J. Math. Biol. 41, 123–142 (2000)
Digital Object Identifier (DOI):
10.1007/s002850000036

Mathematical Biology

Thomas Nagylaki

Geographical invariance and the strong-migration limit in subdivided populations

Received: 5 October 1999 / Revised version: 1 February 2000 /
Published online: 4 July 2000 –

Dedicated to the memory of Gustave Malécot (1911–1998)

Abstract. Invariance under population subdivision and the strong-migration limit are investigated for digenic samples in neutral models. The monoecious, diploid population is subdivided into a finite number of panmictic colonies that exchange gametes. The backward migration matrix is arbitrary, but time independent and ergodic (*i.e.*, irreducible and aperiodic). Results are derived for the distribution of the place and time of coalescence, for the probability of identity in the model of infinitely many alleles, and for the distribution of the number of nucleotide differences in the model of infinitely many sites without recombination.

1. Introduction

Although many, perhaps most, natural populations are distributed in space and mate at random only locally, the genetic consequences of this spatial structure are often, perhaps usually, undetectable. Frequently, the failure to demonstrate statistically significant genetic differentiation in space merely reflects lack of sufficient power to overcome sampling variation. It is also possible that the relevant evolutionary forces are uniform in the entire habitat; *e.g.*, the locus under consideration may be at an overdominant equilibrium under strong, spatially uniform selection. This paper addresses two more interesting possibilities and their interrelation. First, if the measures or relations used to analyze the data are invariant under population structure, then they will not exhibit its genetic effects. Second, and more likely, if migration is sufficiently strong to dominate the other evolutionary forces, such as selection and random drift, which can cause genetic differentiation, then the population will appear to be panmictic.

The theoretical investigation of geographical invariance was pioneered by Maruyama, whose work is reviewed, discussed, and extended in Nagylaki (1982). More recent studies have concerned the intrademe nucleotide diversity (Slatkin 1987; Strobeck 1987; Nagylaki 1998) and the coalescence time of two genes (Herbots 1994; Nagylaki 1998; Wilkinson-Herbots 1998).

T. Nagylaki: Department of Ecology and Evolution, The University of Chicago, 1101 East 57th Street, Chicago, IL 60637-1573, USA

This work was supported by National Science Foundation grant DEB-9706912.

Key words: Migration – Random genetic drift – Coalescent – Effective population number

The strong-migration limit has been analyzed by Nagylaki (1980, 1983, 1998), Notohara (1993, 1997, 1999), Nordborg (1997), and Bahlo and Griffiths (1999). The investigation here was suggested by the results in Nagylaki (1980) and Notohara (1993, 1997).

In this paper, we further explore invariance under population subdivision and the strong-migration limit for digenic samples in neutral models. Our invariance relations are exact. We obtain a fairly easy proof of simple, explicit strong-migration limits by appealing to the separation of time scales: migration reduces genetic differentiation rapidly, at rates of order unity, whereas mutation and random drift act slowly, at rates of the order of the reciprocal of the effective population number. Our invariants evolve on the slow time scale and can be evaluated explicitly in the limit. The strong-migration limits follow directly from a lemma based on these observations and stated and proved in the Appendix.

In Sect. 2, we examine the distribution of the place and time of coalescence and study the corresponding probability-generating functions and moments. Sections 3 and 4 are devoted to the models of infinitely many alleles and sites, respectively; we investigate the probability of identity in state in the former and the probability-generating function and moments of the number of nucleotide differences in the latter. In each of Sects. 2–4, we formulate recursion relations, deduce invariance relations, and derive the strong-migration limit. In Section 5, we extend Sect. 3 from gametic to diploid migration, and, as a further example of invariance, show that the conditional fixation time of a single, suitably sampled allele and the time to the most recent common ancestor of the entire population have the same probability distribution. In Section 6, we briefly summarize and discuss our results.

2. Coalescence probabilities and times

In Sect. 2.1, we describe our model and derive recursion relations for the joint distribution of the place and time of coalescence and for its marginal distributions. In Sect. 2.2, we deduce invariance relations for these distributions and for the moments of the coalescence time. We obtain a simple, explicit strong-migration limit for the joint distribution in Sect. 2.3.

2.1. The recursion relations

Generations are discrete and nonoverlapping; the monoecious, diploid population is subdivided into a finite number of panmictic colonies that exchange gametes in a fixed pattern.

At the beginning of the life cycle, every one of the N_i adults in deme i produces the same very large number of gametes, which then disperse independently. Complete random union of gametes follows. Therefore, a proportion $1/N_i$ of the zygotes whose gametes originate in deme i are produced by self-fertilization. Population regulation returns the number of individuals in deme i to N_i. Thus, random genetic drift operates through population regulation.

Before deriving our recursion relations, we introduce some essential concepts and parameters.

Let m_{ij} designate the probability that a gamete in deme i after dispersion was produced in deme j. In the absence of selection, it is reasonable to assume that the backward migration matrix $M = (m_{ij})$ is constant (Nagylaki 1992, p. 135). We posit also that M is ergodic, *i.e.*, irreducible and aperiodic (Gantmacher 1959, pp. 50, 80, 88). Irreducibility guarantees that the descendants of individuals in each deme are able eventually to reach every other deme. Aperiodicity precludes pathological cyclic behavior. Given irreducibility, the biologically trivial condition that individuals have positive probability of remaining in some deme, *i.e.*, that $m_{ii} > 0$ for some i, suffices for aperiodicity (Feller 1968, p. 426). Of course, M must be stochastic:

$$\sum_j m_{ij} = 1 \ . \tag{2.1}$$

Let N_{T} and κ_i represent the total population number and the proportion of adults in deme i, respectively:

$$N_{\mathrm{T}} = \sum_i N_i, \quad \kappa_i = N_i/N_{\mathrm{T}} \ , \tag{2.2a}$$

$$0 < \kappa_i < 1, \quad \sum_i \kappa_i = 1 \ . \tag{2.2b}$$

By the ergodicity of the nonnegative stochastic matrix M, the eigenvalue 1 of M is simple and exceeds all other eigenvalues in absolute value; we can choose the left eigenvector ν corresponding to this unit eigenvalue to have only positive components (Gantmacher 1959, Chapter 13). Thus, the conditions

$$0 < \nu_i < 1, \quad \sum_i \nu_i = 1, \quad \nu^{\mathrm{T}} M = \nu^{\mathrm{T}} \ , \tag{2.3}$$

where the superscript T signifies matrix transposition, determine ν uniquely. Note that ν is the unique stationary distribution of the Markov chain with transition matrix M. This implies that if a gene is sampled from the population (with any distribution, *e.g.*, from deme i), the probability that its ancestral gene was in deme j t generations ago converges to ν_j as $t \to \infty$. The vector ν depends only on the relative migration rates, *i.e.*, if we replace m_{ij} by cm_{ij} for every i and j such that $i \neq j$, then ν is unaltered (Nagylaki 1998).

Conservative migration patterns are those that do not change the subpopulation numbers; in this case, and only in this case, we have $\nu = \kappa$ (Nagylaki 1980). Conservative migration has many simple intuitive properties that do not always hold for arbitrary migration (Nagylaki 1980, 1982, 1983, 1985, 1986, 1992, pp. 135–136, 151; Nordborg 1997). In our model, the subpopulation numbers N_i refer to adults. However, since the number of gametes in each deme before dispersion is proportional to N_i, it is also true that the gametic numbers are unchanged by conservative migration, and only by conservative migration.

The vectors κ and ν jointly determine our most important parameter, the migration effective population number N_e, defined by (Nagylaki 1980)

$$N_e = \beta N_T, \quad \beta = \left(\sum_i \nu_i^2 / \kappa_i \right)^{-1} . \tag{2.4}$$

We have $\beta \leq 1$ and hence $N_e \leq N_T$, with equality if and only if migration is conservative (Nagylaki 1980). This effective population number replaces the actual total population number in the strong-migration limit (Nagylaki 1980, 1983) and in certain aspects of geographical invariance (Nagylaki 1982, 1994, 1998). Observe that N_e is independent of the genetic model. If two distinct gametes are chosen at random, each with probability distribution ν, then the probability that the gametes are descended from the same parent is $1/N_e$ (Nordborg 1997; Nagylaki 1998).

We are now prepared for our analysis.

We first define the indicator random variable $I_{k,ij}$. If the most recent common ancestor (MRCA) of two distinct, homologous genes chosen at random from adults just before gametogenesis, one from deme i and the other from deme j, was in deme k, then $I_{k,ij} = 1$; otherwise $I_{k,ij} = 0$. If T_{ij} denotes the number of generations to the MRCA (*i.e.*, the coalescence time of the two genes), then

$$p_{k,ij}(t) = \mathscr{P}(I_{k,ij} = 1 \cap T_{ij} \leq t) \tag{2.5}$$

represents the joint probability that coalescence occurs in deme k and in at most t $(= 0, 1, 2, \ldots)$ generations. Clearly,

$$p_{k,ij}(0) = 0 \ . \tag{2.6}$$

The probabilities $p_{k,ij}(t)$ satisfy the recursion relation

$$\begin{aligned} p'_{k,ij} = \sum_{\ell,n:\ell\neq n} m_{i\ell} m_{jn} p_{k,\ell n} + \sum_{\ell:\ell\neq k} m_{i\ell} m_{j\ell} \left[\left(1 - \frac{1}{2N_\ell} \right) p_{k,\ell\ell} + \left(\frac{1}{2N_\ell} \right) (0) \right] \\ + m_{ik} m_{jk} \left[\left(1 - \frac{1}{2N_k} \right) p_{k,kk} + \left(\frac{1}{2N_k} \right) (1) \right] \ , \end{aligned} \tag{$*$}$$

in which the prime signifies the next generation. To understand $(*)$, observe first that the (ℓ, n) term in the first sum is the joint probability that the parental genes of the two genes sampled from demes i and j were in demes ℓ and n, respectively $(m_{i\ell} m_{jn})$, and that they coalesce in deme k in at most t generations $(p_{k,\ell n})$. Next, note that the ℓ term in the second sum is the joint probability that the parental genes were both in deme ℓ $(m_{i\ell} m_{j\ell})$ and that they coalesce in deme k in at most t generations; since $\ell \neq k$, the second probability is $p_{k,\ell\ell}(t)$ if the parental genes are distinct and 0 if they coincide. The last term is the joint probability that the parental genes were both in deme k $(m_{ik} m_{jk})$ and that they coalesce in deme k in at most t generations; the second probability is $p_{k,kk}(t)$ if the parental genes are distinct and 1 if they coincide. The recursion $(*)$ simplifies immediately to

$$p'_{k,ij} = \sum_{\ell,n} m_{i\ell} m_{jn} p_{k,\ell n} - \sum_{\ell} (2N_\ell)^{-1} m_{i\ell} m_{j\ell} p_{k,\ell\ell} + (2N_k)^{-1} m_{ik} m_{jk} \ . \tag{2.7}$$

Also of interest is the generating function ($0 \leq s < 1$)

$$P_{k,ij}(s) = \sum_{t=0}^{\infty} p_{k,ij}(t)s^t \quad . \tag{2.8}$$

Summing (2.7) from $t = 0$ to ∞ and appealing to (2.6), we obtain

$$P_{k,ij}(s) = s\left[\sum_{\ell,n} m_{i\ell}m_{jn}P_{k,\ell n}(s) - \sum_{\ell}(2N_\ell)^{-1}m_{i\ell}m_{j\ell}P_{k,\ell\ell}(s) + \left(\frac{1}{2N_k}\right)\frac{m_{ik}m_{jk}}{1-s}\right]. \tag{2.9}$$

The probability that coalescence occurs in deme k is

$$q_{k,ij} = \mathscr{P}(I_{k,ij} = 1) = \lim_{t\to\infty} p_{k,ij}(t); \tag{2.10}$$

by (2.7), this satisfies

$$q_{k,ij} = \sum_{\ell,n} m_{i\ell}m_{jn}q_{k,\ell n} - \sum_{\ell}(2N_\ell)^{-1}m_{i\ell}m_{j\ell}q_{k,\ell\ell} + (2N_k)^{-1}m_{ik}m_{jk} \quad . \tag{2.11}$$

Since our stochastic process is a finite, absorbing Markov chain, we have $\mathscr{P}(T_{ij} < \infty) = 1$, and therefore

$$\sum_k q_{k,ij} = 1 \tag{2.12}$$

for every i and j, which is consistent with (2.11). Notohara (1990) derived the equation corresponding to (2.11) for a sample of arbitrary size in a continuous-time model.

The probability that coalescence occurs in more than t generations is

$$r_{ij}(t) = \mathscr{P}(T_{ij} > t) = 1 - \sum_k p_{k,ij}(t) \quad , \tag{2.13}$$

with

$$r_{ij}(0) = 1 \quad . \tag{2.14}$$

Summing (2.7) over k gives

$$r'_{ij} = \sum_{\ell,n} m_{i\ell}m_{jn}r_{\ell n} - \sum_{\ell}(2N_\ell)^{-1}m_{i\ell}m_{j\ell}r_{\ell\ell} \quad . \tag{2.15}$$

From either (2.9) or (2.15) we deduce that the generating function

$$R_{ij}(s) = \sum_{t=0}^{\infty} r_{ij}(t)s^t \tag{2.16}$$

satisfies

$$R_{ij}(s) = 1 + s\left[\sum_{\ell,n} m_{i\ell}m_{jn}R_{\ell n}(s) - \sum_{\ell}(2N_\ell)^{-1}m_{i\ell}m_{j\ell}R_{\ell\ell}(s)\right] \quad . \tag{2.17}$$

The probability-generating function of the coalescence times is (Feller 1968, p. 265)

$$\chi_{ij}(s) = \mathscr{E}(s^{T_{ij}}) = 1 - (1-s)R_{ij}(s) \quad . \tag{2.18}$$

Substituting (2.18) into (2.17), we find

$$\chi_{ij}(s) = s\left\{\sum_{\ell,n} m_{i\ell}m_{jn}\chi_{\ell n}(s) + \sum_{\ell}(2N_\ell)^{-1}m_{i\ell}m_{j\ell}[1-\chi_{\ell\ell}(s)]\right\} \quad . \tag{2.19}$$

2.2. Invariance

Here we show that suitably defined averages and probabilities satisfy equations whose form is independent of population subdivision.

Let the random variables ξ, η, and ζ denote demes. We sample two distinct, homologous genes from the entire population and from the same deme according to the distributions

$$\mathscr{P}[(\xi,\eta) = (i,j)] = \nu_i\nu_j, \tag{2.20a}$$

$$\mathscr{P}[\zeta = i] = \beta\left(\frac{\nu_i^2}{\kappa_i}\right) \equiv b_i \quad . \tag{2.20b}$$

Then for any variables a_{ij} we define the averages (Nagylaki 1982)

$$\overline{a} \equiv \mathscr{E}(a_{\xi\eta}) = \sum_{i,j} a_{ij}\nu_i\nu_j, \tag{2.21a}$$

$$\overline{a}^0 \equiv \mathscr{E}(a_{\zeta\zeta}) = \sum_{i} a_{ii}b_i \quad . \tag{2.21b}$$

The vector b is normalized by (2.4). If migration is conservative, then $\nu = \kappa$, $\beta = 1$, and $b = \kappa$, so the sampling and averaging are according to the proportions κ of the population in each deme.

We derive our invariance formulas by averaging our recursion relations according to (2.21a). Applying (2.21a), (2.3), (2.2a), (2.20b), (2.4), and (2.21b) to (2.7) leads to

$$\overline{p}_k' = \overline{p}_k + \left(\frac{1}{2N_\mathrm{e}}\right)(b_k - \overline{p}_k^0) \quad . \tag{2.22}$$

Thus, the change $\Delta\overline{p}_k$ is only of order $1/(2N_\mathrm{e})$ even if migration is rapid. Observe that

$$\overline{p}_k(t) = \mathscr{P}(I_{k,\xi\eta} = 1 \cap T_{\xi\eta} \le t) \quad , \tag{2.23a}$$

$$\overline{p}_k^0(t) = \mathscr{P}(I_{k,\zeta\zeta} = 1 \cap T_{\zeta\zeta} \leq t) \ . \tag{2.23b}$$

Either by recalling (2.8) and summing (2.22) or by averaging (2.9) we find

$$2N_e(1-s)\overline{P}_k(s) = s\left[\frac{b_k}{1-s} - \overline{P}_k^0(s)\right] \ . \tag{2.24}$$

We let $t \to \infty$ in (2.22) and recall (2.10) to conclude that for two distinct genes sampled from the same deme according to (2.20b), the probability that coalescence occurs in deme k is

$$\overline{q}_k^0 = \mathscr{P}(I_{k,\zeta\zeta} = 1) = b_k \ . \tag{2.25}$$

If migration is conservative, this implies that for two distinct genes sampled from the same deme, chosen at random with probabilities κ, the probability that coalescence occurs in deme k is simply κ_k, the proportion of adults in deme k.

From (2.15) and (2.17) we deduce easily the relations

$$\overline{r}' = \overline{r} - \left(\frac{1}{2N_e}\right)\overline{r}^0 \ , \tag{2.26}$$

$$(1-s)\overline{R}(s) = 1 - \left(\frac{s}{2N_e}\right)\overline{R}^0(s). \tag{2.27}$$

Note that $\overline{r}$ and $\overline{r}^0$ have simple interpretations:

$$\overline{r}(t) = \mathscr{P}(T_{\xi\eta} > t), \quad \overline{r}^0(t) = \mathscr{P}(T_{\zeta\zeta} > t) \ . \tag{2.28}$$

Of most interest is the probability-generating function (2.18) because

$$\overline{\chi}(s) = \mathscr{E}(s^{T_{\xi\eta}}), \quad \overline{\chi}^0(s) = \mathscr{E}(s^{T_{\zeta\zeta}}) \ . \tag{2.29}$$

Averaging (2.18) gives

$$\overline{R}(s) = \frac{1-\overline{\chi}(s)}{1-s}, \quad \overline{R}^0(s) = \frac{1-\overline{\chi}^0(s)}{1-s} \ , \tag{2.30}$$

and we substitute this into (2.27) to infer

$$2N_e(1-s)\overline{\chi}(s) = s[1-\overline{\chi}^0(s)] \ . \tag{2.31}$$

From (2.31) we can derive directly invariance relations for the moments of the coalescence time. Differentiating (2.31) with respect to s and setting $s = 1$, we get immediately (Nagylaki 1998)

$$\mathscr{E}(T_{\zeta\zeta}) = 2N_e \ . \tag{2.32}$$

For the higher moments, we use Leibniz' formula (Abramowitz 1964, p. 12) to differentiate (2.31) n $(= 2, 3, \ldots)$ times and set $s = 1$:

$$2N_e n\frac{d^{n-1}\overline{\chi}}{ds^{n-1}}(1) = \frac{d^n\overline{\chi}^0}{ds^n}(1) + n\frac{d^{n-1}\overline{\chi}^0}{ds^{n-1}}(1) \ . \tag{2.33}$$

But the derivatives in (2.33) are factorial moments; *e.g.*,

$$\frac{d^n\overline{\chi}}{ds^n}(1) = \mathscr{E}\left[T_{\xi\eta}(T_{\xi\eta}-1)\dots(T_{\xi\eta}-n+1)\right] \equiv \mathscr{E}\left[(T_{\xi\eta})_n\right] \ . \tag{2.34}$$

Therefore, we obtain ($n = 2, 3, \dots$)

$$2N_{\mathrm{e}}n\mathscr{E}\left[(T_{\xi\eta})_{n-1}\right] = \mathscr{E}\left[(T_{\zeta\zeta})_n\right] + n\mathscr{E}\left[(T_{\zeta\zeta})_{n-1}\right] \ . \tag{2.35}$$

In particular, setting $n = 2$ and invoking (2.32), we find

$$4N_{\mathrm{e}}\mathscr{E}(T_{\xi\eta}) = 2N_{\mathrm{e}}(2N_{\mathrm{e}}+1) + \mathrm{Var}(T_{\zeta\zeta}) \ . \tag{2.36}$$

This is easily checked for panmixia.

The invariance relation (2.35) simplifies considerably for large subpopulation numbers. Suppose $N_i \to \infty$ for every i, with the migration matrix M (and therefore ν) and κ fixed. Then (2.32) suggests the scaling

$$T_{ij} = [2N_{\mathrm{e}}\widetilde{T}_{ij}] \ , \tag{2.37}$$

where the brackets signify the greatest integer. This reduces (2.32) and (2.35) to the asymptotic result ($n = 1, 2, \dots$)

$$n\mathscr{E}(\widetilde{T}_{\xi\eta}^{n-1}) \sim \mathscr{E}(\widetilde{T}_{\zeta\zeta}^{n}) \tag{2.38}$$

as $N_{\mathrm{e}} \to \infty$, which was proved by Herbots (1994, p. 132) for conservative migration at rates of order $1/N_{\mathrm{T}}$. Hence, the first two moments satisfy

$$\mathscr{E}(\widetilde{T}_{\zeta\zeta}) \sim 1, \quad 2\mathscr{E}(\widetilde{T}_{\xi\eta}) \sim 1 + \mathrm{Var}(\widetilde{T}_{\zeta\zeta}) \ . \tag{2.39}$$

2.3. The strong-migration limit

Here we establish an approximation for $p_{k,ij}(t)$ under the assumption that migration dominates random drift.

From (2.2a) and (2.4) we see at once that

$$\frac{1}{2N_\ell} = \frac{\beta}{2N_{\mathrm{e}}\kappa_\ell} \ . \tag{2.40}$$

We let $N_{\mathrm{e}} \to \infty$ with M and κ fixed. Defining the small parameter ϵ, the scaled time τ, and new variables $\pi_{k,ij}(\tau)$ by

$$\epsilon = \frac{1}{2N_{\mathrm{e}}}, \quad t = [2N_{\mathrm{e}}\tau], \quad \pi_{k,ij}(\tau) = p_{k,ij}(t) \tag{2.41}$$

shows that (2.7) has the form (A1) with $A = B = 0$ and $C = D = 1$. In view of (2.6), the initial condition is $\overline{p}_k(0) = 0$. Therefore, Lemma A implies that

$$\pi_{k,ij}(\tau) = \mathscr{P}(I_{k,ij} = 1 \cap T_{ij} \le [2N_{\mathrm{e}}\tau]) \to b_k(1 - e^{-\tau}) \tag{2.42}$$

as $N_{\mathrm{e}} \to \infty$ with $\tau > 0$ and fixed.

Thus, in the strong-migration limit, the place and time of coalescence are mutually independent, and their distributions are independent of the demes sampled. Letting $\tau \to \infty$ and summing over k in (2.42), we find

$$\mathscr{P}(I_{k,ij} = 1) \to b_k \ , \tag{2.43a}$$

$$\mathscr{P}(T_{ij} \le [2N_e\tau]) \to 1 - e^{-\tau} \tag{2.43b}$$

as $N_e \to \infty$. The limit (2.43a) agrees with (2.25); the limit (2.43b) is the panmictic exponential distribution on a time scale of $2N_e$ generations.

Notohara (1993) studied samples of arbitrary size in a continuous-time model. He established (2.43a) and derived the strong-migration limit of the mean coalescence time.

3. The model of infinitely many alleles

Suppose every allele mutates to new alleles at the same rate u $(0 < u < 1)$. Mutation may occur at any time between gametogenesis and population regulation.

3.1. The recursion relations

Let $f_{ij}(t)$ designate the probability that two distinct, homologous genes chosen at random from adults just before gametogenesis in generation t, one from colony i and the other from colony j, are the same allele. These probabilities satisfy (Malécot 1951, 1975; Nagylaki 1976, 1980, 1983; Sawyer 1976)

$$f'_{ij} = v\left[\sum_{\ell,n} m_{i\ell}m_{jn}f_{\ell n} + \sum_{\ell}(2N_\ell)^{-1}m_{i\ell}m_{j\ell}(1 - f_{\ell\ell})\right] \ , \tag{3.1}$$

where $v = (1-u)^2$.

Comparing (3.1) with (2.19) shows that at equilibrium the probabilities of identity are given by

$$\hat{f}_{ij} = \chi_{ij}(v) = \mathscr{E}(v^{T_{ij}}) \ , \tag{3.2}$$

which is probabilistically obvious. If we define $\theta = 4N_e u$ and let $N_e \to \infty$ with θ fixed, from (3.2) and (2.37) we get (*cf.* Hudson 1990; Herbots 1994; Wilkinson-Herbots 1998; Bahlo and Griffiths 1999)

$$\hat{f}_{ij} \sim \mathscr{E}(e^{-\theta\tilde{T}_{ij}}) \ . \tag{3.3}$$

Although the coalescence times are more fundamental than the probabilities of identity, the latter were introduced and investigated first. Thus, in many cases, the probability-generating function (3.2) and the Laplace transform (3.3) of the coalescence times can be deduced directly from the large literature on $\hat{f}_{ij}$, reviewed in Nagylaki (1983, 1986, 1989).

3.2. Invariance

Averaging (3.1) yields our invariance relation (Nagylaki 1982)

$$\overline{f}' = v\left[\overline{f} + \left(\frac{1}{2N_e}\right)(1 - \overline{f}^0)\right] . \tag{3.4}$$

Note that $\overline{f}$ generally evolves slowly even if migration is rapid. At equilibrium, this gives (Nagylaki 1982)

$$\hat{\overline{f}} = \frac{v(1 - \hat{\overline{f}}^0)}{2N_e(1 - v)} \approx \frac{1 - \hat{\overline{f}}^0}{\theta} \quad , \tag{3.5}$$

in which the approximation is valid if $u \ll 1$.

To interpret $\overline{f}(t)$ and $\overline{f}^0(t)$ probabilistically, let the indicator variable $J_{ij}(t) = 1$ if the genes sampled from demes i and j in generation t are the same allele; otherwise $J_{ij}(t) = 0$. Then

$$\overline{f}(t) = \mathscr{P}\left[J_{\xi\eta}(t) = 1\right], \quad \overline{f}^0(t) = \mathscr{P}[J_{\zeta\zeta}(t) = 1] \ . \tag{3.6}$$

3.3. The strong-migration limit

We let $N_e \to \infty$ with M, κ, and θ fixed. Recalling (2.40) and (2.41) and setting $\phi_{ij}(\tau) = f_{ij}(t)$, we see that (3.1) has the form (A1) with $A = \theta$, $B = C = 1$, and $D = 0$. Therefore, Lemma A implies that

$$\phi_{ij}(\tau) \to \frac{1}{1+\theta} + \left[\overline{f}(0) - \frac{1}{1+\theta}\right] e^{-(1+\theta)\tau} \tag{3.7}$$

as $N_e \to \infty$ with $\tau > 0$ and fixed, which is precisely the panmictic result (Malécot 1946, 1948; Kimura and Crow 1964). Note that the appropriate average must be used for the initial condition.

This strong-migration limit was derived for $\tau \gg 1$ in Nagylaki (1980) by a perturbation argument, but the constant that multiplies the exponential was not evaluated.

4. The model of infinitely many sites

Here we investigate a gene or DNA sequence that consists of infinitely many nucleotide sites without recombination. Thus, we posit that the mutation rate per site is so low that mutation occurs at each site at most once and then only at monomorphic sites. This approximation requires that the proportion of polymorphic sites be much less than one. Let u denote the total mutation rate per gene. We assume that mutation occurs between gametic dispersion and population regulation.

4.1. The recursion relations

Let $K_{ij}(t)$ designate the number of nucleotide differences between two distinct, homologous genes chosen at random from adults just before gametogenesis in generation t, one from colony i and the other from colony j. We seek a recursion relation for the probability-generating function

$$w_{ij}(s,t) = \mathscr{E}[s^{K_{ij}(t)}] \ . \tag{4.1}$$

Let $\widetilde{K}_{ij}(t)$ denote the number of nucleotide differences after gametic dispersion but before mutation in generation t. The number of nucleotide differences added in generation t by mutation, $L_{ij}(t)$, has a Poisson distribution with mean $2u$. The random variables $\widetilde{K}_{ij}(t)$ and $L_{ij}(t)$ are mutually stochastically independent.

Since

$$K_{ij}(t+1) = \widetilde{K}_{ij}(t) + L_{ij}(t) \ , \tag{4.2}$$

we have

$$\begin{aligned} w_{ij}(s,t+1) &= \mathscr{E}[s^{\widetilde{K}_{ij}(t)}]\mathscr{E}[s^{L_{ij}(t)}] \\ &= e^{2u(s-1)}\left\{ \sum_{\ell,n:\ell\neq n} m_{i\ell}m_{jn}\mathscr{E}[s^{K_{\ell n}(t)}] \right. \\ &\quad \left. + \sum_{\ell} m_{i\ell}m_{j\ell}\left[\left(1-\frac{1}{2N_\ell}\right)\mathscr{E}[s^{K_{\ell\ell}(t)}] + \left(\frac{1}{2N_\ell}\right)(1)\right]\right\} \ . \end{aligned} \tag{4.3}$$

Substituting (4.1) into (4.3) and simplifying, we obtain immediately

$$\begin{aligned} w_{ij}(s,t+1) = e^{2u(s-1)} &\left\{ \sum_{\ell,n} m_{i\ell}m_{jn}w_{\ell n}(s,t) \right. \\ &\left. + \sum_{\ell} (2N_\ell)^{-1} m_{i\ell}m_{j\ell}[1-w_{\ell\ell}(s,t)]\right\} \ . \end{aligned} \tag{4.4}$$

We can easily relate (4.4) to (3.1). Denoting the solution of (3.1) as $f_{ij}(v,t)$ and defining $f_{ij}(\cdot,0) = w_{ij}(s,0)$ for every i and j, we see at once that

$$w_{ij}(s,t) = f_{ij}(e^{-2u(1-s)},t) \tag{4.5}$$

for every i, j, and t, *i.e.*, the mutation rate in (3.1) must be replaced by

$$1 - e^{-u(1-s)} \sim u(1-s) \ , \tag{4.6}$$

where the asymptotic form holds as $u \to 0$. Thus, many results for (4.4) can be deduced at once from the large literature on (3.1), reviewed in Nagylaki (1983, 1986, 1989). Recalling (3.2), at equilibrium we get

$$\hat{w}_{ij}(s) = \hat{f}_{ij}(e^{-2u(1-s)}) = \chi_{ij}(e^{-2u(1-s)}) = \mathscr{E}(e^{-2u(1-s)T_{ij}}) \ , \tag{4.7}$$

in which the last expression follows directly from the Poisson input of mutations (*cf.* Griffiths 1981; Bahlo and Griffiths 1999).

Notohara (1997) derived the differential equation corresponding to (4.4) and the formula corresponding to (4.7) for a sample of arbitrary size in a continuous-time model.

4.2. Invariance

By either averaging (4.4) or inserting (4.5) into (3.4), we obtain

$$\overline{w}(s,t+1)=e^{2u(s-1)}\left\{\overline{w}(s,t)+\left(\frac{1}{2N_\mathrm{e}}\right)[1-\overline{w}^0(s,t)]\right\}\ , \tag{4.8}$$

where

$$\overline{w}(s,t)=\mathscr{E}[s^{K_{\xi\eta}(t)}],\quad \overline{w}^0(s,t)=\mathscr{E}[s^{K_{\zeta\zeta}(t)}]\ . \tag{4.9}$$

Thus, $\overline{w}$ is the slow dependent variable. At equilibrium, (4.8) reduces to

$$2N_\mathrm{e}[e^{2u(1-s)}-1]\hat{\overline{w}}(s)=1-\hat{\overline{w}}^0(s)\ . \tag{4.10}$$

We can derive an invariance relation for the factorial moments of $\hat{K}_{\xi\eta}$ and $\hat{K}_{\zeta\zeta}$ by invoking Leibniz' formula to differentiate (4.10) n times and setting $s=1$, as in Section 2.2. We find ($n=1,2,\ldots$)

$$\mathscr{E}[(\hat{K}_{\zeta\zeta})_n]=-2N_\mathrm{e}\sum_{k=1}^{n}\binom{n}{k}(-2u)^k\mathscr{E}[(\hat{K}_{\xi\eta})_{n-k}]\ , \tag{4.11}$$

in which $(\hat{K}_{\xi\eta})_0=1$. In particular, the mean number of nucleotide differences is (Nagylaki 1998)

$$\mathscr{E}(\hat{K}_{\zeta\zeta})=\theta\ , \tag{4.12}$$

where $\theta=4N_\mathrm{e}u$. Taking $n=2$, we deduce

$$2\theta\mathscr{E}(\hat{K}_{\xi\eta})=\theta(\theta-1+2u)+\mathrm{Var}(\hat{K}_{\zeta\zeta})\ . \tag{4.13}$$

As a check, after a change in notation we can easily verify that Li's (1976) results for the island model satisfy (4.13). It is also easy to verify that (4.13) holds for a panmictic population.

In the limit $N_\mathrm{e}\to\infty$ with θ fixed, only $k=1$ contributes to (4.11):

$$\mathscr{E}[(\hat{K}_{\zeta\zeta})_n]\sim n\theta\mathscr{E}[(\hat{K}_{\xi\eta})_{n-1}]\ . \tag{4.14}$$

We can confirm (4.14) by observing that the asymptotic form of (4.10) is

$$\theta(1-s)\hat{\overline{w}}(s)\sim 1-\hat{\overline{w}}^0(s)\ . \tag{4.15}$$

For $n=2$, from (4.14) we get (4.13) without the $2u$.

4.3. The strong-migration limit

We let $N_e \to \infty$ with M, κ, and θ fixed. Recalling (2.40) and (2.41) and setting $\omega_{ij}(s,\tau) = w_{ij}(s,t)$, we see that (4.4) has the form (A1) with $A = \theta(1-s)$, $B = C = 1$, and $D = 0$. Therefore, Lemma A implies that

$$\omega_{ij}(s,\tau) \to \hat{\omega}(s) + [\overline{w}(s,0) - \hat{\omega}(s)]e^{-\tau/\hat{\omega}(s)} \tag{4.16a}$$

as $N_e \to \infty$ with s and τ fixed and $\tau > 0$, where

$$\hat{\omega}(s) = \frac{1}{1+\theta(1-s)} \,. \tag{4.16b}$$

The appropriate average must be used for the initial condition. This result can be confirmed at once by observing that, in the strong-migration limit, the substitution (4.6) instructs us to replace θ by $\theta(1-s)$, which is precisely the connection between (3.7) and (4.16).

The panmictic equilibrium formula (4.16b) is due to Watterson (1975); we owe the time-dependent solution for panmixia to Li (1977). Notohara (1997) derived the strong-migration limit at equilibrium for a sample of arbitrary size in a continuous-time model.

5. Some related results

The diffusion approximation has been used to establish invariance relations (Nagylaki 1982) and the strong-migration limit (Nagylaki 1980) in some models with finitely many alleles that include selection. Here we analyze two further examples with the methods developed in this paper. First, we treat the model of infinitely many alleles in the more complicated and realistic case of diploid migration. Second, we prove that the conditional fixation time of a single, suitably sampled allele and the time to the MRCA of the entire population have the same probability distribution.

5.1. Diploid migration

We investigate the model of infinitely many alleles with diploid migration by modifying Section 3. We assume that there is no selfing.

Let $I_{ij}(t)$ denote the probability that two homologous genes chosen at random from distinct adults just before gametogenesis in generation t, one from deme i and the other from deme j, are the same allele. We designate by $J_i(t)$ the probability that the two genes of an adult chosen at random from deme i just before gametogenesis in generation t are the same allele. These probabilities satisfy (Sawyer 1976; Nagylaki 1983)

$$I'_{ij} = v\left[\sum_{\ell,n} m_{i\ell}m_{jn}I_{\ell n} + \sum_{\ell}(2N_\ell)^{-1}m_{i\ell}m_{j\ell}(1+J_\ell-2I_{\ell\ell})\right], \tag{5.1a}$$

$$J'_i = v\sum_{\ell} m_{i\ell}I_{\ell\ell} \,. \tag{5.1b}$$

We average (5.1a) according to (2.21):

$$\overline{I}' = v\left[\overline{I} + \frac{1}{2N_{\rm e}}(1 + \overline{J}^0 - 2\overline{I}^0)\right] \ . \tag{5.2}$$

Thus, $\overline{I}$ is the slow variable in this model. The special use of (5.2) for conservative migration was derived in Nagylaki (1985). At equilibrium, we get

$$\hat{\overline{I}} = \frac{1 + \hat{\overline{J}}^0 - 2\hat{\overline{I}}^0}{2N_{\rm e}(1-v)} \ . \tag{5.3}$$

Our proof of the strong-migration limit is an extension of that in the Appendix. We let $N_{\rm e} \to \infty$ with M, κ, and θ fixed; recall (2.40) and (2.41); and define

$$\widetilde{I}_{ij}(\tau) = I_{ij}(t), \quad \widetilde{J}_i(\tau) = J_i(t) \ . \tag{5.4}$$

Then (5.1) becomes

$$\widetilde{I}_{ij}(\tau+\epsilon) = (1-\tfrac{1}{2}\theta\epsilon)^2\Bigg\{\sum_{\ell,n} m_{i\ell}m_{jn}\widetilde{I}_{\ell n}(\tau) + \beta\epsilon\sum_{\ell}\kappa_\ell^{-1}m_{i\ell}m_{j\ell}[1+\widetilde{J}_\ell(\tau)-2\widetilde{I}_{\ell\ell}(\tau)]\Bigg\}, \tag{5.5a}$$

$$\widetilde{J}_i(\tau+\epsilon) = (1-\tfrac{1}{2}\theta\epsilon)^2\sum_{\ell} m_{i\ell}\widetilde{I}_{\ell\ell}(\tau) \ . \tag{5.5b}$$

We assume that there exist $I^*_{ij}(\tau)$ and $J^*_i(\tau)$, independent of ϵ, such that

$$\widetilde{I}_{ij}(\tau) \to I^*_{ij}(\tau), \quad \widetilde{J}_i(\tau) \to J^*_i(\tau) \ , \tag{5.6a}$$

$$\epsilon^{-1}[\overline{\widetilde{I}}(\tau+\epsilon) - \overline{\widetilde{I}}(\tau)] \to \frac{d\overline{I}^*}{d\tau}(\tau) \tag{5.6b}$$

as $\epsilon \to 0$ with $\tau > 0$ and fixed. In the limit $\epsilon \to 0$, from (5.5a) we get

$$I^*_{ij}(\tau) = \sum_{\ell,n} m_{i\ell}m_{jn}I^*_{\ell n}(\tau), \quad \tau > 0 \ . \tag{5.7}$$

By the argument below (A8), we infer that

$$I^*_{ij}(\tau) = I^*(\tau), \quad \tau > 0, \tag{5.8}$$

independent of i and j. We now let $\epsilon \to 0$ in (5.5b) and appeal to (5.8):

$$J^*_i(\tau) = I^*(\tau), \quad \tau > 0 \ . \tag{5.9}$$

We rewrite (5.2) in the form

$$\overline{\widetilde{I}}(\tau+\epsilon) = (1-\tfrac{1}{2}\theta\epsilon)^2\{\overline{\widetilde{I}}(\tau) + \epsilon[1 - \overline{\widetilde{J}}^0(\tau) - 2\overline{\widetilde{I}}^0(\tau)]\} \ . \tag{5.10}$$

As $\epsilon \to 0$, we obtain from (5.10) the differential equation

$$\frac{d\overline{I}^*}{d\tau} = -\theta\overline{I}^* + 1 + \overline{J}^{0*} - 2\overline{I}^{0*}, \quad \tau > 0 \ , \tag{5.11}$$

and (5.8) and (5.9) reduce (5.11) to

$$\frac{dI^*}{d\tau} = 1 - (1+\theta)I^*, \quad \tau > 0 \ . \tag{5.12}$$

It remains only to determine the initial condition for (5.12). Taking $\tau = 0$ in (5.10), we see that

$$\overline{\widetilde{I}}(\epsilon) = \overline{\widetilde{I}}(0) + O(\epsilon) \ , \tag{5.13}$$

whence

$$I^*(0+) = \overline{I}^*(0+) = \overline{\widetilde{I}}(0) = \overline{I}(0) \ . \tag{5.14}$$

We conclude that $\widetilde{I}_{ij}(\tau) \to I^*(\tau)$ and $\widetilde{J}_i(\tau) \to I^*(\tau)$ for every i and j as $\epsilon \to 0$ with $\tau > 0$ and fixed, where

$$I^*(\tau) = \frac{1}{1+\theta} + \left[\overline{I}(0) - \frac{1}{1+\theta}\right] e^{-(1+\theta)\tau} \ . \tag{5.15}$$

This result was proved for $\tau \gg 1$ in Nagylaki (1983) by a perturbation argument, but the constant that multiplies the exponential was not evaluated. The simple, explicit formula (5.15) demonstrates that only the probability of identity I_{ij} between individuals affects the strong-migration limit. Since the deme sizes are very large, the expected homozygosity J_i does not enter.

5.2. The conditional fixation time

Here we change some of the above notation. Let n designate the number of demes and write $L_k = 2N_k$ for the number of genes in deme k. We introduce the vector $e^{(\ell)}$ with components $e_k^{(\ell)} = \delta_{k\ell}$ for each $\ell, k = 1, 2, \ldots, n$, where $\delta_{k\ell}$ signifies the Kronecker delta ($\delta_{k\ell} = 1$ if $k = \ell$; otherwise $\delta_{k\ell} = 0$).

Suppose there are i_k A alleles in deme k and form the vector of gene numbers $i = (i_1, i_2, \ldots, i_n)$. We write the probability that t generations in the past the ancestral gene numbers were $j = (j_1, j_2, \ldots, j_n)$ as $h_{ij}^{(t)}$. Then the probability distribution function of the time $\widetilde{T}$ to the MRCA of the entire population is given by

$$\mathscr{P}(\widetilde{T} \le t) = \sum_{\ell=1}^{n} h_{Le^{(\ell)}}^{(t)} \ , \tag{5.16}$$

in which $L = (L_1, L_2, \ldots, L_n)$.

Now we assume that initially there is exactly one copy of the A allele in the population and it is in deme ℓ. If $X_k(t)$ represents the number of A alleles in deme k t generations after A has appeared in the population, we have $X(0) = e^{(\ell)}$. The allele A will be absorbed (*i.e.*, fixed or lost) in finite time with probability one. Let T and T^* denote the absorption and conditional fixation time of A, respectively.

The fixation probability of A is the average of its initial frequencies with respect to the stationary distribution ν (Nagylaki 1980, pp. 110, 113):

$$\mathscr{P}[X(T) = L \mid X(0) = e^{(\ell)}] = \frac{\nu_\ell}{L_\ell} \ . \tag{5.17}$$

We evaluate the probability distribution of T^* as follows:

$$\begin{aligned}\mathscr{P}[T^* \le t \mid X(0) = e^{(\ell)}] &= \mathscr{P}[X(t) = L \mid X(T) = L \quad \text{and} \quad X(0) = e^{(\ell)}] \\ &= \frac{\mathscr{P}[X(t) = L \mid X(0) = e^{(\ell)}]}{\mathscr{P}[X(T) = L \mid X(0) = e^{(\ell)}]} \\ &= \frac{h^{(t)}_{Le^{(\ell)}}}{L_\ell} \Bigg/ \frac{\nu_\ell}{L_\ell} = \frac{h^{(t)}_{Le^{(\ell)}}}{\nu_\ell} \ . \end{aligned} \tag{5.18}$$

From (5.16) and (5.18) we infer

$$\sum_{\ell=1}^{n} \nu_\ell \mathscr{P}[T^* \le t \mid X(0) = e^{(\ell)}] = \mathscr{P}(\widetilde{T} \le t) \ . \tag{5.19}$$

Writing T^*_ℓ for the conditional fixation time of a single A allele in deme ℓ and defining the random variable λ so that $\mathscr{P}(\lambda = \ell) = \nu_\ell$, we obtain

$$\mathscr{P}(T^*_\lambda \le t) = \mathscr{P}(\widetilde{T} \le t) \ , \tag{5.20}$$

i.e., T^*_λ and $\widetilde{T}$ have the same probability distribution.

If migration is conservative, then $\nu = \kappa$ and therefore (5.20) shows that the conditional fixation time of a new mutant (which appears in deme ℓ with probability κ_ℓ) and the time to the MRCA have the same distribution. Of course, (5.20) applies directly to a panmictic population.

6. Discussion

Here we summarize and discuss our main results.

We have formulated exact models with discrete, nonoverlapping generations. In most of the literature on coalescents in a subdivided population, a continuous-time model is used, either directly or as a limit. This requires large deme sizes and migration rates of the order of reciprocal deme sizes. Without large deme sizes, the elegance and power of coalescents is lost and the investigation of samples of more than two genes becomes too difficult. Although requiring migration rates to be much less than one may sometimes be biologically restrictive, some analyses may be unfeasible without this further approximation. For the strong-migration limit, however, this approximation is conceptually unnecessary, and one hopes that results can be derived for samples of arbitrary size without it.

The probability distribution, $p_{k,ij}(t)$ in (2.5), of the place and time of the MRCA of two genes satisfies the recursion relation (2.7). From this follow the equations (2.11) for the distribution, $q_{k,ij}$ in (2.10), of the place of coalescence and (2.19) for the probability-generating function, $\chi_{ij}(s)$ in (2.18), of the coalescence time T_{ij}.

If suitably defined averages or probabilities satisfy an equation whose form is independent of population subdivision, we say there is invariance. Our averages are defined in (2.21). Note the crucial role of the stationary distribution ν of the backward migration matrix M. If migration is conservative, the averaging is simply with respect to the proportions κ of individuals in each deme. Our general invariance relation is (2.22), which yields the simple, explicit formula (2.25) for the distribution of the place of coalescence of two distinct genes sampled from the same deme according to (2.20b). The invariance relation for the probability-generating function of the coalescence times is (2.31). This yields the invariance relation (2.35) for the factorial moments; the first two moments satisfy (2.32) and (2.36), respectively. In a large population, the moment relations reduce to (2.38).

We say migration is strong if it dominates all the other evolutionary forces. The probabilities $p_{k,ij}(t)$ have the simple, explicit strong-migration limit (2.42). The place and time of coalescence are asymptotically independent and have the distributions (2.43a) and (2.43b), respectively. The only effect of population subdivision on (2.43b) is that the migration effective population number N_e determines the time scale.

For the model of infinitely many alleles, the probabilities of identity in state, $f_{ij}(t)$, satisfy the recursion and invariance relations and strong-migration limit (3.1), (3.4), and (3.7), respectively. Observe that in the simple, explicit result (3.7), the appropriate average must be used for the initial condition, and population subdivision enters only through N_e. Equations (3.2) and (3.3) permit the calculations of the probability-generating function and the Laplace transform of the coalescence times from the extensive literature on the equilibrium probability of identity.

In the model of infinitely many sites without recombination, the probability-generating function of the number of nucleotide differences between two genes, $w_{ij}(s,t)$, satisfies the recursion and invariance relations (4.4) and (4.8), respectively. According to (4.5), one can obtain $w_{ij}(s,t)$ from the probability of identity by a simple substitution. At equilibrium, the invariance relation reduces to (4.10), which yields (4.11) for the factorial moments; the first two moments satisfy (4.12) and (4.13), respectively. In a large population with weak mutation, the factorial moments satisfy (4.14) at equilibrium. The strong-migration limit is (4.16); again, the initial condition is the appropriate average, and population subdivision enters only through N_e.

For the model of infinitely many alleles with diploid migration, the recursion and invariance relations and strong-migration limit are (5.1), (5.2), and (5.15), respectively. Note that, with the appropriate initial condition, the limit (5.15) is identical to the limit (3.7) for gametic dispersion.

Finally, (5.20) shows that the conditional fixation time of a single, suitably sampled allele and the time to the MRCA of the entire population have the same probability distribution. If migration is conservative, the sampling becomes identical to the appearance of a new mutant in the population.

Acknowledgements. I am very grateful to Prof. Simon Tavaré for a helpful discussion related to Section 5.2.

Appendix. The strong-migration limit

The strong-migration limits in Sections 2–4 follow immediately from the lemma below. Note that the simple limit (A4) is independent of i and j.

Lemma A. *Suppose the variables $g_{k,ij}(t)$ satisfy the recursion relation*

$$g'_{k,ij} = (1-A\epsilon)\sum_{\ell,n} m_{i\ell}m_{jn}g_{k,\ell n} + \beta\epsilon\sum_{\ell}\kappa_\ell^{-1}m_{i\ell}m_{j\ell}(B - Cg_{k,\ell\ell}) + D\beta\epsilon\kappa_k^{-1}m_{ik}m_{jk} + O(\epsilon^2) \tag{A1}$$

as $\epsilon \to 0$ with M and κ fixed, where A, B, C, and D are constants. Let

$$t = [\tau/\epsilon], \quad g_{k,ij}(t) = \gamma_{k,ij}(\tau), \tag{A2}$$

and assume that there exist $\gamma^*_{k,ij}(\tau)$, independent of ϵ, such that

$$\gamma_{k,ij}(\tau) \to \gamma^*_{k,ij}(\tau), \quad \epsilon^{-1}[\overline{\gamma}_k(\tau+\epsilon) - \overline{\gamma}_k(\tau)] \to \frac{d\overline{\gamma}^*_k}{d\tau}(\tau) \tag{A3}$$

as $\epsilon \to 0$ with $\tau > 0$ and fixed. Then

$$\gamma^*_{k,ij}(\tau) = \hat{\gamma}_k + [\overline{g}_k(0) - \hat{\gamma}_k]e^{-(A+C)\tau} \ , \tag{A4a}$$

in which

$$\hat{\gamma}_k = \frac{B + Db_k}{A + C} \ . \tag{A4b}$$

Proof. We use (2.21), (2.3), and (2.20b) to average (A1):

$$\overline{g}'_k = (1 - A\epsilon)\overline{g}_k + \epsilon(B - C\overline{g}^0_k + Db_k) + O(\epsilon^2) \ . \tag{A5}$$

Thus, $\overline{g}_k$ is our slow dependent variable. In terms of τ, the recursions (A1) and (A5) become

$$\gamma_{k,ij}(\tau+\epsilon) = (1-A\epsilon)\sum_{\ell,n} m_{i\ell}m_{jn}\gamma_{k,\ell n}(\tau) + \beta\epsilon\sum_{\ell}\kappa_\ell^{-1}m_{i\ell}m_{j\ell}[B - C\gamma_{k,\ell\ell}(\tau)] + D\beta\epsilon\kappa_k^{-1}m_{ik}m_{jk} + O(\epsilon^2), \tag{A6}$$

$$\overline{\gamma}_k(\tau+\epsilon) = (1-A\epsilon)\overline{\gamma}_k(\tau) + \epsilon[B - C\overline{\gamma}^0_k(\tau) + Db_k] + O(\epsilon^2) \ . \tag{A7}$$

Of course, $\gamma_{k,ij}(\tau)$ depends also on ϵ.

Now let $\epsilon \to 0$ and posit (A3). From (A6) we get

$$\gamma^*_{k,ij}(\tau) = \sum_{\ell,n} m_{i\ell}m_{jn}\gamma^*_{k,\ell n}(\tau), \quad \tau > 0 \ . \tag{A8}$$

By the ergodicity of M, the nonnegative, stochastic Kronecker-product matrix $M \otimes M$ in (A8) has a simple maximal eigenvalue 1, and the corresponding right eigenvector has equal components (Gantmacher 1959, Chapter 13). Therefore,

$$\gamma^*_{k,ij}(\tau) = \gamma^*_k(\tau), \quad \tau > 0 \ , \tag{A9}$$

independent of i and j.

Rearranging (A7) yields

$$\epsilon^{-1}[\overline{\gamma}_k(\tau+\epsilon) - \overline{\gamma}_k(\tau)] = -A\overline{\gamma}_k(\tau) + B - C\overline{\gamma}^0_k(\tau) + Db_k + O(\epsilon) \ , \tag{A10}$$

whence we obtain

$$\frac{d\overline{\gamma}^*_k}{d\tau} = -A\overline{\gamma}^*_k + B - C\overline{\gamma}^{0*}_k + Db_k, \quad \tau > 0 \ , \tag{A11}$$

and (A9) reduces this to

$$\frac{d\gamma^*_k}{d\tau} = -(A+C)\gamma^*_k + B + Db_k, \quad \tau > 0 \ . \tag{A12}$$

It remains only to determine the initial condition for the simple differential equation (A12). Taking $\tau = 0$ in (A7), we see that

$$\overline{\gamma}_k(\epsilon) = \overline{\gamma}_k(0) + O(\epsilon), \tag{A13}$$

whence

$$\gamma^*_k(0+) = \overline{\gamma}^*_k(0+) = \overline{\gamma}_k(0) = \overline{g}_k(0) \ . \tag{A14}$$

The solution of (A12) with the initial condition (A14) is precisely (A4), which completes the proof.

References

Abramowitz, M.: Elementary analytical methods. In: Abramowitz, M., Stegun, I.A. (eds.) Handbook of Mathematical Functions, pp. 9–63. Washington: National Bureau of Standards 1964

Bahlo, M., Griffiths, R.C.: Coalescence time for two genes from a subdivided population. Submitted for publication (1999)

Feller, W.: An Introduction to Probability Theory and Its Applications, Vol. I, 3rd edn. New York: Wiley 1968

Gantmacher, F.R.: The Theory of Matrices, Vol. II. New York: Chelsea 1959

Griffiths, R.C.: The number of heterozygous loci between two randomly chosen completely linked sequences of loci in two subdivided population models. J. Math. Biol. **12**, 251–261 (1981)

Herbots, H.M.: Stochastic models in population genetics: genealogy and genetic differentiation in structured populations. Dissertation, University of London 1994

Hudson, R.R.: Gene genealogies and the coalescent process. In: Futuyma, D.J., Antonovics, J. (eds.) Oxford Surveys in Evolutionary Biology (vol. 7, pp. 1-44) Oxford: Oxford University Press 1990

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

Li, W.-H.: Distribution of nucleotide differences between two randomly chosen cistrons in a subdivided population: the finite island model. Theor. Pop. Biol. **10**, 303–308 (1976)

Li, W.-H.: Distribution of nucleotide differences between two randomly chosen cistrons in a finite population. Genetics **85**, 331–337 (1977)

Malécot, G.: La consanguinité dans une population limitée. C.R. Acad. Sci. Paris **222**, 841–842 (1946)
Malécot, G.: Les mathématiques de l'hérédité. Paris: Masson 1948
Malécot, G.: Un traitement stochastique des problèmes linéaires (mutation, linkage, migration) en Génétique de Population. Ann. Univ. Lyon Sci. Sec. A **14**, 79–117 (1951)
Malécot, G.: Heterozygosity and relationship in regularly subdivided populations. Theor. Pop. Biol. **8**, 212–241 (1975)
Nagylaki, T.: The decay of genetic variability in geographically structured populations. II. Theor. Pop. Biol. **10**, 70–82 (1976)
Nagylaki, T.: The strong-migration limit in geographically structured populations. J. Math. Biol. **9**, 101–114 (1980)
Nagylaki, T.: Geographical invariance in population genetics. J. Theor. Biol. **99**, 159–172 (1982)
Nagylaki, T.: The robustness of neutral models of geographical variation. Theor. Pop. Biol. **24**, 268–294 (1983)
Nagylaki, T.: Homozygosity, effective number of alleles, and interdeme differentiation in subdivided populations. Proc. Natl. Acad. Sci. USA **82**, 8611–8613 (1985)
Nagylaki, T.: Neutral models of geographical variation. In: Tautu, P. (ed.) Stochastic Spatial Processes (Lecture Notes in Mathematics, vol. 1212, pp. 216–237) Berlin: Springer 1986
Nagylaki, T.: Gustave Malécot and the transition from classical to modern population genetics. Genetics **122**, 253–268 (1989)
Nagylaki, T.: Introduction to Theoretical Population Genetics (Biomathematics, vol. 21) Berlin: Springer 1992
Nagylaki, T.: Geographical variation in a quantitative character. Genetics **136**, 361–381 (1994)
Nagylaki, T.: The expected number of heterozygous sites in a subdivided population. Genetics **149**, 1599–1604 (1998)
Nordborg, M.: Structured coalescent processes on different time scales. Genetics **146**, 1501–1514 (1997)
Notohara, M.: The coalescent and the genealogical process in geographically structured population. J. Math. Biol. **29**, 59–75 (1990)
Notohara, M.: The strong-migration limit for the genealogical process in geographically structured populations. J. Math. Biol. **31**, 115–122 (1993)
Notohara, M.: The number of segregating sites in a sample of DNA sequences from a geographically structured population. J. Math. Biol. **36**, 188–200 (1997)
Notohara, M.: A perturbation method for the genealogical process in a geographically structured population with strong migration. Submitted for publication (1999)
Sawyer, S.: Results for the stepping-stone model for migration in population genetics. Ann. Prob. **4**, 699–728 (1976)
Slatkin, M.: The average number of sites separating DNA sequences drawn from a subdivided population. Theor. Pop. Biol. **32**, 42–49 (1987)
Strobeck, C.: Average number of nucleotide differences in a sample from a single subpopulation: a test for population subdivision. Genetics **117**, 149–153 (1987)
Watterson, G.A.: On the number of segregating sites in genetical models without recombination. Theor. Pop. Biol. **7**, 256–276 (1975)
Wilkinson-Herbots, H.M.: Genealogy and subpopulation differentiation under various models of population structure. J. Math. Biol. **37**, 535–585 (1998)