Journal of Evolutionary Biology

Cumulative effects of founding events during colonisation on genetic diversity and differentiation in an island and stepping-stone model

V. Le Corre and A. Kremer*

Laboratoire de Génétique et Amélioration des arbres forestiers, Institut National de la Recherche Agronomique, B.P. 45, F-33611 Gazinet Cedex, France

Key words: Colonisation; founder effect; identity-by-descent; island model; stepping-stone model.

Abstract

This paper investigates the cumulative effect of founding events on the genetic differentiation and the within-population heterozygosity in a metapopulation increasing its size by colonisation. Two contrasting models are considered: first, an island model, where migrants and colonists are taken at random from the entire metapopulation, and second, a linear stepping-stone model, where migrants and colonists are sampled from a limited neighbourhood. The genetic consequences of a range expansion depend on the relative magnitudes of the number of colonists and migrants, in a way similar to extinction and colonisation processes (Wade and McCauley, 1988). The cumulative effect of founding events, resulting most often in a transient increase in genetic differentiation and a gradual loss of within-population heterozygosity, also depends on the age-structure that is established during colonisation. It is the highest when colonists are sampled from recently founded populations and migrants are exchanged among populations of similar ages. The genetic consequences of a range expansion are therefore far more pronounced and lasting in the linear stepping-stone model than in the island model. These two models, however, represent the two extremes between which real populations will fall.

* Author for correspondence, e-mail: Antoine.Kremer@pierroton.inra.fr

Introduction

Many ecological studies in the past few years have emphasised the fact that natural populations of plants and animals are dynamic systems subject to variations both in time and space (Bascompte and Solé, 1995). As a consequence, population genetics studies often have to deal with non-equilibrium states which have only rarely been investigated by classical theories. Founding events are likely to create persistent non-equilibrium structures (Boileau et al., 1992), and can arise by many ways: for example by recolonization of a habitat left unoccupied after the extinction of a previous population, by the introduction of a species in a new territory, or during the expansion of natural ranges of species following environmental disturbances such as the last post-glacial climatic warming.

Following the pioneer work of Slatkin (1977), more recent studies (Wade and McCauley, 1988; Whitlock and McCauley, 1990; Wade et al., 1994) have described some theoretical models for an island metapopulation undergoing extinctions compensated for by recolonisations. Extinction and recolonisation were shown to result in the establishment of an age-structure that may greatly affect the level of genetic differentiation among populations. When the rates of extinction and recolonisation are identical and constant over time, a new genetic equilibrium is obtained, which is different from the equilibrium in the classical island model. This kind of model is not fully adequate for the description of the range expansion of species, where founding events occur successively in the absence of extinctions. The study of the evolution of coalescence times seems a promising approach for the genetic characterisation of such processes (Austerlitz et al., 1996). However this parameter is directly obtainable from the analysis of common genetic markers like allozymes. Our aim in this study is to derive simple models for the probability of identity among alleles in a metapopulation that increases in size by colonization, without extinctions. First, an island metapopulation model is described, where new populations are founded by colonists taken at random from the entire metapopulation. Second, a one-dimensional stepping-stone model is described, where vacant sites are colonised by individuals originating from nearby populations. Recurrence equations are formulated for the change in the probability of identity-by-descent (i.b.d.) of alleles within or among populations. Various genetic parameters may be used to describe the genetic structure of a metapopulation. In this study, two parameters will be calculated: first, the genetic differentiation among populations, F_{st}, second the within-population heterozygosity, Ho. Since founding events are likely to increase the genetic divergence among populations (Slatkin, 1977, Wade and McCauley, 1988), the value of F_{st} will indicate to what extent a colonisation process may globally modify the partition of genetic diversity, within and between populations. A loss of within population heterozygosity is also expected to occur at each founding event (Nei et al., 1975), resulting in a decrease of genetic diversity from the most ancient to the most recent populations. Colonisation processes will therefore be characterised also by the resulting pattern in within-population heterozygosity.

1. General features for the models

We consider a set of populations, each containing N reproducing individuals, that exchange migrants at a rate m constant over time. Mutations are negligible compared to migration, and we assume selective neutrality among alleles. There is an infinite number of vacant sites open to colonisation. New populations are founded in each generation by the arrival of k colonists in vacant sites. Both migrants and colonists are diploid individuals which are sampled in populations and distributed among sites before the random sampling of gametes for reproduction. The k colonists that found a new population immediately reproduce in panmixia to reach the size N.

2. Colonisation in an island metapopulation model

2.1. The model

In this model, individuals which move (either migrants if they move to an extant population, or colonists if they fall in a vacant site) are taken at random from the entire metapopulation. The initial metapopulation, before the beginning of the colonisation, is assumed to be at genetic equilibrium and the number of populations is sufficiently large that only genes within the same population are significantly related. The genetic differentiation among population, F_{st}, is therefore equal to the mean probability of i.b.d. of two alleles chosen within a population (Whitlock and McCauley, 1990). Furthermore, the probability that two colonists or two migrants entering a given site come from the same source population is negligible.

As a consequence, i.b.d. of two alleles has the following origins:

– twice the same allele is considered;

– within a newly founded population, two different alleles carried by the same colonist individual are considered (with probability $1/(2k-1)$);

– within an established population, either two different alleles from non-migrant individuals (probability $(1-m)^2$), or two different alleles carried by the same migrant individual (probability $m^2(1/(2Nm-1))$) are considered.

The island metapopulation model does not deal with the spatial distribution of populations. However, colonisation creates an age-structure, that will be taken into account by considering the values of $fo(t1, t2)$, the probability of i.b.d. measured at time $t2$ within populations that were founded at time $t1$. The heterozygosity measured at time $t2$ within a population founded at time $t1$, will be calculated as $Ho(t1, t2) = 1 - fo(t1, t2)$.

The probability of i.b.d. within newly founded populations, $fo(t1, t1)$, is given by a formula similar to formula (4) in Whitlock and McCauley (1990):

$$fo(t1, t1) = \frac{1}{2k} + \frac{1}{2k}\left[\frac{1}{2N} + \left(1 - \frac{1}{2N}\right)fo(\cdot, t1 - 1)\right] \qquad (1)$$

where N is the effective size of populations, k is the effective number of diploid colonists and $fo(\cdot, t1 - 1)$ is the mean probability of i.b.d. within populations at the preceding generation. The first term represents the fraction of identity-by-descent resulting from sampling twice the same allele in the new population, whereas the second term represents the identity by descent resulting from sampling two distinct alleles which were carried by the same colonist individual.

Once populations are founded, they will exchange migrants at a rate m with other populations. The probability of i.b.d. of two alleles chosen within any population will decrease at each generation after random mating, according to the following equation (with the same notations as above):

$$fo(t1, t2 + 1) = \frac{1}{2N} + \left(1 - \frac{1}{2N}\right)\left[(1 - m)^2 fo(t1, t2)\right.$$

$$\left. + m^2\left(\frac{1}{2Nm - 1}\right)fo(\cdot, t2)\right],$$

for $t2 \geq t1$, and $2Nm \neq 1$. 　　　　　　　　　　　　　　　　　　(2)

The first term again represents the identity by descent obtained by twice sampling the same allele in the population, whereas the second term accounts for the identity by descent between two distinct alleles which are either both non-migrant, or which were both carried by the same migrant individual.

Note that if $2Nm = 1$, the equation becomes

$$fo(t1, t2 + 1) = \frac{1}{2N} + \left(1 - \frac{1}{2N}\right)[(1 - m)^2 fo(t1, t2)]$$ 　　　　　(2')

2.2. Change in the genetic differentiation among populations

The value of genetic differentiation among populations at time $t2$ is obtained by averaging the probabilities of i.b.d. within populations across populations of all ages:

$$F_{st}(t2) = fo(\cdot, t2) = \sum_{t=0}^{t2} p(t, t2)fo(t, t2),$$ 　　　　　　　　　(3)

where $p(t, t2)$ is the relative size of the age-class t, i.e. the proportion of the populations which were founded at time t relative to the total number of populations at time $t2$.

We noted that a general expression for the change with time in the relative size of the age-class t is:

$$p(t, t2) = [1 - p(t2, t2)]p(t, t2 - 1),$$ 　　　　　　　　　　　(4)

where $p(t2, t2)$ is the proportion of the number of newly founded populations relative to the total number of populations at time $t2$.

Using equation (4) to take into account the change with time in the age-structure of the metapopulation, and equations (1) and (2) to obtain the probability of i.b.d. within populations of a given age, a simple recurrence was obtained from (3) for the value of genetic differentiation among populations:

$$F_{st}(t2) = A\ F_{st}(t2 - 1) + B,$$ (5)

where

$$A = \left(1 - \frac{1}{2N}\right)\left[p(t2, t2)\frac{1}{2k} + (1 - p(t2, t2))\right.$$

$$\times \left.\left[(1 - m)^2 + m^2 \frac{1}{2Nm - 1}\right]\right] \text{ if } 2Nm \neq 1,$$

$$\text{or } A = \left(1 - \frac{1}{2N}\right)\left[p(t2, t2)\frac{1}{2k} + (1 - p(t2, t2))[(1 - m)^2]\right] \text{ if } 2Nm = 1.$$

and $$B = \frac{1}{2N} + p(t2, t2)\left[\frac{1}{2k} - \left(1 - \frac{1}{2k}\right)\frac{1}{2N}\right].$$

As seen from these expressions, the genetic differentiation among populations will be determined by three different factors: 1, the strength of the bottleneck occurring at each colonisation event, which depends on the number k of colonists and the final size N of the population; 2, the homogenising action of gene flow, which depends on the proportion m of migrants exchanged among populations; and 3, the dynamics of the colonisation process, determining at each generation the proportion $p(t2, t2)$ of the colonisation events relative to the number of existing populations.

2.3. Conditions for an equilibrium value of genetic differentiation

As colonisation is going on, the genetic differentiation will tend towards a limit whose expression can be derived from equation (5). However, two cases are to be considered.

First, the relative proportion of the newly founded populations, $p(t2, t2)$ may decrease over generations, so that it tends to become negligible compared to the number of already colonised sites. In this case, the genetic differentiation tends towards the asymptotic value:

$$F_{st}(\infty) = \frac{1}{2N - (2N - 1)[(1 - m)^2 + m^2(1/(2Nm - 1))]} \text{ if } 2Nm \neq 1,$$

$$\text{or } F_{st}(\infty) = \frac{1}{2N - (2N - 1)[(1 - m)^2]} \text{ if } 2Nm = 1$$ (6)

When $m \ll 1$, these expression simplify, with the same degree of accuracy as that used by Wright (1940) in his analysis of the classical island model, to:

$$F_{st}(\infty) = \frac{1}{1 + 4Nm}$$

which is the equilibrium value of genetic differentiation in an island metapopulation model with neither colonisation nor extinction. However, before this value is asymptotically met, colonisation results in a transient change in genetic differentiation. Assuming that the initial metapopulation was at genetic equilibrium, we find (from equation 5) that the genetic differentiation among colonised sites is transiently increased compared to the intital genetic differentiation, $1/(1 + 4Nm)$, when $k < 1 + 2Nm$, i.e. when the number of individuals colonising each site is equal to or smaller than twice the number of migrant individuals exchanged among populations. Not surprisingly, this condition is the exact equivalent of the condition that holds for the value of genetic differentiation in a metapopulation undergoing extinction and recolonisation at constant, identical rates (Wade and McCauley, 1988; Whitlock and McCauley, 1990).

Second, it is biologically feasible that the proportion of colonising events $p(t2, t2)$ tends toward a constant value, π, instead of continuously decreasing during colonisation. This will happen when the number of colonised sites remains proportional at each generation to the number of extant populations. In this case, the genetic differentiation among colonised sites will attain an equilibrium value, which for $m \ll 1$, is approximated by:

$$F_{st}(eq) = \frac{1 + \pi\left[\frac{1}{2k}(2N + 1) - 1\right]}{1 + 4Nm + \pi\left[\frac{1}{2k} + \left(1 - \frac{1}{2k}\right)2N - 1 - 4Nm\right]} \tag{7}$$

$F_{st}(eq)$ is strictly higher than the genetic differentiation before the beginning of the colonisation, $1/(1 + 4Nm)$, when $k < 1 + 2Nm$, which is the same condition as above.
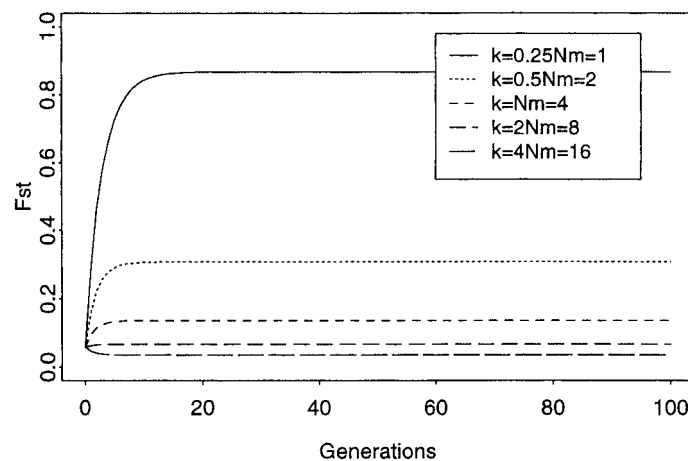
Fig. 1. Change over time in the genetic differentiation F_{st} during colonisation in the "outbreak" island model for different values of k, the number of colonists. Other parameters are $N = 100$ and $m = 0.04$. The metapopulation is geometrically increasing in size at a rate $\alpha = 1$.
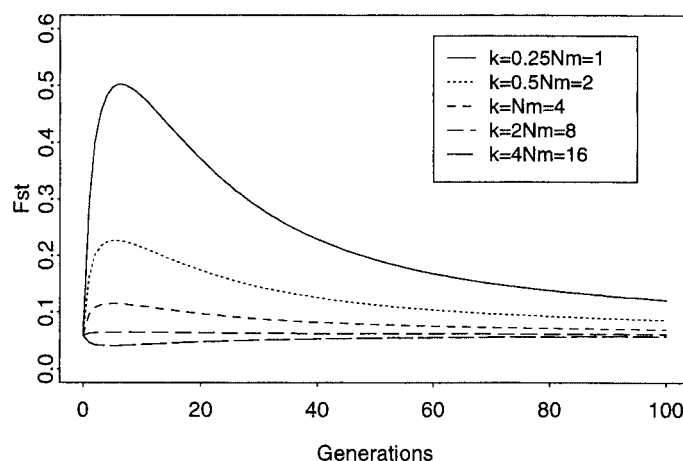
Fig. 2. Change over time in the genetic differentiation F_{st} during colonisation among colonised sites in the "linear diffusion" island model for different values of k, the number of colonists. Other parameters are $N = 100$ and $m = 0.04$. The metapopulation is linearly increasing in size at a rate $x = 1$.

2.4. Featuring different colonisation dynamics

The island model may be applied to different kinds of colonisation processes by determining for each of them the appropriate expression as a function of time of $p(t2, t2)$, the relative proportion at each generation of the newly founded populations. Some simple cases may be described for illustration.

First, as noticed above, a constant ratio is obtained when the number of colonisation events increases proportionally to the number of populations. In this case, the size of the metapopulation will grow geometrically at a rate α, and $p(t2, t2)$ will be equal at each time $t2$ to $\alpha/(1 + \alpha)$. This kind of outbreak demographical increase may characterise initial phases of expansion of an organism in a non-restrictive environment. An example is the establishment of the collared dove in the Netherlands and Great Britain (Hengeveld, 1989).

More often biological invasions proceed at a roughly constant rate, by steady diffusion at the species' range boundaries (Hengeveld, 1989). If spread occurs in a single direction, as along shores, roads, or other one-dimensional habitats, we may expect the number of colonised sites to increase linearly with time at a rate α. The relative proportion of the newly founded populations, $p(t2, t2)$ will therefore be equal to $\alpha/(1 + \alpha t2)$. If colonisation proceeds at a constant rate in all directions from a starting point, the square root of the number of populations will increase linearly with time (Hengeveld, 1989), and $p(t2, t2)$ will be equal to $\alpha(2\alpha t2 - \alpha + 2)/(1 + \alpha t2)^2$.

Numerical results are presented for two simple contrasting cases: first, an "outbreak" model with a geometrical increase with time in the number of populations (Fig. 1), second, a "linear diffusion" model with a linear increase with time in the number of populations (Fig. 2).
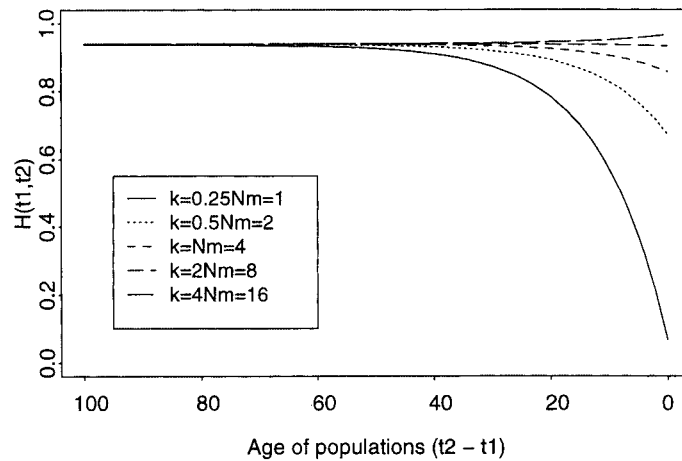
Fig. 3. Within-population heterozygosity after 100 generations of colonisation as a function of populations' ages in the "outbreak" island model for different values of k, the number of colonists. Other parameters are $N = 100$ and $m = 0.04$. The metapopulation increased geometrically in size during colonisation at a rate $\alpha = 1$.

2.5. Relationship between migration and colonisation

An important feature in colonisation is the relationship between the number k of individuals colonising empty sites and the number Nm of individuals moving between extant populations (Wade and McCauley, 1988; Whitlock and McCauley, 1990).
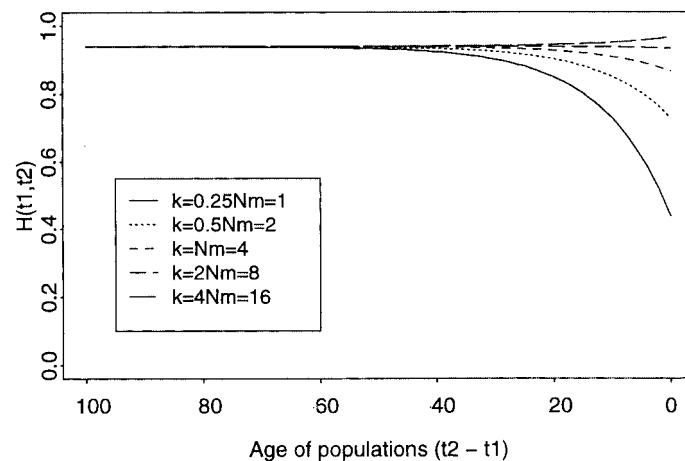
Fig. 4. Within-population heterozygosity after 100 generations of colonisation as a function of populations' ages in the "linear diffusion" island model for different values of k, the number of colonists. Other parameters are $N = 100$ and $m = 0.04$. The metapopulation increased linearly in size during colonisation at a rate $\alpha = 1$.

For $k < 1 + 2Nm$, the genetic differentiation increases during the first generations of colonisation, until founder effects are compensated for by the homogenising action of gene flow among colonised sites. The maximum value of differentiation reached is highly dependent on the relationship between k and Nm, and increases non-linearly with the ratio of k/Nm (Figs. 1 and 2). The genetic differentiation then remains at a stable equilibrium if the age-structure among populations remains unchanged (Fig. 1).

Otherwise F_{st} will asymptotically decrease towards $1/(1 + 4Nm)$ as the age-structure evolves towards a predominance of the populations founded first (Fig. 2). For $k > 2Nm + 1$, there is a transient decrease in the value of genetic differentiation. As already noted by Slatkin (1977, 1985), colonisation is a form of gene flow, and the founding of populations with a high number of colonists taken from the whole metapopulation can increase homogenisation.
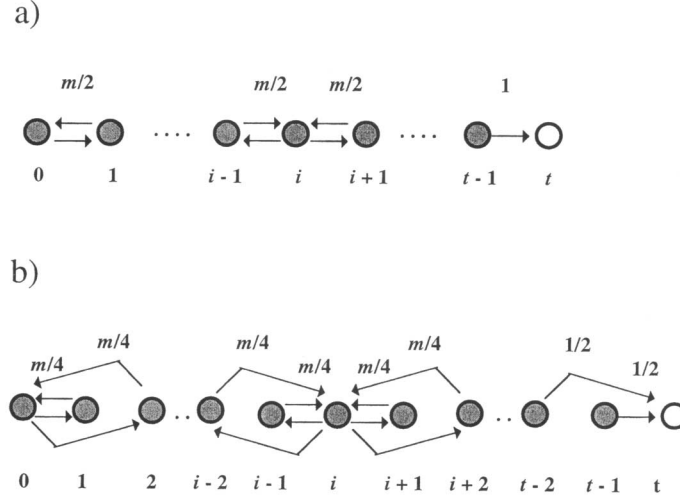
Similarly, the within-population heterozygosity, Ho, gradually increases during colonisation for $k > 2Nm + 1$, whereas it gradually decreases when the opposite condition is met (Figs. 3 and 4). However, significant changes in heterozygosity are observed only within the most recent populations. The loss of diversity is higher in the "outbreak" model (Fig. 3) than in the "linear diffusion" model (Fig. 4), because of the different age-structures: in the "outbreak" model, the pools of colonists and migrants contain higher proportions, of individuals originating from recently founded populations, which allows for a higher cumulative effect of founding events.

3. Colonisation in a linear-stepping stone model

Island metapopulation models have the advantage of their great simplicity, but are not realistic. It is well known that the dispersal ability of individuals is limited, and that short distance movements are usually predominant. Individuals which colonise vacant sites as well as migrants entering populations are more likely to come from some neighbouring populations than to be drawn at random from the entire metapopulation, as is assumed in the island model. More realistic are stepping-stone models, in which individuals move between nearby populations only.

3.1. The model

For the sake of simplicity of analytical derivations, we will consider only the one-dimensional model, in which populations are arranged linearly. Before the beginning of the colonisation (at time $t = 0$), there is only one population and an infinite number of vacant sites, so that the linear stepping-stone array is bounded at only one of its extremities. The probability of i.b.d. within the first population is initially assigned a value of zero. At each generation, a new population is founded from k individuals by colonisation of the vacant site immediately adjacent to the last

a)

m/2   m/2   m/2   1

○→←● ···· ●→←●→←● ···· ●→○
0   1   i-1   i   i+1   t-1   t

b)

          m/4   m/4   m/4   1/2
   m/4 ╱   ╲   ╱   m/4 m/4 ╱   ╲   1/2
●→←● ● · ●→←●→←● ● · ● ●→○
0   1   2   i-2 i-1 i   i+1 i+2   t-2 t-1   t

Fig. 5. Examples of linear stepping-stone models. a: strict stepping-stone model; b: stepping-stone model for ns, the number of populations that produce migrants, equal to 4.

founded population. Populations at time t are numbered from 0 to t, according to their date of founding. From the point of view of the colonisation dynamics, this model is identical to the "linear diffusion" model treated above, with a colonisation rate α equal to 1.

The change with time in the probability of i.b.d. between two alleles randomly drawn in the ith and j^{th} populations in the linear arrangement, resulting from gene flow and random mating, is given by the classical recurrence equation of Malécot (1950):

$$f_{ij}(t) = \sum_l m_{il} m_{jl} \left[\frac{1}{2N} + \left(1 - \frac{1}{2N}\right) f_{ll}(t-1) \right] + \sum_l \sum_{p \neq l} m_{il} m_{jp} f_{lp}(t-1). \quad (8)$$

Where m_{il} is the proportion, in the ith populations, of the migrants coming from the lth populations, with $i = 0 \ldots t-1$, and $l = 0 \ldots t-1$ (since at time t, there are in our model t populations which exchange migrants and 1 newly colonised site).

Similarly, we can define c_{ij}, the proportion in a newly founded population i, of the colonists which come from the jth population. At time t, the probability of i.b.d. between the newly founded population (which is the tth population in the array) and the other populations is

$$f_{tj}(t) = \sum_l c_{tl} c_{jl} \left[\frac{1}{2N} + \left(1 - \frac{1}{2N}\right) f_{ll}(t-1) \right] + \sum_l \sum_{p \neq l} c_{tl} m_{jp} f_{lp}(t-1) \quad (9)$$

and the probability of i.b.d. within the newly founded population is
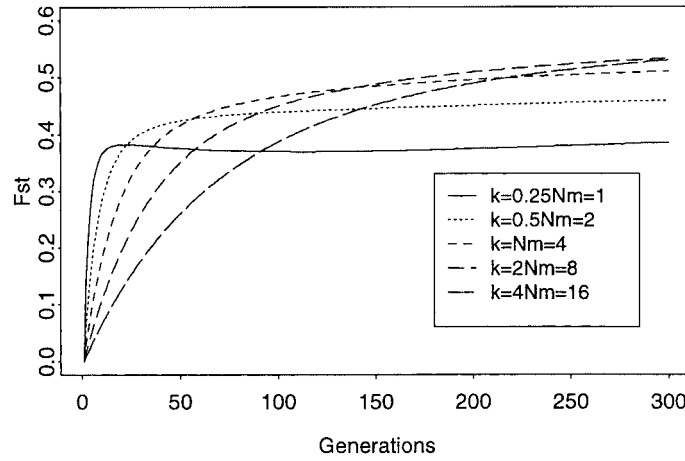
Fig. 6. Change over time in the genetic differentiation among colonised sites in the strict linear stepping-stone model for different value of k, the number of colonists. Other parameters are $N = 100$ and $m = 0.04$

$$f_{ll}(t) = \frac{1}{2k} + \left(1 - \frac{1}{2k}\right)\left\{\sum_l c_{ll}^2 \left[\frac{1}{2N} + \left(1 - \frac{1}{2N}\right)f_{ll}(t-1)\right]\right.$$

$$\left. + \sum_l \sum_{p \neq l} c_{ll}c_{lp}f_{lp}(t-1)\right\}$$ (10)

All those equations may be written in a matrix form, so that the model runs by recursively constructing the matrix $F(t)$ which contains all $f_{ij}(t)$ values (see Appendix). The genetic differentiation at time t is calculated as $(f_0(t) - f(t))/(1 - f(t))$,
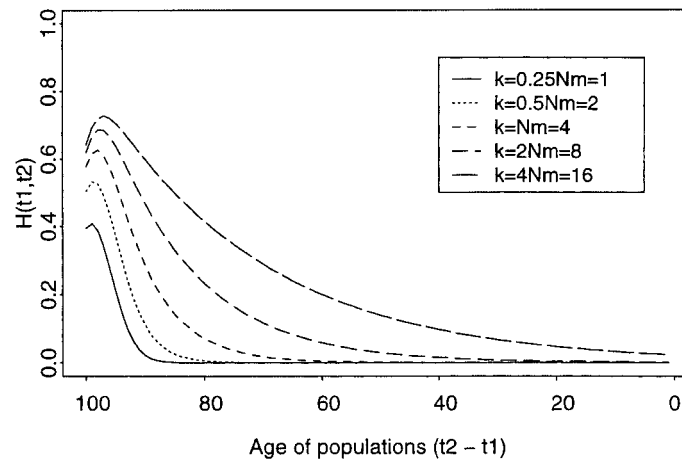
Fig. 7. Within-population heterozygosity after 100 generations of colonisation as a function of populations' age in the strict linear stepping-stone model for different value of k, the number of colonists. Other parameters are $n = 100$ and $m = 0.04$.
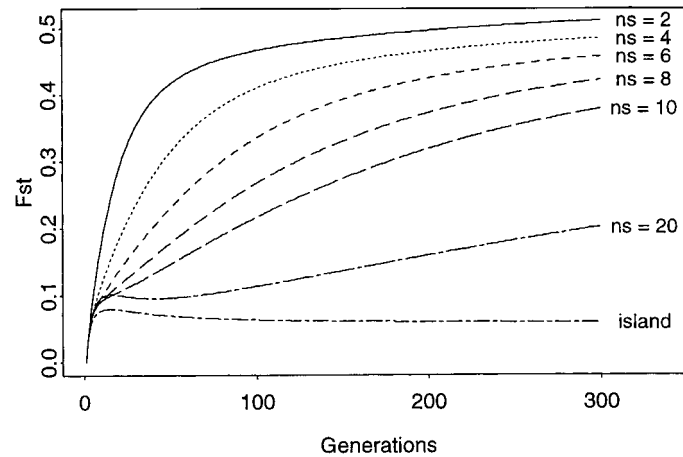
Fig. 8. Effect of dispersal distances in the linear stepping-stone model: the change over time in genetic differentiation is shown for migrants and colonists being sampled among increasing numbers *ns* of populations. Parameters are $N = 100$ and $k = Nm = 4$.

where $f(t)$ is the mean over all elements of the matrix $F(t)$ and $f_0(t)$ is the mean over its diagonal elements. Within-population heterozygosities at time t are the opposites of the diagonal elements of the matrix $F(t)$.

3.2. Effect of the relationship between k and Nm in a strict stepping-stone model

The model was initially run for the strict stepping stone case, i.e. when individuals move between adjacent populations only, so that each population receives migrants from its two neighbours, and each site is colonised by individuals coming from the last founded population (Fig. 5a). For small numbers of colonists, the loss in genetic variability is very rapid, and most populations, except the populations located near the starting point of colonisation, are fixed. Figure 7 shows that the time until fixation is approximately a linear function of the ratio k/Nm. The increase in F_{st} (Fig. 6) is also very rapid, until new populations become fixed.

In the strict-stepping-stone model, the cumulative effect of founding events is very important, since each vacant site is colonised by individuals coming from the last founded population, and gene flow is among populations very close in age. This cumulative effect will act until all variability is lost, which explains why the F_{st} increases more gradually but during a longer time when the number k of colonists is higher. In the island model, cumulative effects of founding events were much more reduced, first because of the equal participation of all populations in the formation of colonists, second because of a greatest homogenisation by gene flow.
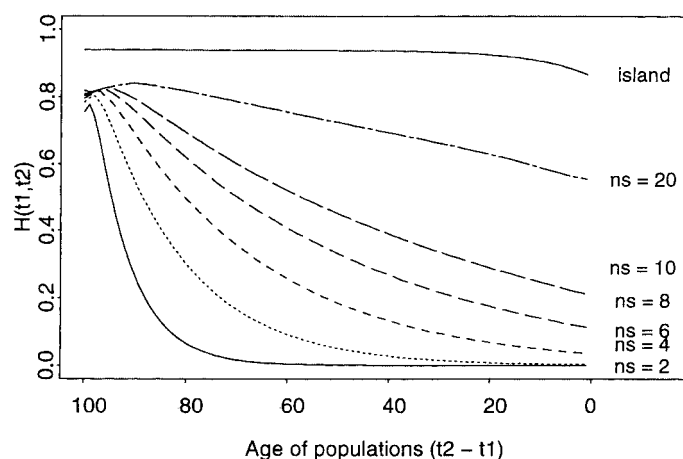
3.3. Effect of dispersal distances

It is worth noting that the island model may be considered as a particular case of the stepping stone model, in which each population would receive an equal number of migrants from all other populations. This is achieved when the migration rates among the t populations existing at each time t are equal to:

$$m_{ii} = 1 - m \quad \text{for} \quad i = 0 \ldots t - 1$$

$$m_{ij} = m/t \quad \text{for} \quad i = 0 \ldots t - 1, \quad j = 0 \ldots t - 1 \quad \text{and} \quad i \neq j$$

In real populations, dispersal will probably be intermediate between the strict stepping-stone and the island models. The effect of the distance at which individuals can move may be tested by modifying the number of neighbouring populations from which each site can receive migrants or colonists. In the strict stepping stone model, this number was 2 (except for the sites at the borders of the metapopulation, for which it is 1). Other numbers were tested, always assuming a symmetrical distribution and identical contributions of the source populations, and that sites at the borders receive less individuals than the others (Fig. 5). The main effect of large dispersal distances is to slow down both the increase in genetic differentiation and the loss of diversity, by preventing a strong cumulative effect of founder events (Figs. 8 and 9). However, it is only when moving individuals are sampled among more than about twenty populations that the stepping-stone model approximately meets the island model.

Fig. 9. Effect of dispersal distances in the stepping-stone model: values of within populations heterozygosities after 100 generations of colonisation are shown for migrants and colonists being sampled among increasing numbers ns of populations. Parameters are $N = 100$ and $k = Nm = 4$.

4. Discussion

4.1. Cumulative effect of founding events during colonisation

The genetic effects of one founding event have been extensively studied, mainly in the perspective of speciation processes (Barton and Charlesworth, 1984; Carson and Templeton, 1984; Templeton, 1980), and were shown to vary according to the number of founding individuals, their total genetic diversity, and the rapidity of the increase in population size after founding (Nei et al., 1975). At the metapopulation level, the genetic consequences of extinction and recolonisation depend on the relationship between the number k of colonists founding new populations and the number Nm of migrants exchanged among extant populations, and the genetic differentiation F_{st} is increased whenever $k < 1 + 2Nm$ (Wade and McCauley, 1988; Whitlock and McCauley, 1990). In the present study, the same condition was found to hold for the direction of departure from equilibrium of the genetic differentiation in an island metapopulation increasing its size by colonisation in the absence of extinctions. Our colonisation models, however, especially emphasise the cumulative effect of founding events during a continuous range expansion. This cumulative effect depends primarily on the strength of each founding event, i.e. on the number of colonists, and on the counterbalancing action of gene flow, i.e. on the number of migrants. But it is also determined by the relative contributions of populations of different ages to the constitution of both the migrant and colonist pools. The distance of dispersal, determining how many populations produce migrants and colonists, and the dynamics of colonisation, determining the age-structure among populations, are therefore important factors. The cumulative effect of founding events is particularly important when the sampling of colonists occurs from recently founded populations and gene flow is restricted to populations of similar ages. Such a cumulative effect was previously suggested by Boileau and Hébert (1991) to explain the loss of diversity which accompanied the post-glacial dispersal of the pond-dwelling copepod *Heterocope septentrionalis*, although its high reproduction rate tends to maintain high levels of diversity at each founding event. It is also noticeable that the effective size of the population bottleneck for introduced species is often estimated from the present genetic diversity by considering only the sampling of colonists at the first founder generation. Such estimates may be biased downwards, because they do not include the cumulative effect of further founding events during the colonisation of the new territory (Eastal, 1985; Baker and Moeed, 1987).

4.2. Island versus stepping-stone models

Among models of metapopulation structure, the island model represents the extreme in long distance gene flow, whereas the strict stepping-stone model is the extreme in short distance gene flow. It is also noteworthy that more divergence among populations is maintained in one-dimensional habitats compared to two-

dimensional ones (Crow and Aoki, 1984), so that the one dimensional stepping-stone model considered here for simplicity is also an extreme case for the genetic structure among populations.

The genetic consequences of colonisation were shown to differ dramatically between the island and linear stepping-stone models. The strict linear stepping-stone model resulted in the strongest cumulative effect of founder events, whereas this cumulative effect is weakest in the island model. Real populations will certainly fall between the two extremes. The island model may apply to colonisation events at a local scale (see e.g. McCauley et al., 1995), or to initial phases of range expansion, and especially in the case of species having a high dispersal ability. Long-range expansions over larger areas, like the last post-glacial recolonisations, will probably be more realistically represented via one- or two-dimensional stepping-stone models.

4.3. Results from experimental studies

Studies or recent invasions have often shown an increase in genetic differentiation after colonisation, as for the land snail *Theba pisana* on Rottnest island, Australia (Johnson, 1988), or for the walnut husk fly, (*Rhagoletis completae*) in California (Berlocher, 1984). A decrease in F_{st} values after colonisation was rarely described, except for the fresh water snail *Biomphalaria straminea*, which has been rapidly expanding in Hong Kong since its introduction in the early 1970s (Woodruff et al., 1985). This decrease, in agreement with our island model, may be a consequence of colonisation events implying egg masses that each contained up to 30 individuals, and of gene flow over long distances made possible by human and animal dispersal.

In contrast to these recent events, only minor changes in variability are presently found among populations of animal species (mostly birds) introduced by man in New Zealand, Australia and various oceanic islands at the end of the nineteenth century (Baker, 1992; Eastal, 1985; Merilä et al., 1996; Parkin and Cole, 1985; Ross, 1983; Taylor and Gorman, 1975), certainly because a longer period has elapsed since colonisation and because these studies concern animals with high dispersal abilities.

Colonisation processes in plants are more complex than in animals, because colonisation occurs through seeds, whereas gene flow is mainly through pollen, and the mechanisms of dispersal may be quite different for the two kinds of propagules. It may be noted, however, that the effect of a discrepancy between the numbers of migrants and colonists, but also between their dispersal distances may easily be introduced in the stepping-stone model. Geographical patterns of diversity resulting from colonisation are likely to be found in temperate forest tree species, because of their wide present natural ranges resulting from broad-scale post-glacial recolonisations (Strauss et al., 1992). However, most of these species show a very low genetic differentiation for nuclear genes, probably because of extensive gene flow by pollen (Hamrick and Godt, 1990). No discernible patterns of change in the intrapopulation diversity were shown to occur in relation to past migration routes for the

European beech *Fagus grandifolia* in Italy (Leonardi and Menozzi, 1995), nor for
the lodgepole pine *Pinus taeda* in western America (Cwynar and McDonald, 1987).
Yet, for the more genetically differentiated coastal variety of Douglas-fir (*Pseudot-
suga menziesii*) in western America, within population diversity was shown to be
highest in the supposed region of the glacial refuge, and to decrease northwards and
southwards (Li and Adams, 1989).

4.4. Applicability of the models

The applicability of colonisation models may be limited by the complexity of
colonisation dynamics in natural populations, which, for example, may not proceed
at a constant rate (Hengeveld, 1989). The island metapopulation model, which has
no spatial structure, may paradoxically be successfully used for reconstructing
various colonisation dynamics, because colonisation is a time process as well as a
spatial process. Stepping-stone models are more realistic from the point of view of
dispersal distances, but will rapidly become intractable if complex colonisation
dynamics are to be taken into account. Simulations may in this case be an
interesting alternative (Ibrahim et al., 1996; Le Corre et al., 1997).

Another difficulty is that colonisation histories are often poorly known, so that
ages of populations are difficult to estimate. For forest trees, population ages may
be estimated through the use of radiocarbon-dated pollen fossil records (Cwynar
and McDonald, 1987). Molecular genetic analyses may also be suitable for the
reconstruction of past colonisation dynamics, particularly for low dispersal species
and for rapidly mutating, non-recombinant genetic markers such as animal mito-
chondrial DNA (Thorpe, 1984). Because of their rapid evolution and high number
of alleles, microsatellites may also be useful for assessing an age-structure among
populations (Bowcock et al., 1994; Goldstein et al., 1995), although their particular
mode of evolution necessitates cautious interpretations. In combination with theo-
retical models such as those presented here, a better knowledge of the age structure
of populations will allow the comparison of observed patterns of genetic diversity
with the patterns expected due to colonisation. This will help to distinguish between
the effect of founder events and the influence of other factors such as differentiating
selection or isolation by distance.

References

Austerlitz, F., B. Jung-Muller, B. Godelle and P. H. Gouyon. 1997. Evolution of coalescence times,
     genetic diversity and structure during colonization. Theor. Pop. Biol. 51: 148–164.
Baker, A. J. 1992. Genetic and morphometric divergence in ancestral European and descendent New
     Zealand populations of chaffinches (*Fringilla coelebs*). Evolution 46: 1784–1800.
Baker, A. J. and A. Moeed. 1987. Rapid genetic differentiation and founder effect in colonisation
     populations of common mynas (*Acridotheres tristis*). Evolution 41: 525 538.
Barton, N. H. and B. Charlesworth. 1984. Genetic revolutions, founder effects, and speciation. Ann.
     Rev. Ecol. System. 15: 133–164.

Bascompte, J. and R. V. Solé. 1995. Rethinking complexity: modelling spatiotemporal dynamics in ecology. Trends Ecol. Evol. 10: 361–366.

Berlocher, S. H. 1984. Genetic changes coinciding with the colonisation of California by the walnut husk fly, *Rhagoletis completa*. Evolution 38: 906–918.

Boileau, M. G., P. D. N. Hébert and S. S. Schwartz. 1992. Non-equilibrium gene frequency divergence: persistent founder effects in natural populations. J. evol. biol. 5: 25–39.

Boileau, M. G. and P. D. N. Hébert. 1991. Genetic consequences of passive dispersal in pond-dwelling copepods. Evolution 45: 721–733.

Bowcock, A. M., A. Luiz-Linares, J. Tomfohrde, E. Minch, J. R. Kidd and L. L. Cavalli-Sforza. 1994. High resolution of human evolutionary trees with polymorphic microsatellites. Nature 368: 455–457.

Carson, H. L. and A. P. Templeton. 1984. Genetic revolutions in relation to speciation phenomena: the founding of new populations. Ann. Rev. Ecol. Syst. 15: 97–131.

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

Cwynar, L. C. and G. M. McDonald. 1987. Geographical variation of lodgepole pine in relation to population history. Am. Nat. 129: 463–469.

Easteal, S. 1985. The ecological genetics of introduced populations of the giant toad *Bufo marinus*. II. Effective population size. Genetics 110: 107–122.

Goldstein, D. B., A. Ruiz Linares, L. L. Cavalli-Sforza and M. W. Feldman. 1995. Genetic absolute dating based on microsatellites and the origin of modern humans. Proc. Natl. Acad. Sci. U.S.A. 92: 6723–6727.

Hamrick, J. L. and M. J. Godt. 1990. Allozyme diversity in plant species, pp 23–42. *In* A. H. D. Brown, M. T. Clegg, A. L. Kahler and B. S. Weir (Eds.), Plant Population Genetics, Breeding and Genetic Resources. Sinauer Associates. Sunderland, Massachusetts.

Hengeveld, R. 1989. Dynamics of Biological Invasions. Chapman and Hall. London.

Ibrahim, K., R. A. Nichols and G. M. Hewitt. 1996. Spatial patterns of genetic variation generated by different forms of dispersal during range expansion. Heredity 77: 282–291.

Johnson, M. S. 1988. founder effects and geographic variaton in the land snail *Theba pisana*. Heredity 61: 133–142.

Le Corre, V., N. Machon, R. J. Petit and A. Kremer. 1997. Colonisation with long distance seed dispersal and genetic structure of maternally inherited genes in forest trees: a simulation study. Genetical Research 69: 117–125.

Leonardi, S. and P. Menozzi. 1995. Genetic variability of *Fagus sylvatica* L. in Italy: the role of postglacial recolonization. Heredity 75: 35–44.

Li, P. and W. T. Adams. 1989. Range-wide patterns of allozyme variation in Douglas-fir (*Pseudotsuga menziesii*). Can. J. For. Res. 19: 149–161.

McCauley, D. E., J. Raveill and J. Antonovics. 1995. Local founding events as determinants of genetic structure in a plant metapopulation. Heredity 75: 630–636.

Merilä, J., M. Björklund and A. J. Baker. 1996. The successful founder: genetics of introduced *Carduelis chloris* (greenfinch) populations in New Zealand. Heredity 77: 410–422.

Nei, M., T. Maruyama and P. Chakraborty. 1975. The bottleneck effect and genetic variability in populations. Evolution 29: 1–10.

Parkin, D. T. and S. R. Cole. 1985. Genetic differentiation and rates of evolution in some introduced populations of the House Sparrow, *Passer domesticus* in Australia and New Zealand. Heredity 54: 15–23.

Ross, H. A. 1983. Genetic differentiation of starling (*Sturnus vulgaris*: Aves) populations in New Zealand and Great Britain. J. Zool. Lond. 201: 351–362.

Slatkin, M. 1977. Gene flow and genetic drift in a species subject to frequent local extinctions. Theor. Pop. Biol. 12: 253–262.

Strauss, S. H., J. Bousquet, V. D. Hipkins and Y. P. Hong. 1992. Biochemical and molecular genetic markers in biosystematic studies of forest trees. New Forests 6: 125–158.

Taylor, C. E. and G. C. Gorman. 1975. Population genetics of a "colonising" lizard: natural selection for allozyme morphs in *Anolis grahami*. Heredity 35: 241–247.

Templeton, A. P. 1980. The theory of speciation via the founder principle. Genetics 94: 1011–1038.

Thorpe, R. S. 1984. Primary and secondary transition zones in speciation and population differentiation: a phylogenetic analysis of range expansion. Evolution 38: 233–243.

Wade, M. J., M. L. McKnight and H. B. Shaffer. 1994. The effects of kin-structured colonisation on nuclear and cytoplasmic genetic diversity. Evolution 48: 1114–1120.

Wade, M. J. and D. E. McCauley. 1988. Extinction and recolonization: their effects on the genetic differentiation of local populations. Evolution 42: 995–1005.

Whitlock, M. C. and D. E. McCauley. 1990. Some population genetic consequences of colony formation and extinction: genetic correlation within founding groups. Evolution 44: 1717–1724.

Woodruff, D. S., M. Mulvey and M. W. Yipp. 1985. Population genetics of *Biomphalaria straminea* in Hong Kong. J. Hered. 76: 355–360.

Wright, S. 1940. Breeding structure of populations in relation to speciation. Am. Nat. 74: 232–248.

Appendix

Matrix recurrence equations for the stepping-stone model

At time t, the model is composed of t established populations (numbered 0 to $t - 1$) and of one newly colonised site which is numbered t. The probability of identity between population i and population j is noted $f_{ij}(t)$. Recurrence equations (8), (9) and (10) can be written under matrix form using the following notations:

$F(t - 1)$ is the $(t \times t)$ matrix containing all $f_{ij}(t - 1)$ elements for $i = 0 \ldots t - 1$ and $j = 0 \ldots t - 1$, and $F'(t - 1)$ is its value after migration;

$Fo(t - 1)$ is the $(t \times t)$ matrix containing the $f_{ii}(t - 1)$ elemens on its diagonal, and zeros elsewhere;

F_t is the row vector of length t containing the $f_{tj}(t)$ elements for $j = 0 \ldots t - 1$;

F_{tt} is the element $f_{tt}(t)$;

$M(t - 1)$ is the $(t \times t)$ matrix containing all m_{ij} coefficients for $i = 0 \ldots t - 1$ and $j = 0 \ldots t - 1$;

$C(t - 1)$ is the row vector of length (t) containing the c_{tj} coefficients for $j = 0 \ldots t - 1$;

Colonisation is modelled using the matrix forms of equations (9) and (10):

$$F_t = 1/2N \ C(t - 1) \cdot (I - Fo(t - 1)) \cdot M(t - 1)^T + C(t - 1) \cdot F(t - 1) \cdot M(t - 1)^T$$

$$F_{tt} = 1/2N \ C(t - 1) \cdot (I - Fo(t - 1)) \cdot C(t - 1)^T + C(t - 1) \cdot F(t - 1) \cdot C(t - 1)^T$$

Migration is modelled using the matrix form of equation (8):

$$F'(t - 1) = 1/2N \ M(t - 1) \cdot (I - Fo(t - 1)) \cdot M(t - 1)^T + M(t - 1) \cdot F(t - 1) \cdot M(t - 1)^T$$

The matrix $F(t)$ is constructed by adding the row F_t, the column Ft^T and the last diagonal element F_{tt} to the matrix $F'(t - 1)$.