Genet. Res., Camb. (1988), **52**, *pp.* 213–222 *Printed in Great Britain*

The coalescent in two partially isolated diffusion populations

NAOYUKI TAKAHATA
National Institute of Genetics, Mishima, Shizuoka-ken 411, Japan

(*Received 14 August 1987 and in revised form 8 March 1988*)

Summary

The n_0 coalescent of Kingman (1982*a*, *b*) describes the family relationships among a sample of n_0 individuals drawn from a panmictic species. It is a stochastic process resulting from n_0-1 independent random events (coalescences) at each of which n ($2 \leqslant n \leqslant n_0$) ancestral lineages of a sample are descended from $n-1$ distinct ancestors for the first time. Here a similar genealogical process is studied for a species consisting of two populations with migration between them. The main interest is with the probability density of the time length between two successive coalescences and the spatial distribution of $n-1$ ancestral lineages over two populations when n to $n-1$ coalescence takes place. These are formulated based on a non-linear birth and death process with killing, and are used to derive several explicit formulae in selectively neutral population genetics models. To confirm and supplement the analytical results, a simulation method is proposed based on the underlying bivariate Markov chain. This method provides a general way for solving the present problem even when an analytical approach appears very difficult. It becomes clear that the effects of the present population structure are most conspicuous on 2 to 1 coalescence, with lesser extents on n to $n-1$ ($3 \leqslant n$) coalescence. This implies that in a more general model of population structure, the number of populations and the way in which a sample is drawn are important factors which determine the n_0 coalescent.

1. Introduction

The family relationships among a sample of n_0 members drawn from a particular generation of a population, called the n_0 coalescent by Kingman (1982*a*), play a central role in describing the genealogical behaviour of generations of individuals. As Kingman (1982*a*, *b*) and Tavaré (1984) pointed out, a genealogical approach is certainly interesting and important in its own right, but it is also very useful in producing a wide variety of classical results in the mathematical theory of population genetics. The power and elegance of the theory rely on *equivalance*, *exchangeability*, or *neutrality* among individuals and have been best demonstrated in single-locus multiple allele systems in a panmictic population (Griffiths 1980; Kingman 1982*a*, *b*; Tavaré 1984 and references therein).

However, little attempt, it seems, has been made to extend the theory to geographically structured species. This paper deals with two partially isolated populations and formulates such a population structure in terms of genealogical relationships among a sample of members drawn from these populations. Each population is assumed to be a diffusion population or diffusion time scale approximations are assumed to be valid. The outline of a general mathematical formulation and the main results for the model of population structure with symmetrical migration and equal population sizes are given in section 2. There a problem emerges concerning the spatial distribution of $n-1$ ($2 < n \leqslant n_0$) ancestral lineages over the two populations when a pair of n ancestral lineages coalesce to a common ancestor. This distribution may be derived from coalescence time, or waiting time with which n ancestral lineages were descended for the first time from $n-1$ distinct ancestors. This conjecture is confirmed in section 3 by explicitly considering the underlying probabilistic model, a continuous time bivariate Markov chain. It is demonstrated that the chain provides an efficient way of stimulating the model for arbitrary sample sizes. It seems very difficult to explore the n_0 coalescent and related processes in a general way, but for small samples it is feasible. Such mathematically tractable cases are presented in detail in sections 4 and 5, in which it is also shown how to superimpose mutations on genealogical processes. Wright's (1931) island model requires an extension of

the symmetrical population structure model to an asymmetrical one with respect to migration and population sizes. This situation is briefly considered in section 6.

The results of this paper indicate that the speed of the n_0 coalescent is largely influenced not only by migration rate but also by the number of populations composing a species and the way in which n_0 individuals are drawn. Throughout the paper, population structure is assumed not to change in time, but the theory developed here can be applied with slight modifications to the case where two populations under study were divided from an ancestral species at a given past time.

2. Formulation and results

A species considered here is monoecious and haploid, and consists of two partially isolated populations X and Y. There are constant N selectively equivalent individuals in each population. Suppose for the moment that generations are discrete and non-overlapping, and for convenience they are counted backward chronologically from the time at which n_0 individuals are randomly sampled without replacement from the species ($T=0$). Suppose also that out of n_0 individuals j_0 are drawn from X and $k_0 = n_0 - j_0$ from Y. This initial sample configuration is described by an integer j_0 in $S_{n_0} = \{0, 1, 2, \ldots, n_0\}$. When we consider the composition of this sample with respect to an ancestral species at generation T, there would be some subset of the sample which descended from an individual, T generations ago; any two individuals in this subset shared this common ancestor. For given T, there may be n such subsets ($1 \leqslant n \leqslant n_0$). Then the configuration of n ancestral lineages T ago can be described again by an integer j in $S_n = \{0, 1, 2, \ldots, n\}$.

Assume that migration occurs immediately after reproduction in each generation, in which $N_I = Nm$ individuals are exchanged at random without replacement between X and Y. Here we have taken an opportunity to define the proportion of immigrants m. There are N_I immigrants and $N_R = N - N_I$ residents accompanied with per-generation migration. Let $i \in S_n$ be the configuration of n ancestral lineages before migration, certain generations ago. The configuration after migration is determined by two probabilities. One is the probability that k ancestral lineages come from i in X from which N_R individuals are sampled to form X without migration. Denote this probability by u_{ik}, which is given by the hypergeometric distribution

$$\binom{i}{k}\binom{N-i}{N_R-k}\Big/\binom{N}{N_R}$$

[see Feller 1970, pp. 43–47]. The other is the probability that l ancestral lineages come from $n-i$ in Y from which N_I individuals are sampled to form X with migration. Denote this probability by $v_{n-i,l}$, which is given by

$$\binom{n-i}{l}\binom{N-n+i}{N_I-l}\Big/\binom{N}{N_I}.$$

Then the probability of $j \in S_n$ after migration is given by $\Sigma u_{ik} v_{n-i,j-k}$ where the sum is taken over the possible range of k. For given $j \in S_n$, this convolution provides the transition probability that the configuration was $i \in S_n$ before migration, which we denote by M_{ji}.

In what follows, we assume that m and N^{-1} are much smaller than 1 and therefore any of their higher order products can be neglected. Then M_{ji} takes a simple form given by

$$\left.\begin{aligned} M_{jj} &= 1 - nm + O(m/N), \\ M_{j,j+1} &= (n-j)m + O(m/N), \\ M_{j,j-1} &= jm + O(m/N), \\ M_{ji} &= O(m/N) \quad \text{for} \quad |i-j| > 1, \end{aligned}\right\} \tag{1}$$

where $O(x)$ stands for the order of magnitude of x.

Genealogy or lines of descent in a randomly mating population has been studied by Watterson (1975, 1984), Griffiths (1980), Kingman (1982*a*, *b*), and others among whom Tavaré (1984) provides a comprehensive review. A key quantity that was formulated in these studies is the probability G_{jk} that j individuals drawn from the population without replacement have k distinct parents. In a single population with large size N and in the absence of mutation, it is approximately given by

$$\left.\begin{aligned} G_{jk} &= 1 - \binom{j}{2}\Big/N + O(N^{-2}) && \text{if } k = j \\ &= \binom{j}{2}\Big/N + O(N^{-2}) && \text{if } k = j-1 \\ &= O(N^{-2}) && \text{otherwise} \end{aligned}\right\} \tag{2}$$

in units of generations (Kingman 1982*b*; see also Watterson 1975; Gladstien 1978).

We wish to derive a formula pertinent to the sojourn time in S_n under the joint effects of migration modelled by (1) and random sampling drift occurring in each population which is modelled by (2). Let T_n be a random time at which the number of distinct ancestral lineages becomes n ($1 \leqslant n < n_0$, $T_{n_0} = 0$), and g_j be the probability that n ancestral lineages with configuration $j \in S_n$ have n distinct parents with the same configuration. Given n,

$$1 - \binom{j}{2}\Big/N$$

is the probability that j ancestral lineages in X have j distinct parents and

$$1 - \binom{n-j}{2}\Big/N$$

https://doi.org/10.1017/S0016672300027683 Published online by Cambridge University Press

is the probability that $n-j$ ancestral lineages in Y have $n-j$ distinct parents so that g_j is given by

$$g_j = 1-\left\{\binom{j}{2}+\binom{n-j}{2}\right\}\Big/N+O(N^{-2}) \tag{3}$$

for $j = 0, 1, 2, \ldots, n$.

Now we consider a Markov chain

$$\{Z(t),\ t = 0, 1, 2, \ldots\}$$

starting at T_n and confined in S_n; the (j, i) element of the transition matrix A of $Z(t)$, irrespective of T_n, is given by

$$A_{ji} = M_{ji} g_i. \tag{4}$$

From (3),

$$\sum_{i=0}^{n} A_{ji} < 1$$

so that $Z(t)$ is a sub-Markov chain, or a Markov chain with killing; when the process enters into S_{n-1} or two ancestral lineages in S_n are descended from a common ancestor, it is terminated. The sojourn time in S_n may be computed in the following way. Let $p_i(n, t)$ be the probability that n ancestral lineages have configuration $i \in S_n$ at T_n+t, given an initial (spatial) configuration in S_n at T_n. These probabilities $p_i(n, t)$ $(i = 0, 1, 2, \ldots, n)$ satisfy a set of difference equations

$$p_i(n, t+1) = \sum_{j=0}^{n} p_j(n, t)\, A_{ji}. \tag{5}$$

We now define the probability of waiting (coalescence) time, $q(n, t)$, with which n ancestral lineages have $n-1$ distinct ancestors for the first time. Noting that $\sum_i p_i(n, t)$ is a monotonic decreasing function of t and that the difference between two successive generations corresponds to the probability that n ancestral lineages are descended from S_{n-1} exactly at the (T_n+t)th generation, it must hold that

$$q(n, t) = \sum_{i=0}^{n} \{p_i(n, t-1)-p_i(n, t)\}, \tag{6}$$

which plays the same role as $G_{nn}^{t-1}-G_{nn}^{t}$ for a panmictic population (see (2)).

Since our concern is with the case $N \to \infty$ with keeping $Nm = O(1)$, we replace (5) and (6) by their continuous time version with the scaled time $\tau = t/(2N)$. Let B be the infinitesimal generator which corresponds to A. We then obtain

$$\left.\begin{aligned} &B_{jj} = -(\tfrac{1}{2}n\theta+\alpha_j), \\ &B_{j,j+1} = \tfrac{1}{2}(n-j)\,\theta, \quad B_{j,j-1} = \tfrac{1}{2}j\theta, \\ &B_{ji} = 0 \quad \text{for } |i-j| > 1, \end{aligned}\right\} \tag{7}$$

where $\theta = 4Nm$ and $\alpha_j = \{j(j-1)+(n-j)(n-j-1)\}$. With this infinitesimal generator, (5) becomes

$$\frac{dp_i(n, \tau)}{d\tau} = \sum_{j=0}^{n} p_j(n, \tau)\, B_{ji}, \tag{8}$$

and (6) corresponds to the probability density of $\tau = (T_{n-1}-T_n)/(2N)$

$$q(n, \tau) = -\sum_{i=0}^{n} \frac{dp_i(n, \tau)}{d\tau} \tag{9}$$

which can be rewritten as

$$q(n, \tau) = \sum_{j=0}^{n} \alpha_j p_j(n, \tau) \tag{10}$$

from (7) and (8).

It is noteworthy that in the present formulation, genealogical processes in partially isolated populations are not pure death processes but birth and death processes with killing. Karlin & Tavaré (1982) developed a method for solving general linear birth and death processes, but our process is non-linear so that their method may not be directly used; it is not easy to derive the spectral representation of B for arbitrary n. Nevertheless a general prescription of the problem, as described below, is available.

Consider the Laplace transform of $p_j(n, \tau)$ and $q(n, \tau)$, defined by

$$\tilde{p}_j(n, s) = \int_0^\infty e^{-s\tau} p_j(n, \tau)\, d\tau, \quad \tilde{q}(n, s) = \int_0^\infty e^{-s\tau} q(n, \tau)\, d\tau.$$

Eqs. (8) and (10) are equivalent to

$$\sum_{j=0}^{n} \tilde{p}_j(n, s)\,(s\delta_{ij}-B_{ji}) = p_i(n, 0), \tag{11}$$

$$\tilde{q}(n, s) = \sum_{j=0}^{n} \alpha_j \tilde{p}_j(n, s) \tag{12}$$

where $\delta_{ij} = 0$ for $i \neq j$ and 1 for $i = j$. Substituting the solutions of (11) for (12) provides the Laplace transform of the probability density of n to $n-1$ coalescence time. In our model, as in the previous work, each coalescence occurs independently and the n_0 coalescent (Kingman 1982*a*) is a result of n_0-1 coalescences. Thus the probability density of the total sojourn time in the n_0 coalescent is given by the inverse Laplace transform of the product

$$\prod_{n=2}^{n_0} \tilde{q}(n, s). \tag{13}$$

However, the spatial (initial) distribution of $p_j(n, \tau)$ cannot be given *a priori* except for $n = n_0$ and must be specified whenever coalescence takes place. This distribution can be inspected in the following way. From (7), (11) and (12), we see

$$\tilde{q}(n, 0) = \sum_{j=0}^{n} \alpha_j \tilde{p}_j(n, 0) = 1. \tag{14}$$

In the above, $\tilde{p}_j(n, 0)$ is the total sojourn time in state $j \in S_n$, and (14) corresponds to the fact that n ancestral lineages derive from $n-1$ distinct ancestors in a finite length of time with probability 1. This suggests that

https://doi.org/10.1017/S0016672300027683 Published online by Cambridge University Press

the spatial distribution in S_{n-1} when a coalescence occurs may be given by

$$\left.\begin{aligned} &p_0(n-1,0) = n(n-1)\tilde{p}_0(n,0),\\ &p_j(n-1,0) = (n-j)(n-j-1)\tilde{p}_j(n,0)\\ &\qquad +j(j+1)\tilde{p}_{j+1}(n,0), \quad (j=1,2,\ldots,n-2),\\ &p_{n-1}(n-1,0) = n(n-1)\tilde{p}_n(n,0). \end{aligned}\right\} \quad (15)$$

A proof of (15) is given in the Appendix and the validity is exemplified in the following sections.

3. Simulation results

To supplement the analytical results, a simulation method is proposed based on the underlying bivariate Markov chain whose state space is two dimensional lattice points. A lattice point (j,k) stands for the situation where j distinct ancestral lineages among a sample of size n_0 resides in X and k in Y $(1 \leqslant j+k \leqslant n_0)$. The infinitesimal generator Q in this state space can be derived from the same consideration for (7), but defines a conservative birth and death process, i.e.

$$\sum_l \sum_m Q_{(j,k)(l,m)} = 0.$$

It is given by

$$\left.\begin{aligned} &q_{(j,k)} = -Q_{(j,k)(j,k)}\\ &\qquad = \tfrac{1}{2}(j+k)\theta + j(j-1) + k(k-1),\\ &Q_{(j,k)(j-1,k+1)} = \tfrac{1}{2}j\theta, \quad Q_{(j,k)(j+1,k-1)} = \tfrac{1}{2}k\theta,\\ &Q_{(j,k)(j-1,k)} = j(j-1), \quad Q_{(j,k)(j,k-1)} = k(k-1)\\ &Q_{(j,k)(l,m)} = 0 \quad \text{otherwise.} \end{aligned}\right\}$$

When $l+m = j+k$, changes are due to migration while when it decreases by one, changes are due to a coalescence between a pair of ancestral lineages in either of two populations.

Suppose that the process (sample path) just entered in state (j,k) $(j+k=n)$. The holding time to (j,k) is exponentially distributed with mean $1/q_{(j,k)}$ (e.g. Karlin & Taylor 1981, pp. 145–149). At the end of this wait, the path jumps into a new state (l,m) with probability $Q_{(j,k)(l,m)}/q_{(j,k)}$. In this new state, the path waits there a random time interval whose distribution law is an appropriate exponential and then jumps again, etc.

Our concern is with the sojourn time in state $j \in S_n$, and with the probabilities of the first arrival states in S_{n-1} when a coalescence takes place. They are denoted by $\tilde{p}_j(n,0)$ and $p_j(n-1,0)$ as in (15). In all simulations, the initial configuration in S_n was set as $j = [n/2]$ for given n ancestral lineages where $[x]$ stands for the integer part of x. Thus the path always started at $([n/2], n-[n/2])$. In each repeat, an exponential random number with the mean specified by the initial configuration was first generated and recorded. Then a uniform random number was generated to determine which state is attained according to the probability law discussed above. If the number of distinct lineages decreases by one, the path was immediately terminated with scoring one to an arrival state. Otherwise the path was continued. When the path arrives at a particular configuration in S_n more than once, the sojourn time there was computed as the total sum of independent exponential random numbers generated with the same mean. On the other hand, the spatial distribution in S_{n-1} was calculated by dividing the total scores by the number of repeats. This requires a number of repeats so that we generated 10^5 independent sample paths for a set of parameters. Both sides of (15) were then compared and the results for n to $n-1$ coalescence are given in Table 1.

It is to be noted that when we are interested in the whole process of the n_0 coalescent, the birth and death process with (16) is continued until all the members drawn from a species descend from a single common ancestor. The probability density of the total sojourn time in the n_0 coalescent can be obtained from the sojourn times in $n_0 \to n_0-1 \to \ldots \to 1$ coalescences. The number of subsets of a sample with respect to an arbitrary time $T/(2N)$ can also be computed by the same simulation procedure. This simulation method is very efficient and can be used for the case of any sample size for which the present analytical approach is difficult as seen below.

4. Coalescences for small samples

When n is small, it is not so tedious to solve (11) and (12). In this section, some explicit results for $n = 2$, 3 and 4 are presented. To facilitate computation, however, it is still effective to make use of a symmetry in (11). Let $r_i(n,\tau) = p_i(n,\tau) + p_{n-i}(n,\tau)$ for $0 \leqslant i \leqslant l-1$ $(l = [n/2])$, and let $r_i(n,\tau) = p_i(n,\tau)$ when $n = 2l$ and $r_i(n,\tau) = p_i(n,\tau) + p_{l+1}(n,\tau)$ when $n = 2l+1$. In terms of Laplace transforms of $r_i(n,\tau)$, (11) lead to, when $n = 2l$,

$$\left.\begin{aligned} &-\frac{\theta}{2}(2l+1-i)\tilde{r}_{i-1} + (\alpha_i + \theta l + s)\tilde{r}_i - \frac{\theta}{2}(i+1)\tilde{r}_{i+1} = r_i(n,0),\\ &\qquad (0 \leqslant i \leqslant l-2),\\ &-\frac{\theta}{2}(l+2)\tilde{r}_{l-2} + (\alpha_{l-1} + \theta l + s)\tilde{r}_{l-1} - \frac{\theta}{2}l\tilde{r}_{l-1} = r_{l-1}(n,0),\\ &-\frac{\theta}{2}(l+1)\tilde{r}_{l-1} + (\alpha_l + \theta l + s)\tilde{r}_l = r_l(n,0), \end{aligned}\right\} \quad (17)$$

https://doi.org/10.1017/S0016672300027683 Published online by Cambridge University Press

Table 1. *Simulation results of sojourn times in S_n and probabilities of initial configurations in S_{n-1} (the initial configuration in S_n is in state $j = [n/2]$ with probability 1)*

		$j \in S_{n-1}$ for a and b, $j \in S_n$ for c									
n		0	1	2	3	4	5	6	7	8	d
					$\theta = 4Nm = 0{\cdot}1$						
2	a	0·500	0·500	—	—	—	—	—	—	—	10·93 (1·000)
	b	0·500	0·500	—	—	—	—	—	—	—	
	c	0·250	10.43	0·248	—	—	—	—	—	—	
3	a	0·23	0·976	0·001	—	—	—	—	—	—	0·491 (0·332)
	b	0·22	0·975	0·001	—	—	—	—	—	—	
	c	0·004	0·465	0·002	0*	—	—	—	—	—	
4	a	0	0·502	0·498	0	—	—	—	—	—	0·245 (0.167)
	b	0	0·498	0·497	0	—	—	—	—	—	
	c	0	0·004	0·237	0·004	0	—	—	—	—	
5	a	0	0·252	0·743	0·004	0	—	—	—	—	0·123 (0·100)
	b	0	0·250	0·730	0·005	0	—	—	—	—	
	c	0	0	0·119	0·002	0	0	—	—	—	
6	a	0	0·002	0·497	0·499	0·002	0	—	—	—	0·082 (0·067)
	b	0	0·002	0·490	0·491	0·002	0	—	—	—	
	c	0	0	0·001	0·080	0·001	0	0	—	—	
7	a	0	0·001	0·335	0·661	0·004	0	0	—	—	0·054 (0·048)
	b	0	0·001	0·328	0·649	0·004	0	0	—	—	
	c	0	0	0	0·054	0·001	0	0	0	—	
8	a	0	0	0·002	0·494	0·502	0·002	0	0	—	0·042 (0.036)
	b	0	0	0·002	0·497	0·497	0·002	0	0	—	
	c	0	0	0	0	0·041	0	0	0	0	
					$\theta = 4Nm = 1{\cdot}0$						
2	a	0·500	0·500	—	—	—	—	—	—	—	1·970 (1·000)
	b	0·497	0·501	—	—	—	—	—	—	—	
	c	0·249	1·471	0·250	—	—	—	—	—	—	
3	a	0·128	0·833	0·039	—	—	—	—	—	—	0·441 (0.332)
	b	0·128	0·827	0·038	—	—	—	—	—	—	
	c	0·021	0·319	0·094	0·006	—	—	—	—	—	
4	a	0·010	0·489	0·492	0·010	—	—	—	—	—	0·224 (0·167)
	b	0·010	0·489	0·489	0·010	—	—	—	—	—	
	c	0·001	0·023	0·177	0·023	0	—	—	—	—	
5	a	0·003	0·278	0·678	0·040	0	—	—	—	—	0·120 (0·100)
	b	0·003	0·277	0·671	0·040	0	—	—	—	—	
	c	0	0·007	0·097	0·014	0·001	0	—	—	—	
6	a	0	0·018	0·483	0·482	0·017	0	—	—	—	0·080 (0·067)
	b	0	0·018	0·477	0·477	0·017	0	—	—	—	
	c	0	0	0.006	0·067	0·006	0	0	—	—	
7	a	0	0·008	0·340	0·620	0·032	0·001	0	—	—	0·054 (0·048)
	b	0	0·008	0·336	0·617	0·032	0·001	0	—	—	
	c	0	0	0·003	0·047	0·004	0	0	0	—	
8	a	0	0	0·017	0·484	0·480	0·018	0	0	—	0·041 (0·036)
	b	0	0	0·017	0·480	0·480	0·017	0	0	—	
	c	0	0	0	0·002	0·036	0·002	0	0	0	

[a] Probabilities of the first arrival states in S_{n-1} when coalescence takes place, corresponding to the initial distributions $p_j(n-1,0)$ in (15).
[b] Expected probabilities of the first arrival states which are computed by the right-hand sides of (15) and sojourn times given in rows c.
[c] Sojourn times in states $(j, n-j)$ until n to $n-1$ coalescence takes place, corresponding to $\tilde{p}_j(n,0)$ in (15).
[d] Total sojourn times in state n in units of $t/2N$. The values in parentheses are expected waiting times in a panmictic population of size $2N$ (see Kingman, 1982*a*, *b*).
* All values are smaller than 5×10^{-4}.

https://doi.org/10.1017/S0016672300027683 Published online by Cambridge University Press

and when $n = 2l+1$,

$$\left.\begin{aligned} &-\frac{\theta}{2}(2l+2-i)\tilde{r}_{i-1}+\left\{\alpha_i+\frac{\theta}{2}(2l+1)+s\right\}\tilde{r}_i+\frac{\theta}{2}(i+1)\tilde{r}_{i+1}=r_i(n,0), \text{ or} \\ &\qquad (0\leqslant i\leqslant l-1), \\ &-\frac{\theta}{2}(l+2)\tilde{r}_{l-1}+\left(\alpha_l+\frac{\theta}{2}l+s\right)\tilde{r}_l=r_l(n,0). \end{aligned}\right\} \quad (18)$$

In the above equations, the arguments of Laplace transforms of $r_i(n,\tau)$ were dropped. With $\tilde{r}_i(n,s)$, (12) becomes

$$\tilde{q}(n,s)=\sum_{j=0}^{[n/2]}\alpha_j\tilde{r}_j(n,s), \quad (19)$$

and the initial distribution in S_{n-1} is given by, when $n = 2l$,

$$\left.\begin{aligned} &r_i(n-1,0)=(n-i)(n-i-1)\tilde{r}_i(n,0)+i(i+1)\tilde{r}_{i+1}(n,0), \\ &r_{l-1}(n-1,0)=l(l+1)\tilde{r}_{l-1}(n,0)+l(l-1)\tilde{r}_l(n,0), \end{aligned}\right\} \quad (0\leqslant i\leqslant l-2), \quad (20)$$

and when $n = 2l+1$

$$\left.\begin{aligned} &r_i(n-1,0)=(n-i)(n-i-1)\tilde{r}_i(n,0)+i(i+1)\tilde{r}_{i+1}(n,0), \\ &r_i(n-1,0)=l(l+1)\tilde{r}_i(n,0). \end{aligned}\right\} \quad (0\leqslant i\leqslant l-1), \quad (21)$$

Case of $n = 2$. Despite its simplicity, this case is instructive and practically important to demonstrate the effects of population subdivision on coalescence. For convenience, denote $r_i(2,0)$ by r_i. From (17) we readily have

$$\left.\begin{aligned} &\tilde{r}_0(2,s)=|B_2|^{-1}\{r_0(\theta+s)+r_1\theta\}, \\ &\tilde{r}_1(2,s)=|B_2|^{-1}\{r_0\theta+r_1(\theta+2+s)\} \end{aligned}\right\} \quad (22)$$

in which $|B_2|$ is the determinant of the matrix

$$B_2=\begin{bmatrix} 2+\theta+s & -\theta \\ -\theta & \theta+s \end{bmatrix}, \quad (23)$$

or $|B_2| = 2\theta+2(1+\theta)s+s^2$. Thus (19) becomes

$$\tilde{q}(2,s)=2|B_2|^{-1}(\theta+r_0 s). \quad (24)$$

Note that $\tilde{q}(2,0) = 1$ and that $\tilde{q}(2,s) = 0$ if $\theta = 0$ and $r_0 = 0$ as they should be. Coalescence takes place with probability 1 unless there is no migration and unless two members are drawn from different populations.

The inverse Laplace transform of (24) is the probability density of waiting time (coalescence time), given by

$$q(2,\tau)=2e^{-(1+\theta)\tau}[\{\cosh(a\tau)-a^{-1}\sinh(a\tau)\}r_0 + \theta a^{-1}\sinh(a\tau)r_1], \quad (a=\sqrt{(1+\theta^2)}) \quad (25)$$

and the mean and variance are

$$M_2=1+r_1\theta^{-1}, \quad (26)$$

$$V_2=1+\theta^{-1}+(1-r_0^2)\theta^{-2} \quad (27)$$

by directly differentiating (24). These results converge to those for a single population of size $2N$ as θ increases;

$\tilde{q}(2,s)\to(1+s)^{-1}$, M_2 and $V_2\to 1$

(see, for example, Hudson, 1983; Tajima, 1983). On the other hand, as θ decreases, $q(2,\tau)$ may approach 0 and therefore M_2 infinity, depending on the initial configuration, but the variance always approaches infinity.

Case of $n = 3$. The situation is as simple as that for $n = 2$. The matrix corresponding to (18) takes the form of

$$B_3=\begin{bmatrix} 6+\frac{3}{2}\theta+s & -\frac{1}{2}\theta \\ -\frac{3}{2}\theta & 2+\frac{1}{2}\theta+s \end{bmatrix} \quad (28)$$

and therefore the determinant is

$$|B_3| = 6(2+\theta)+2(4+\theta)s+s^2.$$

Again abbreviating $r_i(3,0)$ by r_i $(i = 0, 1)$, we have

$$\left.\begin{aligned} &\tilde{r}_0(3,s)=|B_3|^{-1}\{(2+\tfrac{1}{2}\theta+s)r_0+\tfrac{1}{2}\theta r_1\}, \\ &\tilde{r}_1(3,s)=|B_3|^{-1}\{\tfrac{3}{2}\theta r_0+(6+\tfrac{3}{2}\theta+s)r_1\} \end{aligned}\right\} \quad (29)$$

and

$$\tilde{q}(3,s)=|B_3|^{-1}\{6(2+\theta)+2(1+2r_0)s\}. \quad (30)$$

Thus the probability density of 3 to 2 coalescence time is given by

$$q(3,\tau)=e^{-(4+\theta)\tau}(C_1e^{-b\tau}+C_2e^{b\tau}), \quad (31)$$

$$C_1=1+2r_0+2b^{-1}\{3r_0-(1+\theta)r_1\},$$

$$C_2=1+2r_0-2b^{-1}\{3r_0-(1+\theta)r_1\},$$

$$b=\sqrt{(4+2\theta+\theta^2)},$$

and the mean and variance are

$$\left.\begin{aligned} &M_3=\frac{1+\theta+2r_1}{3(2+\theta)}, \\ &V_3=\frac{9-4r_0r_1+5\theta+\theta^2}{9(2+\theta)^2}. \end{aligned}\right\} \quad (32)$$

https://doi.org/10.1017/S0016672300027683 Published online by Cambridge University Press

As θ decreases,

$$\left.\begin{aligned}&\tilde{q}(3,s)\to\frac{2\{6+(1+2r_0)s\}}{(2+s)(6+s)},\\&M_3\to\tfrac{1}{2}-\tfrac{1}{3}r_0,\quad V_3\to\tfrac{1}{4}-\tfrac{1}{9}r_0r_1\end{aligned}\right\}\qquad(33)$$

and as θ increases,

$$\left.\begin{aligned}&\tilde{q}(3,s)\to\frac{3}{3+s},\\&M_3\to\tfrac{1}{3},\quad V_3\to\tfrac{1}{9},\end{aligned}\right\}\qquad(34)$$

which are equivalent to those for a panmictic population of size $2N$.

For $n\geqslant 3$, coalescence does occur within a finite length of time with probability 1 even if $\theta=0$. This is because at least one of the two populations always contains more than one ancestral lineage of the members drawn from the species. As remarked in section 2, a problem occurs when coalescence takes place; we must now determine the initial configuration in S_2, which is given by

$$\left.\begin{aligned}&r_0(2,0)=6\tilde{r}_0(3,0)=\frac{\theta+4r_0(3,0)}{2(2+\theta)},\\&r_1(2,0)=2\tilde{r}_1(3,0)=\frac{\theta+4r_1(3,0)}{2(2+\theta)}\end{aligned}\right\}\qquad(35)$$

from (21). Note that trivial relationships such as $r_i(2,0)=r_i(3,0)$ when $\theta=0$ and $r_0(2,0)=r_1(2,0)=\frac{1}{2}$ when $\theta=\infty$ are met in (35). In Table 1, we have assumed that $r_1(3,0)=1-r_0(3,0)=1$. In this case, $r_0(2,0)=1-r_1(2,0)=\theta/[2(2+\theta)]$ and becomes 0·024 for $\theta=0{\cdot}1$ and 0·167 for $\theta=1$ which are precisely the same values in Table 1 (adding two terms with $j=0$ and 2 for $n=3$). Thus we claim that for a given initial configuration in S_3, the Laplace transform of 3 to 1 coalescence time in (13) has the form of

$$4\{|B_2||B_3|\}^{-1}\left\{\theta+\frac{\theta+4r_0(3,0)}{2(2+\theta)}s\right\}\{3(2+\theta)+(1+2r_0(3,0))s\}\qquad(36)$$

from (24), (30) and (35).

Case of $n=4$. The situation becomes a little complicated. From (17), we have the matrix

$$B_4=\begin{bmatrix}12+2\theta+s & -\frac{1}{2}\theta & 0\\ -2 & 6+2\theta+s & -2\theta\\ 0 & -\frac{3}{2}\theta & 4+2\theta+s\end{bmatrix}\qquad(37)$$

and the determinant

$$|B_4|=48(6+6\theta+\theta^2)+8(18+11\theta+\theta^2)s+2(11+3\theta)s^2+s^3.$$

Again abbreviating $r_i(4,0)$ by r_i $(i=0,1,2)$, the solutions of (17) can be written as

$$\left.\begin{aligned}&\tilde{r}_0(4,s)=|B_4|^{-1}\left[\{24+20\theta+\theta^2+2(5+2\theta)s+s^2\}r_0+\tfrac{1}{2}\theta(4+2\theta+s)r_1+\theta^2r_2\right],\\&\tilde{r}_1(4,s)=|B_4|^{-1}[2\theta(4+2\theta+s)r_0+\{4(2+\theta)(6+\theta)+4(4+\theta)s+s^2\}r_1+2\theta(12+2\theta+s)r_2],\\&\tilde{r}_2(4,s)=|B_4|^{-1}\left[3\theta^2r_0+\tfrac{3}{2}\theta(12+2\theta+s)r_1+\{3(26+12\theta+\theta^2)+2(2\theta+9)s+s^2\}r_2\right]\end{aligned}\right\}\qquad(38)$$

and therefore

$$\tilde{q}(4,s)=|B_4|^{-1}[48(6+6\theta+\theta^2)+4\{15(2+\theta)r_0+3(8+3\theta)r_1+(18+7\theta)r_2\}s+2(2+4r_0+r_1)s^2].\qquad(39)$$

We do not give the explicit form of 4 to 3 coalescence time, nor the mean and variance. They are too complicated and we content ourselves with the fact that (39) reduces $\tilde{q}(4,s)\to 6/(6+s)$ as θ increases. It is noteworthy, however, that the mean and variance formulae for $n=4$ are necessary to compute, for example, the variance of the number of nucleotide differences in pairwise comparisons among multiple genes sampled from two populations. This line of study was made by Takahata & Nei (1985), and we study, in the next section, a similar problem concerning the number of nucleotide differences between genes based on the infinite site model (Kimura, 1971) with no recombination (Watterson, 1975).

Before going further, we briefly examine the initial configuration in S_3 when 4 to 3 coalescence occurs. From (20), we obtain

$$\left.\begin{aligned}&r_0(3,0)=\frac{(24+20\theta+\theta^2)r_0+\theta(2+\theta)r_1+\theta^2r_2}{4(6+6\theta+\theta^2)},\\&r_1(3,0)=\\&\frac{\theta(4+3\theta)r_0+(24+22\theta+3\theta^2)r_1+(24+24\theta+3\theta^2)r_2}{4(6+6\theta+\theta^2)}.\end{aligned}\right\}\qquad(40)$$

If $\theta=0$, $r_0(3,0)$[initial configuration (0, 3) or (3, 0) for $n=3$] is the same as $r_0(4,0)$ [initial configuration (0, 4) or (4, 0) for $n=4$] and $r_1(3,0)$ [initial configuration (1, 2) or (2, 1) for $n=3$] $=1-r_0(4,0)=r_1(4,0)+r_2(4,0)$. If on the other hand $\theta=\infty$, $r_0(3,0)=\frac{1}{4}$ and $r_1(3,0)=\frac{3}{4}$. In Table 1, we have assumed that $r_2(4,0)=1$ and $r_0(4,0)=r_1(4,0)=0$. Thus $r_0(3,0)=1-r_1(3,0)=\theta^2/[4(6+6\theta+\theta^2)]$. The theoretical value of $r_0(3,0)$ becomes 0·004 for $\theta=0{\cdot}1$ and 0·019 for $\theta=1$, which agree very well with the simulation results.

5. Applications

We apply some of the above results to other problems in genealogy. The Laplace transform $\tilde{q}(n,s)$ plays a central role in solving these problems. Consider a single locus with no intragenic recombination and assume the infinite site model of Watterson (1975). Let ν be the mutation rate per locus per unit time of $\tau=T/(2N)$. For a given time span τ, the number of nucleotide changes d_n that accumulate at this locus is Poisson distributed;

$$P\{d_n=k\,|\,\tau\}=\frac{(\nu\tau)^k}{k!}e^{-\nu\tau}.\qquad(41)$$

https://doi.org/10.1017/S0016672300027683 Published online by Cambridge University Press

The generating function $h(z\,|\,\tau)$ of d_n conditioned on τ is given by

$$h(z\,|\,\tau) = e^{-\nu\tau(1-z)}. \tag{42}$$

Assume that τ is a random variable whose probability density is given by $q(n,\tau)$ in (10) and consider the unconditional generating function of (41);

$$\int_0^\infty h(z\,|\,\tau)\,q(n,\tau)\,d\tau = \int_0^\infty e^{-\nu\tau(1-z)}\,q(n,\tau)\,d\tau. \tag{43}$$

The above formula is exactly the same as the Laplace transform of $q(n,\tau)$ with parameter $\nu(1-z)$. Thus the generating function of $P\{d_n = k\}$ is given by

$$\tilde{q}(n,\nu(1-z)). \tag{44}$$

If we are interested in the number of segregating sites in a sample (Watterson, 1975), we first count the nucleotide changes that occur in all gene lineages during n to $n-1$ coalescence. Since these n lineages evolve independently, the generating function of this probability is given by

$$\tilde{q}(n,n\nu(1-z)) = \int_0^\infty h^n(z\,|\,\tau)\,q(n,\tau)\,d\tau \tag{45}$$

and therefore the generating function for the number of segregating sites during the n_0 coalescent takes the form of

$$\prod_{n=2}^{n_0} \tilde{q}(n,n\nu(1-z)). \tag{46}$$

As an example, consider the case of $n_0 = 2$. Formula (46) then becomes

$$\frac{\theta + 2\nu r_0(2,0)\,(1-z)}{\theta + 2\nu(1+\theta)\,(1-z) + 2\{\nu(1-z)\}^2}, \tag{47}$$

and $P\{d_2 = k\}$ is the coefficient of z^k in (47). In particular, $P\{d_2 = 0\}$ is the probability of *homozygosity*, given by

$$P\{d_2 = 0\} = \frac{\theta + 2\nu r_0(2,0)}{\theta + 2\nu(1+\theta) + 2\nu^2}, \tag{48}$$

which reduces well known results, $[1+2\nu]^{-1}$ when $\theta = \infty$ (Kimura & Crow, 1964) and

$$[1 + 2\nu\{1 + (1+\nu)/\theta\}]^{-1}$$

when $r_0(2,0) = 0$ (Nagylaki, 1983; Takahata, 1983). Furthermore, (47) leads to

$$P\{d_2 = k\} = \frac{\nu^k}{(1+\nu)^{k+1}} \quad \text{for} \quad \theta = 0 \quad \text{and} \quad r_0(2,0) = 1, \tag{49}$$

$$= \frac{(2\nu)^k}{(1+2\nu)^{k+1}} \quad \text{for} \quad \theta = \infty. \tag{50}$$

Equations (49) and (50) correspond to those for a panmictic population of size N and $2N$ respectively. Although we could have anticipated these results, it is to be noted that they do not necessarily delimit the range of d_2. In fact, when $\theta = 0$ and $r_0(2,0) = 0$,

$$P\{d_2 = k\} = 0$$

for any finite value of k but $P\{d_2 = \infty\} = 1$. This can be seen also in the mean and variance of nucleotide differences,

$$\left.\begin{aligned} \bar{d}_2 &= 2\nu[1 + \{1 - r_0(2,0)\}\theta^{-1}] \quad \text{(Slatkin 1987)}, \\ V(d_2) &= \bar{d}_2 + 4\nu^2[1 + \theta^{-1} + \{1 - r_0(2,0)^2\}\theta^{-2}], \end{aligned}\right\} \tag{51}$$

which are directly inspected from (26) and (27).

Another application of $\tilde{q}(n,s)$ may be as follows. In some occasions, we are interested in the number of *residence* changes of an individual lineage during n to $n-1$ coalescence. Consider the simplest case of $n = 2$. We first count the number of residence changes in both lineages, K. When two individuals are sampled from a same population, K must be even. Recalling that transition from state $(2,0)$ or $(0,2)$ to $(1,1)$ occurs with probability $\theta/(2+\theta)$ and the reverse occurs with probability 1, we have

$$P(K = 2j) = \left(\frac{\theta}{2+\theta}\right)^{2j} \frac{2}{2+\theta}, \quad (j = 0, 1, \ldots)$$

and thus the generating function of k is given by $2/[2+\theta(1-z^2)]$. Similarly, when two lineages are sampled from different populations, K must be odd and the generating function becomes $2z/[2+\theta(1-z^2)]$. The probability of $K_1 = k$ (the number of residence changes of an individual lineage) conditioned on $K = 2j$ is given by a binomial distribution

$$\binom{2j}{k}\left(\frac{1}{2}\right)^{2j},$$

because two lineages change their residence equally likely. Thus the conditional generating function of K_1 has the form of $\frac{1}{2}(1+z)^{2j}$ and therefore the unconditional one becomes

$$\frac{2\{r_0(2,0) + r_1(2,0)\,(1+z)/2\}}{2 + \theta\{1 - ((1+z)/2)^2\}}.$$

The mean and variance are

$$\tfrac{1}{2}\{r_1(2,0) + \theta\}, \quad \tfrac{1}{4}\{(2 - r_1(2,0))\,r_1(2,0) + 3\theta + \theta^2\},$$

respectively. This derivation was suggested by R. Hudson.

The corresponding formulation, however, becomes cumbersome as n increases. An approximation may then be given as follows. Given T generations, k residence changes occur with probability

$$\binom{T}{k} m^k (1-m)^{T-k}$$

approximately. [Here we ignored the fact that at least two lineages must reside in a single population when a coalescence occurs.] This probability can be well approximated by a Poisson distribution for we have assumed a small value of m. The generating function becomes

$$\sum_{k=0}^{T} \binom{T}{k} (mz)^k (1-m)^{T-k} = \exp[-\tfrac{1}{2}\theta\,\tau(1-z)] \quad \text{for} \quad \text{large } N. \tag{52}$$

https://doi.org/10.1017/S0016672300027683 Published online by Cambridge University Press

Thus if τ is again a random variable with density $q(n,\tau)$, the unconditional generating function of (52) becomes $\tilde{q}(n,\theta/2(1-z))$. In the case of $n=2$, we have

$$\tilde{q}(2,\tfrac{1}{2}\theta(1-z)) = \frac{2+r_0(2,0)(1-z)}{2+(1+\theta)(1-z)+\theta(1-z)^2/4} \quad \text{for} \quad \theta \neq 0 \qquad (53)$$

and the mean and variance are, respectively,

$$\tfrac{1}{2}\{r_1(2,0)+\theta\}, \quad \tfrac{1}{4}\{(4-r_1(2,0))\,r_1(2,0)+3\theta+\theta^2\}. \qquad (54)$$

These are in fairly good agreement with the exact solutions. It may be said that residence changes do not occur so frequently even during 2 to 1 coalescence time and even for a critical value of $\theta=1$ as a panmictic condition. Sampling from a single locality may not contain immigrants even though migration is frequent. When $\theta=0$, it is clear that $\tilde{q}(2,0)=1$ so that a lineage is confined in a single population.

6. The island model

We have studied coalescence processes in a species which consists of two populations, each having N selectively equivalent individuals and exchanging immigrants in a symmetrical way, This model does not include Wright (1931) island model. To incorporate it into the present framework of theory, we must take account of different migration rates between two populations of different sizes. Let N and N' be these population sizes and assume that Nm individuals from X and $N'm'$ from $Y(Nm=N'm')$ are exchanged between two populations. We have

$$\left.\begin{aligned} g_j &= 1-\left\{\binom{j}{2}\epsilon+\binom{n-j}{2}\epsilon'\right\}+O(\epsilon), \\ \epsilon &= N^{-1}, \quad \epsilon' = N'^{-1} \end{aligned}\right\} \qquad (55)$$

instead of (3). The continuous time version of A in (4) can be readily obtained, but scaling time is not useful. Then the infinitesimal generator B has the following components in the original time scale;

$$\left.\begin{aligned} B_{jj} &= -\{jm+(n-j)m' \\ &\qquad +\tfrac{1}{2}j(j-1)\epsilon+\tfrac{1}{2}(n-j)(n-j-1)\epsilon'\}, \\ B_{j,j+1} &= (n-j)m', \quad B_{j,j-1}=jm, \\ B_{ji} &= 0 \quad \text{for} \quad |i-j|>1. \end{aligned}\right\} \qquad (56)$$

To be complete, we present some results corresponding to (11) and (12) with (56) for the case of $n=2$. Abbreviate $p_i(2,0)$ by p_i $(i=0,1,2)$. We then have

$$\left.\begin{aligned} \tilde{p}_0(2,s) &= |B_2|^{-1}[\{2m^2+(m+m')\epsilon+(3m+m'+\epsilon)s+s^2\}p_0+m(2m+\epsilon+s)p_1+2m^2p_2], \\ \tilde{p}_2(2,s) &= |B_2|^{-1}[2m'^2p_0+m'(2m'+\epsilon'+s)p_1+\{2m'^2+(m+m')\epsilon'+(m+3m'+\epsilon')s+s^2\}p_2], \\ |B_2| &= 2m^2\epsilon'+2m'^2\epsilon+(m+m')\epsilon\epsilon'+\{2(m+m')^2+(m+3m')\epsilon+(3m+m')\epsilon'+\epsilon\epsilon'\}s \\ &\qquad +(3m+3m'+\epsilon+\epsilon')s^2+s^3. \end{aligned}\right\} \qquad (57)$$

The formula of $\tilde{p}_1(2,s)$ is unnecessary because $\tilde{q}(2,s)=\epsilon'\tilde{p}_0(2,s)+\epsilon\tilde{p}_2(2,s)$ from (12) with $n=2$. A little algebra leads to

$$\begin{aligned} \tilde{q}(2,s) = |B_2|^{-1}[&2m^2\epsilon'+2m'^2\epsilon+(m+m')\epsilon\epsilon' \\ &+\{\epsilon'(3m+m'+\epsilon)p_0+(m\epsilon'+m'\epsilon)p_1 \\ &+\epsilon(m+3m'+\epsilon')p_2\}s+(\epsilon'p_0+\epsilon p_2)s^2], \end{aligned} \qquad (58)$$

which is reduced to (24) when $\epsilon=\epsilon'$ and $m=m'$. For other quantities, we have, for example,

$$\left.\begin{aligned} M_2 &= [2(m+m')^2+(m+3m')\epsilon p_0+\{(2m+m')\epsilon' \\ &\qquad +(m+2m')\epsilon+\epsilon\epsilon'\}p_1+(3m+m')\epsilon' p_2/ \\ &\qquad [2m^2\epsilon'+2m'^2\epsilon+(m+m')\epsilon\epsilon'], \\ \bar{d}_2 &= 2\nu M_2, \quad P\{d_2=0\}=\tilde{q}(2,2\nu), \end{aligned}\right\} \qquad (59)$$

ν = per-generation mutation rate per locus.

Wright's island model with infinitely many populations may correspond to the case of $\epsilon'=0$, $m'=0$ and $p_2=1$ (two members drawn from an *island*), in which case we have

$$\tilde{q}(2,s)=\frac{\epsilon}{2m+\epsilon+s}$$

so that

$$\left.\begin{aligned} q(2,t) &= \frac{1}{N}\exp\left\{-\left(2m+\frac{1}{N}\right)t\right\}, \\ M_2 &= \infty, \quad P\{d_2=0\}=[1+2(\nu+m)N]^{-1}. \end{aligned}\right\} \qquad (60)$$

Equations (60) imply that coalescence does not occur in a finite length of generations because two members from the *island* might have come from the *continent* and when this happens, their most recent common ancestor must have existed infinitely many generations ago. No coalescence of this case is a consequence of large N in the continent and *no* time scaling.

7. Conclusions

A general effect of population structure on genealogical relationships among a sample of n_0 members drawn from a species is to spin out coalescence events at each of which a pair of the ancestral lineages of the sample derive from a common ancestor. This is equivalent to saying that population structure generally increases genetic variation in a species, because more mutations can accumulate independently in different lineages with longer coalescence times. This rather trivial conclusion is flavoured as follows. The effect of population structure can markedly prolong 2 to 1 coalescence time, but only slightly n to $n-1$ coalescence time $(3 \leqslant n \leqslant n_0)$. Such a strong effect on 2 to 1 coalescence stems from a low migration rate and two ancestral lineages that happen to occur in different populations. By contrast, the weak effect on n to $n-1$

https://doi.org/10.1017/S0016672300027683 Published online by Cambridge University Press

coalescence is due to the fact that at least two ancestral lineages must have been in a population any time. Thus the conclusion is true only in a species which is subdivided into two as modelled here. If a species is made up of c populations, the effect of such population structure should be manifest up to c to $c-1$ coalescence. This is because c ancestral lineages can be distributed over c populations separately in which case there is no possibility that two out of c such lineages coalesce to an ancestor until they happen to move in a same population. When migration is infrequent, initial sampling mainly determines the distribution of c ancestral lineages over c populations, and therefore can greatly affect genealogical processes. In this sense, an extension of the present population structure model is interesting and practically important, but beyond the scope of this paper.

We have assumed that population structure is stable throughout time, but if we deal with an incipient stage of speciation, this assumption would not be appropriate. However, the formulae of $q(n,\tau)$ and $\tilde{q}(n,s)$ in (10) and (12) would be still useful to handle such a situation. For instance, Takahata & Nei (1985) studied the variance of nucleotide differences among genes sampled from two species which have been completely isolated since their separation from an ancestral species. If we allow for migration between these descendant species and want to study the same problem, we need to use those formulae or the simulation method developed in this paper.

Most of this work was done while I was on leave in La Réunion, Mauritius and Seychelles under a grant for scientific research abroad from Monbusho, Japan. The environment and colleagues, Drs S. Ishiwa, A. Fukatami and Y. Fuyama, there offered me a valuable time to write this paper. Thanks are due to Drs M. Kimura, A. Shimizu, Y. Ogura, R. Hudson and two anonymous referees for their helpful comments on an early version of this paper.

Appendix

Equations (15) give the distribution of the state in S_{n-1} at which a coalescence (killing) occurs. From (16) we have the probability density that an event due either to migration or to killing occurs at time τ from current state $j \in S_n$ as

$$p_j(n,\tau)(n\theta/2+\alpha_j) \qquad \text{(A 1)}$$

where $\alpha_j = j(j-1)+(n-j)(n-j-1)$ as in text, and the probability that the event is a killing as

$$\alpha_j/(n\theta/2+\alpha_j). \qquad \text{(A 2)}$$

Further, the killing event may be subdivided into two possibilities: either the coalescence occurred in population X, or in population Y. The probabilities of these two events are $j(j-1)/\alpha_j$ and

$$(n-j)(n-j-1)/\alpha_j,$$

respectively. In order to get $p_j(n-1,0)$, compute first the probability that at time τ an event occurs which results in a killing and a new state $j \in S_{n-1}$. The state at which the killing occurred must have been $j+1 \in S_n$ or $j \in S_n$. The first of these events has probability

$$p_{j+1}(n,\tau)(n\theta/2+\alpha_{j+1})\frac{\alpha_{j+1}}{(n\theta/2+\alpha_{j+1})}\frac{j(j+1)}{\alpha_{j+1}}$$

while the second has probability

$$p_j(n,\tau)(n\theta/2+\alpha_j)\frac{\alpha_j}{(n\theta/2+\alpha_j)}\frac{(n-j)(n-j-1)}{\alpha_j}.$$

Combining these and simplifying give the required probability as

$$j(j+1)p_{j+1}(n,\tau)+(n-j)(n-j-1)p_j(n,\tau). \qquad \text{(A 3)}$$

Finally, integrating (A 3) from $\tau = 0$ to ∞ leads to (15).

References

Feller, W. (1970). *An Introduction to Probability Theory and Its Applications*. New York: John Wiley.

Gladstien, K. (1978). The characteristic values and vectors for a class of stochastic matrices arising in genetics. *SIAM Journal of Applied Mathematics* **34**, 630–642.

Griffiths, R. C. (1980). Lines of descent in the diffusion approximation of neutral Wright–Fisher models. *Theoretical Population Biology* **17**, 37–50.

Hudson, R. R. (1983). Testing the constant-rate neutral allele model with protein sequence data. *Evolution* **37** (1), 203–217.

Karlin, S. & Taylor, L. L. (1981). *A Second Course in Stochastic Processes*. New York: Academic Press.

Karlin, S. & Tavaré, S. (1982). Linear birth and death processes with Killing. *Journal of Applied Probability* **19**, 477–487.

Kimura, M. (1971). Theoretical foundations of population genetics at the molecular level. *Theoretical Population Biology* **2**, 174–208.

Kimura, M. & Crow, J. F. (1964). The number of alleles that can be maintained in a finite population. *Genetics* **49**, 725–738.

Kingman, J. F. C. (1982*a*). The coalescent. *Stochastic Processes and their Applications* **13**, 235–248.

Kingman, J. F. C. (1982*b*). On the genealogy in large populations. *Journal of Applied Probability* **19** A, 27–43.

Nagylaki, T. (1983). The robustness of neutral models of geographic variation. *Theoretical Population Biology* **24**, 268–294.

Slatkin, M. (1987). The average number of sites separating DNA sequences drawn from a subdivided population. *Theoretical Population Biology* **32**, 42–49.

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

Takahata, N. (1983). Gene identity and genetic differentiation of populations in the finite island model. *Genetics* **104**, 497–512.

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

Tavaré, S. (1984). Line-of-descent and genealogical process, and their applications in population genetics models. *Theoretical Population Biology* **26**, 119–164.

Watterson, G. A. (1975). On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology* **7**, 256–276.

Watterson, G. A. (1984). Lines of descent and the coalescent. *Theoretical Population Biology* **26**, 77–92.

Wright, S. (1931). Evolution of Mendelian populations. *Genetics* **16**, 97–159.