THEORETICAL POPULATION BIOLOGY **49**, 369–386 (1996)
ARTICLE NO. 0018

Distinguishing Migration from Isolation Using the Variance of Pairwise Differences

JOHN WAKELEY*

Department of Population Genetics, National Institute of Genetics, Mishima, Japan

Received May 19, 1995

Two demographic scenarios are considered: two populations with migration and two populations that have been completely isolated from each other for some period of time. The variance of the number of differences between pairs of sequences in a single sample is studied and forms the basis of a test of the isolation model. The migration model is one possible alternative to isolation. The isolation model is rejected when the proposed test statistic, which involves the variances of pairwise difference within and between populations, is larger than some critical value. The power and realized significance of the test are investigated using simulations, and an example using mitochondrial DNA illustrates its application.

1. INTRODUCTION

1.1. *A DNA Data Set*

In a recent study of mitochondrial DNA in the threespine stickleback, Ortí *et al.* (1994) sequenced 747 base pairs of the cytochrome *b* gene in 36 individuals from samples covering much of the range of this widespread species. This revealed the existence of two major clades, one comprising mainly western Pacific samples (Japan) and the other comprising mainly northeastern Pacific samples (Alaska and British Columbia). Nested within the northeastern Pacific clade was a group containing samples from the northern Atlantic and one haplotype from Los Angeles, California. The western and northeastern Pacific clades were separated by a minimum of 18 substitutions, yet one haplotype from British Columbia was identical to a common Japanese haplotype and two Alaskan samples clustered with the western Pacific clade.

From this pattern, Ortí *et al.* (1994) inferred two major demographic events: a recent origin of Atlantic stickleback populations from a lineage resembling the Los Angeles sample, and an ancient isolation and divergence between western and northeastern Pacific sticklebacks followed

* E-mail: jwakeley@rci.rutgers.edu. Current address: Rutgers University, Department of Biological Sciences, P.O. Box 1059, Busch Campus, Piscataway, NJ 08855–1059.

0040-5809/96 $18.00

by secondary contact. The present analysis focuses only on the latter comparison; the sample from Los Angeles and those from the Atlantic are omitted. Specifically, the hypothesis that the patterns of divergence within the Pacific are due solely to the isolation and subsequent divergence of western and eastern populations is tested. Under this hypothesis, haplotypes found in both areas represent retained ancestral lineages. Ortí *et al.* (1994) did not consider this hypothesis, but the result of the test is interesting. It illustrates the application of the method developed here and contributes something to our understanding of stickleback historical biogeography in that one hypothesis, at least, can be rejected with statistical confidence.

1.2. *Quantifying Divergence within and between Populations*

There are two commonly-used methods of summarizing the extent of differences among sequences in a sample: counting the number of segregating, or polymorphic, sites, and calculating the average number of pairwise differences. Because, for the models considered here, the relationship between the number of segregating sites and demographic or genetic parameters is not well defined, e.g., Tajima (1989), the present work focuses on pairwise differences.

If a number of sequences, n, are sampled from a population and $k_{ii'}$ is the number of differences observed between sequence i and sequence i', then

$$k = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{i'=i+1}^{n} k_{ii'} \tag{1}$$

is the average number of pairwise differences. The statistical properties k are well known. For a diploid, randomly-mating, population of constant effective size N, Tajima (1983) showed that the expectation of k is $4Nu$, or θ, and the variance is

$$\mathrm{Var}(k) = \frac{n+1}{3(n-1)}\theta + \frac{2(n^2+n+3)}{9n(n-1)}\theta^2, \tag{2}$$

where u is the mutation rate in a neutral, infinite sites model with no recombination (Kimura, 1969; Watterson, 1975).

The present work concerns samples from a sub-divided population. When a number of sequences, n_X and n_Y, are sampled from two sub-populations called X and Y, the divergence among sequences within each population are calculated using (1), separately for X and Y. These quantities have been referred to as d_X and d_Y (Nei and Li, 1979; Takahata and Nei, 1985), and this notation is followed here. Accordingly, from this sample we

can also compute the average number of pairwise differences between sequences from different populations,

$$d_{XY} = \frac{1}{n_X n_Y} \sum_{i=1}^{n_X} \sum_{j=1}^{n_Y} k_{ij}, \tag{3}$$

where k_{ij} is the number of differences observed between sequence i from population X and sequence j from population Y. In addition,

$$d = d_{XY} - \frac{d_X + d_Y}{2} \tag{4}$$

can be used as a measure of the excess number of nucleotide differences between the two populations, e.g., Nei and Li (1979).

1.3. *Two Contrasting Explanations*

We can calculate d_X, d_Y, and d_{XY} (and thus d) from any sample of sequences from two populations. For the data described above, $n_X = 14$, $n_Y = 9$, $d_X = 8.1$, $d_Y = 1.9$, and $d_{XY} = 16.1$, so $d = 11.1$, where populations X and Y are the eastern and western Pacific populations, respectively. In turn, by assuming some population genetic model for the two populations, we can use the observed values of d_X, d_Y, and d_{XY} to estimate its parameters. However, since many different models can give identical predictions for d_X, d_Y, and d_{XY}, we cannot use these to distinguish between alternative explanations for observed levels of polymorphism within and between populations.

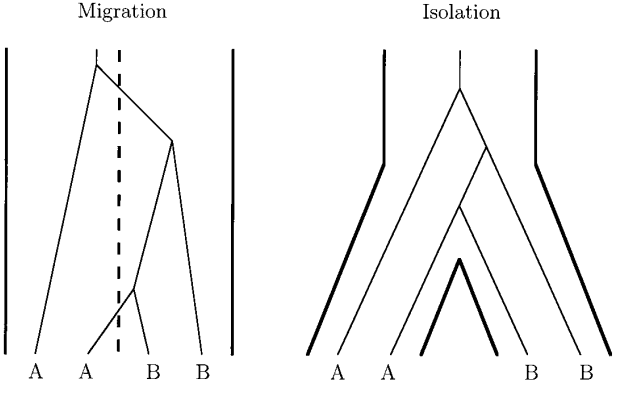

FIG. 1. Depiction of the migration and isolation models. Thick lines represent population boundaries, with dashes indicating that gene flow is possible. Thin lines show one possible genealogical history of a sample of two sequences from each population.

Figure 1 depicts two such models. The first, labelled "Migration" in Fig. 1, is a two-population version of the finite-island model (Kimura and Weiss, 1964). The populations are assumed to have been exchanging migrants at a constant rate for a long enough time that the expectations of d_X, d_Y, and d_{XY} are close to their equilibrium values. Takahata (1983) has studied the rate of approach to equilibrium for the finite-island model. Migration rates are assumed to be symmetric and equal to m per haploid individual per generation. Under the second model, labelled "Isolation" in Fig. 1, the present populations are assumed to be descended from a single population that split into two at some time, t generations, in the past. In both models, there is random mating within each population, generations are nonoverlapping, and the effective sizes of each population have remained constant through time.

Below, migration and isolation are juxtaposed as, possibly, equally good explanations of patterns observed in sequence data. Following a demonstration that neither average pairwise differences nor genealogical structure will suffice to distinguish migration from isolation, it is shown that the variances of differences among pairs in a single sample can be used for this purpose. From that point on, the isolation model is considered the null hypothesis, and a test of the model is developed. Equilibrium migration is just one of many possible alternatives, but it is the one of interest here. The test of isolation is designed with this alternative in mind.

1.4. *The Expectations of Pairwise Differences*

By making some further assumptions, it is straightforward to show that these two models are indistinguishable using d_X, d_Y, and d_{XY}. The first assumption is that all variation is neutral. The second is that the sequences under consideration are long enough that every new mutation occurs at a previously unmutated site. The rate of mutation is u, per sequence per generation. The third assumption is that there is no recombination within the sequences. The fourth is that, for the isolation model, each of the three populations, the ancestor and the two descendents, is of diploid (haploid) effective size N ($2N$), whereas, for the migration model, each of the two populations is of diploid (haploid) effective size $N/2$ (N). The final assumption is that the sequences are sampled randomly.

With these assumptions, it has been known for some time that $E(d_X) = E(d_Y) = \theta$ and $E(d_{XY}) = \theta + \theta/(2M)$ in the migration case, where $\theta = 4Nu$ and $M = 2Nm$ (Nei and Feldman, 1972; Li, 1976; Griffiths, 1981; Slatkin, 1987; Strobeck, 1987; Notohara, 1990; Hey, 1991). Thus, $E(d) = \theta/(2M)$ for this model. Under isolation, $E(d_X) = E(d_Y) = \theta$ and $E(d_{XY}) = \theta + \theta T$, where $T = t/(2N)$ (Kimura, 1969; Watterson, 1975; Li, 1977; Gillespie and Langley, 1979; Nei and Tajima, 1981; Takahata and Nei, 1985), so that $E(d) = \theta T$. The assumptions made above provide a basis for comparing

migration and isolation which is used below. That is, when $T = 1/(2M)$ these two models give identical values for $E(d_X)$, $E(d_Y)$, and $E(d_{XY})$. However, this also means that when we sample some sequences from two populations, we can fit either model to the observed values of d_X, d_Y, and d_{XY}, and estimate the parameters θ and M or T, but we cannot use d_X, d_Y, and d_{XY} to distinguish migration from isolation.

1.5. *The Probabilities of Genealogies*

The genealogy of a sample contains information about demographic history (Slatkin, 1989). However, recent work indicates that topological relationship, by themselves, will also not serve to distinguish migration from isolation (Slatkin and Maddison, 1989; Takahata and Slatkin, 1990). Figures 2 and 3 illustrate why this is true. Figure 2 shows the four possible topologies for a sample of two sequences from each of two populations. After graphs shown by Tajima (1983) for the isolation model and Takahata and Slatkin (1990) for the migration model, Fig. 3 plots the probabilities of the genealogies shown in Fig. 2 over a broad range of values of M or $1/(2T)$. That is, along the horizontal axes in Fig. 3, the expectations of d_X, d_Y, and d_{XY} are the same for both models.

The similarlity between these two sets of curves explains why Slatkin and Maddison (1989) and Takahata and Slatkin (1990) did not find a way to distinguish migration from isolation using only genealogical structure. The relationships among the four probabilities are quite similar over the entire range of M or $1/(2T)$. As the migration rate decreases, or, equivalently, as the time of separation increases, the probability of intrapopulation monophyly approaches one. At the other extreme, all four probabilities converge on the values expected in a single population.

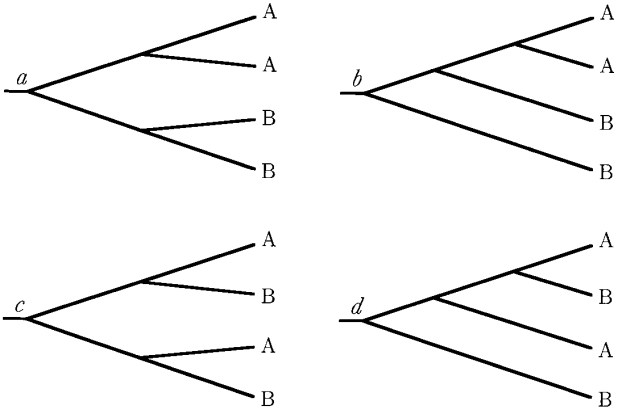

FIG. 2. The four possible topologies for a sample of two sequences from each of the two populations.

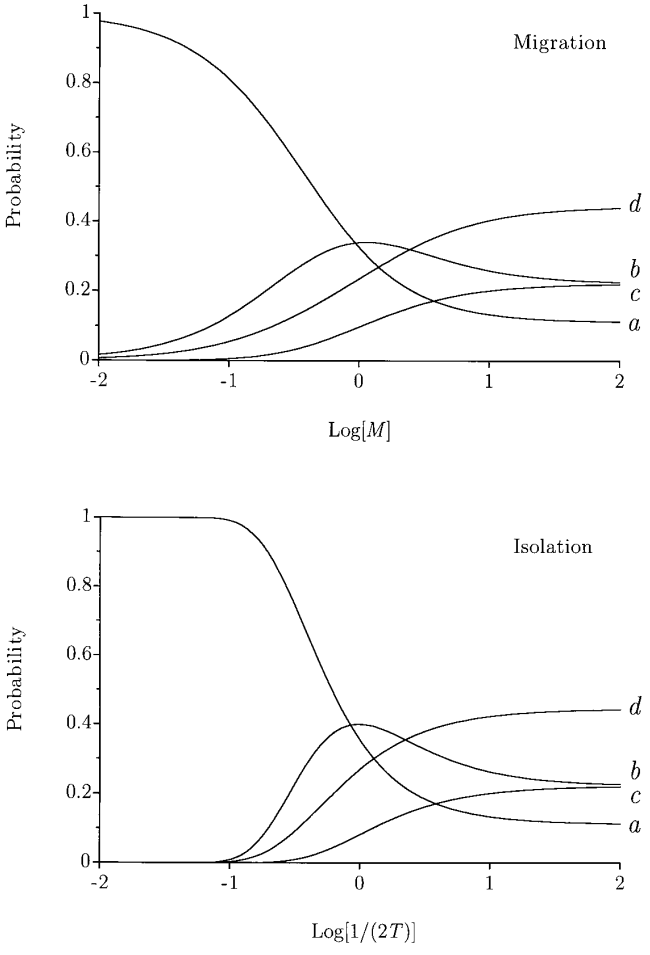

FIG. 3. The probabilities of each of the four topologies shown in Fig. 2 under migration and isolation. Along the horizontal axes, the expectations of d_X, d_Y, and d_{XY} are identical for both models. The logarithms are base 10.

1.6. *The Variances of Pairwise Differences*

Wakeley (1996) showed that the variances of d_X, d_Y, and d_{XY} are greater under migration than under isolation, indicating that a test to distinguish these two situations could be developed. However, the variances derived by Wakeley (1996) under migration and by Takahata and Nei (1985) under isolation are not immediately suitable for such a test. They are the variances of d_X, d_Y, and d_{XY} expected among identical-size samples taken from

several independent populations that all have the same evolutionary parameters, i.e., identical values of θ and M or T.

Except when the effective population size is very small, these are also the variances expected among unlinked loci in a sample from one population, since the histories of unlinked loci are essentially independent (Griffiths, 1981; Hudson, 1983). However, even if data from several loci were available so we could calculate such variances, these previously published formulas would apply only if the loci all had the same underlying values of θ and M or T. This is true also of the variance for a single, randomly mating, population given by (2). Thus, the method introduced here is based on the variance calculable from a single sample of sequences at a single locus.

2. Methods

2.1. *The Variance among Pairs in A Single Sample*

There are $n(n-1)/2$ nonidentical pairwise comparisons possible among sequences in a sample of size n. From these, we can calculate the average number of pairwise differences, as in (1), and we can calculate a variance

$$s^2 = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{i'=i+1}^{n} (k_{ii'} - k)^2, \tag{5}$$

where k is defined by (1). This quantity is similar to the variance proposed by Sved (1968), which was also considered by Brown *et al.* (1980) and by Hudson (1987). The difference is that those authors sum over all n^2 possible pairs.

A somewhat simpler form of (5), from which the expectation of s^2 can be computed easily, is

$$s^2 = \left(\frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{i'=i+1}^{n} k_{ii'}^2 \right) - k^2. \tag{6}$$

For a single, randomly mating population of constant effective size and in which all variation is neutral,

$$E(s^2) = E(k_{ii'}^2) - E(k^2) \tag{7}$$

$$= \mathrm{Var}(k_{ii'}) - \mathrm{Var}(k). \tag{8}$$

Equation (7) is true because $E(k_{ii'}^2)$ is identical for every pair of sequences. Since $E(k_{ii'}) = E(k)$, (8) follows by the addition of $E(k)^2$ to and the subtraction of $E(k_{ii'})^2$ from (7). The second term in (8) is Tajima's (1983)

variance, presented above as (2), and the first term is $\theta + \theta^2$ from the same formula with $n = 2$. Thus,

$$E(s^2) = \frac{2(n-2)}{3(n-1)}\theta + \frac{7n^2 - 11n - 6}{9n(n-1)}\theta^2 \tag{9}$$

is the expectation of (5) for a constant-sized, neutral population.

Unlike $\mathrm{Var}(k)$, when $n = 2$, $E(s^2)$ equals zero since we cannot calculate a variance from only one pair of sequences. The statistic, s^2, quantifies the variation in coalescence times in a given sample. It makes use of additional aspects of a sample's genealogy than the mean value, k, thus increasing the amount of information we can glean from a sample of sequences.

The method is used to produce (9) can be applied to the migration and isolation models as well. For populations X and Y, (5) through (8) apply directly, the only differences being in notation (d_X or d_Y instead of k) and that the variances derived by Wakeley (1996) under migration and by Takahata and Nei (1985) under isolation have to be used in the analogue of (8). These intrapopulation measures are referred to as s_X^2 and s_Y^2. For interpopulation differences, define

$$s_{XY}^2 = \frac{1}{n_X n_Y} \sum_{i=1}^{n_X} \sum_{j=1}^{n_Y} (k_{ij} - d_{XY})^2, \tag{10}$$

which, by the same logic, has expectation $\mathrm{Var}(k_{ij}) - \mathrm{Var}(d_{XY})$.

Compared under migration and isolation, the expectations of s_X^2, s_Y^2, and s_{XY}^2 show a similar pattern to that reported in Wakeley (1996) for $\mathrm{Var}(d_X)$, $\mathrm{Var}(d_Y)$, and $\mathrm{Var}(d_{XY})$. For a given value $M = 1/(2T)$, all three statistics are generally greater under migration than under isolation, the differences becoming more pronounced as the migration rate decreases or, equivalently, as the time of separation increases. As M increases (T decreases), all three converge on the values expected in a single, randomly mating population. Since s_X^2, s_Y^2, and s_{XY}^2 measure the variance in times to common ancestry among the sequences in a sample, this is equivalent to saying that there is greater variance in coalescence times under migration than under isolation.

Thus, s_X^2, s_Y^2, and s_{XY}^2 can be used to distinguish migration from isolation. Here, in coming up with a test, the isolation model is chosen as the null hypothesis. Because s_X^2, s_Y^2, and s_{XY}^2 are smaller under isolation than under migration, it seems likely that their variances, or the variance of any test statistic involving them, will also be smaller. A test of isolation will likely have greater power to distinguish migration from isolation than one where the null model is migration.

2.2. *A Test of the Isolation Model*

A simple test would be as follows: for given values of n_X, n_Y, d_X, d_Y, and d_{XY}, reject isolation if s_X^2, s_Y^2, and/or s_{XY}^2 are too large. However, the strength of correlation between d_X and d_Y, which estimate θ in both models, and s_X^2, s_Y^2, and s_{XY}^2 makes this test ineffective. For example, in the simulations below, under migration with $\theta = 10.0$, $M = 0.5$, and $n_X = n_Y = 10$, the correlation between d_X and s_X^2 is 0.86 and the correlation between d_X and s_{XY}^2 is 0.64. Similarly high values are found under isolation; large values of the variances tend to be associated with large values of the means. Using the above test, the null hypothesis of isolation is almost never rejected, even when it is false.

The derivation of the expected covariances between d_X, d_Y, or d_{XY} and any proposed test statistics involving s_X^2, s_Y^2, or s_{XY}^2 is impractical, given the complexity of computing just $\mathrm{Var}(d_X)$, $\mathrm{Var}(d_Y)$, and $\mathrm{Var}(d_{XY})$ (Takahata and Nei, 1985; Wakeley, 1996). Thus, a large part of the present work involved screening great numbers of possible test statistics for good performance in simulations. Three classes of statistics were investigated: ones related to correlation coefficients, ones involving ratios of observed variances, and ones involving ratios of observed and expected variances. Statistics related to the correlation coefficients of intra- and interpopulation differences showed the best performance over a broad range of parameter values for both migration and isolation. This regrettably *ad hoc* procedure yielded some useful statistics, and the results for overall best one are presented below.

The intra population coefficients of correlation, s_X/d_X and s_Y/d_Y, distinguished migration from isolation better than the interpopulation coefficient, s_{XY}/d_{XY}, when the migration rate was high or the time of separation was short, but the opposite was true when the migration rate was low or the time of separation was long. The weighted average of these quantities, where each was weighted by the number of pairwise comparisons involved, performed better than any of these quantities alone. A less intuitively obvious statistic, s_{XY}/k, where k is the average number of pairwise differences among all sequences, always performed better than s_{XY}/d_{XY}.

The overall best statistic was

$$\frac{1}{n(n-1)}\left[n_X(n_X-1)\frac{s_X}{d_X} + n_Y(n_Y-1)\frac{s_Y}{d_Y} + 2n_X n_Y \frac{s_{XY}}{k}\right], \tag{11}$$

which is called Ψ below. It is worth noting that several other statistics performed only marginally worse than Ψ. In (11), n is the total sample size, $n_X + n_Y$, and s_X, s_Y, and s_{XY} are the square roots of s_X^2, s_Y^2, and s_{XY}^2, respectively. Figure 4 plots the approximate expectation of Ψ under

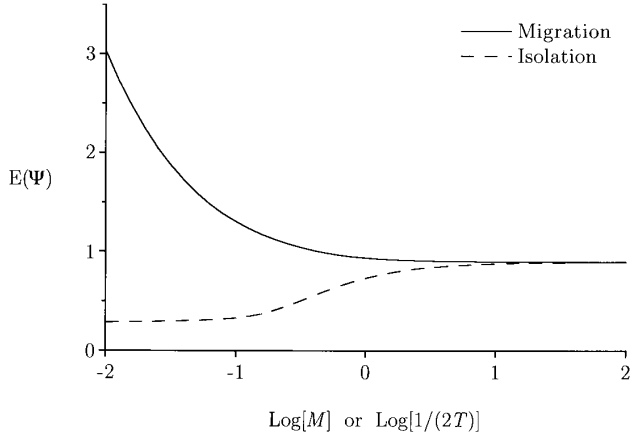

FIG. 4. The approximate expectations of the test statistic, Ψ, compared under migration and isolation. These curves are approximate since they were obtained by simply replacing d_X, d_Y, d_{XY}, s_X^2, s_Y^2, and s_{XY}^2 in (11) with their expected values.

migration and isolation over the same range of parameter values used in Fig. 3. Figure 4 shows that the values under migration and isolation become increasingly similar as M or $1/(2T)$ increases but diverge at the other extreme, giving some indication of when the test might perform well.

2.3. *Simulations*: *Power and Realized Significance*

Simulations were done to assess the utility of the proposed test statistic, Ψ, over a broad range of values of all relevant parameters. Three programs produced all of the results presented below: one to simulate migration, one to simulate isolation, and one to perform the test.

The isolation program was essentially the same as the routine "make_tree" given in Hudson (1990), but with three populations (ancestral plus two descendant) rather than one. The usual coalescent process proceeded independently in each of the two descendant populations until time T in the past, measured in units of $2N$ generations, when the remaining ancestral sequences were joined into a single population. The migration program followed the general approach outlined by Hudson (1983) for simulating the genealogy of a sample when one of several possible kinds of events might happen in any given generation. Here, those events were common-ancestor events within each population and migration events. Both programs employed the infinite-sites mutation model with no recombination.

The program that performed the test took observed values n_X, n_Y, d_X, d_Y, d_{XY}, s_X^2, s_Y^2, and s_{XY}^2 as input. It then estimated θ as $(d_X+d_Y)/2$ and T as $2d/(d_X+d_Y)$ and simulated 1000 replicate data sets under isolation

with those parameter values to determine the approximate 5% cutoff value for Ψ. The cutoff was then compared to the observed value of Ψ, and the null hypothesis rejected or not. If d was negative, T was set to zero. Since Ψ is undefined when either d_X or d_Y equals zero, all results are conditional upon observing at least some variation in both populations.

The simulations comprised two steps. For a given set of parameter values under migration or isolation, 1000 replicate data sets were generated and d_X, d_Y, d_{XY}, s_X^2, s_Y^2, and s_{XY}^2 calculated from each. The test was then applied to each of these 1000 data sets. The fraction of data sets for which isolation can be rejected estimates the probability of rejection for each set of starting parameters and for each underlying model. If the underlying model is migration, this probability represents our power to distinguish migration from isolation. If the underlying model is isolation, it represents what can be called the realized significance of the test. Clearly, we want the power to be as great as possible and the realized significance to be close to whatever significance level is used, in this case 5%.

3. Results

3.1. *Performance in Simulations*

Simulation results are presented in Figs. 5 through 8. These show the power and realized significance of the test when the migration rate or time of separation, the length of the sequences, the numbers of sequences sampled

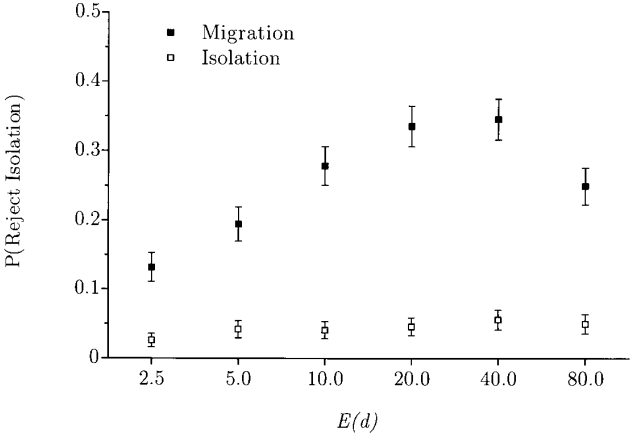

FIG. 5. The power and realized significance of the test for varying levels of differentiation between the two populations. For all points, $E(d_X) = E(d_Y) = 10.0$ and $n_X = n_Y = 10$. For the migration model, the points left to right left correspond to M equals 2, 1, $\frac{1}{2}$, $\frac{1}{4}$, $\frac{1}{8}$ and $\frac{1}{16}$. For the isolation model, T equals $1/(2M)$.

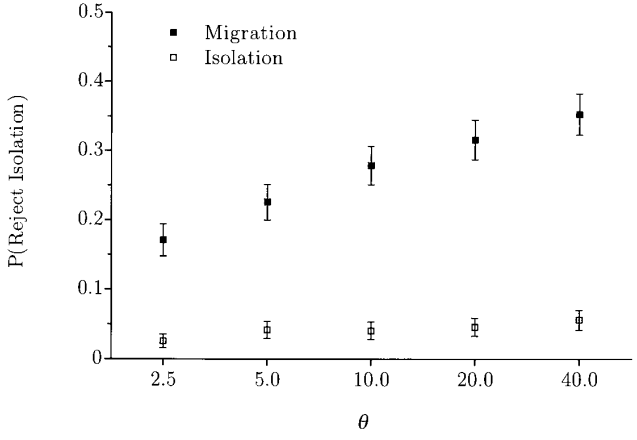

FIG. 6. The power and realized significance of the test for different values of the parameter θ or $4Nu$. For all points, $E(d) = \theta$ ($M = 0.5$ or $T = 1.0$), and $n_X = n_Y = 10$.

from each population, and the number of independent loci are varied. All four figures have the point where $E(d_X) = E(d_Y) = 10.0$, $E(d) = 10.0$, and $n_X = n_Y = 10$ in common. Error bars depict 95% confidence intervals for the probability of rejecting isolation. For multiple loci, the sum of Ψ-values was used as the test statistic. The realized significance is close to 5% in every case and is never much above this value, but the power to reject isolation when migration is the true model varies substantially with all parameters.

Figure 5 shows that the power to reject isolation when the underlying model is migration depends on the migration rate, rising from about

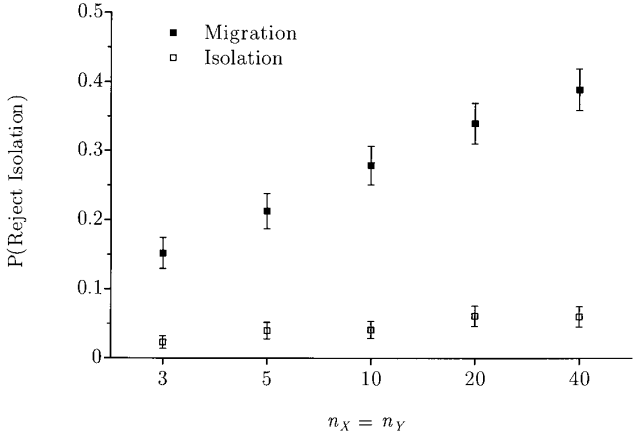

FIG. 7. The power and realized significance of the test for different numbers of sequences sampled. For all points, $E(d_X) = E(d_Y) = E(d) = 10.0$.

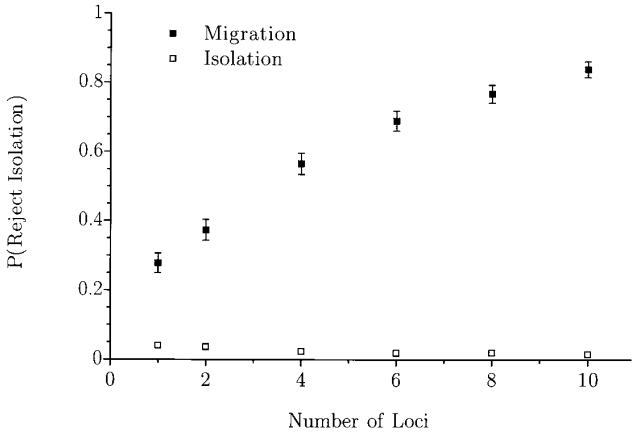

FIG. 8. The power and realized significance of the test for different numbers of independent loci examined. For all points and each locus, $E(d_X) = E(d_Y) = E(d) = 10.0$, and $n_X = n_Y = 10$.

13% when M equals 2 to about 35% when M equals $\frac{1}{8}$. However, as the migration rate decreases further, this trend reverses. Figure 6 shows that the larger θ is, which can be interpreted as increasing the sequence length, the greater is the power to reject isolation. Figure 7 demonstrates a significant increase in power with the number of sequences sampled from each population. In Figs. 5, 6, and 7, the power to distinguish migration from isolation never exceeds 40%. Figure 8 shows the effect of varying the number of independent loci from 1 to 10.

3.2. *Application to the DNA Data*

The DNA data of Ortí *et al.* (1994) give 5.0 as an estimate of θ for the sequences examined and 2.2 as an estimate of T, the time of separation of the western and the eastern Pacific populations, measured in units of $2N$ generations. Also, for these data, $s_X^2 = 82.4$, $s_Y^2 = 1.2$, and $s_{XY}^2 = 57.4$, giving a value if 0.82 for Ψ. Recall that population X is the eastern Pacific and Y is the western Pacific, $n_X = 14$, $n_Y = 9$, $d_X = 8.1$, $d_Y = 1.9$, and $d_{XY} = 16.1$. Ten thousand replicate data sets, simulated under isolation using the parameter values estimated from the data, indicate that the chance of observing a value of Ψ greater than or equal to 0.82 is about 0.013. The hypothesis that the isolation model can explain the similarities and differences between these two populations is rejected.

4. DISCUSSION

For the same values of the average coalescence times, quantified by d_X, d_Y, and d_{XY}, under migration it is possible that both very recent and very

ancient common ancestors of the sequences in the sample exist. Under isolation, however, the range of possible coalescence times is restricted. This is depicted in Fig. 1 and explains why statistics such as Ψ, which quantify the variation in times to common ancestry, are generally larger under migration than under isolation and can be used to distinguish these two demographic situations.

According to Fig. 4, the greatest difference between migration and isolation should occur when the migration rate is very small or, equivalently, when the time of separation is very large. At this extreme, however, Fig. 3 shows that the probability of intrapopulation monophyly (tree *a* in Fig. 2) approaches one. When only a single migration event occurs, migration would appear indistinguishable from isolation, so the increase in and subsequent erosion of power, pictured in Fig. 5, occurs in spite of the great differences in expectation for Ψ under migration and isolation.

Figure 3 shows that the probability of tree *a* approaches one more quickly under isolation as T gets large under migration as M gets small. When $\log[1/(2T)]$ and $\log[M]$ equal -1, which is equivalent to T being 5 and M being 0.1, the chance of observing any of the trees *b*, *c*, or *d* is less than 1% under isolation, but 19% under migration. In this range, a test based only on the topology of the sample would have power comparable to that of the present test. However, for the intermediate values of M and T used in the simulations, Fig. 3 shows that topology alone will likely not serve to distinguish migration from isolation.

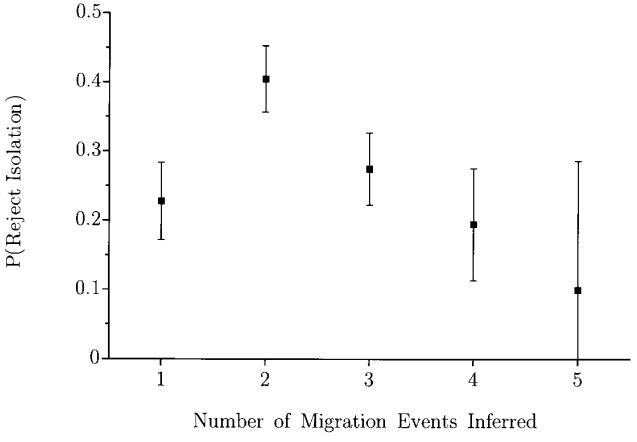

FIG. 9. The power of the test as a function of the minimum number of migration events that can be inferred from the true tree relating the sequences. $E(d_X) = E(d_Y) = E(d) = 10.0$ and $n_X = n_Y = 10$. A total of 1000 replicates were done and the numbers corresponding to each point, from right to left, were 216, 400, 278, 93, and 10. Six migration events were inferred for the remaining three replicates; isolation was rejected for one of these three (not shown).

As expected from the definitions of s_X^2, s_Y^2, and s_{XY}^2, the present test makes use of both the topology and the branch lengths of the sample genealogy. Figure 9 shows the relationship between the power of the present test and one measure of topological structure, the minimum number of migration events that can be inferred from a genealogy (Slatkin and Maddison, 1989; 1990). In producing Fig. 9, the true genealogy of each sample was used rather than one estimated from the data, $E(d_X) = E(d_Y) = E(d) = 10.0$, and $n_X = n_Y = 10$. With these parameter values, there is a clear peak at two inferred migration events, i.e., trees similar to the one pictured under "Migration" in Fig. 1. However, there is considerable power even for trees showing intra population monophyly. It seems likely that, in this case, isolation can only be rejected when there are multiple, undetected migration events which inflate s_X^2 and/or s_Y^2.

Increasing the numbers of sequences from each population and surveying multiple loci both increase the chance of there being more than one migration event in the history of the sample. While the effect of sampling more loci appears stronger, the effect of increasing n_X and n_Y is dramatic. Figure 7 shows a nearly linear increase in power with the logarithm of sample size. This situation is quite unlike the problem of estimating the number of heterozygous nucleotide sites, where a sample of 5 to 10 sequences is sufficient (Tajima, 1983). The more sequences we sample, the more migration events we can expect to detect and the greater chance we have of observing genealogies that are inconsistent with the isolation model.

Here, migration was the only alternative model considered. In fact, the test using Ψ is simply a test of whether the variances of pairwise differences are too large relative to the average values for the isolation model to be true. Other alternatives than equilibrium migration could lead to significantly large values of Ψ. These include, but are not limited to, balancing selection at or near the locus under consideration and subdivision within either the descendent or ancestral populations. Without outside information, such factors cannot be discounted as causes of significant results. Deviations from the model that this test should be robust to include recombination, multiple mutation events at single sites, and any other factors that tend to decrease the variances of pairwise differences more than they affect the means. Of course, such deviations will also decrease the already low power of the test.

For the threespine stickleback, Ortí *et al.* (1994) infer an isolation event followed by secondary contact between western and eastern Pacific populations. Like migration, this historical scenario would also tend to inflate s_X^2, s_Y^2, and s_{XY}^2 relative to the simple isolation model. Stickleback demographic history appears to have been influenced greatly by glaciations and the existence of ice-free refugia during the Peistocene (Hocutt and Wiley, 1986). In addition, individuals have been observed in the open northern

Pacific Ocean, suggesting that long-distance migration is possible (Quinn and Light, 1988). To complicate matters further, sticklebacks display three kinds of life histories—fully marine, resident freshwater, and anadromous—and the western Pacific haplotypes that Ortí *et al.* (1994) discovered in Alaska and British Columbia were found only in freshwater habitats.

It is valid, then, to question the applicability of the migration model considered here to the threespine stickleback. When the present test is applied, with migration as the null hypothesis, the chance of observing a value of Ψ greater than or equal to 0.82 is estimated to be 0.32. The data appear consistent with the migration model. However, the test statistic, Ψ, was chosen to maximize performance under the null hypothesis of isolation with migration as an alternative. There is no reason to suppose that Ψ will be the best statistic for other purposes.

Both equilibrium and nonequilibrium processes seem likely to have played a part in stickleback history. Given the low power of the present test to simply distinguish migration from isolation, disentangling more complicated demographic models may be quite difficult. Figure 8 implies that the most important factor affecting our ability to distinguish between alternative models is the number of loci studied. However, the data that produced Fig. 8 are the result of simulations done under unrealistic assumptions, namely, that the loci are independent and that no intralocus recombination occurs. Most genetic loci are found in the nucleus of cells and nuclear DNA undergoes considerable recombination *e.g.*, Schaeffer and Miller (1993). Figure 8 implies that a full accounting of recombination in the development of other tests, similar to this one, would certainly be worthwhile.

A program, written in the C programming language, that performs the test introduced here is available from the author upon request. Please send an electronic-mail address if you have one.

Acknowledgments

I am greatly indebted to Fumio Tajima and Montgomery Slatkin for helpful discussions of this work and for comments on the manuscript. Funding was provided by Post-doctoral Fellowship P94168 from the Japan Society for the Promotion of Science. Partial support was provided by NIH Grant GM40282 to M. Slatkin.

References

Brown, A. H. D., Feldman, M. W., and Nevo, E. 1980. Multilocus structure of natural populations of *Hordeum spontaneum*, *Genetics* **96**, 523–536.

Gillespie, J. H., and Langley, C. H. 1979. Are evolutionary rates really variable? *J. Mol. Evol.* **13**, 27–34.

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

HEY, J. 1991. A multi-dimensional coalescent process applied to multi-allelic selection models and migration models, *Theor. Popul. Biol.* **39**, 30–48.

HOCUTT, C. H., AND WILEY, E. O. 1986. "The Zoogeography of North American Freshwater Fishes," Wiley, New York.

HUDSON, R. R. 1983. Properties of a neutral allele model with intragenic recombination, *Theor. Popul. Biol.* **23**, 183–201.

HUDSON, R. R. 1987. Estimating the recombination parameter of a finite population model without selection, *Genet. Res. Cambridge* **50**, 245–250.

HUDSON, R. R. 1990. Gene genealogies and the coalescent process, *in* "Oxford Surveys in Evolutionary Biology" (D. J. Futuyma and J. Antonovics, Eds.), Vol. 7, Oxford Univ. Press, Oxford, UK.

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

KIMURA, M., AND WEISS, G. H. 1964. The stepping stone model of population structure and the decrease of genetic correlation with distance, *Genetics* **49**, 561–576.

LI, W.-H. 1976. Distribution of nucleotide difference between two randomly chosen cistrons in a subdivided population: The finite island model, *Theor. Popul. Biol.* **10**, 303-308.

LI, W.-H. 1977. Distribution of nucleotide difference between two randomly chosen cistrons in a finite population, *Genetics* **85**, 331–337.

NEI, M., AND FELDMAN, M. W. 1972. Identity of genes by descent within and between populations under mutation and migration pressure, *Theor. Popul. Biol.* **3**, 460–465.

NEI, M., AND LI, W.-H. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases, *Proc. Natl. Acad. Sci.* **76**, 5269–5273.

NEI, M., AND TAJIMA, F. 1981. DNA polymorphism detectable by restriction endonucleases, *Genetics* **97**, 145–163.

NOTOHARA, M. 1990. The coalescent and the genealogical process in geographically structured population, *J. Math. Biol.* **29**, 59–75.

ORTI, G., BELL, M. A., REIMCHEN, T. E., AND MEYER, A. 1994. Global survey of mitochondrial DNA sequences in the threespine stickleback: Evidence for recent migrations, *Evolution* **48**, 608–622.

QUINN, T. P., AND LIGHT, J. T. 1988. The occurence of threespine sticklebacks in the open North Pacific Ocean: Migration or drift? *Can. J. Zool.* **67**, 2850–2852.

SCHAEFFER, S. W., AND MILLER, E. L. 1993. Estimation of linkage disequilibirum and the recombination parameter determined from segregating nucleotide sites in the alcohol dehydrogenase region of *Drosophila pseudoobscura*, *Genetics* **135**, 541–552.

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

SLATKIN, M. 1989. Detecting small amounts of gene flow from phylogenies of alleles, *Genetics* **121**, 609–612.

SLATKIN, M., AND MADDISON, W. P. 1989. A cladistic measure of gene flow inferred from the phylogenies of alleles, *Genetics* **123**, 603–613.

SLATKIN, M., AND MADDISON, W. P. 1990. Detecting isolation by distance using phylogenies of genes, *Genetics* **126**, 249–260.

STROBECK, C. 1987. Average number of nucleotide differences in an sample from a single subpopulation: A test for population subdivision, *Genetics* **117**, 149–153.

SVED, J. A. 1968. The stability of linked systems of loci with a small population size, *Genetics* **59**, 543–563.

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

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

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

TAKAHATA, N., AND NEI, M. 1985. Gene genealogy and variance of interpopulational nucleotide differences, *Genetics* **110**, 325–344.

TAKAHATA, N., AND SLATKIN, M. 1990. Genealogy of neutral genes in two partially isolated populations, *Theor. Popul. Biol.* **38**, 331–350.

WAKELEY, J. 1996. The variance of pairwise nucleotide differences in two populations with migration, *Theor. Popul. Biol.* **49**, 39–57.

WATTERSON, G. A. 1975. On the number of segregating sites in genetical models without recombination, *Theor. Popul. Biol.* **7**, 256–276.