npg

# Effective size in simple metapopulation models

F Rousset

*Laboratoire Génétique et Environnement, Institut des Sciences de l'Évolution, CC065, USTL, Place E. Bataillon, 34095 Montpellier Cedex 05, France*

A coalescent argument is used to derive the effective size in simple models with recurrent local extinctions. Several alternative methods of derivation of this result are given and compared to earlier analyses of this problem. The different methods described in this paper all give the same result, which differs from earlier ones. For two published sets of estimates of demographic parameters, metapopulation structure appears to result in a moderate reduction of effective size relative to total adult population size.
*Heredity* (2003) **91,** 107–111. doi:10.1038/sj.hdy.6800286

## Introduction

To determine the effects of local extinctions and recolonizations on genetic diversity and effective size, Slatkin (1977) defined two models, the 'migrant pool' and 'propagule pool' models. In the former, recolonizers come from the whole metapopulation; in the latter, they preferentially come from a single population. These models were subsequently investigated by a number of works (eg Wade and McCauley, 1988; Whitlock and McCauley, 1990; Whitlock and Barton, 1997; Pannell and Charlesworth, 1999). However, there are several inaccuracies in the methods previously used to compute effective size in these models. This short paper discusses several methods for computing effective size. All of them provide a single expression (equation (7)), which has not been previously given. First, a simple coalescent argument for this result is provided. Then, alternative derivations are used to explain discrepancies with earlier works. The quantitative importance of these discrepancies is briefly discussed. A Mathematica (Wolfram, 1999) notebook performing the computations described in this paper is available on request.

## A simple coalescent argument

Effective size is defined to give the asymptotic rate of coalescence of pairs of genes. For pairs of genes in different demes, this rate can be deduced by a two-step argument. First, the two ancestral lineages must gather in the same deme. Then, they may coalesce or separate again in different demes.

### No local extinctions

In the island model without local extinctions, this argument develops as follows. Let $m$ be the dispersal probability. The probability that two genes in different demes come from a single deme in the previous generation is

$$\frac{1 - (1 - m)^2}{n_d} + O(1/n_d{}^2) \qquad (1)$$

where $n_d$ is the number of demes. Thus, on a timescale of $n_d$ generations, the rate at which genes in different demes come in a single deme is $1-(1-m)^2$. Then, the ancestral lineages coalesce immediately on this time scale if they have the same parent (probability: $1/N$, where $N$ is the number of adults, here haploid, per deme), or if they coalesce in this deme in a recent past, rather than separate in different demes again. The probability of the latter event may be written in terms of Wright's $F_{ST}$ measure of population structure, as $(1-1/N)F_{ST}$, since in the island model, $F_{ST}$ is approximately (to $O(1/n_d)$) the probability of recent coalescence of genes within demes (Hudson, 1998; Rousset, 2002). Then, the overall rate of coalescence is

$$(1 - (1 - m)^2)\left(\frac{1}{N} + \frac{N - 1}{N}F_{ST}\right) \qquad (2)$$

per $n_d$ generations. For Wright's island model,

$$F_{ST} = (1 - m)^2/((1 - m)^2 + N(1 - (1 - m)^2)),$$

and the whole expression can be written as $(1-F_{ST})/N$. Hence, the effective size is $(1-F_{ST})/(Nn_d)$ as already given by Wright (1943).

This derivation of effective size is in line with other computations of effective size in 'structured coalescents with two time scales' (eg Nordborg, 2001; Wakeley and Aliacar, 2001; Nordborg and Krone, 2002). Here the separation of time scales is obtained for a large number of demes, and the expression for effective size is correct to leading order in $1/n_d$. Wakeley and Aliacar (2001) have already used such an approach (together with additional approximations) to obtain an expression for the effective size of a metapopulation. Their result is consistent with the more general one presented below.

Using a time scale of $n_d$ generations to derive the effective size is in line with such previous arguments. Alternatively, we can compute the probability that two

*Correspondence: F Rousset, Laboratoire Génétique et Environnement, Institut des Sciences de l'Évolution, CC065, USTL, Place E. Bataillon, 34095 Montpellier Cedex 05, France*
*E-mail: rousset@isem.univ-montp2.fr*

108

lineages coalesce in a given 'current' generation as the sum over $t$ of probabilities that (1) the two ancestral lineages gathered in the same deme $t$ generations earlier (probability $(1-(1-m)^2)/n_d$), and (2) given $t$, they coalesce in the current generation rather than separate again in different demes. $1/N+(N-1)F_{ST}/N$ then appears as the sum over $t$ of probabilities of the second event.

### Local extinctions

With local extinctions, the two demes where the genes are sampled may have both become extinct in the previous generation (probability $e^2$), or neither became extinct (probability $(1-e)^2$), or only one did (probability $2e(1-e)$). If neither became extinct, the probability that the two lineages come from a single deme in the previous generation is

$$\frac{1-(1-m)^2}{(1-e)n_d} + O(1/n_d^2) \tag{3}$$

where $(1-e)n_d$ is the number of parental demes that contribute to the next generation. If one or both demes became extinct, the probability is simply $1/((1-e)n_d)$.

Taking the different cases into account, the overall rate of coalescence is

$$\frac{1}{N_e} \sim \frac{1}{n_d}\left((1-e)(1-(1-m)^2)+2e+\frac{e^2}{1-e}\right)$$
$$\times \left(\frac{1}{N}+\frac{N-1}{N}F_{ST}\right) \tag{4}$$

where $\sim$ means that the computation is correct to leading order in $1/n_d$. This can be written as

$$\frac{1}{N_e} \sim \frac{1}{n_d}\frac{1-(1-m)^2(1-e)^2}{1-e}Q^R \tag{5}$$

where $Q^R \equiv 1/N+(N-1)F_{ST}/N$ is the 'identity by descent' among gametes produced by adults within a deme, that is the relatedness of such gametes relative to gametes produced in different demes. $Q^R$ can be computed, for example, as in Wade and McCauley (1988), or can be deduced from the recursions detailed below. One then obtains that for an infinite number of demes

$$Q^R = \frac{1/N+e/k-e/kN}{1-(1-1/N)[(1-m)^2(1-e)+e\phi(1-1/k)]} \tag{6}$$

where $k$ is the number of recolonizers, and where $\phi=0$ for the migrant pool model and $\phi=(1-m)^2$ for the propagule pool model. Hence,

$$\frac{1}{N_e} \sim \frac{1}{n_d}$$
$$\frac{(1-(1-m)^2(1-e)^2)(1/N+e/k-e/(kN))}{(1-e)(1-(1-1/N)[(1-m)^2(1-e)+e\phi(1-1/k)])}. \tag{7}$$

This result is consistent with approximation (23) derived by Wakeley annd Aliacar (2001) for $\phi=0$, large $N$, and small $m$ and $e$. Further, equation (7) is also obtained by the following argument, which is longer but allows step-by-step comparison with earlier analyses of Pannell and Charlesworth (1999).

## Matrix formulation

The asymptotic rate of coalescence is obtained as $(1-\lambda)^{-1}$, where $\lambda$ is the largest eigenvalue of the matrix which describes the decrease of gene diversities in the absence of mutation (eg Hill, 1972; Ewens, 1982; Whitlock and Barton, 1997). To construct this matrix, we first consider a system of recursions for probabilities of identity within and among demes, comparable to those of Slatkin (1977) and Pannell and Charlesworth (1999). These recursions include mutation as those of Slatkin, but of course they describe the decrease of gene diversities when mutation rate is set to zero.

The life cycle considered in these models is as follows (Slatkin, 1977). In the absence of extinction, events occur in the following order: gamete production, dispersal, and population regulation, where $N$ adult offspring survive. In each generation any deme can independently become extinct with some probability $e$. Extinction occurs before reproduction, so that the adults do not contribute anything to the next generation. An extinct deme is immediately recolonized by $k$ colonizers, which reproduce immediately so that there will always be $N$ adults in the next generations in all demes, recolonized or not. Thus, recolonizers experience two rounds of dispersal and reproduction within one 'generation', that is within the time only one round is considered for demes that do not become extinct. This assumption was thought to simplify the computation of $F_{ST}$, but can be relaxed (see Discussion). In the 'migrant pool' model, colonizers in the propagules are independently sampled from all other demes. In the 'propagule pool' model, propagules of $k$ colonizers are formed in nonextinct demes after gamete production and dispersal, and each extinct deme is recolonized by the $k$ members of a single propagule. As noticed by Pannell and Charlesworth (1999), it is actually not required that the extinct and colonized habitats are the same: it is only assumed that a constant number of demes become extinct and that an equal number of habitats are colonized in each generation. For conciseness, I consider $N$ and $k$ where previous authors considered $2N$ and $2k$ genes. Let $Q_i$ be the probability of identity of genes from different adults, within demes ($Q_1$) and in different demes ($Q_2$). Here we consider the probability of allelic identity in the infinite allele model. Let $\mu$ be the mutation rate. We write the recursions for next-generation identities $Q_1'$ and $Q_2'$ as

$$Q_1' = (1-\mu)^2\left((1-e)\left(a\left(Q_1+\frac{1-Q_1}{N}\right)\right.\right.$$
$$\left.\left.+(1-a)Q_2\right)+eA\right) \tag{8}$$

$$Q_2' = (1-\mu)^2\left((1-e)\left(1-\frac{en_d}{n_d-1}\right)\left(b\left(Q_1+\frac{1-Q_1}{N}\right)\right.\right.$$
$$\left.\left.+(1-b)Q_2\right)+CB_1\right) \tag{9}$$

where

$$a \equiv (1-m)^2+\frac{m^2}{(1-e)n_d-1}, \tag{10}$$

$$b \equiv \frac{1-a}{(1-e)n_d - 1}, \tag{11}$$

$$A \equiv \frac{1}{k} + \left(1 - \frac{1}{k}\right)\left(\phi\left(Q_1 + \frac{1-Q_1}{k}\right)\right.$$
$$\left. + (1-\phi)Q_2\right), \tag{12}$$

$$B_1 \equiv \frac{1}{(1-e)n_d}\left(Q_1 + \frac{1-Q_1}{N}\right)$$
$$+ \left(1 - \frac{1}{(1-e)n_d}\right)Q_2, \tag{13}$$

$$C \equiv \frac{2n_d e(1-e)}{n_d - 1} + \frac{e(n_d e - 1)}{n_d - 1}. \tag{14}$$

Here:

- Equations (8) and (9) are as in Slatkin (1977) and Pannell and Charlesworth (1999) (except for obvious typos in the latter). For $n_d \to \infty$, equilibrium $F_{ST}$ is the solution $Q_1$ of the recursion deduced from equation (8) with $Q_2 = 0$:

$$Q' = (1-e)(1-m)^2 Q^R + e$$
$$\times \left(\frac{1}{k} + \left(1 - \frac{1}{k}\right)\phi Q^R\right). \tag{15}$$

This yields equation (6).
- $C$ is a shorthand for a term already considered by these authors.
- Equation (10) is modified from Slatkin (1977) as suggested by Pannell and Charlesworth (1999).
- Equation (10) is modified from Pannell and Charlesworth's equation (A3) so as to be consistent with the exact recursions in Nagylaki (1983) for $e = 0$. Nevertheless, this difference does not affect results for $F_{ST}$ nor for effective size to leading order in $1/n_d$, because equation (11) and their equation (A3) are identical to first order in $1/n_d$: they differ only by terms of order $1/n_d^2$. Likewise, the $O(1/n_d)$ term in equation (10) does not affect $F_{ST}$ nor effective size to leading order in $1/n_d$. Thus, this term, which represents the probability that two genes immigrating in the same deme come from a common parental deme, can also be neglected in later analyses.
- In equation (12), I use the notation $\phi$ for the probability that two colonizers have parent(s) in the same deme. If propagule pools are formed after dispersal of gametes (consistently with Slatkin's verbal description of the life cycle), then $\phi = (1-m)^2 + O(1/n_d)$. If propagule pools are formed from locally produced gametes, then $\phi = 1$. In the migrant pool model, $\phi = O(1/n_d)$. It may be checked that the $O(1/n_d)$ terms in $\phi$ have no bearing on the results, and they will be ignored below.

There are some inconsistencies among different analyses of the propagule pool model. Slatkin analyzed two different scenarios: 'model I' and 'model II' (respectively, an islands-mainland model without dispersal between the islands, and a finite island model with migration). As noticed by Wade and McCauley (1988), $F_{ST}$ should be the same in both models when $n_d \to \infty$. For the propagule pool model,

equation (15) of Wade and McCauley (1988) is consistent with equation (6) of Slatkin (1977); they imply $\phi = (1-m)^2$. But Slatkin's and Pannell and Charlesworth's systems of recursions for model II yield different results; they imply $\phi = 1$. Of course, both cases can be considered, provided they are distinguished.

The probability of common origin of colonizers $\phi$ was first considered by Whitlock and McCauley (1990), but I was unable to match their equations with the above ones. In particular, their equation (4) for identity among gametes implies that there are two successive reproductions at recolonization if $\phi = 1$ (as seen from terms of order $1/(kN)$ in the recursions, and consistently with Slatkin's description of the life cycle), while there is only one reproduction at recolonization if $\phi = 1$ (only terms of order $1/k$ appear, so gametes are not produced by $N$ adults at some stage). Thus, their equations may not correspond to a well-defined life cycle.
- The main discrepancy with earlier expressions is in equation (13), which has the factor $Q_1 + (1-Q_1)/N$ instead of $Q_1$ in Slatkin (1977) and Pannell and Charlesworth (1999). This difference takes into account that when genes from different demes originate from parent(s) in a single deme, they may actually have the same parent and coalesce. This does not affect $F_{ST}$ values, but it does affect $N_e$ values.

From the above equations, one can derive the probabilities $g_{ij}$ that a pair of genes of type $i$ ($i = 1$ for pairs within a deme and 2 for pairs in different demes) derives without coalescence from a pair of genes of type $j$. They are

$$g_{11} = \left(1 - \frac{1}{N}\right)\left((1-e)a + e\left(1 - \frac{1}{k}\right)\phi\right), \tag{16}$$

$$g_{12} = (1-e)(1-a) + e\left(1 - \frac{1}{k}\right)(1-\phi), \tag{17}$$

$$g_{21} = \left(1 - \frac{1}{N}\right)\left((1-e)\left(1 - \frac{en_d}{n_d - 1}\right)b\right.$$
$$\left. + C\frac{1}{(1-e)n_d}\right), \tag{18}$$

$$g_{22} = (1-e)\left(1 - \frac{en_d}{n_d - 1}\right)(1-b)$$
$$+ C\left(1 - \frac{1}{(1-e)n_d}\right). \tag{19}$$

$N_e$ is obtained as $(1-\lambda)^{-1}$, where $\lambda$ is the largest eigenvalue of $\mathbf{G} \equiv (g_{ij})$. To leading order in $1/n_d$, the expression for $N_e$ can be deduced from a perturbation approximation for $\lambda$ near the limit $n_d \to \infty$. Classical expressions for perturbation approximations (eg Horn and Johnson, 1985; Charlesworth, 1994; Caswell, 2001) here take the form

$$\frac{1}{N_e} \sim -d\lambda = -\frac{\mathbf{y}(d\mathbf{G})\mathbf{e}}{\mathbf{y} \cdot \mathbf{e}} \tag{20}$$

where $\mathbf{y}$ and $\mathbf{e}$ are left and right eigenvectors associated with the largest eigenvalue (here 1) or the unperturbed

matrix (here $\lim_{n_d \to \infty} G$); and d$\mathbf{G}$ are the terms of order $1/n_d$ in $\mathbf{G} - \lim_{n_d \to \infty} \mathbf{G}$.

By this method, one obtains again equation (7). It differs from the result implied by the same method but with Pannell and Charlesworth's system of recursions, as expected from the difference in the definition of $B_1$ (equation (13)). Differences between the rate of coalescence predicted from the above recursions and from Pannell and Charlesworth's ones are only of order $e(e+m)$, and the maximum differences I have found numerically were by factors of $\approx 1.5$ to $2.2$ for $k = N$, large $m$, and $e$ from $0.25$ to $0.7$ (details not shown). Approximations comparable to those of their Table 2 for total diversity $\pi_T$ can be derived from equation (7). It appears that a factor $(1 + Ne/k)$ is missing from the denominator of their central and right-hand approximations for $\pi_T$ when $e \leqslant m$. This may imply some reductions in effective size.

## Alternative coalescent arguments

Alternative derivations of effective size also allow comparison with Whitlock and Barton's methods. First, the rate of coalescence may be obtained as the probability $\xi$ that two different lineages that have not already coalesced are in gametes from the same deme, times the probability $\pi_c$ that two gametes from one deme coalesce within one generation. Here, distinguishing whether the deme has just been recolonized or not,

$$\pi_c = (1-e)\frac{1}{N} + e\left(\frac{1}{N} + \frac{N-1}{N}\frac{1}{k}\right)$$
$$= \frac{1}{N} + e\frac{N-1}{N}\frac{1}{k} \tag{21}$$

and $\xi$ is the equilibrium solution of the recursion

$$\xi \sim \xi'(1 - 1/N)[(1-m)^2(1-e) + e\phi(1 - 1/k)]$$
$$+ \frac{1}{n_d}\frac{1 - (1-m)^2(1-e)^2}{1-e}. \tag{22}$$

Here $\xi'$ is the probability $\xi$ considered one generation later, and the factor of $\xi'$ is the probability that two gametes produced in the same deme originate from two different gametes produced in a single deme one generation before (compare with the denominator of $Q^R$ in equation (6). The remainder is the probability that genes in different demes come from gametes in the same deme (see equations (4) and (5)). It is straightforward to check that $\xi\pi_c$ is $1/N_e$ as given by equation (7).

A variant of this argument is to compute $\xi$ as $(1 - F_{ST})\rho$ where $\rho$ is the equilibrium solution of the recursion

$$\rho \sim \rho'((1-e)(1-m)^2 + e\phi) + \frac{1}{n_d}$$
$$\times \frac{1 - (1-m)^2(1-e)^2}{1-e} \tag{23}$$

obtained from equation (2) by ignoring the coalescence terms $1/N$ and $1/k$. The rationale for the computation of $\rho$ is given in the Appendix. This approach again yields equation (7), but now $1/N_e$ is expressed in the form $(1 - F_{ST})\rho\pi_c$, which allows a comparison with an argument on p. 434 of Whitlock and Barton (1997). They derive $1/N_e$ from their equation (13), in the form

$$\left(\sum_{x=0}^{\infty} e(1-e)^x \vartheta_x^2\right)(1 - F_{ST})/N.$$

However, if effective size was of this form, then their $\vartheta_x$ should depend on $k$ (as $\rho\pi_c$ does). This is not the case. The resulting formula tends to overestimate effective size, possibly by a factor of 100 or more for $\phi = 1$ and $e \gg m$ (details not shown). It also conflicts with the approximation $1/N_e \approx 2(m+e)F_{ST}/n_d$ given by Whitlock and Barton (1997) and further considered by Pannell and Charlesworth (1999). Much the same can be said of their equation (22), which is correct only when $k \to \infty$ (for $\phi = 0$). A possible explanation for these discrepancies is that results are derived from their equation (3), which does not hold in Slatkin's models (see the Appendix).

The approximation $1/N_e \approx 2(m+e)F_{ST}/n_d$ is valid, but in need of a general argument. This approximation can be deduced simply by expressing $Q^R$ as a function of $F_{ST} \equiv Q$ using equation (15), plugging the result in equation (5), and simplifying for small $m$ and $e$.

## Discussion

It should be a relief to everyone that effective size can be obtained by the simple coalescent argument leading to equation (4). Such arguments efficiently yield expressions for effective size in more complex metapopulations with variable deme size (Rousset, in press). However, the coalescent argument has been obscured by earlier analyses (except Wakeley And Aliacar, 2001), which conflict with the present results. Previous recursions for probabilities of identity in Slatkin (1977), Whitlock and McCauley (1990) and Pannell and Charlesworth (1999) are inconsistent with Slatkin's life cycle and do not correspond to another well-defined life cycle. These discrepancies affect expressions for $F_{ST}$ in Whitlock and McCauley (1990) and Whitlock and Barton (1997) and for effective size in Whitlock and Barton (1997) and Pannell and Charlesworth (1999). Quantitatively, effective size differs slightly from the expression resulting from Pannell and Charlesworth's system of recursions, and may differ substantially from equation (22) of Whitlock and Barton (1997) (for $\phi = 0$) or from results based on their equation (13).

Expectedly, the present results support the intuitive conclusion that extinctions reduce the effective size, which previous works had reached. The simple coalescent argument easily yields equation (5), which shows that propagule size $k$ and probability of common origin $\phi$ affect effective size only through their effects on $Q^R$, that is on $F_{ST}$. Also as expected, lower $k$ and higher $\phi$ reduce the effective size.

The assumption that two successive reproduction events occur right after extinction when only one occurs in nonextinct demes may seem unnatural and is easily relaxed (eg Whitlock et al, 1993), but results will then depend on additional assumptions about the life cycle, that is whether demes of $k$ colonizers produce as much juveniles as demes of $N$ individuals. If so, equation (5) is still valid, giving $N_e$ in terms of the identity $Q^R$ among gametes produced within a deme. $Q^R$ obeys a recursion

of the form

$$Q^R = (1-e)\left(\frac{1}{N} + \frac{N-1}{N}(1-m)^2 Q^R\right) \\ + e\left(\frac{1}{k} + \left(1-\frac{1}{k}\right)\phi Q^R\right) \qquad (24)$$

which yields

$$Q^R = \frac{(1-e)1/N + e/k}{1-(1-e)(1-m)^2(1-1/N) - e\phi(1-1/k)}. \qquad (25)$$

A concrete illustration of the different formulas is obtained by applying equation (5) to two sets of estimates of demographic parameters from the literature. Whitlock (1992) estimated $2N = 21.7$ (genes copies), $m = 0.31$, $\phi = 0.5$, $e = 0.1$, $2k = 10.6$ (gene copies) in the beetle *Bolithoterus cornutus*. The ratio $N_e/(Nn_d)$ is 0.67 or 0.72 whether an intercalary generation is assumed at recolonization or not. Ingvarsson *et al* (1997) estimated $2N = 22.2$ (genes copies), $m = 0.366$, $\phi = 0.5$, $e = 0.255$, $2k = 8$ (gene copies) in the beetle *Phalacrus substriatus*. The ratio $N_e/Nn_d$ is likewise 0.35 or 0.40. Thus, the overall effect of population structure seems to be a moderate reduction of effective size, whatever formula is used. Substantially larger reductions in effective size may occur for lower numbers $k$ of colonizers relative to $N$. How often this occurs is an empirical question.

## Acknowledgements

I thank J Pannell for very helpful comments on the manuscript. I also thank M Slatkin for comments and M Whitlock for some answers. This is paper ISEM 03-024.

## Appendix

**Migration matrix approach and rationale for $\rho$:** We can write the recursions (8) and (9) in the form

$$\mathbf{Q}' = (1-\mu)^2 \mathbf{A}\mathbf{Q} + (1-\mu)^2 \tilde{\mathbf{A}}\begin{pmatrix} 1-Q_1 \\ 0 \end{pmatrix} \qquad (A.1)$$

where $\mathbf{Q}$ is the column vector $(Q_1, Q_2)$. Here $1-Q_1$ is the gain in identity when a coalescence event occurs, hence the elements $i1$ of $\tilde{\mathbf{A}}$ must be the probabilities of coalescence of a pair $i$ of genes in the previous generation; and $\mathbf{A}$ are the remaining factors of the $Q$'s in recursions (8) and (9), which describe the transitions of a pair of lineages between the states 'within the same deme' and 'in different demes'. However, coalescence events are ignored in the definition of $\mathbf{A}$. A recursion of the form of equation (A.1) reduces to equation (3) of Whitlock and Barton (1997) if $\tilde{\mathbf{A}} = \mathbf{A}$ and if $\mathbf{A}$ can be written as the direct product with itself of the migration matrix for single genes, but neither condition holds here.

The elements of the left 1 eigenvector $\boldsymbol{\varepsilon}$ of $\mathbf{A}$ give the probabilities that ancestral pairs of genes are within the same deme or in different demes, given the $\mathbf{A}$ matrix of transition probabilities, where coalescence events are ignored. The rationale for the computation of $\rho$ from equation (23) is to give $\boldsymbol{\varepsilon}$ in the form $\boldsymbol{\varepsilon} \sim (\rho, 1-\rho)$.

Premultiplying equation (26) $\boldsymbol{\varepsilon}$ yields, for $\mu = 0$,

$$\boldsymbol{\varepsilon} \cdot \mathbf{Q}' = \boldsymbol{\varepsilon} \cdot \mathbf{Q} + (\boldsymbol{\varepsilon} \cdot \tilde{\mathbf{A}})_1 (1-Q_1) \qquad (A.2)$$

hence the second term on the right is the absolute reduction in gene diversity in one generation. Thus the relative reduction in diversity per generation is

$$1 - \frac{1}{N_e} = \frac{1 - \boldsymbol{\varepsilon} \cdot \mathbf{Q}'}{1 - \boldsymbol{\varepsilon} \cdot \mathbf{Q}} = 1 - (\boldsymbol{\varepsilon} \cdot \tilde{\mathbf{A}})_1 \frac{1 - Q_1}{1 - \boldsymbol{\varepsilon} \cdot \mathbf{Q}} \\ \sim 1 - (\boldsymbol{\varepsilon} \cdot \tilde{\mathbf{A}})_1 (1 - F_{ST}). \qquad (A.3)$$

As $(\boldsymbol{\varepsilon} \cdot \tilde{\mathbf{A}})_1 \sim \rho\pi_c$, we recover the derivation of $1/N_e$ as $\rho\pi_c(1-F_{ST})$.

## References

Caswell H (2001). *Matrix Population Models*. Sinauer: Sunderland, MA.

Charlesworth B (1994). *Evolution in Age-structured Populations*, 2nd edn. Cambridge University Press: Cambridge.

Ewens WJ (1982). On the concept of the effective population size. *Theor Popul Biol* **21**: 373–378.

Hill WG (1972). Effective size of populations with overlapping generations. *Theor Popul Biol* **3**: 278–289.

Horn RA, Johnson CR (1985) Matrix Analysis. Cambridge University Press: Cambridge.

Hudson RR (1998). Island models and the coalescent process. *Mol Ecol* **7**: 413–418.

Ingvarsson PK, Olsson K, Ericson L (1997). Extinction–recolonization dynamics in the mycophagous beetle *Phalarus substriatus*. *Evolution* **51**: 187–195.

Nagylaki T (1983). The robustness of neutral models of geographical variation. *Theor Popul Biol* **24**: 268–294.

Nordborg M (2001). Coalescent theory. In: Balding DJ, Bishop M, Cannings C (eds.) *Handbook of Statistical Genetics*, John Wiley & Sons: Chichester, UK. pp 179–212.

Nordborg M, Krone SM (2002). Separation of time scales and convergence to the coalescent in structured populations. In: Slatkin M, Veuille M (eds.) *Modern developments in Theoretical Population Genetics*, Oxford University Press: Oxford. pp 194–232.

Pannell JR, Charlesworth B (1999). Neutral genetic diversity in a metapopulation with recurrent local extinction and recolonization. *Evolution* **53**: 664–676.

Rousset F (2002). Inbreeding and relatedness coefficients: what do they measure? *Heredity* **88**: 371–380.

Rousset F (in press). *Genetic Structure and Selection in Subdivided Populations*. Princeton University Press: Princeton, NJ

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

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

Wakeley J, Aliacar N (2001). Gene genealogies in a metapopulation. *Genetics* **159**: 893–905.

Whitlock MC (1992). Nonequilibrium population structure in forked fungus beetles: extinction, colonization, and the genetic variance among populations. *Am Nat* **139**: 952–970.

Whitlock MC, Barton NH (1997). The effective size of a subdivided population. *Genetics* **146**: 427–441.

Whitlock MC, McCauley DE (1990). Some population genetic consequences of colony formation and extinction: genetic correlations within founding groups. *Evolution* **44**: 1717–1724.

Whitlock MC, Phillips PC, Wade MJ (1993). Gene interaction affects the additive genetic variance in subdivided populations with migration and extinction. *Evolution* **47**: 1758–1769.

Wolfram S (1999). The *Mathematica Book*, 4th edn. Wolfram Media/Cambridge University Press: Cambridge.

Wright S (1943). Isolation by distance. *Genetics* **28**: 114–138.