J. Genet., Vol. 75, Number 1, April 1996, pp. 81–89. © Indian Academy of Sciences

Pairwise differences under a general model of population subdivision

JOHN WAKELEY
Department of Biological Sciences, Rutgers University, New Brunswick, NJ 08903, USA
Address for correspondence: Nelson Biological Labs, P.O. Box 1059, Busch Campus, Piscataway, NJ 08855–1059, USA

Abstract. A number of different migration and isolation models of population subdivision have been studied. In this paper I analyse a general model of two populations derived from a common ancestral population at some time in the past. The two populations may exchange migrants, but they may also be completely isolated from each other. I derive the expectation and variance of the number of differences between two sequences sampled from the two populations. These are then compared to the corresponding results from two other much-used models: equilibrium migration and complete isolation.

Keywords. Migration; isolation; pairwise differences.

1. Introduction

Populations in nature may rarely conform to the commonly used Wright–Fisher model (Fisher 1930; Wright 1931). One of the most prevalent deviations is when a population is divided into a number of subpopulations between which there may be genetic exchange. A number of different migration models have been studied, including the infinite island model (Wright 1931), the finite island model (Maruyama 1970), and the stepping-stone model (Kimura and Weiss 1964). The two-population migration model, in particular, has received a lot of attention (Nei and Feldman 1972; Li and Nei 1977; Tajima 1989). In all of these migration models, the rate of migration is generally assumed to have been constant for a long period of time, so that the populations are in equilibrium with respect to mutation and random genetic drift.

Two subpopulations can resemble each other either as a result of migration between them or because they have recently descended from a common ancestral population. Thus a different kind of model of population subdivision is one where there is no genetic exchange between subpopulations. In this case, which can be called isolation, the differentiation or resemblance of subpopulations is due solely to divergence after an isolation event. The two-population version of this class of models is the best studied (Takahata and Nei 1985; Hudson *et al.* 1987). This is a nonequilibrium model, but has often been compared to two-population equilibrium migration (Slatkin and Maddison 1989; Takahata and Slatkin 1990). In practice migration and isolation are quite difficult to distinguish (Wakeley 1996a).

In this paper I analyse and present some results for a general model of population subdivision. The two-population isolation model and the two-population equilibrium migration model are special cases of this. The results presented are the expectation and variance of the numbers of differences between a pair of sequences sampled from the two populations.

2. Model and statistics

In this model a single, random-mating population splits into two subpopulations which, after the split, exchange migrants. The two descendent populations are assumed to be each of haploid effective size N, and the ancestral population is assumed to be of size $2N$. At time T in the past, which is measured in units of $2N$ generations, the ancestral population split exactly in half, forming the two descendants. After that the migration rate between the two is constant, so that m is the probability that a particular individual emigrated from the other population in the previous generation. Thus $M = 2Nm$ is the effective number of migrants per generation in the entire population. In all of the following it is assumed that mutation occurs at rate u per sequence per generation, and according to the infinite-site model of Kimura (1969). Further, all mutations are of negligible selective advantage or disadvantage, i.e. they are neutral (Kimura 1968). Generations are nonoverlapping and it is assumed that m and $1/N$ are small enough that terms containing m^2, m/N and $1/(N^2)$ can be ignored.

When $M = 0$, this model reduces to a two-population isolation model. However, the model that results is different from the isolation models of Takahata and Nei (1985) and Hudson *et al.* (1987), in which the total population size changes instantaneously at time T. In the present model the total population size is constant for all time. As T approaches infinity, the present model reduces to the two-population equilibrium migration model studied by many workers.

The quantities to be studied below are the expectation and the variance of the numbers of differences between two randomly sampled sequences. If the two sequences are from the same population, call these quantities $\mathrm{E}(k_{ii'})$ and $\mathrm{Var}\,(k_{ii'})$, and if the first sequence is from one population and the second sequence is from the other, call them $\mathrm{E}(k_{ij})$ and $\mathrm{Var}\,(k_{ij})$. $\mathrm{E}(k_{ii'})$ and $\mathrm{E}(k_{ij})$ are also the expectations of the average numbers of pairwise differences within and between populations in samples of more than two sequences. However, it is important to note that $\mathrm{Var}(k_{ii'})$ and $\mathrm{Var}(k_{ij})$ apply only to samples of size two (Tajima 1983).

3. Theory and results

The four quantities are derived by considering the history of the two sampled sequences, taken either from the same or from different populations. These two cases or states can be called AA and A|A, respectively, and are also the two possible configurations of the ancestors of two sequences in the two populations, given that two distinct ancestors exist. The third possibility is that the two sequences are descended from a single common ancestral sequence, i.e. that they have coalesced. This situation can be represented as (AA).

Thus, with the assumptions outlined above, the history of a sample of two sequences during the last $T \times 2N$ generations can be described by a three-state Markov process. One of the states, (AA), is absorbing and the other two are transient. Let $\mathbf{P}\,(1)$ be the following matrix of single-generation transition probabilities.

$$\begin{array}{ll|ccc|}
 & \multicolumn{1}{l}{} & \multicolumn{3}{c}{\text{Previous state}} \\
 & \multicolumn{1}{l}{} & \text{AA} & \text{A}|\text{A} & \multicolumn{1}{c}{\text{(AA)}} \\
 & \text{AA} & 1-1/N-2m & 2m & 1/N \\
\text{Current state} & \text{A}|\text{A} & 2m & 1-2m & 0 \\
 & \text{(AA)} & 0 & 0 & 1
\end{array} \qquad (1)$$

That is, $\mathbf{P}_{i\to j}(1)$ is the probability that a sample, currently in state i, was in state j in the previous generation. Then, $\mathbf{P}_{i\to j}(t)$ is the same probability, but for generation t in the past, and $\mathbf{P}(t)$ is the t-generation transition matrix.

The assumption of a Markov process means that $\mathbf{P}(t)$ is equal to $[\mathbf{P}(1)]^t$. This can be represented in spectral form as

$$\mathbf{P}(t) = \lambda_0^t \mathbf{r}_0 \mathbf{l}'_0 + \lambda_1^t \mathbf{r}_1 \mathbf{l}'_1 + \lambda_2^t \mathbf{r}_2 \mathbf{l}'_2, \qquad (2)$$

where the λ's, $\mathbf{r}$'s and $\mathbf{l}$'s are the eigenvalues and right and left eigenvectors of $\mathbf{P}(1)$ (given in the Appendix), which are normalized so that $\mathbf{l}'_i \mathbf{r}_i = 1$. It is convenient to use a continuous-time approximation to equation (2). Let λ_i^* be equal to $\lambda_i - 1$. Then,

$$\mathbf{P}(t) = e^{\lambda_0^* t} \mathbf{r}_0 \mathbf{l}'_0 + e^{\lambda_1^* t} \mathbf{r}_1 \mathbf{l}'_1 + e^{\lambda_2^* t} \mathbf{r}_2 \mathbf{l}'_2 \qquad (3)$$

approximates equation (2) very well, as long as λ_i^* is small. Here λ_i^* is of the same order of magnitude as m and $1/N$.

In calculating $\mathrm{E}(k_{ii'})$ and $\mathrm{E}(k_{ij})$, the expected times back to the common ancestor of two sequences must be computed. These expected times, which can be called $\mathrm{E}_{\mathrm{AA}}(t)$ and $\mathrm{E}_{\mathrm{A|A}}(t)$, will in general differ depending on whether the sample is presently in state AA or A|A. They are given by

$$\mathrm{E}_{\mathrm{AA}}(t) = \int_0^\infty t f_{\mathrm{AA}}(t)\mathrm{d}t \qquad (4)$$

and

$$\mathrm{E}_{\mathrm{A|A}}(t) = \int_0^\infty t f_{\mathrm{A|A}}(t)\mathrm{d}t, \qquad (5)$$

where $f_{\mathrm{AA}}(t)\mathrm{d}t$ and $f_{\mathrm{A|A}}(t)\mathrm{d}t$ are the probabilities of coalescence into a common ancestral sequence at time t in the past. Once these are calculated, then $\mathrm{E}(k_{ii'}) = 2u\mathrm{E}_{\mathrm{AA}}(t)$, as does $\mathrm{E}(k_{jj'})$, and $\mathrm{E}(k_{ij}) = 2u\mathrm{E}_{\mathrm{A|A}}(t)$.

For two sequences to have a common ancestor in the immediately previous generation, they must both be in the same population now. For generation t in the past, the probabilities of this are $\mathbf{P}_{\mathrm{AA}\to\mathrm{AA}}(t)$ and $\mathbf{P}_{\mathrm{A|A}\to\mathrm{AA}}(t)$. When two sequences are in the same population, the probability of a coalescent event is just $\mathbf{P}_{\mathrm{AA}\to(\mathrm{AA})}(1)$ or $1/N$. Thus the probabilities of coalescence at generation t in the past, starting in states AA and A|A, respectively, are $\mathbf{P}_{\mathrm{AA}\to\mathrm{AA}}(t)\mathrm{d}t/N$ and $\mathbf{P}_{\mathrm{A|A}\to\mathrm{AA}}(t)\mathrm{d}t/N$. Switching variables, so that time, now called τ, is measured in units of $2N$ generations,

$$f_{\mathrm{AA}}(\tau)\mathrm{d}\tau = 2\mathbf{P}_{\mathrm{AA}\to\mathrm{AA}}(\tau)\mathrm{d}\tau \qquad (6)$$

and

$$f_{\mathrm{A|A}}(\tau)\mathrm{d}\tau = 2\mathbf{P}_{\mathrm{A|A}\to\mathrm{AA}}(\tau)\mathrm{d}\tau. \tag{7}$$

Equations (6) and (7) apply only to the portion of the sample's history after time T. According to the model, before T the two subpopulations were actually a single random-mating population of size $2N$. Thus, prior to T, the history of the sample follows the standard coalescence model, studied by Hudson (1983) and Tajima (1983). Dividing up the integrals (4) and (5) and simplifying somewhat,

$$\mathrm{E}_{\mathrm{AA}}(\tau) = 2\int_0^T \tau\mathbf{P}_{\mathrm{AA}\to\mathrm{AA}}(\tau)\mathrm{d}\tau + \mathbf{P}_{\mathrm{AA}\to\{\mathrm{AA\,or\,A|A}\}}(T)[T+1] \tag{8}$$

and

$$\mathrm{E}_{\mathrm{A|A}}(\tau) = 2\int_0^T \tau\mathbf{P}_{\mathrm{A|A}\to\mathrm{AA}}(\tau)\mathrm{d}\tau + \mathbf{P}_{\mathrm{A|A}\to\{\mathrm{AA\,or\,A|A}\}}(T)[T+1]. \tag{9}$$

The second terms in equations (8) and (9) are just the probabilities that the two sequences have not coalesced by time T, in the past, multiplied by the expectation of τ, given that this is the case. The probabilities in equations (8) and (9) are got using equation (3), together with the values given in the appendix, and after a change of variables from t to τ.

Evaluating the integrals in equations (8) and (9) and simplifying,

$$\mathrm{E}_{\mathrm{AA}}(\tau) = 1 - \frac{1}{\sqrt{1+4M^2}}\left[\exp\{-(1+2M-\sqrt{1+4M^2})T\} - \exp\{-(1+2M+\sqrt{1+4M^2})T\}\right] \tag{10}$$

and

$$\mathrm{E}_{\mathrm{A|A}}(\tau) = 1 + \frac{1}{2M} - \frac{1}{4M}\left[\left(1+\frac{1}{\sqrt{1+4M^2}}\right)\exp\{-(1+2M-\sqrt{1+4M^2})T\} + \left(1-\frac{1}{\sqrt{1+4M^2}}\right)\exp\{-(1+2M+\sqrt{1+4M^2})T\}\right] \tag{11}$$

are the expected times to common ancestry of two sequences sampled from the same and from different populations respectively.

The variances of the times to common ancestry are derived by calculating $\mathrm{E}_{\mathrm{AA}}(\tau^2)$ and $\mathrm{E}_{\mathrm{A|A}}(\tau^2)$. Analogous to equations (8) through (11),

$$\mathrm{E}_{\mathrm{AA}}(\tau^2) = 2\int_0^T \tau^2\mathbf{P}_{\mathrm{AA}\to\mathrm{AA}}(\tau)\mathrm{d}\tau + \mathbf{P}_{\mathrm{AA}\to\{\mathrm{AA\,or\,A|A}\}}(T)[T^2+2T+2] \tag{12}$$

and

$$\mathrm{E}_{\mathrm{A|A}}(\tau^2) = 2\int_0^T \tau^2\mathbf{P}_{\mathrm{A|A}\to\mathrm{AA}}(\tau)\mathrm{d}\tau + \mathbf{P}_{\mathrm{A|A}\to\{\mathrm{AA\,or\,A|A}\}}(T)[T^2+2T+2] \tag{13}$$

become

$$\begin{aligned} \mathrm{E}_{\mathrm{AA}}(\tau^2) = 2 + \frac{1}{2M} \\ &- \left[\frac{1+2M(3+2T)+\sqrt{1+4M^2}}{4M\sqrt{1+4M^2}}\right] \\ &\times \exp\{-(1+2M-\sqrt{1+4M^2})T\} \\ &+ \left[\frac{1+2M(3+2T)-\sqrt{1+4M^2}}{4M\sqrt{1+4M^2}}\right] \\ &\times \exp\{-(1+2M+\sqrt{1+4M^2})T\} \end{aligned} \tag{14}$$

and

$$\begin{aligned} \mathrm{E}_{\mathrm{A|A}}(\tau^2) = 2 + \frac{3}{2M} + \frac{1}{2M^2} \\ &- \left[\frac{2M^2+(1+3M+2MT)(1-\sqrt{1+4M^2})}{4M^2\sqrt{1+4M^2}}\right] \\ &\times \exp\{-(1+2M-\sqrt{1+4M^2})T\} \\ &+ \left[\frac{2M^2+(1+3M+2MT)(1+\sqrt{1+4M^2})}{4M^2\sqrt{1+4M^2}}\right] \\ &\times \exp\{-(1+2M+\sqrt{1+4M^2})T\}. \end{aligned} \tag{15}$$

The variances of times to common ancestry for a single pair of sequences are, then, $\mathrm{Var}_{\mathrm{AA}}(\tau) = E_{\mathrm{AA}}(\tau^2) - [\mathrm{E}_{\mathrm{AA}}(\tau)]^2$ and $\mathrm{Var}_{\mathrm{A|A}}(\tau) = \mathrm{E}_{\mathrm{A|A}}(\tau^2) - [\mathrm{E}_{\mathrm{A|A}}(\tau)]^2$.

The total length of the genealogy of the two sequences is 2τ, which has mean and variance $2\mathrm{E}(\tau)$ and $4\mathrm{Var}(\tau)$ respectively. Because u is small, the number of mutations in one time interval, i.e. $2N$ generations, is Poisson-distributed with mean $2Nu$. Then, since the total number of mutations on the entire genealogy is just the sum of these numbers over all intervals, the mean numbers of differences between the two sequences are $\mathrm{E}(k_{ii'}) = \theta \mathrm{E}_{\mathrm{AA}}(\tau)$ and $\mathrm{E}(k_{ij}) = \theta \mathrm{E}_{\mathrm{A|A}}(\tau)$, where $\theta = 4Nu$. Furthermore, the variances of the numbers of differences are $\mathrm{Var}(k_{ii'}) = \theta \mathrm{E}_{\mathrm{AA}}(\tau) + \theta^2 \mathrm{Var}_{\mathrm{AA}}(\tau)$ and $\mathrm{Var}(k_{ij}) = \theta \mathrm{E}_{\mathrm{A|A}}(\tau) + \theta^2 \mathrm{Var}_{\mathrm{A|A}}(\tau)$.

As mentioned above, the present model has the two-population isolation and two-population equilibrium migration models as special cases. Of course, a single random-mating population is another special case. Thus, when T is equal to zero or as M approaches infinity, $\mathrm{E}(k_{ii'}) = \mathrm{E}(k_{ij}) = \theta$, and $\mathrm{Var}(k_{ii'}) = \mathrm{Var}(k_{ij}) = \theta + \theta^2$, which is the single-population result (Watterson 1975). For a given value of T, as M approaches zero, $\mathrm{E}(k_{ij}) = \theta(T+1)$ and $\mathrm{Var}(k_{ij}) = \theta(T+1) + \theta^2$, e.g. Takahata and Nei (1985), and

$$\mathrm{E}(k_{ii'}) = \theta \frac{(1+\mathrm{e}^{-2T})}{2} \tag{16}$$

and

$$\mathrm{Var}(k_{ii'}) = \theta \frac{(1+\mathrm{e}^{-2T})}{2} + \theta^2 \left[\frac{(1-\mathrm{e}^{-4T})}{4} + (T+1)\mathrm{e}^{-2T}\right], \tag{17}$$

which are the values expected in the resulting isolation model. Equations (16) and (17) differ from the corresponding expressions in Takahata and Nei (1985) because here each descendent population is of size N and the ancestor is of size $2N$, whereas in their model both are of size $2N$ so the single-population result obtains. For a given value of M, as T approaches infinity, $\mathrm{E}(k_{ii'}) = \theta$ and $\mathrm{E}(k_{ij}) = \theta + \theta/(2M)$ (Slatkin 1987; Strobeck 1987), and

$$\mathrm{Var}(k_{ii'}) = \theta\left(1 + \theta + \frac{\theta}{2M}\right) \tag{18}$$

and

$$\mathrm{Var}(k_{ij}) = \theta\left(1 + \theta + \frac{\theta}{2M}\right) + \frac{\theta}{2M}\left(1 + \frac{\theta}{2M}\right) \tag{19}$$

(Wakeley 1996b), which are the results for two-population equilibrium migration. The behaviour of $\mathrm{E}(k_{ii'})$ and $\mathrm{Var}(k_{ii'})$, and $\mathrm{E}(k_{ij})$ and $\mathrm{Var}(k_{ij})$, over ranges of M and T, is shown in figures 1 and 2.

4. Discussion

It is clear from equation (10) and figure 1 that $\mathrm{E}(k_{ii'})$ depends on both M and T in addition to θ. This contrasts with the results of Slatkin (1987) and Strobeck (1987) who found that

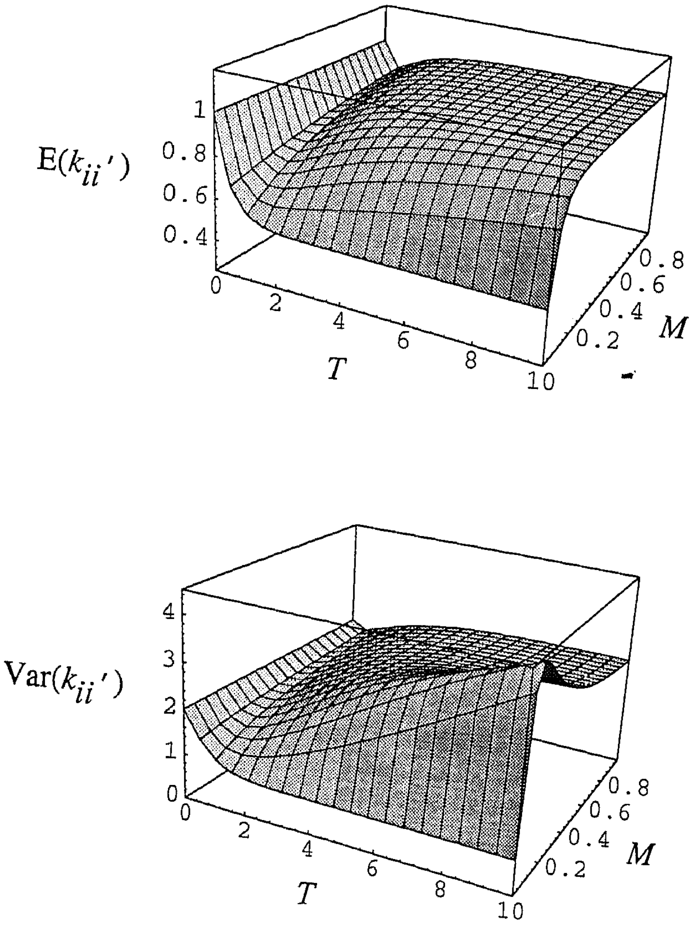

Figure 1. $\mathrm{E}(k_{ii'})$ and $\mathrm{Var}(k_{ii'})$ plotted for a range of T and M, when θ is equal to one. Note that the vertical axes are different in the two panels.

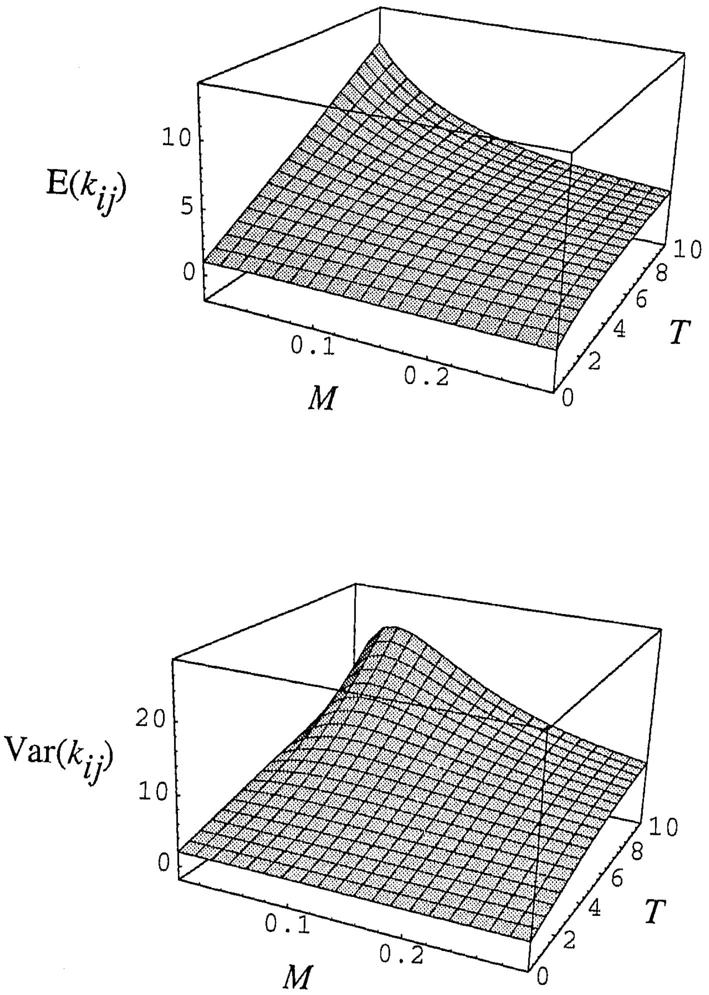

Figure 2. E(k_{ij}) and Var(k_{ij}) plotted for a range of T and M, when θ is equal to one. Note that the vertical axes are different in the two panels, and that horizontal axes are switched relative to those in figure 1.

E($k_{ii'}$) $= \theta$ and is, thus, independent of M in symmetric migration models at equilibrium. Figure 1 shows that, for any M greater than zero, there is a dip in E($k_{ii'}$) at low values of T. Then, as T increases, E($k_{ii'}$) approaches its asymptotic value of θ, which in this case equals one. E($k_{ii'}$) is smaller over a certain range of T owing to an increased chance of occurrence of a within-subpopulation coalescent event more recent than any migration event, but fewer than $T \times 2N$ generations ago, when the original population split. This effect lessens as either T or M, or both, become large. The same phenomenon accounts for the decrease in E($k_{ii'}$), for any positive T, as it approaches equation (16) as M decreases to zero.

The variance of $k_{ii'}$ shows both of these behaviours as well, but also exhibits a third phenomenon. For a given large value of T, there is a peak in Var($k_{ii'}$) as M increases from zero. This can be understood as a balance of two forces. First, as T gets large, Var($k_{ii'}$) approaches expression (18) and so should increase with $1/(2M)$. However, for any T, Var($k_{ii'}$) will converge to expression (17) as M goes to zero. For large $T(\geqslant 2$ in figure 1), expression (17) approaches the single-subpopulation value of $\theta/2 + (\theta/2)^2$. These results for E($k_{ii'}$) and Var($k_{ii'}$), especially for small M, demonstrate how fragile the equilibrium results really are.

The properties of $E(k_{ij})$ and $\mathrm{Var}(k_{ij})$, pictured in figure 2, are less complicated. Notice that the M and T axes are switched in figure 2 relative to figure 1. $E(k_{ij})$ approaches its limiting values, which were discussed above, monotonically. $\mathrm{Var}(k_{ij})$ displays some more interesting behaviour. As with $\mathrm{Var}(k_{ii'})$, the peak here is explained by a balance between the approaches to two limiting values: expression (19) as T goes to infinity and $\theta(T+1)+\theta^2$ as M goes to zero. Again, the behaviour for small M is very different from that under equilibrium migration.

The model analysed above has been called a general model throughout. However, much more general models are possible, including ones with differing subpopulation sizes and asymmetric migration. Further, only two subpopulations were considered here while populations in nature may seldom comprise just two groups. In fact, Takahata (1995) has recently proposed a very general model that is similar to the present one, but which allows for any number of subpopulations. In addition, his model includes the possibility of extinction and recolonization of subpopulations, as well as a change in population size at the time the subpopulations were formed. However, Takahata (1995) was concerned only with expected times to common ancestry, and so did not investigate the variances. The more restricted model used here was chosen so that its relationship to the two-population isolation and two-population equilibrium migration models could be illustrated.

Acknowledgements

I am greatly indebted to the late Dr Motoo Kimura for inspiration in this and in all of my work. This project was funded by PHS GM 17745-01 from the NIH.

Appendix

The eigenvalues and normalized eigenvectors of $\mathbf{P}(1)$.

$$\lambda_0 = 1, \quad \lambda_1 = 1 - \frac{(1 + 4Nm + \sqrt{1 + 16N^2m^2})}{2N},$$

$$\lambda_2 = 1 - \frac{(1 + 4Nm - \sqrt{1 + 16N^2m^2})}{2N}$$

$$\mathbf{r}'_0 = (1, 1, 1)$$

$$\mathbf{r}'_1 = \left(\frac{-(1 + \sqrt{1 + 16N^2m^2})}{C_+}, \frac{4Nm}{C_+}, 0 \right)$$

$$\mathbf{r}'_2 = \left(\frac{-(1 - \sqrt{1 + 16N^2m^2})}{C_-}, \frac{4Nm}{C_-}, 0 \right)$$

$$C_\pm = 1 + 16N^2m^2 \pm (1 + 4Nm)\sqrt{1 + 16N^2m^2}$$

$$\mathbf{l}'_0 = (0, 0, 1)$$

$$\mathbf{l}'_1 = \left(\frac{-(1 + 4Nm + \sqrt{1 + 16N^2m^2})}{2}, \frac{-(1 - 4Nm - \sqrt{1 + 16N^2m^2})}{2}, 1 \right)$$

$$\mathbf{l}'_2 = \left(\frac{-(1 + 4Nm - \sqrt{1 + 16N^2m^2})}{2}, \frac{-(1 - 4Nm + \sqrt{1 + 16N^2m^2})}{2}, 1 \right)$$

The properties of $\mathrm{E}(k_{ij})$ and $\mathrm{Var}(k_{ij})$, pictured in figure 2, are less complicated. Notice that the M and T axes are switched in figure 2 relative to figure 1. $\mathrm{E}(k_{ij})$ approaches its limiting values, which were discussed above, monotonically. $\mathrm{Var}(k_{ij})$ displays some more interesting behaviour. As with $\mathrm{Var}(k_{ii'})$, the peak here is explained by a balance between the approaches to two limiting values: expression (19) as T goes to infinity and $\theta(T+1)+\theta^2$ as M goes to zero. Again, the behaviour for small M is very different from that under equilibrium migration.

The model analysed above has been called a general model throughout. However, much more general models are possible, including ones with differing subpopulation sizes and asymmetric migration. Further, only two subpopulations were considered here while populations in nature may seldom comprise just two groups. In fact, Takahata (1995) has recently proposed a very general model that is similar to the present one, but which allows for any number of subpopulations. In addition, his model includes the possibility of extinction and recolonization of subpopulations, as well as a change in population size at the time the subpopulations were formed. However, Takahata (1995) was concerned only with expected times to common ancestry, and so did not investigate the variances. The more restricted model used here was chosen so that its relationship to the two-population isolation and two-population equilibrium migration models could be illustrated.

Acknowledgements

I am greatly indebted to the late Dr Motoo Kimura for inspiration in this and in all of my work. This project was funded by PHS GM 17745-01 from the NIH.

Appendix

The eigenvalues and normalized eigenvectors of $\mathbf{P}(1)$.

$$\lambda_0 = 1, \quad \lambda_1 = 1 - \frac{(1 + 4Nm + \sqrt{1 + 16N^2m^2})}{2N},$$

$$\lambda_2 = 1 - \frac{(1 + 4Nm - \sqrt{1 + 16N^2m^2})}{2N}$$

$$\mathbf{r}'_0 = (1, 1, 1)$$

$$\mathbf{r}'_1 = \left(\frac{-(1 + \sqrt{1 + 16N^2m^2})}{C_+}, \frac{4Nm}{C_+}, 0 \right)$$

$$\mathbf{r}'_2 = \left(\frac{-(1 - \sqrt{1 + 16N^2m^2})}{C_-}, \frac{4Nm}{C_-}, 0 \right)$$

$$C_\pm = 1 + 16N^2m^2 \pm (1 + 4Nm)\sqrt{1 + 16N^2m^2}$$

$$\mathbf{l}'_0 = (0, 0, 1)$$

$$\mathbf{l}'_1 = \left(\frac{-(1 + 4Nm + \sqrt{1 + 16N^2m^2})}{2}, \frac{-(1 - 4Nm - \sqrt{1 + 16N^2m^2})}{2}, 1 \right)$$

$$\mathbf{l}'_2 = \left(\frac{-(1 + 4Nm - \sqrt{1 + 16N^2m^2})}{2}, \frac{-(1 - 4Nm + \sqrt{1 + 16N^2m^2})}{2}, 1 \right)$$

References

Fisher R. A. 1930 *The genetical theory of natural selection* (Oxford: Clarendon Press)

Hudson R. R. 1983 Testing the constant-rate neutral allele model with protein sequence data. *Evolution* 37: 203–217

Hudson R. R., Kreitman M. and Aguade M. 1987 A test of neutral molecular evolution based on nucleotide data. *Genetics* 116: 153–159

Kimura M. 1968 Evolutionary rate at the molecular level. *Nature* 217: 624–626

Kimura M. 1969 The number of heterozygous nucleotide sites maintained in a finite population due to the steady flux of mutations. *Genetics* 61: 893–903

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

Li W.-H. and Nei M. 1977 Persistence of common alleles in two related populations or species. *Genetics* 86: 901–914

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

Nei M. and Feldman M. W. 1972 Identity of genes by descent within and between populations under mutation and migration pressure. *Theoret. Popul. Biol.* 3: 460–465

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

Slatkin M. and Maddison W. P. 1989 A cladistic measure of gene flow inferred from the phylogenies of alleles. *Genetics* 123: 603–613

Strobeck C. 1987 Average number of nucleotide differences in a 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

Tajima F. 1989 DNA polymorphism in a subdivided population: the expected number of segregating sites in the two-population model. *Genetics* 123: 229–240

Takahata N. 1995 A genetic perspective on the origin and history of humans. *Annu. Rev. Ecol. Syst.* 26: 343–372

Takahata N. and Nei M. 1985 Gene genealogy and variance of interpopulational nucleotide differences. *Genetics* 110: 325–344

Takahata N. and Slatkin M. 1990 Genealogy of neutral genes in two partially isolated populations. *Theoret. Popul. Biol.* 38: 331–350

Wakeley J. 1996a Distinguishing migration from isolation using the variance of pairwise differences. *Theoret. Popul. Biol.* 49: 369–386

Wakeley J. 1996b The variance of pairwise nucleotide differences in two populations with migration. *Theoret. Popul. Biol.* 49: 39–57

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

Wright S. 1931 Evolution in Mendelian populations. *Genetics* 16: 97–159