Theoretical Population Biology **59**, 133–144 (2001)
doi:10.1006/tpbi.2000.1495, available online at http://www.idealibrary.com on IDEAL®

The Coalescent in an Island Model of Population Subdivision with Variation among Demes

John Wakeley

Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, Massachusetts 02138

Received February 14, 2000

A simple genealogical structure is found for a general finite island model of population subdivision. The model allows for variation in the sizes of demes, in contributions to the migrant pool, and in the fraction of each deme that is replaced by migrants every generation. The ancestry of a sample of non-recombining DNA sequences has a simple structure when the sample size is much smaller than the total number of demes in the population. This allows an expression for the probability distribution of the number of segregating sites in the sample to be derived under the infinite-sites mutation model. It also yields easily computed estimators of the migration parameter for each deme in a multi-deme sample. The genealogical process is such that the lineages ancestral to the sample tend to accumulate in demes with low migration rates and/or which contribute disproportionately to the migrant pool. In addition, common ancestor or coalescent events tend to occur in demes of small size. This provides a framework for understanding the determinants of the effective size of the population, and leads to an expression for the probability that the root of a genealogy occurs in a particular geographic region, or among a particular set of demes.

1. INTRODUCTION

Populations of organisms are often subdivided. However, they do not likely conform to any of the models of population subdivision commonly invoked to explain genetic variation, at least not for long periods of time. These well-studied models include the island model (Wright, 1931; Maruyama, 1970) and the one- and two-dimensional stepping stone models (Kimura and Weiss, 1964). Although few would argue that these are probably poor representations of the complicated and fluid demography of natural populations, the analysis of even these simple models of population subdivision is difficult. Nevertheless, following the recent formalization of the "structured coalescent," progress has been made on both analytical and computational fronts. The structured coalescent provides a basis for studying the genealogical history of a sample from a population that is divided into discrete demes among which there is some pattern of migration. Wilkinson-Herbots (1998) gives an excellent discussion of this model, which is a generalization of Kingman's (1982a, b) "n-coalescent."

Due to the complicated nature of the genealogical process in a subdivided population, analytic expressions have typically been confined to samples of size 2 (Notohara, 1990, 1997; Wilkinson-Herbots, 1998). These results, which can be applied to pairwise sequence comparisons, are recognized for their connection to Wright's (1951) F-statistics and often yield simple direct estimators of migration rates (Slatkin, 1991). However, it is well known that estimates of demographic parameters based on pairwise comparisons have less-desirable statistical properties than estimates made from larger samples (Tajima, 1983; Donnelly and Tavaré, 1995). It is difficult to obtain results for larger samples under the structured coalescent because the rate of coalescence during the history of the sample depends not only on the total number of lineages which existed at a given time, as in

0040-5809/01 $35.00

Kingman's coalescent, but also on their distribution among demes. Restrictive assumptions often need to be made even to obtain results for two sequences. For example, whenever there is more than just a few demes in the population, it is typically assumed that migration is isotropic, *i.e.*, that the same rate and pattern of migration holds for every deme. Examples of isotropic migration models include the symmetric island model and the one- and two-dimensional stepping stone models without ends; see Strobeck (1987).

On the computational side, we have seen the development of Monte Carlo maximum likelihood (MCML) methods of ancestral inference (Nath and Griffiths, 1996; Beerli and Felsenstein, 1999). Because these techniques use the all of the data at once, they give better estimates of demographic parameters than methods based on pairwise summary statistics. Another advantage over pairwise methods is that MCML methods allow arbitrary migration schemes to be studied, even if the number of demes is not small. However, depending on the model assumed, there can be a great many parameters to estimate. In addition, it is not clear how to account for the fact that genealogical histories depend on population-wide patterns of migration, whereas samples are often taken only from a small subset of demes. At present, MCML methods for the structured coalescent assume that the sampled demes represent the entire population; see the online documentation Bahlo and Griffiths' (2000) GENETREE (http://www.maths.monash.edu.au/$\sim$mbahlo/mpg/gtree.html) and Beerli and Felsenstein's MIGRATE (http://evolution.genetics.washington.edu/lamarc.html).

For some patterns of migration, the structured coalescent simplifies considerably if the number of demes in the population is much larger than the sample size of sequences. This was demonstrated in Wakeley (1998, 1999) for the symmetric island model and is here shown to be true when the characteristics of demes vary across the population. Under this large-number-of-demes approximation, the genealogical history of any sample can be described in detail. The key is to recognize that there are two kinds of migration events: ones in which a migrant's source deme contains one or more lineages ancestral to the sample and ones in which it contains none. Migration events to "occupied" demes occur at a much lower rate than migration events to unoccupied demes because the fraction of demes that are occupied by ancestral lineages will be small when the number of demes in the population is large. When at least one deme contains multiple lineages, coalescent events may also occur, and these happen at a rate roughly comparable to migration events to unoccupied demes. Thus, there is a separation of time scales between coalescent events and migration events to unoccupied demes on the one hand, and migration events to occupied demes on the other.

Let $\boldsymbol{n}=(n_1, n_2, ..., n_d)$ denote a sample taken from d different demes, where n_i is the sample size from the ith deme. The total sample size is $n=\sum_{i=1}^{d} n_i$. This notation differs from that of Wakeley (1998) but is the same as that of Wakeley (1999). The history of the sample back to a single common ancestor requires exactly $n-1$ coalescent events and at least $d-1$ migration events to occupied demes. There is no limit to the number migration events to unoccupied demes, and these will comprise the vast majority of events in the history of the sample. During the very recent history of the sample, because migration events to occupied demes are rare, the first n_i-1 events within each deme ($1 \leqslant i \leqslant d$) will be a mixture of coalescent events and migration events to unoccupied demes. I have called this the scattering phase because it leaves each remaining lineage in a separate deme (Wakeley, 1999). Never again during the history of the sample will there be more than two lineages in a single deme. Subsequent coalescent events will occur singly and will be widely separated in time because now each of them requires at least one migration event to an occupied deme. This much longer part of the history is called the collecting phase. A tractable ancestral process results as there is a simple probabilistic description of the scattering phase, and the collecting phase is a Kingman-type coalescent process (Wakeley, 1998, 1999).

Here I consider the genealogical history of a sample from an island model in which migration rates, contributions to the migrant pool, and deme sizes vary across the population. This general island model can produce a range of well-known features of subdivided populations, such as source-sink dynamics (Pulliam, 1988) and variation in levels of polymorphism among geographic regions, and is thus more biologically reasonable than the typical symmetric version. A large-number-of-demes approximation like the one discussed above is found to hold. Briefly, the history of a sample from d different demes depends on the values of Nm for those demes, defined below, and on the effective size of the entire population. The effective size of the population is a function of the distribution of migration rates, contributions to the migrant pool, and deme sizes across the population. Thus, populations whose demography changes over time can also be modeled. Samples larger than 2 can be studied easily, and there is no problem of having to estimate parameters of unsampled demes because these are included in the effective size of the population. Expressions for the probability distribution and the expected value of the number of segregating sites in a sample are derived, and these are used to describe the

character of genetic variation in the population. The probability that a coalescent event occurs among a particular set of demes is also obtained.

2. VARIATION AMONG DEMES

Wright's (1931) island model, with a finite number of demes (Maruyama, 1970), can be extended to allow deme sizes and two different aspects of migration to vary among demes. Consider a population of diploid, monoecious organisms which is subdivided into D different demes. Let N_i denote the number of individuals in deme i, so that deme i contains $2N_i$ copies of each genetic locus. As is typical in such models, all of what follows will apply equally well to haploid organisms with deme sizes $2N_i$. The total number of gene copies in the entire population is $2\bar{N}D = \sum_i 2N_i$. That is, $\bar{N}$ is the arithmetic mean deme size across the entire population. There is panmixia within demes, generations are non-overlapping, and there is no selection. Neutral, infinite-sites mutations occur at rate u per sequence per generation.

The following life cycle, with gametic migration, would produce the special case of the structured coalescent that is studied here. In some circumstances, diploid migration will produce the same results (Nagylaki, 1998). Each adult individual contributes a large number of gametes to its own deme's gamete pool, and to a migrant gamete pool. Demes contribute differentially to the migrant gamete pool, and not necessarily in proportion to their sizes. In addition, each deme feels migration pressure differently: the next generation of individuals in deme i is obtained by randomly sampling $2N_i - M_i$ gametes from its own gamete pool and M_i gametes from the migrant gamete pool. Random union of gametes produces N_i diploid descendents, and these replace the members of the previous generation. Note that some of the M_i migrants that enter deme i will have originated in deme i. This was not the case in Wakeley (1998, 1999), but the difference is negligible when D is large.

This migration scheme will not generally be conservative in that it will not necessarily preserve the deme sizes (Nagylaki, 1980; Strobeck, 1987). We assume instead that deme sizes are regulated independently of migration. The model allows for some demes to be overproducers and others to be underproducers depending on their contributions to the migrant gamete pool. In addition, it allows for both high turnover and low turnover demes, according to the proportion of the deme that is replaced by migrants each generation. The overproducers and/or low turnover demes are "sources," and the underproducers and/or high turnover demes are "sinks" with regard to genetic variation in the population. Of course, the relative deme sizes might be those that would be maintained by the migration pattern, so the model does not exclude the possibility of conservative migration.

We assume that variation in contributions to the migrant pool and in the acceptance of migrants is structured as follows. There are K different kinds of demes in the population. Demes of type i, together, account for a fraction α_i of the migrant gamete pool and comprise a proportion β_i of the total number of demes. Thus, $\sum_{i=1}^{K} \alpha_i$ and $\sum_{i=1}^{K} \beta_i$ are both equal to one. Further, each deme of type i gets a proportion m_i of its gametes from the migrant gamete pool each generation. These backward migration probabilities, m_i, will be referred to simply as migration rates. As implied above, every deme in the population can have a unique size; there is no restriction of deme sizes into classes. In terms of notation, it is important to remember that N_i and M_i refer to the values for a particular deme, and thus are indexed from 1 to D, whereas α_i, β_i, and m_i denote values that apply to each deme of a given type, and are indexed from 1 to K.

2.1. The Scattering Phase

Consider the recent history of coalescent events within demes and migration events to unoccupied demes for the sample $\boldsymbol{n} = (n_1, n_2, ..., n_d)$. Let $\boldsymbol{n}' = (n'_1, n'_2, ..., n'_d)$ denote the ancestral sample at the end of the scattering phase. It is important to remember that each of these $n' = \sum_{i=1}^{d} n'_i$ lineages is in a separate deme. With variation among demes, now M_i is the migration parameter for deme i and Eq. (1) in Wakeley (1999) becomes

$$P[n'_i \mid n_i] = \frac{S_{n'_i}^{(n_i)} (2M_i)^{n'_i}}{(2M_i)_{(n_i)}}, \qquad (1)$$

where $x_{(r)} = x(x+1)\cdots(x+r-1)$, and $S_j^{(i)}$ is an unsigned Stirling number of the first kind (Abramowitz and Stegun, 1964). Events in different demes are independent, and the probability function of the ancestral sample at the end of the scattering phase is given by $P[\boldsymbol{n}' \mid \boldsymbol{n}] = \prod_{i=1}^{d} P[n'_i \mid n_i]$.

The demes in the population are labeled, 1 through D, arbitrarily as a matter of convenience. Thus, numbering the sampled demes 1 through d places no restrictions on the values of their migration parameters, M_i $(1 < i < d)$. These may or may not be a random sample of demes, and their values of M may or may not be representative of the variation among demes in the entire population. In any

case, variation in M among the sampled demes will result in different expectations for the scattering phase. Demes with very small M will likely end the scattering phase with $n'_i = 1$, demes with very large M will likely end the scattering phase with $n'_i = n_i$, and others will be intermediate between these two extremes. The duration of the scattering phase is negligible in comparison to that of the collecting phase (Wakeley, 1998).

2.2. The Collecting Phase

The overall character of the collecting phase is unchanged by variation among demes. A coalescent event among the $n' = \sum_{i=1}^{d} n'_i$ lineages cannot occur until a migration event places one of the lineages into a deme occupied by another lineage. Every migration event that occurs has a very small chance of being a migration event to an occupied deme. When there is no variation among demes, this probability is equal to $(n'-1)/D$. Here it will be a function of the α_i and β_i and the distribution of lineages among the classes of demes, but it will still be proportional to $1/D$. When D is large, the lineages will migrate extensively before two of them end up the same deme and have a chance to coalesce. This is only a chance: a coalescent event will occur between them with probability $(2M_i+1)^{-1}$, where M_i is the migration parameter of the deme they are in, otherwise one of the lineages will move to a different deme. The collecting process is punctuated by these scattering events until one of them results in a coalescent event. Then the number of lineages decreases by one and the process continues.

In this section, it is shown that the usual, Kingman-type coalescent process applies to the n' collecting-phase lineages, but with an effective size that depends on the distributions of α, β, m, and N among demes. This is done by first deriving the waiting time for a migration event to an occupied deme, and then using the fact that the number of times this must occur before two lineages coalesce follows a geometric distribution.

2.2.1. *The Waiting Time for a Migration Event to an Occupied Deme.* With K different classes of backward migration probabilities, m_i, the values for the n' lineages will change over time. At the end of the scattering phase, each lineage will be in a separate deme. We can call this collection of migration rates $\{m^{(1)}, m^{(2)}, ..., m^{(n')}\}$, where each is equal to one of the K different m_i. These may or may not be a random sample from the population. The total rate of migration will be equal to $\sum_{i=1}^{n'} m^{(i)}$, and given that a migration event occurs, the probability that the ith lineage is the migrant is equal to $m^{(i)}/\sum_{j=1}^{n'} m^{(j)}$. Thus, migrants will preferentially be found among the lineages in high-m demes. However, the source deme of every migrant will be a random draw from the distribution, α, *i.e.*, will have migration rate m_i with probability α_i. If D is large enough, an equilibrium between this sampling process and the tendency of lineages to reside longer in low-m demes will be reached before a migration event to an occupied deme occurs.

Following work on the coalescent with partial selfing (Nordborg and Donnelly, 1997), Möhle (1998a) developed a convergence theorem for Markov processes that occur on different time scales which is immediately applicable to the present situation. Möhle (1998b) used this result to study convergence to the coalescent in two-sex population models, and Nordborg (1999) has recently used Möhle's theorem to include both balancing selection and partial selfing in a coalescent model. The Markov process considered here is the movement of the n' lineages among the different classes of demes, up to the time the first migration event to an occupied deme occurs. Results are derived for the case of $K=2$ classes of demes, but these are readily extended to larger values of K.

When $K=2$, the n' lineages may have anywhere from zero to n' migration rates equal to m_1, the others being equal to m_2. Let states 1 through $n'+1$ of the chain correspond to their being n' through zero lineages with migration rate m_1. There are two more states, $n'+2$ and $n'+3$, which represent migration events to an occupied deme that has migration rate m_1 and m_2, respectively, and these two states are absorbing. Ignoring terms of order m_1^2, m_2^2, D^{-2}, and smaller, the matrix, $\mathbf{\Pi}_D$, of the single-generation transition probabilities has entries

$$\pi_{i,\,i+1} = [n'-(i-1)]\, m_1\alpha_2 \left[1 - \frac{i-1}{\beta_2 D}\right] \tag{2}$$

$$\pi_{i,\,i-1} = (i-1)\, m_2\alpha_1 \left[1 - \frac{n'-(i-1)}{\beta_1 D}\right] \tag{3}$$

$$\pi_{i,\,n'+2} = \frac{\alpha_1}{\beta_1 D}[n'-(i-1)] \times \{[n'-(i-1)-1]\, m_1 + (i-1)\, m_2\} \tag{4}$$

$$\pi_{i,\,n'+3} = \frac{\alpha_2}{\beta_2 D}(i-1)\{[n'-(i-1)]\, m_1 + (i-2)\, m_2\} \tag{5}$$

$$\pi_{i,\,i} = 1 - \pi_{i,\,i+1} - \pi_{i,\,i-1} - \pi_{i,\,n'+2} - \pi_{i,\,n'+3} \tag{6}$$

in row i ($1 \leqslant i \leqslant n'+1$). All other entries in these rows are equal to zero. Rows $i=n'+2$ and $i=n'+3$ have zeros everywhere except $\pi_{ii}=1$. To explain, the term $\pi_{i,\,i+1}$ is the probability, $[n'-(i-1)]\, m_1$, that one of the lineages

currently in rate-m_1 demes is a migrant times the probability, α_2, that the migrant lineage came from a rate-m_2 deme times the probability, $1-(i-1)/(\beta_2 D)$, that the deme it came from is unoccupied. Equations (3) through (6) are obtained by similar considerations. We can write $\mathbf{\Pi}_D = \mathbf{A} + \mathbf{B}/D$, where $\mathbf{A} = \lim_{D\to\infty} \mathbf{\Pi}_D$ and $\mathbf{B} = \lim_{D\to\infty} D(\mathbf{\Pi}_D - \mathbf{A})$. Then Möhle's (1998a) theorem states that if $\mathbf{P} = \lim_{t\to\infty} \mathbf{A}^t$ exists, the discrete process described by $\mathbf{\Pi}_D$ converges on a continuous-time Markov process with time measured in units of D generations and infinitesimal generator $\mathbf{G} = \mathbf{PBP}$.

The matrix $\mathbf{A}$ can be written

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_1 & 0 \\ 0 & \mathbf{I} \end{pmatrix}, \tag{7}$$

where $\mathbf{I}$ is the (two-dimensional) identity matrix, and the $(n'+1)\times(n'+1)$ submatrix $\mathbf{A}_1$ has row i entries

$$a_{i,\,i+1} = [n' - (i-1)]\, m_1\alpha_2 \tag{8}$$

$$a_{i,\,i-1} = (i-1)\, m_2\alpha_1 \tag{9}$$

$$a_{i,\,i} = 1 - a_{i,\,i+1} - a_{i,\,i-1} \tag{10}$$

and zeros everywhere else. Solving for the stationary distribution $\mathbf{p} = \mathbf{pA}_1$ subject to the condition $\sum_{i=1}^{n'+1} p_i = 1$, we have

$$\mathbf{P} = \begin{pmatrix} \mathbf{P}_1 & 0 \\ 0 & \mathbf{I} \end{pmatrix} \tag{11}$$

in which all the rows of $\mathbf{P}_1$ are identical and have $p_i = \binom{n'}{i-1} f_1^{n'-(i-1)} f_2^{i-1}$ in column i, where

$$f_i = \frac{\alpha_i \div m_i}{\alpha_1/m_1 + \alpha_2/m_2}, \qquad i = 1, 2. \tag{12}$$

Thus, the limiting distribution of the number of the n' lineages that have migration rate equal to m_1 is binomial with parameters n' and f_1, where f_1 is equal to the average proportion of time a single lineage spends in demes with migration rate m_1.

The matrix $\mathbf{B}$ is easily derived from (2) through (6), and we write $\mathbf{G} = \mathbf{PBP}$ as

$$\begin{pmatrix} \mathbf{G}_1 & \mathbf{G}_2 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \mathbf{P}_1 & 0 \\ 0 & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{B}_1 & \mathbf{B}_2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{P}_1 & 0 \\ 0 & \mathbf{I} \end{pmatrix}. \tag{13}$$

The $(n'+1)\times 2$ submatrix $\mathbf{B}_2$ has entries $b_{i,1} = D\pi_{i,\,n'+2}$ from (4) in its first column and $b_{i,2} = D\pi_{i,\,n'+3}$ from (5) in its second column, and $\mathbf{B}_2$ is preserved in the product $\mathbf{BP}$. The two columns of the $(n'+1)\times 2$ submatrix $\mathbf{G}_2 = \mathbf{P}_1\mathbf{B}_2$ are the rates at which migration events occur to occupied demes of type 1 and 2, respectively. Because all the rows of $\mathbf{P}_1$ are identical, so are all the rows of $\mathbf{G}_2$. Thus, in the limiting continuous-time process, the rates of migration to occupied demes of type 1 and 2 are independent of distribution of the lineages among the two types of demes at the start of the collecting phase. These rates are given by

$$\begin{aligned} g_i &= \sum_{j=1}^{n'+1} p_j b_{j,\,i} \\ &= 2\binom{n'}{2} \frac{(\alpha_i/\beta_i)(\alpha_i/m_i)}{(\alpha_1/m_1 + \alpha_2/m_2)^2}, \qquad i = 1, 2. \end{aligned} \tag{14}$$

The waiting time to the first migration event to an occupied deme is exponential with rate $g_1 + g_2$.

The extension to K classes of demes is straightforward. Now the probability that a lineage is in a deme of type i, (12), becomes

$$f_i = \frac{\alpha_i/m_i}{\sum_{j=1}^{K} \alpha_j/m_j} \tag{15}$$

and the distribution of the lineages among deme types is multinomial with parameters $(n'; f_1, f_2, ..., f_K)$. The average over this multinomial distribution of the rate of migration to occupied demes of type i, (14), becomes

$$g_i = 2\binom{n'}{2} \frac{(\alpha_i/\beta_i)(\alpha_i/m_i)}{[\sum_{j=1}^{K} \alpha_j/m_j]^2} \tag{16}$$

for $i = 1, ..., K$. The waiting time for a migration event to an occupied deme is exponential with rate $\sum_{i=1}^{K} g_i$. Given that such an event has occurred, the probability that the deme is of type i is equal to

$$f_i^* = \frac{(\alpha_i/\beta_i)(\alpha_i/m_i)}{\sum_{j=1}^{K} (\alpha_j/\beta_j)(\alpha_j/m_j)}. \tag{17}$$

We make the following definition: $E_f[x] = \sum_{i=1}^{K} x_i f_i$ is the average of x over the distribution (15). Then, when time is measured in units of D generations and there are n' collecting-phase lineages, the rate of migration to occupied demes is equal to

$$2\binom{n'}{2} E_f[m]\, E_f[\alpha/\beta]. \tag{18}$$

The distribution (15) is the "apparent" distribution of deme types, from the perspective of a lineage tracing its history back in time.

2.2.2. *The Time to a Coalescent Event during the Collecting Phase.* During the collecting phase, whenever a migration event to an occupied deme occurs, there is a chance that a coalescent event will result and a chance that one of the lineages will exit the deme. The parameter which determines the probabilities of these two events is the product of the population size and the migration rate of the deme that contains the two lineages. Equation (17) gives the probability that the migration rate of the deme is m_i. In the present model, the sizes of demes can vary independently of the type of deme. Given a migration event to an occupied deme of type i, every type-i deme has an equal chance of being the one which contains the two lineages. Thus, the population size of the deme will be a random draw from the class of demes i.

The overall chance that a migration event to an occupied deme of type i results in coalescence is equal to

$$\langle (4Nm+1)^{-1}\rangle_i = \frac{1}{D\beta_i} \sum_{\{j:\, j\in\Omega_i\}} (4N_j m_i + 1)^{-1}, \qquad (19)$$

where Ω_i is the set of labels of the $D\beta_i$ demes of type i. The global average probability that a coalescent event follows a migration event to an occupied deme is equal to

$$\sum_{i=1}^{K} \langle (4Nm+1)^{-1}\rangle_i f_i^* = \frac{E_f[(\alpha/\beta)\langle (4Nm+1)^{-1}\rangle]}{E_f[(\alpha/\beta)]} \qquad (20)$$

and the number of migration events to occupied demes that must occur before coalescence is geometrically distributed with probability of success equal to (20). The waiting time for a migration event to an occupied deme is exponential with rate (18). It follows that—*e.g.*, see Appendix A in Wakeley (1999)—the time to a coalescent event among the n' lineages is exponentially distributed with rate

$$2\binom{n'}{2} E_f[m]\, E_f[(\alpha/\beta)\langle (4Nm+1)^{-1}\rangle]. \qquad (21)$$

When a coalescent event occurs, the number of lineages decreases by one and the process continues.

Thus, the collecting phase is a Kingman-type coalescent, when time is measured in units of

$$2N_e = \frac{D}{2E_f[m]\, E_f[(\alpha/\beta)\langle (4Nm+1)^{-1}\rangle]} \qquad (22)$$

generations. All the usual results for the coalescent, for example, in Tavaré (1984), apply to the collecting-phase sample, n', if we use the mutation parameter $\theta = 4N_e u$ with N_e as defined in (22). An alternative representation of (22) is

$$2N_e = \frac{D(E_\alpha[m^{-1}])^2}{2E_\alpha[\alpha(\beta m)^{-1}\langle (4Nm+1)^{-1}\rangle]}, \qquad (23)$$

where $E_\alpha[x] = \sum_{i=1}^{K} x_i\alpha_i$. Thus, the effective size of the population can be seen as a function of the averages of relevant quantities either over the actual distribution of source deme types as in (23) or over the apparent distribution of source deme types in which each type is weighted by the mean time a lineage stays in that type of deme.

3. THE DETERMINANTS OF EFFECTIVE POPULATION SIZE

This section describes how the various characteristics of a subdivided population influence its effective size. It is important to note that this is oriented toward the collecting phase of the genealogy. A full description of the history of a sample must also take the scattering phase into account. This is reserved for Section 4 below.

3.1. No Variation among Demes

The first and most important point is that the effective size of a subdivided population depends inversely on the migration rates among demes. When there is no variation among demes, we have $\bar{N} = N_i = N$, $m_i = m$, and $\alpha_i = \beta_i$ for all i, in the present model, and

$$2N_e = 2\bar{N}D\left(1 + \frac{1}{2M}\right), \qquad (24)$$

where $M = 2Nm$. When M is large, (24) is close to the value for a panmictic population, $2\bar{N}D$, and as M becomes small the effective size increases dramatically.

The parameter, M, above differs from that in Wakeley (1998, 1999) by the factor $(D-1)/D$, because here migrants can re-enter the deme which produced them.

As (24) is the standard for a subdivided population, the next three sections examine the ratio

$$\frac{N_e^{(\text{var } X)}}{N_e^{(\text{no var})}} = \frac{N_e^{(\text{var } X)}}{2\bar{N}D[1+1/(2\bar{M})]}, \tag{25}$$

where X represents some quantity which varies across the population, and $\bar{M}$ is the arithmetic average of demic migration parameters across the population. When (25) is equal to one, variation in X does not alter the effective size of the population relative to a reference population without variation that has $M=\bar{M}$.

3.2. Variation in Deme Sizes Only

When some demes are large and some are small, the effective population size is reduced relative to what would be expected in a population without variation in which $M=2\bar{N}m$. Here we have $m_i=m$ and $\alpha_i=\beta_i$ for all i and after some simplification,

$$\frac{N_e^{(\text{var } N)}}{N_e^{(\text{no var})}} = \frac{(4\bar{N}m+1)^{-1}}{\frac{1}{D}\sum_{i=1}^{D}(4N_im+1)^{-1}} \leqslant 1 \tag{26}$$

with equality only if there is no variation in deme size. The inequality in (26) is true because the harmonic mean of $4Nm+1$ over demes is less than the arithmetic mean, as long as the deme sizes differ.

The collecting process whereby lineages eventually migrate to occupied demes does not depend on the sizes of the demes; for example, see (18). Thus, the effect of variation in deme sizes during the collecting phase is manifest only through differences in the probability of a coalescent event between two lineages once they are in the same deme. The chance of coalescent event is greater if they are in a low-Nm deme than if they are in a high-Nm deme. Equation (26) shows that the decrease in coalescence time caused by low-Nm demes more than offsets the increase caused by high-Nm demes. This effect of course depends on the distribution of deme sizes, but it also depends on their magnitudes. If the deme sizes are small enough that $4N_im \ll 1$ for all i, then (26) is near unity; there is no chance for the effect of variation in deme size to be felt because every migration event to an occupied deme results in coalescence.

3.3. Variation in Migration Rates Only

In this case, we have $\bar{N}=N_i=N$, and $\alpha_i=\beta_i$ for all i, and (23) becomes

$$2N_e = \frac{D(E_\beta[m^{-1}])^2}{2E_\beta[m^{-1}(4Nm+1)^{-1}]}. \tag{27}$$

Comparing this to the value in a population that has no variation and a migration rate equal to the average migration rate of population with variation, $\bar{m}=\sum_{i=1}^{K} m_i\beta_i$, we have

$$\frac{N_e^{(\text{var } m)}}{N_e^{(\text{no var})}} = \frac{\bar{m}(4N\bar{m}+1)^{-1}}{\tilde{m}^2 E_\beta[m^{-1}(4Nm+1)^{-1}]} \tag{28}$$

in which $\tilde{m}=1/E_\beta[m^{-1}]$ is the harmonic mean of migration rates over the population.

Two limiting cases are instructive: $4Nm_i \ll 1$ and $4Nm_i \gg 1$ for all i. In the first case, as the probability of a coalescent event for two lineages that are in the same deme approaches one, (28) becomes

$$\frac{N_e^{(\text{var } m)}}{N_e^{(\text{no var})}} = \frac{\bar{m}}{\tilde{m}} \geqslant 1 \tag{29}$$

with equality only if all the m_i are identical. In the second case, using the fact that $(4Nm+1)^{-1}=(4Nm)^{-1}+O[(4Nm)^{-2}]$, as $4Nm$ grows (28) becomes

$$\frac{N_e^{(\text{var } m)}}{N_e^{(\text{no var})}} = \frac{\widetilde{m^2}}{\tilde{m}^2} \leqslant 1 \tag{30}$$

in which $\widetilde{m^2}=1/E_\beta[m^{-2}]$. Expression (30) is less than or equal to one for the same reason that variances are never negative. From the definitions of $\tilde{m}$ and $\widetilde{m^2}$, (30) is equivalent to $(E_\beta[m^{-1}])^2/E_\beta[m^{-2}]$, and the average of squared terms is always greater than or equal to the square of the average.

Variation in migration rates among demes alters the effective size of the population both because lineages tend to spend more of their history in low-m demes, and because variation in m causes variation in Nm among demes. The latter decreases the effective size—see Section 3.2—but the former has the opposite effect. Variation in m among demes leads to the "apparent" distribution of migration rates, (15), whose mean, $\tilde{m}$, is smaller

than the actual mean of migration rates among demes, $\bar{m}$. This increases the waiting time for a migration event to an occupied deme. Thus, we have (29) for the case when the effect of variation in Nm among demes is negligible because every migration event to an occupied deme results in a coalescent event. Equation (30) shows that the effect of differences in coalescence probabilities dominates when $4Nm$ is large for every deme, resulting in a decrease in effective population size. Between these two extremes, the effective size of the population can be either larger or smaller than that of a population in which all demes have migration rate $m = \bar{m}$. The details depend on the exact character of variation in migration rates among demes.

3.4. Variation in Relative Contributions to the Migrant Pool Only

Here we have $\bar{N} = N_i = N$, and $m_i = m$ for all i, and (23) becomes

$$2N_e = 2\bar{N}D\left(1 + \frac{1}{2M}\right)\frac{1}{E_\alpha[\alpha/\beta]} \tag{31}$$

which we compare to (24). We appeal again to the relationship between the harmonic mean and the arithmetic mean:

$$\frac{1}{E_\alpha[\alpha/\beta]} = \frac{1}{E_\alpha[(\beta/\alpha)^{-1}]} \leqslant E_\alpha[\beta/\alpha] = 1. \tag{32}$$

Thus, variation in relative contributions to the migrant pool decreases the effective size of a subdivided population. A simple example of this is when half of the population accounts for all of the migrants. In this case, $\alpha_1 = 1$, $\alpha_2 = 0$, $\beta_1 = 1/2$, $\beta_2 = 1/2$, and $E_\alpha[\alpha/\beta] = 2$; the effective size decreases by one-half. As with differences in N and m among demes, the effect of variation in relative contributions to the migrant pool does not disappear as the product Nm increases. In all three cases, the effective size of the population is smaller than that of a panmictic population of the same total size. This is a well-known consequence of the fact that migration is not conservative (Nagylaki, 1980; Notohara, 1993). Here, with $N_i = N$ and $m_i = m$ for all i, the only way that migration can be conservative is if $\alpha_i = \beta_i$ for all i, that is, if all demes contribute equally to the migrant pool.

4. THE DISTRIBUTION OF THE NUMBER OF SEGREGATING SITES IN A SAMPLE

By the end of the scattering phase the original sample, $\boldsymbol{n} = (n_1, ..., n_d)$, has assumed one of $\prod_{i=1}^{d} n_i$ possible configurations, $\boldsymbol{n}' = (n_1', ..., n_d')$. The probability of a configuration is given by $P[\boldsymbol{n}' \mid \boldsymbol{n}] = \prod_{i=1}^{d} P[n_i' \mid n_i]$, in which $P[n_i' \mid n_i]$ is given by (1). The total sample size at the end of the scattering phase is equal to $n' = \sum_{i=1}^{d} n_i'$. If we assume that neutral infinite-sites mutations occur at rate u per sequence per generation, and that there is no intra-locus recombination, then all the usual coalescent results, for example in Tavaré (1984), apply to this ancestral sample when $\theta = 4N_e u$ and N_e is given by (22). Note that these results will apply directly to the sample $\boldsymbol{n} = (n_1 = 1, ..., n_d = 1)$, whose scattering phase has only one possible outcome. Results for the other samples are averaged over all possible outcomes of the scattering phase. Specifically, if $S(\boldsymbol{n})$ is the number of segregating sites in the sample, then

$$P[S(\mathbf{n}) = k] = \sum_{\boldsymbol{n}'} P[S(\boldsymbol{n}') = k]\, P[\boldsymbol{n}' \mid \boldsymbol{n}]. \tag{33}$$

Summing over all possible values of k gives the corresponding equation for the expectation of $S(\boldsymbol{n})$,

$$E[S(\boldsymbol{n})] = \sum_{\boldsymbol{n}'} E[S(\boldsymbol{n}')]\, P[\boldsymbol{n}' \mid \boldsymbol{n}]. \tag{34}$$

The sums in (33) and (34) can be evaluated analytically when integral representations of $P[S(\boldsymbol{n}')]$ and $E[S(\boldsymbol{n}')]$ are used. In the latter case we have

$$E[S(\boldsymbol{n})] = \theta \int_0^1 (1-x)^{-1}\left[1 - \frac{\prod_{i=1}^{d}(2M_i x)_{(n_i)}}{x\prod_{i=1}^{d}(2M_i)_{(n_i)}}\right] dx. \tag{35}$$

This is an extension of Eq. (38) in Wakeley (1998) to the different notation for the sample and to the more general model used here. Similarly, using the unnumbered equation on page 153 of Tavaré (1984), (33) becomes

$$P[S(\boldsymbol{n}) = k] = \int_0^\infty \frac{(\theta x)^k e^{-\theta x}}{k!\, a_x^2 e^x} \prod_{i=1}^{d} \frac{(2M_i a_x)_{(n_i)}}{(2M_i)_{(n_i)}} \times \left(\sum_{i=1}^{d}\sum_{j=0}^{n_i-1} \frac{2M_i a_x}{2M_i a_x + j} - 1\right) dx \tag{36}$$

in which $a_x = 1 - e^{-x}$. Equations (35) and (36) must be integrated numerically. Alternatively, we could use (34) together with Watterson's (1975) Eq. (1.4a), and (33) together with Tavaré's (1984) Eq. (9.5). A second representation of $E[S(\boldsymbol{n})]$ is possible, by summing (36) over all possible values of k, but (35) is simpler.

It is evident from Section 3 that populations with different patterns of variation in α, β, m, and N among demes can have different effective sizes. A host of possible scenarios might be imagined, and a few of these are taken up in Section 5. Here, I focus on two interesting points about the probability distribution of the number of segregating sites in a sample, and consider these only in the context of a sample from a single deme. First, in the case of no variation among demes, the scaling of the effective size by the single migration parameter, as shown in (24), causes the distribution of the number of segregating sites in a sample from a single deme to flatten out as M decreases, even though the expected number remains more or less constant. Second, when there is variation in α, β, m, and/or N among demes, and there need be no correspondence between the value of M for the sampled deme and the effective size of the population, then both the mean and the overall shape of the distribution change with M.

Figure 1a shows the distribution of the number of segregating sites, S_{20}, in a sample of $n = 20$ sequences all from the same deme when $\theta = 5[1 + 1/(2M)]$, for three different values of the migration parameter, M. Under strong migration ($M = 10.0$), the distribution is nearly

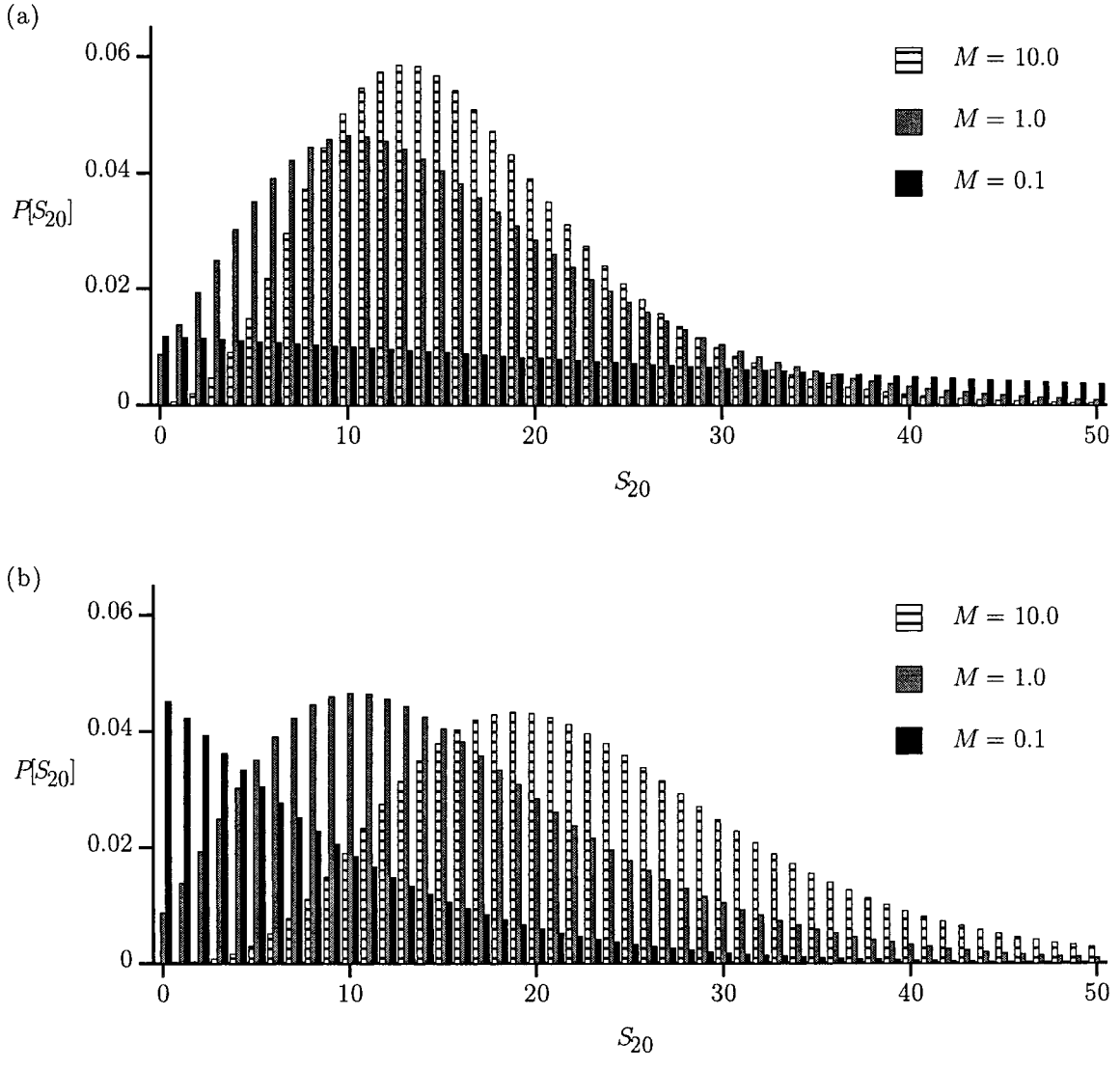

FIG. 1. The probability distribution of the number of segregating sites in the sample described in the text, calculated using Eq. (36). In (a), $\theta = 5[1 + 1/(2M)]$, and in (b), $\theta = 7.5$, for all three values of M.

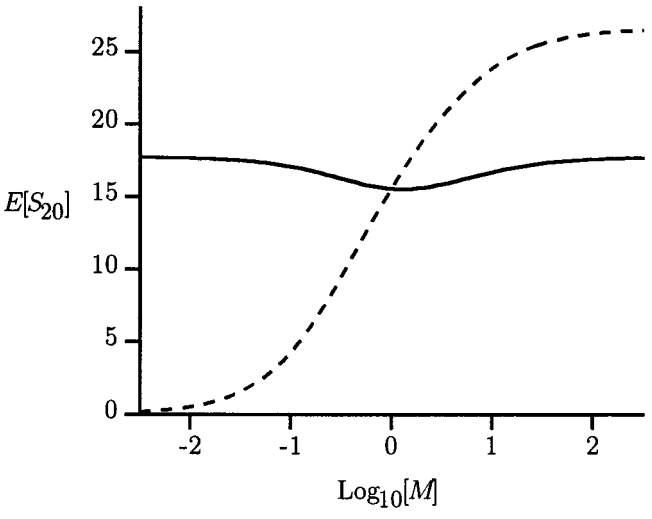

FIG. 2. The expected number of segregating sites in the sample described in the text, calculated using Eq. (35). For the solid curve, $\theta = 5[1 + 1/(2M)]$, and for the dashed curve, $\theta = 7.5$, for all M.

the same as in the panmictic case; compare this to Fig. 2 of Hudson (1990). As M decreases, the distribution spreads out dramatically. Surprisingly, the means of these distributions—16.67, 15.55, and 17.08 for $M = 0.1$, $M = 1.0$, and $M = 10.0$, respectively—are not drastically different from the panmictic value of 17.74. They would be identical for all M in samples of $n = 2$ (Li, 1976; Slatkin, 1987; Strobeck, 1987) or $n = 3$ (Wakeley, 1998), and for any sample size as M approaches infinity. For this single-deme sample of $n = 20$, the solid line in Fig. 2 shows that the expected value of S_{20} is smaller than the panmictic value when M is between about 0.1 and 10.0. However, Fig. 1a shows that changes in M have a much greater effect on the higher moments of the distribution than they do on the mean.

Figure 1b shows the distribution of the number of segregating sites for the same three values of M, again for a sample of $n = 20$ sequences all from the same deme, but now with $\theta = 7.5$. That is, the distributions in 1a and 1b are identical for the case of $M = 1.0$. When there is no hard and fast connection between the characteristics of sampled demes and the effective size of the population, increasing M pushes the distribution toward large values of S_{20}, and decreasing M has the opposite effect. The different values of M in this case might correspond to different demes sampled. The effect of differences in M on the expected number of segregating sites is shown in Fig. 2 (dashed line). Samples from demes with larger migration parameters will be more polymorphic because they will tend to contain more migrants, and thus have a larger ancestral sample during the collecting phase.

The dashed curve in Fig. 2 illustrates an important point about subdivided populations with large numbers of demes. If θ for the entire population is finite, then the per-deme θ's must be infinitesimally small. Thus, the dashed curve in Fig. 2 approaches zero as the migration parameter decreases. The alternative, which is to assume that the per-deme θ's are not small, requires that the total population θ be infinite. This would not be a biologically reasonable assumption for most sequence data, as it would predict an infinite number of segregating sites. It might, however, be useful for analyzing allelic data under the infinite-alleles model. In this case, cf. Slatkin (1982), migration becomes equivalent to mutation because every migrant, lineage will be of a unique allelic type.

5. DISCUSSION

In the general island model studied here, the history of the sample depends on patterns of migration, on relative contributions to the migrant pool, and on the sizes of demes. When the number of demes in the population is large, the genealogical history of a sample follows a simple structure. Simulations suggest that the number of demes need only be three to four times the sample size of sequences for this approximation to hold (Wakeley, 1998). Both this model and the symmetric island model share the fundamental property that samples within demes are more closely related than samples between demes. However, both also have the shortcoming that they cannot generate the important and often observed phenomenon of isolation by distance (Wright, 1951). Even so, the present model has a geographical flavor to it that the symmetric island model lacks entirely. We can imagine that geographic regions differ in their distributions of α, m, and N among demes. Because samples from demes with small values of the migration parameter, M, will be less polymorphic than samples from demes with large values of M, samples from different geographic regions may differ in this respect as well. There does not have to be a one-to-one correspondence between regions and classes of demes for this to be true.

A second aspect of the present model that is an improvement over the symmetric island model is that, during the history of a sample, lineages will not be uniformly distributed across the population. Instead, they will spend more time in the classes of demes that have small migration rates and/or account for a large fraction of all migrants. Continuing with the notion of different geographic regions, this model predicts that common ancestor events will be more likely to occur in

some regions than in others. Again there is no need for a one-to-one correspondence between regions and classes of demes, but for simplicity I will assume that this is the case. Given that a coalescent event happens, the probability that it occurs in a deme of type i can be calculated using (20)

$$\frac{\langle (4Nm+1)^{-1}\rangle_i (\alpha_i/\beta_i)(\alpha_i/m_i)}{\sum_{j=1}^{K}\langle (4Nm+1)^{-1}\rangle_j (\alpha_j/\beta_j)(\alpha_j/m_j)}. \qquad (37)$$

Thus, coalescent events tend to occur in deme types, or regions, that have small migration rates, that are the source of a disproportionate number of migrants, and that contain demes of small sizes. Further, given that such an event occurs, the probability that it occurs in deme j is equal to

$$\frac{(4N_j m_i+1)^{-1}}{\sum_{\{k:\, k\in\Omega_i\}}(4N_k m_i+1)^{-1}}, \qquad j\in\Omega_i. \qquad (38)$$

Within demic types, or regions, smaller demes are more likely than larger ones to be the location of the common ancestor events. Equations (37) and (38) apply equally to all (collecting-phase) coalescent events, including the one which defines the root of the entire genealogy. The probability that a common ancestor event occurs in a particular deme would be the product of (37) and (38), but this might be quite small because there are a large number of demes.

The results derived here yield simple estimators of demic migration parameters. Let π_i be the average number of pairwise differences between sequences in the sample from deme i, and let π_b be the average number pairwise differences between sequences from different demes. Expectation of π_b, the average number pairwise differences between sequences from different demes, is equal to the expected number of segregating sites in a sample of size 2 from two different demes: $E[\pi_b]=\theta$. This is a special case of the sample, $n_i=1$ for all i, whose genealogy is identical to that of a sample from a panmictic population with mutation parameter $\theta=4N_e u$. The quantity π_b is an unbiased estimator of θ, although others are possible. The expectation of π_i is an average over the (two) possible outcomes of the scattering phase:

$$E[\pi_i]=E[\pi_b]\frac{2M_i}{2M_i+1}. \qquad (39)$$

Then, using π_i and π_b, we can estimate the migration parameter for deme i:

$$\hat{M}_i=\frac{1}{2}\left(\frac{\pi_i}{\pi_b-\pi_i}\right). \qquad (40)$$

while better estimators could certainly be devised, for instance, by making estimates from samples larger than 2 as in Wakeley (1998), (40) is exceedingly easy to calculate. Estimates, $\hat{M}_i$ from (40), may also be useful as initial guesses in the MCML methods discussed above, even if the assumptions of present the model are violated.

As in Wakeley (1999), under this model it is possible to allow for changes in the effective size of the population over time because the collecting phase is a Kingman-type coalescent process. Note that estimates of M_i using (40) will be robust to changes in the effective size of the population because (39) will remain true as long as the changes do not occur rapidly enough to alter the scattering phase. Here, changes in effective size can result from differences in the number of demes and in the sizes of demes, as well as differences in the patterns of migration and in relative contributions to the migrant pool among demes. The results of Section 3 suggest that such changes can influence the effective size of the population in surprising ways. For example, if there is a period of time during which one set of demes is the source of all or nearly all the migrant pool, then the ancestry of samples from any deme in the population will trace back to this source region. We typically imagine that the effect of this would be a period of reduced effective population size, or a "bottleneck." However, if the migration rates in the source region are much smaller than those in other regions, especially during the time of its increased contribution to the migrant pool, it is possible that the overall effect would be a period of increased rather than reduced effective population size. Because of the influence of deme sizes and migration rates on the effective size of the population, the total replacement of the population by individuals from a restricted set of demes need not impose a bottleneck on the population. If an overly simple model of population subdivision were applied in this case the results would be misleading.

In deriving Eq. (36), it was assumed that intra-locus recombination does not occur. Recombination can easily be included in the model if it is treated identically to the way in which mutation was included here. We simply define the population recombination rate, $R=4N_e r$, where r is the rate of recombination between the two ends of the sequence and N_e is given by (22) or (23). However, we would also need to adjust R to account for the fact that a recombination event will be unobservable if, looking back, the two products of recombination coalesce

before one of them migrates. Nordborg (2000) recently discovered this phenomenon in the context of a partially-selfing population. Coalescent models of partial selfing (Nordborg and Donnelly, 1997) are similar in structure to the migration models considered here; individual organisms and rates of selfing in those models are analogous to demes and rates of migration here. Thus, the present work implies that samples from partially selfing populations in which individuals vary in their rates of selfing and in their contributions to the gene pool will have a similar genealogical structure to the one described here.

ACKNOWLEDGMENTS

I thank Magnus Nordborg and Simon Tavaré for leading me to Möhle's theorem, and Nicolas Aliacar and Jon Wilkins for helpful discussions. This work was supported by Grant DEB-9815367 from the National Science Foundation.

REFERENCES

Abramowitz, M., and Stegun, I. A. 1964. "Handbook of Mathematical Functions," Dover, New York.

Bahlo, M., and Griffiths, R. C. 2000. Inference from gene trees in a subdivided population, *Theor. Popul. Biol.* **57**, 79–95.

Beerli, P., and Felsenstein, J. 1999. Maximum-likelihood estimation of migration rates and effective population numbers in two populations using a coalescent approach, *Genetics* **152**, 763–773.

Donnelly, P., and Tavaré, S. 1995. Coalescents and genealogical structure under neutrality, *Ann. Rev. Genet.* **29**, 401–421.

Hudson, R. R. 1990. Gene genealogies and the coalescent process, *in* "Oxford Surveys in Evolutionary Biology" (D. J. Futuyma and J. Antonovics, Eds.), Vol. 7, Oxford Univ. Press, Oxford.

Kimura, M., and Weiss, G. H. 1964. The stepping stone model of population structure and the decrease of genetic correlation with distance, *Genetics* **49**, 561–576.

Kingman, J. F. C. 1982a. The coalescent, *Stochastic Process. Appl.* **13**, 235–248.

Kingman, J. F. C. 1982b. On the genealogy of large populations, *J. Appl. Probab.* **19A**, 27–43.

Li, W.-H. 1976. Distribution of nucleotide difference between two randomly chosen cistrons in a subdivided population: The finite island model, *Theor. Popul. Biol.* **10**, 303–308.

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

Möhle, M. 1998a. A convergence theorem for Markov chains arising in population genetics and the coalescent with partial selfing, *Adv. Appl. Probab.* **30**, 493–512.

Möhle, M. 1998b. Coalescent results for two-sex population models, *Adv. Appl. Probab.* **30**, 513–520.

Nagylaki, T. 1980. The strong-migration limit in geographically structured populations, *J. Math. Biol.* **9**, 101–114.

Nagylaki, T. 1998. The expected number of heterozygous sites in a subdivided population, *Genetics* **149**, 1599–1604.

Nath, H. B., and Griffiths, R. C. 1996. Estimation in an island model using simulation, *Theor. Popul. Biol.* **50**, 227–253.

Nordborg, M. 1999. The coalescent with partial selfing and balancing selection: An application of structured coalescent processes, *in* "Statistics in Molecular Biology and Genetics," (F. Seillier-Moiseiwitsch, Ed.), pp. 56–76, IMS Lecture Notes-Monograph Series, Vol. 33, Institute of Mathematical Statistics, Hayward, CA.

Nordborg, M. 2000. Linkage disequilibrium, gene trees and selfing: An ancestral recombination graph with partial selfing, *Genetics* **154**, 923–929.

Nordborg, M., and Donnelly, P. 1997. The coalescent process with selfing, *Genetics* **146**, 1185–1195.

Notohara, M. 1990. The coalescent and the genealogical process in geographically structured population, *J. Math. Biol.* **29**, 59–75.

Notohara, M. 1993. The strong migration limit for the genealogical process in geographically structured populations, *J. Math. Biol.* **31**, 115–122.

Notohara, M. 1997. The number of segregating sites in a sample of DNA sequences from a geographically structured population, *J. Math. Biol.* **36**, 188–200.

Pulliam, H. R. 1988. Sources, sinks, and population regulation, *Am. Nat.* **135**, 652–661.

Slatkin, M. 1982. Testing neutrality in a subdivided population, *Genetics* **100**, 533–545.

Slatkin, M. 1987. The average number of sites separating DNA sequences drawn from a subdivided population, *Theor. Popul. Biol.* **32**, 42–49.

Slatkin, M. 1991. Inbreeding coefficients and coalescence times, *Genet. Res. Camb.* **58**, 167–175.

Strobeck, C. 1987. Average number of nucleotide differences in an sample from a single subpopulation: A test for population subdivision, *Genetics* **117**, 149–153.

Tajima, F. 1983. Evolutionary relationship of DNA sequences in finite populations, *Genetics* **105**, 437–460.

Tavaré, S. 1984. Lines-of-descent and genealogical processes, and their application in population genetic models, *Theor. Popul. Biol.* **26**, 119–164.

Wakeley, J. 1998. Segregating sites in Wright's island model, *Theor. Popul. Biol.* **53**, 166–175.

Wakeley, J. 1999. Non-equilibrium migration in human history, *Genetics* **153**, 1863–1871.

Watterson, G. A. 1975. On the number of segregating sites in genetical models without recombination, *Theor. Popul. Biol.* **7**, 256–276.

Wilkinson-Herbots, H. M. 1998. Genealogy and subpopulation differentiation under various models of population structure, *J. Math. Biol.* **37**, 535–585.

Wrights, S. 1931. Evolution in Mendelian populations, *Genetics* **16**, 97–159.

Wrights, S. 1951. The genetical structure of populations, *Ann. Eugenics* **15**, 323–354.