



**Phylogenetic Hypotheses Under the Assumption of Neutral Quantitative- Genetic Variation**

Michael Lynch

*Evolution*, Vol. 43, No. 1. (Jan., 1989), pp. 1-17.

Stable URL:

<http://links.jstor.org/sici?sici=0014-3820%28198901%2943%3A1%3C1%3APHUTAO%3E2.0.CO%3B2-5>

*Evolution* is currently published by Society for the Study of Evolution.

---

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/ssevol.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

---

The JSTOR Archive is a trusted digital repository providing for long-term preservation and access to leading academic journals and scholarly literature from around the world. The Archive is supported by libraries, scholarly societies, publishers, and foundations. It is an initiative of JSTOR, a not-for-profit organization with a mission to help the scholarly community take advantage of advances in technology. For more information regarding JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

# EVOLUTION

INTERNATIONAL JOURNAL OF ORGANIC EVOLUTION

PUBLISHED BY

THE SOCIETY FOR THE STUDY OF EVOLUTION

---

Vol. 43

January 1989

No. 1

---

*Evolution*, 43(1), 1989, pp. 1-17

## PHYLOGENETIC HYPOTHESES UNDER THE ASSUMPTION OF NEUTRAL QUANTITATIVE-GENETIC VARIATION

MICHAEL LYNCH

*Department of Ecology, Ethology, and Evolution, University of Illinois,  
Shelford Vivarium, 606 E. Healey St., Champaign, IL 61820*

*Abstract.*—There are many situations in which the only available characters for reconstructing phylogenies are morphological. Those traits that are subject only to the forces of mutation and random genetic drift can be used to obtain unbiased estimates of phylogenetic relationships. However, the accurate recovery of a phylogeny from information on neutral characters requires the procurement of data for a large number of independent traits, individuals, and populations. Phylogenetic trees fit to data from more than five species will almost always contain topological errors, even with very large data sets. The population-genetic consequences of the neutral model are reviewed, and some statistical methods for testing whether the diversification of a phylogeny is compatible with such a model are outlined. The theory is then applied to a very large data set on cranial morphology in modern man. The results are consistent with the hypothesis that interracial differences in human skull dimensions are a simple consequence of random drift and mutation.

Received January 19, 1988. Accepted August 5, 1988

Phylogenetic trees are essentially hypotheses, and as such, the predictions they make depend not only on the accuracy of information used to construct them but also on the assumptions about evolutionary processes that underlie the tree-fitting procedure. When such assumptions are not explicitly stated, as is often the case in numerical taxonomy, the resultant tree is a classification scheme, not a phylogenetic hypothesis. Because evolution is probabilistic, it is highly desirable that the statistical properties of phylogenetic hypotheses be known. This subject has received substantial attention (Edwards and Cavalli-Sforza, 1964; Neyman, 1971; Cavalli-Sforza and Piazza, 1975; Thompson, 1975; Cavender, 1978, 1981; Tateno et al., 1982; Felsenstein, 1983, 1985a; Nei et al., 1983, 1985; Templeton, 1983, 1985; Penny and Hendy, 1986).

Since phylogenetic diversification is a genetic phenomenon, attempts to reconstruct phylogenies from information on extant

species ought to rely on rules that are consistent with population-genetic properties. In the case of gene-frequency data, distance statistics with an explicit relationship to population-genetic theory can often be constructed. Nei's (1972) genetic distance,  $D$ , which is equivalent to the average number of gene substitutions per locus, is an example of one such measure. The evolutionary and statistical properties of  $D$  are fairly well understood for the case in which the dynamics of gene-frequency change are due solely to mutation and random genetic drift (Kimura, 1983). When these conditions are met, the expected equilibrium rate of gene substitution per locus per generation is equal to the mutation rate. The expected value of  $D$  will therefore increase linearly with time, providing a basis for estimating branch lengths as well as branch points in a phylogeny. At least one other mechanism, fluctuating selection, could lead to the existence of a stochastic molecular clock. Its implications for the reconstruction of phyloge-

nies have been evaluated only recently (J. Felsenstein, unpubl.).

Despite the burgeoning availability of molecular biological techniques, morphological data still play a central role in the development of phylogenetic hypotheses and are unlikely ever to be supplanted entirely. Unlike isozymes and DNA, morphological characters are often permanently preserved in the fossil record or as museum specimens. Since morphometric traits are generally influenced by a large number of loci, morphological analysis also has the advantage of integrating information over a large portion of the genome, although too much should not be made of this (Rogers and Harpending, 1983; Felsenstein, 1986). Moreover, while nucleotides and amino acids are constrained to discrete character states, metric traits may vary continuously over a very wide range of phenotypes.

There are serious limitations to morphological data as well. Unlike nucleotide and amino acid sequences, morphological expression is often highly dependent upon the environmental setting and on age. Since selection operates directly on phenotypes and not genotypes, it is also reasonable to expect the evolution of mean phenotypes to be less clocklike than that of molecules. Stabilizing selection and character displacement can result in unexpectedly low or high rates of divergence for extended periods of time. These kinds of problems have not discouraged many investigators from drawing evolutionary inferences from morphological data, but naturally, the question as to how reliable such data are for reconstructing phylogenies has arisen repeatedly. This debate has spawned a large number of tree-fitting strategies, ranging from clustering and divisive techniques to the evaluation of character-state matrices from the standpoint of certain parsimony or compatibility philosophies (reviewed in Felsenstein [1982]).

Recently, several attempts have been made to determine which tree-fitting procedure is most efficient in recovering correct phylogenies. The results have been rather discouraging for all methods except in the case of low numbers of species with distant branch points (Kidd et al., 1974; Astolfi et al., 1981; Tateno et al., 1982; Nei et al.,

1983; Fiala and Sokal, 1985; Tateno and Tajima, 1986). It is even less clear what phylogenetic analysis can tell us about the mechanisms of diversification in isolated lineages. Most of the existing computational methods are only loosely defined in terms of evolutionary processes. Parsimony, for example, is a particular viewpoint concerning the expected evolutionary trajectories of character states, but it is not a viewpoint that can easily be translated into population-genetic terms. Cavalli-Sforza and Edwards (1965, 1967) and Cavalli-Sforza and Piazza (1975) outlined a Brownian-motion model for the estimation of evolutionary trees, which in principle can be used to test the hypothesis that the diversification of a phylogeny is a simple consequence of random genetic drift (Felsenstein, 1973). Unfortunately, the Brownian-motion model is almost never utilized by those directly involved in phylogeny estimation.

In this paper, the Brownian-motion model is explicitly defined from a quantitative-genetic perspective under the assumption that mutation and random genetic drift are the only operable evolutionary forces. This is a somewhat different approach than that taken by Cavalli-Sforza, Edwards, and Piazza, who chose to treat mutation as having a negligible influence on the similarity of species. A drift-mutation model is the simplest possible evolutionary scenario for the phenotypic divergence in a phylogeny. It warrants serious consideration, since systematists often focus their attention on characters that *appear* to have little adaptive significance. Following the presentation of the basic theory and methodology, an example of its application will be illustrated with a large data set on cranial variation in modern man.

### *Theoretical Background*

*The Neutral Model of Phenotypic Evolution.*—Formulations from quantitative genetics provide a fundamental basis for reconstructing phylogenies under the assumption of neutral phenotypic variation. Although this theory applies to reproductively isolated populations, races, or species, I will refer to species throughout for the sake of consistency. When a base population is split into two or more reproductively iso-

lated species, the mean phenotypes of the latter will gradually diverge via random genetic drift. Although the initial divergence will be caused largely by the random fixation of alleles that existed in the base population, this has limits. Provided the effective population size ( $N$ ) is fairly large, as it should be when an entire species is considered, then the expected between-species variance resulting from  $t$  generations of drift of genes from the base population is very close to

$$\sigma_b^2(\bar{z}, t) = 2\sigma_w^2(0) \left[ 1 - \exp\left(\frac{-t}{2N}\right) \right] \quad (1)$$

where  $\sigma_w^2(0)$  is the additive genetic variance in the base population (Wright, 1951; Crow and Kimura, 1970). The maximum between-species variance that can arise from initial variation is therefore  $2\sigma_w^2(0)$ .

Additional divergence arises from the random fixation of mutations that appear subsequent to the isolation event. For mutations with additive effects, the between-species variance resulting from this process is approximately

$$\sigma_m^2(\bar{z}, t) = 4NV_m \left\{ \frac{t}{2N} - \left[ 1 - \exp\left(\frac{-t}{2N}\right) \right] \right\} \quad (2)$$

where  $V_m$  is a measure of the mutational rate of input of genetic variance per population per generation (see appendix by Dempster in Bailey [1959]; Lande, 1975; Chakraborty and Nei, 1982; Lynch and Hill, 1986).

The expected level of within-population genetic variance for a neutral character with an additive genetic basis is  $2NV_m$  (Clayton and Robertson, 1955), and for moderate levels of dominance the expectation is not much different (Lynch and Hill, 1986). Substituting  $2NV_m$  for  $\sigma_w^2(0)$  in Equation (1), and summing Equations (1) and (2), the total amount of divergence after  $t$  generations of isolation is found to be simply  $\sigma^2(\bar{z}, t) = 2V_m t$ . Figure 1 shows that the proportional divergence that results from mutations subsequent to an isolation event is a function of  $t/N$ . If the number of generations of divergence is on the order of  $N$  or greater, a

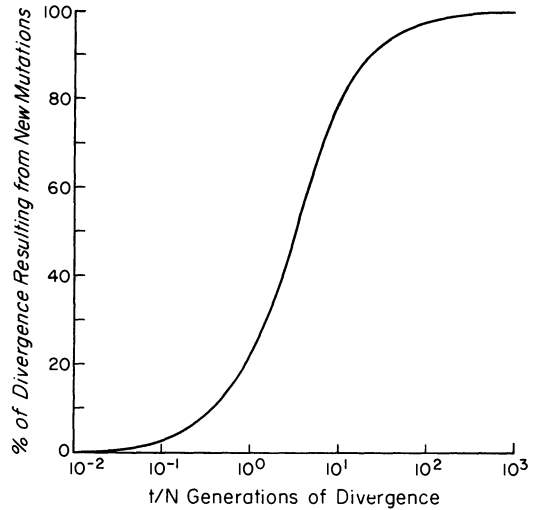


FIG. 1. The expected proportion of the between-species variance resulting from mutational variance arising subsequent to an isolation event.  $N$  is the effective population size, and  $t$  is the number of generations of divergence. Thus, when  $t = N$  generations, the contribution from new mutations is approximately 20%.

substantial amount of between-species variance will be attributable to new mutation.

The expected neutral divergence rate of  $2V_m$  has several useful properties. First, the buildup of the between-species variance is expected to be linear and independent of the sizes of the base and propagule populations. Independence from population size ( $N$ ) arises, as in the neutral theory of molecular evolution (Kimura, 1983), for the simple reason that the number of mutations arising at any locus is proportional to  $2N$  while the probability of fixation of a neutral mutation is its initial frequency,  $1/2N$ . Second, although the asymptotic divergence rate  $2V_m$  was first derived under the assumption of additivity of mutational effects, this is now known to be an unnecessary restriction (Lynch and Hill, 1986). The result applies regardless of the level of dominance and is independent of the degree of linkage of polygenes underlying quantitative traits. This is because both dominance and linkage disequilibrium are transient properties of mutant alleles en route to loss or fixation. Third, the result applies regardless of the form of the distribution of mutational effects.

The between-species variance is an appropriate morphological distance measure

for constructing phylogenetic trees under the neutral hypothesis. There are two reasons for this. First, as noted above, the expectation  $\sigma^2(\bar{z}, t)$  is proportional to divergence time. Second, the divergence along different branches is completely independent and additive (cf. Cavalli-Sforza and Piazza, 1975). The use of the between-species variance as a distance measure is of obvious appeal, since it is readily extracted from the mean squares of a one-way analysis of variance. When multiple characters have been measured, a logical estimate of interspecific distance would be the average of the between-species components of variance for all the characters. Although  $V_m$  may vary extensively between characters (Lynch, 1988), the pooled distance measure for  $k$  characters will still have an expected value

$$2t \sum_{i=1}^k \frac{V_{m_i}}{k}$$

that is proportional to divergence time.

This approach can be generalized further to include between-species covariances for quantitative traits. The expected increment in the between-species component of covariance for neutral characters is  $2V_{m_{ij}}$  where  $V_{m_{ij}}$  is the mutational rate of input of genetic covariance between characters  $i$  and  $j$  (Lande, 1979; Lynch and Hill, 1986). Thus, a generalized distance measure in applications of the neutral theory would be

$$\sum_{i=1}^k \sum_{j=i+1}^k \left( \frac{|C_{ij}|}{\frac{k(k+1)}{2}} \right)$$

where  $|C_{ij}|$  refers to the absolute value of  $ij$ th element of the between-species variance-covariance matrix. In principle, this measure could be refined further by use of appropriate weights for the  $C_{ij}$ , such as the inverses of the within-population genetic variances, but such information is not usually available in phylogenetic studies.

While the expected between-species variance  $\sigma^2(\bar{z}, t)$ , for a neutral character is asymptotically  $2V_m t$  per generation under a very wide range of conditions, realized divergences,  $\text{Var}(\bar{z}, t)$ , will always deviate somewhat from this value. The variable ef-

fects of mutations, their random origin, and independent fixation in different populations can all result in considerable variance in the amount of divergence. The case for purely additive mutations has been worked out in Lynch and Hill (1986), where it is shown that when the distribution of mutational effects is normal with mean zero,  $\sigma^2[\text{Var}(\bar{z}, t)] \approx 2(2V_m t)^2$ . Under these conditions, the realized divergence will be distributed as chi-square (Freund, 1971 pp. 212–216) with a coefficient of variation of approximately  $\sqrt{2}$ . This is the expected variance in  $\text{Var}(\bar{z}, t)$  due to the random occurrence of mutations alone and does not include the additional and often substantial variance due to sampling error by the investigator.

*Phylogenetic Information From Neutral Quantitative Traits.*—The process of phenotypic divergence along the branches of a phylogeny can be most easily envisioned by considering a three-species tree (Fig. 2). In this case, there is only one possible topology: a pair of neighbor species joined to an out-group. Two isolation events are implicit in the illustrated tree, with 4 being the most recent common ancestor of 1 and 2 and with 5 being the root of the tree. In the following, it is assumed that samples of species 1, 2, and 3 are contemporaneous but that species 4 and 5 are unobserved, as is usually the case. Let the mean phenotype of species  $i$  be denoted by  $\mu_i$ . The expected mutational distance for species pair  $(i, j)$  is  $T(i, j) = V_m t'(i, j)$ , where  $t'(i, j)$  is the total time along the branch segments between  $i$  and  $j$ . Further, let  $D(i, j)$  be the actual amount of between-species variance that develops over this period, and let  $D'(i, j)$  be the between-species variance recorded by the investigator. The difference between the expected and actual distances is a function of the random occurrence and fixation of mutations. The difference between the actual and observed distances is a function of the sampling variance.

Consider an arbitrary pair of species  $(i, j)$ . The distance that develops between these species can be defined in terms of a series of independent evolutionary changes. For example, the distance that develops between species 1 and 3 in Figure 2 is a function of the events that occur over branch

segments (3, 5), (4, 5), and (1, 4). The events within these segments are independent because nodes 4 and 5 represent isolation events. Let the actual evolutionary change that occurs along the  $x$ th segment between species  $i$  and  $j$  be  $L_x(i, j) = \mu_{x_1} - \mu_{x_2}$  where  $x_2$  is always the most distant species on the segment from species  $i$ . For species pair (1, 3), for example,  $L_1(1, 3) = \mu_1 - \mu_4$ ,  $L_2(1, 3) = \mu_4 - \mu_5$ , and  $L_3(1, 3) = \mu_5 - \mu_3$ . With  $n$  branching segments, the total evolutionary change between species  $i$  and  $j$  is therefore

$$L(i, j) = \sum_{x=1}^n L_x(i, j) = \mu_i - \mu_j.$$

With a three-species tree, there are four branch segments, so four independent evolutionary changes must be considered. Under the assumption that the effects of mutations are normally distributed with mean zero, which will be adhered to below, each of these changes will be normally distributed with mean zero and variance  $2T_x(i, j)$ . The latter follows from the observation that  $\sigma^2[L_x(i, j)] = E[L_x^2(i, j)] = E[(\mu_{x_1} - \mu_{x_2})^2]$ , this final quantity being twice the expected variance between species  $x_1$  and  $x_2$ .

In terms of the  $n$  branch segments between species  $i$  and  $j$ , the distance is defined to be

$$D(i, j) = \frac{\left[ \sum_{x=1}^n L_x(i, j) \right]^2}{2}.$$

The expected distance is therefore

$$\begin{aligned} T(i, j) &= \sum_{x=1}^n \frac{E[L_x^2(i, j)]}{2} \\ &= \sum_{x=1}^n T_x(i, j) \end{aligned}$$

which implies additivity of the distances over the branch segments from  $i$  to  $j$ . On the other hand, the variance of distance  $D(i, j)$  is not a linear function of the variances of the individual branch segments, but

$$\begin{aligned} \sigma^2[D(i, j)] &= \sigma^2 \left\{ \frac{\left[ \sum_{x=1}^n L_x(i, j) \right]^2}{2} \right\} \\ &= 2 \left[ \sum_{x=1}^n T_x(i, j) \right]^2. \end{aligned}$$

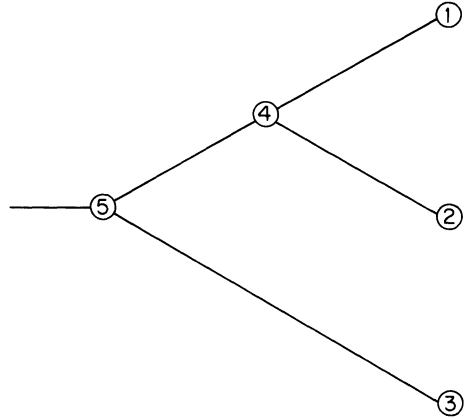


FIG. 2. A hypothetical phylogeny with three extant species (1, 2, and 3) and two ancestral species (4 and 5).

Thus, the distance between any two species on the tree is distributed as chi-square with expectation  $T(i, j)$  and variance  $2T^2(i, j)$ .

The problem of sampling aside, the random occurrence and fixation of mutations impose an inherent limitation on the probability of recovering the proper phylogenetic topology and branch lengths from extant species. The fundamental problem is again most easily envisioned by reference to a three-species tree (Fig. 2). In the illustrated example, although  $T(1, 3) = T(2, 3) > T(1, 2)$ , the possibility exists that the realized distances will be inconsistent with the expectations, i.e.,  $D(1, 3) < D(1, 2)$  or  $D(2, 3) < D(1, 2)$ . The fulfillment of either of these conditions would encourage the post facto recognition of species 1 or 2, rather than 3, as the out-group. On the other hand, if both of these inequalities were met, but  $D(1, 3) \approx D(2, 3)$ , then the best-fit tree would retain species 3 as an out-group while assigning a negative value to the fitted distances  $\hat{D}(3, 5)$  or  $\hat{D}(4, 5)$ . An example of a perfectly additive tree with a negative branch segment is given in Figure 3.

While the objection may be raised that such a tree is unrealistic, since a variance cannot be negative, it must be emphasized that the  $\hat{D}(i, j)$  are estimates of the expected distances  $T(i, j)$  and, as such, are not constrained to be positive. The arbitrary replacement of negative  $\hat{D}(i, j)$  by zeros, as is often done in distance Wagner procedures (Farris, 1972) will result in upwardly biased

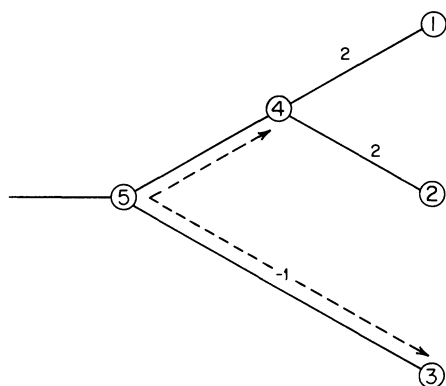


FIG. 3. An example of a purely additive tree with a negative distance (-1) fit between species 3 and 4. Such a tree would arise if  $D(1, 2) = 4$  and  $D(1, 3) = D(2, 3) = 1$ .

distance estimates (Kidd et al., 1974; Tatenno et al., 1982; Tatenno and Tajima, 1986), although it might not influence the likelihood of recovering the correct topology.

The ability to recover correctly a topology depends on the relative position of the internal nodes. Returning to Figure 2, it can be seen that the condition  $D(1, 3) < D(1, 2)$ , which by symmetry is the same as  $D(2, 3) < D(1, 2)$ , requires that  $[D(4, 5) + D(3, 5)] < D(2, 4)$ . Figure 4 shows that the probability of this occurring declines from 0.5 to 0.0 as node 4 approaches the branch tip. The decline is rather slow, however. For any single trait, the probability of the realized  $D(1, 3)$  being less than  $D(1, 2)$  is greater than 0.1, unless the expected position of node 4 is extremely close to the branch tip. Thus, for most phylogenies, inferences drawn from single traits have very high probabilities of being misinformative. This, perhaps, comes as no surprise. However, a substantial problem remains even when the distances are averaged over a large number of characters. With 100 measured traits, for example, if  $T(1, 4)/T(3, 4) > 0.8$ , then the probability of  $D(1, 3) < D(1, 2)$  will exceed 0.1.

*Additional Error Due to Sampling.*—With finite sample sizes, the observed mean phenotypes ( $\bar{z}_i$ ) will generally deviate from the parameters ( $\mu_i$ ). However, an unbiased estimator of  $D(i, j) = (\mu_i - \mu_j)^2/2$  is obtainable from a simple one-way analysis of variance.

Consider the ideal situation in which two taxa have equal within-species variance,  $\sigma_p^2$ ,

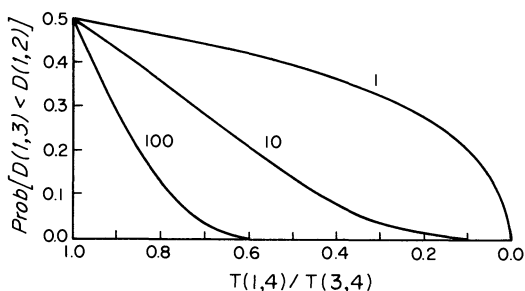


FIG. 4. The probability of  $D(1, 3) < D(1, 2)$  when species 3 is the true out-group as in Figure 2 and when  $D(i, j)$  is averaged over 1, 10, or 100 characters. The abscissa is a measure of the relative position of ancestral species 4,  $T(1, 4)/T(3, 4) = 1$  implying a trifurcation.

and the experimental design is balanced. The within-species mean square is

$$MS_w = \frac{\sum_{k=1}^n (z_{ik} - \bar{z}_i)^2 + \sum_{k=1}^n (z_{jk} - \bar{z}_j)^2}{2(n-1)}$$

where  $z_{xk}$  represents the phenotype for the  $k$ th individual of the  $x$ th species,  $\bar{z}_x$  is the sample mean of the  $x$ th species, and  $n$  is the sample size.  $MS_w$  will be distributed as chi-square with expectation  $\sigma_p^2$  and variance  $\sigma_p^4/n$ . With only two species, the between-species mean square is simply

$$MS_b = \frac{(\bar{z}_i - \bar{z}_j)^2}{2}$$

with expectation  $\sigma_p^2 + nD(i, j)$ . Thus, an unbiased estimate of  $D(i, j)$  is

$$D'(i, j) = \frac{MS_b - MS_w}{n}$$

The form of the distribution of  $MS_b$  is complicated, but its variance is

$$\sigma^2(MS_b) = 4n\sigma_p^2 \left[ \frac{\sigma_p^2}{n} + D(i, j) \right]$$

Since the mean squares are independently distributed,

$$\begin{aligned} \sigma^2[D'(i, j)] &= \frac{\text{Var}(MS_w) + \text{Var}(MS_b)}{n^2} \\ &\approx \frac{4\sigma_p^2}{n} \left[ \frac{\sigma_p^2}{n} + D(i, j) \right]. \end{aligned}$$

Thus, the sampling variance of the between-species variance depends largely on the

sampling variance of the mean phenotypes,  $\sigma_p^2/n$ . The coefficient of variation of  $D'(i, j)$  is  $[4\phi(\phi + 1)]^{1/2}$ , where  $\phi = \sigma_p^2/nD(i, j)$ . For  $\phi = 0.001, 0.01, 0.1, 1,$  and  $10$ , this takes on values of  $0.06, 0.20, 0.66, 2.83,$  and  $20.98$ . Clearly, unless sample sizes for interspecific distances are large enough so that the sampling variance of the mean is much less than the distance, say 1% or less, estimates of  $D(i, j)$  obtained from single characters are likely to be highly inaccurate. This source of error can be further reduced by pooling estimates of  $D(i, j)$  over several characters, since  $\sigma^2[D'(i, j)]$  will then be inversely proportional to the number of characters.

*Potential Utility of Neutral Quantitative Traits for Recovering Phylogenies*

Special emphasis has been put on the random occurrence of mutations and the incomplete sampling of populations because they produce inaccuracies in the estimates of the true distances between species. In order to ascertain the implications of these sources of error for the ability to reconstruct phylogenies from observed mean phenotypes, it is necessary to rely on computer simulation. Starting with a phylogenetic tree with known expected distances,  $T(i, j)$ , and the theoretical sampling distributions outlined above, random sets of observed distances,  $D'(i, j)$ , can be generated. These observed values can then be used to construct an estimated tree for which the fitted values,  $\hat{D}(i, j)$ , are most compatible with the observations. By repeating this procedure many times, it becomes possible to evaluate the probability of recovering the correct tree topology as well as to determine the variance of the fitted branch lengths.

Many iterative procedures exist for the construction of phylogenetic trees from distance data (Edwards and Cavalli-Sforza, 1965; Cavalli-Sforza and Edwards, 1967; Fitch and Margoliash, 1967; Farris, 1972; Sneath and Sokal, 1973; Li, 1981). Generally, there is no guarantee that such procedures will identify the best-fit tree, since they do not examine all possible topologies. Such an examination is impractical with large numbers of taxa. For example, with only eight species, there are 35,280 possible topologies. In this study, the simulated trees

contained no more than five species, and it was possible to obtain best-fits to all possible tree topologies (3, 15, and 105 for 3, 4, and 5 species) by the least-squares method of Chakraborty (1977). Under the assumption of constant expected evolutionary rates, this method is equivalent to Felsenstein's (1973) restricted maximum-likelihood procedure. The final tree was considered to be the one that minimized the average percentage standard deviation,

$$APSD = \left\{ \sum_{i=1}^{n-1} \sum_{j=i+1}^n \frac{[D'(i, j) - \hat{D}(i, j)]^2}{D'(i, j)} \right\} \times 100. \\ \left\{ \frac{n(n-1)}{2} - 1 \right\}^{1/2}$$

Some disagreement exists over the use of 2 rather than 1 or 0 as the power of  $D'(i, j)$  in this expression, but since the variance of the distance is proportional to the expected distance squared under the neutral theory, the use of 2 seems to be justified.

Under the neutral theory, the expected distances from the root of the tree to all of the branch tips are equal, a constraint that is incorporated into the fitting procedure. For example, with a three-species analysis, there are only three possible trees: 1, 2, or 3 being the out-group. For the tree in which species 3 is taken to be the out-group,  $\hat{D}(1, 2) = D'(1, 2)$  and  $\hat{D}(1, 3) = \hat{D}(2, 3) = [D'(1, 3) + D'(2, 3)]/2$ . The root is then a distance  $\hat{D}(1, 3)/2$  back from species 3.

All of the trees studied were replicated to a high degree (at least 1,000 replications per starting condition). Algorithms for generating random normal and random chi-square variates were taken from Knuth (1969), and the distributions of all simulated  $D'(i, j)$  were found to be in close accordance with the theoretical expectations outlined above. With a high degree of replication, the mean fitted distances always converged on the expected distances,  $T(i, j)$ .

Before proceeding to the results, the assumptions underlying the simulated process of evolutionary divergence should be restated. 1) All measured characters are assumed to be free of the forces of selection and to have an expected divergence rate of  $2V_m$  per

TABLE 1. Summary statistics for least-squares fits to three-species trees when characters are assumed to diverge via drift and mutation. Species 3 is assumed to be the outgroup.  $T(1, 2)$  is the expected distance between species 1 and 2, and  $\sigma_p^2/n$  is the sampling variance of the mean phenotypes.  $T(1, 3)$  (the expected distance between species 1 and 3) and  $T(2, 3)$  (the expected distance between species 2 and 3) are assumed to equal 2. Two thousand simulations were performed for each set of conditions. CL refers to the upper 5 and 1% confidence limits for the APSD (average percentage standard deviation) under the neutral hypothesis. Negative fit implies that the fitted distance  $\hat{D}(4, 5) < 0$ .

$T(1, 2)$	Number of traits	$\frac{\sigma_p^2}{n} = 0$					$\frac{\sigma_p^2}{n} = 0.5$				
		APSD			Percentage wrong topology	Percentage negative fits	APSD			Percentage wrong topology	Percentage negative fits
		$\bar{x}$	SD	CL			$\bar{x}$	SD	CL		
0.5	10	11.5	9.6	30, 41	8.4	7.8	17.3	17.1	45, 80	19.6	17.4
	25	8.0	6.3	19, 26	1.0	0.8	11.6	9.7	30, 39	6.6	6.4
	50	5.6	4.4	14, 18	0.0	0.0	8.5	6.7	21, 31	0.6	0.6
	100	3.9	3.0	9, 13	0.0	0.0	5.8	4.6	14, 21	0.0	0.0
1.0	10	11.9	10.1	32, 48	36.5	30.5	14.4	14.3	41, 63	44.6	35.0
	25	8.7	6.7	21, 29	19.8	18.6	10.2	8.2	25, 37	27.2	24.3
	50	7.2	5.3	16, 22	6.5	6.3	8.4	6.2	19, 25	15.4	14.4
	100	5.3	4.2	13, 17	1.4	1.4	6.9	5.0	16, 21	3.8	3.8
1.5	10	10.0	8.9	28, 41	59.0	43.6	12.2	12.1	32, 55	58.9	42.4
	25	6.6	5.4	17, 23	52.7	41.4	7.4	6.3	19, 28	55.3	42.0
	50	4.7	3.7	13, 17	45.6	39.2	6.0	5.0	16, 21	51.0	41.3
	100	4.4	3.4	11, 14	33.2	30.2	4.9	3.8	11, 16	39.4	34.5

generation. 2)  $V_m$  is taken to be constant across species as well as across characters, the latter being simply a matter of scale. 3) The distribution of mutational effects is assumed to be normal with mean zero. 4) Genotype  $\times$  environment interaction is assumed to be negligible, so that estimates of  $D(i, j)$  are purely a function of genetic differences.

Three things stand out in the results for three-species trees (Table 1). First, the probability of recovering the correct topology is highly dependent on the position of the internal node (4 in Fig. 2). If this is halfway along the branch [ $T(1, 2) = 1.0$  in Table 1], then roughly 40 characters would have to be measured without sampling error for the probability of establishing the correct topology to exceed 90%. If the sampling variance is 25% of  $T(1, 3)$ , then 60 traits would have to be measured to attain the same degree of precision. If the internal node is such that  $T(1, 2) = 0.75T(1, 3)$ , then the probability of recovering the correct topology is less than 70%, even if 100 characters are measured with no sampling error.

A second point of interest is that the frequency with which fitted trees with negative evolutionary change arise is on the order of the probability of recovering the wrong to-

polo. Thus, unless the internal node is close to the branch tips and unless very many characters are measured, apparent convergence will arise rather frequently in collections of characters that evolve by the joint interaction of drift and mutation.

Third, while the APSD provides a reasonable criterion for the best-fit tree, it provides little, if any, information on whether the best-fit tree is the correct one. In general, the mode of the distribution of APSD is less than 5%, with a long tail to the right (Fig. 5), so that the upper 5% confidence limit is on the order of 10–30%. Figure 5 provides two examples in which the distributions of APSD are quite similar, while the frequencies of obtaining the wrong topology are very different.

As the number of species in an analysis increases, the average distance between nodes in the tree must decline, and consequently the likelihood of establishing the correct phylogeny diminishes. Table 2 summarizes some results of simulations for four- and five-species trees with topologies of the forms given in Figure 6. Even if the means of 50 characters are measured with no error, the probability of recovering the wrong topology often exceeds 30%, and the frequency with which trees with negative branch

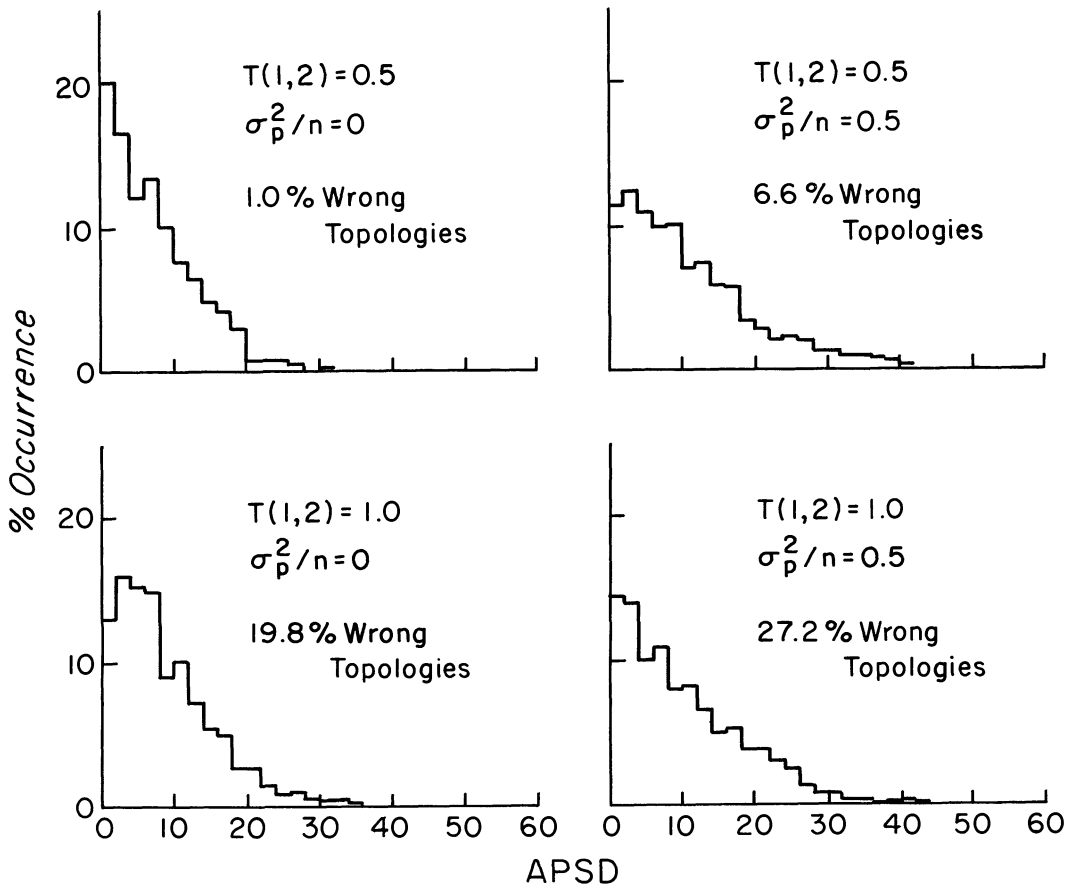


FIG. 5. Frequency distributions for the APSD obtained from 2,000 replicates for four of the simulations given in Table 1. Each case is based on 25 independent characters.

lengths arise is of the same order. Again, the APSD provides little insight into whether the correct topology has been recovered.

*Analysis of Three Major Races of Man*

One of the most extensive collections of data on quantitative characters for any

species is Howells' (1970, 1973a, 1973b) survey of metric cranial characters in museum specimens for 31 well-defined populations of man. Sample sizes for each population are on the order of 50 for each sex, and Howells performed 61 measurements on each cranium. I have used a subset of

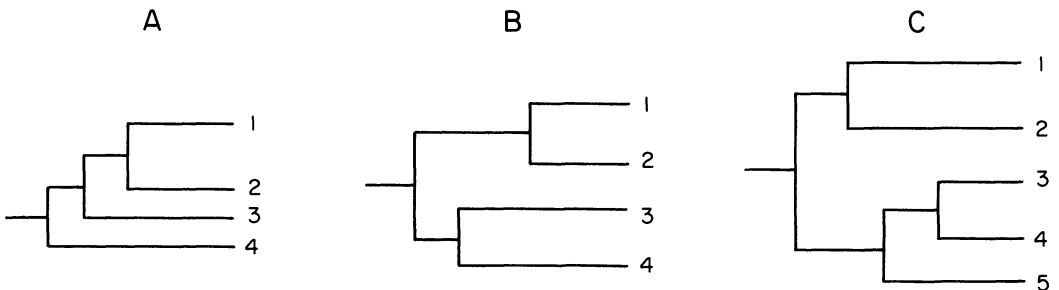


FIG. 6. Forms of the three tree topologies referred to in Table 2.

TABLE 2. Summary statistics (as in Table 1) for some four- and five-species trees with topologies of the forms in Figure 6. The expected distance between species 1 and 4 is 2.0 in all cases, and the sampling variance is zero. One thousand simulations were performed for each set of conditions.

Topol- ogy	Expected distances				10 characters						50 characters					
					APSD			Per- centage wrong topolo- gies	Per- centage negative fits	APSD			Per- centage wrong topolo- gies	Per- centage negative fits		
	$\bar{x}$	SD	CL	$\bar{x}$	SD	CL										
A	0.5	1.0	2.0	—	16.0	8.0	30, 40	57.3	49.7	8.7	3.9	15, 19	11.7	11.0		
	0.5	1.5	2.0	—	16.0	7.9	30, 40	68.1	55.0	7.7	3.5	14, 16	36.4	32.4		
	1.0	1.5	2.0	—	14.9	7.7	28, 35	83.0	71.8	8.3	3.6	15, 17	57.9	52.2		
B	0.5	2.0	1.0	—	17.4	9.4	35, 44	44.1	39.6	9.0	4.4	17, 21	8.8	8.8		
	0.5	2.0	1.5	—	16.2	8.5	33, 42	64.8	51.9	7.3	3.6	14, 17	47.3	40.5		
	1.0	2.0	1.5	—	15.6	8.6	31, 41	76.9	67.6	8.9	4.0	16, 19	53.2	48.1		
C	0.5	2.0	0.5	1.0	18.7	8.5	34, 45	62.5	57.1	9.6	4.0	16, 21	10.9	10.8		
	0.5	2.0	1.0	1.5	16.9	6.8	29, 35	83.3	71.5	9.5	3.7	16, 20	63.1	56.4		
	1.0	2.0	0.5	1.0	17.0	7.4	31, 39	76.0	70.8	10.1	4.0	17, 21	21.4	21.1		
	1.0	2.0	1.0	1.5	16.6	6.8	29, 39	88.1	79.0	9.9	3.8	16, 19	62.9	58.5		

the total data collection (Table 3) to construct a phylogenetic hypothesis for the evolutionary relationship among the Caucasian, Mongoloid, and Negroid races and to provide a rough test as to whether the divergence in cranial morphology among these three major groups is consistent with the neutral hypothesis. In total, 76,433 measurements have been employed in this analysis. All of the sample populations were geographically or temporally distributed such

TABLE 3. Populations and sample sizes for the data set used to construct an evolutionary tree for the Caucasian, Negroid, and Mongoloid races of man (from Howells [1973a, pers. comm.]). Detailed descriptions of the sources of Caucasian and Negroid skulls are given in Howells (1973a).

Population	Males	Females
Caucasians:		
Medieval Norse (Oslo)	55	55
Medieval Hungary (Zalavar)	54	45
Berg (Carinthia, Austria)	56	53
Egypt (Gizeh Dynasties)	58	53
Negroid:		
Tetia (Kenya)	34	49
Dogon (Mali)	48	53
Zulu (South Africa)	55	47
Bushmen (South Africa)	41	49
Andaman Islands	26	28
Mongoloid:		
North Japan	55	32
South Japan	50	41
Hainan Chinese	45	38
Atayal (Taiwan)	29	18
Ainu	48	38

that recent interracial gene flow seems very unlikely (Howells 1970, 1973a, 1973b), although it cannot be ruled out entirely.

The distance between each pair of populations in Table 3 was estimated by the mean of the between-population variances computed for each of the 61 characters. Prior to this pooling of the data, each character was given equal weighting by adjusting the mean of the pairwise between-population variances to one. Such a treatment does not account for the likely nonindependence of the 61 characters. Unfortunately, the raw data necessary for the computation of the between-population genetic covariances were not directly available to me. As noted above, however, because the absolute values of all of the elements of the between-population variance-covariance matrix are expected to increase linearly with time under the neutral hypothesis, the confinement of the analysis to the diagonal elements should not bias the relative measures of distance.

The average distance between Caucasians and Mongoloids (0.78) is less than that between Caucasians and Negroids (1.18) and Mongoloids and Negroids (1.32), the difference between the latter two distances being nonsignificant (Table 4). The least-squares tree is given in Figure 7 (upper left). There are two ways to test for the significance of the branching order. First, after fitting trees to all possible permutations of the sample populations (4 Caucasians, 5

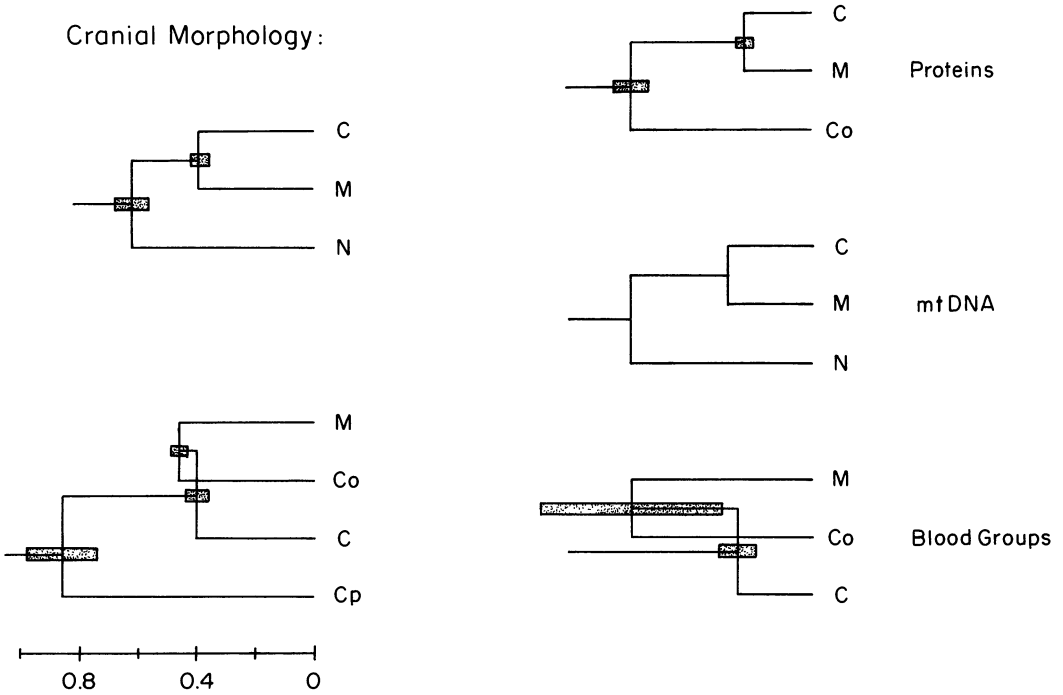


FIG. 7. Left) Least-squares trees fit to average between-population variances for cranial morphology in major races of man (C = Caucasians, Co = Congoids, Cp = Capoids, M = Mongoloids, N = Negroids). Bars denote standard errors. The inward branching at the upper node of the lower tree denotes a negative branch segment. Right) Least-squares trees fit to genetic distance data for proteins and blood groups (Nei and Roychoudhury, 1981) and mtDNA (Cann et al., 1987).

Mongoloids, 5 Negroids), a chi-square test can be used to test for heterogeneity among the observed numbers of the three tree types. Of the 100 trees generated, 48 identified Ne-

groids as the out-group, 27 identified Caucasians as the out-group, and 25 identified Mongoloids as the out-group. This yields  $X^2 = 9.65$ , which with one degree of free-

TABLE 4. Summary statistics for the observed ( $D'$ ) and fitted ( $\hat{D}$ ) morphological distances for analyses of three and four major races of man. The degrees of freedom used to estimate the standard errors of  $D'$  were taken to be  $(N_i - 1)(N_j - 1)$ , where  $N_i$  and  $N_j$  are the numbers of populations for races  $i$  and  $j$ . The variance of the distance expected under the neutral model was approximated by  $2(\hat{D} + \hat{\sigma}_p^2)^2/(n - 1)$ , where  $\hat{\sigma}_p^2$  is the observed sampling variance of the means within populations and  $n = 61$  is the number of characters; its associated standard error is approximately  $[\{2(\hat{D} + V_w)^2/(n - 1)\}/(2/d.f.)]^{1/2}$ .

Races	Observed distance			Variance of between-population distance			Fitted distance	
	$\bar{x}$	SE	d.f.	Observed	Expected	SE	$\hat{D}$	SE
Three races:								
Caucasian-Mongoloid	0.78	0.05	12	0.027	0.022	0.009	0.78	0.05
Caucasian-Negroid	1.18	0.14	16	0.218	0.055	0.019	1.25	0.12
Mongoloid-Negroid	1.32	0.17	16	0.471	0.055	0.019	1.25	0.12
Four races:								
Caucasian-Mongoloid	0.78	0.05	12	0.027	0.024	0.010	0.81	0.05
Caucasian-Congoid	0.84	0.09	4	0.035	0.024	0.017	0.81	0.05
Caucasian-Capoid	1.68	0.18	2	0.065	0.101	0.101	1.71	0.24
Mongoloid-Congoid	0.93	0.08	8	0.052	0.031	0.015	0.93	0.08
Mongoloid-Capoid	1.91	0.35	4	0.513	0.101	0.072	1.71	0.24
Congoid-Capoid	1.53	0.35	2	0.243	0.101	0.101	1.71	0.24

dom is significant at the 0.005 level. Because the populations within the three major groups are not entirely independent, however, this level of significance is inflated to an unknown degree (Felsenstein, 1985*b*).

A second approach is to follow the logic of Chakraborty (1977) and Nei et al. (1985) and compute the standard errors of the branch points. The standard error of the node between Caucasians and Orientals is simply half the standard error of their  $D'$ , i.e., 0.025 (Table 4). Since the fitted distances  $\hat{D}(C, N) = \hat{D}(M, N) = [D'(C, N) + D'(M, N)]/2$ , the variance of the node distance between Negroids and Caucasians-Mongoloids is  $\{\text{Var}[D'(C, N)] + \text{Var}[D'(M, N)] + 2\text{Cov}[D'(C, N), D'(M, N)]\}/16$ . The variances of the distances are the squares of the standard errors given in Table 4, while the covariance should be approximately  $[(1.25 - 0.39)/1.25]\{\text{Var}[D'(C, N)] + \text{Var}[D'(M, N)]\}/2$ , which is the variance of the divergence between Negroids and the node between Caucasians and Mongoloids. The standard error of the branch point at the base of the tree is therefore 0.07, providing further evidence that the distance between the two nodes is significant.

The next question of interest is whether the observed data support the hypothesis that interracial variation in cranial morphology is a simple product of drift and mutation. Two observations are consistent with the neutral model: 1) the approximate equality of the mean Caucasian-Negroid and Mongoloid-Negroid divergences and 2) the close agreement between the observed and expected variance in estimated distances between Caucasian and Mongoloid populations (Table 4). On the other hand, the Caucasian-Negroid and Mongoloid-Negroid distances are four and nine times more variable than expected under the neutral model. Moreover, assuming  $\hat{D}(C, M) = 0.78$  and  $\hat{D}(C, N) = \hat{D}(M, N) = 1.25$  to be the true divergences, the neutral model predicts that the "correct" topology (Negroids as the out-group) should arise 80% of the time, given the observed number of characters and sample sizes. However, as noted above, this topology only arose in 48 of 100 comparisons.

Thus, although the divergence in cranial morphology between Caucasians and Mon-

goloids is quite compatible with the neutral hypothesis, the excessive variation in the divergences between Caucasians-Mongoloids and Negroids is inconsistent with the same hypothesis. Figure 8, a Fitch-Margoliash tree fit to the 14 populations using Felsenstein's program KITSCH, suggests the cause of the problem. Unlike the previous tree, Figure 8 is not necessarily a least-squares fit, and negative branch lengths have been suppressed in obtaining it. Nevertheless, it strongly suggests that the Negroids are not a monophyletic clade. The Bushmen and Andaman Islanders are distantly related to each other but still comprise a single lineage that is highly distinct from all remaining populations.

This separation of dark-skinned humans into two lineages is not without precedence. Coon (1965) partitioned Negroids into Congoids and Capoids (Bushman, Hottentots, and Sandawe). He did not formally include the Andamanese in the Capoids, but he did note several similarities between them and Bushmen: short stature, juvenile appearance, shortage of body hair, and steatopygia (fat buttocks). Data on isozymes (Nei and Roychoudhury, 1981) and mitochondrial-DNA restriction-fragment polymorphisms (Cann et al., 1987) also support the division of Negroids into Congoids and Capoids.

In order to determine whether the ambiguity of the previous results could have been an artifact of treating Negroids as a monophyletic lineage, the data were reanalyzed from the standpoint of four major divisions. The results suggest that Caucasians, Congoids, and Mongoloids are equidistant from each other, but about twice as distant from the Capoids (Table 4, Figure 7 [lower left]). There are no significant differences among the three mean observed distances to the Capoids, consistent with the neutral model. Moreover, five out of six of the observed variances of interracial distances are well within two standard errors of the expectation under the neutral model. The exception is the excessive variation in Mongoloid-Capoid distances, due primarily to a large distance between the Ainu and Andamanese. This comparison of observed and expected variances in interracial distances is still by no means totally satisfactory. The expected values are likely to be

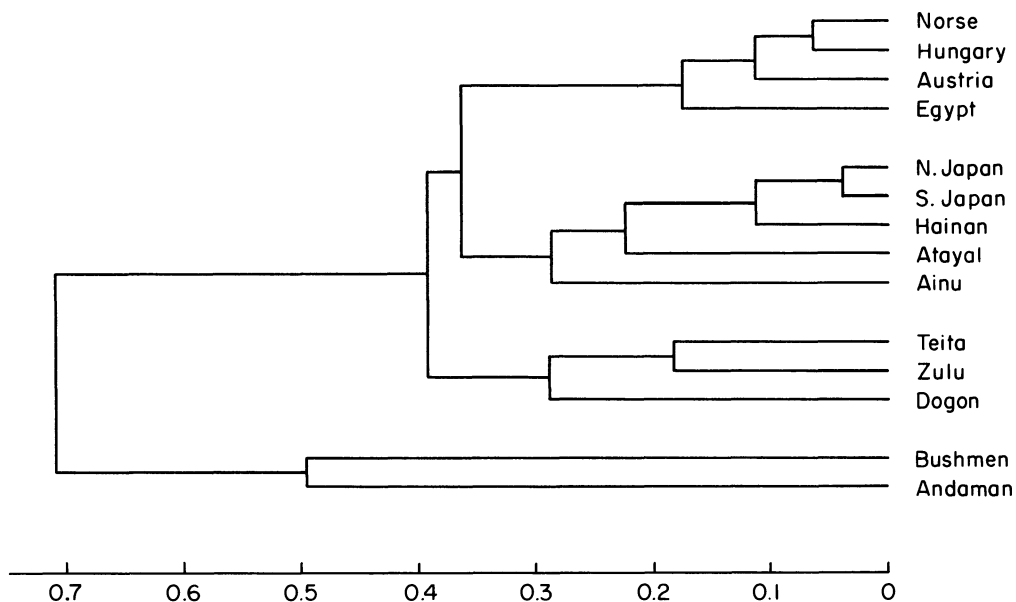


FIG. 8. A Fitch-Margoliash tree fit to the distance matrix for the populations listed in Table 3. The scale is the same as that in Figure 7.

somewhat too low due to the fact that the 61 measured characters are not genetically independent, as assumed. The observed values are likely to underrepresent the magnitude of dispersion of  $D'$  along branch segments, because the populations sampled within the major races are not likely to have been independent since the origin of their respective races.

With these limitations in mind, it is clear that Howells' unusually large data set, when subdivided into four distinct lineages, provides no strong basis for rejecting the hypothesis that the divergence of cranial morphology in the races of man is a simple consequence of random drift and polygenic mutation. The observation that multivariate distance measures of dental morphology in groups of man are strongly correlated with independent estimates of divergence times (Turner, 1986) lends additional credence to the neutral hypothesis for morphological diversification. These findings do not eliminate the possibilities that migration, selection, or environmental effects are responsible for the observed patterns. Indeed, since scenarios involving these influences can be concocted to explain virtually any pattern of variation, it is almost impossible to reject (or accept) them in a statistical sense. On

the other hand, the neutral hypothesis could be subjected to further scrutiny by performing measurements on collections of skulls from ancestral populations if they ever become available in sufficient numbers. Under the neutral model, such populations, on average, should be equidistant from all of their descendent lines.

It is of interest to inquire as to the rate of polygenic mutation that is required to account for the observed divergence in cranial morphology. Recall that the expected variance between isolated populations is  $2V_m t$  for neutral characters. The fossil evidence and molecular data indicate that modern man originated between 200,000 and 100,000 years B.P. Letting the generation time be 20 years, then the four-race analysis places the root of the tree at about 7,500 generations B.P. followed by a Caucasian-Congoid-Mongoloid trifurcation at about 3,750 generations B.P. Averaging over all combinations of characters and population pairs, the mean value of  $D'(i, j)/MS_w t$  is  $3.2 \times 10^{-4}$ . Since  $MS_w$  is an estimate of the within-population phenotypic variance, the quantity  $3.2 \times 10^{-4}$  may be compared with existing estimates of the standardized rate of polygenic mutation,  $V_m/V_E$ , where  $V_E$  is the within-population environmental vari-

ance, is necessarily less than  $MS_w$ . A survey of several species of invertebrates, vertebrates, and plants indicates that the normal limits to  $V_m/V_E$  are  $10^{-4}$  to  $10^{-2}$  (Lynch, 1988). Since the heritabilities of human cranial characters average approximately 0.5 (Bernhard et al., 1980), reasonable limits to  $D'(i, j)/MS_{w,t}$  under the neutral hypothesis are therefore  $5 \times 10^{-5}$  to  $5 \times 10^{-3}$ . Thus, the observed level of phenotypic divergence in the human cranium is quantitatively as well as qualitatively consistent with the neutral hypothesis. A similar analysis based on total tooth area on the human line of descent over the past 100,000 years (Brace et al., 1987) yields  $D'(i, j)/MS_{w,t} = 2.6 \times 10^{-4}$  (Lynch, unpubl.), which is quite consistent with the cranial data.

Finally, it is useful to compare the least-squares trees based on cranial measurements with those obtained previously with distances based on biochemical data. In an extensive survey of the literature, Nei and Roychoudhury (1981) have obtained average genetic distances between Caucasians, Congoids, and Mongoloids for proteins (62 loci) and blood groups (23 loci). Unlike the morphological analysis, the protein data suggest a much later split between Caucasians and Mongoloids than between Congoids and Caucasian-Mongoloids (Fig. 7). Both Caucasians and Mongoloids are equidistant from the Congoids on the basis of isozymes (Nei and Roychoudhury, 1981), consistent with the neutral theory. However, it is premature to say a conflict exists between the morphological and protein analyses, since the variances of the protein distances have not been compared with the neutral expectations.

The blood-group distances have relatively large standard errors and fail to establish any significant difference between branch points for Caucasians, Congoids, and Mongoloids, as in the morphological analysis. Nei and Roychoudhury (1981) argued that the evolutionary tree generated by protein data is more reliable than that based on blood groups because of uncertainty about the genetic basis of some of the latter and because a number of blood-group variants are known to be subject to selection. Certainly, because of their large standard errors, the blood-group data are inherently less in-

formative than the protein data, but it remains to be demonstrated that the latter are less biased than the former.

Cann et al. (1987) have reported the average genetic distances between the mitochondrial genomes of Caucasians, Mongoloids, and Negroids (including at least one Bushman). The resultant least-squares tree, standard errors for which are unavailable, is very similar to the one based on cranial morphology (Fig 7). A tree of the same general form has been suggested on the basis of restriction-site polymorphisms in the  $\beta$ -globin gene cluster (Wainscoat et al., 1986).

#### DISCUSSION

The message of this paper, which focuses on quantitative characters, is essentially the same as that emphasized in several recent theoretical investigations with discrete characters (Tateno et al., 1982; Nei et al., 1983; Fiala and Sokal, 1985; Tateno and Tajima, 1986). Provided that the measured characters are neutral and the study species have not been significantly affected by gene flow or by genotype  $\times$  environment interaction or covariance, then it is in principle possible to reconstruct phylogenetic trees. The price, however, is very high. The accurate assessment of the topology and branch lengths of a phylogeny requires that the number of characters and individuals measured per species be very large and that the fitted tree involve no more than four or five taxa. Even with enormous amounts of data, phylogenetic trees fitted to the distances among six or more species will almost always contain topological errors if the mode of divergence has been random drift and mutation. Very low values for distortion indices such as the APSD indicate only that the fitted distances are close to the observed distances and have very little bearing on whether the correct phylogeny has been recovered.

For neutral characters, the ability to reconstruct a phylogeny does not depend on the total evolutionary time for the lineage under consideration or on the total time elapsed between internal branch points. In other words, a tree of a particular form is equally difficult to recover whether the time scale is  $10^4$ ,  $10^5$ , or  $10^6$  years. This failure

of discriminatory power to increase with evolutionary time results because both the expected divergence and its standard error are proportional to time. The critical factors that determine whether phylogenies can be resolved from neutral characters (morphological or molecular) are therefore the relative positions of the internal nodes with respect to each other and to the branch tips.

Population genetics and systematics have developed largely as independent lines of endeavor. This is unfortunate, since phenomena such as convergence and parallelism, which thwart attempts to reconstruct the history of a phylogeny, are consequences of population-genetic processes such as random drift, selection, and migration. The results in this paper indicate that phylogenetic trees estimated by a least-squares or maximum-likelihood procedure can shed some light on the mechanisms underlying phylogenetic diversification. There are limits to what can be accomplished, however, when the only available data come from extant populations. While a rejection of the neutral hypothesis is tantamount to verifying that other evolutionary forces such as selection or migration have been important determinants of phenotypic divergence, acceptance of the null hypothesis does not eliminate the possible operation of these other evolutionary forces. It should also be noted that the neutral model is concerned only with the phenotypic changes that occur along the branch segments of a phylogeny. The isolation and speciation events that create the nodes of a tree are fundamentally separate issues.

An advantage of the least-squares approach is that it yields unbiased estimates of branch lengths under a neutral hypothesis. Most of the tree-building procedures that are usually used by systematists do not have this property. A disadvantage of the least-squares procedure is that with more than seven or so species, the computational time required for a solution makes it unfeasible to check all possible topologies. This is presumably the reason why most systematists follow alternative strategies. However, it may be questioned whether this cost of the least-squares procedure is really a serious problem. As noted above, with seven or more species, there will inevitably be to-

pological errors in the fitted tree, regardless of the fitting procedure.

Recently, Schluter (1984) presented a measure of evolutionary distance for quantitative characters for use in phylogeny reconstruction under a hypothesis that is the antithesis of the one presented here. Schluter's distance for a single character is

$$B(i, j) = \frac{|\bar{z}_i - \bar{z}_j|}{h^2 V_p}$$

where  $h^2$  is the heritability and  $V_p$  the within-species phenotypic variance.  $B(i, j)$  is the total net force of selection necessary to account for the different mean phenotypes in species  $i$  and  $j$ . Provided the additive genetic variance,  $h^2 V_p$ , and the average net force of directional selection per generation remain constant, then  $B(i, j)$  is expected to be proportional to divergence time. Schluter (1984) points out a strong correlation between a multivariate measure of  $B(i, j)$  and independent estimates of Rogers' genetic distance of isozymes in the Darwin's finches.

Schluter's (1984) measure of "selection distance" is very closely related to the relevant measure for the neutral model,  $D(i, j)$ . If the scale is adjusted so that the genetic variance is equal to 1, then  $B^2(i, j) = 2D(i, j)$ . Thus, the major difference between  $D(i, j)$  and  $B(i, j)$  is the expected time scale of divergence. The between-species variance is expected to increase linearly under the neutral model but quadratically under the selection model.

At first sight, this suggests a simple test to determine whether observed morphological divergence is more concordant with the neutral model or with the constant-selection model. If the divergence times,  $t(i, j)$ , are known from independent sources of information, a logarithmic regression of the fitted distances on  $t(i, j)$  should yield a slope close to one in the former case but close to two in the latter. If direct estimates of divergence times are not available, a similar regression analysis could be performed on a biochemical distance. Nei's (1972) distance, for example, is expected to scale with divergence time, provided that most of the amino acid (or nucleotide) substitutions are effectively neutral. Unfortunately, the power of such a test is expected to be low because

of the large variance of distance statistics, resulting from the random nature of the drift-mutation process.

Moreover, on closer consideration, it appears that the realization of the assumption of constant selection differentials between species would eliminate completely the possibility of identifying the branch points in a phylogeny. Consider the simple phylogeny in Figure 2, and let the net selection differentials operating on the three species be  $B_1$ ,  $B_2$ , and  $B_3$ , so that  $B(i, j) = B_i - B_j$ . Since species 1 and 2 are equidistant from species 3 in evolutionary time, the constant-selection model implies  $B(1, 3) = B(2, 3)$  and hence  $B_2 = B_3$ . Species 2 and 3 are therefore constrained to be morphologically identical under this model. With additional species, the implicit constraints of a constant-selection model become much more complex and stringent. Thus, the very fact that different species within lineages can be identified on the basis of morphology is inconsistent with the hypothesis of constant selection. These observations, as well as the relation of  $B(i, j)$  to  $D(i, j)$ , raise the possibility that the rough correspondence between  $B(i, j)$  and molecular distance in the Darwin's finches may be compatible with a hypothesis for divergence via random genetic drift.

No evolutionary biologist, including Schluter, would argue that selection differentials are absolutely constant in time. The problem is that as soon as this is admitted, an infinite number of scenarios arises, and it becomes impossible to reject the selection model in a general sense. Thus, the neutral model, which is now reasonably well understood, appears to be a more logical point of departure than a selection model in analyzing the phenotypic divergence in a phylogeny. With the neutral model, one at least has the possibility of rejecting the null hypothesis.

#### ACKNOWLEDGMENTS

This work has been supported in part by NSF grants BSR 83-06072, BSR 86-00487, and SUB U MICHX98674. I am very grateful to W. W. Howells for access to unpublished data, and to J. Felsenstein, W. W. Howells, and D. Schluter for very helpful comments on an early draft.

#### LITERATURE CITED

- ASTOLFI, P., K. K. KIDD, AND L. L. CAVALLI-SFORZA. 1981. A comparison of methods for reconstructing evolutionary trees. *Syst. Zool.* 30:156-169.
- BAILEY, D. W. 1959. Rates of subline divergence in highly inbred strains of mice. *J. Hered.* 50:26-30.
- BERNHARD, W., A. HANCKE, G. BRAUER, AND V. P. CHOPRA. 1980. Quantitative genetical analysis of morphological characters of the human head and face. *J. Hum. Evol.* 9:621-626.
- BRACE, C. L., K. R. ROSENBERG, AND K. D. HUNT. 1987. Gradual change in human tooth size in the late Pleistocene and post-Pleistocene. *Evolution* 41: 705-720.
- CANN, R. L., M. STONEKING, AND A. C. WILSON. 1987. Mitochondrial DNA and human evolution. *Nature* 325:31-36.
- CAVALLI-SFORZA, L. L., AND A. W. F. EDWARDS. 1965. Analysis of human evolution, pp. 923-933. *In* S. J. Geerts (ed.), *Genetics Today, Proceedings of the 11th International Congress of Genetics*, Vol. 3. Pergamon, Oxford, U.K.
- . 1967. Phylogenetic analysis: Models and estimation procedures. *Amer. J. Hum. Genet.* 19: 233-257.
- CAVALLI-SFORZA, L. L., AND A. PIAZZA. 1975. Analysis of evolution: Evolutionary rates, independence and treeness. *Theoret. Popul. Biol.* 8:127-165.
- CAVENDER, J. A. 1978. Taxonomy with confidence. *Mathemat. Biosci.* 40:27-280. [Erratum, 1979. 44: 308].
- . 1981. Tests of phylogenetic hypotheses under generalized models. *Mathemat. Biosci.* 54:217-229.
- CHAKRABORTY, R. 1977. Estimation of time of divergence from phylogenetic studies. *Can. J. Genet. Cytol.* 19:217-223.
- CHAKRABORTY, R., AND M. NEI. 1982. Genetic differentiation of quantitative characters between populations or species. I. Mutation and random genetic drift. *Genet. Res.* 39:303-314.
- CLAYTON, G. A., AND A. ROBERTSON. 1955. Mutation and quantitative variation. *Amer. Natur.* 89:151-158.
- COON, C. S. 1965. *The Living Races of Man*. Knopf, N.Y.
- CROW, J. F., AND M. KIMURA. 1970. *An Introduction to Population Genetics Theory*. Harper and Row, N.Y.
- EDWARDS, A. W. F., AND L. L. CAVALLI-SFORZA. 1964. Reconstruction of evolutionary trees, pp. 67-76. *In* V. H. Heywood and J. McNeill (eds.), *Phenetic and Phylogenetic Classification*. Systematics Association Publication Number 6. Systematics Association, London, U.K.
- . 1965. A method for cluster analysis. *Biometrics* 21:362-375.
- FARRIS, J. S. 1972. Estimating phylogenetic trees from distance matrices. *Amer. Natur.* 106:645-668.
- FELSENSTEIN, J. 1973. Maximum-likelihood estimation of evolutionary trees from continuous characters. *Amer. J. Hum. Genet.* 25:471-492.
- . 1982. Numerical methods for inferring evolutionary trees. *Quart. Rev. Biol.* 57:379-404.
- . 1983. Statistical inference of phylogenies. *J. Roy. Stat. Soc. A* 146:246-272.

- . 1985a. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* 39:783–791.
- . 1985b. Phylogenies and the comparative method. *Amer. Natur.* 125:1–15.
- . 1986. Population differences in quantitative characters and gene frequencies: A comment on papers by Lewontin and Rogers. *Amer. Natur.* 127:731–732.
- FIALA, K. L., AND R. R. SOKAL. 1985. Factors determining the accuracy of cladogram estimation: Evaluation using computer simulation. *Evolution* 39:609–622.
- FITCH, W. M., AND E. MARGOLISH. 1967. Construction of phylogenetic trees. *Science* 155:279–284.
- FREUND, J. E. 1971. *Mathematical Statistics*, 2nd Ed. Prentice-Hall, Englewood Cliffs, NJ.
- HOWELLS, W. W. 1970. Anthropometric grouping analysis of Pacific peoples. *Arch. Phys. Anthropol. Ocean.* 5:192–217.
- . 1973a. Cranial variation in man. *Pap. Peabody Mus. Archeol. Ethnol. Harvard Univ.* 67:1–259.
- . 1973b. *The Pacific Islanders*. Weidenfeld and Nicolson, London, U.K.
- KIDD, K. K., P. ASTOLFI, AND L. L. CAVALLI-SFORZA. 1974. Error in the reconstruction of evolutionary trees, pp. 121–136. *In* J. F. Crow and C. Denniston (eds.), *Genetic Distance*. Plenum, N.Y.
- KIMURA, M. 1983. *The Neutral Theory of Molecular Evolution*. Cambridge Univ. Press, Cambridge, U.K.
- KNUTH, D. E. 1969. *The Art of Computer Programming*, Vol. 2. Seminumerical Algorithms. Addison-Wesley, Reading, MA.
- LANDE, R. 1975. Natural selection and random genetic drift in phenotypic evolution. *Evolution* 30:314–334.
- . 1979. Quantitative genetic analysis of multivariate evolution, applied to brain: body size allometry. *Evolution* 33:402–416.
- LI, W.-H. 1981. Simple methods for constructing phylogenetic trees from distance matrices. *Proc. Nat. Acad. Sci. USA* 78:1085–1089.
- LYNCH, M. 1988. The rate of polygenic mutation. *Genet. Res.* 51:137–148.
- LYNCH, M., AND W. G. HILL. 1986. Phenotypic evolution by neutral mutation. *Evolution* 40:915–935.
- NEI, M. 1972. Genetic distance between populations. *Amer. Natur.* 106:283–292.
- NEI, M., AND A. K. ROYCHOUDHURY. 1981. Genetic relationships and evolution of human races. *Evol. Biol.* 14:1–59.
- NEI, M., J. C. STEPHENS, AND N. SAITOU. 1985. Methods for computing the standard errors of branching points in an evolutionary tree and their application to molecular data from humans and apes. *Molec. Biol. Evol.* 2:66–85.
- NEI, M., F. TAJIMA, AND Y. TATENO. 1983. Accuracy of estimated phylogenetic trees from molecular data. II. Gene frequency data. *J. Molec. Evol.* 19:153–170.
- NEYMAN, J. 1971. Molecular studies of evolution: A source of novel statistical problems, pp. 1–27. *In* S. S. Gupta and J. Yackel (eds.), *Statistical Decision Theory and Related Topics*. Academic Press, N.Y.
- PENNY, D., AND M. HENDY. 1986. Estimating the reliability of evolutionary trees. *Molec. Biol. Evol.* 3:403–417.
- ROGERS, A. R., AND H. C. HARPENDING. 1983. Population structure and quantitative characters. *Genetics* 105:985–1002.
- SCHLUTER, D. 1984. Morphological and phylogenetic relations among the Darwin's finches. *Evolution* 38:921–930.
- SNEATH, P. H. A., AND R. R. SOKAL. 1973. *Numerical Taxonomy*. Freeman, San Francisco, CA.
- TATENO, Y., M. NEI, AND F. TAJIMA. 1982. Accuracy of estimated phylogenetic trees from molecular data. I. Distantly related species. *J. Molec. Evol.* 18:387–404.
- TATENO, Y., AND F. TAJIMA. 1986. Statistical properties of molecular tree construction methods under the neutral mutation model. *J. Molec. Evol.* 23:354–361.
- 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.
- . 1985. The phylogeny of the hominoid primates: A statistical analysis of the DNA-DNA hybridization data. *Molec. Biol. Evol.* 2:420–433.
- THOMPSON, E. A. 1975. *Human Evolutionary Trees*. Cambridge Univ. Press, Cambridge, U.K.
- TURNER, C. G., II. 1986. Dentochronological separation estimates for Pacific rim populations. *Science* 232:1140–1142.
- WAINSCOAT, J. S., A. V. S. HILL, A. L. BOYCE, J. FLINT, M. HERNANDEZ, S. L. THEIN, J. M. OLD, J. R. LYNCH, A. G. FALUSI, D. J. WEATHERALL, AND J. B. CLEGG. 1986. Evolutionary relationships of human populations from an analysis of nuclear DNA polymorphisms. *Nature* 319:491–493.
- WRIGHT, S. 1951. The genetical structure of populations. *Ann. Eugen.* 15:323–354.

Corresponding Editor: J. M. Cheverud