



**Statistical Significance of the Matrix Correlation Coefficient for Comparing Independent Phylogenetic Trees**

Francois-Joseph Lapointe; Pierre Legendre

*Systematic Biology*, Vol. 41, No. 3 (Sep., 1992), 378-384.

Stable URL:

<http://links.jstor.org/sici?sici=1063-5157%28199209%2941%3A3%3C378%3ASSOTMC%3E2.0.CO%3B2-K>

*Systematic Biology* is currently published by Society of Systematic Biologists.

---

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/ssbiol.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.

---

For more information on JSTOR contact [jstor-info@umich.edu](mailto:jstor-info@umich.edu).

©2003 JSTOR

- mals. Pages 57–95 in International colloquium on the evolution of lower and specialized mammals, Part 1 (G. Vandebroek, ed.). Kon. Vlaamse Acad. Wetensch., Lett. Schone Kunstten Belgie, Brussels.
- SIMPSON, G. G. 1971. Concluding remarks. Pages 181–198 in Early mammals (D. M. Kermack and K. L. Kermack, eds.). Zool. J. Linn. Soc. 50(Suppl. 1).
- VAN VALEN, L. 1960. Therapsids as mammals. *Evolution* 14:304–313.
- WIBLE, J. R. 1991. Origin of Mammalia: The craniodental evidence reexamined. *J. Vertebr. Paleontol.* 11:1–28.

Received 6 June 1992; accepted 18 June 1992

*Syst. Biol.* 41(3):378–384, 1992

## Statistical Significance of the Matrix Correlation Coefficient for Comparing Independent Phylogenetic Trees

FRANÇOIS-JOSEPH LAPOINTE AND PIERRE LEGENDRE

*Département de Sciences biologiques, Université de Montréal, C.P. 6128, Succursale A, Montréal, Québec H3C 3J7, Canada*

Ultrametric trees or dendrograms (=phenograms of some authors; Sneath and Sokal, 1973) have received much attention in phylogenetics and are used to portray the relationships among members of a given taxonomic group in the presence of constant evolutionary rates, corresponding to a molecular clock hypothesis (Blanken et al., 1982). However, additive trees or cladograms provide a better representation of phylogenetic distances in the presence of unequal evolutionary rates among lineages (Tateno et al., 1982). Numerous algorithms permit derivation of ultrametric or additive trees from distance matrices or directly from character matrices (Felsenstein, 1982; Gordon, 1987; Swofford and Olsen, 1990; Penny et al., 1992). With these methods arose the problems of efficiency (Milligan, 1981). Which algorithm is the best? What is a good phylogenetic tree? How should the efficiency of a particular method be measured?

The next logical step after characterization of the various methods is to compare the stability of the constructed phylogenetic trees under different reconstruction techniques (Sokal et al., 1992). Consensus

tree methods (Day, 1986, and references therein) and consensus indices (Rohlf, 1982) were derived to measure the similarity between phylogenies or to produce a compromise solution reflecting the common agreement of several trees. More recently, procedures have been proposed to compare trees statistically (Page, 1988; Lapointe and Legendre, 1990, 1992), and significance tables have been published for many consensus indices (Shao and Rohlf, 1983; Shao and Sokal, 1986). The present paper proposes tables of critical values of the cross-product matrix correlation coefficient (Rohlf, 1982), designed to allow statistical comparison of trees while taking into account the metric information embedded in the branches of a phylogeny.

### TREE PROPERTIES

The relationships among the objects of a set  $W = \{1, 2, \dots, n\}$  that represents a given taxonomic group can be portrayed in a matrix  $\mathbf{D}$  of pairwise distances. Such a matrix of metric distances must satisfy four minimal conditions (cf. Sneath and Sokal, 1973):

$$d_{ij} = 0 \tag{1}$$

for  $i = j$ ,

$$d_{ij} \geq 0 \tag{2}$$

for  $i \neq j$ ,

$$d_{ij} = d_{ji} \tag{3}$$

for  $i \neq j$ , and

$$d_{ik} \leq d_{ij} + d_{jk} \tag{4}$$

for every  $i, j$ , and  $k$ . Furthermore, patristic matrices **A** that represent additive distances must also satisfy the four-point condition (Buneman, 1971):

$$d_{ij} + d_{kl} \leq \max[(d_{ik} + d_{jl}), (d_{il} + d_{jk})] \tag{5}$$

for every  $i, j, k$ , and  $l$ , whereas ultrametric matrices **U** containing ultrametric distances should meet an additional property, called the ultrametric condition (Hartigan, 1967; Johnson, 1967):

$$d_{ij} \leq \max[d_{ij}, d_{jk}] \tag{6}$$

for every  $i, j$ , and  $k$ . Any matrix **A** or **U** that uniquely defines a tree can be represented (Fig. 1) in the form of either a cladogram (=additive tree) or a dendrogram (=ultrametric tree). For each pair of objects  $i$  and  $j$  in a tree,  $a_{ij}$  or  $u_{ij}$  is defined as the sum of branch lengths along the path connecting these two objects. Given a distance matrix **D**, one can represent the relations among the  $n$  elements of **D** in the form of a phylogenetic tree. This reduces to obtaining an additive or ultrametric representation of **D**. A phylogenetic tree reconstruction operation thus acts as a transformation from metric to path-length distances; even though most cladograms are directly derived from a character matrix, the resulting tree remains additive.

MATRIX CORRELATION

The efficiency of a particular method is reflected in its ability to uncover the phylogenetic structure of a distance matrix. Numerous indices have been suggested to measure the fit of a tree to the original distance matrix (Gordon, 1987). The matrix

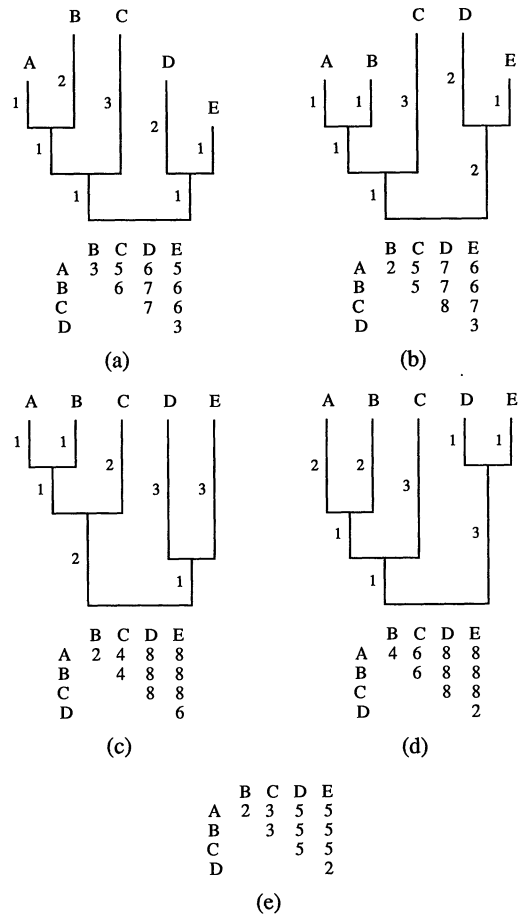


FIGURE 1. A pair of cladograms, (a) and (b), and a pair of dendrograms, (c) and (d), with their corresponding path-length matrices. The cardinality matrix (e) is the same for all four trees.

correlation coefficient, which is simply the Pearson cross-product correlation coefficient computed over the pairs of values in the patristic or ultrametric and the original distance matrices, has been suggested by Sokal and Rohlf (1962) to evaluate clustering results:

$$\frac{\sum (d_{ij} - \bar{d})(\delta_{ij} - \bar{\delta})}{\sqrt{\sum (d_{ij} - \bar{d})^2 \sum (\delta_{ij} - \bar{\delta})^2}} \tag{7}$$

for  $i \neq j$ , where the  $d_{ij}$  are the input distance

values and the  $\delta_{ij}$  are the path-length distances. This measure is equivalent to the standardized form of the Mantel (1967) statistic.

This widely used coefficient has often been suggested as a means of selecting the optimal tree for the distances in **D**. In this type of situation, the correlation coefficient cannot be tested for statistical significance because matrix **A** or matrix **U** is not independent from distance matrix **D**, from which it is derived. However, the same correlation coefficient can also be used to measure the agreement between matrices derived from different tree reconstruction methods to reflect the consensus of the corresponding solutions (Rohlf, 1982). Like any standard correlation coefficient, one could be interested in assessing the statistical significance of the correlation value, either as a recovery or as a stability index. For the stability index, the correlation compares distinct path-length matrices (**A** or **U**) derived from the same matrix **D** using different methods. Again, the significance of the matrix correlation coefficient cannot be assessed because the matrices under comparison are not independent, thus invalidating the test (Hubert and Baker, 1977). The test is correct only when comparing trees independently derived from different distance matrices; it is not appropriate to evaluate the ability of distinct methods to produce similar phylogenies given the same initial matrix **D**.

Besides the independence of the matrices, the usual test of the Pearson correlation also requires independence of the observations to satisfy its application conditions. The original (in **D**), the patristic (in **A**), or the ultrametric (in **U**) distances are not independent from one another, however, because the triangular inequality (Equation 4), the four-point condition (Equation 5), or the ultrametric inequality (Equation 6) imposes constraints on the distance values of these matrices. The usual test of significance involving independent matrices always overestimates the correct degrees of freedom when comparing trees. Therefore, we need corrected tables of significance of the

matrix correlation coefficient to measure the consensus of independent patristic or ultrametric matrices representing cladograms or dendrograms.

#### CORRELATION OF PATH-LENGTH MATRICES

When comparing a pair of actual phylogenetic trees uniquely represented by a pair of path-length matrices, one evaluates whether these trees are more similar than those that could be obtained by chance alone. This is a one-tailed test. The null hypothesis is that the two matrices under comparison are not more correlated than matrices associated with random trees would be (Lapointe and Legendre, 1990, 1992). The best way to evaluate the significance of the actual correlation value is to use a reference distribution of the statistic constructed using all possible tree comparisons. The exact distribution cannot be constructed, however, because the number of trees increases in a nonlinear way as a function of the number of taxa  $n$  (Phipps, 1975; Felsenstein, 1978; Murtagh, 1984) and quickly becomes intractable. The test must be formulated, therefore, by reference to a randomized null distribution based on the comparison of trees sampled equiprobably from the population of cladograms or dendrograms bearing  $n$  taxa.

Lapointe and Legendre (1991) proposed a uniform algorithm to generate ultrametric matrices that represent random dendrograms. The method proceeds by generating a random topology followed by random labeling and random allocation of cluster heights sampled from a uniform [0, 1] distribution. An extension of this method has been used to generate random cladograms by adding a random ultrametric tree and a random star tree (Lapointe and Legendre, 1992). The method used in the present paper involves generation of a random ultrametric tree followed by additional randomization of the branch lengths to obtain an additive tree by adding (or subtracting) a random length to (or from) each terminal branch of the tree under the constraint that all branches of the cladogram must have a positive length. Here, we use these two procedures to gen-

erate tables of statistical significance for the matrix correlation coefficient involving cladogram and dendrogram comparisons.

#### CORRELATION OF TOPOLOGICAL MATRICES

When comparing trees with the Pearson correlation coefficient, one evaluates whether the path-length matrices under consideration are linearly correlated. However, a pair of trees with similar topologies can have different path-length distances; identical topologies may even be declared statistically different when considering branch lengths (Fig. 1). Comparing topological relationships among taxa is a problem distinct from the comparison of path-length distances, however. To compare topologies, one should consider topological distance matrices (Phipps, 1971) instead of path-length matrices, replacing the sum of branch lengths by the number of branches along the path connecting two objects. Ultrametric topological distances (=cardinality matrices) and additive topological distances uniquely correspond to the trees they represent (Smolenskii, 1969), although cladogram and dendrogram topologies have an identical distribution under a strictly dichotomous Markovian model (Page, 1991). Generation of topological distance matrices thus proceeds by generation of random ultrametric matrices that are transformed into cardinality matrices (Fig. 1e), in which every  $d_{ij}$  represents the cardinality of the smallest set of objects that includes both  $i$  and  $j$ . The topological comparison of cladograms and dendrograms is based on the same null hypothesis, that their topologies are not more similar than random cardinality matrices would be. Tables of significance could be generated by comparing such random matrices.

#### GENERATING TABLES OF SIGNIFICANCE

To produce tables of critical correlation values, cladograms and dendrograms were generated at random, using the Lapointe and Legendre (1990, 1991, 1992) algorithms, for increasing numbers of taxa ranging from 4 to 100. The null distribu-

TABLE 1. Critical values of the matrix correlation coefficient used to compare independent cladograms.

n	Level of significance ( $\alpha$ )				
	0.50	0.25	0.10	0.05	0.01
4	-0.056	0.374	0.684	0.808	0.955
5	-0.005	0.306	0.541	0.651	0.838
6	-0.045	0.240	0.429	0.548	0.718
7	0.074	0.211	0.389	0.483	0.691
8	0.045	0.209	0.358	0.456	0.605
9	-0.018	0.186	0.346	0.422	0.580
10	-0.016	0.173	0.334	0.404	0.551
11	0.048	0.156	0.290	0.373	0.512
12	0.034	0.150	0.279	0.352	0.500
13	-0.022	0.131	0.271	0.347	0.465
14	-0.067	0.120	0.242	0.318	0.453
15	-0.016	0.118	0.238	0.307	0.420
20	0.011	0.115	0.224	0.276	0.375
25	-0.077	0.103	0.186	0.250	0.343
30	-0.017	0.092	0.178	0.225	0.321
35	-0.053	0.089	0.169	0.213	0.285
40	0.003	0.088	0.168	0.209	0.274
45	0.005	0.085	0.152	0.203	0.269
50	-0.035	0.081	0.141	0.177	0.246
75	-0.008	0.060	0.123	0.160	0.237
100	-0.002	0.055	0.108	0.136	0.213

tion of the matrix correlation coefficient was then constructed by pairing phylogenetic trees to compute the correlation values. Empirical simulations (Lapointe, 1990) have shown that 500 correlation values were more than enough to construct a satisfactory reference distribution. In the present work, we compared 1,000 pairs of cladograms or dendrograms for each row of the tables. For each number of taxa, 2,000 random patristic or ultrametric matrices were thus generated, paired at random, and correlated. Table 1 provides information about the critical values of the matrix correlation coefficient at different levels of significance for the comparison of pairs of cladograms, and Table 2 contains the same information for dendrogram comparisons. Table 3 was computed to give critical values of the matrix correlation coefficient when comparing a dendrogram to a cladogram, and Table 4 was obtained by comparing cardinality matrices representing random phylogenetic trees. One can interpolate the correlation values linearly to obtain an approximate critical value for any particular number of objects not presented in these tables. The 0.001 level was not

TABLE 2. Critical values of the matrix correlation coefficient used to compare independent dendrograms.

n	Level of significance ( $\alpha$ )				
	0.50	0.25	0.10	0.05	0.01
4	-0.189	0.216	0.671	0.882	0.992
5	-0.089	0.185	0.506	0.689	0.949
6	-0.034	0.157	0.381	0.547	0.791
7	-0.058	0.131	0.321	0.465	0.745
8	-0.037	0.113	0.262	0.371	0.572
9	-0.031	0.111	0.238	0.351	0.530
10	-0.013	0.102	0.206	0.282	0.445
11	-0.020	0.083	0.190	0.245	0.432
12	-0.028	0.068	0.167	0.231	0.340
13	-0.025	0.058	0.153	0.212	0.325
14	-0.012	0.060	0.147	0.190	0.297
15	-0.018	0.056	0.126	0.172	0.277
20	-0.012	0.043	0.098	0.140	0.204
25	-0.007	0.035	0.083	0.101	0.170
30	-0.008	0.027	0.071	0.093	0.163
35	-0.003	0.027	0.058	0.080	0.123
40	-0.003	0.022	0.054	0.069	0.104
45	-0.001	0.023	0.047	0.059	0.088
50	-0.003	0.020	0.042	0.056	0.084
75	-0.002	0.014	0.029	0.036	0.056
100	-0.001	0.009	0.020	0.026	0.046

TABLE 3. Critical values of the matrix correlation coefficient used to compare a cladogram to a dendrogram.

n	Level of significance ( $\alpha$ )				
	0.50	0.25	0.10	0.05	0.01
4	-0.031	0.398	0.715	0.824	0.943
5	-0.035	0.218	0.490	0.670	0.841
6	-0.051	0.169	0.376	0.502	0.708
7	-0.032	0.128	0.315	0.400	0.654
8	-0.014	0.121	0.260	0.363	0.571
9	-0.093	0.120	0.259	0.330	0.500
10	-0.011	0.099	0.212	0.284	0.443
11	0.036	0.094	0.195	0.263	0.386
12	-0.035	0.088	0.190	0.240	0.353
13	-0.012	0.086	0.163	0.223	0.339
14	-0.075	0.077	0.159	0.219	0.311
15	-0.063	0.076	0.150	0.203	0.307
20	-0.041	0.053	0.108	0.143	0.249
25	0.003	0.046	0.093	0.120	0.195
30	-0.021	0.039	0.087	0.114	0.163
35	-0.008	0.036	0.068	0.096	0.145
40	0.004	0.031	0.063	0.079	0.138
45	-0.017	0.030	0.059	0.076	0.116
50	-0.011	0.025	0.051	0.069	0.102
75	0.022	0.020	0.038	0.053	0.095
100	0.018	0.016	0.030	0.043	0.061

computed here, however; more comparisons would have been required to obtain accurate critical values for this level.

#### EXAMPLE

To compare the dendrograms and cladograms depicted in Figure 1, the first step is to compute the matrix correlation coefficient between their associated path-length matrices. The significance of this statistic is evaluated by examining the tables of critical values under the preselected level of significance (say  $\alpha = 0.05$ ) and the corresponding number of taxa (five in this case). If the actual value of the statistic is larger than or equal to the critical value from the random distribution, then the null hypothesis of randomness of the relation between the two phylogenetic trees is to be rejected. In our example, we obtained a correlation of 0.926 between the two cladograms and a correlation of 0.684 between the two dendrograms. Therefore, the cladograms under comparison are more similar than expected by chance alone, and

TABLE 4. Critical values of the matrix correlation coefficient based on topological comparisons of independent dendrograms.

n	Level of significance ( $\alpha$ )				
	0.50	0.25	0.10	0.05	0.01
4	-0.158	0.316	0.700	0.999	1.000
5	-0.072	0.144	0.500	0.818	1.000
6	-0.068	0.167	0.349	0.474	0.912
7	-0.044	0.089	0.324	0.431	0.880
8	-0.061	0.080	0.275	0.424	0.798
9	-0.044	0.074	0.224	0.391	0.637
10	-0.049	0.062	0.214	0.321	0.618
11	-0.045	0.051	0.216	0.298	0.599
12	-0.049	0.040	0.190	0.289	0.585
13	-0.040	0.045	0.167	0.275	0.577
14	-0.039	0.036	0.139	0.262	0.564
15	-0.041	0.029	0.141	0.249	0.541
20	-0.028	0.026	0.110	0.180	0.359
25	-0.023	0.022	0.113	0.182	0.360
30	-0.020	0.019	0.092	0.178	0.327
35	-0.019	0.018	0.088	0.155	0.311
40	-0.016	0.016	0.080	0.131	0.253
45	-0.014	0.015	0.070	0.119	0.233
50	-0.012	0.017	0.064	0.108	0.201
75	-0.007	0.016	0.062	0.091	0.188
100	-0.006	0.016	0.055	0.083	0.180

the two dendrograms are statistically different, at the given significance level. When comparing topologies, all tests are significant ( $r = 1.00$ ) at the 0.01 level, however, illustrating the difference between topological and path-length comparisons. The magnitude of type II error, i.e., of accepting the null hypothesis when it is wrong, remains unknown because we have not conducted a power analysis of these tests (Kraemer and Thieman, 1987; Cohen, 1988).

#### DISCUSSION

We have generated specific tables of significance values to measure the consensus of cladograms and dendrograms, based on matrix correlations. Although our tables are based on  $[0, 1]$  distance matrix correlations, they can also be used to compare trees presented on a similarity scale, as long as both path-length matrices are either similarity or distance. The correlation coefficient is invariant over the rescaling and transformation required to convert similarity values into distances, which is also true for distances not bounded by 0 and 1.

The method proposed to construct the reference distribution of the matrix correlation coefficient could be repeated for other consensus indices, for larger numbers of objects, and even for other types of trees. An interesting extension would be to consider nonparametric indices such as the Spearman rank correlation coefficient or the Procrustes statistic (Gower, 1971, 1975). In the case of dendrograms, the computation of a Spearman rank correlation on ultrametric matrices is identical to a Pearson correlation computed over ranked dendrograms; for the Pearson correlation, the ultrametric distances are replaced by their ranks in the ultrametric matrix. The Procrustes problem, however, consists of translating, rotating, reflecting, and scaling the distance matrix configurations to minimize a goodness-of-fit criterion (Gower, 1975); the "stress" statistic then represents a consensus index that can be tested using a standard permutation procedure. These nonparametric approaches are quite appealing for overcoming the problems

encountered when comparing independent phylogenies based on different distance metrics and derived using distinct methods.

#### ACKNOWLEDGMENTS

We acknowledge the helpful comments of Sam Rogers on earlier drafts of this manuscript. This work was supported by a NSERC scholarship to F.-J. Lapointe and by NSERC grant No. OGP0007738 to P. Legendre. This is contribution No. 385 of the Groupe d'Ecologie des Eaux Douces, Université de Montréal.

#### REFERENCES

- BLANKEN, R. L., L. C. KLOTZ, AND A. G. HINNEBUSCH. 1982. Computer comparison of new and existing criteria for constructing evolutionary trees from sequence data. *J. Mol. Evol.* 19:9-19.
- BUNEMAN, P. 1971. The recovery of trees from measures of dissimilarity. Pages 387-395 in *Mathematics in archaeological and historical sciences* (F. R. Hodson, D. G. Kendall, and P. Tautu, eds.). Edinburgh Univ. Press, Edinburgh.
- COHEN, J. 1988. *Statistical power analysis for the behavioral sciences*, 2nd edition. Erlbaum, Hillsdale, New Jersey.
- DAY, W. H. E. 1986. Foreword: Comparison and consensus of classifications. *J. Classif.* 3:183-185.
- FELSENSTEIN, J. 1978. The number of evolutionary trees. *Syst. Zool.* 27:27-33.
- FELSENSTEIN, J. 1982. Numerical methods for inferring evolutionary trees. *Q. Rev. Biol.* 57:379-404.
- GORDON, A. D. 1987. A review of hierarchical classification. *J. R. Stat. Soc. A* 150:119-137.
- GOWER, J. C. 1971. Statistical methods of comparing different multivariate analyses of the same data. Pages 138-149 in *Mathematics in archaeological and historical sciences* (F. R. Hodson, D. G. Kendall, and P. Tautu, eds.). Edinburgh Univ. Press, Edinburgh.
- GOWER, J. C. 1975. Generalized Procrustes analysis. *Psychometrika* 40:33-51.
- HARTIGAN, J. A. 1967. Representation of similarity matrices by trees. *J. Am. Stat. Assoc.* 62:1140-1158.
- HUBERT, L. J., AND F. B. BAKER. 1977. The comparison and fitting of given classification schemes. *J. Math. Psychol.* 16:233-253.
- JOHNSON, S. C. 1967. Hierarchical clustering schemes. *Psychometrika* 32:241-254.
- KRAEMER, H. C., AND S. THIEMAN. 1987. *How many subjects? Statistical power analysis in research*. Sage, London.
- LAPOINTE, F.-J. 1990. La comparaison statistique de classifications hiérarchiques: Évaluation de trois tests de comparaison par permutations. Pages 255-266 in *Recueil des textes des présentations du colloque sur les méthodes et domaines d'application de la statistique*. Bureau de la statistique du Québec, Québec.
- LAPOINTE, F.-J., AND P. LEGENDRE. 1990. A statistical

- framework to test the consensus of two nested classifications. *Syst. Zool.* 39:1-13.
- LAPOINTE, F.-J., AND P. LEGENDRE. 1991. The generation of random ultrametric matrices representing dendrograms. *J. Classif.* 8:177-200.
- LAPOINTE, F.-J., AND P. LEGENDRE. 1992. A statistical framework to test the consensus among additive trees (cladograms). *Syst. Biol.* 41:158-171.
- MANTEL, N. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Res.* 27:209-220.
- MILLIGAN, G. W. 1981. A Monte-Carlo study of 30 internal criterion measures for cluster-analysis. *Psychometrika* 46:187-195.
- MURTAGH, F. 1984. Counting dendrograms: A survey. *Discrete Appl. Math.* 7:191-199.
- PAGE, R. D. M. 1988. Quantitative cladistic biogeography: Constructing and comparing area cladograms. *Syst. Zool.* 37:254-270.
- PAGE, R. D. M. 1991. Random dendrograms and null hypotheses in cladistic biogeography. *Syst. Zool.* 40:54-62.
- PENNY, D., M. D. HENDY, AND M. A. STEEL. 1992. Progress with methods for constructing evolutionary trees. *Trends Ecol. Evol.* 7:73-79.
- PHIPPS, J. P. 1971. Dendrogram topology. *Syst. Zool.* 20:306-308.
- PHIPPS, J. P. 1975. The numbers of classifications. *Can. J. Bot.* 54:686-688.
- ROHLF, F. J. 1982. Consensus indices for comparing classifications. *Math. Biosci.* 59:131-144.
- SHAO, K., AND F. J. ROHLF. 1983. Sampling distributions of consensus indices when all bifurcating trees are equally likely. Pages 132-137 in *Numerical taxonomy* (J. Felsenstein, ed.). NATO Advanced Studies Institute, Series G (Ecological Sciences), No. 1. Springer-Verlag, Berlin.
- SHAO, K., AND R. R. SOKAL. 1986. Significance tests of consensus indices. *Syst. Zool.* 35:582-590.
- SMOLENSKII, Y. A. 1969. A method for linear recording of graphs. *U.S.S.R. Comput. Math. Phys.* 2:396-397.
- SNEATH, P. H. A., AND R. R. SOKAL. 1973. *Numerical taxonomy*. W. H. Freeman, San Francisco.
- SOKAL, R. R., J. KIM, AND F. J. ROHLF. 1992. Character and OTU stability in five taxonomic groups. *J. Classif.* 9:117-140.
- SOKAL, R. R., AND F. J. ROHLF. 1962. The comparison of dendrograms by objective methods. *Taxon* 9:33-40.
- SWOFFORD, D. L., AND G. J. OLSEN. 1990. Phylogeny reconstruction. Pages 411-501 in *Molecular systematics* (D. M. Hillis and C. Moritz, eds.). Sinauer, Sunderland, Massachusetts.
- TATENO, Y., M. NEI, AND F. TAJIMA. 1982. Accuracy of estimated phylogenetic trees from molecular data. I. Distantly related species. *J. Mol. Evol.* 18:387-404.

Received 11 April 1991; accepted 12 August 1991

*Syst. Biol.* 41(3):384-391, 1992

## Geographic Origin of Human Mitochondrial DNA Revisited

MARK STONEKING, STEPHEN T. SHERRY, AND LINDA VIGILANT

*Department of Anthropology, Pennsylvania State University,  
University Park, Pennsylvania 16802, USA*

In a Points of View section in a previous issue of *Systematic Biology*, Maddison et al. (1992) presented a reanalysis of the DNA sequences of the human mitochondrial DNA (mtDNA) control region that were originally determined and analyzed by Vigilant et al. (1991). Maddison et al. arrived at the following three conclusions: (1) the parsimony analysis performed in the original study was inadequate, resulting in a biased sample of most-parsimonious trees (MPTs); (2) the geographic-states test described in the original study is flawed on several grounds; and (3) the

chimpanzee control region sequences that were used as outgroups may be too distantly related to human control region sequences to provide a meaningful placement of the ancestor on the human tree. We consider each of these conclusions in turn.

### THE PROBLEM WITH PARSIMONY ANALYSES

We agree with the first conclusion reached by Maddison et al., i.e., that the original parsimony analysis was inadequate and found a biased set of trees. This