mason and Peter Smouse for their help with English usage. F.A. and S.M. have a Formation Complémentaire Par la Recherche grant from the French Ministère de l'Agriculture. This work was funded in part by the "Programme National Dynamique de la Biodiversité et Environnement" of the French Centre National de la Recherche Scientifique. Some simulations were performed on the SP2 and SGI computers of the Centre des Ressources Informatiques of the Université de Paris-Sud (Orsay).

#### LITERATURE CITED

- Andersen, B. G., and H. W. Borns, Jr., 1994 *The Ice Age World: An Introduction to Quaternary History and Research With Emphasis on North America and Northern Europe During the Last 2.5 Million Years*. Scandinavian University Press, Oslo.
- Austerlitz, F., B. Jung-Muller, B. Godelle and P.-H. Gouyon, 1997 Evolution of coalescence times, genetic diversity and structure during colonization. *Theor. Popul. Biol.* **51**: 148–164.
- Barbujani, G., A. Magnani, E. Minch and L. L. Cavalli-Sforza, 1997 An apportionment of human diversity. *Proc. Natl. Acad. Sci. USA* **94**: 4516–4519.
- Barton, N. H., 1997 Population genetics: a new apportionment of human diversity. *Curr. Biol.* **7**: R757–R758.
- Billington, H. L., 1991 Effect of population size on genetic variation in a dioecious conifer. *Conserv. Biol.* **5**: 115–119.
- Bosch, C. A., 1971 Redwoods: a population model. *Science* **174**: 345–349.
- Bossema, I., 1979 Jays and oaks: an eco-ethological study of a symbiosis. *Behaviour* **70**: 1–117.
- Buongiorno, J., and B. R. Michie, 1980 A matrix model of uneven-aged forest management. *Forest Sci.* **26**: 609–625.
- Buongiorno, J., J. L. Peyron, F. Houllier and M. Bruciamacchie, 1995 Growth and management of mixed-species, uneven-aged forests in the French Jura: implications for economic returns and tree diversity. *Forest Sci.* **41**: 397–429.
- Cavalli-Sforza, L. L., P. Menozzi and A. Piazza, 1994 *The History and Geography of Human Genes*. Princeton University Press, Princeton, NJ.
- Darley-Hill, S., and W. C. Johnson, 1981 Dispersal of acorns by blue jays (*Cyanocitta cristata*). *Oecologia* **50**: 231–232.
- Dong, J., and D. B. Wagner, 1994 Paternally-inherited chloroplasts in *Pinus*: estimation of diversity and population subdivision and tests of disequilibrium with a maternally inherited mitochondrial polymorphism. *Genetics* **136**: 1187–1194.
- Ennos, R. A., 1994 Estimating the relative rates of pollen and seed migration among plant populations. *Heredity* **72**: 250–259.
- Fallour, D., D. Fady and F. Lefevre, 1997 Study on isozyme variation in *Pinus pinea* L.: evidence for low polymorphism. *Silvae Genet.* **46**: 201–207.
- Hamrick, J. L., and M. J. W. Godt, 1996 Effects of life history traits on genetic diversity in plants. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* **351**: 1291–1298.
- Hamrick, J. L., M. J. W. Godt and S. L. Sherman-Broyles, 1992 Factors influencing levels of genetic diversity in woody plant species. *New For.* **6**: 95–124.
- Harcombe, P. A., 1987 Tree life tables: simple birth, growth and death data encapsulate life histories and ecological roles. *BioScience* **37**: 557–568.
- Hays, J. D., J. Imbrie and N. J. Shackleton, 1976 Variations in the earth's orbit: pacemakers of the ice ages. *Science* **194**: 1121–1132.
- Huntley, B., 1990 European vegetation history: paleovegetation maps from pollen data—13000 yr BP to present. *J. Quat. Sci.* **5**: 103–122.
- Huntley, B., and H. J. B. Birks, 1983 *An Atlas of Past and Present Pollen Maps for Europe: 0–13000 Years Ago*. Cambridge University Press, Cambridge, United Kingdom.
- Kremer, A., 1994 Diversité génétique et variabilité des caractères phénotypiques chez les arbres forestiers. *Genet. Sel. Evol.* **26** (Suppl. 1): 105–123.
- Latta, R. G., and J. B. Mitton, 1997 A comparison of population differentiation across four classes of gene markers in limber pine (*Pinus flexilis* James). *Genetics* **146**: 1153–1163.
- Le Corre, V., 1997 Organisation de la diversité génétique et histoire postglaciaire des chênes blancs européens: approche expérimentale et par simulation. Ph.D. thesis, Institut National Agronomique Paris-Grignon.
- Le Corre, V., and A. Kremer, 1998 Cumulative effects of founding events during colonisation on genetic diversity and differentiation in an island and stepping-stone model. *J. Evol. Biol.* **11**: 495–512.
- Le Corre, V., S. Dumolin-Lapègue and A. Kremer, 1997a Genetic variation at allozyme and RAPD locus in sessile oak *Quercus Petraea* (Matt.) Liebl.: the role of history and geography. *Mol. Ecol.* **6**: 1–11.
- Le Corre, V., N. Machon, R. J. Petit and A. Kremer, 1997b Colonization with long-distance seed dispersal and genetic structure of maternally inherited genes in forest trees: a simulation study. *Genet. Res.* **69**: 117–125.
- Lefkovich, L. P., 1965 The study of population growth in organisms grouped by stages. *Biometrics* **21**: 1–18.
- Leslie, P. H., 1945 On the use of matrices in certain population mathematics. *Biometrika* **33**: 183–212.
- Lewis, N. A., 1942 On the generation and growth of a population. *Sankhya* **6**: 93–96.
- Mariette, S., M. Lefranc, P. Legrand, D. Taneyhil, N. Frascaria-Lacoste *et al.*, 1997 Genetic variability in wild cherry populations in France. Effects of colonizing processes. *Theor. Appl. Genet.* **94**: 904–908.
- Maruyama, T., 1971 Analysis of population structure II. Two-dimensional stepping-stone models of finite length and other geographically structured populations. *Ann. Hum. Genet.* **35**: 179–196.
- McCauley, D. E., 1994 Contrasting the distribution of chloroplast DNA and allozyme polymorphism among local populations of *Silene alba*: implications for studies of gene flow in plants. *Proc. Natl. Acad. Sci. USA* **17**: 8127–8131.
- McCauley, D. E., 1997 The relative contribution of seed and pollen movement to the local genetic structure of *Silene alba*. *J. Hered.* **88**: 257–263.
- McCauley, D. E., 1998 The genetic structure of a gynodioecious plant: nuclear and cytoplasmic genes. *Evolution* **52**: 255–260.
- Nei, M., 1973 Analysis of the gene diversity in subdivided populations. *Proc. Natl. Acad. Sci. USA* **70**: 3321–3323.
- Orive, M. E., 1993 Effective population size in organisms with complex life-histories. *Theor. Popul. Biol.* **44**: 316–340.
- Pailard, D., 1997 The timing of Pleistocene glaciations from a simple multiple-state climate model. *Nature* **391**: 378–381.
- Petit, R. J., A. Kremer and D. B. Wagner, 1993a Geographic structure of chloroplast DNA polymorphisms in European white oaks. *Theor. Appl. Genet.* **87**: 122–128.
- Petit, R. J., A. Kremer and D. B. Wagner, 1993b Finite island model for organelle and nuclear genes in plants. *Heredity* **71**: 630–641.
- Petit, R. J., E. Pineau, B. Demesure, R. Bacilieri, A. Ducouso *et al.*, 1997 Chloroplast DNA footprints of postglacial recolonization by oaks. *Proc. Natl. Acad. Sci. USA* **94**: 9996–10001.
- Raspé, O., and A.-L. Jacquemart, 1998 Allozyme diversity and genetic structure of European populations of *Sorbus aucuparia* L. (Rosaceae: Maloideae). *Heredity* **81**: 537–545.
- Raybould, A. F., J. Goudet, R. J. Mogg, C. J. Gliddon and A. J.

- Gray, 1996 Genetic structure of a linear population of *Beta vulgaris* ssp. *maritima* sea beet revealed by isozyme and RFLP analysis. *Heredity* **76**: 111–117.
- Slatkin, M., 1991 Inbreeding coefficients and coalescence times. *Genet. Res.* **58**: 167–175.
- Templeton, A. R., and D. A. Levin, 1979 Evolutionary consequences of seed pools. *Am. Nat.* **114**: 232–249.
- Usher, M. B., 1966 A matrix approach to the management of renewable resources, with special reference to selection forests. *J. Appl. Ecol.* **3**: 355–367.
- Usher, M. B., 1969 A matrix model for forest management. *Biometrics* **25**: 309–315.
- Ward, D. W., D. O. F. Skibinski and M. Woodward, 1992 Protein heterozygosity, protein structure and taxonomic differentiation. *Evol. Biol.* **26**: 73–152.
- Westerbergh, A., and A. Saura, 1994 Genetic differentiation in endemic *Silene* Caryophyllaceae on the Hawaiian Islands. *Am. J. Bot.* **81**: 1487–1493.
- Wright, S., 1951 The genetical structure of populations. *Ann. Eugen.* **15**: 323–354.

Communicating editor: R. G. Shaw

## APPENDIX A

Regulation in each population occurs as follows. We denote by  $\mathbf{Y}(t)$  the vector of the numbers of individuals in the population that would be potentially recruited in each size class from  $t$  to  $t + 1$  if their was no regulation, and by  $\mathbf{S}(t)$  the vector of the numbers of potential surviving individuals that remain in each size class during the interval  $t$  to  $t + 1$ .  $\mathbf{Y}'(t) = \alpha(t)\mathbf{Y}(t)$  and  $\mathbf{S}'(t) = \beta(t)\mathbf{S}(t)$  are the corresponding vectors after regulation.  $\alpha(t)$  and  $\beta(t)$  are regulation parameters comprised between 0 and 1, which depend on the total stand basal area of the population  $G(t)$ ; they are adjusted so that  $G(t)$  grows logistically to an equilibrium value.

$S_i(t) = P_{ii}N_i(t)$  is the number of potential surviving trees in size class  $i$  during the interval  $t$  to  $t + 1$  and the corresponding stand basal area is  $G(\mathbf{S}(t)) = \sum_{i=1}^k G_i S_i(t)$ , with  $G_i$  the stand basal area of a tree of class  $i$ ,  $\mathbf{S}(t) = (S_1(t), S_2(t), \dots, S_k(t))$ .  $Y_1(t) = \sum_{i=1}^k f_i N_i(t)$  is the number of individuals that are potentially recruited in the first size class,  $Y_{i+1}(t) = P_{i+1}N_i(t)$  is the number of trees that potentially pass from size class  $i$  to size class  $i + 1$ , and the corresponding stand basal area is  $G(\mathbf{Y}(t)) = \sum_{i=1}^k G_i Y_i(t)$ , with  $\mathbf{Y}(t) = (Y_1(t), Y_2(t), \dots, Y_k(t))$ .

$\mathbf{S}'(t)$  and  $\mathbf{Y}'(t)$  are the corresponding vectors after regulation and  $N_i(t + 1) = \mathbf{S}'(t) + \mathbf{Y}'(t)$ .  $\alpha(t)$  and  $\beta(t)$  are the coefficients of regulation for  $\mathbf{Y}(t)$  and  $\mathbf{S}(t)$ :  $\mathbf{Y}'(t) = \alpha(t)\mathbf{Y}(t)$  and  $\mathbf{S}'(t) = \beta(t)\mathbf{S}(t)$ .  $\Delta G(t) = G(\mathbf{S}(t)) + G(\mathbf{Y}(t)) - G(t)$  is the potential increase of stand basal area between  $t$  and  $t + 1$  and  $\Delta G'(t) = G(\mathbf{S}'(t)) + G(\mathbf{Y}'(t)) - G(t)$  is the real increase of stand basal area. Regulation is introduced by limiting this increase:  $\Delta G'(t) = \Delta G(t)(1 - G(t)/G_{\text{eq}})$  with  $G_{\text{eq}}$  the equilibrium  $G$  for the population.

This gives only one relation between  $\alpha(t)$  and  $\beta(t)$ . To calculate them completely, another relation has to be chosen arbitrarily; this will determine the proportion

of the two categories of individuals (the ones that remain in the same class and the ones that move to the next class) that is affected by density dependence. The probabilities of transition between size classes are strongly affected by the basal area of the stand (Buongiorno *et al.* 1995). Therefore, for simplicity, we fixed  $\beta(t)$  at 1, making density dependence act only on the number of individuals that move to the next class and on the number of offspring. This leads to the relation

$$\alpha(t) = \frac{(G(t) + \Delta G'(t) - G(\mathbf{S}(t)))}{G(\mathbf{Y}(t))}.$$

We chose here the following parameters:  $k = 25$ ,  $f = 250$ ,  $G_{\text{eq}} = 91,912$ ,  $P_{11} = 0$ ,  $P_{ii} = 0.54$  for  $2 \leq i \leq 24$ ,  $P_{25,25} = 0.98$ ,  $P_{12} = 0.1$ ,  $P_{i,i+1} = 0.4$  for  $2 \leq i \leq 24$ ,  $G_1 = 0.01$ ,  $G_i = 0.5$  for  $2 \leq i \leq 8$ ,  $G_i = 1$  for  $9 \leq i \leq 16$ ,  $G_i = 2$  for  $17 \leq i \leq 24$ ,  $G_{25} = 3$ . These values are somewhat arbitrary; they were adjusted principally to fulfill the main characteristics of tree species as indicated in the main text.

## APPENDIX B: ITERATION METHOD FOR COMPUTING THE PROBABILITIES OF IDENTITY BY DESCENT

Migration is defined by two matrices  $M_{\text{seed}}$  and  $M_{\text{poll}}$ , which are set according to Figure 2, A or B.  $M_{\text{seed}}(p, q)$  and  $M_{\text{poll}}(p, q)$  denote the proportion of seeds and pollen, respectively, that arrive in  $q$  coming from  $p$ . Demography behaves as follows for each population  $p$  and each class  $i$ :

$$N_{pi}(t + 1) = P_{ii}^c N_{pi}(t) + P_{i,i-1}^c N_{p,i-1}(t) \quad (i > 1),$$

$$N_{p,1}(t + 1) = \sum_{q=1}^d N_{q,k}(t) M_{\text{seed}}^c(q, p),$$

where the  $c$  superscript denotes in each case the values corrected by density dependence. Thus, if we denote by  $P_c(p, i, q, j)$  the probability for a cytoplasmic gene of an individual of population  $p$  and class  $i$  at time  $t + 1$  to have been in population  $p$  and class  $i$  at time  $t$ , this probability is

$$P_c(p, i, p, i) = \frac{P_{ii}^c N_{pi}(t)}{N_{pi}(t + 1)} \left. \vphantom{\frac{P_{ii}^c N_{pi}(t)}{N_{pi}(t + 1)}}} \right\} \text{for whatever } p \text{ and } i > 1,$$

$$P_c(p, i, p, i - 1) = \frac{P_{i,i-1}^c N_{p,i-1}(t)}{N_{pi}(t + 1)}$$

$$P_c(p, 1, p, k) = \frac{N_{q,k}(t) M_{\text{seed}}^c(q, p)}{N_{p,1}(t + 1)} \quad \text{for whatever } p \text{ and } q,$$

$$P_c(p, i, q, j) = 0 \quad \text{for any other combination of } p, i, q, \text{ and } j.$$

Then the IBD for two cytoplasmic genes in any population and any class can be deduced as

$$f_{p,i,p',i'}^c(t+1) = \sum_{\substack{1 \leq q,q' \leq d \\ 1 \leq j,j' \leq k}} P_c(p,i,q,j) P_c(p',i',q',j') \\ \times [(1 - \alpha_{q,q',i,i',j,j'}^c) f_{q,j,q',j'}^c(t) + \alpha_{q,q',i,i',j,j'}^c],$$

where  $\alpha_{q,q',i,i',j,j'}^c$  denotes the probability that two genes are IBD because one is the parent of the other or that they both had the same parent at time  $t$ ,

$$\alpha_{q,q',i,i',j,j'}^c = \frac{(1 - \mu)^2}{N_{q,k}(t)} \text{ if } q = q' \text{ and } j = j' = k \text{ and } i = i' = 1,$$

$$\alpha_{q,q',i,i',j,j'}^c = \frac{1 - \mu}{N_{q,k}(t)} \text{ and } \begin{cases} i = 1, i' = k, j = k, j' = k \\ \text{or} \\ i = k, i' = k, j = 1, j' = k \end{cases},$$

$$\alpha_{q,q',i,i',j,j'}^c = 0 \text{ otherwise,}$$

where  $\mu$  denotes the mutation rate.

When reproduction is not involved, nuclear genes behave like cytoplasmic genes. Thus, if we denote by  $P_n(p, i, q, j)$  the probability for a nuclear gene of an individual of population  $p$  and class  $i$  at time  $t + 1$  to have been in population  $q$  and class  $j$  at time  $t$ , this probability is

$$P_n(p,i,q,j) = P_c(p,i,q,j)$$

for whatever  $p,q,j$  and whatever  $i > 1$ .

For the juvenile class ( $i = 1$ ), an individual inherits

half of his genes maternally and half paternally. When received maternally, the distribution of probability of origin for this gene is the same as for a cytoplasmic gene. When received paternally, we have to take into account pollen flow: for whatever population  $q$  and  $s$ , the proportion of pollen from  $s$  in the pollen cloud of  $q$  is

$$P_{\text{poll}}(s,q) = \frac{N_{s,k}(t) M_{\text{poll}}(s,q)}{\sum_{s'=1}^d N_{s',k}(t) M_{\text{poll}}(s',q)}$$

and then the probability for an individual of population  $p$  and class 1 to inherit a gene paternally from population  $q$  and class  $k$  will be

$$P_{\text{pat}}(p,1,q,k) = \sum_{s=1}^d P_c(p,1,s,k) P_{\text{poll}}(s,q).$$

Altogether

$$P_n(p,1,q,k) = \frac{P_c(p,1,q,k) + P_{\text{pat}}(p,1,q,k)}{2}.$$

Then as above,

$$f_{p,i,p',i'}^n(t+1) = \sum_{\substack{1 \leq q,q' \leq d \\ 1 \leq j,j' \leq k}} P_n(p,i,q,j) P_n(p',i',q',j') \\ \times [(1 - \alpha_{q,q',i,i',j,j'}^n) f_{q,j,q',j'}^n(t) + \alpha_{q,q',i,i',j,j'}^n],$$

where  $\alpha_{q,q',i,i',j,j'}^n$  is the same as  $\alpha_{q,q',i,i',j,j'}^c$  except that  $N_{q,k}(t)$  is replaced by  $2N_{q,k}(t)$ .