A Quantitative Approach to Baraminology With Examples from the Catarrhine Primates

D. Ashley Robinson and David P. Cavanaugh*

Abstract

Quantitative methods for identifying holobaramins have yet to be introduced into the field of baraminology. In this report we examine some quantitative methods which may be applied to a variety of biological data to empirically estimate the identity of holobaramins. Organismal relationships are based on a measure of dissimilarity called baraminic distance. A set of diagnostic statistics is described that allows the researcher to assess the completeness, variation, resolving power, and associations within a data set. Bootstrapped dendrograms are constructed to identify clusters of organisms, which are subsequently evaluated for phylogenetic discontinuity by comparing baraminic distance variation, and by correlating sets of baraminic distances. Using this approach both related monobaraminic groups and unrelated apobaraminic groups can be identified. The described methods are illustrated using data from humans and nonhuman primates, a group assumed by baraminologists to be polybaraminic. We have found that baraminic distances based on hemoglobin amino acid sequences, 12S-rRNA sequences, and chromosomal data were largely ineffective for identifying the Human holobaramin. Baraminic distances based on ecological and morphological characters, however, were quite reliable for distinguishing humans from nonhuman primates.

Introduction

The identification of specially created groups of organisms requires a panel of theoretically-defined membership criteria to detect phylogenetic discontinuity. Criteria such as the ability to hybridize are used to lump species into related monobaramins. In contrast, criteria such as morphological and paleontological gaps are used to split species into phylogenetically unrelated apobaramins. Successively refining the monobaraminic and apobaraminic boundaries converges on a taxon called the holobaramin. This unit is defined as a complete set of organisms related by common descent, and is an estimate of a created biological unit (ReMine, 1990; Wise, 1990; 1992; Robinson, 1997). In this paper we examine some multivariate statistical procedures, along with standard phenetic and cladistic procedures, and illustrate how these quantitative methods might be applied to baraminic membership criteria to detect phylogenetic discontinuity. The methods should also provide baraminologists with a tool for characterizing the relative performance of different membership criteria. The catarrhine primates have been selected for analysis because valid baraminic methodology must be capable of distinguishing between biologically similar yet phylogenetically distinct species such as humans and nonhuman primates.

*D. Ashley Robinson, 665 Idlewild Circle, Apt. B10, Birmingham, AL 35205, email: daro@aol.com.
David P. Cavanaugh, 27329 Alberta Drive, Harvest, AL 35749, email: dcavanau@ro.com
Received 27 January 1997; Revised 10 November 1997

Quantitative Baraminology

Phenetic and Cladistic Methodology Applied

Phenetics and cladistics represent the two major quantitative methodologies used in modern biosystematics. Wide interest in biosystematics earlier this century led to a reintroduction and extensive development of both disciplines. Michel Adanson, a creationist and contemporary of Linnaeus, is often credited for founding phenetic taxonomy, while Robert Sokal and Peter Sneath reintroduced it in the 1950s (Sokal and Sneath, 1963). Phenetics uses a measure of biological similarity, often summarized by a resemblance coefficient, as a basis for classification. Since phenetic methods classify by similarity only, they are not expected to generate evolutionary classifications (Sneath and Sokal, 1973). Baraminologists have also suggested that similarity alone does not necessarily imply ancestry (Wise, 1990). It would seem phenetic methods are thus intrinsically favorable to a typological view of nature (see Sokal, 1962).

In general, however, properly weighted similarity has been shown to be a reliable basis for inferring phylogenetic relationships (Mayr, 1969). Species that are highly similar are often related. Within the context of baraminology we would therefore expect that species highly dissimilar are often unrelated. We suggest that a number of baraminic membership criteria including the ecological, morphological, and molecular criteria can be examined quantitatively by measuring discontinuity in terms of dissimilarity.

Cladistic methods were thought to be first used by Plato (Panchen, 1992), and were introduced to biosystematics in 1950 by William Hennig (Hennig, 1950). This approach focuses on the branching patterns of individual characters, according to parsimony or compatibility considerations (Felsenstein, 1982), and generates an inherently evolutionary classification. Wise (1992) suggested that baraminologists utilize cladistic procedures to infer intraholobaraminic relationships. Cladistics further provides the best method for examining the homoplasy and synapomorphy membership criteria.

Phenetic and cladistic methodologies should be considered complementary. Phenetic techniques are ideal for detecting general biological gaps, whereas cladistic methods allows focus on specific character details. Although most of the methods described in this paper are phenetic, we have used cladistic methods to investigate the homoplasy criterion and to compare the hierarchical patterns generated by standard phenetic and cladistic clustering algorithms.

Selecting Organisms

In order to define the baraminology of a group of organisms it is necessary to identify both monobaraminic relationships uniting the group and apobaraminic divisions separating the group from other species. We recommend that taxonomy, or preferably a hypothesized phylogeny, be utilized as a guide for selecting organisms in studies of baraminology. Taxonomic and hypothesized monophyletic groups are most likely to contain all the members of a monobaraminic group.

The selection of outgroups is necessary to test for apobaraminic divisions separating the group of interest from the rest of the biota. Outgroups should represent the nearest taxonomic or hypothesized monophyletic neighbor to the group of interest. For example, if the group of interest formed a monobaraminic family, then it would be preferable to select members of the superfamily to which the family belonged as an outgroup. The discovery of a phylogenetic gap separating a monobaraminic family from its apobaraminic superfamily would fully resolve the holobaraminic identity of the family. In cases where the researcher discovers the outgroup is monobaraminic with the group of interest it would be necessary to select an additional outgroup, taxonomically more distant, and reconduct the analysis. The selection of similar outgroups also improves the comparability of a data set. For example, it would be unwarranted to select turtles as an outgroup for studies of primate baraminology because much of their biology is not comparable.

Selecting and Coding Characters

Holobaramins are defined polythetically because emphasis is placed on using multiple criteria for identification, as opposed to monothetic identification which uses a single criterion such as the ability to hybridize. With the exception of the Scriptural criterion no single data set is sufficient to define the holobaramin. Moreover, it is not known which features of an organism are reliable for determining their baraminic affinity. During the development of baraminology it is therefore desirable to evaluate the performance of a wide variety of biological data for resolving baraminic relationships. Creation biologists have long theorized that specially created *min* contain a recognizable set of morphological, physiological, and genetic characters (Marsh, 1971; 1976). A representative sample of a species' biology is therefore expected to provide baraminically informative data.

Membership criteria as defined by ReMine (1990) and Wise (1990; 1992) refer to broad classes of evidence. We use the term criterion in that general sense, and use the conventional term character to represent the actual datum collected by an investigator. For example, body weight, mandibular shape, and the number of lumbar vertebrae are considered different characters of the morphological criterion. Biosystematic characters exist as discrete or continuous variables. Discrete variables with a finite number of character states such as present or absent are often used in biosystematic studies. Multistate characters such as tan, stripped, or spotted fur would also be categorized as discrete variables because the number of possible character states are limited. Discrete variables are coded by simply assigning a number to a particular character state. For example, the presence or absence of a tail can be coded as 1 and 0 respectively. Since the phenetic measure of resemblance used in this paper is based on the number of mismatched characters (described below), it is not necessary to code the data in a logically ordered manner. The magnitude of the difference between character state 1 and 0 is the same as the difference between state 1 and 8 since both are mismatches. The second type of character used in biosystematics are continuous variables. Such characters can theoretically occupy an infinite number of states such as 124.7 mm and 64.9 kg. Continuous variables are generally subject to clinal variation. For example, body size is a continuous variable that commonly differs according to an environmental gradient. In order to avoid error due to clinal variation it is important to obtain a representative sample of the character from different places on the gradient. Practical comparisons among continuous variables can be made when the data are recoded as discrete variables. The present study employs Sturges' rule (Daniel, 1995) to assign continuous variables a limited number of character states based on the range of variation present in a character. The number of character states are calculated as:

$$K = 1 + 3.322\ (\log n) \qquad (1)$$

where n equals the number of values in the data set. The width of a character state is given as:

$$W = R/K \qquad (2)$$

where R equals the range of the continuous variable (maximum value – minimum value).

It is ideal that characters selected for biosystematic studies represent statistically independent information. For ex-

ample, using both the presence or absence of folivory and the percent of foliage in the diet as different characters does not provide independent evidence of relationship because the latter is clearly dependent on the former. Baraminologists have reason to suspect that organisms are too well designed to consider many characters absolutely independent of each other. Note the analogies given in Scripture such as I Corinthians 12:12-26 suggest that no morphological characters within a given individual are absolutely independent. Moreover, passages such as Psalms 139:13, 15 suggest that the biological structures within an individual are knit or woven together by the Creator. The revealed fact that all individuals have ultimately originated from a common Creator further suggests some degree of dependence, albeit immeasurable in many cases, between the biological characters of all organisms. For example, human behavior resulting in environmental pollution can effect the ecology and even morphology of aquatic insects. Including correlated characters in biosystematic studies will accentuate the overall relationships implied by the characters. We believe all attempts to quantitate baraminic relationships between species will inherently suffer, in a statistical sense, from the inability to select absolutely independent characters for comparison.

The Coefficient of Baraminic Distance

The proportion of character mismatches is used as a basis for evaluating organismal relationships. The coefficient of baraminic distance is calculated as:

$$d_{ij} = \frac{m_{ij}}{n_{ij}} \qquad (3)$$

where m is the number of mismatched characters between the ith and jth organism, and n is the number of compared characters. Characters that are not directly comparable between two organisms such as inapplicable and missing data should be excluded from the calculation. This equation is the complementary form of the simple matching coefficient introduced by Sokal and Michener (1958), and has been used successfully in numerous biological contexts.

As a coefficient, baraminic distance represents an estimate of a true biological distance (D_{ij}). The true distance would theoretically measure the deviation in the plan used by the Creator when originally designing specially created *min*. Biological variation caused by deterministic and stochastic processes, plus incomplete sampling of baraminically informative characters, would cause the coefficient to differ from the true distance.

Characters can be weighted during the calculation of baraminic distances. During the development of baraminology we recommend that all characters be given equal weight. In other words, all characters contribute 0 (match) or 1 (mismatch) to the baraminic distance numerator. Such a natural weighting scheme allows the data to dictate the classification (see Sneath and Sokal, 1973). In artificial weighting schemes the value of a desired weight substitutes for the character mismatch. After it is learned which characters are more reliable for identifying holobaramins it might be advisable to give these characters greater weight. Such *a posteriori* weighting schemes are commonly used in biosystematics (Mayr, 1969). It should be noted, however, that when a large number of characters are analyzed phenetically, artificial weighting has little impact on the resulting classification unless the weights are extreme (Sneath and Sokal, 1973). An example of baraminic distance calculations has been provided in Figure 1.

A

	Character*				
Taxon	A	B	C	D	E
1	0	0	2	0	0
2	0	0	1	2	0
3	1	0	2	3	1
4	1	1	0	0	?
5	2	1	1	1	0

*Characters A, B, and C form criterion I; Characters D and F form criterion II.

B

	1	2	3	4	5
1	-	2/5	3/5	3/4	4/5
2	0.400	-	4/5	4/4	3/5
3	0.600	0.800	-	3/4	5/5
4	0.750	1.000	0.750	-	3/4
5	0.800	0.600	1.000	0.750	-

C

	1	2	3	4	5
1	-	1.5/5	2/5	3/4	3.5/5
2	0.300	-	3/5	3.5/4	2.5/5
3	0.400	0.600	-	2.5/4	4/5
4	0.750	0.875	0.625	-	2.5/4
5	0.700	0.500	0.800	0.625	-

Figure 1. Example of the baranunic distance calculations. Panel A represents a hypothetical data matrix composed of taxon rows and character columns. Panel B depicts a pairwise baraminic distance matrix (lower diagonal) and mismatch matrix (upper diagonal). Artificially weighted baraminic distances, where criterion II is given half the weight of criterion I, are presented in Panel C.

Diagnostic Statistics

A set of statistics has been compiled to summarize the completeness, variation, resolving power, and associations within a given data set. As mentioned by Wise (1992) the accuracy of a baraminic hypothesis is dependent on the reliability of the data used to determine that relationship. Diagnostic statistics might therefore be useful in choosing between competing hypotheses that are based on data sets of different quality.

Character relevance (*a*) measures the completeness of a data set and is given as:

$$a_i = x/n \quad (4)$$

where *x* is the number of organisms to which character *i* is applicable, and *n* is the total number of organisms (Sneath and Sokal, 1973). Criterial relevance (*A*) represents the average of all *a* values for a given criterion. We recommend, rather arbitrarily, that a 95% or greater relevance be obtained in the combined data set since lower values are indicative of missing and/or inapplicable data.

Character diversity (*c*) measures the variation within a data set and is calculated as:

$$c = 1 - \sum x_i^2 \frac{n}{n-1} \quad (5)$$

where *x* is the frequency of the *i*th character state, *n* is the number of organisms, and $n/(n\text{-}1)$ is a correction for small samples (Selander et al., 1986). Criterial diversity (*C*) is calculated as the average of all *c* values for a given criterion. Character diversity is an estimate of the probability that two randomly selected organisms will have dissimilar character states for a given character, while criterial diversity estimates the probability of a mismatch using the average character of a criterion. The average baraminic distance (d_{avg}) offers an alternative method for measuring variation.

Baraminic signal (*S1*) is a chi-square statistic (Daniel, 1995) adapted here for use with character mismatch matrices. The statistic can be used to assess the resolving power of a data set and is calculated as:

$$S1 = \sum \frac{(d_o - d_e)^2}{d_e}. \quad (6)$$

The observed distance (d_o) is simply the number of pairwise mismatches (m_{ij}) between two organisms. Note that the number of mismatches between different pairs of species can be based on different numbers of total compared characters depending on the relevance of the data set. The expected distance (d_e) between each pair of organisms is obtained as the product of the marginal total of character mismatches (i.e. sum of the rows) for the *i*th and *j*th organism, divided by the total number of character mismatches (i.e. sum of row marginal sums) within the data set. Baraminic signal can be tested with a null hypothesis of homogeneity using $[(n^2 - n)/2] - 2n$ degrees of freedom, where *n* equals the number of species. The justification for the degrees of freedom are as follows: (1) Since one-half of a square symmetric distance matrix is redundant, $(n^2\text{-}n)/2$ degrees of freedom are lost. (2) Since the diagonal of a distance matrix represents the distance from an organism to itself (i.e. zero distance), *n* degrees of freedom are lost. (3) Since the rows of the distance matrix are summed, *n* degrees of freedom are lost. Note that the column sums would equal the row sums and have been dropped from the calculation. The corresponding degrees of freedom are therefore not lost as would be the case in a normal contingency table. Species that represent a homogeneous group without statistically significant baraminic signal present exhibit similar baraminic distances and low chi-square values. Two or more groups of organisms well separated on the basis of baraminic distances display large chi-square values.

The Pearson product-moment correlation coefficient (*r*) (Daniel, 1995) can be used to investigate the associations between different criteria. Statistically significant positive correlations between baraminic distances calculated from different biological data sets suggest the data sets detect similar relationships. Correlations between ordered sets of pairwise baraminic distances calculated for each criteria are tested with a null hypothesis of no correlation. Since the pairs of a given distance matrix are not statistically independent, it is not strictly valid to estimate the significance of criterial correlation coefficients in a classical manner. Mantel (1967) provided a solution to this problem by comparing the cross-products of analogous cells in two matrices by the expected value calculated from random permutations. The Bonferroni correction calculated as $(0.05/n)$, where *n* equals the number of pairwise comparisons, was used to establish a probability level (*P*.0009 in this study) for rejecting the null hypothesis of no correlation (Douglas and Endler, 1982). An example of the diagnostic statistics calculations has been provided in Figure 2.

Identifying Baraminic Units

A protocol for identifying baraminic units is here proposed. The first step identifies subgroups of organisms using standard cluster analyses. Alternatively, taxonomy or a hypothesized phylogeny can be used to define subgroups for comparison. The second step evaluates the subgroups for phylogenetic discontinuity by comparing baraminic distance variation and correlating sets of baraminic distances. Baraminic distance variation analyses can involve intragroup versus intergroup comparisons and/or intragroup versus control group comparisons.

Cluster analysis has proven to be an especially useful multivariate method in systematics for hypothesizing the hierarchical arrangement of living creatures. Since a nested hierarchy can be constructed for any given collection of objects, a dendrogram does not necessarily imply its components are evolutionarily related. The fact that a creationist, Linnaeus, developed the current hierarchical system of biological classification indicates that macroevolutionary theory is unnecessary in biosystematics (ReMine, 1993), and that biological hierarchy is compatible with a creationist view of nature.

One caveat of inferring biological hierarchy from distances is that unequal rates of evolution can produce am-

biguous results (Felsenstein, 1982). We therefore utilize the neighbor-joining algorithm (Saitou and Nei, 1987), which generates a single most parsimonious tree free from the assumption that taxa have evolved at a constant rate. We recommend that bootstrapping (Felsenstein, 1985), a process whereby significant clusters of species are recovered after repeated randomization and reanalysis of the data, be utilized to place confidence limits on the branching patterns within a dendrogram. Since a bootstrap-supported cluster of species may often be divisible into numerous holobaramins, baraminologists must include an additional step of testing clusters for phylogenetic discontinuity. It is much simpler to assume all creatures are related, as macroevolutionists do, and disregard the challenges associated with defining phylogenetically isolated groups of organisms.

Comparisons of baraminic distance variation provides one method for detecting possible phylogenetic discontinuity. For groups yielding more than 30 comparisons, a standard confidence interval for a population mean (Daniel, 1995) can be constructed as:

$$d \pm z\frac{s}{\sqrt{n}} \quad (7)$$

where d equals the average baraminic distance of the group, z represents the desired reliability coefficient, s equals the standard deviation, and n denotes the number of comparisons. Comparisons of less than 30 should replace z with the corresponding critical value from the t distribution.

Two different types of baraminic distance variation analyses can be performed. One type of analysis compares intragroup with intergroup baraminic distance variation. Overlapping confidence intervals suggest the intragroup baraminic distance variation is not significantly different than the intergroup baraminic distance variation, and may indicate a monobaraminic relationship uniting the groups. A gap between the intragroup and intergroup confidence intervals indicates the range of baraminic distance variation is significantly different, which may support an apobaraminic division between the groups.

Baraminic distance variation can also be used to detect putative similarity thresholds. This type of analysis determines whether intragroup baraminic distance variation is greater or less than that occurring within a control group. Previous creationist work has suggested that one family within the Catarrhini, the Cercopithecidae or old world monkeys, form a basic type on the basis of hybridization (Hartwig-Scherer, 1993). If similarity thresholds exist within the Catarrhini, then the variation between humans and nonhuman primates (the control group) should be greater than the variation occurring within a related monobaraminic group such as the Cercopithecidae (the intragroup).

Baraminic units can be separately elucidated with a correlation analysis of the baraminic distances of each pair of organisms. Since this analysis is not a matrix-level comparison (as is the case for the criterial correlation analysis), the significance of Pearson product-moment correlation coefficients can be estimated in a classical manner with $n - 2$ degrees of freedom, where n equals the number of species.

A

Character	*a*
A	1.000
B	1.000
C	1.000
D	1.000
E	0.800

A = 0.960

B

	Frequency of indicated character state				
Character	0	1	2	3	*c*
A	0.400	0.400	0.200		0.550
B	0.600	0.400			0.350
C	0.200	0.400	0.400		0.550
D	0.400	0.200	0.200	0.200	0.650
E	0.750	0.250			0.167

C = 0.453

C

	1	2	3	4	5
1	-	2	3	3	4
2	2.294	-	4	4	3
3	2.647	2.868	-	3	5
4	2.294	2.485	2.868	-	3
5	2.267	2.868	3.309	2.868	-

S1 = 3.880

D

	1	2	3	4	5
1	-	0.500	1.000	0.000	0.500
2	0.333	-	1.000	1.000	0.500
3	0.333	0.667	-	1.000	1.000
4	1.000	1.000	0.667	-	1.000
5	1.000	0.667	1.000	0.667	-

r = 0.106

Figure 2. Example of the diagnostic statistics calculations using the unweighted data matrix of Figure 1. Panels A and B represent relevance and diversity calculations respectively. Panel C illustrates baraminic signal calculations with expected (lower diagonal) and observed (upper diagonal) character mismatches. Panel D provides a pairwise matrix of baraminic distances based on criterion I (lower diagonal) and criterion II (upper diagonal) with the corresponding product-moment correlation coefficient.

Positive correlations indicate the compared organisms share a similar set of distances that differ from other organisms by a similar magnitude. This condition may therefore be diagnostic of a monobaraminic relationship. In contrast, negative correlations indicate the compared species contain antithetical patterns of baraminic distances. Taxa similar to species X would be dissimilar to species Y and vice versa, which suggests species X and Y may be classified as apobaraminic. The baraminic relationship between organisms with uncorrelated baraminic distances is considered unresolved, but a monobaraminic affinity or apobaraminic discontinuity can be cautiously hypothesized with well supported trends.

Although correlations of baraminic distances can greatly aid holobaraminic identification, one caution is advised. The ability to define a positive or negative correlation is dependent on the species selected for analysis. By including and excluding various taxa the geometric relationships defined by their baraminic distances can change, the resolution and specificity of the analysis can be altered, the sign of the correlation coefficient can vary, and thus the baraminic classification of the organisms can differ. An example of the organismal correlation analysis can be seen in Figure 3.

Materials and Methods

We selected 204 characters (Appendix I) representing 11 species (Table I) within the Catarrhini infraorder. The complete data matrix is available upon request. A variety and quantity of biological data was selected with the intent of maximizing criterial relevance (*A*). Only polymorphic characters were included in the analysis because invariable attributes are not useful for distinguishing organisms. We categorized the data into four general criteria including ecological, morphological, chromosomal, and molecular characters. Since Scriptures clearly imply that humans were specially created (Genesis 1:26-27, 2:7, 22), and thus phylogenetically distinct from other organisms, we utilize the human-nonhuman primate relationship as a control.

The CLUSTAL W program (Thompson, Higgins, and Gibson, 1994) was used to align the molecular sequence data. Calculations of baraminic distance, relevance, diversity, and signal were accomplished with a Macintosh computer program developed by the first author. The MANTEL 3.0 program of the R package (Legendre and Vaudor, 1991) was used to perform Mantel's test for estimating the significance of criterial correlations. The DATA DESK 3.0 statistical package (Odesta Corporation, Northbrook, Illinois) was used for the organismal correlation analyses. A neighbor-joining dendrogram (Saitou and Nei, 1987) based on baraminic distances was generated using the NEIGHBOR program of the PHYLIP 3.54 computer package (Felsenstein, 1989). The PAUP 3.1.1 computer program (Swofford, 1993) was used for the cladistic analysis. All characters were unordered and unweighted. A 50% majority rule consensus tree was constructed using the heuristic search option, random addition of taxa, MAXTREES set to 100, and TBR branch swapping parameters. Both the phenetic and cladistic dendrograms were statistically evaluated with 200 bootstrap iterations.

Table I. List of species included in this study.

Taxon	Common Name	Code
Infraorder Catarrhini		
Superfamily Cercopithecoidea		
Family Cercopithecidae		
Macaca mulatta	macaque	Mmu
Cercoce bus torquatus	mangabey	Cto
Paplo cynocephalus	baboon	Pcy
Mandrill us sphinx	mandrill	Msp
Cercopitbecus aethiops	green monkey	Cae
Superfamily Hominoidea		
Family Hylobatidae		
Hylobates lar	gibbon	HIa
Family Pongidae		
Pongo pygmaeus	orangutan	Ppy
Pan paniscus	pygmy chimpanzee	Ppa
Pan troglodytes	chimpanzee	Ptr
Gorilla gorilla	gorilla	Ggo
Family Hominidae		
Homo sapiens	human	Hsa

Results

Evaluation of Criteria

The combined data set was applicable to an average of 95.5% of the species (Table II). The more relevant criteria tended to contain more characters. For example, there was little missing data among the 139 molecular characters, whereas the 18 ecological characters were only applicable among an average of 78.8% of the organisms. Species differed on average among 43.9% of their characters with a 34.8% probability of a mismatch at the average character. The more diverse criteria tended to contain the least number of characters and vice versa, which may suggest the data most often collected was simple rather than complex variables. The probability of a mismatch between two organisms ranged from 62.7% when comparing the four chro-

Table II. Summary of data used to characterize the catarrhine primates.

	No. of	Diagnostic Statistics			
Criteria	Characters	*A*	*C*	d_{avg}	*S1*
Combined	204	0.955	0.348	0.439	743.471**
Ecological	18	0.788	0.465	0.514	60.276*
Morphological	43	0.915	0.419	0.505	262.172**
Chromosomal	4	0.886	0.627	0.715	25.105
Molecular	139	0.990	0.303	0.404	548.333**

* $P < 0.005$
** $P < 0.0005$

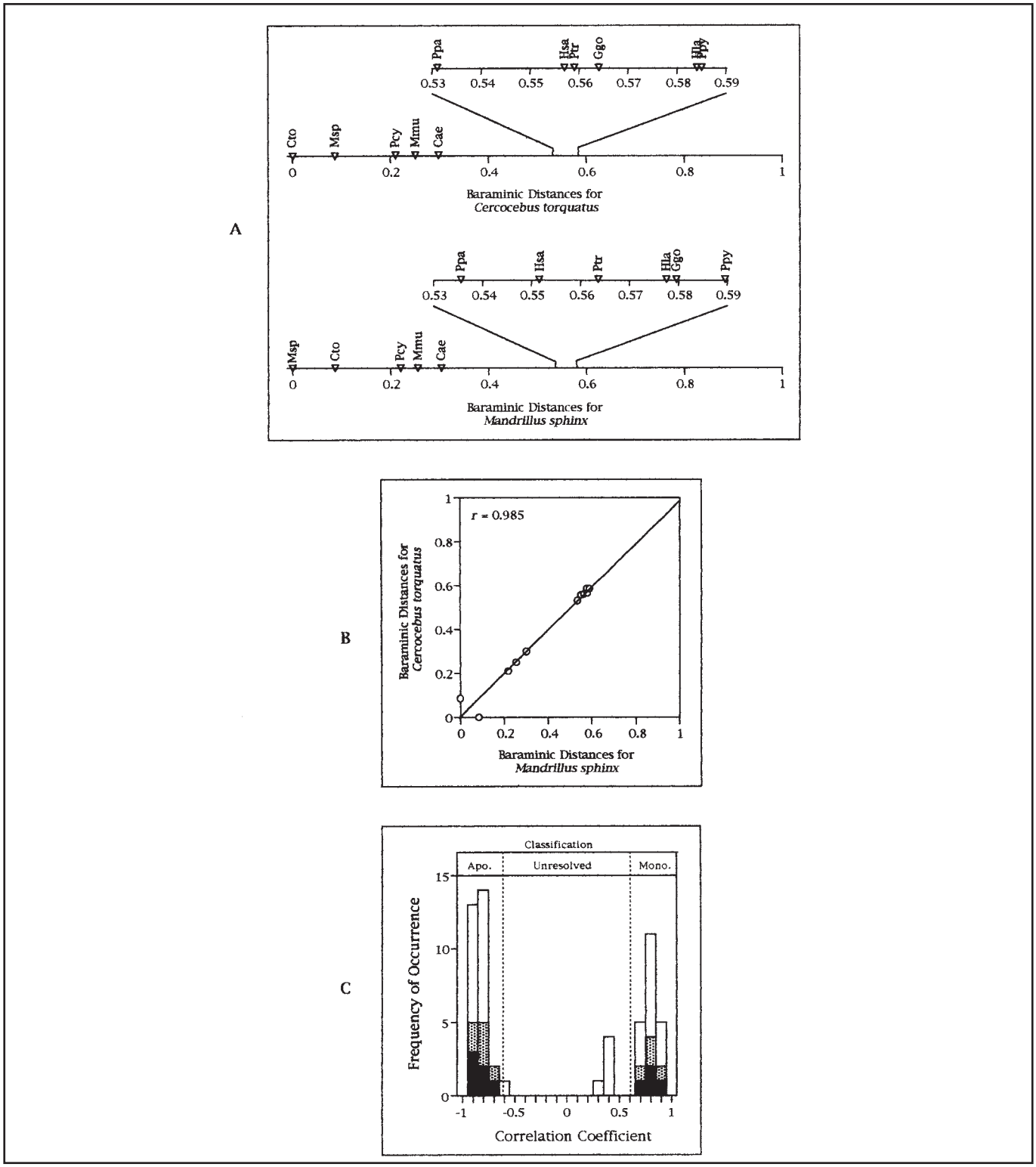

Figure 3. Example of the organismal correlation analysis. Panel A represents the baraminic distances of two cercopithecid species with other catarrhine primates. Panel B illustrates the significant positive correlation between their baraminic distances. Panel C summarizes all baraminic distance correlations within the catarrhine primates. The dotted lines denote critical values for a 95% probability that the correlation is not zero. Black bars represent *Mandrillus sphinx* data, spotted bars denote *Cercoce bus torquatus* data, and white bars represent data for the remaining catarrhine primates. The two cercopithecid species are positively correlated with the other cercopithecids confirming their monobaraminic relationship (Hartwig-Scherer, 1993), and negatively correlated with all other examined species suggesting an apobaraminic relationship.

mosomal characters to 30.3% when comparing the molecular data. The combined data set contained a highly significant ($P < 0.0005$) level of baraminic signal, which suggested a heterogeneous assemblage of organisms was sampled. Baraminic signal was also significantly present among the ecological, morphological, and molecular data. Note that criteria with only modest baraminic signal can be combined to increase the net signal. This is one justification for analyzing a composite data matrix (given that the characters are baraminically informative) and generating a systematic hypothesis based on the total evidence available. It should be further noted that an inverse relationship existed between criterial diversity and signal. Among the more diverse criteria it is possible that a larger number of character states resulted in an increased proportion of noise. The baraminic distances calculated from the morphological characters were significantly associated ($P < 0.0009$) with the distances calculated from ecological, chromosomal, and molecular characters (Table III). The ecological baraminic distances were also correlated with the chromosomal baraminic distances. These results demonstrate the nonindependence of widely different characters from phylogenetically distinct species such as humans and nonhuman primates.

Table III. Product-moment correlation matrix of baraminic distances calculated for each pair of criteria.

	Ecological	Morphological	Chromosomal	Molecular
Ecological	–			
Morphological	0.857*	–		
Chromosomal	0.429*	0.581*	–	
Molecular	0.391	0.603*	0.344	–

* $P < 0.0009$

Cluster Analyses

Congruence existed between the phenetic and cladistic dendrograms (Figure 4). The Cercopithecid monobaramin was supported with 100% bootstrap support in both analyses. Furthermore, the relative branching order of the old world monkeys was stable in both analyses. Phenetic clustering algorithms such as the neighbor-joining method can therefore be as useful as cladistics for defining intraholobaraminic relationships. The gibbon was distinguished from the cercopithecids and pongids on both dendrograms. Within the pongids the only topological discrepancy related to the position of the orangutan. Note that only minimal bootstrap support separated humans from the pongids (Figure 4). Three groups corresponding to the cercopithecid, hylobatid, and pongid clusters were selected for subsequent comparisons with the humans.

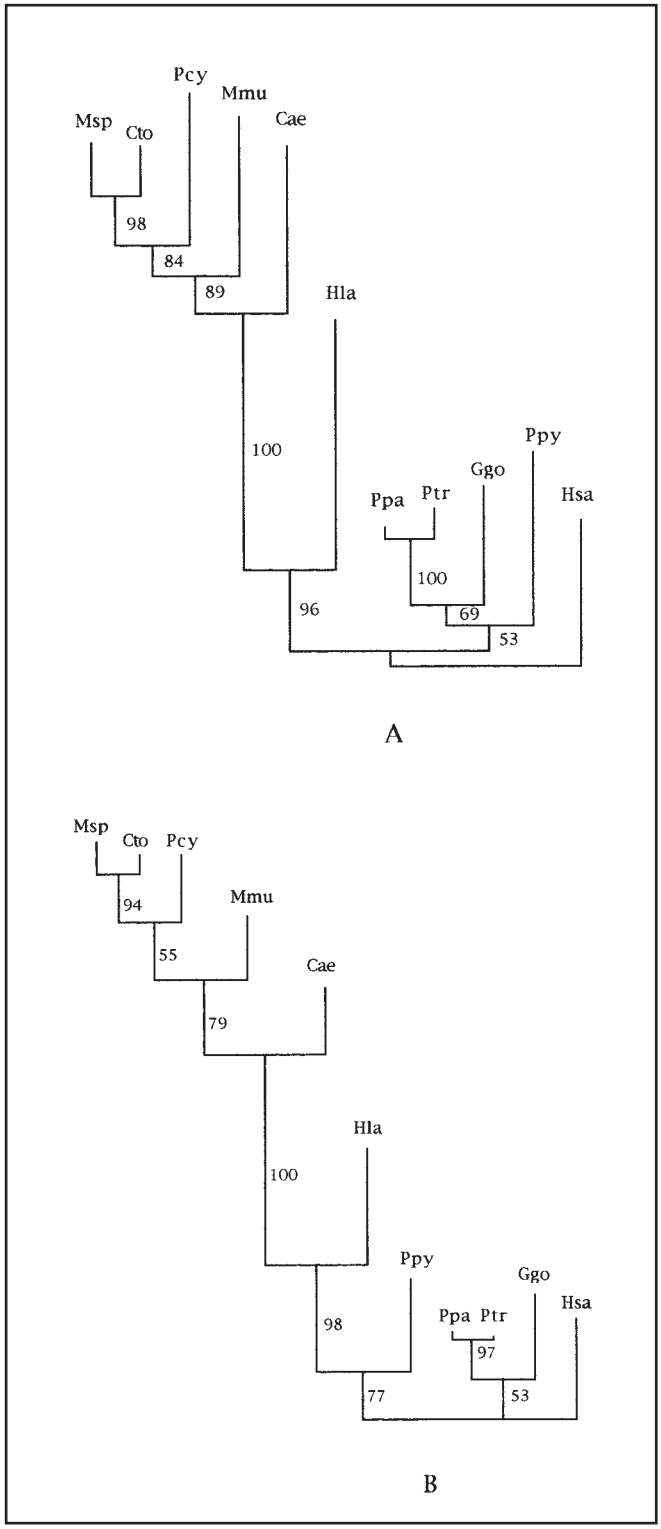

Figure 4. Outgroup-rooted neighbor-joining (A) and cladistic (B) dendrograms. Bootstrap values supporting branches with at least 50% confidence are presented.

Baraminic Distance Variation Analyses

A pairwise matrix of baraminic distances has been compiled in Table IV . The distances ranged from 1.6% between the two chimpanzee species, to 63.6% between the gibbon and macaque. The average distance between humans and nonhuman primates was 43.3%. The combined data set was insensitive to a putative similarity threshold (an intragroup

Table IV. Baraminic distance matrix listing the proportion (lower diagonal) and number (upper diagonal) of character mismatches.

	Mmu	Cto	Pcy	Msp	Cae	Hia	Ppy	Ppa	Ptr	Ggo	Hsa
Mmu	–	47/188	53/189	48/188	58/193	119/187	125/202	99/186	114/202	123/202	114/192
Cto	0.25	–	39/186	16/187	56/188	101/173	110/188	95/179	105/188	106/188	102/183
Pcy	0.280	0.210	–	41/186	64/187	108/174	117/189	103/178	114/189	118/189	112/183
Msp	0.255	0.086	0.220	–	57/188	100/173	111/188	96/179	106/188	109/188	101/183
Cae	0.301	0.298	0.342	0.303	–	100/178	119/194	95/180	107/194	114/194	103/185
Hia	0.636	0.584	0.621	0.578	0.562	–	76/188	67/171	82/188	79/188	76/178
Ppy	0.619	0.585	0.619	0.590	0.613	0.404	–	41/187	52/204	53/204	64/194
Ppa	0.532	0.531	0.579	0.536	0.528	0.392	0.219	–	3/187	26/187	35/182
Ptr	0.564	0.559	0.603	0.564	0.552	0.436	0.255	0.016	–	38/204	44/194
Ggo	0.609	0.564	0.624	0.580	0.588	0.420	0.260	0.139	0.186	–	55/194
Hsa	0.594	0.557	0.612	0.552	0.557	0.427	0.330	0.192	0.227	0.284	–

versus control group comparison) since the range of baraminic distance variation between humans and pongids overlapped with that found in the Cercopithecid monobaramin (Table V). For example, the average intracercopithecid baraminic distance was 23.3% with a 95% confidence interval of 30.0% to 16.6%. The average human-pongid baraminic distance was 25.8% with a 95% confidence interval of 35.5% to 16.1%. Since these intervals overlapped, the differences between humans and pongids did not differ significantly from that found within a monobaramin. Because the combined data set was influenced mainly by the molecular data, it was not surprising that the hemoglobin amino acid sequences and the 12S-rRNA sequences were also incapable of distinguishing humans on the basis of baraminic distance variation. The chromosomal data also failed to mark humans as distinct. A phyletic division between humans and nonhuman primates was, however, strongly supported by the variation in morphological and ecological baraminic distances. These criteria defined gaps in baraminic distances of 20.4% and 45.7% respectively (see Table V).

Organismal Correlation Analyses

Fifteen organismal correlation matrices have been summarized in Figure 5. Panels A-C reflect comparisons using the combined data set. With all catarrhines included in the analysis (panel A) the majority of comparisons were either significantly positive or negative. A lack of resolution predominated when only the hominoidea (panel B) and pongidae (panel C) were included in the analysis. These data confirm the baraminic signal results that suggested greater resolving power was obtained from the larger data set. The specificity of the combined data set, however, was poor because humans were significantly and positively correlated with the pongids in panel A, suggesting an incorrect monobaraminic relationship. The same trends were observed with organismal correlations based on the chromosomal (panels J–L) and molecular (panels M–O) data. These results agree with the analyses of baraminic distance variation, which demonstrated the combined, chromosomal, and molecular data could not reliably distinguish humans from nonhuman primates. It is very encouraging to note that correlations of baraminic distances from ecological (panels D–F) and morphological (panels G–I) data provided adequate specificity for distinguishing the Human holobaramin. All ecological comparisons and 15/19 morphological comparisons yielded negative correlations between humans and nonhuman primates, suggesting a correct apobaraminic relationship. In summary, both the baraminic distance variation analyses and organismal correlation analyses confirmed the reliability of the ecological and morphological criteria for defining gaps between humans and nonhuman primates.

Table V. Comparisons of intragroup with control group baraminic distance variation.

	Average Baraminic Distance (±95% C.I.) of Humans Versus			Cercopithecidae
Criteria	Cercopithecidae	Hylobatidae	Pongidae	Only
Combined	0.574 (0.607, 0.541)	0.427	0.258 (0.355, 0.161)	0.233 (0.300, 0.166)
Ecological	0.745 (0.802, 0.688)	0.800	0.767 (0.935, 0.599)	0.096 (0.142, 0.050)
Morphological	0.755 (0.779, 0.731)	0.707	0.628 (0.951, 0.305)	0.058 (0.101, 0.015)
Chromosomal	0.900 (1.000, 0.622)	1.000	0.750 (1.000, 0.291)	0.364 (0.703, 0.025)
Molecular	0.512 (0.552, 0.472)	0.290	0.108 (0.192, 0.024)	0.285 (0.364, 0.206)
n	5	1	4	11

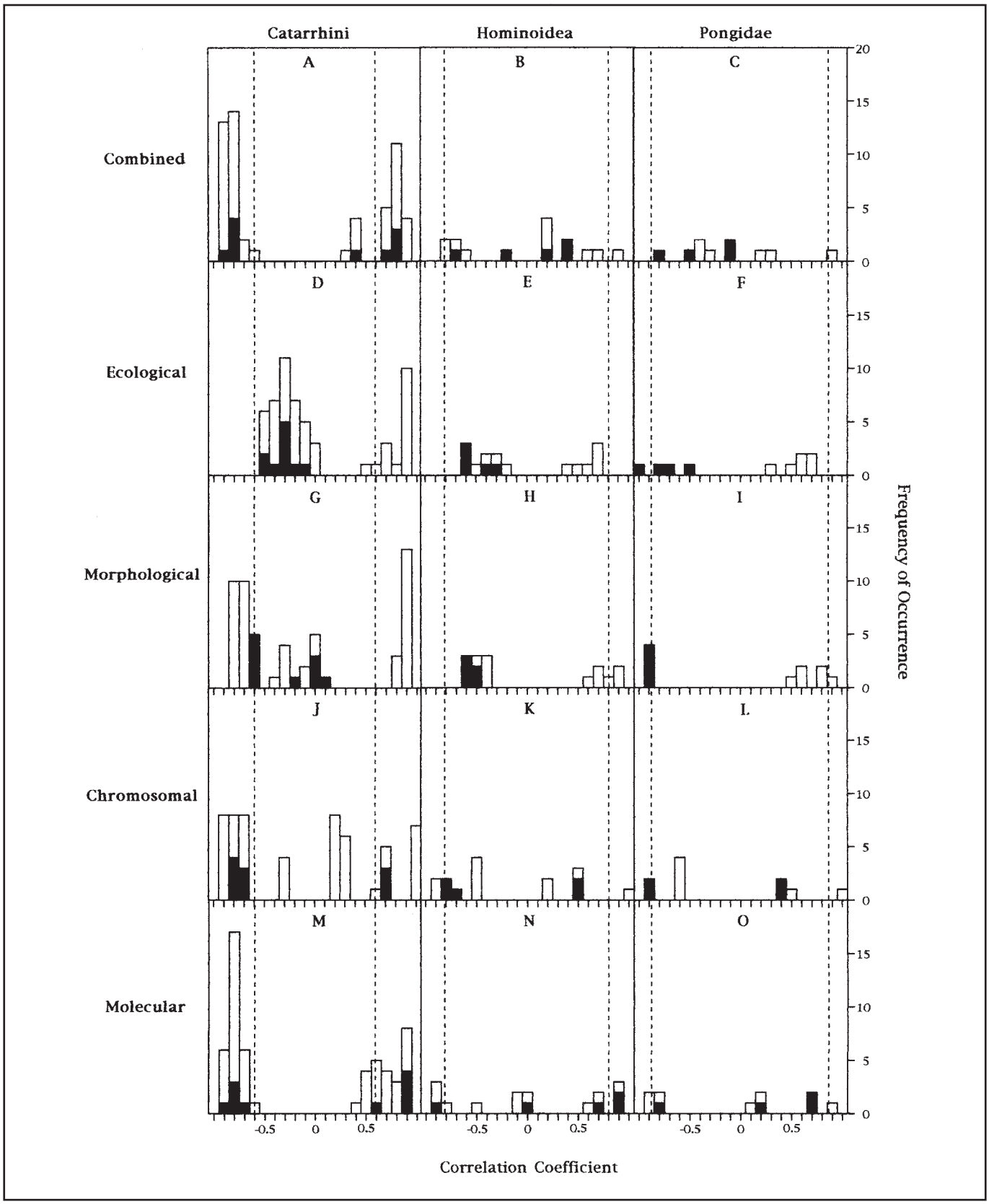

Figure 5. Exhaustive summary of the organismal correlation analyses based on different groups of taxa and different data sets. The dotted lines represent critical values for a 95% probability that the correlation is not zero. Black bars denote humans-nonhuman primate comparisons, and white bars represent comparisons among nonhuman primates.

Evaluation of the Homoplasy Criterion

Homoplasies are readily identified as characters that occur in disjoint branches of a cladogram. We utilize the homoplasy index, which measures the goodness-of-fit of a data set to a cladogram, to quantitate and compare homoplasy. The index is calculated as:

$$H.I. = 1 - \frac{(S - C)}{L} \quad (8)$$

where S equals the total number of character states, C equals the number of characters, and L equals the cladogram length. The homoplasy index is the complementary form of the consistency index (Kluge and Farris, 1969), and is scaled from 0 to 1 in order of increasing homoplasy. Bias resulting from a direct relationship between the index and the cladogram length can limit its utility in comparative studies (see Archie, 1989 and Farris, 1989 but also De Queiroz and Wimberge, 1993).

In evolutionary theory similar selective pressures are often suggested to be the driving force that leads to homoplasy. Another explanation for homoplasy offered by Scherer (1993) suggests that scattering characters from a complex ancestor via the capacity for interspecific hybridization might result in a mosaic network of characters. Lammerts and Howe (1974) have also implicated hybridization as a mechanism for propagating diversity within the created kind. According to these hypotheses, large numbers of homoplasies within a holobaramin may be indicative of frequent gene flow. It should be noted that homoplasy could also be the result of errors in data collection or cladogram construction, or it could be due to the expression of latent genetic information in disjoint lineages. Since members of different holobaramins are not expected to be reproductively compatible, homoplasy between holobaramins would highlight the baraminically uninformative characters. Such attributes could have originated by the activity of a common Creator, but could also represent an interholobaraminic transfer of genetic information via an unknown vector.

The majority of homoplasies in the present survey were found within the molecular data set (Table VI), which confirmed the baraminic distance variation and organismal correlation data that suggested the hemoglobin and 12S-rRNA data was largely baraminically uninformative. Within the Cercopithecidae only molecular characters were homoplasious. This result was somewhat surprising. Since many of the cercopithecids are known to hybridize (Hartwig-Scherer, 1993), we might have expected a larger amount of homoplasy within this group. The ecological and morphological characters exhibited comparable levels of homoplasy as found in other studies (De Queiroz & Wimberge, 1993).

Note that homoplasies can exist within monobaraminic groups (Robinson, 1997). The molecular criterion actually demonstrated more homoplasy within the Cercopithecid monobaramin, than the ecological and chromosomal criteria exhibited within the Catarrhine apobaramin (Table VI). The number of homoplasies between any given pair of species was roughly proportional to their baraminic distance (Figure 6). The highest frequency of homoplasy occurred between cercopithecoids and hominoids, which were the most divergent taxa. Humans exhibited morphological and molecular homoplasies with the Cercopithecidae and Hylobatidae, but no homoplasies with the Pongidae. In order for homoplasy to have been a reliable indicator of baraminic relationships, it would have been desirable to see a spike in homoplasy frequency between humans and pongids. We therefore recommend that the homoplasy criterion be examined in greater detail before using it to identify baraminic units.

Table VI. Comparisons of homoplasy within the Catarrhine apobaramin and Cercopithecid monobaramin.

	Catarrhini		Cercopithecidae	
Criteria	Length	H.I.	Length	H.I.
Combined	380	0.208	128	0.125
Ecological	49	0.143	7	0.000
Morphological	92	0.130	7	0.000
Chromosomal	13	0.000	3	0.000
Molecular	234	0.291	118	0.195

Note: Length refers to the total number of character changes required by the indicated cladogram, and H.I. is an abbreviation for the homoplasy index. The cercopithecid values were obtained by pruning all other taxa from the cladograms.

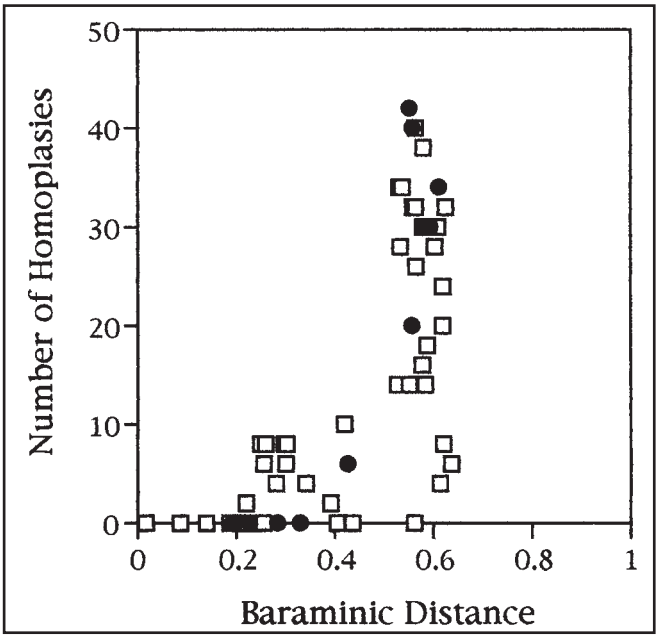

Figure 6. Plot of pairwise baraminic distances versus the number of pairwise homoplasies detected in the cladistic analysis. Black circles denote human-nonhuman primate comparisons, and white squares represent comparisons among nonhuman primates.

Discussion

Character selection, not the method of analysis, is expected to be the primary factor affecting baraminic hypotheses. False conclusions can be reached unless baraminically informative data has been sampled. Since we have no *a priori* knowledge regarding which characters are more reliable for identifying holobaramins, it is important to evaluate the reliability of a wide variety of biological data for inferring baraminic relationships.

New insight into the value of different criteria has been provided by quantitative phenetic and cladistic testing of the Catarrhine primates. Investigations of the molecular criterion questioned the use of hemoglobin amino acid sequences and 12S-rRNA sequences for identifying holobaramins. Although these data readily distinguished humans from cercopithecids they did not efficiently distinguish humans from other hominoids. The variation in hemoglobin amino acid sequences may be simply too limited to provide adequate specificity in baraminic studies. Variation in 12S-rRNA sequences are also limited possibly because of functional constraints (consider that 12S-rRNA must interact properly with DNA, RNA, and protein to form a functional ribosome).

Robinson (1997) demonstrated that the sequence variation among six mitochondrial tRNA genes between humans and chimpanzees was within the range of variation of monobaraminic turtles. Since appreciable levels of variation in hemoglobin, 12S-rRNA, and tRNA sequences may occur at taxonomic levels higher than the holobaramin, we recommend caution in using these specific molecules to identify baraminic units. In contrast, protein-coding genes such as cytochrome *b* seem to provide adequate specificity for identifying baraminic units (Robinson, 1997; Robinson and Cavanaugh, 1997). The chromosomal criterion was of limited utility in this survey possibly because too few characters were sampled. Consequently, baraminic distances calculated from criteria composed of only a few characters should be interpreted with caution.

Jones (1982) and Wise (1992) encouraged the use of ecological data in baraminic studies. Robinson (1997) questioned the specificity of the ecological and trophic criteria because a variety of habitats and trophic categories were found among monobaraminic turtles. The present survey demonstrated that the baraminic distances calculated from several ecological characters provided some of the strongest evidence for discontinuity. These data highlight the importance of using numerous characters to represent a criterion. No single character should be expected to identify a holobaramin.

The morphological criterion provided another reliable data set for defining discontinuity. It is interesting to note that the ecological and morphological criteria were the most adept at distinguishing humans and the most highly correlated, indicating that the data sets in strongest agreement were the most reliable.

Appendix I

Ecological characters: gestation length, weaning age, length of estrous cycle, female age at first breeding, male and female age at sexual maturity, lifespan, interbirth interval, arboreal and terrestrial habitats, foliovory, single and multiple male breeding groups, monogamy, population group size and density, percent foliage in diet, home range size (Clutton-Brock and Harvey, 1979; Harvey and Clutton-Brock, 1985)

Morphological characters: male and female body weight, neonatal body weight, adult and neonatal brain weight, incisor crown compressed, mesial groove and sexual dimorphism of canines, premolar compressed and single cusp, two cusps, hypoconulid, buccal cingulum, talonid, third molar size and width, hypocone, lingual cingulum, protoconule, second molar size, mandibular depth and shape, choanal shape, inter-orbital distance, length of ulnar olecranon, ulnar-carpal articulation, astragalo-calcaneal joint, tail, number of lumbar, sacral, and caudal vertebrae, arm-trunk, leg-trunk, appearance and ontogeny of ischial callosities, foot-free hind limb, tarsus-foot, longest metatarsal-foot, longest free toe-foot, hallucal-foot, hallucal-third radius, axony (Ankel, 1972; Delson and Andrews, 1975; Lessertisseur and Jouffroy, 1975; Harvey and Clutton-Brock, 1985)

Chromosomal characters: diploid number, number of submetacentric and acrocentric chromosomes, Y chromosome morphology (Chiarelli, 1975)

Molecular characters: 25 polymorphic sites for the alpha and beta chains of hemoglobin, 114 polymorphic sites for 12S-rRNA (GenBank/EMBL accession numbers: P01925, P02026, P01933, P02031, P01931, P02030, P02028, P01926, P02025, P06635, P01923, P02024, P01922, P02023, 360949, 360950, 223012, L35203, L35204, L35184, L35196, L35207, L35201, L35209, D38115, X99256, X93340, V00710)

Acknowledgements

We thank Dr. Kurt P. Wise for critical reading of an earlier manuscript, and Dr. Eugene F. Chaffin for editorial remarks. We also thank Todd C. Wood, Walter J. ReMine, and Dr. Frank Maas for helpful comments during various stages of the development of this paper.

References

CRSQ- Creation Research Society Quarterly

Ankel, F. 1972. Vertebral morphology of fossil and extant primates. In Tuttle, R. (editor), *The functional and evolutionary biology of primates*. Aldine. Chicago, IL. pp.223-240.

Archie, J. W. 1989. Homoplasy excess ratios: New indices for measuring levels of homoplasy in phylogenetic systematics and a critique of the consistency index. *Systematic Zo-*

ology 38(3):253-269.

Chiarelli, B. 1975. The study of primate chromosomes. In Tuttle, R. (editor), *Primate functional morphology and evolution*. Mouton Publ. The Hague. pp.103-127.

Clutton-Brock, T. H. and P. H. Harvey. 1979. Home range size, population density and phylogeny in primates. In Bernstein, I.S. and E. O. Smith (editors), *Primate ecology and human origins: Ecological influences on social organization*. Garland STPM Press. New York. pp.201-214.

Daniel, W. W. 1995. *Biostatistics: A foundation for analysis in the health sciences*, sixth edition. John Wiley and Sons, Inc. New York.

De Queiroz, A. and P. H. Wimberge. 1993. The usefulness of behavior for phylogeny estimation: Levels of homoplasy in behavioral and morphological characters. *Evolution* 47(1):46-60.

Delson, E. and P. Andrews. 1975. Evolution and interrelationships of the Catarrhine primates. In Luckett, W.P. and F. S. Szalay (editors), *Phylogeny of the primates: A multidisciplinary approach*. Plenum Press. New York, NY. pp.405-446.

Douglas, M. E. and J. A. Endler. 1982. Quantitative matrix comparisons in ecological and evolutionary investigations. *Journal of Theoretical Biology* 99:777-795.

Farris, J. S. 1989. The retention index and homoplasy excess. *Systematic Zoology* 38(4):406-407.

Felsenstein, J. 1982. Numerical methods for inferring evolutionary trees. *The Quarterly Review of Biology* 57(4):379-404.

———. 1985. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* 39:783-791.

———. 1989. PHYLIP: Phylogeny Inference Package (version 3.2). *Cladistics* 5: 164-166.

Hartwig-Scherer, S. 1993. Hybridisierung und Artbildung bei den Meerkatzenartigen (Primates, Cercopithecoidea). In Scherer, S. (editor), *Typen des Lebens*. Pascal Verlag. Berlin. pp. 245-257.

Harvey, P. H. and T. H. Clutton-Brock. 1985. Life history variation in primates. *Evolution* 39(3):559-581.

Hennig, W. 1950. *Grundzuge einer theorie der phylogenetischen systematik*. Deutscher Zentralverlag. Berlin. 370 pp.

Jones, A. J. 1982. A creationist critique of homology. *CRSQ* 19:156-175.

Kluge, A. G. and J. S. Farris. 1969. Quantitative phyletics and the evolution of anurans. *Systematic Zoology* 18:1-32.

Lammerts, W. E. and G. F. Howe. 1974. Plant succession studies in relation to microevolution. *CRSQ* 10:208-28.

Legendre, P. and A. Vaudor. 1991. *The R package: multidimensional analysis, spatial analysis*. Departement de Sciences Biologiques, Universitede Montreal. 142 pp.

Lessertisseur, J. and F. K. Jouffroy. 1975. Comparative osteometry of the foot of man and facultatively bipedal primates. In Tuttle, R. (editor), *Primate functional morphology and evolution*. Mouton Publ. The Hague. pp.327-340.

Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Research 27:209-220.*

Marsh, F. L. 1971. The Genesis kinds in the modern world. In W. E. Lammerts (editor), *Scientific studies in special creation*. Baker Book House. Grand Rapids, MI. pp. 136-155.

———. 1976. *Variation and fixity in nature*. Pacific. Omaha, NE. 150 pp.

Mayr, E. 1969. *Principles of systematic zoology*. McGraw-Hill Book Co. New York, NY. pp. 428.

Panchen, A. L. 1992. *Classification, evolution, and the nature of biology*. Cambridge Univ. Press. Cambridge. 403 pp.

ReMine, W. J. 1990. Discontinuity systematics: A new methodology of biosystematics relevant to the creation model. In Walsh, R. E. and C. L. Brooks (editors). *Proceedings of the second international conference on creationism*. Volume 2. Creation Science Fellowship. Pittsburgh, PA. pp. 207-213.

———. 1993. *The biotic message: Evolution versus message theory*. St. Paul Science. St. Paul, MN. 538 pp.

Robinson, D. A., 1997. A mitochondrial DNA analysis of the testudine apobaramin. *CRSQ* 33:262-272.

Robinson, D. A., and D. P. Cavanaugh. 1997. Evidence for a holobaraminic origin of the cats. *CRSQ* Submitted.

Saitou, N. and M. Nei. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* 4:406-425.

Scherer, S. 1993. Basic types of life. In S. Scherer (editor), *Typen des Lebens*. Pascal Verlag. Berlin. pp. 11-30.

Selander, R. K., D. A. Caugant, H. Ochman, J. M. Musser, M. N. Gilmour, and T. S. Whittam. 1986. Methods of multilocus enzyme electrophoresis for bacterial population genetics and systematics. *Applied and Environmental Microbiology* 51:873-884.

Sneath, P. H. A. and R. R. Sokal. 1973. *Numerical taxonomy*. W. H. Freeman & Co. San Francisco, CA. 573 pp.

Sokal, R. R. 1962. Typology and empiricism in taxonomy. *Journal of Theoretical Biology* 3:230-267.

Sokal, R. R., and C. D. Michener. 1958. A statistical method for evaluating systematic relationships. *University of Kansas Science Bulletin* 38:1409-1438.

Sokal, R. R., and P. H. A. Sneath. 1963. *Principles of numerical taxonomy*. W. H. Freeman & Co. San Francisco, CA. 359 pp.

Swofford, D. L. 1993. *PAUP: Phylogenetic Analysis Using Parsimony* (version 3.1.1). Champaign, Illinois Natural History Survey.

Thompson, J. D., D. G. Higgins and T. J. Gibson. 1994. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* 22:4673-4680.

Wise, K. P. 1990. Baraminology: A young-earth creation biosystematic method. In Walsh, R. E. and C. L. Brooks (editors), *Proceedings of the second international conference on creationism*. Volume 2. Creation Science Fellowship. Pittsburgh, PA. pp. 345-358.

Wise, K. P. 1992. Practical baraminology. *Creation Ex Nihilo Technical Journal* 6(2):122-137.