Genet. Res., Camb. (1996), **67**, *pp.* 147–158 *With 6 text-figures*

Estimating gene flow in island populations

BRUCE RANNALA*†[1] AND J. A. HARTIGAN[2]
[1] *Department of Biology, Center for Computational Ecology*
[2] *Department of Statistics, Yale University, New Haven, Connecticut, 06520, USA*

(*Received 24 May 1995, and in revised form 23 October 1995*)

Summary

A new method is presented for estimating the rate of gene flow into island populations using the distribution of alleles in samples from a number of islands. The pseudo maximum likelihood estimator (PMLE) that we derive may be applied to species with either discrete or continuous generation times. For Wright's discrete-generation island model, the method provides an estimate of $\theta = 2Nm$ where N is the (haploid) population size on each island and m is the fraction of individuals replaced by immigrants in each generation. For a continuous-generation island model, the corresponding parameter θ is the ratio of the immigration rate ϕ to the individual birth rate λ. Monte Carlo simulations are used to compare the statistical properties of the PMLE with those of two alternative estimators of θ derived from Wright's F-statistics. The PMLE is shown to have greatest efficiency (least mean square error) in most cases for a wide range of sample sizes and parameter values. The PMLE is applied to estimate θ using mtDNA haplotypes and allozymes for subdivided populations of African elephants and Channel Island foxes.

1. Introduction

The degree of reproductive isolation and rates of genetic exchange among subpopulations in nature are of critical importance for understanding many biological processes including speciation, genetic differentiation, and the maintenance of genetic diversity under population fragmentation. Previous studies have examined the expected frequency distributions of neutral alleles among subpopulations for several general patterns of migration, including an island model (Wright, 1931; Maruyama, 1970; Rannala & Hartigan, 1995) in which allele frequencies among immigrants are constant, or a one-dimensional or two-dimensional stepping-stone model (Malécot, 1948; Kimura, 1953; Kimura & Weiss, 1964; Maruyama, 1971) in which migrants are exchanged primarily between adjacent subpopulations.

An extension of these efforts to determine the theoretical distribution of neutral alleles among subpopulations, under particular models of gene flow, uses the theory-based predictions to infer the level of gene flow among natural subpopulations based on the observed distribution of allele frequencies (Wright, 1969; Lewontin, 1974; Slatkin, 1994). In particular, much research has been focused on procedures for estimating the composite parameter Nm (the number of individuals replaced by migrants per population per generation) using molecular genetic data for a collection of semi-isolated populations conforming to Wright's (1931) island model of population structure (Slatkin, 1985; Wehrhahn & Powell, 1987; Slatkin & Barton, 1989; Cockerham & Weir, 1993; Slatkin, 1994).

In this paper, we develop a general statistical method for estimating levels of gene flow in a neutral island model of population structure with fixed allele frequencies among immigrants and either discrete nonoverlapping generations with a constant rate of immigration and fixed population size (Wright, 1931) or continuous overlapping generations with constant immigration and birth rates (Rannala & Hartigan, 1995). We use the term *island model* to denote models of the genetic structure among semi-isolated populations in which the allele frequencies among immigrants are assumed constant and the effects of mutation negligible. These assumptions are important for the theory presented in this paper, and are implicit in previous papers on this topic (i.e. Barton & Slatkin,

* Author to whom correspondence and reprint requests should be sent.
† Current address: Department of Integrative Biology, University of California, Berkeley, California, 94720, USA.

https://doi.org/10.1017/S0016672300033607 Published online by Cambridge University Press

1986; Wehrhahn & Powell, 1987; Cockerham & Weir, 1993).

We will be interested in estimating what we term the 'scale' parameter θ under either a discrete or a continuous generation island model of population structure. For Wright's (1931) diploid discrete-generation island model $\theta = 4Nm$, where N is the population size and m is the fraction of individuals in each population replaced by immigrants in each generation. For the continuous-generation island model of Rannala & Hartigan (1995) $\theta = \phi/\lambda$, where ϕ is the immigration rate into an island and λ is the individual birth rate. In this paper, we assume θ is constant for each of a collection of independent islands as this is the situation considered in previous studies (i.e. Wehrhahn & Powell, 1987; Slatkin & Barton, 1989; Cockerham & Weir, 1993). In general, an assumption of constant θ among islands is not essential to the theory.

Three general classes of estimator are considered: (1) maximum likelihood estimator (MLE); (2) pseudo maximum likelihood estimator (PMLE; Gong & Samaniego, 1981); and (3) method of moments estimator (MME). The first approach using maximum likelihood advances the earlier work on MLE's of θ by Barton *et al.* (1983), Wehrhahn & Powell (1987), Wehrhahn (1989), and Slatkin & Barton (1989) while the second approach using PMLE's is entirely new. The MME is shown to be equivalent (in the two-allele case) to the well-known estimator $\hat{\theta} = 1/F_{ST} - 1$ derived from Wright's F_{ST} statistic (i.e. Wright 1969; Slatkin & Barton, 1989). Statistical properties of MME's and PMLE's are studied using Monte Carlo methods and the PMLE method is applied to analyse two empirical data sets. PMLE's of the parameter θ are obtained for subdivided populations of (1) Channel Island foxes and (2) African elephants using mtDNA Restriction Fragment Length Polymorphism (RFLP) and allozyme data.

2. The Dirichlet distribution

The limiting distribution of allele frequencies on each island, under a discrete-generation (Wright, 1931) or a continuous-generation (Rannala & Hartigan, 1995) island model of population structure is given by the finite Dirichlet distribution, a multidimensional generalization of the beta distribution. The finite Dirichlet distribution with scale parameter θ has density (see Johnson & Kotz, 1972),

$$P(\bar{\alpha}) = \Gamma(\theta) \prod_{i=1}^{k} \frac{\alpha_i^{\theta p_i - 1}}{\Gamma(\theta p_i)}, \tag{1}$$

where $\bar{\alpha} = \alpha_1, \alpha_2, \ldots, \alpha_k$ is the frequency array for k allele types on an island, and p_i is the frequency of the ith allele type among immigrants. The Dirichlet distribution arises as either the asymptotic distribution of allele frequencies at equilibrium (obtained by a diffusion approximation; see Wright, 1949, 1969) on islands for a multiallelic version of Wright's (1931) discrete-generation island model, or as the limiting distribution of allele frequencies on islands under the continuous-generation island model (Rannala & Hartigan, 1995).

The marginal distribution of the allele types in a random sample of size N, averaging over all elements of $\bar{\alpha}$, is the compound multinomial-Dirichlet (see Mosimann, 1962) given by,

$$\mathrm{P}(n_1, \ldots, n_k \mid N, \theta, \mathbf{p}) = \binom{N+\theta-1}{N}^{-1} \prod_{i=1}^{k} \binom{n_i + \theta p_i - 1}{n_i}. \tag{2}$$

This is the distribution of allele types for an island population (or a sample taken without replacement) of size N under the continuous-generation island model (Rannala & Hartigan, 1995) or for a random sample (with replacement) from an island population at equilibrium under the multiallelic generalization of Wright's island model when population size is large. These distributions will be applied below in deriving estimators of θ using the method of moments and maximum likelihood.

3. Method of moments estimator of θ

Wright (1969) proposed the following estimator of θ, for a discrete-generation island model of population structure at equilibrium, based on his F_{ST} statistic,

$$\hat{\theta} = \frac{1}{F_{ST}} - 1. \tag{3}$$

The statistic F_{ST} is defined (for a two-allele model) as,

$$F_{ST} = \frac{\sigma_p^2}{\bar{p}(1-\bar{p})}, \tag{4}$$

where σ_p^2 is the variance of allele frequency among populations and $\bar{p}$ is the average allele frequency in the collection of populations as a whole (see Wright, 1969). It can be shown (see Appendix) that the estimator given by eqn 3 arises as a special case of a more general approach to estimating the 'scale' parameter θ of the compound multinomial-Dirichlet distribution using the method of moments. The Wright approach to estimating θ using F-statistics for a two-allele model has been extended to allow for multiple alleles (see Slatkin & Barton, 1989; Cockerham & Weir, 1993) and two such methods for estimating θ are evaluated using Monte Carlo simulation in Section 6.

4. Maximum likelihood estimator of θ

Wehrhahn & Powell (1987) derived a MLE of θ based on a diffusion approximation for the asymptotic distribution of allele frequency on islands under Wright's (1931) two-allele island model. Their approach is based on the marginal beta distribution of allele frequency under this model. These authors did not take into account the multinomial process of

https://doi.org/10.1017/S0016672300033607 Published online by Cambridge University Press

sampling alleles from each island and their approach is only appropriate for estimates based on a large sample of individuals. Wehrhahn (1989) has extended this approach to include finite sampling effects using the beta-binomial p.m.f. as the likelihood.

MLE's of θ were also considered by Barton *et al.* (1983) and Slatkin & Barton (1989). The likelihood function proposed by these authors is only approximate and the exact likelihood for the case of two alleles and a sample of N individuals from a single island is the compound beta-binomial p.m.f. To estimate θ using samples from multiple independent islands the complete likelihood function is taken as the product of the observed likelihoods for samples from each island (or equivalently the sum of the log-likelihoods). Similarly, for several independent loci, the likelihood is taken as the product of the observed likelihoods over loci.

To treat multiple allele types both Slatkin & Barton (1989) and Wehrhahn & Powell (1987) derived the joint likelihood by taking the product of independent beta distributions over allele types (or equivalently the sum of the log-likelihoods). The marginal distributions of families of unique allele types at a single locus are only asymptotically independent however, and this approach may be inaccurate when few alleles are present among immigrants. The exact likelihood function for a sample from a population at a single locus composed of multiple allele types is given by the compound multinomial-Dirichlet p.m.f. considered above. We now derive the exact likelihood function for the distribution of allele types in a sample of individuals from each of a collection of independent islands, with each island receiving immigrants from a source with constant allele frequencies and scale parameter θ.

(i) *MLE for a single locus*

In this section, we present a general MLE of θ that may be applied for samples of any size from each of I islands. Initially, we consider a sample of alleles from a single locus. Let $\mathbf{n}$ be a matrix with dimension $I \times k$ and elements n_{ij}, where this is the number of alleles of the jth type obtained in a sample from the ith island, where $i = 1, \ldots, I$; $j = 1, \ldots, k$; I is the number of islands; and k is the number of unique alleles among immigrants. Let $\mathbf{N} = N_1, \ldots, N_I$ be the vector of island sample sizes, where N_i is the number of individuals sampled from the ith island (without replacement). Let θ be the scale parameter of the finite Dirichlet distribution (see above), and $\mathbf{p} = p_1, \ldots, p_k$ be the vector of allele frequencies among immigrants.

Assuming each immigrant is of allele type j with probability p_j, the likelihood of the observed $\mathbf{n}$, conditional on $\mathbf{N}$ is,

$$f(\mathbf{n} \mid \theta, \mathbf{p}) = \prod_{i=1}^{I} \left\{ \binom{N_i + \theta - 1}{N_i}^{-1} \prod_{j=1}^{k} \binom{n_{ij} + \theta p_j - 1}{n_{ij}} \right\}. \quad (5)$$

To obtain MLE's of $\mathbf{p}$ and θ the likelihood score function is obtained from the likelihood by a logarithmic transformation,

$$L = \sum_{i=1}^{I} \left\{ \sum_{j=1}^{k} \sum_{l=1}^{n_{ij}-1} \log(\theta p_j + l) - \sum_{l=0}^{N_i-1} \log(\theta + l) \right\}. \quad (6)$$

This equation may be numerically maximized, subject to the constraints,

$$\theta \geqslant 0 \quad \text{and} \quad \sum_{j=1}^{k} p_j = 1. \quad (7)$$

It may be shown (see Levin & Reed, 1977) that if an equal sample of size N is obtained from each island this likelihood has at most one local maximum. This maximum occurs for finite θ if,

$$\chi^2 > \sum_{j=1}^{k} \frac{n_{+j}/N}{p_j} - I, \quad (8)$$

and for $\theta = \infty$ otherwise, where

$$\chi_2 = \sum_{i=1}^{I} \sum_{j=1}^{k} (n_{ij} - N p_j)^2 / N p_j, \quad (9)$$

and $n_{+j} = \sum_{i=1}^{I} n_{ij}$. By applying Equation 8 above using an initial method of moments estimate of $\mathbf{p}$ it is possible to predict whether a unique local maximum exists for finite θ before attempting to locate the MLE of θ.

(ii) *MLE for multiple loci*

In the sequel, we assume that ξ loci are sampled and that the haplotypes of the individuals for the different loci are in linkage equilibrium. Let n_{lij} be the number of observed haplotypes with the jth allele type at the lth locus in a sample of N_{li} individuals for that locus from the ith island population. The likelihood function is,

$$f(\mathbf{n} \mid \theta, \mathbf{p}) = \prod_{l=1}^{\xi} \prod_{i=1}^{I} \left\{ \binom{N_{li} + \theta - 1}{N_{li}}^{-1} \prod_{j=1}^{k_l} \binom{n_{lij} + \theta p_{lj} - 1}{n_{lij}} \right\}. \quad (10)$$

The likelihood score function is,

$$\sum_{l=1}^{\xi} \sum_{i=1}^{I} \left\{ \sum_{j=1}^{k_l} \sum_{\gamma=0}^{n_{lij}-1} \log(\theta p_{lj} + \gamma) - \sum_{\gamma=0}^{N_{li}-1} \log(\theta + \gamma) \right\}. \quad (11)$$

This equation may be numerically maximized subject to the constraints,

$$\theta \geqslant 0 \quad \text{and} \quad \sum_{j=1}^{k_l} p_{lj} = 1 \quad \text{for all} \quad l = 1, \ldots, \xi. \quad (12)$$

In general, maximization of the complete likelihood will be a $1 + \sum_{i=1}^{\xi} k_i - \xi$ dimensional problem and may be very difficult for large numbers of loci and alleles. Since a pseudo maximum likelihood estimate of θ may be obtained with much less computational effort and is asymptotically greater than 90% efficient (see Chuang & Cox, 1985) we will not consider the complete

https://doi.org/10.1017/S0016672300033607 Published online by Cambridge University Press

likelihood further in this paper and will focus instead on the properties of a pseudo maximum likelihood estimator (PMLE; Gong & Samaniego, 1981).

5. Pseudo maximum likelihood estimator of θ

The likelihood function given in eqn 5 above depends on the parameters θ and $\mathbf{p}$. Since we are interested only in obtaining an estimate of θ, the vector $\mathbf{p}$ represents a set of nuisance parameters that complicate the maximization of the complete likelihood function. A useful technique for simplifying estimation problems involving many nuisance parameters is pseudo maximum likelihood estimation (PMLE; Gong & Samaniego, 1981). In PMLE nuisance parameters are replaced by consistent estimates and the reduced system of likelihood equations are solved to obtain estimates of the parameters of interest. Gong & Samaniego (1981) and Parke (1986) have shown that under fairly standard regularity conditions PMLE's are consistent and asymptotically normal.

In the case of the multinomial-Dirichlet distribution considered in the previous section, Chuang & Cox (1985) have considered PMLE's. Their analyses suggest that PMLE's possess uniformly higher asymptotic relative efficiencies (ARE; see Casella & Berger, 1990) than MME's. In addition, when only an estimate of the scale parameter θ is needed, the PMLE is generally greater than 90% efficient relative to the MLE, and so comes very close to the MLE in precision with much less computational effort.

The PMLE of the parameter θ for the multinomial-Dirichlet distribution is calculated as follows: (1) estimate the nuisance parameters $\mathbf{p}$ using the MME,

$$\hat{p}_i = \frac{\sum_{j=1}^{I} n_{ij}}{N'}, \quad i = 1, 2, \ldots, k. \tag{13}$$

where $N' = \sum_{j=1}^{I} \sum_{i=1}^{k} n_{ij}$. This MME is a consistent estimator of $\mathbf{p}$ since the expectation of $\bar{\alpha}$ under the Dirichlet distribution is $\mathbf{p}$; (2) maximize the likelihood function given by eqn 5 with respect to θ with $\hat{\mathbf{p}}$ substituted for $\mathbf{p}$. This maximization is then one-dimensional in θ and a local maximum exists in regions where the likelihood is convex (i.e. the second derivative is negative) and the first derivative with respect to θ equals zero,

$$\frac{\partial L}{\partial \theta} = \sum_{i=1}^{I} \left\{ \sum_{j=1}^{k} \sum_{a=0}^{n_{ij}-1} \frac{1}{\theta + a\hat{p}_j^{-1}} - \sum_{a=0}^{N_i-1} \frac{1}{\theta + a} \right\}. \tag{14}$$

The extrema of θ may be located numerically by finding roots in θ for the following equation,

$$\sum_{i=1}^{I} \sum_{j=1}^{k} \sum_{a=0}^{n_{ij}-1} \frac{1}{\theta + a\hat{p}_j^{-1}} = \sum_{i=1}^{I} \sum_{a=0}^{N_i-1} \frac{1}{\theta + a}. \tag{15}$$

The PMLE of θ may also be found by a direct graphical analysis. For the case of multiple independent loci, the PMLE is obtained by differentiating the likelihood score function of eqn 11 with respect to θ and setting the resulting equation equal to zero to obtain,

$$\frac{\partial L}{\partial \theta} = \sum_{l=1}^{\xi} \sum_{i=1}^{I} \left\{ \sum_{j=1}^{k_l} \sum_{a=0}^{n_{lij}-1} \frac{1}{\theta + a\hat{p}_{lj}^{-1}} - \sum_{a=0}^{N_{li}-1} \frac{1}{\theta + a} \right\}. \tag{16}$$

The extrema of θ may be found numerically by solving the following equation for the roots in θ,

$$\sum_{l=1}^{\xi} \sum_{i=1}^{I} \sum_{j=1}^{k_l} \sum_{a=0}^{n_{lij}-1} \frac{1}{\theta + a\hat{p}_{lj}^{-1}} = \sum_{l=1}^{\xi} \sum_{i=1}^{I} \sum_{a=0}^{N_{li}-1} \frac{1}{\theta + a}, \tag{17}$$

where

$$\hat{p}_{lj} = \sum_{i=1}^{I} n_{lij}/N', \quad j = 1, 2, \ldots, k_l,$$

$$\text{and} \quad N' = \sum_{i=1}^{I} \sum_{j=1}^{k_l} n_{lij}.$$

For the numerical analysis presented in this paper, eqns 15 and 17 were solved for θ using algorithms based on Newton's method contained in the MATHEMATICA computer program (Wolfram Research, Inc., 1992). A MATHEMATICA package for estimating θ using the PMLE method is available upon request from the authors or on the internet World-Wide Web (WWW) from `http://mw511.biol.berkeley.edu/homepage.html`.

6. Monte Carlo simulations

Computer simulations were used to study the statistical properties of the PMLE and MME estimators considered in this paper. Replicate island populations structured according to a stochastic BDI process were generated using the Polya urn scheme described below. The genetic model considered a haploid species with k allele types among immigrants in frequencies $\mathbf{p} = p_1, p_2, \ldots, p_k$. Properties of the estimators over several ranges of parameter values were studied. We examined the influenced of the number of unique allele types among immigrants (k), the distribution of allele frequencies among immigrants ($\mathbf{p}$), the number of islands sampled (I), the number of individuals sampled from each island (N), and the magnitude of the true parameter θ on the accuracy of estimates of θ obtained using MME's and PMLE's.

The effect of the number of loci sampled on statistical accuracy, under the multi-locus model, was not studied using simulations because the number of parameter combinations to be considered becomes too large. If the distribution of allele frequencies among immigrants is identical for all loci, the statistical effect of adding either more loci, or more islands, to the sample is identical. The properties of the estimators were compared for 48 combinations of parameter values in total (see Table 1).

We compared the statistical properties of three estimators: (1) the MME proposed by Slatkin &

https://doi.org/10.1017/S0016672300033607 Published online by Cambridge University Press

Barton (1989) based on a ratio of the G_{ST} statistic of Nei (1973),

$$\hat{\theta}_{SB} = \frac{1}{G_{ST}} - 1, \tag{18}$$

calculated using the formula suggested by Cockerham & Weir (1993) which we refer to as SB; (2) the estimator of Cockerham & Weir (1993) based on a ratio of the G_{CA} statistic of Crow & Aoki (1984),

$$\hat{\theta}_{CW} = \frac{1}{G_{CA}} - 1, \tag{19}$$

which we refer to as CW; and (3) the pseudo maximum likelihood estimator (PMLE) considered in Section 5. The methods used to calculate CW and SB are described in the paper by Cockerham & Weir (1993).

(i) *Polya urn scheme for simulating island populations*

An island population of size n was generated from n arrivals of a birth and immigration process by using uniform (0, 1) random variables to decide the type of event (birth or immigration) at each arrival. The probability that a new arrival is an immigrant in a population of size N is,

$$\frac{\theta}{N+\theta}. \tag{20}$$

The arriving immigrant has an allele of type i with probability p_i where $i = 1, 2, \ldots, k$. The probability that a new arrival is a birth in a population of size N is,

$$\frac{N}{N+\theta}. \tag{21}$$

If the arrival is a birth its allele type is chosen with probability equal to the relative frequency of individuals of marker types $i = 1, 2, \ldots, k$ in the population immediately prior to the birth. A program written in C to perform these simulations is available from the authors upon request or on the WWW from `http://mw511.biol.berkeley.edu/homepage.html`.

(ii) *Evaluating estimators of* θ

The statistical properties of the estimators were compared using two optimality criteria: (1) the mean square error (MSE) defined as,

$$\mathrm{MSE}(\hat{\theta}) = \frac{1}{R}\sum_{i=1}^{R}(\theta - \hat{\theta}_i)^2, \tag{22}$$

where θ is the expected value of the parameter under the model, $\hat{\theta}_i$ is an estimate of θ from the ith replicate dataset using a particular statistic, and R is the total number of simulated datasets ($R = 1000$ in all cases), and (2) the bias of $\hat{\theta}$ defined as (see Casella & Berger, 1990),

$$\begin{aligned}\mathrm{BIAS}(\hat{\theta}) &= E[\hat{\theta}] - \theta, \\ &= \sqrt{[\mathrm{MSE}(\hat{\theta}) - VAR(\hat{\theta})]}.\end{aligned} \tag{23}$$

The observed difference between θ and $\bar{\theta} = 1/R\sum_{i=1}^{R}\hat{\theta}_i$ was used to estimate $\mathrm{BIAS}(\hat{\theta})$ for the simulated datasets according to the following formula,

$$\mathrm{BIAS}(\hat{\theta}) = \bar{\theta} - \theta. \tag{24}$$

To reduce the influence of extreme outliers on the results of comparisons, we used the following criterion to remove extreme values from the simulation distributions for each estimator prior to calculations of the summary statistics,

$$|\hat{\theta}_i - \bar{\theta}| < 3\sigma_{\hat{\theta}}, \tag{25}$$

where $\sigma_{\hat{\theta}} = \sqrt{[\sum_{i=1}^{R}(\bar{\theta} - \hat{\theta}_i)^2]}/\sqrt{R}$. An application of this criterion always reduced the total number of simulated datasets by 23 or less and substantially improved the variance of our estimates of $\mathrm{MSE}(\hat{\theta})$. In the two cases where outliers were most extreme ($k = 2$, $I = 10$, $\theta = 10$, $N = 10$, $\mathbf{p} = 0{\cdot}5$: $0{\cdot}5$ and $\mathbf{p} = 0{\cdot}25$: $0{\cdot}75$), we used $2\sigma_{\hat{\theta}}$ as our criterion for inclusion.

To facilitate comparisons of MSE and BIAS for different ranges of the true parameter θ, we used the following transformation to 'normalize' all estimates to have expectation one,

$$\hat{\theta}_{\mathrm{Norm}} = \frac{\hat{\theta}}{\theta}. \tag{26}$$

The significance of observed differences in MSE among estimators was tested using the following statistic to perform pairwise comparisons,

$$D_{12} = \frac{1}{R'}\sum_{i=1}^{R'}\{[\hat{\theta}_{i(1)} - \theta]^2 - [\hat{\theta}_{i(2)} - \theta]^2\}, \tag{27}$$

where R' is the total number of simulated datasets satisfying criterion 25 above for both estimators 1 and 2, and $\hat{\theta}_{i(1)}$ and $\hat{\theta}_{i(2)}$ are estimates of θ obtained for the ith simulated dataset using estimators 1 and 2, respectively. This statistic is asymptotically normally distributed with standard error (S.E.),

$$\mathrm{S.E.}(D_{12}) = \frac{\sqrt{\left(\frac{1}{R'}\sum_{i=1}^{R'}\{D_{12} - [\hat{\theta}_{i(1)} - \theta]^2 + [\hat{\theta}_{i(20} - \theta]^2\}^2\right)}}{\sqrt{R'}}. \tag{28}$$

The statistic D_{12} was considered significant at the $\alpha = 0{\cdot}05$ level if the 95% confidence interval derived from eqn 28 did not include zero (where $CI = 1{\cdot}96 \times \mathrm{S.E.}(D_{12})$).

(iii) *Simulation results*

The results of the Monte Carlo simulations are summarized in Figs 1–4 and Table 1. To study the influence of population sample size, either $N = 10$ or

https://doi.org/10.1017/S0016672300033607 Published online by Cambridge University Press

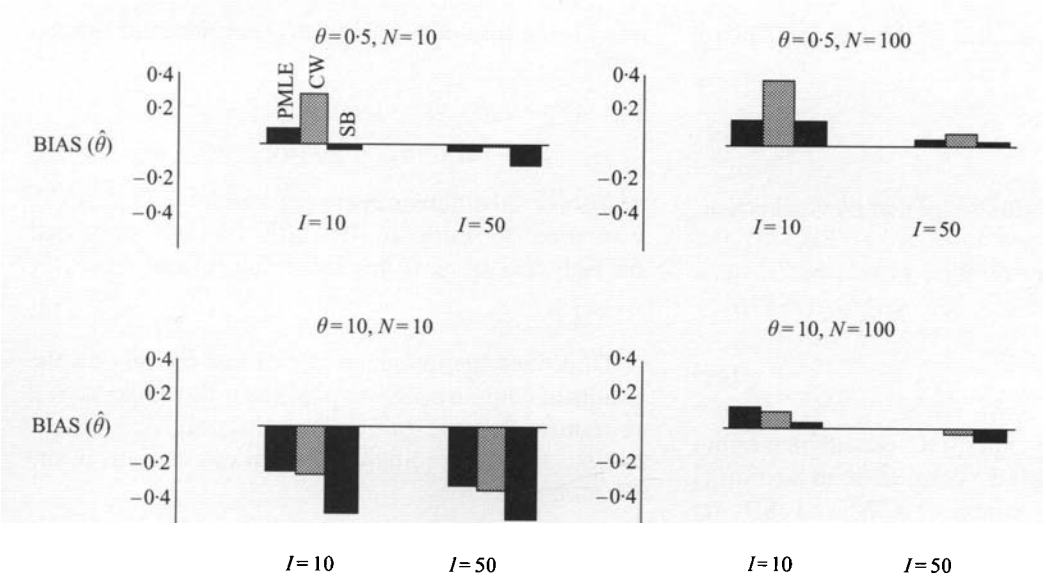

Fig. 1. The average normalized bias of estimates of parameter θ for eight combinations of $R = 1000$ simulations with $k = 10$ and $\mathbf{p} = 0{\cdot}1, \ldots, 0{\cdot}1$. The histograms from left to right represent estimators PMLE, CW and SB. I denotes the number of islands sampled, θ is the scale parameter (see text), N is the sample size per island, k is the number of allele types among immigrants, and $\mathbf{p}$ is the array of allele frequencies among immigrants.

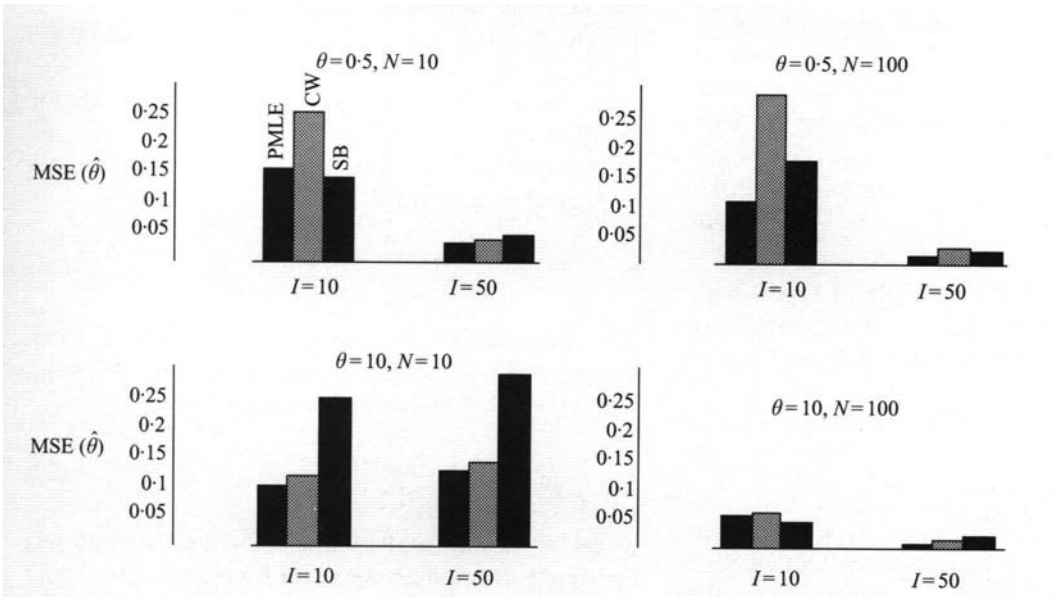

Fig. 2. The average normalized mean square error of estimates of parameter θ for eight combinations of $R = 1000$ simulations with $k = 10$ and $\mathbf{p} = 0{\cdot}1, \ldots, 0{\cdot}1$. The histograms from left to right represent estimators PMLE, CW and SB. I denotes the number of islands sampled, θ is the scale parameter (see text), N is the sample size per island, k is the number of allele types among immigrants, and $\mathbf{p}$ is the array of allele frequencies among immigrants.

$N = 100$ individuals were generated to form each simulated island population, for each set of parameter values, by an independent birth-immigration process (see above). To study the influence of the number of islands sampled, we simulated either 10 islands or 50 islands for each set of parameter values. The value of θ was set to either 0·5, 2, or 10. We studied both the two-allele model, with equal ($\mathbf{p} = 0{\cdot}5 : 0{\cdot}5$) or skewed ($\mathbf{p} = 0{\cdot}25 : 0{\cdot}75$) allele frequencies, and the multi-allelic model with $k = 10$ allele types and equal ($\mathbf{p} = 0{\cdot}1 : 0{\cdot}1 : 0{\cdot}1 : 0{\cdot}1 : 0{\cdot}1 : 0{\cdot}1 : 0{\cdot}1 : 0{\cdot}1 : 0{\cdot}1 : 0{\cdot}1$) or skewed ($\mathbf{p} = 0{\cdot}5 : 0{\cdot}1 : 0{\cdot}05 : 0{\cdot}05 : 0{\cdot}05 : 0{\cdot}05 : 0{\cdot}05 : 0{\cdot}05 : 0{\cdot}05 : 0{\cdot}05$) allele frequencies.

Fig. 1 shows the effect of the number of islands sampled (I) on BIAS($\hat{\theta}$) for the three estimators PMLE, CW and SB. All the estimators are biased over some range of the parameter space. Increasing I has a large effect in reducing BIAS($\hat{\theta}$), while the effect of increasing N is smaller. The expected bias may be either positive or negative, depending on the combination of parameter values, and may change from a positive to a negative value with an increase in I or θ (see Fig. 1).

Fig. 2 shows the effect of the number of islands sampled (I) on MSE($\hat{\theta}$). Increasing I has a large effect in reducing MSE($\hat{\theta}$) for all three estimators, except in the case that θ is large and N is small. In general, sample size N has its greatest effect on BIAS($\hat{\theta}$)

https://doi.org/10.1017/S0016672300033607 Published online by Cambridge University Press

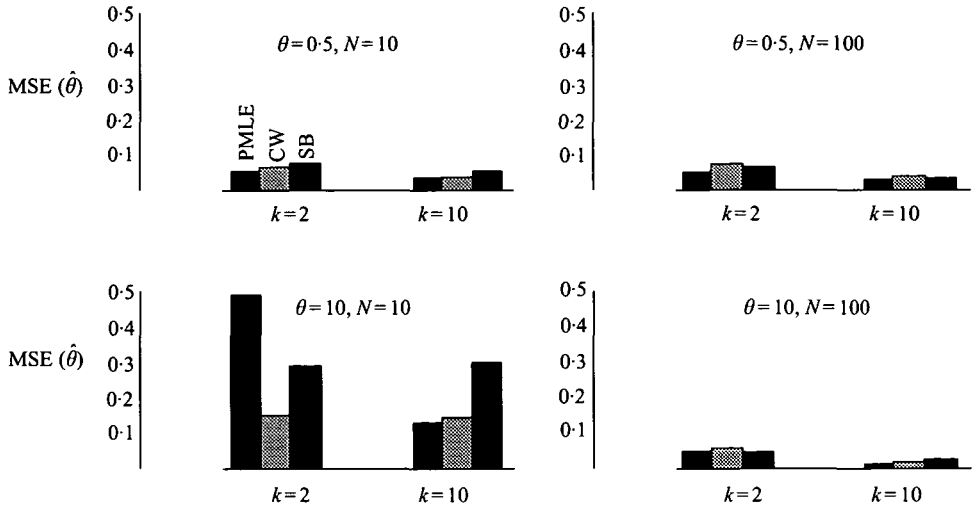

Fig. 3. The average normalized mean square error of estimates of parameter θ for eight combinations of $R = 1000$ simulations with $\mathbf{p} = 0{\cdot}5, 0{\cdot}5$ (for $k = 2$), $\mathbf{p} = 0{\cdot}1, \ldots, 0{\cdot}1$ (for $k = 10$), and $I = 50$. The histograms from left to right represent estimators PMLE, CW and SB. I indicates the number of islands sampled, θ is the scale parameter (see text), N is the sample size per island, k is the number of allele types among immigrants, and $\mathbf{p}$ is the array of allele frequencies among immigrants.

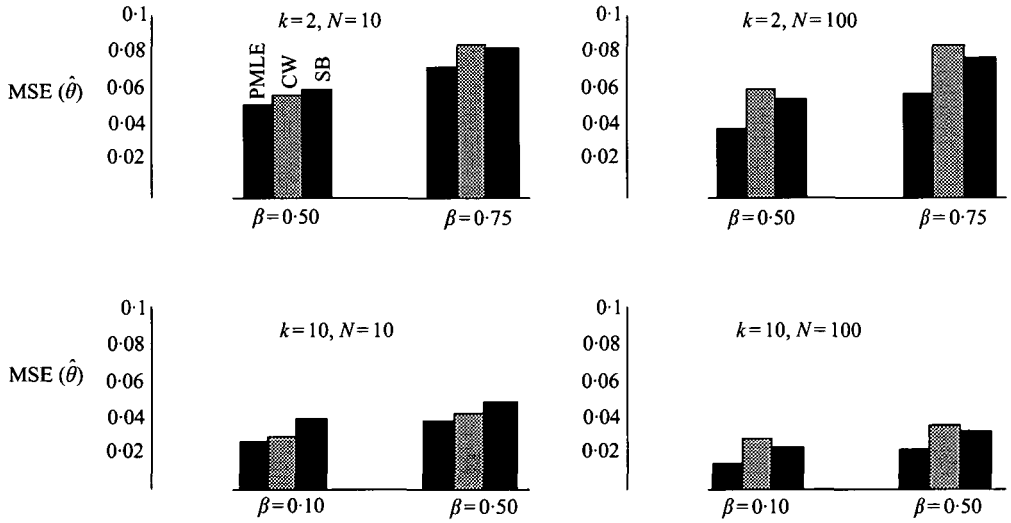

Fig. 4. The average normalized mean square error of estimates of parameter θ for each combinations of $R = 1000$ simulations with $\theta = 0{\cdot}5$ and $I = 50$. β is the frequency of the most common allele. For $\beta = 0{\cdot}50$ and $k = 2$, $\mathbf{p} = 0{\cdot}5, 0{\cdot}5$, and for $\beta = 0{\cdot}75$ and $k = 2$, $\mathbf{p} = 0{\cdot}75, 0{\cdot}25$. For $\beta = 0{\cdot}10$ and $k = 10$, $\mathbf{p} = 0{\cdot}10, \ldots, 0{\cdot}10$, and for $\beta = 0{\cdot}50$ and $k = 10$, $\mathbf{p} = 0{\cdot}50, 0{\cdot}10, 0{\cdot}05, \ldots, 0{\cdot}05$. The histograms from left to right represent estimators PMLE, CW and SB. I indicates the number of islands sampled, θ is the scale parameter (see text), N is the sample size per island, k is the number of allele types among immigrants, and $\mathbf{p}$ is the array of allele frequencies among immigrants.

and MSE($\hat{\theta}$) when θ is large. Fig. 3 shows the effect of the number of alleles among immigrants (k) on MSE($\hat{\theta}$). An increase in k always reduces MSE($\hat{\theta}$). For the case of large θ and small N, the MSE($\hat{\theta}$) is greatly reduced by an increase in k for the PMLE, but not the other estimators. Fig. 4 shows the effect of the distribution of allele frequencies on the MSE($\hat{\theta}$). An increase in the frequency of the most common allele (β) always results in an increase in MSE($\hat{\theta}$) for the three estimators considered under either a $k = 2$ allele model or a $k = 10$ allele model (see Fig. 4).

In general, the number of islands sampled (I) appears to have a much greater influence on MSE($\hat{\theta}$) and BIAS($\hat{\theta}$) than the sample size per island (N) except in the case that θ is large. This may be noted by examining Figs 1 and 2. This finding is consistent with the earlier results of Slatkin & Barton (1989) who also noted the important influence of the number of islands sampled on the accuracy of estimators of Nm. The effect on MSE($\hat{\theta}$) of additional loci should be similar. Table 1 shows the results of comparisons of MSE($\hat{\theta}$) for the three estimators studied using the D-statistic given by eqn 27 above. For a small sample of individuals ($N = 10$) the PMLE is most efficient (i.e. has least MSE) in 63% of cases, CW is most efficient in 29% of cases, and SB is most efficient in 21% of

https://doi.org/10.1017/S0016672300033607 Published online by Cambridge University Press

Table 1. *Pairwise comparison of mean square error (MSE) among estimators*

	$\theta = 0{\cdot}5$		$\theta = 2$		$\theta = 10$	
	$I = 10$	$I = 50$	$I = 10$	$I = 50$	$I = 10$	$I = 50$
$k = 2$, $\mathbf{p} = 0{\cdot}5, 0{\cdot}5$						
$N = 10$	+	+	+	−	−	−
$N = 100$	+	+	+	+	+	+
$k = 2$, $\mathbf{p} = 0{\cdot}75, 0{\cdot}25$						
$N = 10$	+	+	+	−	−	−
$N = 100$	+	+	+	+	+	+
$k = 10$, $\mathbf{p} = 0{\cdot}1, \ldots, 0{\cdot}1$						
$N = 10$	−	+	+	+	+	+
$N = 100$	+	+	+	+	−	+
$k = 10$, $\mathbf{p} = 0{\cdot}5, 0{\cdot}1, 0{\cdot}05, \ldots, 0{\cdot}05$						
$N = 10$	+	+	+	+	−	−
$N = 100$	+	+	+	+	+	+

The significance of pairwise D-statistics (see text) for estimators PMLE, CW and SB are listed in each box. A minus sign in a box indicates that the PMLE was significantly less efficient (at the $\alpha = 0{\cdot}05$ level) than either the CW, the SB, or both, for the parameter values indicated. A plus sign indicates that neither of the estimators CW or SB were significantly greater in efficiency than the PMLE. The sample size per island is N; I is the number of islands sampled; k is the number of allele types; $\mathbf{p}$ is the array of allele frequencies among immigrants; θ is the actual value of the scale parameter (see text); PMLE is the pseudo maximum likelihood estimator; CW is the method of moments estimator (MME) of Cockerham & Weir (1993); SB is the MME of Slatkin & Barton (1989).

cases. For a large sample size ($N = 100$), PMLE is most efficient in 96% of cases, CW is most efficient in no cases, and SB is most efficient in 17% of cases.

In terms of MSE the PMLE of θ appears to be the best choice regardless of whether a large or a small sample of individuals is taken. The efficiency of the PMLE relative to the other estimators increases as the sample size per island increases. The simulations of Chuang & Cox (1985) comparing the PMLE of θ for a compound multinomial-Dirichlet with the MME of θ proposed by Brier (1980) also showed the PMLE to be superior.

7. Numerical examples

To illustrate the application of the PMLE derived in this paper to interpret genetic data from natural populations, we analysed mitochondrial DNA (mtDNA) Restriction Fragment Length Polymorphism (RFLP) data and allozyme data from subdivided populations of two species: Channel Island foxes and African elephants. The following calculations were performed: (1) numerical estimates were obtained for $\mathbf{p}$ and $\hat{\theta}$ using the MME and the PMLE of eqns 13 and 15 above (for mtDNA data), or eqn 17 (for allozyme data), respectively; (2) standard errors were calculated for $\hat{\theta}$ by simulating 1000 datasets using estimates of $\mathbf{p}$ and $\hat{\theta}$ as parameters for the simulation, and the observed sample sizes, and equating the standard error of $\hat{\theta}$ to the standard deviation of estimates of $\hat{\theta}$ obtained for the simulated datasets.

Table 2. *The mtDNA haplotype distribution among islands for Channel Island foxes*

	Locality					
Haplotype	SMi	SRo	SCr	SNi	SCa	SCl
1	17	0	0	0	0	27
2	0	0	0	22	0	0
3	0	0	0	0	3	0
4	0	7	5	0	17	0
5	5	23	23	0	4	0

Island names are abbreviated as follows: SMi = San Miguel; SRo = Santa Rosa; SCr = Santa Cruz; SNi = San Nicolas; SCa = Santa Catalina; SCl = San Clemente (modified from Wayne *et al.* 1991).

(i) *mtDNA and allozyme variation in Channel Island fox populations*

The morphologically-distinct island fox *Urocyon littoralis* inhabits six of the eight Channel Islands, located off the coast of southern California (Collins, 1982). The island fox is descended from the mainland gray fox *U. cinereoargenteus* (Wayne *et al.* 1991). The Channel Islands are between 40 and 100 km from the mainland. Wayne *et al.* (1991) examined the frequencies of five mtDNA haplotypes in each of six Channel Island fox populations based on a sample of 153 foxes

https://doi.org/10.1017/S0016672300033607 Published online by Cambridge University Press

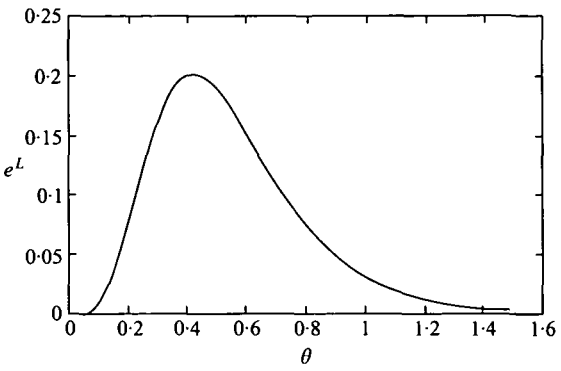

Fig. 5. Likelihood as a function of the scale parameter θ for Channel Island fox populations. Likelihood is multiplied by 10^{-14}.

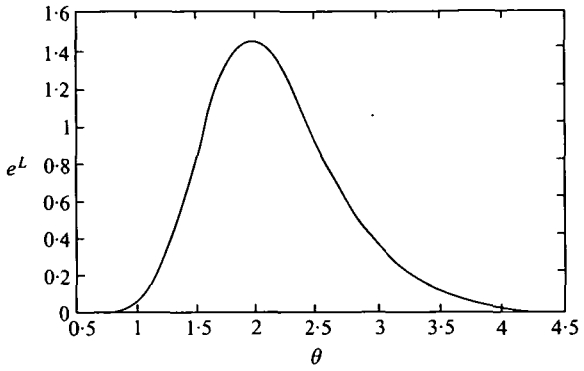

Fig. 6. Likelihood as a function of the scale parameter θ for African elephant populations. Likelihood is multiplied by 10^{-50}.

in total. The observed haplotype frequencies are shown in Table 2. The likelihood as a function of θ is shown in Fig. 5. The estimates obtained from these data are: $\mathbf{p} = 0{\cdot}29: 0{\cdot}14: 0{\cdot}02: 0{\cdot}19: 0{\cdot}36$ and $\hat{\theta} = 0{\cdot}41 \pm 0{\cdot}35$.

Wayne *et al.* (1991) also studied allozymes for Channel Island foxes. The distribution of alleles at four polymorphic allozyme loci were used to estimate θ from their data. The allelic distributions were inferred from the observed allele frequencies and sample sizes in Table 4 of the Wayne *et al.* (1991) paper. The estimate of $\hat{\theta}$ obtained from these data is $\hat{\theta} = 1{\cdot}34 \pm 0{\cdot}42$. The observed difference between the estimate of θ from the mtDNA and allozyme data is $D = 0{\cdot}93$. This is within the 95% confidence interval derived from the standard errors of the estimates ($CI = 1{\cdot}07$).

(ii) *mtDNA variation in African elephant populations*

The African elephant (*Loxodonta africana*), indigenous to eastern and southern Africa, exists in numerous populations fragmented by human development. Georgiadis *et al.* (1994) studied the distribution of 10 mtDNA RFLP haplotypes by sampling a total of 270 elephants from 10 populations in Kenya, Zimbabwe, Botswana and South Africa. The observed haplotype frequencies are shown in Table 3. The likelihood as a function of θ is shown in Fig. 6. The estimates obtained from these data are: $\mathbf{p} = 0{\cdot}07: 0{\cdot}03: 0{\cdot}36: 0{\cdot}22: 0{\cdot}06: 0{\cdot}01: 0{\cdot}01: 0{\cdot}21: 0{\cdot}03: 0{\cdot}01$, and $\hat{\theta} = 1{\cdot}88 \pm 0{\cdot}61$.

8. Discussion

In this paper we have considered a generalized demographic-genetic model for island populations receiving immigrants from a source with constant allele frequencies. Previous studies of the genetics of island populations have focused almost exclusively on the discrete-generation Fisher–Wright demographic model (i.e. Wright, 1951; Weir & Cockerham, 1984; Slatkin & Barton, 1989; Slatkin & Maddison, 1989; Cockerham & Weir, 1993). We have suggested that a

Table 3. *The mtDNA haplotype distribution among populations of African elephants*

	Locality									
Haplotype	Kr	Ma	Sa	Ch	Za	Hw	Se	Ta	Ts	Am
1	0	0	0	0	0	0	0	0	8	11
2	0	0	0	0	0	0	0	6	2	1
3	9	2	2	20	4	3	11	27	2	16
4	25	15	0	8	1	7	0	0	2	1
5	0	0	0	0	1	7	8	0	0	0
6	0	0	0	0	0	0	3	0	0	0
7	0	0	0	0	0	0	0	2	0	0
8	0	0	18	12	21	6	0	0	0	0
9	0	0	6	0	0	1	0	0	0	0
10	0	0	1	0	1	0	0	0	0	0

Population names are abbreviated as follows: Am = Amboselli, Kenya; Ts = Tsavo, Kenya; Ta = Tarangire, Tanzania; Se = Sengwa, Zimbabwe; Hw = Hwang, Zimbabwe; Za = Zambezi, Zimbabwe; Ch = Chobe, Botswana; Sa = Savute, Botswana; Ma = Mashatu, Botswana; Kr = Kruger, South Africa (modified from Georgiadis *et al.* 1994).

https://doi.org/10.1017/S0016672300033607 Published online by Cambridge University Press

Dirichlet distribution of population allele frequencies on islands results for either a continuous or (asymptotically) a discrete generation island model of population structure. This finding provides a theoretical justification for the widespread use of an estimator of Nm, based on Wright's F-statistics, by empiricists for inferring gene flow in species with continuous overlapping generations (reviewed by Slatkin, 1985). Wright's estimator is shown to arise as a ratio estimator of the parameter θ of the Dirichlet using the method of moments (see Appendix).

There are several properties of the continuous-generation island model that necessitate a new interpretation of θ by comparison with traditional interpretations based on a discrete-generation island model (Wright, 1931). Perhaps the most important difference is that the complete replacement of all individuals in the population in each generation of Wright's model is replaced by continuous births, and θ is interpreted as the immigration rate ϕ divided by the individual birth rate λ. In general, an estimate of θ based on molecular genetic data provides no information about ϕ without additional information about λ. The interrelationship between ϕ, λ and θ highlights the importance of studying the population biology (i.e. demography) as well as the genetics in attempting to infer gene flow in natural populations with overlapping generations.

Another way in which the continuous-generation model differs from Wright's model is that the compound multinomial-Dirichlet distribution describes the allele frequency distribution for populations of any age, whereas for Wright's model the populations must have existed long enough to have reached 'equilibrium'. This means that estimators for the discrete-generation model are only valid in situations in which populations are at equilibrium, whereas for continuous-generation species no equilibrium assumption is necessary (other factors such as age-specific birth and immigration rates may change this however).

The results of our simulation study suggest that the pseudo maximum likelihood estimator (PMLE) is the most efficient estimator of θ for an island model of population structure, except possibly in cases where the sample size on each island is very small (e.g. $N = 10$) and the scale parameter θ is large (e.g. $\theta = 10$). All three estimators appear to behaviour badly for large θ, small N, and small I. Consistent with the findings of Slatkin & Barton (1989), the number of islands sampled (I), and the number of loci sampled (ξ) in the multi-locus case, appear to be the most important factors influencing MSE and BIAS for all three estimators of θ considered.

There are a number of additional estimators of θ in the literature that were not considered in this paper. These include estimators of θ using genealogical information based on a cladistic analysis (Slatkin & Maddison, 1989; Hudson *et al.* 1992). The statistical properties of these estimators appear quite similar to those of the MME's based on Wright's F-statistics (see e.g. Hudson *et al.* 1992).

The dependence of θ on birth rate for the continuous-generation model has important consequences when an attempt is made to compare relative levels of gene flow between species with different generation times. For example, in our analysis of mtDNA data for Channel island fox and African elephant populations we estimated for foxes $\hat{\theta} = 0{\cdot}41 \pm 0{\cdot}35$, and for elephants $\hat{\theta} = 1{\cdot}88 \pm 0{\cdot}61$. Since elephants may be expected to have lower individual birth rates than foxes the observed difference in θ might be due either to a difference in birth rate between the two species, a difference in the immigration rate, or both.

A number of additional demographic factors that may influence estimates of θ for species with overlapping generations were not considered in this paper. Most important is age-structure: the continuous-generation island model we have considered, based on a birth, death and immigration process (BDI; Rannala & Hartigan, 1995), assumes that each individual on an island has an equal probability of reproducing during any interval of time. The probability of reproduction does not depend on factors such as age, or genotype. In natural populations, age is obviously an important factor influencing reproduction rate and a more satisfactory model would also take into account the effects of age structure, both for island residents and for immigrants.

One way to generalize the classical island model considered in this paper to provide a better description of many natural populations is to allow θ to vary among islands, while keeping $\mathbf{p}$ constant. It is clear that the variance of individual estimates of θ will be greatly increased by this modification since each estimate of θ is then essentially based on a single observation (island). By examining multiple loci it should be possible to reduce the variance of these estimates to reasonable levels. It is clear that 'generalized' island models of this form deserve more careful study.

9. Conclusions

Much work remains to be done to evaluate the statistical properties of the large number of existing approaches for estimating gene flow in natural populations. In this paper, we have concentrated on estimators that may be derived from a theoretical consideration of the distribution of alleles in populations under an island model of population structure. The maximum likelihood and method of moments estimators of θ that we have considered are based on a Dirichlet distribution of allele frequencies on islands. Future studies should consider the form of the allele frequency distribution on islands under alternative models of population structure such as the 'stepping-stone' model. Such studies have the potential to

https://doi.org/10.1017/S0016672300033607 Published online by Cambridge University Press

greatly increase the range of natural population structures for which exact MLE's might be developed and applied to estimate gene flow.

Based on comparisons of the mean square error (MSE) from our simulation results, the pseudo maximum likelihood method (PMLE) appears to be a better choice for estimating θ under an island model of population structure than either the method of moments estimator (MME) of Cockerham & Weir (1993) or the MME of Slatkin & Barton (1989). We have shown that all three estimators are justifiably applied to estimate θ for species with either discrete or continuous generation times. The number of islands sampled (I) appears to have the greatest influence on bias and MSE, and all three estimators have low efficiency and high bias when θ is large, the sample size (N) is small, and the number of islands sampled (I) is small. The PMLE appears to be asymptotically greater than 90% efficient (Chuang & Cox, 1985) and is much simpler to calculate than the complete MLE. This makes it well suited for analyzing many empirical datasets when a large computing effort may be required to obtain MLE's.

Population genetic studies are frequently carried out with limited demographic information for the study populations and estimates of the number of immigrants per generation are often based exclusively on a genetic analysis (see e.g. Zink & Remsen, 1986; Edwards, 1993; Mercure *et al.*, 1993). It is clear that estimates of immigration from molecular genetic data are not possible for a continuous-generation island model without additional demographic information concerning the individual birth rates on islands. This suggests that a closer collaboration between field ecologists and geneticists is needed in future studies of subdivided populations aimed at predicting gene flow.

We thank W. J. Ewens, M. Slatkin, G. P. Wagner, C. F. Wehrhahn, and an anonymous reviewer, for their comments on an earlier version of this paper. N. Georgiadis and R. Wayne generously offered us the use of their mtDNA data. Financial support for the work of B.R. was provided by grants from the National Science Foundation (DEB-9224949), the John F. Enders Research Grant of Yale University, and Sigma-Xi. Personal support for B.R. was provided by a Natural Sciences and Engineering Research Council (NSERC) of Canada postgraduate scholarship and a fellowship from Yale University. Financial support for J.A.H. was provided by a grant from the National Science Foundation (DMS-8617919). This article is communication no. 26 from the Center for Computational Ecology of the Yale Institute for Biospheric Studies.

Appendix

In this Appendix we show that the gene flow estimator proposed by Wright (1969), based on his F_{ST} statistic, also arises as a method of moments ratio estimator of the parameter θ of the compound multinomial-Dirichlet (CMD) distribution, which is the asymptotic sampling distribution for Wright's island model at equilibrium.

Mosimann (1962) proved the following result for the compound multinomial-Dirichlet distribution which may be applied to our genetic models,

$$\mathbf{Z} = \left(\frac{N+\theta}{1+\theta}\right)\mathbf{Y}, \tag{A 1}$$

where $\mathbf{Z}$ is the variance–covariance matrix of the sizes of families of unique alleles in a sample of size N from a compound multinomial-Dirichlet distribution, and $\mathbf{Y}$ is the variance–covariance matrix of the sizes of families of unique alleles in a sample of size N from a multinomial distribution with parameters equal to the expectations of the Dirichlet distribution (i.e. $Y_{ij} = -Np_ip_j$, $Y_{ii} = Np_i(1-p_i)$, where $i \neq j$). Note that the first-order moments of the Dirichlet distribution are, in the genetic models, equal to the allele frequencies among immigrants.

If N individuals are sampled from each of I independent islands, then an estimator of $C = (N+\theta)/(1+\theta)$ is,

$$\hat{C} = \left(\frac{|\hat{\mathbf{Z}}|}{|\hat{\mathbf{Y}}|}\right)^{1/k-1}, \tag{A 2}$$

where k is the total number of unique alleles, and $\hat{\mathbf{Z}}$ and $\hat{\mathbf{Y}}$ are consistent estimates of $\mathbf{Z}$ and $\mathbf{Y}$. For example, the method of moments estimates,

$$\hat{Y}_{ij} = -N\hat{p}_i\hat{p}_j, \quad i \neq j, \tag{A 3}$$

$$\hat{Y}_{ii} = N\hat{p}_i(1-\hat{p}_i), \tag{A 4}$$

where $\hat{p}_i = 1/I\sum_{j=1}^{I} n_{ij}/N$ and,

$$\hat{Z}_{ij} = \frac{1}{I-1}\sum_{l=1}^{I}(\bar{n}_{i.}-n_{il})(\bar{n}_{j.}-n_{jl}), \quad i \neq j, \tag{A 5}$$

$$\hat{Z}_{ii} = \frac{1}{I-1}\sum_{l=1}^{I}(\bar{n}_{i.}-n_{il})^2, \tag{A 6}$$

where $\bar{n}_{i.} = 1/I\sum_{l=1}^{I} n_{il}$. The determinants of eqn A 2 are calculated using $k-1$ of the allele types to avoid singularity (Mosimann, 1962).

In the case of the two-allele genetic model, an estimator of $\hat{C}$ is then,

$$\hat{C} = \frac{Ns_1^2}{\bar{n}_1(N-\bar{n}_1)}, \tag{A 7}$$

where $\bar{n}_1 = 1/I\sum_{i=1}^{I} n_{1i}$ and $s_1^2 = 1/I-1\sum_{i=1}^{I}(n_{1i}-\bar{n}_1)^2$. For asymptotically large N and I we then obtain,

$$\frac{1}{1+\theta} = \frac{\sigma_p^2}{\bar{p}(1-\bar{p})}, \tag{A 8}$$

where $\sigma_p^2 = s_1^2/N^2$ and $\bar{p} = \bar{n}_1/N$. Solving for θ produces an estimator that is identical in form to the one given above based on Wright's F_{ST} statistic demonstrating that Wright's estimator of θ is a MME ratio estimator of the scale parameter of the CMD distribution.

References

Barton, N. H., Halliday, R. B. & Hewitt, G. M. (1983). Rare electrophoretic variants in a hybrid zone. *Heredity* **50**, 139–146.

Barton, N. H. & Slatkin, M. (1986). A quasi-equilibrium theory of the distribution of rare alleles in a subdivided population. *Heredity* **56**, 409–416.

Brier, S. S. (1980). Analysis of contingency tables under cluster sampling. *Biometrika* **67**, 591–596.

Casella, G. & Berger, R. (1990). *Statistical Inference*. Belmont: Duxbury Press.

Chuang, C. & Cox, C. (1985). Pseudo maximum likelihood estimation for the Dirichlet-multinomial distribution. *Communications in Statistics: Theory and Methodology* **14**, 2293–2311.

Cockerham, C. C. & Weir, B. S. (1993). Estimation of gene flow from F-statistics. *Evolution* **47**, 855–863.

Collins, P. W. (1982). Origin and differentiation of the island fox: A study on evolution in insular populations. Master's thesis, University of California, Santa Barbara, CA.

Crow, J. F. & Aoki, K. (1984). Group selection for a polygenic behavioral trait: Estimating the degree of subdivision. *Proceedings of the National Academy of Sciences, USA* **81**, 6073–6077.

Edwards, S. V. (1993). Mitochondrial gene genealogy and gene flow among island and mainland populations of a sedentary songbird, the grey-crowned babbler (*Pomatostomus temporalis*). *Evolution* **47**, 1118–1137.

Georgiadis, N., Bischof, L., Templeton, A., Patton, J., Karesh, W. & Western, D. (1994). Structure and history of African elephant populations. I. Eastern and southern Africa. *Journal of Heredity* **85**, 100–104.

Gong, G. & Samaniego, F. J. (1981). Pseudo maximum likelihood estimation: Theory and applications. *The Annals of Statistics* **9**, 861–869.

Hudson, R. R., Slatkin, M. & Maddison, W. P. (1992). Estimation of levels of gene flow from DNA sequence data. *Genetics* **132**, 583–589.

Johnson, N. L. & Kotz, S. (1972). *Distributions in Statistics: Continuous Multivariate Distributions*. Wiley and Sons, New York.

Kimura, M. (1953). 'Stepping-stone' model of population. *Annual Report of the National Institute of Genetics, Japan* **3**, 62–63.

Kimura, M. & Weiss, G. H. (1964). The stepping stone model of population structure and the decrease of genetic correlation with distance. *Genetics* **59**, 561–576.

Levin, B. & Reed, J. (1977). Compound multinomial likelihood functions are unimodal: Proof of a conjecture of I. J. Good. *The Annals of Statistics* **5**, 79–87.

Lewontin, R. C. (1974). *The Genetic Basis of Evolutionary Change*. New York: Columbia University Press.

Malécot, G. (1948). *Les mathématiques de l'hérédité*. Freeman, San Francisco, CA. (1969 The mathematics of heredity, D. M. Yermanos (trans.).)

Maruyama, T. (1970). Effective number of alleles in a subdivided population. *Theoretical Population Biology* **1**, 273–306.

Maruyama, T. (1971). Analysis of population structure. II. Two-dimensional stepping stone models of finite length and other geographically structured populations. *Annals of Human Genetics* **35**, 179–196.

Mercure, A., Ralls, K., Koepfli, K. P. & Wayne, R. K. (1993). Genetic subdivision among small canids: Mitochondrial DNA differentiation of swift, kit and Arctic foxes. *Evolution* **47**, 1313–1328.

Mosimann, J. E. (1962). On the compound multinomial distribution, the multivariate β-distribution, and correlations among proportions. *Biometrika* **49**, 65–82.

Nei, M. (1973). Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences USA* **70**, 3321–3323.

Parke, W. R. (1986). Pseudo maximum likelihood estimation: The asymptotic distribution. *The Annals of Statistics* **14**, 355–357.

Rannala, B. & Hartigan, J. A. (1995). Identity by descent in island-mainland populations. *Genetics* **139**, 429–437.

Slatkin, M. (1985). Gene flow in natural populations. *Annual Review of Ecology and Systematics* **16**, 393–430.

Slatkin, M. (1994). Gene flow and population structure. In *Ecological Genetics*, (ed. L. Real), pp. 3–17. Princeton. NJ: Princeton University Press.

Slatkin, M. & Barton, N. H. (1989). A comparison of three indirect methods for estimating average levels of gene flow. *Evolution* **43**, 1349–1368.

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

Wayne, R. K., George, S. B., Gilbert, D., Collins, P. W., Kovach, S. D., Girman, D. & Lehman, N. (1991). A morphological and genetic study of the island fox *Urocyon littoralis*. *Evolution* **45**, 1849–1868.

Wehrhahn, C. F. (1989). Proceedings of the ecological genetics workshop. *Genome* **31**, 1098–1099.

Wehrhahn, C F. & Powell, R. (1987). Electrophoretic variation, regional differences, and gene flow in the coho salmon (*Oncorhynchus kisutch*) of southern British Columbia. *Canadian Journal of Fisheries and Aquatic Sciences* **44**, 822–831.

Weir, B. S. & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution* **38**, 1358–1370.

Wolfram Research, Inc. (1992). *Mathematica*. Wolfram Research, Inc., Champaign, IL. Version 2.2.

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

Wright, S. (1949). Adaptation and selection. In *Genetics, Paleontology and Evolution*, (ed. G. Jepson, G. Simpson & E. Mayr), pp. 365–389. Princeton, NJ: Princeton University Press.

Wright, S. (1951). The genetical structure of populations. *Annals of Eugenics* **15**, 323–354.

Wright, S. (1969). *Evolution and Genetics of Populations. The Theory of Gene Frequencies*, vol. 2. Chicago IL: University of Chicago Press.

Zink, R. M. & Remsen, J. V. (1986). Evolutionary processes and patterns of geographic variation in birds. In *Current Ornithology*, (ed. R. F. Johnston), pp. 1–69, vol. 4. New York: Plenum.