& J. L. Slightom. 1984. The sequence of the gorilla fetal globin genes: evidence for multiple gene conversions in human evolution. Mol. Biol. Evol. 1: 371-389.

Sheldon, F. H. 1987a. Phylogeny of herons estimated from DNA-DNA hybridization data. Auk 104: 97-108.

———. 1987b. Rates of single-copy DNA evolution in herons. Mol. Biol. Evol. 4: 56-69.

Sibley, C. G., & J. E. Ahlquist. 1980. The relationships of the "primitive insect eaters" (Aves: Passeriformes) as indicated by DNA-DNA hybridization. Proc. 17th Intern. Ornithol. Congr.: 1215-1220.

———, & ———. 1981. The phylogeny and relationships of the ratite birds as indicated by DNA-DNA hybridization. Pp. 303-335 *in* Evolution today. Proc. 2nd Intern. Congr. Syst. Evol. Biol. (G. G. E. Scudder and J. L. Reveal, Eds.). Pittsburgh, Pennsylvania, Hunt Inst. Botan. Document., Carnegie-Mellon Univ.

———, & ———. 1984. The phylogeny of the hominoid primates, as indicated by DNA-DNA hybridization. J. Mol. Evol. 20: 2-15.

———, & ———. 1985a. Phylogeny and classification of New World suboscine passerine birds (Passeriformes: Oligomyodi: Tyrannides). Pp. 396-428 *in* Neotropical ornithology (P. A. Buckley, M. S. Foster, R. S. Ridgely, and F. G. Buckley, Eds.). Ornithol. Monogr. No. 36.

———, & ———. 1985b. The relationships of some groups of African birds, based on comparisons of the genetic material, DNA. Pp. 115-161 *in* Proc. Intern. Symp. on African Vertebrates (K.-L. Schuchmann, Ed.). Bonn, Zool. Forschungsinst. and Mus. A. Koenig.

———, & ———. 1987. DNA hybridization evidence of hominoid phylogeny: results from an expanded data set. J. Mol. Evol. in press.

———, ———, & F. H. Sheldon. 1987. DNA hybridization and avian phylogenetics: reply to Cracraft. Pp. 97-125 *in* Vol. 21, Evol. Biol. (M. Hecht, B. Wallace, and G. Prance, Eds.). New York, Plenum Press.

———, G. R. Williams, & J. E. Ahlquist. 1982. The relationships of the New Zealand wrens (Acanthisittidae) as indicated by DNA-DNA hybridization. Notornis 29: 113-130.

Templeton, A. R. 1985. The phylogeny of the hominoid primates: a statistical analysis of the DNA-DNA hybridization data. Mol. Biol. Evol. 2: 420-433.

———. 1986. Further comments on statistical analysis of DNA-DNA hybridization data. Mol. Biol. Evol. 3: 290-295.

Wetmore, A. 1960. A classification for the birds of the world. Smithsonian Misc. Coll. 139(11): 1-37.

Wu, C.-I., & W.-H. Li. 1985. Evidence for higher rates of nucleotide substitution in rodents than in man. Proc. Natl. Acad. Sci. USA 82: 1741-1745.

Received 13 April 1987, accepted 16 April 1987.

Estimation of Phylogeny from Molecular Distance Data: The Issue of Variable Rates

Anthony H. Bledsoe[1]

Must rates of macromolecular evolution be uniform for measurements of amino acid or DNA sequence differences among taxa to be used in phylogenetic reconstruction? The evidence that rates of DNA evolution vary significantly among lineages of many organisms (Britten 1986), including birds (Sheldon 1987), makes this question especially pertinent to avian systematics. Houde (1987) contended that uniformity of rate is necessary for the use of distance data in phylogenetic reconstruction. However, his statement (p. 25) that "Satisfaction of the relative-rate test [of rate uniformity] is a prerequisite for the use of DNA data for phylogenetic reconstructions" reveals a misunderstanding of the nature of distance data and of the methods used to analyze their phylogenetic implications. This misunderstanding contributes to Houde's mistaken idea that variation in rates alone will introduce ambiguity into the reconstruction of phylogenetic branching patterns.

To see that varying rates do not inherently preclude accurate estimates of phylogeny, imagine a monophyletic set of species whose DNA sequences are evolving at the same positive rate, except for two species, which are not sister groups. These have a slow rate of DNA evolution, and thus show a smaller distance between one another than either does to any other species, including their sister groups. It is obvious that the incorrect joining of these two as sister species will produce discrepancies between the original data and any possible set of positive distances among taxa in the reconstructed (and incorrect) topology. If one would measure the level of discrepancy

---

[1] Section of Birds, Carnegie Museum of Natural History, Pittsburgh, Pennsylvania 15213 USA; and Department of Biological Sciences, University of Pittsburgh, Pittsburgh, Pennsylvania 15260 USA.

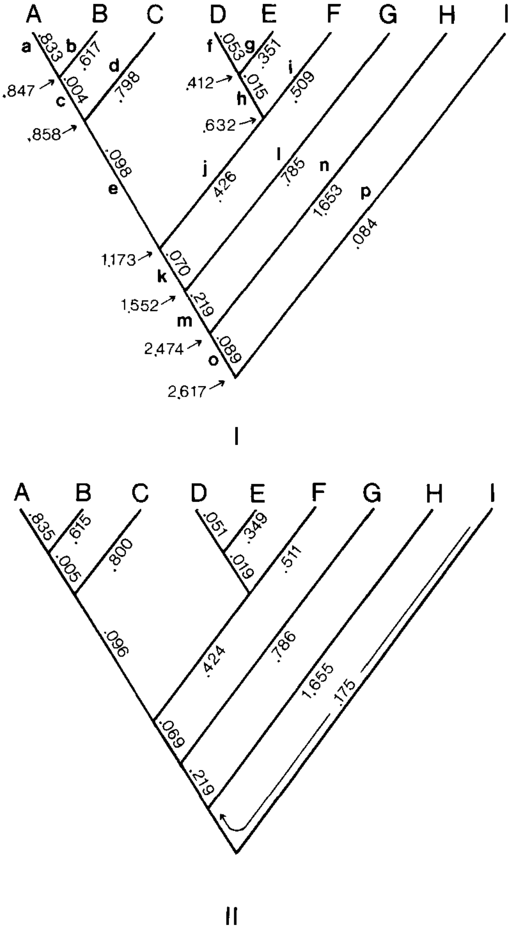

Fig. 1. (I) Tree of 9 hypothetical taxa and (II) tree reconstructed by least-squares analysis of the matrix of distances among them. For I, the numbers next to the arrows are relative dates of divergence for the indicated nodes, and letters correspond to the following relative divergence rates for the indicated stems: a = 0.984, b = 0.728, c = 0.343, d = 0.930, e = 0.310, f = 0.128, g = 0.852, h = 0.070, i = 0.805, j = 0.788, k = 0.184, l = 0.506, m = 0.238, n = 0.668, o = 0.621, p = 0.032. The numbers parallel to the stems of both trees are branch lengths. II is rooted at i and has a residual sum of squares = 0.00021.

between the original and reconstructed distances for a variety of different topologies, including the correct one, it would become clear that the initial reconstruction exhibited a level of discrepancy higher than that of the correct topology.

The problem is to use methods of reconstruction that can identify and correct such discrepant topologies. Clustering methods, such as average linkage procedures, clearly cannot, because they produce a single tree that constrains branch lengths to reflect uniform rates of divergence. Such methods are inappropriate when rate differences exist. However, pairwise methods (reviewed by Felsenstein 1982), when combined with heuristic algorithms to check many different trees, adjust topology and branch lengths to minimize some measure (e.g. the sum of squares) of the discrepancy between the original distances and those calculated from the reconstructed topology. If distances are additive, independent, and error free, such methods regularly will reconstruct branching patterns correctly despite large variations in evolutionary rates.

Figure 1 provides an example of the ability of pairwise methods to recover the correct branching pattern in spite of large differences in evolutionary rates. I constructed a matrix of distances from a tree of 9 hypothetical taxa whose dates of divergence and positive rates of evolution were chosen randomly with a random-number table. The matrix was then analyzed with a least-squares pair-wise method (Cavalli-Sforza and Edwards 1967) using the FITCH option of the computer program PHYLIP (version 2.6, by J. Felsenstein). FITCH includes a heuristic algorithm for assessing many topologies for the lowest least-squares network. For this hypothetical tree, and for the 5 others that I analyzed, the least-squares analysis recovered the correct branching pattern. In 3 instances, the correct branch lengths were also reconstructed. Clearly, variation in rates of evolution alone does not necessarily preclude accurate estimation of phylogeny from distance data.

Unfortunately, it is not clear that existing methods will always recover the pattern of branching accurately, although not because rates are variable. Rather, in estimating branching pattern from a matrix of distances, there are two interrelated problems to solve: (1) to find the minimal topological length given a specific topology, and (2) to identify the topology that gives the minimum length (Fitch and Smith 1982). No efficient method exists to solve these problems simultaneously; available quantitative procedures either yield optimal solutions but are inefficient and are thus limited to very small problems (i.e. few taxa) or are efficient but give solutions not shown to be optimal (Day 1983).

These considerations make it clear why "estimation" is the most appropriate term for describing the results of phylogenetic analysis of distance (or for that matter character-state) data. Such estimates nonetheless can be quite accurate despite substantial variation in evolutionary rates, particularly when error-free and strictly additive distance data (Fig. 1) are used.

Problems in the estimation of branching patterns from distance data are thus less a function of variation in positive evolutionary rates (given that one uses an appropriate method of estimation) than of the inconsistency produced when values fail to meet metric (Sneath and Sokal 1973) or additive (four-point metric; Buneman 1971) criteria. For DNA-DNA hybridization measurements, such failures can result from at least three phenomena. First, experimental error may introduce homoplasy, as well as negate the assumptions of additivity and independence required by pair-wise methods (Felsenstein 1984). Second, net rates of divergence may not be positive (that is, there may be convergence). Arguments against this possibility rely on the small probability of convergence at each of many consecutive nucleotide positions, but a rigorous statistical model has not to my knowledge been developed. The probability argument does not apply to apparent convergence as a result of lateral sequence transfer through introgressive hybridization. Cavalli-Sforza and Edwards (1967) noted that the existence of convergence would also violate the assumption of independence of distances. Third, homologies may not exist between taxa, as the result of sequence deletion or lateral transfer mediated by infectious agents. This would be evident in poor reciprocity of delta $T_{50}H$ (Houde 1987), but would not affect delta T_m (Sheldon 1987) or delta mode (Bledsoe 1984) distances, which do not incorporate a percentage hybridization component.

These possible sources of discrepancy point to the importance of the additivity and independence of distances emphasized by Felsenstein (1982, 1984). If distances are additive and independent metrics, then tree reconstruction becomes an issue of whether distances should be used at all (see Farris 1981 and Felsenstein 1984 for opposing views) and, if they should, of how well heuristic methods estimate topology. The latter is an area where additional research is especially needed. If nonadditivity or nonindependence is caused by experimental error, then better biochemical techniques can be applied to reduce the level of error. That a suitable framework exists for identifying error-induced nonadditivity and nonindependence is apparent from the statistical tests used by Bledsoe (1984), Sheldon (1986, 1987), and Houde (1987); such tests need to be developed more fully, particularly with respect to additivity. However, if biological circumstances of convergence or nonhomology produce dissimilarity measures that are not metrics, then new methods of construction and interpretation of topologies based on molecular comparisons will be required.

I thank R. J. Raikow and D. S. Wood for their helpful comments on this manuscript.

LITERATURE CITED

BLEDSOE, A. H. 1984. The phylogeny and evolution of the New World nine-primaried oscines, as indicated by DNA-DNA hybridization. Ph.D. dissertation, New Haven, Connecticut, Yale Univ.

BRITTEN, R. J. 1986. Rates of DNA sequence evolution differ between taxonomic groups. Science 231: 1393–1398.

BUNEMAN, P. 1971. The recovery of trees from measurements of dissimilarity. Pp. 387–396 *in* Mathematics in the archeological and historical sciences (F. R. Hodson, D. G. Kendall, and P. Tautu, Eds.). Edinburgh, Scotland, Edinburgh Univ. Press.

CAVALLI-SFORZA, L. L., & A. W. F. EDWARDS. 1967. Phylogenetic analysis: models and estimation procedures. Evolution 21: 550–570.

DAY, W. H. E. 1983. Computationally difficult parsimony problems in phylogenetic systematics. J. Theor. Biol. 103: 429–438.

FARRIS, J. S. 1981. Distance data in phylogenetic analysis. Pp. 3–23 *in* Advances in cladistics (V. A. Funk and D. R. Brooks, Eds.). New York, New York Botanical Garden.

FELSENSTEIN, J. 1982. Numerical methods for inferring evolutionary trees. Q. Rev. Biol. 57: 379–404.

———. 1984. Distance methods for inferring phylogenies: a justification. Evolution 38: 16–24.

FITCH, W. M., & T. F. SMITH. 1982. Implications of minimal length trees. Syst. Zool. 31: 68–75.

HOUDE, P. 1987. Critical evaluation of DNA hybridization studies in avian systematics. Auk 104: 17–32.

SHELDON, F. H. 1986. A study of the evolution and phylogeny of the herons (Ardeidae) using DNA-DNA hybridization. Ph.D. dissertation, New Haven, Connecticut, Yale Univ.

———. 1987. Rates of single-copy DNA evolution in herons. Mol. Biol. Evol. 4: 56–69.

SNEATH, P. H. A., & R. R. SOKAL. 1973. Numerical taxonomy. San Francisco, W. H. Freeman and Co.

Received 23 February 1987, accepted 16 April 1987.