Estimating Effective Population Size and Migration Rates From Genetic Samples Over Space and Time

Jinliang Wang* 1 and Michael C. Whitlock†

*Institute of Zoology, Zoological Society of London, London NW1 4RY, United Kingdom and †Department of Zoology, University of British Columbia, Vancouver, British Columbia V6T 1Z4, Canada

Manuscript received July 24, 2002
Accepted for publication October 25, 2002

ABSTRACT

In the past, moment and likelihood methods have been developed to estimate the effective population size (Ne) on the basis of the observed changes of marker allele frequencies over time, and these have been applied to a large variety of species and populations. Such methods invariably make the critical assumption of a single isolated population receiving no immigrants over the study interval. For most populations in the real world, however, migration is not negligible and can substantially bias estimates of Ne if it is not accounted for. Here we extend previous moment and maximum-likelihood methods to allow the joint estimation of Ne and migration rate (m) using genetic samples over space and time. It is shown that, compared to genetic drift acting alone, migration results in changes in allele frequency that are greater in the short term and smaller in the long term, leading to under- and overestimation of Ne, respectively, if it is ignored. Extensive simulations are run to evaluate the newly developed moment and likelihood methods, which yield generally satisfactory estimates of both Ne and m for populations with widely different effective sizes and migration rates and patterns, given a reasonably large sample size and number of markers.

ONE of the major applications of population genetics theory has been in the estimation of genetic and demographic properties of populations. Quantities such as the mutation rate, the migration rate, and the recombination rate are all in principle estimable from population genetic data, at least as products with the effective population size.

The variance effective population size (Ne) is an important quantity in evolutionary biology, as a description of the rate at which genetic variance changes due to genetic drift. A variety of methods have been developed to estimate Ne (Schwartz et al. 1999), including predictive equations based on life history data (reviewed in Caballero 1994; Wang and Caballero 1999), the lethal allelism approach (Dobzhansky and Wright 1941), and a linkage disequilibrium approach (Hill 1981). Several authors have developed methods to estimate Ne for a population from temporal genetic data, that is, data on the frequencies of marker alleles taken at two or more time points from the same populations. In principle, if other population genetic processes such as mutation, migration, and selection are negligible in their effects on allele frequency change over the time frame of the study, then the observed changes in allele frequency can be assumed to be due to the effects of genetic drift. Using what is known about the effects of genetic drift, then, we could estimate Ne from such temporal data. This so-called “temporal approach” was pioneered by Krimbas and Tsakas (1971), with subsequent substantial statistical refinements by others (e.g., Pamilo and Varvio-Aho 1980; Nei and Tajima 1981; Pollak 1983; Tajima and Nei 1984; Waples 1989; Williamson and Slatkin 1999; Anderson et al. 2000; Wang 2001a). These methods have been applied to a large variety of species and populations (e.g., Funk et al. 1999; Labate et al. 1999; Fiumera et al. 1999, 2000; Kantanen et al. 1999; Turner et al. 1999; see also many earlier references cited in Williamson and Slatkin 1999).

A key assumption made by all these approaches, however, is that the only factor involved in allele frequency change is genetic drift. All systematic forces (selection, mutation, and migration) are assumed to be absent. In our view, it is often legitimate to ignore the effects of mutation on the timescale involved in most empirical studies. It also may be reasonable to neglect the effects of selection because direct selection on most markers may be unlikely to be strong enough to cause substantial changes in their frequencies. (The effects of indirect selection acting on other linked loci on the allele frequency change of the marker locus may arguably be included as a “drift” effect and legitimately included in a Ne term.) However, for most populations the effects of migration are not negligible. A major weakness of these temporal methods is that they ignore the effects of migration, which can substantially bias estimates of Ne, either upward or downward (see below). The pur-
pose of this article is to include the effects of migration in the estimation of demographic parameters from temporal genetic data.

Migration can affect the estimation of \( N_e \) in two ways. In the short term, migration can change allele frequencies quite quickly. Given that the temporal methods estimate \( N_e \) by measuring the pace at which allele frequencies change, migration causes a population to behave as if strong drift was changing allele frequencies very rapidly, resulting in an estimate of \( N_e \) that is too low. In the longer term, the problem is reversed. Constant migration and drift would cause the populations to approach an equilibrium level of genetic differentiation and the rate of change of allele frequencies in a deme to slow down to a speed that predicted by the effective size of the whole metapopulation. Thus the effective population size of the deme would be substantially underestimated in the long term. These effects are discussed later in this article in more mathematical detail.

This problem can be avoided by simultaneously estimating \( N_e \) and the migration rate, which is the approach taken in this article. It turns out that while both migration and drift affect allele frequency changes in a population, they do so differently in terms of the direction and magnitude of these changes. Under migration, the allele frequency of a focal population would tend to become increasingly similar to that of the source population. Under drift, however, the allele frequency of a focal population changes purely at random, irrespective of that of the source population. Furthermore, the magnitude of allele frequency change of a focal population depends on the allele frequency difference between the focal and source populations under migration. Migration has no effect on alleles that happen to be at the same frequencies both in the deme in question and in the populations providing those migrants, but migration is expected to change allele frequencies substantially in cases when the source and recipient demes have very different allele frequencies. Genetic drift, on the other hand, is not affected by differences in allele frequencies among demes, but depends only on the local allele frequency. These differences in the patterns of changes of allele frequencies due to migration and drift allow us to estimate them separately, so long as enough data are from multiple loci. Fortunately, these kinds of data are often available with modern biochemical methods.

The estimation of migration rate and migrant number has a long history. The oldest and most widely used genetic method assumes that populations are at equilibrium in a system approximated by Wright’s (1931) island model and then uses estimates of Wright’s \( F_{is} \) to calculate \( 1/(4N_e m + 1) \), where \( m \) is the local immigration rate (see Slatkin 1987). This method has been roundly criticized (Whitlock and McCauley 1999), in particular because of the unrealistic assumptions it requires about the demographic structure of the species. Other approaches include genetic stock identification (Milner et al. 1985), assignment methods (Ranala and Mountain 1997), two-locus descent measures (Vitalis and Couvet 2001), coalescent methods (Nielsen and Wakeley 2001), isolation-by-distance methods (see review by Roussset 2001), and migration matrix-likelihood models (Beerli and Felsenstein 2001). The first two approaches require detailed knowledge about the possible sources of migrants and do not allow for genetic drift. The latter four approaches assume that the populations are at an equilibrium between the effects of genetic drift and migration. There is a need for a method that does not make restrictive assumptions about genetic and demographic equilibria, yet allows simultaneous estimation of the effects of drift and migration.

Here we extend previous temporal methods for estimating the \( N_e \) of a single isolated population so that they can be used to estimate \( N_e \) and \( m \) jointly for a metapopulation. We first develop an expectation-based approach that does not make restrictive assumptions about genetic stock identification (e.g., a small number of demes in island or stepping-stone migration models) were checked by simulations. As an example for the application of our methods, a spatial and temporal genetic data set on Triturus newts (Jehle et al. 2001) was reanalyzed by our methods for the simultaneous estimation of migration and drift.

DEFINITIONS AND ASSUMPTIONS

We use temporal data to estimate the effective size \( (N_e) \) and immigration rate \( (m) \) of a partially isolated population. For brevity, we refer to the population for which parameters are being estimated as the focal population and that from which the migrants come as the source population. Data and parameters for the source population are given the subscript \( A \), and those for the focal population, \( B \). In the first section, we assume that the source population is infinitely large and therefore not changing genetically through the time frame of the study. In the section after that, we assume that both populations are finite and jointly estimate the properties of both. In the latter case, source and focal populations are obviously reciprocal. Throughout we assume dip-
For data from haploid species or loci, the \( N_e \) estimated by these models would be double the actual effective size.

For the infinite-source population model, a minimum of three samples is required to estimate \( N_e \) and \( m \) of the focal population: two temporally spaced ones from the focal population and one taken at any time from the source. For the two small population models, a minimum of four samples is necessary, two temporally spaced ones taken concurrently from each population. Let us denote the size of these samples \( S \), with subscripts \( A \) or \( B \) to describe whether the sample is from the source or focal population, respectively, and a subscript \( t \) to describe the generation in which the sample is taken. Thus \( S_{00} \) is the number of individuals sampled from the focal population in the initial generation. We use \( x \) to refer to the relative frequency of a particular allele in a given sample. Therefore \( x \) is an estimator of the true frequency \( (p) \) of that allele at that time in that population. For convenience, we define \( q = 1 - p \). These quantities, \( x \), \( p \), and \( q \), are subscripted in the same way as \( S \).

Several summary statistics of these quantities turn out to be useful. We define \( \Delta p_{AB} = p_{A0} - p_{B0} \) to be the initial difference between the allele frequencies of the two populations and \( \Delta p_{Bi} = p_{Bi} - p_{B0} \) to be the change in allele frequency by generation \( i \) in deme \( B \). In the derivations of the moment-based estimators, we also use \( \delta_{Bi} \) to refer to the change in generation \( i \) due to genetic drift of the allele frequency in population \( B \). Each of the quantities can refer to population \( A \) if the subscript \( B \) is replaced by \( A \).

We make several assumptions throughout. These are that the alleles under study are not under direct selection, that mutation is a negligible source of allele frequency change, that potential sources of migrants to the focal populations can be identified a priori, that all samples are randomly taken from their populations, and that the sampling event does not change the available pool of reproductive individuals. The latter can be realized by sampling after reproduction, sampling with replacement, or sampling from a population much larger than the sample. This sampling scheme therefore corresponds to Waples’ (1989) plan 1.

Table 1 summarizes most of the terms used in this article.

**Table 1**

<table>
<thead>
<tr>
<th>Term</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>( N_e )</td>
<td>Variance effective population size</td>
</tr>
<tr>
<td>( m )</td>
<td>Immigration rate</td>
</tr>
<tr>
<td>( A )</td>
<td>Source population</td>
</tr>
<tr>
<td>( B )</td>
<td>Focal population</td>
</tr>
<tr>
<td>( S )</td>
<td>Size of a sample</td>
</tr>
<tr>
<td>( t )</td>
<td>Time (in generations)</td>
</tr>
<tr>
<td>( p )</td>
<td>Population allele frequency</td>
</tr>
<tr>
<td>( x )</td>
<td>Sample allele frequency</td>
</tr>
<tr>
<td>( q )</td>
<td>( 1 - p )</td>
</tr>
<tr>
<td>( \Delta p_{AB} )</td>
<td>Initial difference between allele frequencies of the two populations</td>
</tr>
<tr>
<td>( \Delta p_{Bi} )</td>
<td>Change in allele frequency by generation ( i ) in deme ( B )</td>
</tr>
<tr>
<td>( \delta_{Bi} )</td>
<td>Change in ( p ) in generation ( i ) in population ( B ) due to genetic drift</td>
</tr>
<tr>
<td>( E[ ] )</td>
<td>Expectation operator</td>
</tr>
<tr>
<td>( Est[ ] )</td>
<td>Estimate</td>
</tr>
<tr>
<td>( Pr[ ] )</td>
<td>Probability</td>
</tr>
<tr>
<td>( L[ ] )</td>
<td>Likelihood</td>
</tr>
</tbody>
</table>

Immigration from an Infinite-Source Population

**Theoretical background**: Imagine that the focal population is receiving migrants at rate \( m \) per generation from an infinite, unchanging source population. For effectively neutral alleles, changes in allele frequency over a short period of time are a function of the local effective population size and migration rate. Mutations are negligible compared to drift and migration because of their rare occurrence and the short time span. Under these conditions, and assuming that migration precedes population regulation, the allele frequency at generation \( i \) is given by

\[
p_{Bi} = (1 - m)p_{Bi-1} + mp_A + \delta_{Bi}.
\] (1)

Note then that

\[
p_{B1} = p_{B0} + m\Delta p_{AB} + \delta_{B1}.
\] (2)

Expanding across generations, we can find

\[
p_{Bi} = p_{B0} + m\Delta p_{AB} \sum_{j=0}^{i-1} (1 - m)^j + \sum_{j=1}^{i} \delta_{Bj}(1 - m)^{i-j},
\] (3)

so

\[
\Delta p_{B0} = p_{B0} - p_{B0} = \Delta p_{AB}(1 - (1 - m)^i) + \sum_{j=1}^{i} \delta_{Bj}(1 - m)^{i-j}.
\] (4)

Under the assumptions of Wright-Fisher sampling, the \( \delta \) terms are independently binomially distributed. The expected value of all of the \( \delta \) terms is zero, because drift does not change on average the allele frequency. The variance of each of these terms is given by \( p_{B0}(1 - p_{B0})/N_e \) for haploids or \( p_{B0}(1 - p_{B0})/2N_e \) for diploids. In reality, we do not know \( p_{B0} \) for any generation between samples, and we can only estimate \( p_{Bi} \) for generations in which there was a sample. This uncertainty is the root of error in the estimation of statistics derived from these equations. In this article, we first estimate this error by using the approximate moments of the distribution of allele frequencies, and then we describe a likelihood approach to the estimation of \( N_e \) and \( m \).

**Moment approximations**: \( N_e \) and \( m \) affect the first and second moments of the change in allele frequency over time. Under certain circumstances, we can use these moments to estimate \( m \) and \( N_e \). We do this in a two-step process, first estimating \( m \) from the mean change in
allele frequencies at all loci and then using this value in the estimation of \(N_e\) from the second moment of allele frequency change.

Estimating the migration rate: The change in allele frequency is given by Equation 4 above. Taking the expectation of \(\Delta p_{\text{stat}}\), we find

\[
E[\Delta p_{\text{stat}}] = E[\Delta p_{\text{stat}}(1 - (1 - m)')] ,
\]

(5)
because the expected values of all the \(\delta\) terms are zero. Assuming that \(t\) is known without error, we can therefore make a one-locus estimate of \((1 - (1 - m)')\) from \(\Delta p_{\text{stat}}/\Delta p_{\text{stat}}\).

To predict the optimal sampling strategy and to approximate standard errors, we estimate the error variance of an estimate of \(m\). The expected error in estimating a ratio is an approximate function of the error variance of estimating the numerator and the denominator and the error covariance between the two (see Lynch and Walsh 1998, Appendix 1). The error in estimating \((1 - (1 - m))'\) can be approximated from the error in estimating \(\Delta p_{\text{stat}}/\Delta p_{\text{stat}}\). The error variances (\(V_{\text{err}}[\ ]\)) and error covariance (\(\text{COV}_{\text{err}}\)) are given by

\[
V_{\text{err}}[\Delta p_{\text{stat}}] = \frac{p_{AB} q_{AB} + p_{B} q_{B}}{2 S_{\text{eff}}},
\]

\[
V_{\text{err}}[\Delta p_{\text{stat}}] = \frac{p_{AB} q_{AB} + p_{B} q_{B}}{2 S_{\text{eff}}},
\]

\[
\text{COV}_{\text{err}}[\Delta p_{\text{stat}}, \Delta p_{\text{stat}}] = \frac{p_{AB} q_{AB}}{2 S_{\text{eff}}}.
\]

(6)

Note that the 2’s in the denominators drop out when these equations are applied to haploids. The approximation in the first equation here comes from using the initial value of \(pq\) to predict the amount of variance in final allele frequency due to genetic drift.

Using these last equations and Equation A1.19b from Lynch and Walsh (1998), we can find the expected error variance for the estimate of \((1 - (1 - m))'\):

\[
V_{\text{err}}[1 - (1 - m')] = \left(\frac{1}{\Delta p_{\text{stat}}}ight)^2 
\times \left(\frac{p_{AB} q_{AB} + p_{B} q_{B}}{2 S_{\text{eff}}} \left(1 - (1 - m')\right)^2 + \frac{1}{2 S_{\text{eff}}} + \frac{1}{2 S_{\text{eff}}} \left(1 - (1 - m')\right)^2\right).
\]

(7)

If the three sample sizes are roughly equivalent, then the error in estimating \(m\) is influenced more by the allele frequency in population \(B\) than by that in the source population, especially if the product \(mt\) is small. The terms involving only the sample sizes, \(m\), and \(t\) are presumably constant across loci, so the error variance is largely proportional to \(1/(\Delta p_{\text{stat}})^2\). The minimum variance possible when combining \(L\) loci can be obtained by weighting each locus by the inverse of its expected error variance. The best estimator of \(m\) is given approximately by

\[
m = 1 - t \sqrt{1 - \frac{1}{W} \sum_{i=1}^{L} w_i F_i},
\]

(8)

where the weight for locus \(l\) is \(w_l = (\Delta x_{AB})^2, F_i = \Delta x_{AB}/\Delta x_{AB}\), and \(W\) is the sum of \(w_l\) across loci.

Estimating \(N_e\): Under the assumption that the total change in allele frequency per generation is not large (i.e., \(N_e\) is not very small and \(m\) is not very large), each of the terms in the equation describing allele frequency change is roughly independent of the others. The variance in allele frequency change at generation \(i\) due to genetic drift within the focal population is \(V[\delta_{e,i}] = p_{AB} q_{AB}/2N_e\). If the allele frequency is not changing rapidly, then we can assume that the value of \(pq\) is approximately the same each generation. A reasonable estimator of the mean value of \(pq\) is \(\bar{pq} = (x_{AB}(1 - x_{AB}) + x_{AB}(1 - x_{AB}))/2\). This is implicitly assuming that the allele frequency on average takes a linear path from the initial to the final value, but in any case is an ad hoc approximation. Then using Equation 4, we can find

\[
E[\Delta p_{\text{stat}}^2] = \Delta p_{\text{stat}}^2(1 - (1 - m')^2) + M(\bar{pq})^2,\]

(9)

where \(M = (1 - (1 - m')^2)/(2 - m)\). Again, the coefficient of \(N_{e,B}\) of 2 drops out when this formula is used for haploids. From Equation 9 and using the estimate of \(m\), we obtain an estimator of \(1/(2N_{e,B})\),

\[
\text{Est} \left[\frac{1}{2N_{e,B}}\right] = E[\Delta p_{\text{stat}}^2] - \Delta p_{\text{stat}}^2(1 - (1 - m')^2)/M(\bar{pq}),
\]

(10)

where \(M = (1 - (1 - m')^2)/(2 - m)\).

In general we do not have exact information about the change in population allele frequencies, but only an estimate based on sample data. Because of sampling variance, \(E[\Delta x_{AB}^2]\) is biased upward as an estimator of \(\Delta p_{\text{stat}}^2\):

\[
E[\Delta x_{AB}^2] = \Delta p_{\text{stat}}^2 + \frac{p_{AB} q_{AB}}{2 S_{\text{eff}}} + \frac{p_{B} q_{B}}{2 S_{\text{eff}}},
\]

(11)

Thus we can estimate

\[
\text{Est} \left[\frac{1}{2N_{e,B}}\right] = \left[\Delta x_{AB}^2 - \frac{p_{AB} q_{AB}}{2 S_{\text{eff}}} + \frac{p_{B} q_{B}}{2 S_{\text{eff}}} \left(1 - (1 - m')^2\right)\right]/(M \bar{pq}).
\]

(12)

The error variance per locus for estimating \(1/(2N_e)\) is approximately proportional to \(1/\bar{pq}\) (calculations not shown), so a minimum-variance estimator of \(1/(2N_e)\) can be obtained by weighting alleles by \(\bar{pq}\) approximately, estimated by \((x_{AB}(1 - x_{AB}) + x_{AB}(1 - x_{AB}))/2\).

Choosing the right value of \(t\): With a large separation in time between the first and second samples, estimates of \(m\) become problematic. The estimate of \((1 - (1 - m')\)
becomes more and more similar for different values of \( m \) as \( t \) gets larger (\( \lim_{t \to \infty} (1 - (1 - m)^{1/2}/\sqrt{m}) = 0 \)), yet the error variance in estimating this quantity increases linearly with \( t \). Furthermore, as \( t \) increases, the allele frequency change approaches a distribution determined largely by the relative sizes of the effective size and \( m \), rather than by the absolute values of either. When \( t \) is large enough that the populations reach an equilibrium between genetic drift and migration, then \( m \to 0 \), irrespective of the actual true migration rate. This is clear by inspecting the single-locus estimator of \( m \), which is defined as \( \hat{m} = 1 - \sqrt{1 - \Delta x_{e,1}/\Delta x_{e,0}} \). As \( t \to \infty \), \( \Delta x_{e,1}/\Delta x_{e,0} \) no longer increases and \( \Delta x_{e,1}/\Delta x_{e,0} \) is expected to reach a constant smaller than one with time, therefore resulting in \( \hat{m} \to 0 \).

If the time interval is too short and \( m \) too small, then there is less power to measure \( m \) because too few migration events may have occurred. Larger samples with more informative markers are therefore necessary to obtain a good estimate of \( m \) when \( (1 - m)^{1/2} \) is close to one or zero.

Similarly, the estimation of \( N_e \) is also biased if the sampling interval is too long. As \( t \to \infty \), we have \( \hat{m} \to 0 \) shown above and Equation 10 reduces to \( \hat{m} = E[\Delta p_{e,1}]/(t\hat{m}) \). If the source population is infinite in size, then \( E[\Delta p_{e,1}] \) tends to a constant as \( t \to \infty \), and therefore \( \hat{N}_{e,b} \to \infty \). If the source population has a finite size, then \( E[\Delta p_{e,1}] \) continues to increase as \( t \to \infty \) (so long as the alleles are not fixed in the population), but at a slower rate corresponding to the total size of the focal and source populations. In both cases, the estimate of \( N_{e,b} \) approaches the effective size of the whole species rather than that of the focal population. Again, there is less power for estimating \( N_{e,b} \) if \( t \) and \( S/(N_{e,b}) \) are too small. With a small quantity of \( tS/(N_{e,b}) \), the change in allele frequency due to drift would be overwhelmed by that from sampling.

As a numerical example, Figure 1A shows the average estimates of \( 1/(2N_{e,b}) \) and \( m \) over 10,000 replicates plotted against sampling interval \( t \), and Figure 1B shows the coefficients of variation (CV) of the estimates when \( t = 1-12 \) where the estimation is unbiased. (Methods for these simulations are given in the Appendix.) The true values being estimated are \( 1/(2N_{e,b}) = 0.005 \) (\( N_{e,b} = 100 \)) and \( m = 0.2 \) for the focal population, which is assumed to be in a metapopulation containing 10 other populations of the same size (\( N_e = 100 \)) with an island migration model. The effective size of the source population is therefore \( N_e \sim 1000 \), sufficiently large although not infinite. As can be seen, estimates of both \( 1/(2N_{e,b}) \) and \( m \) are close to the true values when \( t \) is small, but approach the true values for the whole species, 0.00045 (\( N_e = 1112 \), calculated using the known parameters) and 0, with increasing \( t \). The maximum number of generations allowed for valid joint estimates of \( 1/(2N_{e,b}) \) and \( m \) and the number of generations required to reach the equilibrium so that the species \( N_e \) is estimated depend on \( m \), the effective sizes of the focal and source populations, and the initial genetic differentiation when the first sample was taken (see below).

For this specific numerical example, the estimation is unbiased when \( \sim t < 12 \). The precision of the estimators changes considerably over this short sampling period (Figure 1B). As expected, the best estimates are obtained at an intermediate value of \( t \) (\( t = 2-4 \) for the example). The optimal sampling interval depends, again, on \( m \), the effective size of the focal population and the initial genetic differentiation when the first
sample was taken. With decreasing $m$ and/or increasing initial differentiation, for example, the optimal $t$ is expected to increase.

In experimental design, it is difficult to determine beforehand the optimal interval between sampling for maximum power of the temporal method, because the power depends partially on the parameter being estimated (see e.g., Equation 7). Without any prior knowledge of the demographic history of the populations in question, it is reasonable to consider a short rather than a long sampling interval. While a too long sampling interval results in biased estimates of $N_e$ and $m$, a too short sampling interval only decreases the precision of the estimates. With powerful statistical methods (e.g., maximum likelihood shown below) and large sample size and marker information, the estimation precision can be improved greatly even if $t$ is as short as one generation.

The consequences of assuming that $m = 0$: Other methods using temporal data for estimating the effective size have assumed that the immigration rate is zero. We can use the equations above to investigate the effects of making this assumption when it is in fact false.

When $m \to 0$, the expected change in allele frequency is zero, $M = 1$, and Equation 12 reduces to

$$\text{Est} \left[ \frac{1}{2N_e} \right] = \frac{\Delta x_0^2}{\Delta x_0^2} - \frac{x_{th} x_{th} / (2 S_0) + x_{th} x_{th} / (2 S_0)}{1/\theta}$$

which is similar to Nei and Tajima’s (1981) estimator using $F_e$.

When $m > 0$, the variance in allele frequency change over the interval between samples is a function of both genetic drift and migration. This can be either greater or smaller than that expected by drift alone, depending on, among other terms, $t$. In the short term, the change in allele frequency due to the joint effects of migration and drift will be larger than that expected by drift alone. If $t = 1$, for example, we have $E[(p_{th} - p_{sh})^2] = \Delta p_{th}^2 m^2 + p_{sh}(1 - p_{sh})/2N_{e,b}$. In this case, therefore, the squared change in allele frequency is increased relative to that due to drift only, which is $p_{sh}(1 - p_{sh})/2N_{e,b}$. Falsely assuming that $m = 0$ would on average cause an overestimate of the amount of drift and an underestimate of $N_{e,b}$. In contrast, as the time between samples gets larger, the change in allele frequency is constrained by migration to a term only somewhat larger than the squared initial difference in allele frequency between the source and focal populations: $\lim_{t \to \infty} E[(p_{th} - p_{sh})^2] = \Delta p_{th}^2 + p_{sh}(1 - p_{sh})/(N_{e,b} m(2 - m))$. With an increasing sampling interval ($t$), therefore, a false assumption of $m = 0$ would cause a downward bias in the estimation of the amount of genetic drift and an upward bias in estimating $N_e$. As the number of generations between samples continues to get larger, however, the estimate would approach the effective size of the whole species, rather than just the effective size of the focal population, no matter whether the assumption of $m = 0$ is made or not. These biases can be indefinitely large.

Figure 2 shows the biases of $N_e$ estimation when a false assumption of $m = 0$ is made. The true parameter values are $N_e = 100$ and $m = 0.1$, and the populations are assumed to be either completely differentiated (no previous migration, case A) or at drift-migration equilibrium (case B, where $F_e = 0.02$) when the initial sample is taken. Ignoring migration results in an overestimation of $1/(2N_e)$ when sampling interval ($t$) is small and underestimation when $t$ is large. Compared with case A, the initial overestimation of $1/(2N_e)$ in case B from ignoring migration is much smaller, because the populations are more similar genetically ($\Delta p_{th}$ small) and therefore migrants have less effect on allele frequency changes. With a larger $m$ and smaller $N_e$, the initial overestimation of $1/(2N_e)$ can be dramatic even when the populations are initially at equilibrium (data not shown). For both cases, the estimation is reasonably good for a large range of sampling interval when immigration is accounted for. However, the sampling interval cannot be too long. Otherwise, $1/(2N_e)$ is increasingly underestimated with $t$, approaching the true value for the whole species.

Likelihood model: Let us first consider a single locus with two alleles, and we focus on a particular allele. For brevity in this subsection, subscript $B$ is dropped for the focal population. To estimate $m$ and $N_e$ of the focal population jointly, at least two temporally spaced samples from the focal and a single sample from the source
population \((A)\) are necessary. Because the three sampling events are independent, the probability of getting a set of samples with allele frequencies denoted by \(x\) is
\[
\Pr[x_0, x_1, x_2 | p_0, p_1] = \Pr[x_0 | p_0] \Pr[x_1 | p_1] \Pr[x_2 | p_2]. \quad (14)
\]
Each of these samples can be assumed to follow a binomial distribution (see Waples 1989 for a discussion of when this is appropriate); therefore
\[
\Pr[x | p] = \frac{2S}{2Sx_1} p^x (1 - p)^{2S(1 - x)},
\]
where \(S\) is the number of diploid individuals sampled. For haploid species, the coefficient 2 of \(S\) should be dropped.

The likelihood \((L)\) of a particular set of values of \(N_e\) and \(m\) is the probability of observing the data given those values. We have, therefore,
\[
L(N_e, m | x_0, x_1, x_2) = \Pr[x_0, x_1, x_2 | N_e, m] = \frac{\prod_{h=0}^{m} \Pr[x_0, x_1, x_2 | p_0, p_1] \Pr[p_0, p_1 | N_e, m]}{\Pr[p_0 | N_e] \Pr[p_1 | N_e]}.
\]
As we can see, the key problem for the likelihood function is to calculate \(\Pr[p_0, p_1 | N_e, m]\). By the definition of conditional probability, we have
\[
\Pr[p_0, p_1 | N_e, m] = \Pr[p_0 | N_e] \Pr[p_1 | N_e, p_0] \Pr[p_0, p_1 | N_e, m]. \quad (16)
\]
Determining the second term in Equation 16, \(\Pr[p_0, p_1 | N_e, m]\), is potentially quite tricky. If we assume, however, that we do not know anything about how close the focal population is to a migration-drift equilibrium nor about the history of these populations, then this can be given reasonably as a product of the uniform distributions of \(p_0\) and \(p_1\), \(\Pr[p_0, p_1 | N_e, m] = \Pr[p_0 | N_e] \times \Pr[p_1 | N_e]\).

As Williamson and Slatkin (1999) point out, the range of the possible values for the uniform prior, \(\Pr[p_0 | N_e]\), will depend on whether only alleles segregating in the initial population are used. There are \(2N_e + 1\) possible configurations for a biallelic locus in a population with size \(N_e\), of which \(2N_e - 1\) are polymorphic. With the uniform prior, therefore, any possible configuration for the starting frequency \(p_0\) has an equal probability of either \(1/(2N_e + 1)\) or \(1/(2N_e - 1)\) if we consider all or only polymorphic configurations, respectively. For the source population, we use a discrete approximation, \(\Pr[p_0 | j / J] = j / (J + 1)\) for \(j = 0, 1, 2, \ldots, J\). The estimation does not change essentially once \(J\) is reasonably large (say, 1000), as expected.

The first term in Equation 16, \(\Pr[p_0 | N_e, m]\), can be calculated by using the transition matrix \(T\) (Ewens 1979). The probability of a transition from the \(j\)th configuration of the parental population at time \(t - 1\) to the \(i\)th configuration of the offspring population at time \(t\) is
\[
T_{ij} = \binom{2N_e}{i} (p^i)(1 - p^i)^{2N_e-i},
\]
where \(p^i = (j/2N_e)(1 - m) + mp_j\) and \((i, j = 0, \ldots, 2N_e)\). This accounts for the deterministic change in allele frequency due to migration as well as the genetic sampling due to drift. Given the initial distribution of allele frequency of the focal population \(P_0\) [a column vector of \(2N_e + 1\) elements with values \(1/(2N_e + 1)\) or \(1/(2N_e - 1)\) as assumed above] and values of \(p_0, N_e\), and \(m\), the distribution after \(t\) generations, \(P_t\), is calculated by the recurrence equation
\[
P_t = T^t P_0. \quad (17)
\]
These are all the parts necessary to calculate the likelihood, above. For multiple independent, biallelic loci, the joint likelihood of \(N_e\) and \(m\) is calculated as the product of the likelihoods for each locus.

For loci with more than two alleles, the calculation of the likelihood function is problematic. The number of possible configurations, \(C = (2N_e + k - 1)!/(2N_e! (k - 1)!)\) for a diploid locus with \(k\) alleles (Feller 1950; Williamson and Slatkin 1999), increases rapidly with \(k\), prohibiting the evaluation of the exact likelihood due to the limitations of computer memory and processing speed. For a moderate value of \(N_e\), even \(k = 3\) poses a serious computational problem for the likelihood method. Anderson et al. (2000) proposed a Monte Carlo method to evaluate the likelihood function with data on multiallelic markers. Their method does not need much storage but is highly demanding computationally and may not converge when \(k\) is large (say, \(k = 10\)) or some alleles are not observed in all samples. To simplify the likelihood computation for multiallelic loci, we follow our previous approach (Wang 2001), which transforms a \(k\) allelic locus into \(k\) biallelic “loci,” each having one of the \(k\) alleles with all the other alleles pooled. The overall log-likelihood is approximated by the sum of the log-likelihoods across loci multiplied by the factor of \((k - 1)/k\) to account for the dependence among such converted loci.

The likelihood method can also use more than two samples from the focal population simultaneously to obtain a mean \(N_e\) and an average \(m\) during the whole sampling period. If samples are taken from the focal population at generations \(t_0, t_1, t_2, \ldots, t_g\), then the joint likelihood of \(N_e\) and \(m\) is
\[
L(N_e, m | x_0, x_1, x_2, \ldots, x_g) = \frac{\prod_{h=0}^{m} \Pr[x_0, x_1, x_2, \ldots, x_g | p_0, p_1, \ldots, p_g] \Pr[p_0, p_1, \ldots, p_g | N_e, m]}{\Pr[p_0 | N_e] \Pr[p_1 | N_e] \cdots \Pr[p_g | N_e, p_{g-1}]}.
\]
Similarly, a continuous growth model, such as exponential growth, could be fitted to the temporal data and joint estimates of initial effective size, growth rate, and immigration rate can be obtained (Wang 2001). A similar model to account for the changes of immigration
rate over the sampling period could be developed using multiple samples.

TWO FINITE DEMES

Up to this point, we have considered the demographic properties of a single deme that receives immigrants from an effectively infinite source, that is, a source so large that over the course of the study it is not changing appreciably in allele frequency. While this is likely to be a reasonable approximation in many cases, there are examples in nature where only a few relatively small demes serve as the only source of migrants to each other. This case is more difficult, both analytically and computationally, but here we consider the case of two small demes as a beginning to such an analysis.

The allele frequency at a neutral locus will change over a single generation in the two populations according to

\[
P_{\alpha,i} = (1 - m_{\alpha}) p_{\alpha,i-1} + m_{\alpha} p_{\beta,i-1} + \delta_{\alpha,i}
\]

\[
p_{\beta,i} = (1 - m_{\beta}) p_{\beta,i-1} + m_{\beta} p_{\alpha,i-1} + \delta_{\beta,i}.
\] (19)

Moment estimator: We can find a moment estimator for the two migration rates for this two-finite-deme case, if we consider the change in the difference in allele frequency between the demes over time. We consider the simple case in which the samples are separated by only a single generation, and so we drop the subscript notation for time. Taking the expectations of Equation 19 above and remembering that the expected values of \(\delta_{\alpha,i}\) and \(\delta_{\beta,i}\) are zero, we find

\[
m_{\alpha} = \frac{\Delta p_{\alpha}}{\Delta p_{AB}}
\]

\[
m_{\beta} = \frac{\Delta p_{\beta}}{\Delta p_{AB}}.
\] (20)

The error variance in estimating migration rates would be proportional to \(1/\Delta p_{AB}^2\), so we can find a low variance estimator of the migration rates by weighting information from each allele by \(\Delta p_{AB}^2\).

Finding an estimator of the effective size of each deme is also possible. The variance of the change in allele frequency over time due to drift for population \(A\) is

\[
V[\delta_{\alpha}] = E[\delta_{\alpha}^2] = \frac{p_{\alpha}^4 (1 - p_{\alpha}^4)}{2N_{e,A}}
\] (21)

where \(p_{\alpha}^4\) is the expected allele frequency after migration, which can be estimated from Equation 19 by setting \(\delta_{\alpha} = 0\) and using the estimated migration rates and initial allele frequencies. Thus the effective size can be estimated in this case in a manner similar to the Waples technique, except instead of using the initial allele frequency, one uses the expected initial frequency after migration. Similar derivations for \(N_{e,B}\) can also be obtained.

For the case of \(t > 1\), moment estimators for both \(m_{\alpha}\) and \(N_{e}\) become much more complicated (especially the latter) and are not shown here.

Likelihood estimator: The likelihood of the effective sizes \((N_{e,A}, N_{e,B})\) and immigration rates \((m_{\alpha}, m_{\beta})\) of demes \(A\) and \(B\) given the temporal data \((x_{A,0}, x_{A,t}, x_{B,0}, x_{B,t})\) is

\[
L(N_{e,A}, N_{e,B}, m_{\alpha}, m_{\beta}) = \sum_{r_{A,0}, r_{B,0}} \times Pr(x_{A,0}| r_{A,0})Pr(x_{A,t}| r_{A,0})Pr(x_{B,0}| r_{B,0})Pr(x_{B,t}| r_{B,0})
\]

\[
\times Pr(x_{A,0}, x_{A,t}, x_{B,0}, x_{B,t}| r_{A,0}, r_{B,0})
\] (22)

In Equation 22, each of the four samples follows a binomial distribution for \(Pr(x|p)\), the initial allele frequency of each population is assumed to be in a uniform distribution, and the allele frequency distribution of each population at generation \(t\) can be calculated by using the transition matrix.

In principle, the exact likelihood could be worked out for any set of the four parameter values, but in practice this would be computationally intensive for anything but small population sizes and small sampling interval \((t)\). The total number of configuration combinations that need to be considered in Equation 22 is \(((2N_{e,A} + 1) (2N_{e,B} + 1))\) for a biallelic locus, which increases exponentially with sampling interval \(t\). Therefore, we use the transition matrix method for the exact calculation of the likelihood for samples separated by only a single generation \((t = 1)\) and use a Monte Carlo approach to approximate the likelihood for samples with \(t > 1\).

Equation 22 could easily be extended to more than two samples from each deme with little increase in computational demands. It could also be extended to more than two demes, but the computational intensity increases prohibitively with the number of demes, because of the exponential increases in configuration combinations and the number of parameters for joint estimation.

**SIMULATION RESULTS**

**Infinite source:** Methods for these simulations are summarized in the APPENDIX. For focal populations with wide ranges of effective size and immigration rates, the maximum-likelihood and moment estimates of \(1/(2N_{e})\) and \(m\) from simulations are summarized in Table 2. The sample sizes used in the simulations are large, but not unrealistic. Indeed, a large sample size is important for the temporal methods; otherwise, the apparent changes in allele frequency would be dominated by sampling error rather than by genetic drift. For most natural populations, the actual census population sizes could be orders of magnitude larger than the effective sizes (FRANKHAM 1995), making both large samples and significant genetic drift possible.
<table>
<thead>
<tr>
<th>Parameters</th>
<th>ML estimates of $1/(2N_e)^a$</th>
<th>ML estimates of $m$</th>
<th>MT estimates of $1/(2N_e)^a$</th>
<th>MT estimates of $m$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>2.5%</td>
<td>97.5%</td>
</tr>
<tr>
<td>$N_e$</td>
<td>$m$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>0.975</td>
<td>0.283</td>
<td>0.465</td>
</tr>
<tr>
<td>0.01</td>
<td>0.996</td>
<td>0.249</td>
<td>0.336</td>
<td>1.567</td>
</tr>
<tr>
<td>0.1</td>
<td>0.999</td>
<td>0.270</td>
<td>0.506</td>
<td>1.599</td>
</tr>
<tr>
<td>0.5</td>
<td>0.908</td>
<td>0.292</td>
<td>0.382</td>
<td>1.563</td>
</tr>
<tr>
<td>100</td>
<td>0</td>
<td>0.982</td>
<td>0.245</td>
<td>0.519</td>
</tr>
<tr>
<td>0.01</td>
<td>0.972</td>
<td>0.254</td>
<td>0.519</td>
<td>1.484</td>
</tr>
<tr>
<td>0.1</td>
<td>0.991</td>
<td>0.265</td>
<td>0.514</td>
<td>1.531</td>
</tr>
<tr>
<td>0.5</td>
<td>0.938</td>
<td>0.293</td>
<td>0.427</td>
<td>1.533</td>
</tr>
<tr>
<td>1000</td>
<td>0</td>
<td>1.047</td>
<td>0.260</td>
<td>0.512</td>
</tr>
<tr>
<td>0.01</td>
<td>1.055</td>
<td>0.241</td>
<td>0.583</td>
<td>1.488</td>
</tr>
<tr>
<td>0.1</td>
<td>0.993</td>
<td>0.239</td>
<td>0.500</td>
<td>1.452</td>
</tr>
<tr>
<td>0.5</td>
<td>1.016</td>
<td>0.280</td>
<td>0.500</td>
<td>1.613</td>
</tr>
</tbody>
</table>

The simulated focal populations had various effective sizes (first column) and immigration rates (second column). Twenty loci, each having an equal number of alleles with initial frequencies drawn from a uniform distribution, were used in the estimation. A single sample from the large source and two samples separated by two generations from the focal population were taken with replacement. For populations with $N_e = 10$, 100, and 1000, the sizes of the three samples are 100, 200, and 1000 diploid individuals, respectively; the numbers of alleles per locus are two, four, and eight, respectively; and the numbers of replicates are 10,000, 1000, and 100, respectively. The estimation is summarized by the mean, standard deviation, and 2.5 and 97.5 percentiles of the estimates of $1/(2N_e)$ and $m$.

*The estimates of $1/(2N_e)$ are expressed relative to the true values to facilitate comparisons across wide ranges of the true parameter $N_e$. 
Figure 4.—Log likelihood as a function of estimates of $N_e$ and $m$ for a particular simulated data set with true parameter values $N_e = 100$ and $m = 0.1$. A single sample from the large source and two samples separated by two generations from the focal population were taken with replacement, the sizes of the three samples being 200 diploid individuals. Twenty loci, each having four alleles, were used in the estimation.

Figure 3.—The frequency distributions of $N_e$ (A) and $m$ (B) estimates from maximum-likelihood (thick lines) and moment (thin lines) estimators over 10,000 replicates. Two samples separated by four generations are taken from the focal population, and one sample is taken from the infinite-source population. Each sample contains 100 diploid individuals, which are genotyped for 10 loci, each having four alleles. The true parameter values are $N_e = 100$, $m = 0.1$, indicated by the dashed vertical lines.

The maximum-likelihood estimator gives unbiased estimates of $1/(2N_e)$ in all cases considered. The coefficient of variation of the estimates is roughly one-quarter, with 95% of the estimates being in the range of $\sim 50$ and 150% of the true values. The moment estimator underestimates $1/(2N_e)$ slightly when $m$ is large and sample size is small relative to $N_e$. Increasing sample size can decrease the bias dramatically. For the case of $N_e = 100$ and $m = 0.5$ in Table 2, for example, the mean ± SD of the moment estimates of $1/(2N_e)$ improves to $0.9 \pm 0.3$ (relative to the true value) when the sample size is 500 individuals. The estimates from the moment estimator are generally more dispersed than those from the maximum-likelihood estimator.

The maximum-likelihood method overestimates $m$ slightly only when $m$ is small in very small populations ($N_e = 10$). This is expected because maximum-likelihood estimates have a lower bound at zero. Increasing sample size or marker information (number of loci and alleles) in a sample decreases the bias and increases precision. In contrast to $1/(2N_e)$, both the moment and the maximum-likelihood methods give similar estimates of $m$. However, some moment estimates of $m$ are negative when $m$ is small.

In general, both maximum-likelihood and moment methods give reasonable estimates of $N_e$ and $m$. The accuracy and precision of the two estimators are similar for $m$, but the likelihood method is better for $N_e$. This is partially because moment methods must use the estimate of $m$ in estimating $N_e$. Such a two-step estimation procedure is expected to be inferior to the joint estimation of $m$ and $N_e$ from the likelihood approach, because the estimation error of $m$ is not accounted for in estimating $N_e$. Also, when $m$ is large and therefore the changes of allele frequencies are rapid, the approximation to the mean value of $pq$ in deriving the moment estimator of $N_e$ may be no longer good enough.

The distribution of the estimates of $N_e$ [or $1/(2N_e)$] generally has a long tail, with a significant proportion of the estimates being much larger (smaller) than the true value. This is similar to the observation made on the case of a single isolated population (Nei and Tajima 1981; Waples 1989; Wang 2001a). For the particular parameter set of $N_e = 100$, $m = 0.1$, the distributions of $N_e$ and $m$ estimates over 10,000 replicates are shown in Figure 3. Again, there is very little difference in the distribution of estimates of $m$ between the two techniques, the mean ± SD being $0.114 \pm 0.048$ for maximum likelihood and $0.115 \pm 0.049$ for the moment estimator. For $N_e$ estimation, however, maximum likelihood is much more accurate and precise than the moment approach. The mean ± SD of estimated $1/(2N_e)$
TABLE 3

Maximum-likelihood and moment estimates of $1/(2N_e)$ and $m$ relative to their true values, using different numbers of samples

<table>
<thead>
<tr>
<th>No. of samples</th>
<th>Relative estimates of $1/(2N_e)$</th>
<th>Relative estimates of $m$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>2</td>
<td>1.013</td>
<td>0.286</td>
</tr>
<tr>
<td>3</td>
<td>1.005</td>
<td>0.190</td>
</tr>
<tr>
<td>4</td>
<td>1.013</td>
<td>0.156</td>
</tr>
<tr>
<td>5</td>
<td>1.020</td>
<td>0.134</td>
</tr>
<tr>
<td>6</td>
<td>1.014</td>
<td>0.123</td>
</tr>
</tbody>
</table>

The estimates of $1/(2N_e)$ and $m$, relative to the true parameter values $1/(2N_e) = 0.01$ and $m = 0.2$, are obtained from the maximum-likelihood and moment estimators using a single sample from the source and different numbers of samples taken from the focal population every two generations. The size of each sample is 50 individuals, and 10 loci, each having 10 alleles, are used in the estimation.

The estimates of $1/(2N_e)$ and $m$ are 102 and 0.107, respectively, with 95% confidence intervals being 65–172 and 0.038–0.191, respectively. Multiple samples over time can increase the estimation precision for the average $1/(2N_e)$ and $m$ during the entire sampling period. Table 3 summarizes the estimates of $1/(2N_e)$ and $m$, relative to the true parameter values $1/(2N_e) = 0.01$ and $m = 0.2$, from maximum-likelihood and moment estimators using two to six temporal samples from the focal population. As expected, increasing the number of samples has no influence on the accuracy, but increases the precision dramatically, as indicated by the decreasing SD and the width between the 2.5 and 97.5 percentiles. The likelihood method gives unbiased estimation of $1/(2N_e)$ and $m$, while the moment estimator results in underestimation of $1/(2N_e)$ and overestimation of $m$. The poor accuracy of the moment estimator is partially due to the highly polymorphic loci (10 alleles per locus) and small sample sizes, which lead to a high proportion of rare alleles in the samples (Wang 2001a).

Quasi-infinite source, the island model and the stepping-stone model: In reality, the source population is not infinite in size, although it could be very large. Moreover, there can be different numbers of source populations for a focal population with migration in different patterns. As long as we can identify the source populations $a$ priori, however, the parameters for the focal population can be estimated approximately using the methods developed for the infinite-source model by treating the source populations collectively as the single “infinite” source. Here we explore numerical examples to show that the infinite-source model is very robust and works well for small source populations with various models of migration.

Consider a metapopulation consisting of 21 populations of equal effective size $N_e$. Each population receives, at each generation, a proportion of $m$ migrants taken equally from each of the remaining 20 populations in the island model or from each of the two neighboring populations in the circular stepping-stone model. For the estimation of $m$ and $N_e$ of a focal population, we take two samples separated by two generations from the focal and each of four source populations (the island model) or each of the two neighboring populations (the stepping-stone model). Each sample had 50 individuals and two alleles per locus for $N_e = 10$ and 100 individuals and eight alleles per locus for $N_e = 100$. The estimates of $1/(2N_e)$ and $m$ relative to their true values are summarized in Table 4 for maximum-likelihood and moment estimators. In all cases considered, maximum likelihood
### Table 4

Relative maximum-likelihood (ML) and moment (MT) estimates of $1/(2N_e)$ and $m$ for quasi-infinite populations

<table>
<thead>
<tr>
<th>Parameters</th>
<th></th>
<th>Relative ML estimates of $1/(2N_e)$</th>
<th></th>
<th>Relative ML estimates of $m$</th>
<th></th>
<th>Relative MT estimates of $1/(2N_e)$</th>
<th></th>
<th>Relative MT estimates of $m$</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$N_e$</td>
<td>$m$</td>
<td>Mean</td>
<td>SD</td>
<td>2.5%</td>
<td>97.5%</td>
<td>Mean</td>
<td>SD</td>
<td>2.5%</td>
</tr>
<tr>
<td>Island model</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.01</td>
<td>1.041</td>
<td>0.314</td>
<td>0.493</td>
<td>1.709</td>
<td>1.353</td>
<td>2.199</td>
<td>0.005</td>
<td>7.271</td>
</tr>
<tr>
<td></td>
<td>0.1</td>
<td>0.916</td>
<td>0.331</td>
<td>0.393</td>
<td>1.699</td>
<td>0.956</td>
<td>0.661</td>
<td>0.002</td>
<td>2.421</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.873</td>
<td>0.346</td>
<td>0.324</td>
<td>1.572</td>
<td>0.902</td>
<td>0.400</td>
<td>0.254</td>
<td>1.999</td>
</tr>
<tr>
<td>100</td>
<td>0.01</td>
<td>1.084</td>
<td>0.245</td>
<td>0.672</td>
<td>1.591</td>
<td>1.018</td>
<td>0.668</td>
<td>0.111</td>
<td>2.533</td>
</tr>
<tr>
<td></td>
<td>0.1</td>
<td>1.012</td>
<td>0.241</td>
<td>0.555</td>
<td>1.529</td>
<td>0.860</td>
<td>0.354</td>
<td>0.229</td>
<td>1.581</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>1.044</td>
<td>0.257</td>
<td>0.604</td>
<td>1.619</td>
<td>0.903</td>
<td>0.244</td>
<td>0.461</td>
<td>1.465</td>
</tr>
<tr>
<td>Circular stepping-stone model</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.01</td>
<td>1.262</td>
<td>0.320</td>
<td>0.595</td>
<td>1.897</td>
<td>1.291</td>
<td>1.805</td>
<td>0.005</td>
<td>6.325</td>
</tr>
<tr>
<td></td>
<td>0.1</td>
<td>1.070</td>
<td>0.326</td>
<td>0.511</td>
<td>1.858</td>
<td>1.131</td>
<td>0.772</td>
<td>0.001</td>
<td>2.947</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.930</td>
<td>0.338</td>
<td>0.348</td>
<td>1.711</td>
<td>1.001</td>
<td>0.477</td>
<td>0.214</td>
<td>1.999</td>
</tr>
<tr>
<td>100</td>
<td>0.01</td>
<td>1.078</td>
<td>0.228</td>
<td>0.690</td>
<td>1.544</td>
<td>1.131</td>
<td>0.789</td>
<td>0.008</td>
<td>2.887</td>
</tr>
<tr>
<td></td>
<td>0.1</td>
<td>1.030</td>
<td>0.235</td>
<td>0.604</td>
<td>1.542</td>
<td>0.995</td>
<td>0.421</td>
<td>0.232</td>
<td>1.894</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.977</td>
<td>0.251</td>
<td>0.511</td>
<td>1.504</td>
<td>0.991</td>
<td>0.337</td>
<td>0.452</td>
<td>1.995</td>
</tr>
</tbody>
</table>

The simulated metapopulation contains 21 populations of equal $N_e$ in an island model or a circular stepping-stone model. The focal populations have various effective sizes (first column) and immigration rates (second column). Twenty loci, each having an equal number of alleles with initial frequencies drawn from a uniform distribution, are used in the estimation. Two samples of equal size separated by two generations are taken from the focal and each of four source populations (the island model) or two neighboring populations (the circular stepping-stone model). For populations with $N_e = 10$ and 100, the sizes of the samples are 50 and 100 diploid individuals, respectively, and the numbers of alleles per locus are two and eight, respectively. In each case, 1000 replicates are run and the estimation is summarized by the mean, standard deviation, and 2.5 and 97.5 percentiles of the estimates of $1/(2N_e)$ and $m$ relative to their true values.
provided reasonable estimates of both $1/(2N_e)$ and $m$, while the moment estimator gives satisfactory estimation of $1/(2N_e)$ and $m$ only when drift is not very weak relative to migration. Otherwise, $1/(2N_e)$ is underestimated, similar to the situation of an infinite large source population shown in Table 2.

We also tested the infinite-source model with the extreme case of two small populations of equal $N_e$ and $m$, shown in Table 5. Although the source population ($A$) is small, the estimation of both $1/(2N_e)$ and $m$ of the focal population ($B$) is still not bad, at least within a factor of two. The moment estimator gives similar results (not shown). Like the results shown above, the worst case occurs with an extremely high migration rate, where populations can quickly reach the drift-migration equilibrium. In this case, increasing sample size improves the estimation considerably. For the situation of $N_e = 100$ and $m = 0.5$ in Table 5, for example, the mean (relative) estimate $\pm$ SD is $0.969 \pm 0.204$ for $1/(2N_e)$ and $0.809 \pm 0.204$ for $m$ when the sample size is increased to 400 individuals.

Two-deme models: As a numerical example for the case of two demes with sampling interval $t = 1$, Figure 5 shows the frequency distributions of the estimates of $N_{e,A}$, $m_A$, $N_{e,B}$, and $m_B$. As can be seen, both estimators give reasonably good estimates of effective size and migration rate, the precision of the likelihood estimator being always higher than that of the moment estimator.

For the case of sampling interval $t > 1$, the moment estimator for $N_e$ becomes quite complicated and the likelihood estimator becomes computationally intensive. In principle, however, both methods are not limited by the sampling interval, except for the case of very large $t$ where populations are approaching an equilibrium between drift and migration (Figures 1 and 2). At intermediate values of $t$, both methods should actually be more powerful than at $t = 1$ (Figure 1B). Because of the computational demands of the Monte Carlo approach, we have run only a few replicates for $t > 1$ for the likelihood method, yielding reasonable estimates of both effective size and migration rate, as expected (data not shown).

**Application to a Real Data Set**

Jehle et al. (2001) used microsatellite markers to investigate the spatial and temporal genetic differentiation of syntopic newts (Triturus cristatus, *T. marmoratus*). Individuals from three ponds in western France, situated 4–10 kilometers apart and inhabited by both species, were sampled and genotyped for eight microsatellite loci. Here we focus on *T. marmoratus* and consider the subpopulations in ponds 1 and 2 only for which temporal samples are available. The subpopulation in pond 3 is ignored because it is much farther away from the other two and has only a single sample. The first sample was collected as sacrificed larvae in 1989 (pond 1) or 1986 (pond 2), and the second was taken nondestructively from both larvae and adults in both ponds in 1998. The allele counts for five polymorphic loci were converted from the original data on allele frequencies and sample sizes listed in the Appendix of Jehle et al. (2001), and the data for adults and larvae in 1998 were combined for the likelihood analysis. The generation interval was estimated to be 6 years, and the sampling interval is therefore $\sim 1.5$ and 2 generations for ponds 1 and 2, respectively (Jehle et al. 2001). Because of the unsynchronized sampling, it is not possible to apply our

---

**Table 5**

Relative maximum-likelihood (ML) estimates of $1/(2N_e)$ and $m$ for a two-deme population under the infinite-source model

<table>
<thead>
<tr>
<th>$m$</th>
<th>$N_e = 10$</th>
<th>$N_e = 100$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>0.01</td>
<td>1.077</td>
<td>0.286</td>
</tr>
<tr>
<td>0.10</td>
<td>0.940</td>
<td>0.312</td>
</tr>
<tr>
<td>0.50</td>
<td>0.845</td>
<td>0.308</td>
</tr>
</tbody>
</table>

The simulated metapopulation contains two populations ($A$ and $B$ as the source and focal populations, respectively) of equal $N_e$ migrating with equal rate $m$. Twenty biallelic loci are used in the estimation. Two samples, each containing 50 (for $N_{e,A} = N_{e,B} = 10$) or 100 (for $N_{e,A} = N_{e,B} = 100$) individuals, separated by two generations are taken from each population, and the two samples from population $A$ are combined as a single sample. In each case, 1000 replicates are run and the estimation is summarized by the mean, standard deviation, and 2.5 and 97.5 percentiles of the estimates of $1/(2N_e)$ and $m$ relative to their true values.
method for the two-small-population model. However, we can use our method for the infinite-source model as an approximation to estimate the effective size and immigration rate for one pond, with the other pond acting as the source population and samples from it pooled. Since only integer numbers for sampling intervals are allowed in likelihood estimation, the analysis for pond 1 was carried out by using a sampling interval of 2, and the estimates of \( N \) and \( m \) \((N_e \text{'s}, \text{'s}')\) were converted by \( N_e = (1.5/2)N \text{'s} \) and \( m = 1 - e^{2/15}\log(1-m') \), respectively. Similar results were obtained by using a sampling interval of 1.

The relative profile log-likelihood curves for the effective size and migration rate of subpopulations in ponds 1 and 2 are shown in Figure 6. The maximum-likelihood estimates are \( N_e = 16.3 \) and \( m = 0.39 \) for pond 1 and \( N_e = 18.0 \) and \( m = 0.19 \) for pond 2. If isolation between ponds is assumed, the likelihood \( N_e \) estimates are 17.1 and 21.9 for ponds 1 and 2, respectively, both being somewhat larger than those allowing migration.

The estimated migration rates between ponds are high. Although newts are found to migrate as juveniles and adults during breeding seasons (Hayward et al. 2000), there is no direct estimate of migration rate between populations in the literature to support or reject the estimates obtained herein. However, we argue that the migration rate could be high for this population. The average observed number of distinctive alleles per microsatellite locus is seven, which seems to be too large to be maintained in isolated populations with \( N_e \) at \( \sim10-20 \). Interspecies introgression may also be partially responsible for the high migration rate estimates. On average \( \sim4\% \) of adult newts from these ponds are \( F_1 \)-hybrids between \( T. \) cristatus and \( T. \) marmoratus (Arntzen and Wallis 1991). Although hybrids were identified morphologically for adults and genetically with diagnostic microsatellites for larvae (Jehle et al. 2000) and excluded from the analysis (Jehle et al. 2001), it is inevitable that some hybrids may be undetected and included in the samples. Coincidentally, pond 2 contains few \( T. \) cristatus (Jehle et al. 2001), and the immigration rate for \( T. \) marmoratus in this pond is estimated to be much lower than that in pond 1, which contains \( \sim75 \) \( T. \) cristatus individuals.

**DISCUSSION**

The estimation of demographic properties of a species from its genetic characteristics is now an old topic. Many methods have been generated by statistical geneticists to estimate population sizes, migration rates, population growth rates, and many other interesting and useful ecological parameters. Some of these methods work quite well most of the time, and some are less successful. All share one property: they make simplifying assumptions about the possible range of demographies that can be considered, and the methods proposed here are no exception. It should be our goal, though, to make only the assumptions that are biologically tenable and avoid those we know to be both important and untrue. One such untenable assumption in previous temporal methods is that there is no migration into the local populations under study. The intent of this article is twofold: to assess and correct for the population genetic effects of immigration on estimates of the effective population size and to estimate the rate of migration itself.

Techniques that use data over a time series are relatively rare in statistical genetics, and this is unfortunate, as it ignores information available from the evolutionary process itself. One exception to this rule is that the effective population size has most often been estimated.
by a variety of techniques that use temporal data. Two or more samples separated by one or more generations allow us to measure how much allele frequencies change over time. With a model for how evolution might be occurring, we can then predict something about the evolutionary properties of the species. Previous estimators of $N_e$ from temporal data have made a significant assumption in this translation from measuring the pattern to inferring the process: they have assumed that all evolutionary processes other than genetic drift are not occurring in the populations under study. This may be reasonable for some genetic markers with respect to selection and mutation, but in most cases the effect of migration into a population is not negligible. What we have seen here is that this assumption can give estimates of $N_e$ with substantial bias (Figure 2). Yet it is possible to correct for this error and in the process estimate another of the elusive demographic properties of populations, the immigration rate.

In this article, we have presented techniques based on both moment estimation and maximum likelihood. As a general rule, the maximum-likelihood method was more successful, having higher accuracy and precision, although in most cases the moment approach was not much worse. The likelihood approach, as usual, is much more computationally intensive, although with modern computers this limitation would not be prohibitive. In most cases, the methods require a large amount of marker information for reasonably good estimates of evolution of the species. Previous estimates of $N_e$ from temporal data have made a significant assumption in this translation from measuring the pattern to inferring the process: they have assumed that all evolutionary processes other than genetic drift are not occurring in the populations under study. This may be reasonable for some genetic markers with respect to selection and mutation, but in most cases the effect of migration into a population is not negligible. What we have seen here is that this assumption can give estimates of $N_e$ with substantial bias (Figure 2). Yet it is possible to correct for this error and in the process estimate another of the elusive demographic properties of populations, the immigration rate.

In this article, we have presented techniques based on both moment estimation and maximum likelihood. As a general rule, the maximum-likelihood method was more successful, having higher accuracy and precision, although in most cases the moment approach was not much worse. The likelihood approach, as usual, is much more computationally intensive, although with modern computers this limitation would not be prohibitive. In most cases, the methods require a large amount of marker information for reasonably good estimates of migration rate and effective size. Like the case of a single isolated population (no migration), the amount of information depends on the total number of independent alleles across loci (total number of alleles − number of loci) in the samples, the sample size, and the number of samples (Wang 2001a). Unlike the single isolated population case, however, the amount of information and thus the power of the methods are no longer simply proportional to sampling interval $t$, but depend on $t$ in a complicated way. Other things being equal, the power of the methods first increases with $t$, reaches a maximum at an intermediate $t$ value, and thereafter decreases with $t$ (Figure 1). If $t$ is too large, the methods tend to estimate the $N_e$ for the whole species rather than for the local population (Figures 1 and 2). Because of these and the fact that the optimal $t$ depends partly on the parameters being estimated, we recommend a short (say, two to
four generations) sampling interval in practice. If we have some prior information indicating a low migration rate, however, \( t \) should be larger to increase the estimation precision. In general, the numbers of individuals and markers required are not outside the range of the better studies of population structure. Most of the results given here, which allow at least precision to the right order of magnitude, use a few hundreds of individuals genotyped for 10–20 loci for a total number of 20–140 alleles across loci. It should be noted that most simulations in the present study were run assuming the equilibrium level of initial genetic differentiation among populations. The power of the methods can be substantially higher than that indicated by these simulations if, in practice, the populations are more differentiated when initial samples are taken.

Many methods have been developed for estimating the effective size, migration rate, or their product \( N_e \). Although it seems desirable to compare these different methods for their precision and accuracy, such a comparison is not possible because they use different types of information and estimate different genetic properties of the populations (see, e.g., Schwartz et al. 1999). For the estimation of effective size, methods employing a single sample generally use the (equilibrium) distribution of some genetic property, such as linkage disequilibrium (Hill 1981), identity disequilibrium (Vitalis and Couvet 2001), or coalescence time (e.g., Beaumont 1999; Beerli and Felsenstein 2001) to estimate the long-term effective population size. In contrast, methods employing two or more temporal samples (e.g., Nei and Tajima 1981; Pollak 1983; Waples 1989; Williamson and Slatkin 1999; Anderson et al. 2000; Wang 2001a; this article) use the allele frequency changes over time under a given genetic model to estimate the short-term average effective size over the sampling interval. An exception is the heterozygosity excess method (Pudovkin et al. 1996; Luikart and Cornuet 1999), which uses a single sample but estimates the effective size of the population just one generation before the sampling. This method, however, has a low precision and applies only to very small populations under strictly random mating.

For the estimation of migration, most previous approaches estimate the composite quantity of effective number of migrants, \( 4N_e m \), assuming an equilibrium among mutation, migration, and drift (e.g., Slatkin 1987; Beerli and Felsenstein 2001). This quantity largely determines the genetic differentiation among populations at equilibrium, but tells us little about the specific effect of either genetic drift (effective size) or migration. These techniques also suffer from a long list of untenable assumptions (Whitlock and McCauley 1999). So far the only approach available to estimate effective size and migration rate jointly is based on the estimates of one- and two-locus identity probabilities (Vitalis and Couvet 2001). It requires only a single sample, but assumes populations at equilibrium and with known parameters for the mating system and linkage relationship among markers. For large populations, tightly linked markers with known recombination rate are necessary for substantial identity disequilibrium and therefore reliable \( N_e \) estimates from this method. In practice, parameters describing mating system and recombination are unknown and, in the most favorable case, estimated from data and the estimates can suffer from large sampling errors. In contrast, our temporal method must use at least two temporal samples from the focal population, but it does not make any assumptions about mating system and genetic status (at equilibrium or not) of the populations when sampling, and it applies to small as well as large populations. We do assume independence among marker loci, but violation of the assumption affects only the precision, not the accuracy of the estimation. With tightly linked markers, the information of the data and thus precision might be overestimated, but no bias should result from dependence among loci.

Another advantage of our temporal approach is that it is applicable to a wide range of population structures and migration patterns, e.g., arbitrary numbers of demes of various sizes in an island or stepping-stone model, as long as the migration sources are known a priori from another source of information. Simulation results indicate that the methods developed assuming an infinitely large source are robust to violations of the assumption. For the case of a single small source population, or for sampling only a few from many small source populations, the methods under the infinite-source model still yield satisfactory estimates for both immigration rate and effective size for the focal population.

A computer program for the infinite-source model, MLNE (maximum-likelihood estimate of \( N_e \)), which implements the maximum-likelihood and moment estimation of both \( N_e \) and \( m \), is available for free download from http://www.zoo.cam.ac.uk/ioz/software.htm.

We thank Mark Beaumont, Cort Griswold, Sally Otto, Dolph Schlüter, and two anonymous reviewers for their careful reading of and helpful comments on the manuscript. Funding was provided by a grant from the Natural Science and Engineering Research Council (Canada).

**LITERATURE CITED**


Dobzhansky, Th., and S. Wright, 1941 Genetics of natural populations. V. Relations between mutation rate and accumulation of lethals in populations of D. pseudoobscura. Genetics 26: 25–51.


Wang, J., and A. Caballero, 1999 Developments in predicting the effective size of subdivided populations. Heredity 82: 212–226.


Whitlock, M. C., and D. E. McCauley, 1999 Indirect measures of gene flow and migration: \( F_{st} \) not equal to 1/(4Nm+1). Heredity 82: 117–125.


Communicating editor: J. B. Walsh

**APPENDIX: SIMULATION AND LIKELIHOOD SEARCH METHODS**

**Simulations:** Simulations were run to check the performance of the moment and likelihood methods over a wide range of the parameter values. The changes in allele frequency in a diploid population of effective size \( N_e \) and immigration rate \( m \) were simulated by Monte Carlo. For each replicate, the initial allele frequencies for each locus were drawn from a uniform distribution. Except where specified, some generations (required to reach \( >95\% \) equilibrium \( F_{st} \) predicted by using the true parameter values) were run to allow populations to reach quasi-equilibrium between drift and migration before samples were taken binomially (sampled with replacement) for estimating the parameters. For equilibrium populations, obviously the number of alleles at a locus observed (maintained) in a focal population covaries with \( N_e \) and \( m \). To compare the performance of the estimation methods across wide ranges of parameter values of \( N_e \) and \( m \), we used a small number of alleles per locus in most simulations. For populations with very small \( m \) and \( N_e \), mutations (at rate 0.001) were incurred to maintain polymorphism. The methods are, however, applicable to nonequilibrium populations and should be slightly more powerful if the genetic differentiation among populations is larger than the equilibrium value, and vice versa.

Since both the likelihood and the moment methods estimate \( 1/(2N_e) \) rather than \( N_e \) directly, and it is \( 1/(2N_e) \) that is important in most applications, we report...
the estimates of \(1/(2N_e)\) and \(m\) from the simulations. Any unbiased estimate of \(1/(2N_e)\) with sampling error would be biased for estimating \(N_e\), and vice versa. To reduce running time, a replicate was terminated once the likelihood estimate of \(N_e\) was determined to be larger than the threshold value of \(10N_e\), and the estimate of \(1/(2N_e)\) is regarded as zero (\(N_e\) infinite) from this replicate. For the moment estimator, a negative estimate of \(N_e\) is taken as \(N_e \to \infty\) and the estimate of \(1/(2N_e)\) is zero. For a given set of parameters, the means and standard deviations of \(1/(2N_e)\) and \(m\) estimates from each estimator were obtained over a large number of replicates (depending on \(N_e\) and thus the computational demand). As a measure of the dispersion of the estimates, the 2.5 and 97.5 percentiles of the distribution of estimates of each parameter from all replicates were also listed for each estimator.

For more than two samples from the focal population, a single estimate of the (harmonic) mean effective size and (arithmetic) mean immigration rate can be obtained from the likelihood estimator. From the moment estimator, \(N_e\) and \(m\) for each period between two adjacent samples can be calculated, and the harmonic and arithmetic means of these over different periods are used as the estimates. More appropriately, averaging moment estimates should take sampling variances into account, which can be obtained by resampling.

**Likelihood searches:** Depending on the data, the likelihood function defined above may have multiple maxima. To search for the global maximum, we could use techniques employing a Monte Carlo method, such as the simulated annealing method (Press et al. 1996; Wang 2001b). However, such methods are computationally intensive and therefore not appropriate for simulation studies where large numbers of replicates are to be run. We choose to use Powell’s quadratically convergent method (Press et al. 1996), which can find quickly a local maximum from a given starting point. Widely varying starting values of \(N_e\) and \(m\), including the estimates from the moment estimator, are used and the largest likelihood is picked as the global maximum. For large \(N_e\), we start the maximum-likelihood searching from the initial point of moment estimates if they are valid or a random nearby point if not. The performance of the maximum-likelihood estimator indicated by the simulations for large \(N_e\) can therefore be slightly conservative.