EVOLUTION

INTERNATIONAL JOURNAL OF ORGANIC EVOLUTION

PUBLISHED BY

THE SOCIETY FOR THE STUDY OF EVOLUTION

Vol. 54 **December 2000** **No. 6**

Evolution, 54(6), 2000, pp. 1839–1854

PERSPECTIVE: GENE DIVERGENCE, POPULATION DIVERGENCE, AND THE VARIANCE IN COALESCENCE TIME IN PHYLOGEOGRAPHIC STUDIES

SCOTT V. EDWARDS[1,2] AND PETER BEERLI[3]

[1]*Department of Zoology and Burke Museum, Box 351800, University of Washington, Seattle, Washington 98195*

[2]*E-mail: edwards@zoology.washington.edu*

[3]*Department of Genetics, Box 357360, University of Washington, Seattle, Washington 98195*

Abstract.—Molecular methods as applied to the biogeography of single species (phylogeography) or multiple codistributed species (comparative phylogeography) have been productively and extensively used to elucidate common historical features in the diversification of the Earth's biota. However, only recently have methods for estimating population divergence times or their confidence limits while taking into account the critical effects of genetic polymorphism in ancestral species become available, and earlier methods for doing so are underutilized. We review models that address the crucial distinction between the gene divergence, the parameter that is typically recovered in molecular phylogeographic studies, and the population divergence, which is in most cases the parameter of interest and will almost always postdate the gene divergence. Assuming that population sizes of ancestral species are distributed similarly to those of extant species, we show that phylogeographic studies in vertebrates suggest that divergence of alleles in ancestral species can comprise from less than 10% to over 50% of the total divergence between sister species, suggesting that the problem of ancestral polymorphism in dating population divergence can be substantial. The variance in the number of substitutions (among loci for a given species or among species for a given gene) resulting from the stochastic nature of DNA change is generally smaller than the variance due to substitutions along allelic lines whose coalescence times vary due to genetic drift in the ancestral population. Whereas the former variance can be reduced by further DNA sequencing at a single locus, the latter cannot. Contrary to phylogeographic intuition, dating population divergence times when allelic lines have achieved reciprocal monophyly is in some ways more challenging than when allelic lines have not achieved monophyly, because in the former case critical data on ancestral population size provided by residual ancestral polymorphism is lost. In the former case differences in coalescence time between species pairs can in principle be explained entirely by differences in ancestral population size without resorting to explanations involving differences in divergence time. Furthermore, the confidence limits on population divergence times are severely underestimated when those for number of substitutions per site in the DNA sequences examined are used as a proxy. This uncertainty highlights the importance of multilocus data in estimating population divergence times; multilocus data can in principle distinguish differences in coalescence time (T) resulting from differences in population divergence time and differences in T due to differences in ancestral population sizes and will reduce the confidence limits on the estimates.

We analyze the contribution of ancestral population size (θ) to T and the effect of uncertainty in θ on estimates of population divergence (τ) for single loci under reciprocal monophyly using a simple Bayesian extension of Takahata and Satta's and Yang's recent coalescent methods. The confidence limits on τ decrease when the range over which ancestral population size θ is assumed to be distributed decreases and when τ increases; they generally exclude zero when $\tau/(4N_e) > 1$. We also apply a maximum-likelihood method to several single and multilocus data sets. With multilocus data, the criterion for excluding $\tau = 0$ is roughly that $l\tau/(4N_e) > 1$, where l is the number of loci. Our analyses corroborate recent suggestions that increasing the number of loci is critical to decreasing the uncertainty in estimates of population divergence time.

Key words.—Coalescent theory, gene flow, intron, maximum likelihood, multilocus data, vicariance biogeography.

Received December 8, 1999. Accepted May 18, 2000.

The biogeographic study of multiple species groups distributed across similar geographic areas, known as comparative phylogeography, is a powerful method for elucidating shared vicariant events and for developing predictive hypotheses in the form of area cladograms (Platnick and Nelson 1978; Cracraft 1988; Avise 1992, 1998; Zink 1997; Walker and Avise 1998). This rapid growth of this research program is evidenced by a recent entire issue of *Molecular Ecology* devoted to the subject (Bermingham and Moritz 1998) and a recent comprehensive review (Avise 2000). Although mo-

lecular clocks are not always reliable indicators of absolute time (Li 1993; Hillis et al. 1996; Ayala 1997), taxon-specific local clocks are probably widespread, can be tested empirically, and for many species without fossil records, hold the only hope of determining the timing of diversification events (Beerli et al. 1996; Sanderson 1998). In addition, application of clocks to biogeography is being broadened by methods for dating gene divergences in the absence of rate constancy (Sanderson 1997). Thus, a common practice is to determine molecular phylogenies for multiple pairs of species across a common biogeographic barrier (or multiple clades across multiple barriers) and to evaluate both the topological and temporal congruence of these phylogenies.

There are a large number of methods for quantifying topological congruence of gene trees (Templeton 1983; Felsenstein 1988; Kishino and Hasegawa 1989; Cunningham 1997); fewer tests have been proposed for comparing temporal congruence of gene trees (e.g., Page 1990, 1994, 1996) or of populations in which those gene trees are embedded (Takahata et al. 1995; Takahata and Satta 1997; Yang 1997). The simplest comparative phylogeographic studies involve gene lineages of several species or population pairs examined across a single biogeographic barrier, such as the Strait of Gibraltar (Busack 1986), the Florida Peninsula (Avise 1992), or the Isthmus of Panama (Knowlton et al. 1993; Knowlton and Weight 1998). Even in phylogenetically straightforward cases in which gene lineages have achieved reciprocal monophyly across the barrier, determination of temporal congruence of the species split is challenging because we cannot observe the population divergence directly, as we can the gene divergence (Nei and Li 1979; Takahata 1986; Wakeley and Hey 1997; Yang 1997). The divergence of populations from the ancestral population is necessarily defined by the cessation of gene flow (Nielsen and Slatkin 2000); whether this cessation coincides with particular geological events or occurs some time after such events is usually not known (Knowlton et al. 1993), but we will assume for this paper, as do many empirical studies, that the cessation of gene flow and the origin of extrinsic geographic isolating barriers are contemporaneous.

The relevant event in most phylogeographic studies is not the split between gene lineages at time T, but the split of the ancestral population of organisms at time τ, which occurred sometime after the gene split (Nei and Li 1979; Wilson et al. 1985). In Figure 1, we illustrate the distinction between T and τ as measured in absolute time (generations) or between $D = \mu T$ and $\gamma = \mu\tau$, the number of substitutions per site since the allelic and population divergences, respectively, where μ is the substitution rate. Typically D is recovered in phylogeographic surveys, but τ is the desired quantity; most theory for estimating τ assumes that μ remains constant (i.e., a molecular clock operates) and focuses primarily on inferring γ from DNA sequence data. If γ and μ are known, then one can estimate τ. If the ancestral population giving rise to the two diverging species or populations was mating at random, then on average the most recent common ancestor (MRCA) of allelic lineages in that ancestral population will be $4N_e$ generations, or Θ mutational events down one lineage, where $\Theta = 4N_e\mu$ and N_e is the effective number of diploid individuals. However, the MRCA of the two alleles that will

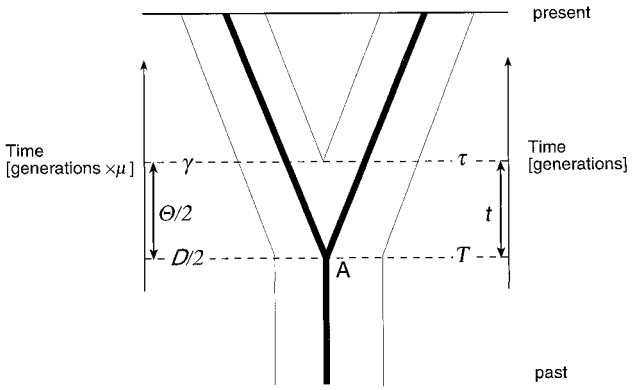

FIG. 1. Schematic diagram showing a single ancestral population splitting into two descendent populations, and associated variables. T is the time in generations since the gene divergence occurred, τ is population divergence, γ is the population divergence τ multiplied by the mutation rate μ, and D is the gene divergence time T scaled by μ. t is the difference between τ and T, and has an expected value of $2N_e$ generations, or $\Theta/2$ mutational events, where $\Theta = 4N\mu$. The bold A marks the most recent common ancestor (MRCA) to which alleles in the two descendent species coalesce, not the MRCA of all the alleles in the ancestral species.

ultimately give rise to alleles in the diverging populations (node A in Fig. 1) will not necessarily be the same as the MRCA for all the alleles in the ancestral population. In a simple speciation scenario, the two alleles that will be the MRCAs of haplotype clades in the two descendent species will be a random pair of alleles from the ancestral population; population genetics theory predicts that the MRCA of this random pair of alleles will occur on average $2N_e$ generations, or $\Theta/2$ mutational units, in the past (Wright 1951; Gillespie and Langley 1979; Nei and Li 1979; Hudson 1990; Tavaré et al. 1997; Fig. 1). On average the value of D (or T, when μ is known) recovered in phylogeographic surveys will overestimate γ (or τ), the parameter of interest, by a quantity equivalent to $\Theta/2$ mutational units, or $2N_e$ generations in absolute time. Throughout this paper, we will use Θ and N_e to refer to parameters of the *ancestral* population from which two species diverged, not the sizes of extant populations.

Although fundamental to the accurate determination of population divergence times, the distinction between T and τ has been overlooked in many recent applications of phylogeography (for reviews of avian studies see Edwards 1997; Avise and Walker 1998; Klicka and Zink 1999). For example, of the many articles in the April 1998 issue of *Molecular Ecology* that used molecular data to infer relative or absolute divergence times, only a few (e.g., Fleischer et al. 1998; Schneider et al. 1998) discussed implications of ancestral polymorphism for these divergence times. The recent neglect of the distinction between gene and population divergence is particularly surprising given the coincident increase in interest in diversification during the Pleistocene (Zink and Slowinski 1995; Riddle 1996, 1998; Klicka and Zink 1997; Bermingham and Martin 1998; da Silva and Patton 1998) and in even more recent events such as the origin of domesticated animals (Vila et al. 1997) and of humans (Vigilant et al. 1991; Dorit et al. 1995; Hammer 1995; Semino et al. 2000), for it is precisely when estimating such recent events that incor-

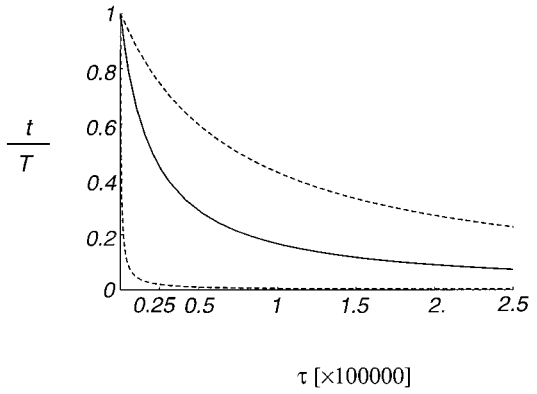

FIG. 2. The fraction of the total interspecific coalescence $T = t + 2N_e$ generations taken up by ancestral divergence coalescence t. An effective population size of 100,000 was assumed. Dashed lines indicate upper and lower 95% confidence intervals.

porating information on ancestral polymorphism is most critical. Despite this, it is phylogeography itself that has been crucial to the modern distinction between gene and population divergences and to the development of new theory modeling this distinction. The problem of correcting for ancestral polymorphism is a long-recognized and troubling challenge to the phylogeography; this paper reviews older and newer approaches to the problem and points the way to some new directions.

Magnitude of the Ancestral Overestimation Problem

The difference between T and τ will become inconsequential as the divergence time between species becomes large, because the discrepancy t between T and τ will be a small fraction of the total gene divergence ($t + \tau = T$). However, as illustrated in Figure 2, when divergence times between populations are short, t can make a substantial contribution to the total gene divergence T between species. Studies that ignore the distinction between τ and T implicitly assume a large value of τ/N_e or that $N_e = 0$, resulting in overestimates of population divergence times or estimates with unrealistically small confidence limits. The expected amount of overestimation of τ by $2N_e$ outlined above and depicted in Figure 2 assumes a randomly mating ancestral population. Such overestimation will influence a number of important issues in the study of molecular evolution and speciation, such as calibration of molecular clocks, the number of speciation events estimated to have occurred before the Pleistocene, the tempo of speciation as inferred from molecular data (Gillespie and Langley 1979; Klicka and Zink 1997; Avise and Walker 1998), and cospeciation studies (Huelsenbeck et al. 2000).

The expected degree of overestimation of τ if the ancestral species is structured into multiple, semi-isolated populations could be much greater than when the ancestral species is unstructured (Wakeley, unpubl. ms.). The total coalescence time of a structured species and the expected coalescence time for a random pair of alleles increases dramatically as the number of populations increases and migration rate between populations decreases, as does the variance (Takahata 1991; Nei and Takahata 1993; Hoelzer 1997). In fact, the equilibrium coalescence time approaches infinity under limited migration, implying that the extent of overestimation could become quite large. Another reason why the usual model for calculating the extent of overestimation may be inadequate is that ancestral coalescence time is generally modeled as an exponential distribution. In this case, the mean overestimation is not an optimal measure of central tendency and extremely high values of overestimation can arise at higher frequencies than if the distribution were normal. The median of the distribution (x) would be a better indication of the usual extent of overestimation, but we will ignore this for the present study.

The geography of speciation of this ancestral species will also influence the extent of overestimation (Moore 1997; Nordborg 1997); for example, in isotropic models, if the alleles that give rise to those in descendant species derive from the same subpopulation, we would not expect the overestimation problem to be exacerbated relative to the problem with random mating ancestor of the same total size, because the coalescence time of such alleles is unaffected by migration (Slatkin 1987). However, if the alleles that give rise to those in descendant species derive from different subpopulations, the extent of overestimation will be greater than that for a random mating ancestor and will depend on a number of factors (Slatkin 1987; Nei and Takahata 1993; Wakeley 1999). Clearly the effect of neglecting to incorporate ancestral coalescence times into estimates of population divergence could be drastic. Ultimately, however, it is difficult to infer ancestral population sizes from those of extant species because we do not know how ancestral species are structured just prior to speciation, unless such information is provided by still segregating ancestral polymorphisms (Wakeley and Hey 1997).

If we assume that the effective population sizes of ancestral species are on average similar to those of extant species, then we can use empirical data on coalescence times within extant species to gauge the severity of ignoring ancestral polymorphism. The ratio of the average pairwise divergences within species to gene divergences separating extant sister species provides a clue as to the magnitude of the ancestral polymorphism problem, a clue that overcomes uncertainty in the absolute mutation rate μ, which will cancel in such a measure (Fig. 1; see Palumbi and Citriano 1998). Average pairwise divergence of mitochondrial haplotypes has recently been summarized for several vertebrate groups (Moore 1995; Bernatchez and Wilson 1998). Moore (1995) estimated the average maximum D for mtDNA within species of birds as 0.07 (which, if the rule of 2%/million years applies, corresponds to approximately 350,000 years). Because this value represents a hypothetical total coalescence time in the ancestor, we halve this estimated value to 0.035 (175,000 years) because the alleles in the ancestral species that ultimately give rise to those found in extant species will on average have an MRCA halfway down the tree. (Thus, the value of 350,000 years, proffered as a generalized ''correction'' of gene divergences in recent ornithological work [Moore 1995; Edwards 1997; Klicka and Zink 1997], is actually an overcompensation by 175,000 years, making the distinction between gene and species divergence smaller). The ratio of Moore's (1995) average for the avian cytochrome *b* gene, 0.0035, to

average sister-species divergence for birds, 0.029 (calculated from restriction fragment length polymorphism [RFLP] and cytochrome *b* data for the 27 avian haplogroups in table 2 of Avise and Walker [1998] that were not already corrected for divergence in the ancestor), is 0.12, suggesting that on average 12% of avian haplogroup divergence is taken up by divergence in the ancestor. Bernatchez and Wilson (1998) recently summarized intra- and interspecific divergence in North American fishes. For 12 species pairs with relevant data, values in their appendices 1 and 2 suggest that this ratio varies from 0.045 (*Onchorhynchus mazu* and *O. mykiss*) to 0.53 (*Lepomis punctatus* and *L. microlophus*), with a mean of 0.21. Although we have not done an exhaustive survey, this line of reasoning suggests that neglecting to incorporate ancestral diversity into estimates of population divergence could in some cases result in overestimation of the latter by over 50%. However, Moore and, to a certain extent, Bernatchez and Wilson focused on taxa within which no appreciable population subdivision was known; therefore, using such species as surrogates for ancestral population sizes could result in underestimation of the latter (Hoelzer 1997; see also Klicka and Zink 1999).

Sources of Variance in Coalescence Times

Infinite Sites Model of Two Diverging Populations

The expected variance in coalescence time of alleles in an ancestral population cannot only be used to erect confidence limits on population divergence times (Nei and Jin 1989), but can also be used to make inferences about the size of the ancestral population (Gillespie and Langley 1979; Takahata 1986). We refer to the situation outlined in Figures 1 and 2 in which alleles in both descendent populations have reached reciprocal monophyly, a condition that will hold after the two descendent populations have been separated by on average $4N_e$ generations (Neigel and Avise 1986; Nei 1987). We are concerned here with the decoupling of gene and population divergence *times,* not with the multiplicity of genealogies across loci or the possibility that gene lineages may not have yet achieved reciprocal monophyly because of the recency of population splitting; these questions have been addressed recently by several groups (Moore 1995; Brower et al. 1996; Avise and Wollenberg 1997; Maddison 1997; Wollenberg and Avise 1998). As stated above, the two allelic lines leading to the MRCAs between the descendent species will consist of a random pair of alleles from the ancestral population, and thus will have a mean coalescence time t in the ancestral population of $2N$ generations, where N is the number of diploid individuals (Fig. 1). Because t follows an exponential distribution (Kingman 1982a,b) and τ is a constant, the variance of time of the MRCA of the entire two-species tree with a gene divergence at $T = \tau + t$ will be

$$\sigma(T) = \sigma(\tau + t) = \sigma(t) = (2N_e)^2 \quad (1)$$

(Hudson 1990; Donnelly and Tavaré 1995). This is the variance in $\tau + t$ contributed by the effects of random drift in the ancestral population. It is also the variance in coalescence time for different loci sampled from the same pair of species (Takahata 1986). DNA sequence data allows us to measure T indirectly as D (Fig. 1), and the stochastic process of mutation will also contribute to the variance of the estimate of D. If sequences are evolving according to a Poisson process under a molecular clock and an infinite sites model, the expected number of differences between alleles in sampled from two descendent species is

$$E(D) = 2\mu n(\tau + t) = 2\mu n(\tau + 2N_e) = n(2\gamma + \Theta), \quad (2)$$

where μ is the mutation rate and n is the number of sites in the DNA sequence, $\gamma = \mu\tau$ and $\Theta = 4N_e\mu$ in the ancestral population. The total variance of the number of substitutions D, including variance associated with ancestral population size, is

$$\begin{aligned}\sigma(D) &= 2\mu n(\tau + 2N_e) + (2\mu n)^2(2N_e)^2 \\ &= n(2\gamma + \Theta + \Theta^2 n). \end{aligned} \quad (3)$$

The variance in the number of substitutions per site $d = D/n$, is

$$\begin{aligned}\sigma(d) = \sigma\left(\frac{D}{n}\right) &= \frac{1}{n^2}\sigma(D) = \frac{2\mu\tau}{n} + \frac{2\mu(2N_e)}{n} + (2\mu)^2(2N_e)^2 \\ &= \frac{2\gamma + \Theta}{n} + \Theta^2 \end{aligned} \quad (4)$$

The variance associated solely with the process of nucleotide substitution is

$$\begin{aligned}\sigma(d) - \sigma(\mu T) &= \frac{2\mu\tau + \Theta}{n} + \Theta^2 - \left(\frac{\Theta}{2}\right)^2 \\ &= \frac{3}{4}\Theta^2 + \frac{2\gamma + \Theta}{n}. \end{aligned} \quad (5)$$

Equations (3–5) show that, as the number of bases sequenced, n, becomes large, the variance in the estimated number of substitutions per site between alleles in two species will asymptote to a value of $3\Theta^2/4$. Although $3\Theta^2/4$ will usually not be a large number, the variance on sequence divergence does not approach 0. Sequencing L multiple independent loci will, however, reduce the variance of D approximately by a factor of L:

$$\sigma(D_L) = \frac{\sigma(D)}{L}. \quad (6)$$

Obviously one can greatly increase the precision with which one estimates average D, and thus T, by sampling multiple loci.

Multiple Pairs of Populations

If all species pairs diverged at exactly the same time τ, we can express the expected ratio ρ of the among-pair variance in coalescence time to the mean coalescence time in the same way that Gillespie and Langley (1979) expressed the ratio of the variance to the mean coalescence time among loci for a single species pair:

$$\begin{aligned}\rho = \frac{\sigma(D)}{E(D)} &= \frac{2\mu n(\tau + 2N_e) + (2\mu n)^2(2N_e)^2}{2\mu n(\tau + 2N_e)} \\ &= \frac{2\mu\tau + \Theta + \Theta^2 n}{2\mu\tau + \Theta}, \end{aligned} \quad (7)$$

which, after rearranging, becomes:

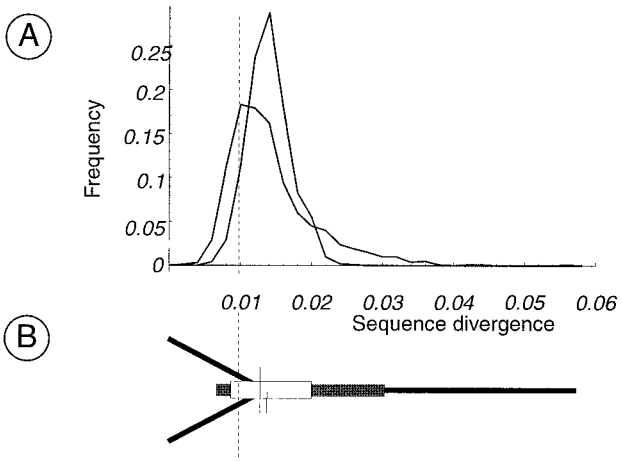

FIG. 3. Contribution to the total variance of $\gamma = \mu\tau$ of mutational variance and coalescent variance. Vertical dashed line indicates the population divergence τ. (A) Histograms showing two distributions of γ. Lower curve: Distribution of γ including both coalescent and mutational variance. 1000 two-tipped trees were generated using the coalescent from a pair of species that diverged $\tau = 10^6$ generations in the past. The mutation rate for the locus was $\mu = 10^{-8}$ per lineage per site per year (equivalent to 2%/million years), making $\gamma = 0.01$. The ancestral population size prior to speciation was $N_e = 2500$, making $\Theta = 0.01$. On each of these trees, 1000 DNA sites with even base composition were allowed to evolve via Kimura's (1980) 2-parameter model. The divergence of each pair of sequences was estimated using DNADIST, and a histogram of these 1000 divergence values was generated. Upper curve: Distribution of γ without taking coalescent variance into account. A single two-tipped tree with common ancestor at exactly $\tau + 2N_e$ generation in the past was used to generate 1000 DNA datasets of 1000 sites as above. These datasets were analyzed and plotted as above. (B) Visualization and summary of 95% confidence interval of γ with (gray bar) and without (white bar) coalescent variance on a two-tipped tree.

$$1 + \frac{\Theta^2 n}{2\gamma + \Theta}. \tag{8}$$

ρ is the ratio of the variance in coalescence times among species pairs with the same ancestral N_e to the mean coalescence time of those pairs, all of which diverged at the same time τ.

Using a simple simulation, we visualized the contribution of mutation and ancestral population size to the total variance in gene divergence expected for two populations that diverged τ generations ago (Fig. 3). Consistent with the above results, the simulation shows that the mutational variance on $\gamma = \mu\tau$ is a moderate fraction of the total variance on γ. Because this simulation used a finite-sites approach, we would not expect the mutation variance to go to zero even if we sequenced an infinite number sites. The distribution of γ including coalescent variance has a long tail, indicating that, given a known τ, we can expect to recover empirically a wide range of actual gene divergences D (Fig. 1). In addition, Figure 3 emphasizes that there is a finite chance that one will recover a gene divergence D that is *less* than the lower 95% confidence interval on γ; thus in some cases the empirically recovered D appears as if it is less than the population divergence time.

Interpretation of Among-Species Variation in Coalescence Times

Variance in D observed among codistributed species in phylogeographic studies is often interpreted as variation in species divergence, or, in some cases, as variation in substitution rates among species pairs (Avise et al. 1992). Such variance could also be interpreted simply as variation in ancestral population size (Avise 2000, pp. 232–235). For example, when confronted with data on mean and variance in D among different loci of species pairs of mammals Gillespie and Langley (1979) and Gillespie (1991) ask what value of ancestral Θ will allow ρ to match the observed, empirical value, and then assess whether this value for Θ could reasonably occur in nature. Gillespie and Langley (1979) suggested that much of the apparent variation among genes in number of substitutions observed could in fact be explained when the variance contributed by ancestral polymorphism was taken into account, in part because the value of Θ required to explain the observed ρ was ''not unreasonable.'' Lynch and Jarrell (1993) also emphasized the importance of ancestral polymorphism in calibrating rates of molecular evolution. We suggest that the null hypothesis in comparative phylogeographic studies should be equivalence of τ between species pairs. One way of rejecting this hypothesis would be to show that the value of Θ required to explain the observed variance is biologically implausible. This is particularly relevant for single locus comparative phylogeographic studies in which reciprocal monophyly has been achieved, because in such cases one cannot in principle distinguish whether variance in D results from variance in τ or from variance in ancestral Θ.

As an another example of this approach, we can make use of a recent dataset on mitochondrial DNA divergences in North American birds (Klicka and Zink 1997). Klicka and Zink assembled RFLP and sequence data for 35 pairs of North American birds, and presented a distribution of coalescence times for these pairs (their fig. 1). Klicka and Zink (1997) were more concerned with what their data implied about absolute divergence times—pre- or post-Pleistocene—than with whether their data implied the *same* divergence time for the various pairs. Nonetheless, it is interesting to ask under what conditions their data is compatible with a similar coalescence time for these avian species pairs. Expressed as percent sequence divergence over all sites, the mean divergence among these pairs was 5.1%, which corresponds to 2.5 million years ago if the 2%/million years rate assumed by Klicka and Zink (1997) is correct. Using the 2%/million years mitochondrial clock, Klicka and Zink estimated a range of coalescence times from 200,000 to 5.55 million years ago. If we naively assume that all these estimates in fact record population divergences at a single vicariant event (a situation that forces the population divergences to have been no older than the minimum estimated T of 200,000 years) and that variance in sequence differences resulting from mutation is negligible, we can ask what ancestral N_e would be compatible with such a wide range of coalescence times. For mtDNA, equation (1) changes to $\sigma(t) = N_f^2$, where N_f is the number of females. The standard deviation of the coalescence times of these 35 spe-

cies pairs is 1.5 million years ago, and, for the Great Plains species, 1.3 million years ago. Assuming there is one generation per year, equation (1), the variance in coalescence time among species pairs, can be explained if we invoke an ancestral N_f for mtDNA of 1.5×10^6 females, or, under equal sex ratios, 3×10^6 individuals (or 2.6×10^6 for Great Plains species). These ancestral sizes explain the variance in coalescence time among pairs without requiring any differences in population divergence time. Given what we know of long-term effective population sizes in birds (Moore 1995), these values seem inordinately large, and population size alone seems unlikely to be the sole source of variance in coalescence time observed among species pairs. This conclusion might be valid, however, if we have reason to believe that avian ancestors that give rise to new clades were widespread and subdivided for long periods of time, a scenario that is somewhat at odds with genetic patterns. Regardless, were we to reject the large ancestral population scenario with a single locus such as mtDNA, we reject it not because a single vicariant event is itself unrealistic, but because of the unrealistically large ancestral population sizes that one would have to invoke to have the scenario plausible. Ultimately, inference of population divergence times from a single locus for which reciprocal monophyly has been achieved is tantamount to making assumptions about the effective size of the ancestral species. Without reference to external data from geology or morphology, the only certain thing one can say with such data is that the species pairs diverged sometime after their respective gene coalescences. In such cases, we simply do not know how much of the total gene divergence is a result of divergence within the ancestral species and how much is a result of divergence since species separation.

The above exercise implies an important caveat to single-locus studies of comparative phylogeography. It reinforces the population genetic result that the expected difference in T among replicate species pairs is not an absolute, but a relative value that depends on the ancestral N_e. Another relevant issue, particularly when but a single allele is sampled in the two descendant species, is whether allelic lineages in the two descendant species will have had time to achieve reciprocal monophyly given an observed D and a particular value for ancestral N_e (when a single allele per locus is sampled, we cannot directly observe whether reciprocal monophyly has been achieved). Some large values of ancestral N_e might make an observation of reciprocal monophyly in the two descendant species unlikely if this ancestral N_e has gone unchanged throughout the history of the divergence. But founder events and other reductions in population size in one or both descendent species could reconcile a large ancestral N_e with observations of reciprocal monophyly in the two descendant species, because such reductions in population size are expected to force coalescent events into a short, recent time window (Slatkin and Hudson 1991; Griffiths and Tavaré 1994; Wakeley and Hey 1997; Kuhner et al. 1998). Although such scenarios may not be parsimonious, they can nonetheless explain variance in D among codistributed species without resorting to scenarios involving differences in τ.

Methods for Estimating Population Divergence Time τ

Methods Ignoring Gene Divergence in the Ancestor

Many researchers assess significant differences in τ by using D and its standard error as a proxy for statistical tests. The standard errors of D for several models of nucleotide change are known (Nei 1987). In addition, several recent methods for designed to assess temporal congruence in D have been proposed. For example, Steel et al. (1996) suggested ways of tightening the confidence intervals for D by incorporating information on the variation in distances among species stemming from a common ancestor. Such methods are most often used, however, to measure τ, not D; as such they are most appropriate for deeper divergences, for which ancestral polymorphism is less of a problem (Fig. 2). Other methods, although not designed to estimate divergence times explicitly, assess the simultaneity of species splits based on the expectation of a correlation in gene divergences when measured across multiple species in two independent clades that have diversified as a result of similar temporal events (for a recent application of this approach, see Bermingham and Martin 1998). Because of their reliance on a correlation coefficient measured from multiple divergence estimates, such methods are more appropriate to multispecies problems, rather than simple pairs of species, but they have the advantage of allowing for rate differences between clades, and specific methods for determining the expected correlation under various null models have been proposed (Hafner et al. 1994; Page 1996). Again, however, these approaches ignore the variation associated with stochastic variance in the ancestral population, although Hafner and Page (1995) acknowledged the potential importance of this factor.

Methods Accounting for Gene Divergence in the Ancestor

Several major methods have been proposed for estimating τ (Table 1). All of these methods take into account possible divergence of alleles in the ancestor when necessary. Methods based solely on divergence of allele *frequencies* in two diverging populations (Cavalli-Sforza 1969; Nei 1972) do not have to address allelic divergence in the ancestor because divergence of populations and of allele frequencies occur over identical time periods. The methods in Table 1 differ in their basic assumptions, requirement of reciprocal monophyly, and suitability for multilocus data. Most of the methods are formulated to handle a single pair of species, although this pair does not necessarily have to be sister species (see Takahata 1986). Several of the methods, such as those of Nei (1972) and Nielsen et al. (1998) are more appropriate for electrophoretic or RFLP markers than for DNA sequences because they rely on information from allelic counts or frequencies, which can be influenced not only by drift but also by high mutation rates.

Methods accommodating reciprocal monophyly

We cannot normally estimate both Θ and τ from a single pair of sequences because in this case there is no information on ancestral Θ or even information on whether reciprocal monophyly has been achieved. When multiple copies of a

TABLE 1. Summary of methods for estimating the population divergence time τ. Only methods that distinguish between the gene divergence and population divergence are listed.

Method	Description	Appropriate for DNA sequence data?	Is reciprocal monophyly desired or required?	Designed for multilocus data?	Major assumptions and (dis)advantages	Major references
1. Genetic distance or F_{ST} methods	Uses divergence in allele frequencies, which occurs over a time frame identical to the divergence of populations	No	Alleles must show variability across populations	Yes	Assumes mutation rate is low; does not take advantage of all information	Cavalli-Sforza 1969; Nei 1972; reviewed in Nielsen et al. 1998
2. Pairwise method	Subtracts average pairwise divergence within species from gene divergence between species	Yes	Not required	No	Assumes ancestral N_e is the average of two descendant N_es	Nei and Li 1979
3. Moment method	τ estimated using mean and variance of coalescence time	Yes	Required	Yes	Assumes all loci have similar mutation rates	Takahata 1986
4. Cladistic method	Uses the number of between-population coalescent events (s) to estimate τ	Yes	Not desired	No	Assumes gene tree is accurate and compares s to results from simulations	Slatkin and Maddison 1989 (see their fig. 8)
5. "Two-tip" maximum-likelihood methods	Finds values of τ and θ that maximize the likelihood of multilocus datasets	Yes	Required	Yes	Requires only one sequence per species/population; can accommodate rate variation among loci (Yang 1997) and finite-sites models (this paper)	Takahata et al. 1995; Takahata and Satta 1997; Yang 1997; this paper
6. Segregating sites method (SITES)	Uses shared polymorphic sites to estimate ancestral θ	Yes	Not desired	No	Requires segregating sites within both descendant populations; descendent populations can differ in size; extendable to multiple populations and loci	Wakeley and Hey 1997; Wang et al. 1997
7. Coalescent likelihood method	Uses differences in counts of alleles between populations to estimate τ and ancestral θ; integrates over all possible allelic configurations in ancestor	No	Alleles must show variability in at least one population	Yes	Assumes non mutation during period of divergence and thus is most appropriate for allozymes, nuclear RFLPs* or SNPs**	Nielsen et al. 1998
8. Coalescent likelihood method	Integrates over all possible haplotype configurations in ancestor to estimate τ	Yes	Not required	Yes	Assumes infinite-sites model	Nielsen 1998
9. Bayesian "two-tip" method	Integrates over all values of ancestral θ given a specified range and distribution	Yes	Required	Yes	Can be used on single-locus data	this paper

* RFLP, restriction fragment length polymorphism; ** SNP, single nucleotide polymorphism.

single locus are sampled from a pair of species, we still cannot estimate both parameters if reciprocal monophyly has been achieved because the information on ancestral Θ, which otherwise can be gleaned from shared polymorphic sites and other signals of incomplete lineage sorting, is lost upon reaching reciprocal monophyly. In contrast, both parameters can be estimated with one locus in situations in which ancestral polymorphisms are still segregating (Wakeley and Hey 1997). Thus, contrary to phylogeographic intuition, with either single-locus or multilocus data it is in some respects easier to estimate τ when gene trees have not reached reciprocal monophyly than when they have. However, the time during which alleles have not achieved reciprocal monophyly is transient and represents a nonequilibirum condition, a situation that may compromise strict application of coalescent theory.

Perhaps the earliest commonly used method for RFLP or sequence data for determining the population divergence time τ was proposed by Nei and Li (1979). Their measure, *d,* which Avise later coined ''net nucleotide divergence,'' is $d = d_{xy} - 0.5(d_x + d_y)$, where d_{xy} is the average divergence between haplotypes between populations *x* and *y,* and d_x and d_y are the average pairwise divergence between haplotypes within populations *x* and *y*. Takahata and Nei (1985) showed that $d_{xy} = 2\mu\tau + \Theta/2$ and if $0.5(d_x + d_y)$ is assumed to represent ancestral $\Theta/2$, then τ can be estimated (see also Takahata and Tajima 1991). The measure was widely used in the RFLP era of mitochondrial phylogeography (Wilson et al. 1985; Avise et al. 1987), but its use has waned in more recent studies (Edwards 1997). The method can be used whether reciprocal monophyly has been achieved, requires examining multiple haplotypes within the two descendant species, and assumes that the ancestral population from which populations *x* and *y* diverged had a size exactly intermediate between those currently exhibited by the locus in question for the two descendant species. This may be a restrictive assumption, particularly as new models for estimating changes in population size are introduced. However, as of now, the only methods for directly estimating the population size of a common ancestral species under reciprocal monophyly are those of Takahata and Satta (1997), Yang (1997), and the finite-sites extension discussed in this paper; other population size change models focus on the scenario of a single evolving lineage (Slatkin and Hudson 1991; Griffiths and Tavaré 1994; Wakeley and Hey 1997; Kuhner et al. 1998). The method of Nielsen et al. (1998), although more appropriate for slowly evolving markers, also estimates ancestral Θ as well as the phylogeny for up to three populations. Because d_x and d_y are reasonable estimates of $\Theta/2$ in the extant species, they will be good estimates $\Theta/2$ in ancestral species when average population size has remained the same, and the estimation method is simple enough that it should be used more frequently.

Nielsen et al.'s (1998) method relies primarily on differences in counts of alleles or haplotypes in diverging populations and as such cannot be categorized as accommodating reciprocal monophyly or not. This method can accommodate situations in which populations show alternate fixations of alleles (akin to reciprocal monophyly), but do require some variability across both populations. It is preferable to other methods that focus on allelic frequencies or counts (Cavalli-Sforza 1969; Nei 1972) because it explicitly models the coalescent process and thus makes more use of the available information. In a different model, Nielsen (1998) used an infinite-sites likelihood method using an extension of the coalescent models of Griffiths and Tavaré (1994) and applied it successfully to two datasets in which reciprocal monophyly had and had not been achieved. The infinite-sites assumption here will probably be adequate for situations in which the mutation rate is small. This particular method is also apparently computationally very slow, making it inapplicable to many datasets (Nielsen 1998). The Nielsen (1998) and Nielsen et al. (1998) methods are computationally different and applicable to different kinds of data, but are similar in taking full account of the lineage sorting process in a coalescent-likelihood framework.

Methods that work better without reciprocal monophyly

Other more recent methods for estimating τ take advantage of site patterns that are still shared between diverging populations since separation from an ancestor. Slatkin and Maddison's (1989) method, which requires a reliable haplotype tree and has mostly been used to estimate gene flow, can in fact be used to estimate divergence time if one is willing to assume an isolation model, that is, that para- or polyphyly of allelic lines is due to recency of population isolation rather than migration (see their fig. 8). Edwards (1993) examined this method for investigating population divergence times. Wakeley and Hey (1997) and Wang et al. (1997) proposed methods for estimating τ that incorporate information from the frequencies of sites that are segregating within descendent populations. Although in principle applicable to situations in which reciprocal monophyly has been achieved, this method is less powerful in this situation because all sites provide essentially the same picture of the genealogical history (J. Wakely, pers. comm.).

Methods requiring reciprocal monophyly

As described above, when a single, reciprocally monophyletic locus is used, it is difficult distinguishing between differences in coalescence time among species pairs resulting from differences in species divergence times and those resulting from differences in ancestral population size. Multilocus approaches to phylogeography and speciation have become more frequent in recent years (Slade et al. 1993; 1998; Friesen et al. 1997; Hilton and Hey 1997) and provide more robust datasets for estimating τ when alleles at all loci have achieved reciprocal monophyly.

The variance in coalescence time of multiple unlinked loci for a pair of species provides valuable information on the size of the ancestral population, which can in turn be used to estimate τ if μ is known (Takahata 1986; Takahata et al. 1995). These methods were a major advance because, by estimating directly the size of the ancestral population, they allowed a distinction to be made between the effects on coalescence time of population divergence and the effects resulting from polymorphism in the ancestral species (Fig. 4). Takahata et al. (1995) developed a likelihood model for analyzing sequence divergence from multiple loci for pairs of

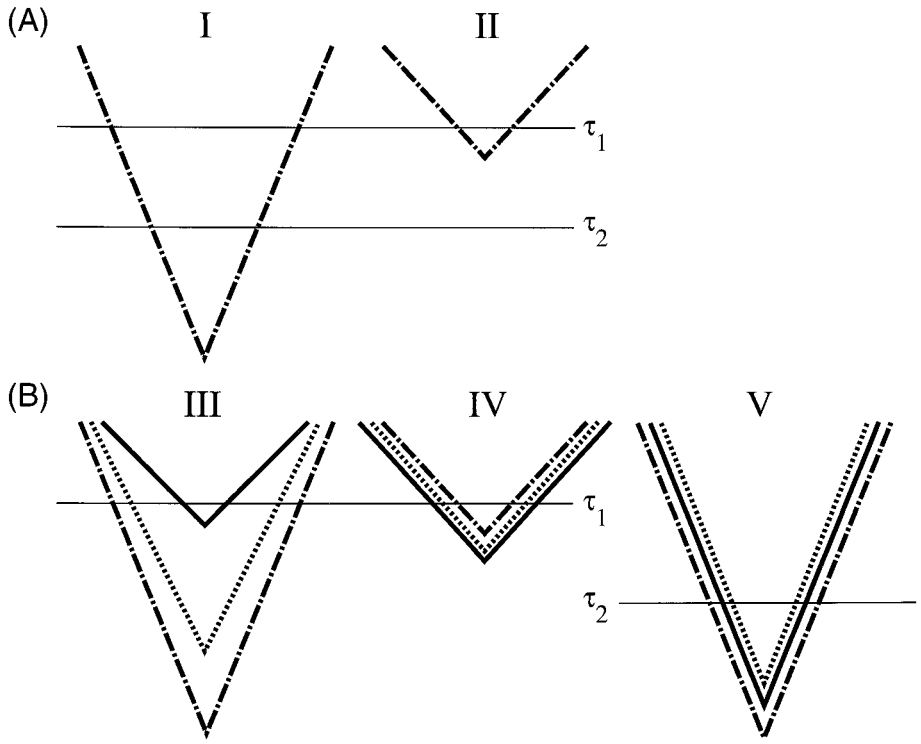

FIG. 4. Advantages of multilocus approaches to estimating population divergence time, using Takahata's (1986) argument. (A) In the single-locus case, when retrieving gene trees I and II, we cannot distinguish between population divergence times τ_1 and τ_2. If divergence times are assumed to be the same when gene trees I and II are recovered, population divergence time is only constrained to be less than τ_1. (B) With multilocus data (gene trees III–V) divergence due to ancestral population size can be assessed by the variance in coalescence time among loci and can be distinguished from divergence due to differences in population divergence time. When variance among loci is large (tree III), population divergence time is likely more recent (τ_1); when variance among loci is small (tree V), population time is likely more ancient (τ_2). Multilocus datasets also narrow confidence limits and the range of compatible population divergence times between codistributed species pairs.

species. This method finds values of ancestral Θ and γ that maximize the probability of a given distribution of numbers of sequence differences at L loci. Takahata and Satta (1997) used this method on silent sites of multilocus data and a substitution rate of 1×10^{-9} substitutions/site/year from pairs of primate species to suggest that the population divergence of chimps and humans occurred approximately 4.5 million years ago, and that the divergence between gorillas and humans was approximately 8 million years ago. Their method is extendable to various models of nucleotide substitution. Because differences in the numbers of substitutions between loci for a given species pair could result not only from the stochastic coalescent variance, but also from variation in substitution rate among loci, it is important to incorporate variation in substitution rate into such models (Yang 1997; Takahata and Satta 1997). Yang (1997) applied a discretized gamma distribution to the evolutionary rates among loci and developed an infinite-sites method for finding the maximum-likelihood values of Θ and γ. Yang showed that estimates of γ were very sensitive to the value of α, the among-locus rate variation parameter, and that ignoring rate variation among loci tends to overestimate Θ when both Θ and γ are being estimated. This is because variation in D among loci is attributed entirely to coalescent variance, rather than to among-locus rate variation, when the latter parameter is ignored. Yang also suggested that the standard errors on these estimates are frustratingly large, particularly for recent divergences, and that large amounts of data from many loci would be required to estimate all population parameters with reasonable accuracy. Knowledge of any of the parameters gleaned from independent phylogenetic analyses, for example, such as the extent of rate variation among loci, generally increased accuracy of estimation of the other parameters.

VISUALIZING THE EFFECTS ON τ OF DIVERGENCE IN THE ANCESTOR

Yang's (1997) infinite sites method for estimating ancestral Θ and γ, as currently implemented, requires only the numbers of differences observed between two species at multiple loci that have achieved reciprocal monophyly, and, as such, assumes a Jukes-Cantor (1969) model of nucleotide change. However, it can easily be extended to a finite-sites model with more complex models of nucleotide substitution. We can evaluate the likelihood for a given divergence time τ and a given ancestral population size Θ for pairs of species (see Yang 1997):

$$L(\tau, \Theta) = \int_{t=0}^{t=\infty} P(t, \Theta)P(D, \tau + t)\, dt$$

$$= \int_{t=0}^{t=\infty} e^{-2t/\Theta}\frac{2}{\Theta}L(D \mid \tau + t)\, dt, \qquad (9)$$

where $P(\tau, \Theta)$ is the coalescence probability (Kingman 1982a,b) and where $L(D \mid \tau + t)$ is the likelihood of the data given the divergence time between the two sequences $\tau + t$. The coalescent probability is the exponential waiting time until a coalescent event, proceeding backward in time, multiplied by the chance of that coalescent event. Between the present time and τ, we assume that the species have remained separate and do not need to consider the possibility that they are in the same population. Therefore, we do not need to consider τ in the calculation of the coalescence probability (cf. Fig. 1). We do need the full time interval $\tau + t$ for the calculation of the likelihood of the data, however, because this time encompasses the entire gene divergence. In the following applications, we employ the Kimura two-parameter model (Kimura 1980) using numerical integration; actual DNA sequences, rather than just the number of differences between species for each locus, are required.

Estimation Using One Locus under Reciprocal Monophyly

In the following we use our finite-sites extension of Yang's (1997) model to visualize the effects of ancestral θ on estimates of τ. Because we cannot estimate both Θ and γ from a single locus in which reciprocal monophyly has been achieved, we need to choose which parameter to estimate. Information on ancestral Θ is certainly of interest, but most biogeographers would like to know the species divergence time τ. If we have some idea about the ancestral Θ, we can use that a priori knowledge and maximize the likelihood function in equation (9) with respect to γ for a constant and known Θ. In Figure 5A–C and Table 2, we illustrate this approach for simulated data and two single-locus datasets. These results show that if we want to assume that ancestral Θ is small on the scale of γ (roughly when $\tau/N_e > 1$ or $\gamma/\Theta > 1$), we can get reasonably precise estimates of γ, but not when this ratio is small. We have plotted the likelihood curves in Figures 5–7 such that the intersections of the curve with the x-axis at -2 mark the 95% confidence limits on the maximum-likelihood estimate. This provides a convenient way to visualize the confidence limits and also to test the null hypothesis of simultaneous diversification of species pairs, which cannot be rejected if the curves for two species pairs overlap above two likelihood units from the maximum (see below). In Figure 5, when Θ is assumed to be large compared to γ, the 95% confidence intervals are very wide and include $\gamma = 0$ (Fig. 5A). Yang (1997) found a similar result even for multilocus data when analyzing a dataset of human DNA sequences. As expected, the discrepancy between γ and D (or τ and T) increases at shorter divergence times (Fig. 5B vs. Fig. 5 C; Table 2) when ancestral Θ is taken into account, and the maximum-likelihood estimate of τ can sometimes be slightly more than the gene divergence due to the stochastic nature of the substitution process (Fig. 3).

In practice, we do not know the precise value of ancestral Θ. One can incorporate uncertainty of Θ into estimates of γ by taking a Bayesian approach. By integrating over all possible Θ, we are estimating the $L(\tau, \Theta)$ of equation (9) for each value of Θ assuming some arbitrary prior distribution of Θ (e.g., Casella and Berger 1990).

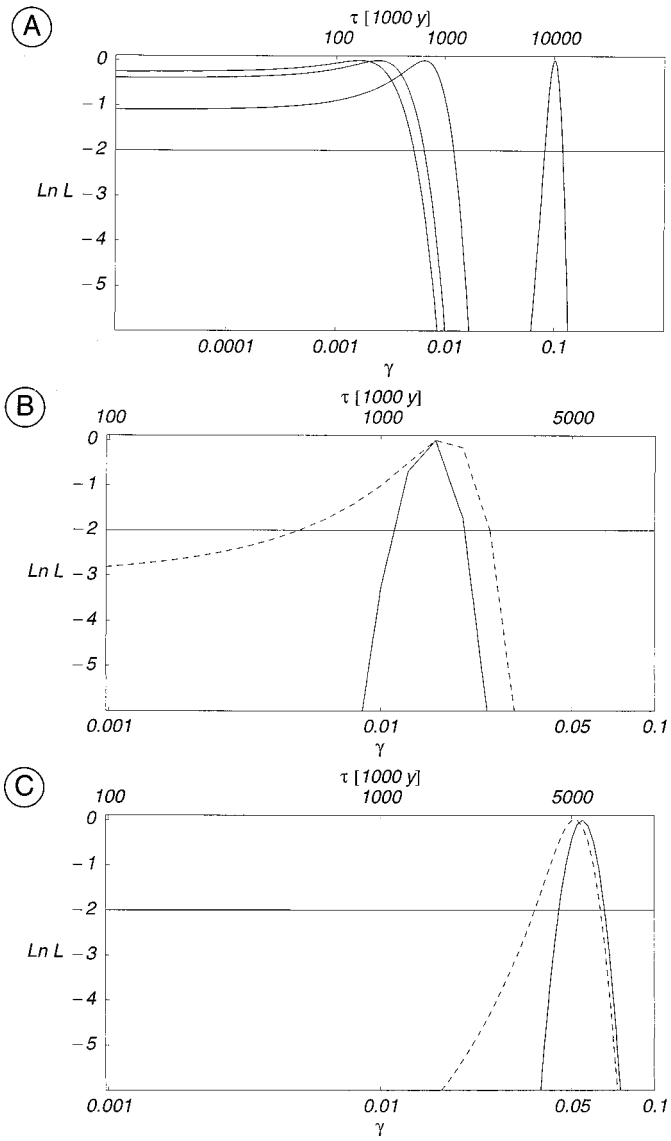

FIG. 5. Likelihood function for the population divergence time γ for a given Θ with one locus. Above each graph are population divergence times τ assuming $\mu = 10^{-8}$ per lineage per site per year as in Figure 3. (A) DNA sequence data was simulated using the same approach as in Figure 3A, using an ancestral population size Θ of 0.01 and a population divergence time, γ, of 0.0001, 0.001, 0.01, and 0.1. (B) Analysis of data from a pair of sequences from two North American grosbeaks (*Pheucticus ludovicianus* and *P. melanocephalus,* Klicka and Zink 1997), assuming two different ancestral population sizes. Dashed line is for an ancestral Θ of 0.01, solid line, of 0.000001. (C) Analysis similar to (B) but for two North American jays (*Cyanocitta cristata* and *C. stelleri*).

$$L(\gamma) = \int_{\Theta=0}^{\Theta=\infty} f(\Theta) \int_{t=0}^{t=\infty} P(t, \Theta)P(D, \gamma + t)\, dt\, d\Theta. \quad (10)$$

The function $f(\Theta)$ can be any reasonable prior distribution of Θ, for example, an exponential distribution with a mean taken from contemporary populations or simply a rectangular distribution, where every Θ falls in some range, say, between zero and 10, with equal probability. In this way, by maximizing equation (10), the most likely value of γ can be es-

TABLE 2. Summary of means and percentiles of population divergence estimates in Figures 5–7. L is the number of loci used for the estimates. The parameters D, Θ, γ, and τ are the gene divergence, 4 × effective population size × mutation rate per site and per generation (μ), the population divergence time × μ. Lower confidence interval (CI) and upper CI enclose the 95% CI of the maximum-likelihood estimate (MLE). For all rows except the last row, to estimate population divergence times in millions of generations using the rule of 2%/million years for animal mtDNA, simply move the decimal place in the estimates of γ two places to the right.

Species pair	L	D/2	Θ	Estimate of γ		
				Lower 95% CI	MLE	Upper 95% CI
Fig. 5A, curves from right to left						
simulated	1	—	0.01	0.00000	0.00166	0.00525
simulated	1	—	0.01	0.00000	0.00245	0.00646
simulated	1	—	0.01	0.00001	0.00646	0.01202
simulated	1	—	0.01	0.08318	0.10233	0.11749
Fig. 5B						
Pheucticus ludovidicanus-Pheucticus melanocephalus	1	0.022	0.000001	0.0126	0.0158	0.0200
P. ludovidicanus-P. melanocephalus (dashed line)	1	0.022	0.01	0.0063	0.0158	0.0200
Fig. 5C						
Cyanositta cristata-Cyanositta stelleri	1	0.054	0.000001	0.0457	0.0550	0.0631
C. cristata-C. stelleri (dashed line)	1	0.054	0.01	0.0380	0.0525	0.0631
Fig. 6A						
curve 1, *C. cristata-C. stelleri*	1	0.054	0–10[1]	0.0000	0.0525	0.0661
curve 2, *C. cristata-C. stelleri*	1	0.054	0–∞[2]	0.0348	0.0575	0.0692
curve 3, *C. cristata-C. stelleri*	1	0.054	0–∞[3]	0.0191	0.0575	0.0692
curve 4, *C. cristata-C. stelleri*	1	0.054	0.00073-0.00623[4]	0.0501	0.0562	0.0708
Fig. 6B (right to left)						
curve 4, *C. cristata-C. stelleri*	1	0.054	0.00073-0.00623[4]	0.0501	0.0562	0.0708
curve 2, *T. bendirei-T. cinereum*	1	0.008	0.00073-0.00623[4]	0.0035	0.0089	0.0126
curve 1, *Oporornis philadephia-Oporornis tolmiei*	1	0.011	0.00073-0.00623[4]	0.0025	0.0079	0.0126
Fig. 7A (curves from left to right)						
simulated	20	—	0.01	0.0001	0.0001	0.0055
simulated	20	—	0.01	0.00019	0.00135	0.00240
simulated	20	—	0.01	0.00794	0.00977	0.01148
simulated	20	—	0.01	0.09772	0.10233	0.10471
Fig. 7B						
Homo-Pan	15	—	(0.00422 0.0281)[5]	0	0.00091	0.00473

[1] Rectangular prior for Θ (range 1–10).
[2] Exponential prior for an average Θ of 0.02.
[3] Prior is $1/\Theta$.
[4] Rectangular prior for Θ using the mean ± 1 standard deviation of values of $D_{max}/2$ in Moore (1995).
[5] Θ was jointly estimated with γ, using a gamma-distributed mutation rate with shape parameter $\alpha = 5$; 95% CIs for Θ are given.

timated while taking into account uncertainty of ancestral Θ. Figure 6 shows equation (10) applied to cytochrome *b* data from North American jays (*Cyanocitta*; Klicka and Zink 1997). The particular prior distribution of Θ influences the 95% confidence limits of the maximum-likelihood estimate much more than it does the maximum-likelihood estimate itself (Fig. 6A, Table 2). Instead of using a completely arbitrary distribution of Θ with large or infinite range (curves 1–3), we can assume that the distribution and range of Θ in contemporary species tells us something about these values in ancestral species. Using a distribution again derived from Moore's (1995) survey considerably reduces the confidence limits around the maximum (Fig. 6A, curve 4). Figure 6A and Table 2 values for this figure also show that the finite-sites approach can yield estimates of γ that are slightly larger than the observed gene divergence, as in Figure 3 (Table 2). In addition, Figure 6B shows that the rank order of gene divergences for a single locus does not necessarily predict the rank order of the population divergences (see also Table 2): whereas the *D*-value for *Oporornis philadelphia-O. tolmiei* is greater than that for *Toxostoma bendirei-T. cinereum*, the maximum-likelihood estimate of γ for the *Toxostoma* clade is greater. Such cases are likely the result of the particular nucleotide configurations in the two species pairs; even so, the confidence limits on γ for both clades overlap broadly.

Estimates Using Multiple (Unlinked) Loci

If we assume that μ is the same for all loci, a likelihood framework permits simple multiplication of the likelihoods of each locus as determined by equation (10), assuming loci are unlinked to each other:

$$L(\gamma) = \prod_{l=1}^{loci} L_l(\gamma). \quad (11)$$

The value of γ that maximizes the product in equation (11) is the maximum-likelihood estimate of τ over all loci, although it will not necessarily be the value that maximizes the likelihood of τ for each single locus. Figure 7a shows this method applied to simulated multilocus DNA data from

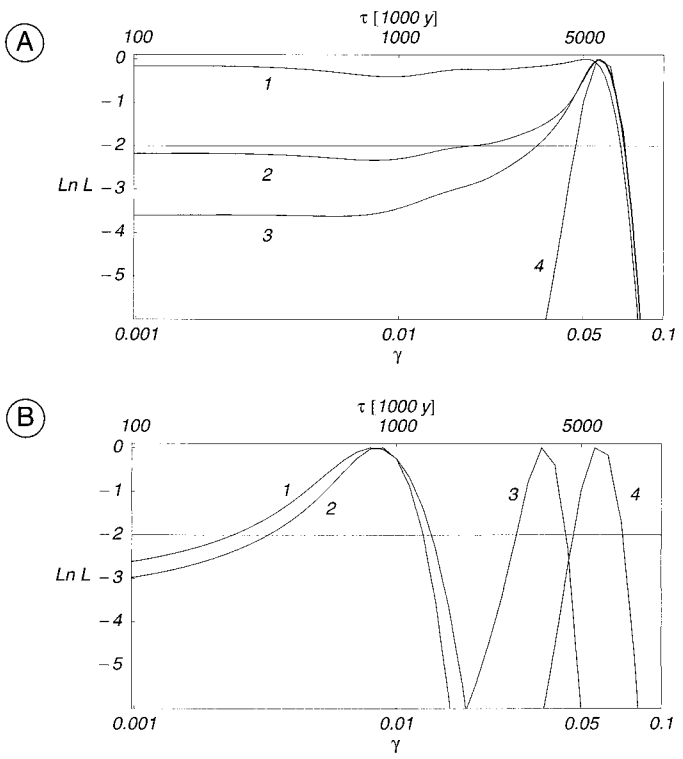

FIG. 6. Bayesian inference of population divergence time from single-locus data. (A) Curves with various prior distributions of Θ: (1) rectangular prior with Θ in the range between 0 and 10; (2) exponential distribution with a mean Θ of 0.02; (3) with an inverse prior for Θ and; (4) rectangular prior using boundaries of Θ specified by contemporary avian population sizes (Moore 1995). (B) Analysis of population divergence times of avian species pairs from Klicka and Zink (1997). All curves use a rectangular prior with boundaries derived from Moore (1995). Species pairs are at follows: (1) *Toxostoma crissale* and *T. lecontei;* (2) *Oporornis philadelphia* and *O. tolmiei;* (3) *Toxostoma cinerea* and *T. bendirei;* and (4) *Cyanocitta cristata* and *C. stelleri.*

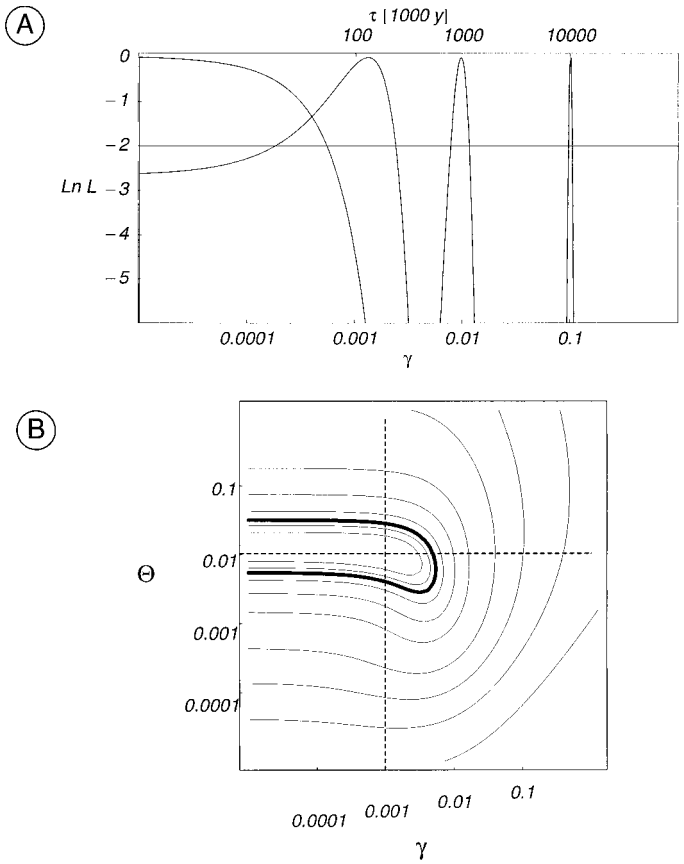

FIG. 7. Multilocus estimates of the number of substitutions per site since population divergence, γ. (A) Simulations showing multilocus estimates (20 loci) of γ when ancestral population size Θ = 0.01 is known. Compare widths of curves with those in Figure 5A, the single-locus case. (B) Multilocus analysis of divergence times of humans and chimpanzees. 15 pairs of DNA sequences, one from *Homo sapiens* and one from *Pan troglodytes,* were aligned (details of this dataset are available on request). We confined analysis to the 2791 third positions of these loci, an action that might affect the value of μ used to convert γ to an estimate of τ (see also Table 2). The intersection of the dashed lines indicate the maximum-likelihood estimate of Θ and γ. The contour lines are at log-likelihood values of −1, −2, −3, −4, −5, −10, −20, −50, −100, and −200. The 95% confidence region is enclosed by the −3 contour line in bold.

two populations that diverged at a known time τ in the past (Table 2). As in the single-locus case, the precision with which γ can be estimated depends primarily on the magnitude of Θ relative to γ. In Figure 7, which involves 20 independent loci, when the ratio γ/Θ is approximately 0.1 or greater, the 95% confidence interval of γ does not include zero. We have explored the conditions under which the 95% confidence interval on γ does not include zero and have found that this conditions holds, approximately, when $l\gamma/\Theta > 1$ (where l is the number of loci). This suggests that a fruitful way of determining effort in estimation of γ would be to use a rough estimates of γ and Θ from preliminary data to guess the approximate number of loci that will satisfy $l\gamma/\Theta > 1$.

It is likely that μ varies across loci, even if μ represents the silent substitution rate of different loci (Wolfe et al. 1989). Like Yang (1997), we investigated the estimate of γ as in equations (10) and (11), this time assuming rate variation among loci following a gamma distribution:

$$L(\gamma) = \prod_{l=1}^{L} \int_{r=0}^{r=\infty} \frac{e^{-r\alpha} r^{\alpha-1}}{\Gamma(\alpha)\alpha^{-\alpha}} L_l(\gamma r)\, dr, \qquad (12)$$

where r is a gamma-distributed modifier of the average mutation rate μ and the variance in r is $1/\alpha$. Using our finite-sites method, we found that for 15 human loci (102–506 sites), assuming $\alpha = 5$, that the likelihood contour plots for divergence time γ included 0 within the 95% confidence limits (Table 2; Fig. 7B). This is the same as saying that, from a strictly genetic perspective, we do not know whether the differences observed between species were drawn from a single population or from two species that had diverged long ago. Yang's (1997) estimate for γ from this same dataset (0.00352 ± 0.00357 SD) also included zero. This sobering result stems from the small number of differences in the loci between humans and chimps (zero to six substitutions per locus) relative to variation among loci in the number of differences, which flows directly from the large ancestral population size. Whereas the confidence limits of θ using Yang's method also included zero for this dataset (0.00354 ±

0.00797), the finite-sites estimate did not (Table 2), a result likely arising from the use of more information from the data in the latter method. With $\alpha = 10$ using the finite sites method, the 95% confidence limit still included zero and the maximum-likelihood surface was very similar to $\alpha = 5$. Realistically, however, estimates of γ will vary more so from these results only as values of α fall well below one, that is, when much of the variation in D among loci can be attributed to among-locus rate variation, rather than to large ancestral population size.

Likelihood Ratio Test of Simultaneous Diversification of Multiple Species Pairs

The methods of Yang (1997) and the methods outlined here suggest several ways of statistically comparing estimates of population divergence between independent pairs of species, a central goal of comparative phylogeography. Yang's method provides standard errors on τ based on the ''curvature'' method. Another method, one that we use here, considers as valid all parameter values within two log-likelihood values of the maximum (Figs. 5–7). Our finite-sites likelihood framework permits the calculation of approximate confidence intervals based on the likelihood ratio test statistic (Casella and Berger 1990). The likelihood ratio of two alternative values of γ is distributed asymptotically as χ^2 with one degree of freedom and we reject the null hypothesis, that there is no difference between the maximum-likelihood estimate (γ_0) and an alternative (γ_1) at the significance level of 0.05 when χ^2 is larger than two:

$$\chi^2_{df=1} \approx -2 \ln\left(\frac{L(\gamma_0)}{L(\gamma_1)}\right). \quad (13)$$

For example, in the single-locus mtDNA data from birds (Fig. 6B), we can reject the hypothesis of simultaneous diversification when the distribution of ancestral Θ is assumed to follow the empirical distribution of Θ found in contemporary avian species (Moore 1995). If, however, Θ is assumed to be very large (Fig. 5A) or to follow a very wide distribution (Fig. 5B), it is difficult to reject the hypothesis of simultaneous diversification. Although the likelihood-ratio test is only approximate, it nonetheless is a useful tool for quantitatively testing ideas on comparative phylogeography, cospeciation, and vicariance biogeography. In the absence of knowledge on ancestral Θ, we suspect it may be difficult to reject simultaneity of many diversification events that on casual inspection of embedded single-locus divergences seem to differ in age.

Prospectus

Bermingham and Moritz (1998) recently reviewed important concepts in the burgeoning field of comparative phylogeography and recommended, among other things, that a more complete integration of coalescent theory and phylogeography would benefit the field. Our review of the distinctions between gene and population divergence and application of single and multilocus methods for estimating population divergence time represent a step toward such integration. In many phylogeographic studies, particularly those for which a single locus has achieved reciprocal monophyly for sister species, the degree to which the gene divergence overestimates the population divergence is unknown. Although the extent of overestimation may be relatively small in many cases, the confidence limits on the population divergence time will exceed those for the number of DNA substitutions (D) along lineages because of drift in the ancestral population. Thus, calculation of confidence limits on D with methods that ignore divergence in the ancestral species will in general underestimate the actual confidence limits on τ, the population divergence, in some cases severely, and generally relegates such methods to deeper divergences for which ancestral polymorphism is less of a problem. Throughout we have purposefully played devil's advocate regarding estimating species divergence times with single loci that have achieved reciprocal monophyly in two descendent species, emphasizing that making any estimate of population divergence time from such data is tantamount to making claims about population size of the ancestral species. The ways in which various taxonomic groups are thought to speciate, such as via founder events or via vicariance of large ancestral populations, can undoubtedly provide useful qualitative guidance to the population sizes of the ancestral species (Chesser and Zink 1994), but ultimately is guesswork, because upon achieving reciprocal monophyly, information on the size of the ancestral species is lost. This difficulty in estimating divergence times with single loci is but another argument for the increased need for multilocus data for estimating divergence times when the distinction between gene divergence (T) and population divergence (τ) is large.

It is ironic that the overestimation problem stems in part from advances in DNA technology that allow us to examine gene genealogies directly. The overestimation problem does not apply to protein electrophoretic data or other data for which allele frequencies alone are used to estimate divergence times (Nei 1987). This is because, at the time of speciation, allele frequencies in the two descendant populations are identical and divergence time and divergence in allele frequencies occur over identical time periods. The overestimation problem does apply, however, to loci such as microsatellite loci, distances for which are usually a function of both allele frequencies and divergence in the loci themselves, that is, in repeat number.

The two-tipped trees we and others have employed are useful for the problems studied here, particularly because many phylogeographic studies sample only a single allele per species, especially when those studies are conducted using a multispecies phylogeny. Sampling multiple alleles within each descendant species slightly improves the ability to infer divergence times between species, but only in indirect ways. Whereas such sampling will help confirm that allelic lineages have in fact reached reciprocal monophyly, it does not provide extra information on the estimate of coalescence time in the ancestor, because the population sizes of descendent species need not be related to those of their ancestors. Sampling within species diversity can be used to provide a bound on the most recent possible divergence time of a species pair, because the coalescent depths of descendant species will usually be the result of independent, postspeciation evolution in those descendent species. Sampling within species diver-

gences will also help gauge the uncertainty in the estimate of interspecific gene divergence D (Fig. 1), because different alleles will tend to give slightly different estimates of D even in the presence of a molecular clock.

Because different loci can evolve at different rates, multilocus data raises the problem of among-locus rate variation as a source of among-locus variation in divergence between species that is additional to the stochastic variance expected in the ancestral coalescent. The simultaneous estimation of α (the among-locus rate variation parameter), γ, and ancestral Θ is a data-hungry process even with multiple loci, and although ideally α can be estimated independently, in reality this too is difficult (Takahata and Satta 1997; Yang 1997). Future studies could gain rough estimates of γ and Θ from a few loci and use the criterion of $l\gamma/\Theta > 1$ to predict the number of additional loci that will be needed to resolve τ adequately.

The analysis of ancestral population sizes should be extended to trees containing more than two species (cf. Takahata et al. 1995; Wang et al. 1997) to test for temporal congruence over the history of replicate multispecies assemblages. As acknowledged by Takahata and Satta (1995) and addressed explicitly by Wang et al. (1997), sorting of ancestral lineages makes it possible that the trees of different genes may not be congruent with one another or with the species tree (e.g., Hilton and Hey 1997); but this may not be a complicating factor for many groups. Estimating confidence limits on population divergence times for entire phylogenies, as has been done for gene divergence times (Nei et al. 1985), will yield increased statistical rigor over simple pairwise comparisons of species (Takahata and Satta 1997; Huelsenbeck et al. 2000) and will permit more robust tests of cospeciation in a variety of contexts. Thus, a move away from single-locus studies to multilocus, multispecies studies should increasingly dominate comparative phylogeography.

Acknowledgments

We thank R. Zink, J. Klicka, and Y. Satta for sending DNA sequence data. For helpful discussion and encouragement, we thank R. Zink, J. Wakeley, M. Nachman, D. Olmstead, Z. Yang, A. Larson, and particularly J. Felsenstein. We thank J. Avise, M. Hare, R. Nielsen, T. Price, T. Uzzell, J. Wakeley, R. Zink, and several anonymous reviewers for comments on the manuscript. Finite sites programs that calculate γ and θ for single-locus and multilocus data under reciprocal monophyly as described in this paper are available from PB (beerli@genetics.washington.edu). This work was supported by NSF grants to SVE (DEB 9707458 and DEB 9815800) and NSF (BIR 9527687) and NIH (GM 51929-02 and GM 01989-02) grants to J. Felsenstein (PB).

Literature Cited

Avise, J. C. 1992. Molecular population structure and the biogeographic history of a regional fauna: a case history with lessons for conservation biology. Oikos 63:62–76.

———. 1998. The history and purview of phylogeography: a personal reflection. Mol. Ecol. 7:371–379.

———. 2000. Phylogeography: the history and formation of species. Harvard Univ. Press, Cambridge, MA.

Avise, J. C., and D. Walker. 1998. Pleistocene phylogeographic effects on avian populations and the speciation process. Proc. R. Soc. Lond. B 265:457–463.

Avise, J. C., and K. Wollenberg. 1997. Phylogenetics and the origin of species. Proc. Natl. Acad. Sci. 94:7748–7755.

Avise, J. C., J. E. Neigel, and J. Arnold. 1984. Demographic influences on mitochondrial DNA lineage survivorship in animal population. J. Mol. Evol. 20:99–105.

Avise, J. C., J. Arnold, R. M. Ball, E. Bermingham, T. Lamb, J. E. Neigel, C. A. Reeb, and N. C. Saunders. 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. Annu. Rev. Ecol. Syst. 18: 489–522.

Avise, J. C., B. W. Bowen, T. Lamb, A. B. Meylan and E. Bermingham. 1992. Mitochondrial DNA evolution at a turtle's pace evidence for low genetic variability and reduced microevolutionary rate in the Testudines. Mol. Biol. Evol. 9:457–473.

Ayala, F. J. 1997. Vagaries of the molecular clock. Proc. Natl. Acad. Sci. USA 94:7776–7783.

Beerli, P., H. Hotz, and T. Uzzell. 1996. Geologically dated sea barriers calibrate a protein clock for Aegean water frogs. Evolution 50:1676–1687.

Bermingham, E., and A. P. Martin. 1998. Comparative mtDNA phylogeography of neotropical freshwater fishes: testing shared history to infer the evolutionary landscape of lower Central America. Mol. Ecol. 7:499–517.

Bermingham, E., and C. Moritz. 1998. Comparative phylogeography: concepts and applications. Mol. Ecol. 7:367–369.

Bernatchez, L., and C. C. Wilson. 1998. Comparative phylogeography of nearctic and palearctic fishes. Mol. Ecol. 7:431–452.

Brower, A. V. Z., R. DeSalle, and A. Vogler. 1996. Gene trees, species trees and systematics: a cladistic perspective. Annu. Rev. Ecol. Syst. 27:423–450.

Busack, S. 1986. Biogeographic analysis of the herpetofauna separated by the Strait of Gibraltar. Nat. Geo. Res. 2:17–36.

Casella, G., and R. L. Berger. 1990. Statistical inference. Duxbury Press, Belmont, CA.

Cavalli-Sforza, L. L. 1969. Human diversity. Pp. 405–416 *in* Proc. 12th int. cong. genet. Vol. 2.

Chesser, R. T., and R. M. Zink. 1994. Modes of speciation in birds: a test of Lynch's method. Evolution 48:490–497.

Cracraft, J. 1986. Origin and evolution of continental biotas: speciation and historical congruence within the Australian avifauna. Evolution 40:977–996.

———. 1988. Deep-history biogeography: retrieving the historical pattern of evolving continental biotas. Syst. Zool. 37:221–236.

Cunningham, C. W. 1997. Can three incongruence tests predict when data should be combined? Mol. Biol. Evol. 14:733–740.

Cunningham, C. W., and T. M. Collins. 1994. Developing model systems for molecular biogeography: vicariance and interchange in marine invertebrates. Pp. 405–433 *in* B. Schierwater, B. Streit, G. P. Wagner, and R. DeSalle, ed. Molecular ecology and evolution: approaches and applications Birkhäuser, Boston.

da Silva, M. N. F., and J. L. Patton. 1998. Molecular phylogeography and the evolution and conservation of Amazonian mammals. Mol. Ecol. 7:475–486.

Donnelly, P., and S. Tavaré. 1995. Coalescents and genealogical structure under neutrality. Annu. Rev. Genet. 29:401–421.

Dorit, R. L., H. Akashi, and W. Gilbert. 1995. Absence of polymorphism at the ZFY locus on the human Y chromosome. Science 268:1183–1185.

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

———. 1997. Relevance of microevolutionary processes for higher level molecular systematics. Pp. 251–278 in D. P. Mindell, ed. Avian molecular systematics and evolution. Academic Press, New York.

Felsenstein, J. 1988. Phylogenies from molecular sequences: inference and reliability. Annu. Rev. Genet. 22:521–565.

Fleischer, R. C., C. E. McIntosh, and C. L. Tarr. 1998. Evolution on a volcanic conveyor belt: using phylogeographic reconstruc-

tions and K-Ar-based ages of the Hawaiian Islands to estimate molecular evolutionary rates. Mol. Ecol. 7:533–545.
Friesen, V. L., B. C. Congdon, H. E. Walsh, and T. P. Birt. 1997. Intron variation in marbled murrelets detected using analyses of single strand conformational polymorphisms. Mol. Ecol. 6: 1047–1058.
Gillespie, J. H. 1991. The causes of molecular evolution. Oxford Univ. Press, Oxford, U.K.
Gillespie, J. H., and C. H. Langley. 1979. Are evolutionary rates really variable? J. Mol. Evol. 13:27–34.
Griffiths, R. C., and S. Tavaré. 1994. Sampling theory for neutral alleles in a varying environment. Phil. Trans. R. Soc. B 344: 403–410.
Hafner, M. S., and R. D. Page. 1995. Molecular phylogenies and host-parasite cospeciation: gophers and lice as a model system. Phil. Trans. R. Soc. B 349:77–83.
Hafner, M. S., P. D. Sudman, F. X. Villablanca, T. A. Spradling, J. W. Demastes and S. A. Nadler. 1994. Disparate rates of molecular evolution in cospeciating hosts and parasites. Science 265:1087–1090.
Hammer, M. F. 1995. A recent common ancestry for human Y chromosomes. Nature 378:376–378.
Hillis, D. M., B. K. Mable and C. Moritz. 1996. Applications of molecular systematics: the state of the field and a look to the future. Pp. 515–543. *in* D. M. Hillis, C. Moritz, and B. K. Mable, eds. Molecular Systematics Sinaur Associates, Inc., Sunderland, MA.
Hilton, H., and J. Hey. 1997. A multilocus view of speciation in the *Drosophila virilis* species group reveals complex histories and taxonomic conflicts. Genet. Res. 70:185–194.
Hoelzer, G. A. 1997. Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees revisited. Evolution 51:622–626.
Hudson, R. R. 1990. Gene genealogies and the coalescent process. Oxf. Surv. Evol. Biol. 7:1–44.
Huelsenbeck, J. P., B. Rannala, and B. Larget. 2000. A Bayesian framework for the analysis of cospeciation. Evolution 54: 352–364.
Johns, G. C., and J. C. Avise. 1998. A comparative summary of genetic distances in the vertebrates from the mitochondrial cytochrome *b* gene. Mol. Biol. Evol. 15:1481–1490.
Kimura, M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16:111–120.
Kingman, J. 1982a. The coalescent. Stochast. Proc. Appl. 13: 235–248.
———. 1982b. On the genealogy of large populations. Pp. 27–43. *in* J. Gani and E. Hannan, eds. Essays in statistical science Applied Probability Trust, London.
Kishino, H., and M. Hasegawa. 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order of the Hominoidea. J. Mol. Evol. 29:170–179.
Klicka, J., and R. M. Zink. 1997. The importance of recent ice ages in speciation: a failed paradigm. Science 277:1666–1669.
———. 1999. Pleistocene effects on North American songbird evolution. Proc. R. Soc. Lond. B. 266:695–700.
Knowlton, N., and L. A. Weigt. 1998. New dates and new rates for divergence across the Isthmus of Panama. Proc. R. Soc. Lond. B 265:2257–2263.
Knowlton, N., L. A. Weigt, L. A. Sol'orzano, D. K. Mills, and E. Bermingham. 1993. Divergence in proteins, mitochondrial DNA, and reproductive compatibility across the Isthmus of Panama. Science. 260:1629–1632.
Kuhner, M. K., J. Yamato, and J. Felsenstein. 1998. Maximum likelihood estimation of population growth rates based on the coalescent. Genetics 149:429–434.
Langley, C. H., and W. M. Fitch. 1974. An examination of the constancy of the rate of molecular evolution. J. Mol. Evol. 3: 161–177.
Li, W. H. 1993. So, what about the molecular clock hypothesis? Curr. Op. Gen. Dev. 3:896–901.
Lynch, M., and P. E. Jarrell. 1993. A method for calibrating molecular clocks and its application to animal mitochondrial DNA. Mol. Biol. Evol. 135:1197–1208.
Maddison, W. P. 1997. Gene trees in species trees. Syst. Biol. 46: 523–536.
Jukes, T. H., and C. R. Cantor. 1969. Evolution of protein molecules. Pp. 21–123 *in* H. N. Munro, ed. Mammalian protein metabolism. Academic Press, New York.
Moore, W. S. 1995. Inferring phylogenies from mtDNA variation: mitochondrial gene trees vs. nuclear gene trees. Evolution 49: 718–726.
———. 1997. Mitochondrial-gene trees versus nuclear-gene trees, a reply to Hoelzer. Evolution 51:627–629.
Nei, M. 1972. Genetic distance between populations. Am. Nat. 106: 283–292.
———. 1987. Molecular evolutionary genetics. Columbia Univ. Press, New York.
Nei, M., and L. Jin. 1989. Variances of the average numbers of nucleotide substitutions within and between populations. Mol. Biol. Evol. 6:290–300.
Nei, M., and W.-H. Li. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc. Natl. Acad. Sci. (USA) 76:5269–5273.
Nei, M., and M. Takahata. 1993. Effective population size, genetic diversity and coalescence time in subdivided populations. J. Mol. Evol. 37:240–244.
Nei, M., J. C. Stephens, and N. Saitou. 1985. Methods for computing the standard errors of branch points in an evolutionary tree and their application to molecular data from humans and apes. Mol. Biol. Evol. 2:66–85.
Neigel, J. E., and J. C. Avise. 1986. Phylogenetic relationships of mitochondrial DNA under various demographic models of speciation. Pp. 515–534 *in* S. Karlin and E. Nevo eds. Evolutionary processes and theory. Academic Press, New York.
Nielsen, R. 1998. Maximum likelihood estimation of population divergence times and population phylogenies under the infinite sites model. Theor. Popul. Biol. 53:143–151.
Nielsen, R., J. L. Mountain, J. P. Huelsenbeck, and M. Slatkin. 1998. Maximum-likelihood estimation of population divergence times and population phylogeny in models without mutation. Evolution 52:669–677.
Nielsen, R., and M. Slatkin. 2000. Likelihood analysis of ongoing gene flow and historical association. Evolution 54:44–50.
Nordborg, M., 1997. Structured coalecent processes on different time scales. Genetics 146:1501–1514.
Page, R. D. M. 1990. Temporal congruence and cladistic analysis of biogeography and cospeciation. Syst. Zool. 39:205–226.
———. 1991. Random dendrograms and null hypotheses in cladistic biogeography. Syst. Zool. 40:54–62.
———. 1994. Maps between trees and cladistic analysis of historical associations among genes, organisms, and areas. Syst. Biol. 43:58–77.
———. 1996. Temporal congruence revisited: Comparison of mitochondrial DNA sequence divergence in cospeciating pocket gophers and their chewing lice. Syst. Biol. 45:151–167.
Palumbi, S., and F. Cipriano. 1998. Species identification using genetic tools: the value of nuclear and mitochondrial gene sequences in whale conservation. J. Hered. 89:459–464.
Platnick, N. I., and G. J. Nelson. 1978. A method for analysis for historical biogeography. Syst. Zool. 27:1–16.
Riddle, B. R. 1996. The molecular phylogeographic bridge between deep and shallow history in continental biotas. Trends Ecol. Evol. 11:207–211.
———. 1998. The historical assembly of continental biotas: late Quaternary range-shifting, areas of endemism, and biogeographic structure in the North American mammal fauna. Ecography 21:437–446.
Sanderson, M. J. 1997. A nonparametric approach to estimating divergence times in the absence of rate constancy. Mol. Biol. Evol. 14:1218–1231.
———. 1998. Estimating rate and time in molecular phylogenies: beyond the molecular clock? Pp. 242–264 *in* P. S. Soltis, D. E. Soltis, and J. J. Doyle, eds. Molecular systematics of plants. Vol. II. DNA sequencing Kluwer, Boston.

Schneider, C. J., M. Cunningham, C. Moritz. 1998. Comparative phylogeography and the history of endemic vertebrates in the wet tropics rainforests of Australia. Mol. Ecol. 7:487–498.

Semino, O., G. Passarino, P. J. Oefner, A. A. Lin, S. Arbuzova, L. E. Beckman, G. De Benedictis, P. Francalacci, A. Kouvatsi, S. Limborska, M. Marcikiae, A. Mika, B. Mika, D. Primorac, A. S. Santachiara–Benerecetti, L. L. Cavalli–Sforza, and P. A. Underhill. 2000. The genetic legacy of paleolithic *Homo sapiens sapiens* in extant Europeans: a Y chromosome perspective. Science 290:1155–1159.

Slade, R. W., C. Moritz, A. Heideman, and P. T. Hale. 1993. Rapid assessment of single-copy nuclear DNA variation in diverse species. Mol. Ecol. 2:359–373.

Slade, R. W., C. Moritz, A. R. Hoelzel, and H. R. Burton. 1998. Molecular population genetics of the southern elephant seal Mirounga leonina. Genetics 149:1945–1957.

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

Slatkin, M., and R. R. Hudson. 1991. Pairwise comparison of mitochondrial DNA sequences in stable and exponentially growing populations. Genetics 129:555–562.

Slatkin, M., and W. P. Maddison. 1989. A cladistic measure of gene flow inferred from the phylogenies of alleles. Genetics 12: 603–613.

Steel, M. A., A. C. Cooper, and D. Penny. 1996. Confidence intervals for the divergence time of two clades. Syst. Biol. 45: 127–134.

Takahata, N. 1986. An attempt to estimate the effective size of the ancestral species common to two extant species from which homologous genes are sequenced. Genet. Res. 48:187–190.

———. 1991. Genealogy of neutral genes and spreading of selected mutations in a geographically structured population. Genetics 129:585–595.

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

Takahata, N., and Y. Satta. 1997. Evolution of the primate lineage leading to modern humans: phylogenetic and demographic inferences from DNA sequences. Proc. Natl. Acad. Sci. (USA) 94: 4811–4815.

Takahata, N., and F. Tajima. 1991. Sampling errors in phylogeny. Mol. Biol. Evol. 8:494–502.

Takahata, N., Y. Satta, and J. Klein. 1995. Divergence time and population size in the lineage leading to modern humans. Theor. Popul. Biol. 48:198–221.

Tamura, K., and M. Nei. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol. Biol. Evol. 10:512–526.

Tavaré, S., D. J. Balding, R. C. Griffiths, and P. Donnelly. 1997. Inferring coalescence times from DNA sequence data. Genetics 145:505–518.

Templeton, A. R. 1983. Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to the evolution of humans and the apes. Evolution 37:221–244.

———. 1993. The ''Eve'' hypothesis: a genetic critique and reanalysis. Am. Anthropol. 95:51–72.

Vigilant, L., M. Stoneking, H. Harpending, K. Hawkes, and A. C. Wilson. 1991. African populations and the evolution of mitochondrial DNA. Science 253:1503–1507.

Vila, C., P. Savolainen, J. E. Maldonado, I. R. Amorim, J. E. Rice, R. L. Honeycutt, K. A. Crandall, J. Lundeberg, and R. K. Wayne. 1997. Multiple and ancient origins of the domestic dog. Science 276:1687–1689.

Wakeley, J., and J. Hey. 1997. Estimating ancestral population parameters. Genetics 145:847–855.

Walker, D., and J. C. Avise. 1998. Principles of phylogeography as illustrated by freshwater and terrestrial turtles in the southeastern United States. Annu. Rev. Ecol. Syst. 29:23–58.

Wang, R. L., J. Wakeley, and J. Hey. 1997. Gene flow and natural selection in the origin of *Drosophila pseudoobscura* and close relatives. Genetics 147:1091–1106.

Wilson, A. C., R. L. Cann, S. M. Carr, M. George, U. B. Gyllensten, K. M. Helm-Bychowski, R. G. Higuchi, S. R. Palumbi, E. M. Prager, R. D. Sage, and M. Stoneking. 1985. Mitochondrial DNA and two perspectives on evolutionary genetics. Biol. J. Linn. Soc. 26:375–400.

Wolfe, K. H., P. M. Sharp, and W. H. Li. 1989. Mutation rates differ among regions of the mammalian genome. Nature 337: 283–285.

Wollenberg, K., and J. C. Avise. 1998. Sampling properties of genealogical pathways underlying population pedigrees. Evolution 52:957–966.

Wright, S. 1951. The genetical structure of populations. Ann. Eugen. 15:323–354.

Yang, Z. 1997. On the estimation of ancestral population sizes of modern humans. Genet. Res. 69:111–116.

Zink, R. M. 1997. Phylogeographic studies of North American birds. Pp. 301–324 *in* D. P. Mindell ed. Avian molecular systematics and evolution. Academic Press, New York.

Zink, R. M., and J. B. Slowinski. 1995. Evidence from molecular systematics for decreased avian diversification in the Pleistocene epoch. Proc. Natl. Acad. Sci. (USA) 92:5832–5835.

Corresponding Editor: J. Neigel