

Spatial and Temporal Aspects of Populations Revealed by Mitochondrial DNA

FRANCIS X. VILLABLANCA

1. Introduction

The evolutionary analysis of DNA sequences bridges phylogenetics and population genetics. Ancient DNA (aDNA) allows the study of extinct genotypes, populations, and species, as well as dichronic comparisons of extant populations and species. Thus aDNA forges an empirical link between history and two inherently historical fields of research. Fortunately, the conceptual frameworks of phylogenetics and population genetics can easily be extended to encompass advances being made in the study of aDNA.

In order to demonstrate and discuss the evolutionary analysis of ancient DNA sequences, we will be investigating an expanded version of the study originally presented by Thomas et al. (1990). Thomas and coworkers sequenced 225 base pairs of the mitochondrial control region (d-loop) from 43 museum specimens and 63 modern specimens of the kangaroo rat *Dipodomys panamintinus* (Rodentia, Heteromyidae). In total, the 106 specimens represent two temporal samples from each of three different geographical populations. Modern samples were taken in 1988, 53 to 77 years after the museum specimens were collected. Expanding this study, we will consider a total of 421 base pairs from 14 modern and 5 museum specimens. These 19 specimens represent unique mitochondrial genotypes found by Thomas et al. and will be used to better resolve the evolutionary relationships among genotypes and quantify gene flow between populations.

To date, *D. panamintinus* is the only species studied for which population size samples of aDNA have been applied toward phylogenetics and population genetics. From this perspective it is exemplary. Yet it should be noted that the results are still tentative and await sequencing of longer fragments and a better understanding of intraspecific geographic variation in *D. panamintinus*. In the sections below we consider methods for reconstructing hypothesized evolutionary relationships among genotypes, populations, and species. The hypotheses include estimates of gene flow between populations and of time since divergence between populations. We also quantify genetic variability at the nucleotide level, and measure its change over time. The discussion of

each of these approaches will be followed immediately by an application to the data from kangaroo rats. Before addressing phylogenetics and population genetics, let us consider the genome under investigation.

2. Mitochondrial DNA and Gene Trees

The mitochondrial genome has been used extensively in evolutionary studies of animals. Probably the single most important reason for its widespread use is its rate and mode of evolution (Aquadro et al. 1984; Wilson et al. 1985; Moritz et al. 1987; Miyamoto and Boyle 1989). Population level studies exploit the genome's rapid evolutionary rate (Brown et al. 1979, 1982); substitutions accumulate rapidly in third codon positions of coding genes and in the control region (Kocher and Wilson 1991). Interspecific and higher level relationships can be estimated using the slower amino acid replacement changes (Brown 1985; Moritz et al. 1987; DeSalle et al. 1987). Protein coding genes may be particularly useful at bridging the intra- and interspecific levels. Unfortunately, plant mitochondrial DNA evolves faster in structure than in point mutations (Wolfe et al. 1987), and as a result evolutionary studies of plants make use of the chloroplast genome instead (Clegg et al. 1991; Golenberg 1991).

Unlike nuclear DNA, the mitochondrial genome is effectively haploid. Its inheritance is generally considered to be strictly maternal (Wilson et al. 1985; Gyllensten et al. 1991). This must be taken into account as we consider what questions can feasibly be asked when employing the mitochondrial genome (Awise et al. 1987; Birky et al. 1983, 1989; Birky 1991). Specifically, we can only draw conclusions with regard to maternal phylogenies and maternal gene flow. Mitochondrial genome patterns may be quite different from those shown by nuclear genes, particularly if the mating or dispersal/philopatry strategy differs strongly between the sexes. In the absence of sex-specific differences, however, the patterns should be like those from any randomly chosen nuclear gene.

Due to its haploid nature the mitochondrial genome does not undergo recombination (Wilson et al. 1985). This is especially important from the perspective of phylogenetics, which generally assumes that gene evolution involves mutation and branching but no reticulation or recombination of gene fragments. Thus, the mitochondrial genome is considered one completely linked locus. An implication of linkage is that selection acting on one gene affects other genes via hitchhiking (Kaplan et al. 1989; Kreitman 1991).

The mode of inheritance and evolution of mitochondrial DNA will both help and hinder an investigation of aDNA. A practical consideration of great importance is that the mitochondrial genes occur in much higher copy numbers than nuclear genes. As a result, it may be over 100 times easier to amplify a mitochondrial sequence. A high copy number also means that the consensus sequence produced by the polymerase chain reaction (PCR) is derived from a larger number of templates, thus avoiding errors due to

template-specific degradation (Pääbo and Wilson 1988). Copy number and retrieval advantages are offset by the fact that mitochondrial DNA is a single locus. Theoretical considerations indicate that several loci need to be studied in order to confidently recover phylogenies (Pamilo and Nei 1988). One locus produces a gene tree. It cannot be expected that gene trees will always be equivalent to species trees, except under specific demographic conditions (Takahata and Slatkin 1990). Thus gene phylogenies from ancient mitochondrial DNA may be easy and reliable, but without phylogenies from nuclear genes such gene trees are only tentative hypotheses of phylogenetic relationships (especially where males and females differ in life history). It is also necessary to distinguish between gene trees and population trees: our analysis will resolve the dichotomous relationships among alleles, but caution is needed in interpreting relationships among populations, which may be reticularly related due to the continued exchange of alleles.

3. Phylogenetics

Let us consider the construction of phylogenetic hypotheses and their applications (reviewed by Swofford and Olsen 1990; Felsenstein 1979, 1988; Wiley 1981). The basic questions are: how are the observed mitochondrial genotypes related genealogically, and how do these relationships compare to the geographic distribution of genotypes. Prior to interpreting our results, we need to evaluate the information content of the nucleotide sequences.

3.1 *Data and Samples*

D. panamintinus occurs primarily around the western periphery of the North American Mojave Desert. The three kangaroo rat populations studied by Thomas et al. (1990) are referred to as A, B, and C, and represent samples from three discontinuous portions of the species' range. These authors present separate phylogenies for the modern and museum genotypes, demonstrating that the results are essentially equivalent. For simplicity, and because most researchers will be combining the two types of data, we will consider a combination of modern and ancient genotypes. All unique modern genotypes will be included ($n = 14$), as will three types found only in museum samples of population B and two samples of a type which, based on 225 base pairs (Thomas et al. 1990), is identical and shared between museum populations B and C.

Extractions were accomplished in solutions of 5% Chelex (Walsh et al. 1991). PCR from liver tissue was performed in either 25- or 50- μ l reactions containing 67 mM Tris (pH 8.8), 16.8 mM AmSO_4 , 2 mM MgCl_2 , 01 M β -mercaptoethanol (final concentration), and 0.5 units of *Taq* polymerase. Primers were kangaroo rat specific (TAS; 5'-CTTGATGTTTCATCGTGC-ATT-3') and generalized vertebrate d-loop primers (H16498; Kocher and Wilson 1991) at 1 mM final concentration. Museum skins were amplified using 10 mM Tris (pH 8.3), 50 mM KCl, 1.5 mM MgCl_2 , 0.001% gelatin, 1.25

units of *Taq*, 2 µg/ml bovine serum albumin, and 40 rather than 30 PCR cycles. A second kangaroo rat specific primer (Dpd4; 5'-TACATTAGAC-ATTAATCCACATAA-3') which amplified a shorter fragment (270 nucleotides rather than 375) was used in combination with H16498 when amplifying from skins. Thermocycling consisted of denaturation at 92°C (60 sec), annealing at 55°C (45 sec), and extension at 72°C (60 sec). All products were sequenced directly in both directions. Figure 1 shows all nucleotide states for variable positions found across 421 base pairs.

Samples of aDNA pose two potentially significant problems to phylogenetic analysis. The first is contamination. Extreme care must be taken to ensure that the relationships being hypothesized are based on amplification of target DNA rather than of contaminants. This can be done simply by limiting the possibility of cross-contamination (Orrego 1990) and by sequence alignment

```

1111111111222222222222222233334
1424457789990011233333489916680
30044720293452503001256015601449
A1. AAGCGTAGCTGCCGACAATTTAACGGACTTGG
A2. ACACACCACCACCGATAACTTTAACGGACTTGA
A3. AAGCATGGCTGCCGACAATTTACGTAAGTGG
A4. AAGCATCACTGCCGATAAATTTAATGGACTTGA
A5. AAGCATCGCTGCCGACGATTTAACGGACATGN
A6. AAGCATAGCTGCCGACAATTTAGCAGTCTCGA
A7. AAGCATAACTGTCGACAATTTAACGGACCTCA
B1. AAGTATCACCACCAATAATTTAACGGACCTGA
B2. GAGTATCACCACCAATAATTTAACGGACCTGA
B3. AAGTATCACTACCAATAATTTAACGGATCTGA
B4. AAGTATCACCACCGATAATTTAACGGACATGA
B5. AAGTACCACCACCAATAATTTAACGGACATGA
B6. AAGTATCACCACCAGTAATTTAACGGACATGA
B7. AAGTATCACCACCAATGATTTAACGGACATGA
C1. AAGCATCATCGCTGATAGCCAAACGGACCTGA
C2. AAGCATCGTCCGCTGATAGCCAAACGGACCTGA
C3. AAGCATAACTGCCGACAACTTAACAGACTTGG
C4. ACGCATCATCGCTGATAATTTAACGGACCTGA
C5. AAGTATCACCACCAGTAATTTAACGGACATGA
D.A. ???TATAATTA CAACAACCAACAG_????
D.H. ???TATAACTACCGACAACTTAACAG_????

```

FIGURE 1. Character states of the 32 variable nucleotide positions from a 421-bp fragment of *Dipodomys panamintinus* control regions. Letters on the left designate the three populations sampled, while letter-number combinations identify unique mitochondrial genotypes within populations. Specimens examined are the first representatives of each genotype as presented in the appendix of Thomas et al. (1990). D.A. and D.H. are the outgroups, *D. agilis* and *D. heermanni*, respectively. Insertions/deletions are indicated by a dash, and undetermined sequences with a question mark. The first variable position is 3, and the last is 409. Position 1 corresponds to position 15307 of the mouse mitochondrial genome (Bibb et al. 1981).

and comparison of amplified DNA with that of known individuals (Thomas et al. 1989). Knowledge of modern levels of variation within a group of organisms will be helpful to assess whether aDNA sequences fall within an expected range of variation. The use of negative controls (extractions and amplifications into which no DNA is added) will reduce the chance that amplification of contaminants will go unnoticed (Pääbo 1990; Thomas et al. 1990). For some primer/template combinations, particularly if primer specificity is poor or if primers are "universal," UV irradiation may also be necessary (Golenberg 1991).

The second potential problem is diagenetic modification resulting in degradation of DNA (Pääbo 1989, 1990). If degradation is randomly scattered across a gene, then the fact that an amplified fragment is a consensus of multiple sequences should compensate for it (Pääbo and Wilson 1988). However, if only one target template is present in the amplification, then any diagenetic modifications will appear in the final sequence. One possible solution (though probably time consuming) is to amplify several highly diluted extractions containing on average zero to one target template each (cf. Golenberg 1991). An alternative is to sequence several extractions per individual. Another alternative (Golenberg 1991) exploits the integrity of double-stranded DNA and repairs the templates prior to amplification. In all of these cases, differences among sequences derived from a single individual could be considered diagenetic modifications. Importantly, diagenetic changes may be less of a problem when considering phylogenetics rather than genetic diversity. Diagenetic changes will likely appear as uniquely derived substitutions (see below); consequently they will have no effect on phylogenetic groupings.

If, on the other hand, the same diagenetic modifications were to occur systematically across samples (i.e., strand breakage and jumping PCR with an adenine insertion at exactly the same site; Pääbo et al. 1990), such modifications would be considered a *synapomorphy* (defined below). The likelihood of this may be small, as is the likelihood that only one character would greatly alter a phylogeny. But, however remote, the possibility does exist.

In the case of *D. panamintinus*, several precautions were taken. Negative controls were used throughout as were cross contamination precautions. At least one kangaroo rat specific primer was always used. PCR conditions were optimized to ensure efficient production of double-stranded product, and therefore a consensus of multiple sequences. Ultimately, sequence alignments, patterns of variation, and phylogenetic placement were used to control for authenticity.

3.2 Phylogenetic Analyses and Results

3.2.1 Homology, Saturation, Synapomorphies, and Outgroups

In order for nucleotide sequences to contain information regarding evolutionary relationships, the character states present in any one position

must be *homologous*. That is, they must be shared due to ancestry or be identical by descent. We can begin by assuming that homology is the basis of similarity (Hennig 1965), but we will need to evaluate this assumption. We may be able to reject homology if we can demonstrate that molecular evolutionary change is saturated (Villablanca and Thomas 1993; Hillis 1991; Hillis and Huelsenbeck 1992).

Molecular evolution tends to follow a single model when we consider one gene and one group of organisms (Brown et al. 1982; Brown and Simpson 1982; Aquadro et al. 1984). For example, vertebrate mitochondrial DNA evolves with a bias toward transitions relative to transversions (*ibid*). If observed replacements deviate extensively from the model, we conclude that some actual changes are going unobserved, as a result of evolution in parallel or via reversals (Brown et al. 1979). Mutational differences will appear to plateau once they reach a point of saturation. Importantly, saturated characters can no longer be considered homologous (Villablanca and Thomas 1993). Figure 2 shows that in kangaroo rat control regions, transitions make up the majority of mutations until the pairwise sequence difference between any pair of sequences reaches about 5%. Beyond that point, transitions continue to occur, yet there is no record of their continued evolution since transversions will begin to be counted instead. The genotypes from *D. panamintinus* considered below differ by a maximum of 2.76% corrected sequence divergence (observed sequence difference: max = 2.7%, mean = 2.06%, st. dev. = 0.6). This does not imply that all transitions are homologous, but it does suggest that as a category of change, transitions will be informative.

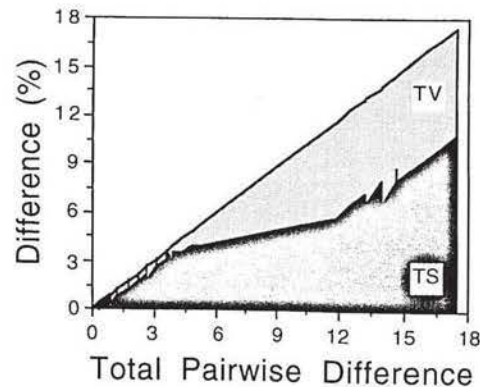


FIGURE 2. Difference in percentage contributed by transitions (TS) and transversions (TV) as a function of total (TS + TV) pairwise difference, indicating the level of pairwise difference at which a transition-biased model is no longer supported. Control region data are from six species of kangaroo rats, including population samples of *D. panamintinus*.

The grouping of organisms based on synapomorphies was one of the most significant conceptual advances of early phylogenetic systematics (Hennig 1965). We can divide evolutionary homologies into shared ancestral, shared derived, and unique (unshared). Only synapomorphies, or derived characters shared between one ancestor and its descendants, will be indicative of phylogenetic relationship. In contrast, shared ancestral characters occur in the ancestors of several mutually exclusive groups and are useless for identifying evolutionary relationships.

Outgroups are used to distinguish ancestral from derived character states (Watrous and Wheeler 1981; Maddison et al. 1984). An outgroup, comprised of at least two species from the sister group, or a more distantly related group, is used to root a phylogenetic tree. Character states shared between the ingroup and the two or more species in the outgroup are considered to be ancestral. Mutations from these character states to some new states would represent shared derived characters of the ingroup.

Phylogenetic studies of relationships among kangaroo rat species (Villablanca and Thomas 1993) using a 432-base-pair fragment of the mitochondrial cytochrome *b* gene (excluding silent first and third position transitions; see also Irwin et al. 1991, Cracraft and Helm-Bychowski 1991), led to the conclusion that the sister group to *D. panamintinus* is a clade containing at least four species. Here we will employ d-loop data from two of these species, *D. agilis* and *D. heermanni*, as an outgroup. Comparison with Thomas et al. (1990) shows that changing the outgroup to one species (*D. californicus*) results in only a minor change among branches that are not fully resolved to begin with. Thus, we are confident that either outgroup will yield similar results; by employing two species in the outgroup we decrease the probability of chance similarity between the outgroup and *D. panamintinus*.

The phylogenetic hypothesis presented here is based on the principle of *parsimony*. Swofford and Olsen (1990) provide a thorough and unbiased review of this and alternative methods for reconstructing phylogenies. Parsimony analysis does not explicitly assume that evolution occurs parsimoniously (all changes at all positions equally likely) but that we should accept the simplest explanation of the data (Wiley 1981, p. 111). As a result, the tree chosen under the parsimony criterion is the shortest tree—the one requiring the fewest number of events (observed and implied) to explain the distribution of characters among the organisms studied (Farris 1983). Fitch, or multistate unordered, parsimony is commonly used for nucleotide analysis (Felsenstein 1988). Though some proponents of parsimony say that its strength lies in the absence of the need to infer a specific model of evolution, we wish to emphasize that the strength of unweighted parsimony is seen when all changes occur at a low and equal rate and when all positions are equally likely to change (Felsenstein 1983, 1988; Villablanca and Thomas 1993). This is the model that parsimony inherently assumes, and the one under which it gives the most reliable results (Felsenstein 1983). In order to incorporate models of molecular evolution into evolutionary analysis, the computer program PAUP (version

3.0; Swofford 1989) allows the user to input a transition/transversion matrix weighting character state changes, as well as to assign weights to different positions. Interest in maximum likelihood methods (PHYLIP; Felsenstein 1983, 1988, 1989) is currently on the rise, primarily due to their greater statistical power, inherent acceptance of explicit models of evolution, and applications to population genetics (ibid: Hudson 1990). The acceptance of weighted parsimony analysis will likely pave the way for increased use of maximum likelihood methods.

3.2.2 Heuristic Parsimony Analyses and Consensus Trees

As the number of individuals being analyzed increases, so does the number of possible trees which need to be evaluated. Theoretically one should look at all possible trees, determine the number of changes required by each, and then choose the shortest tree. For the 22 individuals under consideration here we would have 3.2×10^{23} possible trees (Felsenstein 1978). In order to avoid a computational nightmare, PAUP performs a *heuristic search* (Swofford and Olsen 1990). This is an estimation of the shortest tree, obtained by starting with one individual and adding the others such that the tree length is minimized, then switching branches until one or more shorter or equally parsimonious trees are found, in which case branch swapping begins all over. One replicate of a heuristic search will produce all of the shortest trees which result from branch swapping on a given starting tree. This procedure is replicated ten or more times, with individuals being added to the starting tree in random order for each replicate.

In *D. panamintinus*, 32 of the 421 nucleotide positions vary across genotypes. Twenty of these positions have states shared by two or more genotypes and are thus inferred to be phylogenetically informative. Twelve positions have characters which occur uniquely in only one genotype. Ten replicated heuristic searches resulted in one set of 325 equally parsimonious trees. These trees are summarized in the *consensus tree* presented in Figure 3. The advantage to summarizing data in this way is that if any one of the 325 trees does not resolve a branch, then the branch will appear as part of a polycotomy in the consensus tree (Swofford 1989). The disadvantage is that it suggests some trees which are not observed in the heuristic results, and it conveys less information than any one of the equally parsimonious trees (Carpenter 1988).

Several reasons may exist for lack of phylogenetic resolution. First, the data may not have enough characters to resolve all relationships. This is certainly a problem with *D. panamintinus*, since we are trying to resolve 20 internal branches with exactly 20 "informative" characters. Second, characters may conflict with one another due to parallelisms or reversals (jointly termed *homoplasy*). Homoplastic characters can have several equally parsimonious resolutions. Of the 20 "informative" characters, 7 demonstrate homoplasy. Third, some of the branches connecting genotypes may be very short, with few or no substitutions supporting them. If this problem is suspected, more

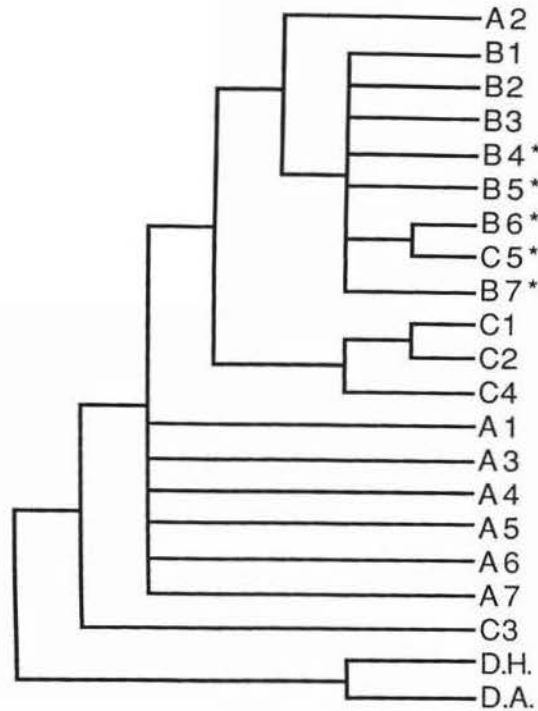


FIGURE 3. Strict consensus of 325 equally short trees obtained from 10 heuristic maximum parsimony searches of the data in Figure 1. All 10 replicates produced the same set of 53-step trees. Asterisks identify genotypes sequenced from museum specimens.

sequence data need to be collected. As indicated at the beginning of this chapter, our results are tentative for exactly this reason.

3.2.3 Exhaustive Parsimony Analyses and Tree Length Distributions

In some cases a disproportionate amount of the homoplasy may be contributed by the outgroup. If the outgroup's level of divergence relative to the ingroup places it in the region of saturation, we may suspect that at least some of the similarities are due to homoplasy rather than homology. Given that in this *D. panamintinus* data the mean sequence difference between the ingroup and the outgroup is 15.6%, we may be better off using a *functional outgroup* (Watrous and Wheeler 1981; Maddison et al. 1984). Regardless of how we root the tree, the relationship between *D. heermanni*, *D. agilis*, and genotype C3 (Fig. 3) is consistent. This relationship holds up even when the tree is rooted using only transversions between the ingroup and *D. merriami* and *D. nitrotoides*, which differ from the ingroup by a mean sequence difference

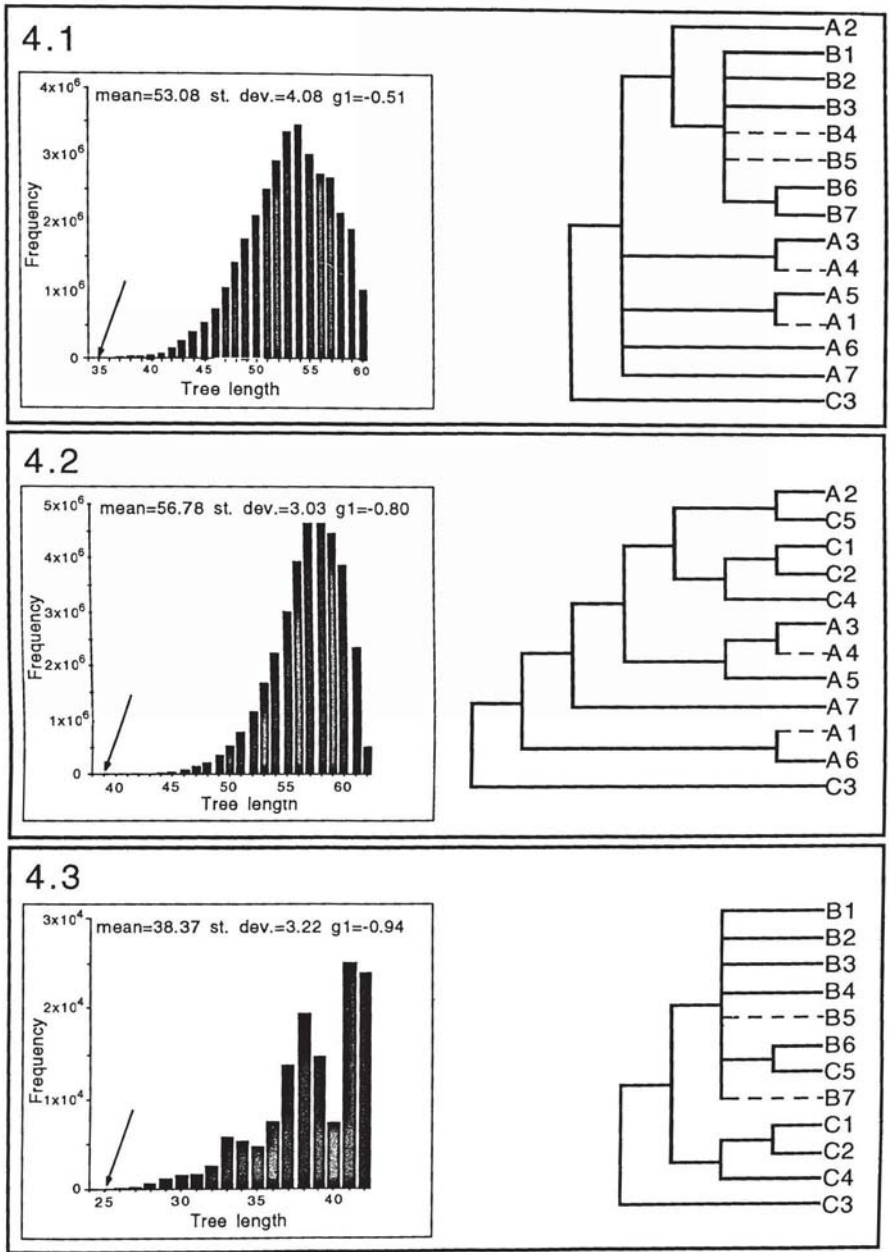
of 22%. Thus it is safe to conclude that genotype C3 would function as an excellent root (outgroup) for the remainder of the genotypes (ingroup). Making C3 the functional outgroup allows a reduction of the homoplasy resulting from the use of too divergent an outgroup.

An *exhaustive search* differs from a heuristic search in that we calculate the length of all possible trees (Swofford 1989). This procedure can currently be conducted for up to 11 genotypes and guarantees finding the shortest tree(s). Above 11 one can analyze random subsets of genotypes (Hillis 1991). An advantage of exhaustive searches is that we can output and analyze the frequency distribution of tree lengths. Hillis (1991) has shown that the distribution of tree lengths and the skewness of this distribution can be used to determine the relative information content of DNA sequences. A data set of randomized sequences will produce trees with a symmetrical length distribution. Data that contain significant phylogenetic information will result in very few trees toward the optimal (shorter) end of the distribution, and thus a skew toward that tail. The level of skewness can be determined by the *g*-statistic. The probability of observing the shortest tree can be inferred from the mean and standard deviation of tree lengths. These statistics are provided by PAUP along with the results of an exhaustive search.

The *D. panamintinus* data have been analyzed by three separate exhaustive searches. Each of the three populations is compared to the two others (A/B, A/C, B/C); the resultant trees and the length distributions are presented in Figure 4. In all cases, most parsimonious trees are between 4 and 5.6 standard deviations from the mean, with significant ($p < 0.01$) *g*-statistics (Hillis and Huelsenbeck 1992). The analysis involves comparison of 10 to 11 genotypes, using 12 to 15 phylogenetically informative positions. Increasing the number of individuals and the sequence length will add statistical power to this test (cf. Hillis and Huelsenbeck 1992).

The phylogenetic results permit several conclusions regarding *D. panamintinus* population structure. But first we must define the terms monophyly, paraphyly, and polyphyly (Wiley 1981 sensu Farris). *Monophyly* means that two groups have no genotypes in common; the groups arise from two separate ancestors. If we were to ignore genotype C5, Figure 4.3 would show two monophyletic groups. *Paraphyly* means that a group arises from one point within another group, and thus the stem group does not contain all of its descendants. In Figure 4.1, population A is paraphyletic relative to population B, since all B genotypes arise from a branch joining them with A2. *Polyphyly* means that individuals within a group arise from various different branches which also give rise to individuals of other groups; in other words, subsets of

FIGURE 4. Tree length distributions and parsimony trees for exhaustive search of genotypes in populations A and B (Fig. 4.1), A and C (Fig. 4.2), and B and C (Fig. 4.3). An arrow indicates the length of the shortest tree, which is compared to the mean and the standard deviation of all possible trees. A negative g_1 -statistic indicates the



degree of skew in the length distribution. Strict consensus trees are presented when more than one tree has the minimum length (Fig. 4.1 summarizes 30 trees; Fig. 4.3 summarizes 5 trees). In order to limit the number of genotypes analyzed to 10 or 11, genotypes which differed by one unique (unshared) difference only were removed from the search and then added to the tree with a dashed line adjacent to the genotypes from which they differ by one step.

individuals within the group are not in sister groups, and thus the most recent common ancestor is outside the group.

If one panmictic population were separated into two, each of these would initially be polyphyletic, with time progress through paraphyly, and eventually become monophyletic (Avice et al. 1987). Monophyly can be taken as evidence of genetic isolation (Hey 1991; Slatkin and Maddison 1989). Paraphyly, particularly in conjunction with fixed differences leading to the derived group, can also be taken as evidence of genetic isolation (Hey 1991; Fig. 4.1). Polyphyly is equivocal since populations may be reproductively isolated but not have been separated long enough to show monophyly (Fig. 4.3). In general, polyphyly suggests genetic exchange. Thus we can conclude that both populations B and C have arisen from genotypes within population A (Figs. 4.1 and 4.2, respectively), and that all populations except A and B are, or recently were, exchanging genes. Below we will estimate the actual levels of gene flow between populations.

3.2.4 Cladistic Estimate of Gene Flow

A phylogeny of alleles in one locus can be extremely useful for estimating the level of gene flow between populations (cf. Slatkin and Maddison 1989). Gene flow is normally calculated as the number of immigrants per generation (m). This value is difficult to calculate; thus genetic exchange between populations is normally expressed as $N_e m$, or the effective population size times the number of immigrants per generation. An important value is $Nm = 1$, below which genetic drift between populations becomes highly probable (Slatkin 1985). Genetic exchange at levels of $Nm > 1$ is sufficient to keep populations from diverging unless selection is extremely strong (ibid.).

Allele frequency data from several polymorphic loci, as well as private allele methods, have been used extensively to calculate F_{st} values from which one can estimate Nm (reviewed in Slatkin 1985). Slatkin and Maddison (1989) present a method based on phylogenies of alleles which is as useful as an F_{st} calculated from 10 protein loci (Slatkin 1991) and avoids problems Lewontin (1985) attributes to estimates of Nm based on F_{st} . Their method assumes that populations follow the island model, and that the true phylogeny for a nonrecombinant DNA fragment is known. The phylogeny is used to determine the minimum number of migration events required to explain the geographic distribution of alleles. A migration event(s) is postulated when we observe a discordance between the geographic and phylogenetic distribution of an allele: the geographic location is treated as a multistate unordered character, and each polymorphic state in a branch tip or internal node is assigned a value of $s = 1$. The sum of migration events (sum of s values) is compared to simulation results presented by Slatkin and Maddison (1989) in order to estimate Nm .

In Figure 4.1 observe that only one node (the one joining the B clade to genotype A2) is polymorphic when comparing populations A and B. An $s \equiv 1$

yields a minimum estimate of $Nm < 0.1$, with a standard deviation of less than 0.4. A comparison between populations A and C (Fig. 4.2) shows $s = 3$ and yields $Nm = 2$, with a standard deviation of about 1. Finally, if we compare populations B and C (Fig. 4.3), we find one polymorphic node and one polymorphic branch tip, which together yield $s = 2$ with a minimum estimate of $Nm = 0.66$ and a standard deviation of about 0.7. Thus we can conclude that these populations are held together by variable amounts of gene flow, with divergence due to genetic drift being most likely in population B. Drift is merely an expression for chance fixation of genotypes in a population due to the effects of random sampling of genotypes from one generation to the next. Either drift or selection or both can be the forces which take a population from polyphyly through paraphyly and ultimately to monophyly. Yet even with an Nm of 0.1 (between A and B), we can only be about 65% certain that all of the genotypes found in each population have arisen from a common ancestor within each population, and not from immigrant genotypes (Slatkin 1989).

3.3 Molecular Clock, Rate of Divergence, and Mutation Rate

The possibility of adding a temporal scale to molecular evolution is extremely intriguing. Used with caution and consideration, a well-calibrated molecular clock may be quite useful. To begin with, we could estimate the time of divergence between lineages. The time scale could then be used to estimate rates of evolution for other features, thus adding a temporal element to comparative biology. Finally, if mutations are neutral or have a selection coefficient of much less than the absolute value of $1/N$, then the substitution rate can be used to estimate the mutation rate (Kimura 1983, p. 47).

Due to degeneracy in the genetic code, genes evolve via mutations affecting amino acid replacement as well as through silent mutations. Silent mutations occur at a much faster rate. As a consequence, it makes little sense to consider a molecular clock based on an overall substitution rate. It may be better to consider the molecular "protein clock" proposed by Zuckerkandl and Pauling (Zuckerkandl 1987) separately from the "neutral clock" proposed by Kimura (1983, p. 100). The Zuckerkandl model is based on the observation that amino acid changes tend to accumulate linearly over time. The Kimura model is based on the conclusion that most silent mutations should be neutral or only slightly deleterious, and thus their accumulation should reflect a fairly constant mutation rate. Factors which cause excessive variation in substitution rates have been discussed by several authors (Kimura 1983; Ohta 1987; Dover 1987; Takahata 1987; Palumbi 1989; Gillespie 1991). The existence of a clock per se is not debated as much as is its accuracy and the statistical distribution of the underlying process. Theoretically, any gene can be used as a clock, so long as its mode of evolution (i.e., degree of constraint and distribution of variation) is understood.

As implied, the Zuckerkandl model is suitable for protein coding genes whereas the Kimura model is best for silent positions in protein coding genes, introns, or structural genes. In order to apply a molecular clock to non protein coding regions, one needs to be fairly confident that most mutations are neutral. This condition is most likely met in rapidly evolving, relatively unconstrained regions, such as introns of nuclear genes or the control region of animal mitochondrial DNA. For the Kimura model to be applicable, fairly similar sequences must be compared. Minimally, one must correct the observed difference for multiple mutations (i.e., corrected divergence; Nei 1987, p. 64). In the worst case it may become difficult even to align the sequences. Molecular clocks assume that all positions evolve at the same rate, or more specifically, a Poisson distributed rate. If the substitution process deviates significantly from this distribution, it will be difficult to apply the proper correction. There is evidence, at least for humans (Kocher and Wilson 1991), that the d-loop does not follow a Poisson distribution. Errors in time estimates based on d-loops would therefore increase at higher levels of divergence.

Calibration of a molecular clock is usually accomplished through the use of well-dated fossil material. For example, the split between the family which contains kangaroo rats and the family containing pocket gophers is dated at 20–23 million years ago (Wood 1935; Hafner and Hafner 1983). A comparison of replacement changes in a 432-bp fragment of cytochrome *b* between two pocket gophers and 16 kangaroo rats (Villablanca and Thomas 1993) yielded an average divergence of 5.48% (st. error 0.32). This corresponds to an average rate of 0.27%/MY. Irwin et al. (1991) estimated that the replacement rate of the entire cytochrome *b* gene in mammals is 0.36%/MY. These two estimates are compatible since the 432-bp fragment is from one of the more conserved regions of the gene (ibid.).

The rate of 0.27%/MY can be used to estimate the time since divergence between species of kangaroo rats. These times can then be used to calibrate the rate of substitution in the more rapidly evolving control region. The observed difference between 24 pairs of kangaroo rat control regions, corrected using Kimura's two-parameter model (eq. 5.5 of Nei 1987) and divided by the time since divergence for each pair of species, yields a mean divergence rate of 3.0%/MY (st. dev. 0.74) for the 421-bp fragment. If we make the reasonable assumption that all control region mutations are neutral, then the rate of substitution per nucleotide site per generation is equal to the mutation rate (Kimura 1983), or 15×10^{-9} (st. dev. 3.6×10^{-9}).

The rate of divergence can be used to infer the history of *D. panamintinus* populations. Considering the historical biogeography of American southwestern deserts plant assemblages, we would conclude that population B has been separated from populations A and C for at most 8,000 to 13,000 years (Spaulding 1990). Yet considering the phylogeny in Figure 1 we obtain a different estimate. All of the B genotypes differ from genotype A2 by three to five mutations (mean 3.9, st. dev. 0.8). This is equivalent to 0.7–1.2% difference, or roughly 200,000 to 400,000 years. Therefore, postglacial vegeta-

tional changes (Spaulding 1990) alone cannot be responsible for the degree of differentiation of population B. Given that on average it takes 2 times the female effective population size ($2N_f$) generations for a population to become monophyletic, in order for genotype C5 (Fig. 4.3) to have been retained from an ancestral population (i.e. para or polyphyly) the mean female population size of B would have been N_f greater than 200,000/2 to 400,000/2. Alternatively, if the population size of B were much less than this (i.e. time for monophyly has been sufficient), genotype C5 would have arisen in population B and flowed from B to population C. Unfortunately, the mean time of $2N_f$ has a very high variance. We will address this issue and test the alternative predictions regarding population B in Section 4.

Some aDNA samples are themselves fossils (more specifically, subfossils). As such, they should be useful for direct calibrations of molecular clocks (see Runnegar 1991): ancient samples stop accumulating mutations at the time they are deposited, while extant lineages continue to evolve. Unfortunately one cannot simply compare the percentage difference between two samples and divide by twice the time since divergence (once for each lineage). A phylogenetic approach is more appropriate. Consider a tree constructed from n_j modern sequences and n_i ancient ones, which contains $n_k = (j - 2 + i)$ internal nodes. If we know the time t in million years since the ancient genotype i was deposited, and we can find the node k furthest from the root which unites genotypes i and j , then the mean substitution rate per million years for one direct calibration is given by

$$\frac{\sum(d_{ij} - 2d_{ik})}{n_i n_j} \times \frac{1}{t_i}$$

where d_{ij} and d_{ik} are the number of substitutions/number of sites between genotypes i and j and between genotype i and node k , respectively, and $n_i n_j$ is the number of pairwise comparisons. This estimate is based on the number of changes inferred from the phylogeny. If sequences are so divergent that a correction measure is needed, then it becomes necessary to infer the ancestral sequences at all nodes and use these sequences to estimate the corrected d_{ik} . Likewise, if estimates are available from population size samples, and intrapopulation variation is significant, it will be necessary to estimate a mean d_{ij} and d_{ik} by subtracting the mean variation within populations from the variation between populations (cf. Wilson et al. 1985).

4. Population Genetics

The origin and maintenance of genetic variation is the central problematica of population genetics (Lewontin 1991). Studies of genetic variation at the allozyme level have not been as illuminating as originally hoped (ibid.), nor have DNA studies employing RFLPs (Wilson et al. 1989). DNA sequences, on the other hand, hold great promise (ibid.; Lewontin 1985). For example,

we already know that much, if not most, variation at the DNA level is selectively neutral (Kimura 1983); selection is limited to variation that wanders too far from neutrality. Importantly, the competing hypotheses of selection and neutrality are both testable (Kreitman 1991; Tajima 1989a). Population parameters as well help determine the amount of variation: effective population size, gene flow, genetic isolation, and population structure are all significant components. The mutation rate, or rate of origination of variation, is also a parameter which can be estimated under specific assumptions (see above). Altogether, the potential impact of DNA studies on population genetics is enormous. Population genetics, in turn, can contribute tremendously to our understanding of the evolution of demes, populations, species, and higher groupings.

4.1 Genetic Diversity at the Nucleotide Level

Genetic variation is currently best understood in the context of the coalescent (reviewed by Hudson 1990; Birky 1991). Put simply: if it takes on average $2N_f$ generations back in time for two mitochondrial genotypes to coalesce onto one node (ancestor), then the quantity $2N_f u$ (or $4N_e u$ for diploid nuclear genes) measures the expected number of nucleotide differences per site between randomly chosen genotypes, given that u is the mutation rate per nucleotide per generation, and N_f is the female effective population size. The parameter $2N_f u$ is termed *theta*. It is important to note that theta equals $4Nu$ when considering diploid nuclear genomes, while for organelle genomes the value of theta is $2Nu$. Theta (or more exactly its estimator π) is a useful measure of variation in its own right, and with additional assumptions can be quantitatively related to many population parameters. For example, π gives the average internode length and mean coalescence time for genotypes in a population (Hudson 1990), and it can be used to calculate F_{st} values between populations (Slatkin 1991) as well as Nm (gene flow times population size; see above, Slatkin 1991). Regrettably, the coalescence times of alleles in a phylogeny are exponentially distributed, while mutations accumulate following a Poisson or negative binomial distribution. Consequently, π has a very large variance (Tajima 1983). For this reason, alternative approaches which estimate coalescence parameters will be very useful, if and when they become available.

4.1.1 Estimating Theta and Its Variance

Estimation of parameters from the coalescent theory requires explicit assumptions: the process is based on an ideal Wright-Fisher model (discussed in Birky 1991; Hudson 1990). It requires us to explicitly assume neutrality, a molecular clock, equilibrium with respect to mutation and random drift, and that the total tree length equals the actual number of mutations. This last assumption is based on an infinite-sites model (Kimura 1983, p. 46), which

states that the probability of multiple mutations at the same nucleotide position is infinitesimal. Invariant positions, rate variations across positions, and saturation are empirical facts which point to the weaknesses of the model. Hudson (1990) presents several alternative models of evolution which exploit the coalescent and allow for processes such as: selection, hitchhiking, and recombination. His work is an excellent review of the utility and power of the coalescent.

Theta can be estimated directly from sequence data using either of two estimators. The first (eq. 10.5 of Nei 1987) uses the observed number of differences (π = nucleotide diversity) to estimate the expected number of differences. The variance of π (eq. 10.9 of Nei 1987, or eq. 16 of Hudson 1990) is a function of the estimated value of theta and the number of sequences examined. The second estimator of theta (Hudson eq. 17 and Nei eq. 10.3) uses the number of polymorphic (segregating) nucleotide sites per sequence (S of Hudson, and p_n of Nei), corrected to be unbiased by sample size.

4.1.2 Nucleotide Diversity in Modern and Museum Populations

Estimates of π , plus or minus one standard error, are presented in Figure 5. Values are obtained from Nei's (1987) equations 10.5 and 10.9. These results show that mean nucleotide diversity within populations is stable over time even though the actual alleles represented in the samples differ from one time point to the next. As a result, museum populations can be used to

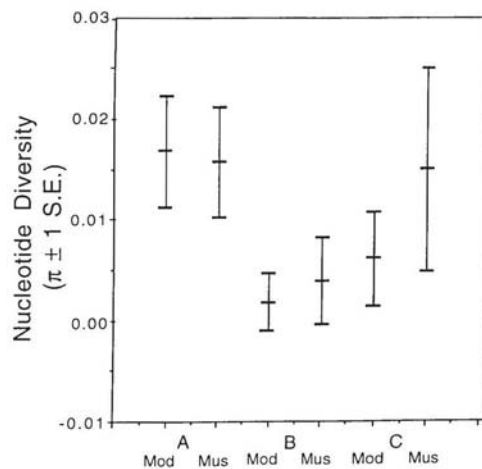


FIGURE 5. Nucleotide diversity ($\pi \pm 1$ st. dev.) among three geographic populations of *D. panamintinus*, comparing modern (Mod) and museum (Mus) samples for each population.

establish base levels of genetic variation in populations. The effect of sample size on confidence limits of π can also be seen in Figure 5. The population with the smallest sample size (Mus C, $n = 8$) is also the one with the largest range, represented by one standard error. The standard errors given are actually underestimates; we are assuming a constant mutation rate across positions, which is clearly unrealistic (Kocher and Wilson 1991).

Based on percentage sequence divergence and a molecular clock we estimated above that population B diverged about 200,000 to 400,000 years ago. It was predicted that either population B contained genotype C5 as a polymorphism retained from the population ancestral to B and C, or that C5 arose within population B and is found in C due to gene flow. The ancestral polymorphism hypothesis predicts $N_f \geq 100,000$ –200,000, while the gene flow hypothesis predicts $N_f < 100,000$ –200,000. Using π and the mutation rate, it is possible to estimate an N_f of 74,500 for population B, which would lead us to reject the ancestral polymorphism hypothesis. Yet, when the variance of π is taken into account, it becomes obvious that it is impossible to say whether the genotype shared between these two populations was present in a population ancestral to both, or whether it arose in one and flowed to the other.

4.1.3 Demonstrating Significant Changes in Nucleotide Diversity over Time

Loss of genetic variability is primarily a function of the effective population size. It is generally assumed that variation is lost at a rate of $\frac{1}{2N_e}$ and $\frac{1}{N_f}$ for diploid and haploid genes, respectively (Wright 1969, p. 212). Thus the only cases in which changes in genetic variation could be observed would be if populations changed in size (ibid.) or if selection acted to change the frequency at an allele and linked loci (Kreitman 1991; Kaplan et al. 1989). Yet even if a population was significantly reduced in size, it would take N generations for it to return to drift-mutation equilibrium. Using temporal samples of aDNA to show changes in genetic variation would require that the time interval (t) between these temporal samples correspond roughly to the new N of the population (i.e. $t = N$) and that the population be held at roughly the same N for the period between samples. As an example, assume a population of $N_e = 1,000$ exists at time $t = 0$, and that the population falls to $N_e = 100$ at time $t = 1$; then a sample from $t = 0$ and $t = 100$ (in generations) would be required in order to estimate the full loss in variability due to the reduction in population size. Samples from any time prior to $t = 100$ would not be as useful, since the population would still be responding to drift.

The large variance of nucleotide diversity further complicates the issue. Fortunately, as the number of alleles (n) in the sample increases, the variance of π decreases (Tajima 1983), though the advantage is hardly noticeable past an n of about 10. It is also noteworthy that the variance of π increases with increasing π (ibid.). To exploit this property of π , one could utilize genes with

low mutation rates when comparing populations in the hundreds of thousands. Since π is a product of N_e and u , decreasing u when N_e is held constant will in turn yield a lower estimate of π and thus a lower variance. Using genes with faster mutation rates is preferable for smaller populations: the estimate of π may be less precise since π and therefore its variance are higher, yet the mean will be accurate if (and only if) the locus evolves quickly and variation is observed.

In order to investigate the utility of nucleotide diversity for estimating reductions in genetic variability, the expected π for kangaroo rat d-loops has been projected over a wide range of N_f values. The results (Fig. 6) are obtained for $u = 1.5 \times 10^{-8}$, 500 base pairs, and 10 alleles sampled from 20 individuals. The figure can be interpreted by finding the mean π value up the N_f scale which lies outside of the 95% confidence interval identified from standard errors of π values down the N_f scale. Generally speaking, the larger population would have to differ by a factor of three to four from the smaller population in order for the mean of the larger population to lie outside two standard errors of the mean of the smaller population.

When the purpose is to test for changes in the level of genetic variability over time, it may be more appropriate to use estimates of theta based on the number of segregating sites (S of Hudson, and p_n of Nei) rather than π . Tajima (1989b) has shown that when a population is in a bottleneck, the current population size affects S more than it affects π , while the size of the original population has a greater effect on π than it has on S . Following the bottleneck, S will recover relatively quickly while π will reflect the bottleneck for a much longer time. In contrast to the comparisons of π values over time (Fig. 5) the numbers of segregating sites show pronounced variation between

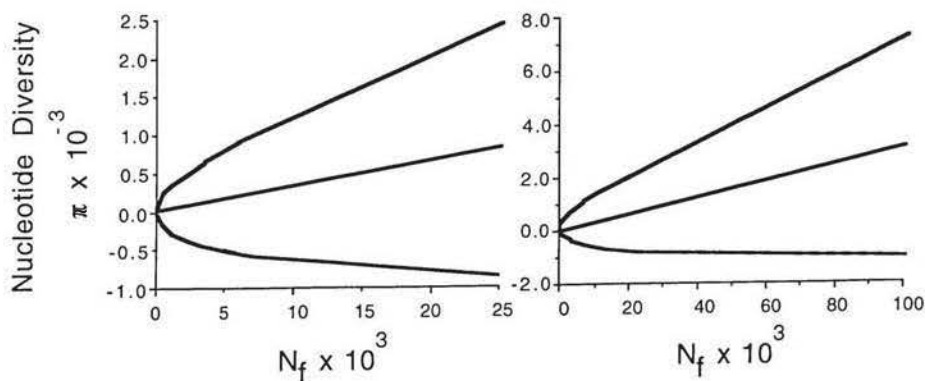


FIGURE 6. Nucleotide diversity (π) as a function of effective population size (N_f). Curves bracketing the central line (mean π) indicate 2 standard errors. The panel on the right is a continuation of the one on the left. Note the difference in scales. See text for discussion.

temporal samples of kangaroo rats. Population A has the most constant S : its value is 10 for both museum and modern samples. Even so, when corrected for sample size (Tajima 1989b), S shows a reduction of 10% over time. Population B shows a decrease of 66%, from 6 to 2. Population C demonstrates a drop of 27%, from 11 to 8. These results clearly illustrate the greater sensitivity of S while suggesting that these populations are not in equilibrium. For a 66% reduction in S (as in population B), a 55% reduction is expected in π (Tajima 1989b); yet Figure 5 shows no such reduction. Following Tajima, the only reasonable interpretation of these results, other than stochastic sampling error, is that π reflects the historical size of the B population, while S more closely reflects the current population size. These results cannot be due to diagenetic processes. If degradation of DNA were a factor, then all museum populations should show a larger number of segregating sites. Population A does not support this notion. Likewise, it could be suspected that even though the S values are corrected for sample size, a bias remains. If this were the case, it would be predicted that modern populations, because of their slightly larger sample sizes, would have larger values of S . The results clearly refute this prediction.

4.2 Changes in Gene Frequency over Time

Another aspect of changing genetic variation over time is the frequency of individual genotypes. Since theta is a function of $2Nu$, it is possible to have different genotypes present in different temporal samples without observing a change in π . Thomas et al. (1990) show a change in the frequency and presence/absence of genotypes across temporal samples. The question is whether this change in frequency is significantly different from what would be expected due to chance alone.

Thomas et al. test for significant frequency shifts by trying to show that the probability of identity by descent between temporal samples is not drastically different from that shown within each temporal sample. Yet this approach does not provide for the estimation of a confidence interval, and thus the interpretation is somewhat subjective. An alternative approach is to test whether the observed frequencies are statistically dependent or independent of the sample time. If the frequency differences are independent of sample origins, then the observed variation can be explained as mere sampling error. A $2 \times n$ contingency table can be applied to the genotype frequencies shown in Figure 7. If the results approach significance, it is necessary to apply a correction for small expected frequencies (Sokal and Rohlf 1981, pp. 735–47). Significance values of the G -test are also shown in Figure 7. None of the tests are significant, though this conclusion requires Williams correction (G_{adj}) in the case of population A. We must conclude that variation in frequencies is simply attributable to sampling error. The possibility does remain that the results are significant in the case of population A; larger samples will be required to test this further.

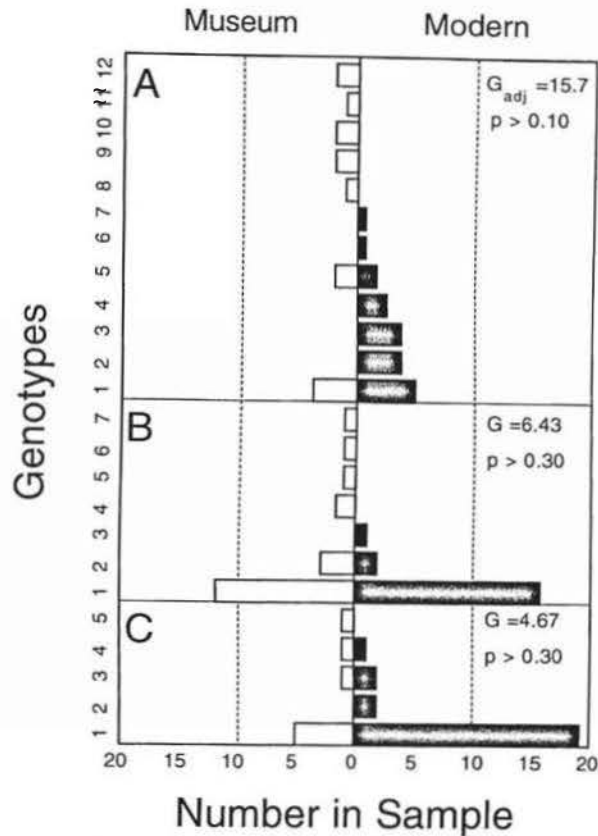


FIGURE 7. Mitochondrial haplotype frequencies for two temporal samples each of three populations of kangaroo rats. Differences in frequency between samples are tested with a $2 \times n$ contingency table. G -statistics and p values are indicated. If the G -statistic is significant, then a G adjusted for small expected values (G_{adj}) is estimated via a Williams correction.

The nature of temporal variation in population A is intriguing. All of the genotypes which are present in the 1911 sample, but absent in the 1988 sample, were found in subadults of the 1911 sample. One possible interpretation is that subadults from neighboring demes dispersed into the deme being sampled, but failed to set up residency and become breeding members of the deme. Though there are alternate interpretations, they all have the same implication: when subadults (roughly equated with nonbreeders) are included in temporal samples, they do not significantly affect the values of π or the G -test, or the conclusions of phylogenetic analysis. Caution might be required in comparing temporal samples taken during different phases of the dispersal cycle in populations where dispersers temporarily increase the number of segregating sites (S) but do not have a high probability of becoming resident breeders.

5. Discussion

5.1 Studies Uniquely Based on aDNA

Many of evolution's most interesting processes have an intrinsically temporal component. Evolutionary analysis of aDNA can afford us the temporal samples needed to study such processes. Yet clearly, the potential applications are still in the process of being identified.

Hybrid zones represent areas of genetic contact between differentiated forms. The processes and dynamics occurring at hybrid zones have a direct impact on speciation and differentiation (Endler 1977; Harrison 1990). Lessa (1992 and pers. comm.) has begun to unlock the temporal information which museum specimens provide. His study investigates a situation in which the population on one side of a hybrid zone has recently come to morphologically resemble the population on the other side. Mitochondrial DNA, in concert with nuclear genes and denaturing gradient gels, allows for the testing of three alternative hypotheses: local extinction with recolonization by a different morph, male-biased gene flow, and morphological convergence without genetic exchange.

Tools to study speciation can also be found in the aDNA tool chest. They offer the potential to directly test long-standing hypotheses regarding the mechanisms and implications of speciation. Some researchers have strongly argued for the importance of the founder flush, or genetic revolution mechanisms, for speciation (Mayr 1954, 1963; Carson 1975, 1982). These hypotheses make explicit predictions regarding the levels and structuring of genetic variation prior to and following speciation. Properly placed temporal samples could provide an avenue for testing such predictions.

Neontological museum specimens have already been applied to the study of evolutionary relationships in extinct species (Thomas et al. 1989). Thomas et al. (1990) and the present study suggest how phylogenetic and population genetic analysis can be extended to rare or endangered species and populations. Additionally, museum collections may contain information which is no longer available from modern populations. One possible example of this is the presence of a shared genotype between museum samples of populations B and C. The genotype (B6 and C5) was found in museum samples with little effort, yet could not be found in modern samples of 19 and 24 individuals. As discussed, this genotype critically affects our interpretation of gene flow between populations B and C.

The possible information contained in archaeological and paleontological specimens is also an area of active research. Several different approaches show promise. The ground-breaking work of Golenberg (cf. Chapter 17) clearly exemplifies two advances. First, as discussed above, given the proper phylogenetic analysis, Miocene samples of *Magnolia* can be used to directly calibrate the molecular clock. Second, these samples lead to specific conclusions regarding the prior geographic distribution of a particular phylogenetic group.

The subfossil material is sufficient to place magnolias in a certain part of North America during the Miocene, while the aDNA from the samples indicates the specific evolutionary lineage involved. Such simultaneous biogeographic and phylogenetic links are usually difficult to establish. The potential of subfossil molecules to reveal the history of organisms and molecules alike is largely unexplored. It seems likely that many interesting studies will be forthcoming in this area.

Phylogenetic reconstructions can be radically affected by the inclusion or exclusion of fossil specimens (Donoghue et al. 1989). There are several reasons for this effect: fossils can identify intermediate and differentiated forms along lineages which later converge; they can resolve short branches along which evolution in rapidly evolving characters can be arrested, so that the record of change is not obliterated; in some cases they are the only hope for resolving the problems of chance similarity between branch tips (cf. Henny and Penny 1989; Sanderson and Donoghue 1989).

5.2 *Biological and Statistical Samples*

Successful investigation of many of the issues discussed above will depend on the biological and statistical distribution of the samples. Though it is impossible to foresee all possible issues, some basic guidelines can be provided here.

Phylogenetic analysis can have profound impact on our perception of populations, species, speciation, and comparative biology. Identifying a population as monophyletic, paraphyletic, or polyphyletic is the key, both in terms of gene flow and differentiation, and in determining whether gene trees correspond to population or species trees (Takahata and Slatkin 1990). There is a clear relationship here: the probability of observing paraphyly and polyphyly increases with increasing sample size (ibid.). There is no better argument against the common practice of using no more than one or two individuals per species in phylogenetic analyses. Conclusions at one hierarchical level must be based on adequate sampling of the next lower level. For example, conclusions regarding *D. panamintinus* can currently be drawn only at the population level, since only individuals within populations have been sampled. In order to draw conclusions regarding subspecies, populations within subspecies would need to be sampled. Similarly, it is clear that in order to make interspecific inferences, one would need to sample populations within species.

Estimates of population parameters may also present particular sampling requirements. Gene flow (Nm) estimates based on phylogenies of alleles are most confidently done from samples of 10 or more alleles (genotypes) per population (Slatkin and Maddison 1989). Since the number of alleles in a population is a function of demographic and historical processes (Slatkin 1987), one is limited to sampling what is available. The number of alleles may be increased by sampling several demes from one population (ibid.).

Confidence in estimates of nucleotide diversity (π) is affected by the number of individuals as well as the number of alleles. The variance of π decreases significantly up to samples of about 10 individuals (Tajima 1983). Above 10 individuals, the improvement is small unless samples are in the neighborhood of 50 individuals (ibid.). Diachronic comparisons of genotype frequency are likewise affected by the number of individuals sampled. Only when samples are large (> 20) can one hope to observe any existing statistical significance. If the expected frequencies of genotypes are low (on the order of 0.05), then a Williams correction is required. This test is conservative (Sokal and Rohlf 1981) and thus requires increased sample sizes. Finally, it seems that diachronic comparisons of genetic variation using π will be most useful when the temporal samples are separated by $t = N \times$ generation time (yrs), where N is the population size following a change in population size. The number of segregating sites (S) may be superior to nucleotide diversity (π) in detecting differences in diachronic comparisons when the number of generations is less than N and when populations are in a bottleneck (Tajima 1989b). Yet if populations are rebounding from a severe bottleneck, then π would be a more sensitive detector of changing levels of genetic variation.

Acknowledgments. This work has been supported by a National Science Foundation Graduate Fellowship and by a grant from the Museum of Vertebrate Zoology. Special thanks are given to Monty Slatkin for helpful discussions and critical review of the manuscript.

References

- Aquadro CF, Kaplan N, Risko KJ (1984) An analysis of the dynamics of mammalian mitochondrial DNA sequence evolution. *Mol Biol and Evol* 5:423–434
- Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel JE, Reeb CA, Saunders NC (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annu Rev Ecol Syst* 18:489–521
- Bibb M, Van Etten RA, Wright CT, Walberg MW, Clayton DA (1981) Sequence and gene organization of the mouse mitochondrial DNA. *Cell* 26:167–180
- Birky CW Jr (1991) Evolution and population genetics of organelle genes: mechanisms and models. In: Selander RK, Clark AG, Whittam TS (eds) *Evolution at the Molecular Level*. Sunderland Mass.: Sinauer, pp. 112–134
- Birky CW Jr, Maruyama T, Fuerst P (1983) An approach to population and evolutionary genetic theory for genes in mitochondria and chloroplasts and some results. *Genetics* 103:513–527.
- Birky CW Jr, Fuerst P, Maruyama T (1989) Organelle gene diversity under migration, mutation and drift: equilibrium expectations, approaches to equilibrium, effects of heteroplasmic cells, and comparisons to nuclear genes. *Genetics* 121:613–627
- Brown WM (1985) The mitochondrial genome of animals. In: MacIntyre RJ (ed) *Molecular Evolutionary Genetics*, New York: Plenum Press, pp. 95–130

- Brown GG, Simpson MV (1982) Novel features of animal mtDNA evolution as shown by sequences of two cytochrome oxidase subunit II genes. *Proc Natl Acad Sci USA* 79:3246–3250
- Brown WM, George M Jr, Wilson AC (1979) Rapid evolution of animal mitochondrial DNA. *Proc Natl Acad USA* 76:1967–1971
- Brown WM, Prager EM, Wang A, Wilson AC (1982) Mitochondrial DNA sequence of primates: tempo and mode of evolution. *J Mol Evol* 18:225–239
- Carpenter JM (1988) Choosing among multiple equally parsimonious cladograms. *Cladistics* 4:291–296
- Carson HL (1975) The genetics of speciation at the diploid level. *Am Naturalist* 109:83–92
- Carson HL (1982) Speciation as a major reorganization of polygenic balances. In: Barigozzi C (ed) *Mechanisms of Speciation*. New York: Alan R. Liss, pp. 411–433
- Clegg MT, Learn GH, Golenberg EM (1991) Molecular evolution of chloroplast DNA. In: Selander RK, Clark AG, Whittam TS (eds) *Evolution at the Molecular Level*. Sunderland, Mass.: Sinauer, pp. 135–149
- Cracraft J, Helm-Bychowski K (1991) Parsimony and phylogenetic inference using DNA sequences: some methodological strategies. In: Miyamoto MM, Cracraft J (eds) *Phylogenetic Analysis of DNA Sequences*. Oxford: Oxford Univ. Press, pp. 184–220
- DeSalle R, Freedman T, Prager EM, Wilson AC (1987) Tempo and mode of sequence evolution in mitochondrial DNA of Hawaiian *Drosophila*. *J Mol Evol* 26:157–164
- Donoghue MJ, Doyle JA, Gauthier J, Kluge AG, Rowe T (1989) The importance of fossils in phylogeny reconstruction. *Annu Rev Ecol Syst* 20:431–460
- Dover G (1987) DNA turnover and the molecular clock. *J Mol Evol* 26:47–58
- Endler JA (1977) *Geographic Variation, Speciation and Clines*. Princeton, N.J.: Princeton Univ. Press
- Farris JS (1983) The logical basis of phylogenetic analysis. In: Platnick NI, Funk VA (eds) *Advances in Cladistics*, vol 2. New York: Columbia Univ. Press, pp. 7–36
- Felsenstein J (1978) The number of evolutionary trees. *Syst Zool* 27:27–33
- Felsenstein J (1979) Alternative methods of phylogenetic inference and their inter-relationship. *Syst Zool* 28:49–62
- Felsenstein J (1983) Parsimony in systematics: biological and statistical issues. *Annu Rev Ecol Syst* 14:313–333
- Felsenstein J (1988) Phylogenies from molecular sequences: inference and reliability. *Annu Rev Genet* 22:521–565
- Felsenstein J (1989) PHYLIP 3.2 Manual. Berkeley, Cal.: Univ. of California Herbarium
- Gillespie JH (1991) *The Causes of Molecular Evolution*. Oxford: Oxford Univ. Press
- Golenberg EM (1991) Amplification and analysis of Miocene plant fossil DNA. *Phil Trans R Soc Lond B* 333:419–426
- Gyllenstein U, Wharton D, Josefsson A, Wilson AC (1992) Paternal inheritance of mitochondrial DNA in mice. *Nature (London)* 352:255–357
- Hafner JC, Hafner MS (1983) Evolutionary relationships of heteromyid rodents. *Great Basin Naturalist Memoirs* 7:3–29
- Harrison RC (1990) Hybrid zones: windows on evolutionary processes. In: Futuyama D, Antonovics J (eds) *Oxford Surveys in Evolutionary Biology*, vol 7. Oxford: Oxford Univ. Press, pp. 69–128
- Hendy MD, Penny D (1989) A framework for the quantitative study of evolutionary trees. *Syst Zool* 38:297–309

- Hennig W (1965) Phylogenetic systematics. *Annu Rev Entomol* 10:97–116
- Hey J (1991) The structure of genealogies and the distribution of fixed differences between DNA sequence samples from natural populations. *Genetics* 128:831–840
- Hillis DM (1991) Discriminating between phylogenetic signal and random noise in DNA sequences. In: Miyamoto MM, Cracraft J (eds) *Phylogenetic Analysis of DNA Sequences*. Oxford: Oxford Univ. Press, pp. 278–294
- Hillis D, Huelsenbeck J (1992) Signal, noise, and reliability in molecular phylogenetic analysis. *Jour Hered* 83:189–195
- Hudson RR (1990) Gene genealogies and the coalescent process. In: Futuyma D, Antonovics J (eds) *Oxford Surveys in Evolutionary Biology*, vol. 7. Oxford: Oxford Univ. Press, pp. 1–44
- Irwin D, Kocher TD, Wilson AC (1991) Evolution of the cytochrome *b* gene of mammals. *J Mol Evol* 32:128–144
- Kaplan NL, Hudson RR, Langley CH (1989) The “hitchhiking effect” revisited. *Genetics* 123:887–899
- Kimura M (1983) *The Neutral Theory of Molecular Evolution*. Cambridge: Cambridge Univ. Press
- Kocher TD, Wilson AC (1991) Sequence evolution of mitochondrial DNA in humans and chimpanzees: control region and a protein coding region. In: Osawa S, Honjo T (eds) *Evolution of Life*. Tokyo: Springer-Verlag, pp. 391–413
- Kreitman M (1991) Detecting selection at the level of DNA. In: Selander RK, Clark AG, Whittam TS (eds) *Evolution at the Molecular Level*. Sunderland, Mass.: Sinauer, pp. 204–221
- Lessa EP (1992) Rapid surveying of DNA sequence variation in natural populations. *Mol Biol Evol* 9:323–330
- Lewontin RC (1985) Population genetics. In: Greenwood PJ, Harvey PH, Slatkin M (eds) *Evolution: Essays in the Honor of John Maynard Smith*. Cambridge: Cambridge Univ. Press, pp. 3–18
- Lewontin RC (1991) Twenty-five years ago in genetics: electrophoresis in the development of evolutionary genetics—milestone or millstone? *Genetics* 128:657–662
- Maddison WP, Donoghue MJ, Maddison DR (1984) Outgroup analysis and parsimony. *Syst Zool* 33:83–103
- Mayr E (1954) Change of genetic environment and evolution. In: Huxley J, Hardy AC, Ford EB (eds) *Evolution as a Process*. New York: MacMillan, pp. 157–180
- Mayr E (1963) *Animal Species and Evolution*. Cambridge, Mass.: Harvard Univ. Press
- Miyamoto MM, Boyle SM (1989) The potential importance of mitochondrial DNA sequences to eutherian mammal phylogeny. In: Fernholm B, Bremer K, Jörnvall H (eds) *The Hierarchy of Life*. Amsterdam: Elsevier, pp. 437–450
- Moritz C, Dowling TE, Brown WM (1987) Evolution of mitochondrial DNA: relevance for population biology and systematics. *Annu Rev Ecol Syst* 18:269–292
- Nei M (1987) *Molecular Evolutionary Genetics*. New York: Columbia Univ. Press
- Ohta T (1987) Very slight deleterious mutations and the molecular clock. *J Mol Evol* 26:1–6
- Orrego C (1990) Organizing a laboratory for PCR work. In: Innis MA, Gelfand DH, Sninsky JJ, White TJ (eds) *PCR Protocols: A Guide to Methods and Applications*. San Diego: Academic Press, pp. 447–454
- Pääbo S (1989) Ancient DNA: extraction, characterization, molecular cloning, and enzymatic amplification. *Proc Natl Acad Sci USA* 86:1939–1943
- Pääbo S (1990) Amplifying ancient DNA. In: Innis MA, Gelfand DH, Sninsky JJ,

- White TJ (eds) *PCR Protocols: A Guide to Methods and Applications*. San Diego: Academic Press, pp. 159–166
- Pääbo S, Wilson AC (1988) Polymerase chain reaction reveals cloning artefacts. *Nature (London)* 334:387–388
- Pääbo S, Irwin DM, Wilson AC (1990) DNA damage promotes jumping between templates during enzymatic amplification. *J Biol Chem* 265:4718–4721
- Palumbi SR (1989) Rates of molecular evolution and the number of nucleotide positions free to vary. *J Mol Evol* 29:180–187
- Pamilo P, Nei M (1988) Relationship between gene trees and species trees. *Mol Biol Evol* 5:568–583
- Runnegar B (1991) Nucleic acid and protein clocks. *Phil Trans R Soc Lond B* 333:391–397
- Sanderson MJ, Donoghue DJ (1989) Patterns of variation in levels of homoplasy. *Evolution* 43:1781–1795
- Slatkin M (1985) Gene flow in natural populations. *Annu Rev Ecol Syst* 16:393–430
- 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 (1991) Inbreeding coefficients and coalescence times. *Genet Res Camb* 58:167–175
- Slatkin M, Maddison WP (1989) A cladistic measure of gene flow inferred from the phylogenies of alleles. *Genetics* 123:602–613
- Sokal RR, Rohlf FJ (1981) *Biometry*, 2nd ed. New York: Freeman, W.H.
- Spaulding WG (1990) Vegetational and climatic development of the Mojave Desert: the last glacial maximum to the present. In: Betancourt JL, Van Devender TR, Martin PS (eds) *Packrat Middens*, Tucson: Univ. of Arizona Press, pp. 166–199
- Swofford DL (1989) Phylogenetic analysis using parsimony (ver. 3.0). Illinois Natural History Survey
- Swofford DL, Olsen GJ (1990) Phylogeny reconstruction. In: Hillis DM, Moritz C (eds) *Molecular Systematics*. Sunderland Mass: Sinauer, pp. 411–501
- Tajima F (1983) Evolutionary relationship of DNA sequences in a finite population. *Genetics* 105:437–460
- Tajima F (1989a) Statistical method for testing the neutral mutation hypothesis of DNA polymorphism. *Genetics* 123:585–595
- Tajima F (1989b) The effect of change in population size on DNA polymorphism. *Genetics* 123:597–601
- Takahata N (1987) On the overdispersed molecular clock. *Genetics* 116:169–179
- Takahata N, Slatkin M (1990) Genealogy of neutral genes in two partially isolated populations. *Theor Popul Biol* 38:331–350
- Thomas RH, Schaffner W, Wilson AC, Pääbo S (1989) DNA phylogeny of the extinct marsupial wolf. *Nature (London)* 340:465–467
- Thomas WK, Pääbo S, Villablanca FX, Wilson AC (1990) Spatial and temporal continuity of kangaroo rat populations shown by sequencing mitochondrial DNA from museum specimens. *J Mol Evol* 31:101–112
- Villablanca FX, Thomas WK (1993) Empirical limits to the phylogenetic utility of DNA sequences. M.S. submitted
- Waish SP, Metzger DA, Higuchi R (1991) Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *BioTechniques* 10:506–513

- Watrous LE, Wheeler QD (1981) The outgroup comparison method of character analysis. *Syst Zool* 30:1-11
- Wiley EO (1981) *Phylogenetics: The Theory and Practice of Phylogenetic Systematics*. New York: Wiley
- Wilson AC, Cann RL, Carr SM, George M, Gyllensten UR, Helm-Bychowski KM, Higuchi RG, Palumbi SR, Prager EM, Sage RD, Stoneking M (1985) Mitochondrial DNA and two perspectives on evolutionary genetics. *Biol J Linn Soc* 26:375-400
- Wilson AC, Zimmer EA, Prager EM, Kocher TD (1989) Restriction mapping in the molecular systematics of mammals: a retrospective salute. In: Fernholm B, Bremer K, Jörnvall H (eds) *The Hierarchy of Life*. Amsterdam: Elsevier, pp. 407-419
- Wolfe KH, Li W-H, Sharp PM (1987) Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proc Natl Acad Sci USA* 84:9054-9058
- Wood AE (1935) Evolution and relationships of the heteromyid rodents with new forms from the Tertiary of western North America. *Ann Carnegie Mus* 24:73-262
- Wright S (1969) *Evolution and the Genetics of Populations*, vol 2. Chicago: University of Chicago Press
- Zuckerkindl E (1987) On the molecular evolutionary clock. *J Mol Evol* 26:34-46