Evolution, 41(6), 1987, pp. 1162–1176

LIKELIHOOD ANALYSIS OF MITOCHONDRIAL RESTRICTION-CLEAVAGE PATTERNS FOR THE HUMAN-CHIMPANZEE-GORILLA TRICHOTOMY

PETER E. SMOUSE
Departments of Human Genetics and Biology, University of Michigan, Ann Arbor, MI 48109

AND

WEN-HSIUNG LI
Center for Demographic and Population Genetics, University of Texas, Houston, TX 77030

Abstract.—In spite of considerable work by many evolutionary biologists, it has not been possible to separate convincingly the human (H) - chimpanzee (C) - gorilla (G) trichotomy into a pair of sequential dichotomies. There are three possible phylogenetic groupings [(HC)G, (HG)C, and (CG)H], and each has its proponents. The evidence remains ambiguous, and instead of choosing among the available phylogenies, it might be better to provide a statement of their relative probabilities. We develop a likelihood analysis of mtDNA restriction-pattern data that can be used to make such probability statements and illustrate it with data on humans, chimpanzees, gorillas, orangutans, and gibbons.

The results of our analyses suggest that the best fitting model is that of the (CG)H grouping, but with a very short time span between the first and second splits. For either the (HC)G or (HG)C hypothesis, a true trichotomy is the best model (no elapsed time between the two splits). Chi-square tests indicate no compelling resolution among the three models, however, and all three retain nontrivial posterior probabilities. We also compare each model with an alternative allowing for rate heterogeneity among lineages, but there is no convincing evidence for such heterogeneity. Our results suggest that, while it may eventually be possible to resolve the trichotomy into a pair of unambiguously ordered (but very close) dichotomies, it is possible that the ancestors of all three taxa (H, C, and G) were still conspecific subsequent to the second split, perhaps no more different than the "major races" of extant *Homo sapiens*.

Received September 29, 1986. Accepted June 15, 1987

In spite of an increasing number of attempts, it has not yet been possible to resolve the human (H) - chimpanzee (C) - gorilla (G) trichotomy into a sequential pair of dichotomies. Evidence exists that supports each of three possible phylogenetic groupings [(HC)G, (HG)C, and (CG)H]. Both the (HC)G and (CG)H hypotheses have numerous proponents. The anatomical analyses of Tuttle (1981), Stern and Susman (1981), and Fleagle et al. (1981), the morphometric analyses of Oxnard (1981) and Aiello (1981), and the fossil and phenetic analyses of Greenfield (1980) and Andrews (1982) all favor the (CG)H split. Treatments of molecular data by Benveniste and Todaro (1976), Wilson et al. (1977), Ferris et al. (1981*a*), Brown et al. (1982), and Kluge (1983), and restriction enzyme analysis of chromosomes by Bianchi et al. (1985) lead to the same choice. On the other hand, the (HC)G hypothesis is supported by morphological analysis (Washburn, 1982), protein sequence data (Goodman et al., 1982), G-banding analysis of chromosomes (Yunis and Prakash, 1982), and DNA-DNA hybridization studies (Sibley and Ahlquist, 1984), although Templeton (1985) has challenged the (HC)G interpretation of Sibley and Ahlquist (1984). Moreover, reanalyses of the mtDNA sequence data produced by Brown et al. (1982) led Nei et al. (1985) to favor the (HC)G hypothesis, in contradiction to the original interpretations of Brown et al. (1982). The (HG)C hypothesis is the least favored of the three, but the molecular data of Ueda et al. (1985) tend to favor that choice, and DNA sequence analysis (Hixson and Brown, 1986) shows that it is as parsimonious as the (HC)G hypothesis.

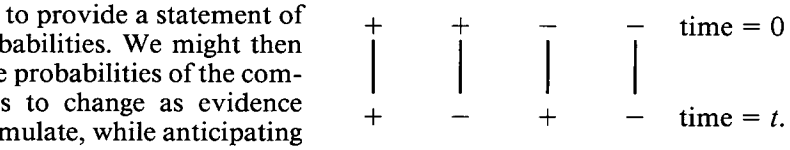
The problem is not only that different sorts of data are being used; different procedures used on the same data yield different answers. Such ambiguities are typical of difficult phylogenetic reconstructions; a pair of dichotomies, closely spaced in time, inevitably lead to weak inference. Although it may eventually be possible to resolve the trichotomy into a sequential pair of dichotomies on the weight of genetic evi-

dence, too few data are available to do so now. Rather than choosing one of the three hypothesized phylogenies as the "winner," it might be better to provide a statement of their relative probabilities. We might then expect the relative probabilities of the competing hypotheses to change as evidence continues to accumulate, while anticipating that all three candidate phylogenies would remain credible possibilities for the forseeable future.

Our concern here is with restriction-map data. The essential mathematical tools for a likelihood analysis are implicit in the papers of Nei and Tajima (1985) and Li (1986). With appropriate probability arguments, the likelihood treatment can be used with other sorts of data, following the same analytic philosophy developed here. Earlier attempts of this same general sort can be found in Kashyap and Subas (1974), Kaplan and Langley (1979), Felsenstein (1981, 1983), and DeBry and Slade (1985). We shall begin with a brief presentation of the probability functions that are particularly appropriate for restriction-site data and will then deploy the likelihood analysis. We shall describe the analysis for three species (human-chimpanzee-gorilla) in considerable detail, dealing first with a single restriction marker and then with the full constellation of markers. We shall then indicate the extension to four species and will use it to include a known "outgroup" (either orangutan or gibbon) in our analysis. We anticipate that the addition of an undisputed outgroup will accentuate the differences among the three probabilities, by providing information about the genetic state of the ancestral node of the trio. We shall use published data on restriction-site patterns for these five taxa (Ferris et al., 1981*a*) to illustrate the theory with some numerical examples and to gain some appreciation for the utility of restriction-site evidence.

Restriction Site Change Probabilities for Three Species

Restriction-Site-Change Probabilities.— Nei and Tajima (1985) and Li (1986) have derived the probabilities of the four types of events that we can encounter with respect to restriction-site changes in a single lineage. Let + denote a site recognized by a particular restriction enzyme and − a site not recognized by that same enzyme. The four possible patterns are



Patterns (+ | +) and (− | −) both represent conservation of state, while patterns (+ | −) and (− | +) represent changes of state. The probability arguments are symmetric in either time direction (forward or backward), so $Pr(+|-) = Pr(-|+)$, but $Pr(+|+) \neq Pr(-|-)$. What this means in practice is that the sharing of a − site by two taxa is not probabilistically equivalent to the sharing of a + state, a fact that has some important (generally unfortunate) consequences for all methods based on sheer counts of shared and differing character states.

To compute the probabilities of the above patterns, we shall use the following two-parameter scheme of nucleotide substitution (Kimura, 1981):

		To			
		A	T	C	G
From	A	γ	β	β	α
	T	β	γ	α	β
	C	β	α	γ	β
	G	α	β	β	γ

where α is the probability of a transitional change per unit time, β is the corresponding probability of a particular type of transversional change, and $\gamma = 1 - \lambda$, with $\lambda = \alpha + 2\beta$. Brown et al. (1982) have found that for higher primate mtDNA, $\alpha \approx 18\beta$. This two-parameter model has been used extensively by Li (1986) and Saitou and Nei (1986). If $\alpha = \beta$, then we can use the one-parameter model of Nei and Tajima (1985). A six-parameter version is also available (Kimura, 1981), but the philosopy of analysis does not depend on the number of parameters.

All of the theory to come is based on the assumption that the nucleotides are equifrequent. Indeed, the symmetry of the nucleotide-substitution matrix employed above leads ultimately to equal probabilities of the four nucleotides. The assumption

cannot be true in general, but it leads to tractable theory. One could presumably develop analogous theory for the case of a permanent deviation from equifrequent nucleotide frequencies, but the required mathematical complexity is daunting. Moreover, Tajima and Nei (1982) have shown that both the sign and magnitude of any bias in the parameter estimates will depend on the specific restriction enzyme in use but that the bias will tend to cancel out if many restriction enzymes are used. Since we are using several different restriction enzymes here, we will follow the usual practice of using the equifrequent theory.

Using the one- and two-parameter models, Nei and Li (1979), Aoki et al. (1981), and Li (1986) have studied the evolutionary change of nucleotides at a single nucleotide site. Under the two-parameter model, the probability that the nucleotide at time t will be the same as that at time 0 is given by

$$P(t) = \frac{[1 + e^{-4\beta t} + 2e^{-2(\alpha+\beta)t}]}{4}. \quad (1)$$

The probability of a transition (A ↔ G or T ↔ C) within t time units is

$$Q(t) = \frac{1 + e^{-4\beta t} - 2e^{-2(\alpha+\beta)t}}{4}. \quad (2)$$

The probability of a particular transversion (A ↔ T, A ↔ C, G ↔ T, or G ↔ A) is

$$S(t) = \frac{1 - e^{-4\beta t}}{4}. \quad (3)$$

The one-parameter model is obtained by substituting $\alpha = \beta$ in (1) and (2), whereas (3) requires no change.

Now, consider a restriction enzyme with a particular recognition sequence of r nucleotides. The probability a that a randomly drawn sequence of r nucleotides is a restriction site is given by (Nei and Li, 1979)

$$a = \left(\frac{1}{4}\right)^r. \quad (4)$$

The probability of encountering the pattern (+ | +) is

$$\Pr(+\,|\,+) = aP^r(t) \quad (5a)$$

(Li, 1986). Similarly, we have

$$\Pr(+\,|\,-) = a[1 - P^r(t)] = \Pr(-\,|\,+), \quad (5b)$$

and

$$\Pr(-\,|\,-) = 1 - a[2 - P^r(t)]. \quad (5c)$$

Note once again that $\Pr(+\,|\,+) \neq \Pr(-\,|\,-)$. Since (for any one restriction enzyme) there are many more nonrecognition than recognition sequences, preservation of a restriction site in the presence of nucleotide substitution is less likely than preservation of its absence. Such considerations lead to the observations that when $t \to 0$, $\Pr(-\,|\,-) \gg \Pr(+\,|\,+) \gg \Pr(+\,|\,-) = \Pr(-\,|\,+)$, whereas when $t \to \infty$, $\Pr(-\,|\,-) \gg \Pr(+\,|\,-) = \Pr(-\,|\,+) \gg \Pr(+\,|\,+)$, so that $\Pr(-\,|\,-)$ is generally much larger than $\Pr(+\,|\,+)$. Also of relevance is the fact that the (+ | +) pattern may be either more or less probable than the (+ | −) or (− | +) patterns, depending on the time depth involved.

Network Probabilities.—The probabilities in (1)–(3) can be used to specify the probability of any pattern of restriction-site states, given the time depth and the choice of phylogeny. Just to illustrate, consider the phylogeny and data set presented in Figure 1a. The probability of this outcome, given the phylogeny imposed, is the sum of the probabilities of the two unrooted networks presented in Figure 1b. Since the exact restriction-site status of the ultimate ancestral node does not enter into the argument, these two networks, denoted as "(1)" and "(2)" in Figure 1b can be redrawn as in the top portion of Figure 1c. We assume in the following that α and β are constant for all lineages and nucleotide positions, unless specified otherwise. Li (1986) shows that the probability of network "(2)" is most easily defined in terms of those of networks "(3)" and "(4)" in Figure 1c:

$$\Pr(1) = aP^{2r}(t_1)[1 - P^r(t_1 + 2t_2)] \quad (6a)$$

$$\Pr(2) = \Pr(3) - \Pr(4) = a[P^r(2t_1) - P^{2r}(t_1)] - a\psi[t_1, t_1, t_1 + 2t_2], \quad (6b)$$

where

$$\psi[t_1, t_1, t_1 + 2t_2] = \xi(t_1, t_1, t_1 + 2t_2) - P^{2r}(t_1)P^r(t_1 + 2t_2), \quad (7a)$$

with $\xi(t_1, t_1, t_1 + 2t_2)$ defined as in (Li, 1986):

$$\xi(t_1, t_1, t_1 + 2t_2) = [Q^2(t_1)Q(t_1 + 2t_2) + 2S^2(t_1)S(t_1 + 2t_2) + P^2(t_1)P(t_1 + 2t_2)]^r. \quad (7b)$$

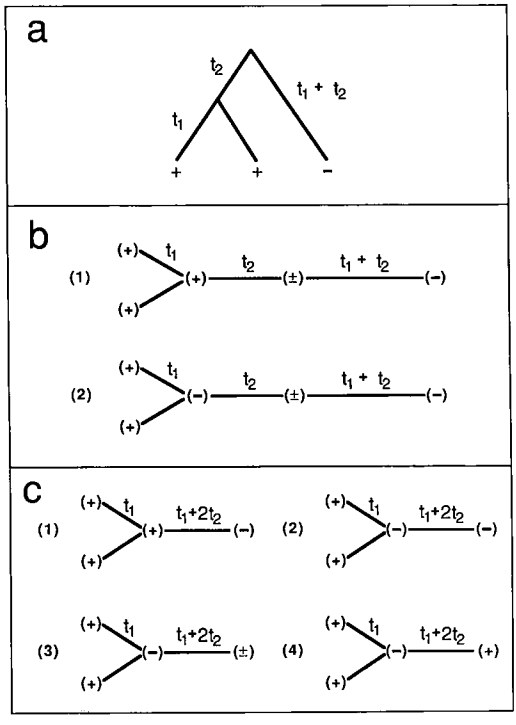


FIG. 1. Example phylogeny with restriction pattern (+ + −) superimposed; a) the rooted phylogeny, b) the two constituent unrooted phylogenies, c) the networks needed to compute the probability of the restriction pattern, given the phylogeny, with ancestral nodes indicated. Restriction-site status is indicated by (+) for recognition and (−) for nonrecognition.

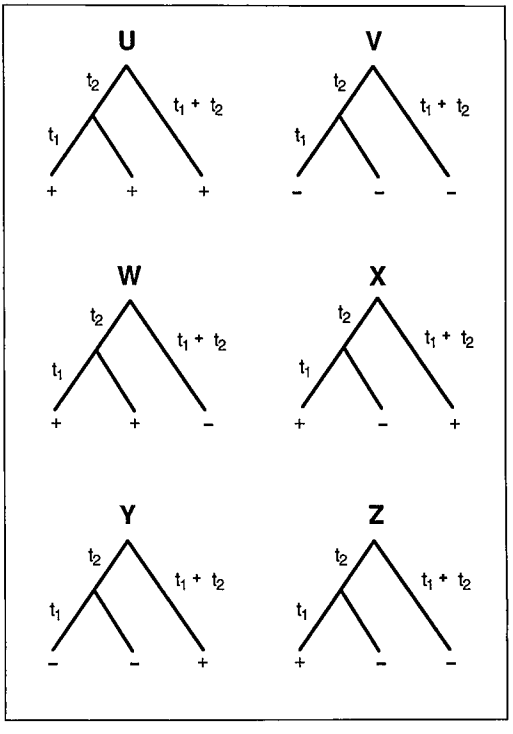


FIG. 2. The six restriction patterns for a phylogeny of three species, with their probabilities (*U*–Z); restriction-site status is indicated by (+) for recognition and (−) for nonrecognition.

Given α, β, t_1, and t_2, the probability of the phylogeny in Figure 1a is thus

$$W(t_1, t_2) = \Pr(\text{data} \mid \text{phylogeny}) = a[P^r(2t_1) - \xi(t_1, t_1, t_1 + 2t_2)]. \quad (8)$$

All six possible realizations are presented in Figure 2. Using procedures similar to the above, we show that:

$$U(t_1, t_2) = a\xi(t_1, t_1, t_1 + 2t_2) \quad (9)$$

$$V(t_1, t_2) = 1 - 3a + aP^r(2t_1) + 2aP^r(2t_1 + 2t_2) - a\xi(t_1, t_1, t_1 + 2t_2) \quad (10)$$

$$X(t_1, t_2) = a[P^r(2t_2 + 2t_1) - \xi(t_1, t_1, t_1 + 2t_2)] \quad (11)$$

$$Y(t_1, t_2) = a[1 - 2P^r(2t_2 + 2t_1) + \xi(t_1, t_1, t_1 + 2t_2)] \quad (12)$$

$$Z(t_1, t_2) = a[1 - P^r(2t_1) - P^r(2t_2 + 2t_1) + \xi(t_1, t_1, t_1 + 2t_2)]. \quad (13)$$

These six functions (*U*, *V*, *W*, *X*, *Y*, and *Z*) are sufficient to evaluate the phylogenies for a trio of species. The restriction patterns (HCG) = (+ + −), (+ − +), (− + +), (− − +), (− + −), and (+ − −) alone provide taxonomic resolution, but the (+ + +) and (− − −) sites provide useful information on the time depth. Note that (+ − +) and (− + +) are equally likely, as are (− + −) and (+ − −).

Likelihood Analysis

Phylogenetic Probabilities.—We present the three possible dichotomous phylogenies for the three species in question in Figure 3a. There are also eight classes of possible restriction-site outcomes, as indicated in Table 1. Given restriction enzymes with the same *r* value (the number of nucleotides in the enzyme recognition sequence), we can compute the probability of the total data set for each of the phylogenies in Figure 3 from the numbers (N_i) of sites in each of the eight restriction classes. Neglecting a combina-

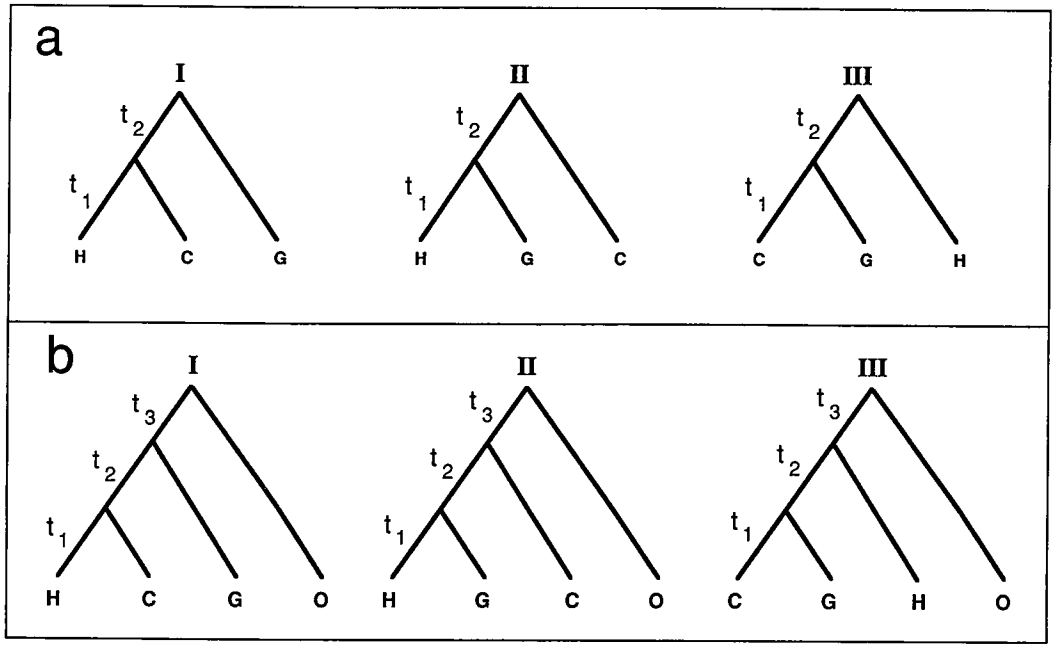


FIG. 3. The three possible phylogenies for human (H), chimpanzee (C), and gorilla (G): a) ignoring a known outgroup, b) including a known outgroup (O).

torial constant that will cancel, these probabilities are:

$$\Pr(\text{data}\,|\,\text{I}) = U^{N_1}V^{N_2}W^{N_3}X^{(N_4+N_5)} \cdot Y^{N_6}Z^{(N_7+N_8)} \quad (14a)$$

$$\Pr(\text{data}\,|\,\text{II}) = U^{N_1}V^{N_2}W^{N_4}X^{(N_3+N_5)} \cdot Y^{N_7}Z^{(N_6+N_8)} \quad (14b)$$

$$\Pr(\text{data}\,|\,\text{III}) = U^{N_1}V^{N_2}W^{N_5}X^{(N_3+N_4)} \cdot Y^{N_8}Z^{(N_6+N_7)}. \quad (14c)$$

Given data from enzymes with different r values, we set (14a)–(14c) up for each r value separately and then multiply the respective sets together.

These three equations presuppose their respective phylogenies. We need to compute the posterior probability of each of the phylogenies, given the restriction-site data:

$$\Pr(\text{I}\,|\,\text{data}) = \frac{\Pr(\text{I})\Pr(\text{data}\,|\,\text{I})}{\Pr(\text{data})}, \quad (15a)$$

$$\Pr(\text{II}\,|\,\text{data}) = \frac{\Pr(\text{II})\Pr(\text{data}\,|\,\text{II})}{\Pr(\text{data})}, \quad (15b)$$

$$\text{PR}(\text{III}\,|\,\text{data}) = \frac{\Pr(\text{III})\Pr(\text{data}\,|\,\text{III})}{\Pr(\text{data})}, \quad (15c)$$

where $\Pr(\text{data}) = \sum_i \Pr(i)\Pr(\text{data}\,|\,i)$: i = I, II, III.

The Prior.—The prior probabilities [$\Pr(i)$: i = I, II, III] can be specified in advance, on the basis of external information. In the context of the current problem, it is reasonable to assume that $\Pr(i) = \frac{1}{3}$: i = I, II, III. There is nothing much to choose among the three possibilities; the trichotomy has stubbornly resisted resolution to this point. Templeton (1983*b*) claims to have resolved it, using data from Ferris et al. (1981*a*), but since we will here analyze data from the same source, we will hold that claim in abeyance, pending the outcome. The choice of a neutral prior cancels the $\Pr(i)$ values from (15a)–(15c), and reduces the problem to a consideration of (14a)–(14c).

Maximizing the Likelihood.—We can view (14a)–(14c) as the probabilities of the data given the model phylogenies, or, by turning the argument around, we can view them as the likelihoods of the model phylogenies, given the data. Treating them as likelihoods, where the data are constant but the probabilities are variables, we can choose values of t_1 and t_2 that optimize (maximum) the likelihood values. Since (14a)–(14c) are products of sums of exponentials, the likelihood scoring algorithms based on explicit

TABLE 1. Numbers of recognition sites in each of eight different pattern classes, separated by r value (in columns headed "recognition length") and the probabilities of obtaining such patterns (in columns headed "phylogeny") under the three possible phylogenies in Figure 3a. Restriction-site status is indicated by + for recognition and − for nonrecognition in human (H), chimpanzee (C), and gorilla (G). The probabilities of the patterns observed are U, V, W, X, Y, and Z, defined as in text Equations (8)–(12).

Restriction-site class	Species			Recognition length (r)			Phylogeny		
	H	C	G	6	16/3	4	I	II	III
1	+	+	+	13	2	4	U	U	U
2	−	−	−	24	6	9	V	V	V
3	+	+	−	3	0	2	W	X	X
4	+	−	+	5	0	0	X	W	X
5	−	+	+	7	1	0	X	X	W
6	−	−	+	6	3	2	Y	Z	Z
7	−	+	−	9	4	0	Z	Y	Z
8	+	−	−	7	8	0	Z	Z	Y

first and second partial derivatives are not particularly convenient for evaluation. Our only real recourse is numeric solution. Fortunately, the data show some regularities, and there are some simplifying shortcuts. The procedure we present below is patterned after the "lod score" procedure developed for genetic linkage analysis by Morton (1955).

First, ignore the (neutral) priors and the shared (common) denominators in (15a)–(15c). That leaves the traditional likelihood functions, normally written in more convenient logarithmic form as

$$\begin{aligned}\log L(\text{I}|\text{data}) = {} & N_1\log(U) \\ & + N_2\log(V) + N_3\log(W) \\ & +(N_4 + N_5)\log(X) \\ & + N_6\log(Y) \\ & + (N_7 + N_8)\log(Z) \quad (16a)\end{aligned}$$

$$\begin{aligned}\log L(\text{II}|\text{data}) = {} & N_1\log(U) \\ & + N_2\log(V) + N_4\log(W) \\ & + (N_3 + N_5)\log(X) \\ & + N_7\log(Y) \\ & + (N_6 + N_8)\log(Z) \quad (16b)\end{aligned}$$

$$\begin{aligned}\log L(\text{III}|\text{data}) = {} & N_1\log(U) \\ & + N_2\log(v) + N_5\log(W) \\ & + (N_3 + N_4)\log(X) \\ & + N_8\log(Y) \\ & + (N_6 + N_7)\log(Z). \quad (16c)\end{aligned}$$

We construct a separate set of two-dimensional log-likelihood tables for each r value, one table each for $\log(U)$, $\log(V)$, $\log(W)$, $\log(X)$, $\log(Y)$, and $\log(Z)$, over the permissible ranges of t_1 and t_2. (The need for separate tables for different r values is explained below.) The same set of tables can be used for all three hypotheses. We then draw matching (t_1, t_2) elements from these various tables, and add the log-likelihoods up in the requisite numbers to obtain a trio of two-dimensional tables for the values of (16a)–(16c). We then choose the optimal (t_1, t_2) elements from the tables for each hypothesis. By using the parameter-optimized forms of (16a)–(16c), we can compare the three hypotheses in (15a)–(15c), each shown to best advantage.

Mitochondrial Restriction-Site Data.—We present in Table 1 the numbers of restriction sites in each of the eight pattern classes, extracted from Ferris et al. (1981*a*). The data presented here are a subset of those in the original paper, and some comments are in order before proceeding with the analysis. We include all those sites that are detectable in at least one of the five species. The procedure presented here thus allows usage of all the available data. By contrast, the analyses of both Templeton (1983*b*) and Li (1986) employ only those sites for which two of the five taxa differ from the rest. We shall have more to say about usage of data in the Discussion.

Hinc II recognizes the sequence GTPyPuAC, and overlaps in recognition sequence with *Hpa* I (GTTAAC) and *Sal* I (GTCGAC). *Ava* I recognizes the sequence CPyCGPuG and overlaps in recognition sequence with *Xho* I (CTCGAG) and *Sma* I (CCCGGG). We have ignored all the *Hpa* I, *Sal* I, *Xho* I, and *Sma* I sites, using instead the *Hinc* II and *Ava* I sites, so as to avoid

TABLE 2. Log-likelihood analysis of the data in Table 1, under the one- and two-parameter models, with maximum likelihood estimates of t_1 and t_2 and posterior probabilities for each of the candidate phylogenies.

Hypothesized phylogeny	Maximum-likelihood values			Posterior probability
	t_1	t_2	Log$L$	
One-parameter model ($\alpha = 1/3 = \beta$):				
I	0.0906	0.0000	−684.710	0.256
II	0.0906	0.0000	−684.710	0.256
III	0.0785	0.0185	−684.068	0.488
Two-parameter model ($\alpha = 0.90$, $\beta = 0.05$):				
I	0.0942	0.0000	−685.097	0.257
II	0.0942	0.0000	−685.097	0.257
III	0.0809	0.0206	−684.461	0.486

any double counting of information. We note that, while *Ava* I and *Hinc* II are six-base cutters, the degeneracy of their recognition sequences alters the r value. Nei and Tajima (1983) use an approximate r value of 16/3 for such enzymes, a practice we shall follow here. The total nucleotide information retained is larger if we maintain the *Hinc* II and *Ava* I sites, even allowing for the degeneracy of their recognition sequences. Finally, *Fnu*D II is a four-base cutter, so we set $r = 4$ for the quartet of *Fnu*D II sites.

The Results.—As pointed out above, each site can be placed in one of eight restriction classes, if we consider just human, chimpanzee, and gorilla. The tallies are presented in Table 1, separated into sets corresponding to six-base cutters ($r = 6$), degenerate six-base cutters ($r = 16/3$), and four-base cutters ($r = 4$). We can anticipate the outcome of the analysis from Table 1. The preferred phylogeny is (III) of Ferris et al. (1981*a*), also favored by the analyses of Templeton (1983*b*), Nei et al. (1985), and Li (1986), all of whom used the same basic data set. Comparing classes 3, 4, and 5, we discover that class 5 is more common than either class 4 or class 3. Class 3 also has representatives from $r = 4$, but the affinity implied by a shared site decreases with decreasing r value. On the basis of pattern classes 3, 4, and 5 alone, we should anticipate that

$$L(\text{III}|\text{data}) > L(\text{II}|\text{data}) > L(\text{I}|\text{data}) \quad (17)$$

and that the ratio of t_1 to t_2 will be larger (more nearly a true trichotomy) as we move to the less likely phylogenies. A comparison of the numbers in pattern classes 6, 7, and 8 indicates that we should expect to see the same general result.

The final results of the three-species analyses are presented in Table 2 for both the one- and two-parameter models and for each of the three phylogenies in Figure 3a. Since time parameters appear only in product form (αt and βt), we arbitrarily set $\alpha = \beta = 1/3$ for the one-parameter model and $\alpha = 0.90$ and $\beta = 0.05$ for the two-parameter model (because $\alpha \approx 18\beta$). That reduces $\lambda = \alpha + 2\beta$ to unity in both cases. The order of model precedence is much as predicted in (17), and the posterior odds of the three models are 1:1:2 (an improvement over the prior odds of 1:1:1), but the differences in log-likelihoods are disappointingly small. Moreover, the most likely value of t_2 is zero for models I and II, and the value for model III is only 0.0206. In addition, for the short evolutionary time periods of interest, the process is quite adequately described by the one-parameter model, as discussed earlier by Nei et al. (1985) and Li (1986). The two-parameter model adds nothing in the way of resolution and changes the total time depth only slightly. Over the time depths under discussion ($\lambda t < 0.11$), the one-parameter model is adequate. With deeper time depth, such as might be necessary with a fourth species (see below), the probability of parallel losses or gains of restriction sites may become substantially higher for the two-parameter model, so that the one- and two-parameter models may not be so nearly equivalent.

A Trichotomous Model.—The failure of these data to yield strong evidence for a pair of dichotomies within the human-chimpanzee-gorilla trio raises the question of how well a trichotomous model of radiation would work. The answer is already implicit in Table 2, because we need merely set $t_2 = 0$ in any of the preceding models, a ploy that reduces all three models to trichotomous form (Fig. 4a). The analytical results are presented in Table 3, where the trichotomous model (0) is contrasted with models I, II, and III. There is very little improvement in

the log-likelihood to be obtained by allowing $t_2 > 0$, and the trichotomous model works about as well as any of the others.

Nonconstant Rates of Evolution.—We have explicitly assumed throughout this paper that the rate of nucleotide substitution ($\lambda = \alpha + 2\beta$) is a constant in all lineages. This assumption can never be more than approximately true in practice. Goodman (1985) has argued that the rate of evolution in the final lineage leading to humans (H) has been lower than in other lineages. If that were true, our treatment could lead to erroneous results. How are we to assess the size of the potential problem? Suppose we let λ vary among lineages, while holding the ratio (α:β) constant. That allows us to alter the phylogeny in Figure 4a to that shown in Figure 4b, where the time depths (here called τ's) of the trichotomous radiation are unequal (somewhat exaggerated here to improve visualization). Time depth is always measured in terms of the product $\lambda\tau$, and we have arbitrarily set $\lambda = 1$ in all of the above. Varying τ is the equivalent of varying λ. If we let τ_1, τ_2, and τ_3 take the values that optimize the likelihood, we obtain the results for model IV in Table 3. The values of the $\hat{\tau}$'s are all quite similar to each other, to the $\hat{\tau}$ value of the strict trichotomy, and to the values of t_1 in Table 2. Moreover, the log-likelihoods are only slightly different. There is no compelling evidence for lineage heterogeneity of evolutionary rates in these data. Note, however, that without some outside evidence concerning either the state of individual restriction sites at the nodes of the phylogeny or the respective time depths, any inference on this point is weak. We assume that there is rate homogeneity for all that follows, barring any convincing evidence to the contrary.

Hypothesis Testing.—We are led to a consideration of hypothesis testing, using differences in log-likelihood values. Twice the difference in log-likelihood values of two models is asymptotically distributed as χ^2, provided that one of the models is a proper submodel (special case) of the other. None of the models I, II, or III is a submodel of the other, so we cannot test the differences among them in the usual χ^2 fashion. On the other hand, all three are special cases of

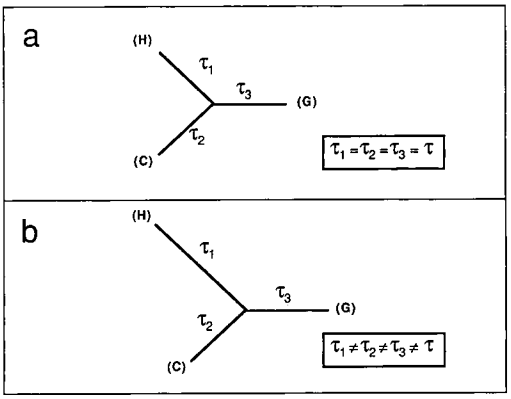


FIG. 4. Trichotomous phylogenetic models for human (H), chimpanzee (C), and gorilla (G): a) rate homogeneity, all τ's equal, b) rate heterogeneity for the same three taxa.

model IV, the rate-heterogeneous trichotomy. Model I sets $\tau_1 = \tau_2 = t_1$ and $\tau_3 = t_1 + 2t_2$; model II sets $\tau_1 = \tau_3 = t_1$ and $\tau_2 = t_1 + 2t_2$; model III sets $\tau_2 = \tau_3 = t_1$ and $\tau_1 = t_1 + 2t_2$. The time estimates (t_i's) for models I, II, and III are simply extracted from Table 1 and translated into those of Table 3. To evaluate model IV, we use (8)–(13) in the altered form:

$$U(\tau_1, \tau_2, \tau_3) = a\xi(\tau_1, \tau_2, \tau_3) \quad (18)$$

$$V(\tau_1, \tau_2, \tau_3) = 1 - 3a + aP^r(\tau_1 + \tau_2) + aP^r(\tau_1 + \tau_3) + aP^r(\tau_2 + \tau_3) - a\xi(\tau_1, \tau_2, \tau_3) \quad (19)$$

$$W(\tau_1, \tau_2, \tau_3) = a[P^r(\tau_1 + \tau_2) - \xi(\tau_1, \tau_2, \tau_3)] \quad (20)$$

$$X_1(\tau_1, \tau_2, \tau_3) = a[P^r(\tau_1 + \tau_3) - \xi(\tau_1, \tau_2, \tau_3)] \quad (21a)$$

$$X_2(\tau_1, \tau_2, \tau_3) = a[P^r(\tau_2 + \tau_3) - \xi(\tau_1, \tau_2, \tau_3)] \quad (21b)$$

$$Y(\tau_1, \tau_2, \tau_3) = a[1 - P^r(\tau_1 + \tau_3) - P^r(\tau_2 + \tau_3) + \xi(\tau_1, \tau_2, \tau_3)] \quad (22)$$

$$Z_1(\tau_1, \tau_2, \tau_3) = a[1 - P^r(\tau_1 + \tau_2) - P^r(\tau_1 + \tau_3) + \xi(\tau_1, \tau_2, \tau_3)] \quad (23a)$$

$$Z_2(\tau_1, \tau_2, \tau_3) = a[1 - P^r(\tau_1 + \tau_2) - P^r(\tau_2 + \tau_3) + \xi(\tau_1, \tau_2, \tau_3)]. \quad (23b)$$

The two forms of X and those of Z are used to distinguish the case where human is (+) and chimpanzee is (−) from the case where the human is (−) and the chimpanzee is (+) in Figure 2, a distinction that was not necessary as long as $\tau_1 = \tau_2$ in the rate-homogeneous models. A contrast of the log-likelihood of any one of models I, II, or III with that of model IV is χ^2 distributed with one degree of freedom. An examination of Table 3 will show that all of these test criteria are less than unity, providing no evidence for $\tau_1 \neq \tau_2$, $\tau_1 \neq \tau_3$, or $\tau_2 \neq \tau_3$, taken pairwise. Note further that the trichotomous model 0 is a proper subset of each of models I, II, and III, and hence also of model IV, obtained by setting $t_2 = 0$. Again, the appropriate differences in log-likelihoods are χ^2 distributed with one degree of freedom. It is also clear from Table 3 that there is no strong evidence for $t_2 > 0$. Recall from Table 2 that models I and II are optimized with $t_2 = 0$, which reduces them explicitly to model 0, while model III (with $t_2 > 0$) is minimally better than model 0. A rate-homogeneous trichotomy, the simplest model we could specify, is compatible with the data.

TABLE 3. A comparison of time estimates and log-likelihood values for five models of the phyletic radiation of human (H), chimpanzee (C), and gorilla (G). For model I, $\tau_1 = \tau_2 = t_1$, and $\tau_3 = t_1 + 2t_2$; for model II, $\tau_1 = \tau_3 = t_1$, and $\tau_2 = t_1 + 2t_2$. The best estimates of t_2 for both models are 0 (Table 2), reducing each to model 0, with $\tau_1 = \tau_2 = \tau_3 = t_1$. For model III, $\tau_2 = \tau_3 = t_1$, and $\tau_1 = t_1 + 2t_2$. For model IV, the τ's can take any positive values.

Hypothesized phylogeny	Maximum-likelihood values τ_1 (H)	τ_2 (C)	τ_3 (G)	log$L$
One-parameter model ($\alpha = 1/3 = \beta$):				
0	0.0906	0.0906	0.0906	−684.710
I	0.0906	0.0906	0.0906	−684.710
II	0.0906	0.0906	0.0906	−684.710
III	0.1155	0.0785	0.0785	−684.068
IV	0.1155	0.0811	0.0760	−684.059
Two-parameter model ($\alpha = 0.90$, $\beta = 0.05$):				
0	0.0942	0.0942	0.0942	−685.097
I	0.0942	0.0942	0.0942	−685.097
II	0.0942	0.0942	0.0942	−685.097
III	0.1221	0.0809	0.0809	−684.461
IV	0.1219	0.0839	0.0778	−684.452

Adding an Outgroup

Motivation.—The point of adding an outgroup is to improve our information on the ancestral node of the trio, treated as ambiguous in all of the preceding. Although we cannot categorically impose a particular (+ or −) state on the ancestral node of the trio for any given restriction site, as would be the standard practice with most parsimony procedures, the availability of an outgroup will tend to make one or the other alternative more probable. Although we simply move the ambiguous node one step deeper in the phylogeny, the extra resolution on the trio should help. There are 15 possible rooted topologies with four species. The probability functions can be constructed and evaluated for all fifteen (Li, 1986), but only a subset of these functions is necessary for our purpose here. It is well established that the trio radiated subsequent to divergence from the outgroup. We have three plausible phylogenies, specifically those shown in Figure 3b. This amounts to the imposition of a strong prior on the Pr(i); we set all Pr(i) = 0 except those for the three phylogenies in Figure 3b, set equal to ⅓ each.

Restriction Patterns.—With either outgroup (orangutan or gibbon), we can define a set of 16 restriction-cleavage-pattern classes, and we present the numbers of sites in each pattern class in Table 4. Using the same sort of network addition and subtraction strategies deployed earlier for a trio of species, Li (1986) has derived the probabilities of these 16 patterns. The mathematical labor involved is too large for useful exposition here, and the resulting functions, denoted as $A(t_1, t_2, t_3)$ through $M(t_1, t_2, t_3)$, are simply presented in the Appendix for easy access. The reader interested in the derivations is referred to Li (1986). From this point forward, the analysis is the same as that for the three species case, albeit more laborious.

Results.—The results of the four-species likelihood analyses are presented in Table 5. It is clear that adding an outgroup does change the relative probabilities of the candidate phylogenies, particularly when the orangutan is used as that outgroup. The order of hypotheses is still that indicated in Table 2, and it does not matter a great deal which outgroup is used, although the or-

TABLE 4. Numbers of DNA recognition sites exhibiting the 12 different restriction classes for a quartet of species, separated by r value (in columns headed "recognition length"), and the probabilities of obtaining such patterns (in columns headed "phylogeny") under the three possible phylogenies; a) using orangutan (O) as the outgroup, b) using gibbon (B) as the outgroup. The probabilities, A–M are all defined in the Appendix.

a. Restriction-site class	Species				Recognition length			Phylogeny		
	H	C	G	O	6	16/3	4	I	II	III
1	+	+	+	+	9	1	4	*A*	*A*	*A*
2	+	+	+	−	4	1	0	*B*	*B*	*B*
3	−	−	−	+	10	4	6	*C*	*C*	*C*
4	−	−	−	−	14	2	3	*D*	*D*	*D*
5	+	+	−	+	2	0	0	*E*	*F*	*F*
6	+	+	−	−	1	0	2	*G*	*H*	*H*
7	+	−	+	+	3	0	0	*F*	*E*	*F*
8	+	−	+	−	2	0	0	*H*	*G*	*H*
9	−	+	+	+	0	0	0	*F*	*F*	*E*
10	−	+	+	−	7	1	0	*H*	*H*	*G*
11	−	−	+	+	0	0	0	*I*	*J*	*J*
12	−	−	+	−	6	3	2	*K*	*M*	*M*
13	−	+	−	+	0	1	0	*J*	*I*	*J*
14	−	+	−	−	9	3	0	*M*	*K*	*M*
15	+	−	−	+	0	2	0	*J*	*J*	*I*
16	+	−	−	−	7	6	0	*M*	*M*	*K*
b. Restriction-site class	Species				Recognition length			Phylogeny		
	H	C	G	B	6	16/3	4	I	II	III
1	+	+	+	+	7	1	4	*A*	*A*	*A*
2	+	+	+	−	6	1	0	*B*	*B*	*B*
3	−	−	−	+	15	3	3	*C*	*C*	*C*
4	−	−	−	−	9	3	6	*D*	*D*	*D*
5	+	+	−	+	0	0	0	*E*	*F*	*F*
6	+	+	−	−	3	0	2	*G*	*H*	*H*
7	+	−	+	+	4	0	0	*F*	*E*	*F*
8	+	−	+	−	1	0	0	*H*	*G*	*H*
9	−	+	+	+	2	1	0	*F*	*F*	*E*
10	−	+	+	−	5	0	0	*H*	*H*	*G*
11	−	−	+	+	1	0	0	*I*	*J*	*J*
12	−	−	+	−	5	3	2	*K*	*M*	*M*
13	−	+	−	+	2	0	0	*J*	*I*	*J*
14	−	+	−	−	7	4	0	*M*	*K*	*M*
15	+	−	−	+	1	2	0	*J*	*J*	*I*
16	+	−	−	−	6	6	0	*M*	*M*	*K*

angutan does provide slightly better discrimination among hypotheses. It is also clear from Table 5 that while t_3, the time between separation of the outgroup and subsequent radiation of the trio, is quite large, t_2 is still negligible under models I and II. With the greater total time depth of the four-species phylogeny, the one- and two-parameter models are more divergent. The central conclusion does not change, however; we are still not quite able to choose compellingly among the three competing hypotheses. It is also important to note that the estimated value of t_3 is larger when the orangutan is used as the outlier than when the gibbon is used, contrary to the conventional wisdom that the gibbon divergence antedates that of the orangutan. Our interpretation of this result is that the greater time depth of the gibbon divergence provides more opportunity for evolutionary reversals, (+ → − → +) or (− → + → −), or for convergent evolution, (+ → −) or (− → +) in two independent lineages, leading to greater phylogenetic ambiguity. The more divergent outgroup conveys less information about the genetic state of the ancestral node of the trio. While an unambiguous outgroup is useful, a closely related outgroup is best.

TABLE 5. Log-likelihood analysis of the data in Table 4, under the one- and two-parameter models, with maximum-likelihood estimates of t_1, t_2, and t_3 and posterior probabilities for each of the three candidate phylogenies.

Hypothesized phylogeny	Maximum-likelihood values				Posterior probability
	t_1	t_2	t_3	log$L$	
One-parameter ($\alpha = 1/3 = \beta$); orangutan as outlier:					
I	0.0922	0.0000	0.0552	−878.194	0.115
II	0.0922	0.0000	0.0552	−878.194	0.115
III	0.0736	0.0304	0.0425	−876.298	0.770
Two-parameter ($\alpha = 0.90$, $\beta = 0.05$); orangutan as outlier:					
I	0.0948	0.0000	0.0625	−879.763	0.147
II	0.0948	0.0000	0.0625	−879.763	0.147
III	0.0757	0.0313	0.0490	−878.192	0.706
One-parameter ($\alpha = 1/3 = \beta$); gibbon as outlier:					
I	0.0926	0.0005	0.0459	−902.673	0.217
II	0.0929	0.0000	0.0461	−902.674	0.217
III	0.0784	0.0223	0.0385	−901.714	0.566
Two-parameter ($\alpha = 0.90$, $\beta = 0.05$); gibbon as outlier:					
I	0.0951	0.0000	0.0537	−903.614	0.233
II	0.0951	0.0000	0.0537	−903.614	0.233
III	0.0800	0.0235	0.0454	−902.782	0.534

DISCUSSION

Choice of Data.—We have tallied in Table 1 all eight permutations of restriction-site status for the three species and have used the observed numbers of these restriction classes in the three-species analyses. Similarly, we have listed all 16 permutations of restriction-site status for a quartet of species in Table 4 and have used the observed numbers in the four-species analyses. The use of (− − −) sites for three species and of (− − − −) sites for a quartet of species requires some comment.

Consider first the three-species case and the recognition sequence for a particular restriction enzyme. The vast majority of the mitochondrial nucleotide sequence is not recognized by this enzyme in any of the species examined, and the number of (− − −) sites is thus very large, though not directly observable. Because the reading frame for recognition is arbitrary, there are as many potential sites as there are nucleotides in the mtDNA (on the order of 16,000). We could simply assign a number to the (− − −) class which is the total nucleotide length of the mtDNA, minus the sum of the other seven restriction-class sizes. Because the reading frames of the "sites" overlap, however, the occurrence of a nonrecognition site in one frame is not independent of the occurrence of a nonrecognition site in the overlapping frames, contrary to assumption of the model. Moreover, since we are using many different restriction enzymes, each with a different recognition sequence, the nonrecognition sites for one enzyme include the recognition sites for all of the others. Recognition and nonrecognition sites for one enzyme are not independent of those for another, again contrary to model assumption. In fact, the dependencies are exceedingly complex if we insist on complete enumeration. The same problems emerge in trying to enumerate the (− − − −) sites for the four-species analyses.

That being the case, where have we obtained the numbers for (− − −) in Table 1 and those for (− − − −) in Table 4? We have defined a "site" as a nucleotide sequence that is recognized (+) by an enzyme in *at least* one of the five species (human, chimpanzee, gorilla, orangutan, and gibbon). The (− − −) sites of Table 1 are recognized as (+) for either the orangutan or the gibbon in Table 4. All of the (− − − −) sites in Table 4a (orangutan as the outgroup) are (− − − +) in Table 4b (gibbon as outgroup), and vice versa. Thus, we tally sites that are (HCGOB) = (− − − − +) and (− − − + −), but do not include the unobservable sites that are (− − − − −). There are undoubtedly

many such sites in the mtDNA genome, but we cannot enumerate them.

What are the practical consequences of treating the data in this fashion? First, we note that ignoring the (−−−−−) sites cannot directly affect the choice of models, since whatever the number of such sites, the probability of a (−−−) site or a (−−−−) site is the same under all phylogenetic models. Second, however, we note that there is an indirect effect on the choice of models, since the numbers of (−−−) sites or (−−−−) sites, in comparison with the numbers of other sites, determines the best (most likely) choice of t_1, t_2, and t_3; since the optimal choice of time values varies with the model, a change in the number of such sites can tip the balance of evidence in favor of one hypothesis or another. Third, this effect is too subtle to have any appreciable impact for the present study.

This last point has been established in two ways. Consider again the three-species case. It is possible to ignore the (−−−) sites in Table 1, computing a conditional likelihood of the three models, given that only seven classes of outcome are observable. That treatment yields estimated t_1 values on the order of ~0.05 in Table 2, instead of the listed values of ~0.09. The estimated values for t_2 changed correspondingly. The resulting posterior probability values and the χ^2 tests changed negligibly. The second treatment used large numbers of (−−−) sites instead of the numbers listed in Table 1. The larger the numbers used, the smaller were the corresponding estimates of the t values. Again, the posterior probabilities and χ^2 tests were not sensitive to the choice. Analogous treatments of the four-species models had the same consequences, major reductions in all of the time parameters, but only subtle changes in the posterior probabilities or χ^2 tests.

Given the insensitivity of the outcome to the precise treatment of "negative" sites, we have not attempted to adjust our results for the unobservable. We should caution, however, that the estimated t values (really λt values), should be viewed as relative to each other; they should not be interpreted as absolute measures of time without calibration against an outside (nongenetic) reference standard.

The Inadvisability of Choosing.—The real value of the likelihood approach and its principle advantage over competing techniques is that instead of merely choosing a "winner" among the candidate phylogenies, we can assess the genetic evidence for each model. Model III [(CG)H] is a posteriori the most probable of the candidates presented, in agreement with earlier analyses of the Ferris et al. (1981*a*) data (Templeton, 1983*b*; Nei et al., 1985; Li, 1986), but there is not much to choose among the three dichotomous models, and we cannot really exclude a trichotomy. Extensive simulations by Li (1986) and Saitou and Nei (1986) show that whenever t_2 is a small fraction of the total time depth, as it is here, the statistical power of likelihood analysis is low and that we will need very large numbers of restriction sites to resolve a pair of sequential dichotomies convincingly.

It is possible that an analogous treatment of mitochondrial nucleotide sequences will yield better resolution, although the sequence data currently available (Anderson et al., 1981; Brown et al., 1982; Hixson and Brown, 1986) do not inspire much confidence on this point. Additional comparative sequence data on nuclear genes are becoming available (e.g., Slightom et al., 1985; Koop et al., 1986) and may hold some promise for increasing resolution. We suspect, however, that our inability to settle the issue with the great mass of data (molecular and otherwise) already available is an indication that the radiation of humans, chimpanzees, and gorillas was very nearly trichotomous.

A reexamination of the data in Table 1 indicates the nature of the problem. Restriction classes 3 [(++)−] and 6 [(−−)+], represented by a total of 16 sites, are most "consistent" with phylogeny I [(HC)G], while classes 4 [(+−)+], 5 [(−+)+], 7 [(−+)−] and 8 [(+−)−], represented by 41 sites, are less consistent, though by no means impossible, given mutational substitutions over a relatively short evolutionary time period. Unfortunately, phylogenies II and III do not fare much better, with only 18/57 and 23/57 "consistent" sites, respectively. Either there has been a great deal of convergent evolution or these restriction patterns represent polymorphisms within the

species ancestral to all three taxa that have become independently fixed in the separate lines of descent only after phyletic radiation. It is important to reiterate here the stricture given by Nei and Tajima (1983) that a mitochondrial phylogeny is not the same as and should not be confused with a species phylogeny. As a matter of fact, there is a certain amount of mtDNA restriction-pattern polymorphism within each of these species (Brown and Goodman, 1979; Brown, 1980; Ferris et al., 1981*b*; Cann et al., 1984). Comparison of individuals other than the three presented in Tables 1 and 4 might well have tipped the scales in favor of models I or II. The sampling noise outweighs the evolutionary signal in this case.

The time estimates (λt or $\lambda\tau$) for all of the three-species phylogenies are less than 0.11, indicating that there has been insufficient time for very much convergent evolution, and it is possible that we are examining the lingering remnants of ancestral conspecific polymorphism (Templeton, 1983*a*; see also Avise and Lansman, 1983; Avise, 1986). The ancestors of all three taxa (H, C, and G) may still have been conspecific subsequent to the second split, perhaps no more different than the "major races" of extant *Homo sapiens*.

Acknowledgments

The authors thank W. Brown, R. DeBry, S. Diehl, T. Dowling, J. Felsenstein, S. Frank, B. Milligan, C. Moritz, J. Neel, P. Pamilo, N. Saitou, L. Sattenspiel, P. Sharp, C. Stephens, A. Templeton, L. Vawter, and an anonymous reviewer for voluminous commentary on earlier drafts of the manuscript. We thank T. Goelz and J. Tworek for the programming and computations. P.E.S. was supported by DoE-86-ER60089 and NIH-R01-GM32589, and W.-H.L. was supported by NSF-BRS-83-03965.

Literature Cited

Aiello, L. C. 1981. The allometry of primate body proportions. Symp. Zool. Soc. Lond. 48:331–358.

Anderson, S., A. T. Bankier, B. G. Barrell, M. H. L. de Bruijn, A. R. Coulson, J. Drouin, C. Eperon, D. P. Nierlich, B. A. Roe, F. Sanger, P. H. Schreider, A. J. H. Smith, R. Staden, and I. G. Young. 1981. Sequence and organization of the human mitochondrial genome. Nature 290:457–464.

Andrews, P. 1982. Hominid evolution. Nature 295: 185–186.

Aoki, K., Y. Tateno, and N. Takahata. 1981. Estimating evolutionary distance from restriction maps of mitochondrial DNA with arbitrary G + C content. J. Molec. Evol. 18:1–8.

Avise, J. C. 1986. Mitochondrial DNA and the evolutionary genetics of higher animals. Phil. Trans. Roy. Soc. Lond. B 312:325–342.

Avise, J. C., and R. A. Lansman. 1983. Polymorphism of mitochondrial DNA in populations of higher animals, pp. 147–164. *In* M. Nei and R. K. Koehn (eds.), Evolution of Genes and Proteins. Sinauer, Sunderland, MA.

Benveniste, R. E., and G. J. Todaro. 1976. Evolution of type-C viral genes: Evidence for an Asian origin of man. Nature 261:101–108.

Bianchi, N. O., M. S. Bianchi, J. E. Cleaveo, and S. Wolff. 1985. The pattern of restriction enzyme-induced banding in the chromosomes of chimpanzee, gorilla and orangutan and its evolutionary significance. J. Molec. Evol. 22:323–333.

Brown, W. M. 1980. Polymorphism in mitochondrial DNA of humans as revealed by restriction endonuclease analysis. Proc. Nat. Acad. Sci. USA 77:3605–3609.

Brown, W. M., and H. M. Goodman. 1979. Quantitation of intrapopulation variation by restriction endonuclease analysis of human mitochondrial DNA, pp. 485–500. *In* P. Borst, I. B. Dawid, S. Weissman, and C. F. Cox (eds.), Extrachromosomal DNA, Symposia on Cellular Biology, Vol. 15. Academic Press, N.Y.

Brown, W. M., E. M. Prager, A. Wang, and A. C. Wilson. 1982. Mitochondrial DNA sequences of primates: Tempo and mode of evolution. J. Molec. Evol. 18:225–239.

Cann, R. L., W. M. Brown, and A. C. Wilson. 1984. Polymorphic sites and the mechanism of evolution in human mitochondrial DNA. Genetics 106:479–499.

DeBry, R. W., and N. A. Slade. 1985. Cladistic analysis of restriction endonuclease cleavage maps within a maximum-likelihood framework. Syst. Zool. 34:21–34.

Felsenstein, J. 1981. Evolutionary trees from DNA sequences: A maximum likelihood approach. J. Molec. Evol. 17:368–376.

———. 1983. Inferring evolutionary trees from DNA sequences, pp. 133–150. *In* B. Weir (ed.), Statistical Analysis of DNA Sequence Data. Dekker, N.Y.

Ferris, S. D., A. C. Wilson, and W. M. Brown. 1981*a*. Evolutionary tree for apes and humans based on cleavage maps of mitochondrial DNA. Proc. Nat. Acad. Sci. USA 78:2432–2436.

———. 1981*b*. Extensive polymorphism in the mitochondrial DNA of apes. Proc. Nat. Acad. Sci. USA 78:6319–6323.

Fleagle, J. G., J. T. Stern, Jr., W. L. Jungers, R. L. Susman, A. K. Vangor, and J. P. Wells. 1981. Climbing: A biomechanical link with brachiation and with bipedalism. Symp. Zool. Soc. Lond. 48: 359–375.

Goodman, M. 1985. Rates of molecular evolution: The hominoid slowdown. BioEssays 3:9–14.

Goodman, M., A. E. Romero-Herrera, H. Dene, J.

CZELUSNIAK, AND R. E. TASHIAN. 1982. Amino acid sequence evidence on the phylogeny of primates and other eutherians, pp. 115–187. *In* M. Goodman (ed.), Macromolecular Sequences in Systematics and Evolutionary Biology. Plenum, N.Y.

GREENFIELD, L. O. 1980. A late divergence hypothesis. Amer. J. Phys. Anthropol. 52:351–365.

HIXSON, J. E., AND W. M. BROWN. 1986. A comparison of the small ribosomal RNA genes from the mitochondrial DNA of the great apes and human: Sequence, structure, evolution and phylogenetic implications. Molec. Biol. Evol. 3:1–18.

KAPLAN, N., AND C. H. LANGLEY. 1979. A new estimate of sequence divergence of mitochondrial DNA using restriction endonuclease mapping. J. Molec. Evol. 13:292–304.

KASHYAP, R. L., AND S. SUBAS. 1974. Statistical estimation of parameters in a phylogenetic tree using a dynamic model of the substitution process. J. Theoret. Biol. 47:75–101.

KIMURA, M. 1981. Estimation of evolutionary distances between homologous nucleotide sequences. Proc. Nat. Acad. Sci. USA 78:454–458.

KLUGE, A. 1983. Reclassification of the great apes, pp. 151–177. *In* R. L. Ciochon and R. S. Corruccini (eds.), New Interpretations of Ape and Human Ancestry. Plenum, N.Y.

KOOP, B. F., M. GOODMAN, P. XU, K. CHAN, AND J. L. SLIGHTOM. 1986. Primate η-globin DNA sequences and man's place among the great apes. Nature 319:234–238.

LI, W-H. 1986. Evolutionary change of restriction cleavage sites and phylogenetic inference. Genetics 113:187–213.

MORTON, N. E. 1955. Sequential tests for the detection of linkage. Amer. J. Hum. Genet. 7:277–318.

NEI, M., AND W-H. LI. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc. Nat. Acad. Sci. USA 76:5269–5273.

NEI, M., J. C. STEPHENS, AND N. SAITOU. 1985. Methods for computing the standard errors of branching points in an evolutionary tree and their application to molecular data from humans and apes. Molec. Biol. Evol. 2:66–85.

NEI, M., AND F. TAJIMA. 1983. Maximum likelihood estimation of the number of nucleotide substitutions from restriction sites data. Genetics 105:207–217.

———. 1985. Evolutionary change of restriction cleavage sites and phylogenetic inference for man and apes. Molec. Biol. Evol. 2:189–205.

OXNARD, C. E. 1981. The place of man among the primates: Anatomical, molecular and morphometric evidence. Homo 32:149–176.

SAITOU, N., AND M. NEI. 1986. The number of nucleotides required to determine the branching order of three species with special reference to the human-chimpanzee-gorilla divergence. J. Molec. Evol. 24: 189–204.

SIBLEY, C. G., AND J. E. AHLQUIST. 1984. The phylogeny of the hominid primates, as indicated by DNA-DNA hybridization. J. Molec. Evol. 20:2–15.

SLIGHTOM, J. L., L. Y. E. CHANG, B. KOOP, AND M. GOODMAN. 1985. Chimpanzee fetal $^{G}\gamma$ and $^{A}\gamma$ globin gene nucleotide sequences provide further evidence of gene conversions in hominid evolution. Molec. Biol. Evol. 2:370–389.

STERN, J. T., AND R. L. SUSMAN. 1981. Electromyography of the gluteal muscles in *Hylobates, Pongo,* and *Pan*: Implications for the evolution of hominid bipedality. Amer. J. Phys. Anthropol. 55: 153–166.

TAJIMA, F., AND M. NEI. 1982. Biases of the estimates of DNA divergence obtained by restriction enzyme technique. J. Molec. Evol. 18:115–120.

TEMPLETON, A. R. 1983*a*. Convergent evolution and nonparametric inferences from restriction data and DNA sequences, pp. 151–179. *In* B. Weir (ed.), Statistical Analysis of DNA Sequence Data. Dekker, N.Y.

——— 1983*b*. Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to the evolution of humans and the apes. Evolution 37:221–244.

———. 1985. The phylogeny of the hominoid primates: A statistical analysis of the DNA-DNA hybridization data. Molec. Biol. Evol. 2:420–433.

TUTTLE, R. H. 1981. Evolution of hominid bipedalism and prehensile capabilities. Phil. Trans. Roy. Soc. Lond. B 292:89–94.

UEDA, S., O. TAKENAKA, AND T. HONJO. 1985. A truncated immunoglobulin ϵ pseudogene is found in the gorilla and man but not in chimpanzee. Proc. Nat. Acad. Sci. USA 82:3712–3715.

WASHBURN, S. L. 1982. Human evolution. Persp. Biol. Med. 25:583–602.

WILSON, A. C., S. S. CARLSON, AND T. J. WHITE. 1977. Biochemical evolution. Ann. Rev. Biochem. 46:573–639.

YUNIS, J. J., AND O. PRAKASH. 1982. The origin of man: A chromosomal pictorial legacy. Science 215: 1525–1530.

Corresponding Editor: M. Slatkin

APPENDIX

The probility functions of the 16 restriction classes listed in Table 4 can be derived from the same sort of network addition and subtraction algorithm used in Figure 1. A pair of examples have been worked in Li (1986) for those who want more detail. The function $A(t_1, t_2, t_3)$ is

$$\begin{aligned} A(t_1, t_2, t_3) = {} & a\xi(t_1, t_1, t_2)P^r(t_1 + t_2) \\ & \cdot P^r(t_1 + t_2 + 2t_3) + aP^{2r}(t_1) \\ & \cdot [\xi(t_2, t_1 + t_2, t_1 + t_2 + 2t_3) \\ & - P^r(t_2)P^r(t_1 + t_2)P^r(t_1 + t_2 + 2t_3)] \\ & + a[Q^2(t_1)\, Q(t_1 + t_2)Q(t_1 + t_2 + 2t_3) \\ & + 2S^2(t_1)S(t_1 + t_2)S(t_1 + t_2 + 2t_3) \\ & + P^2(t_1)P(t_1 + t_2)P(t_1 + t_2 + 2t_3)]^r \\ & - aP^{2r}(t_1)P^r(t_1 + t_2)P^r(t_1 + t_2 + 2t_3) \end{aligned} \tag{A.1}$$

to a very close approximation when t_2 is small, as it is here. A much more elaborate (but marginally more exact) value is presented in Li (1986). Similar treatment yeilds

$$B(t_1, t_2, t_3) = U(t_1, t_2) - A(t_1, t_2, t_3) \tag{A.2}$$

$$\begin{aligned} C(t_1, t_2, t_3) = {} & Y(t_1 + t_2, t_3) \\ & - J(t_1, t_2, t_3) \end{aligned} \tag{A.3}$$

$$D(t_1, t_2, t_3) = V(t_1, t_2) - C(t_1, t_2, t_3) \quad \text{(A.4)}$$

$$E(t_1, t_2, t_3) = a\xi(t_1, t_1, t_1 + 2t_2 + 2t_3) - A(t_1, t_2, t_3) \quad \text{(A.5)}$$

$$F(t_1, t_2, t_3) = a\xi(t_1 + t_2, t_1 + t_2, t_1 + t_2 + 2t_3) - A(t_1, t_2, t_3) \quad \text{(A.6)}$$

$$G(t_1, t_2, t_3) = W(t_1, t_2) - E(t_1, t_2, t_3) \quad \text{(A.7)}$$

$$H(t_1, t_2, t_3) = X(t_1, t_2) - F(t_1, t_2, t_3) \quad \text{(A.8)}$$

$$I(t_1, t_2, t_3) = X(t_1 + t_2, t_3) - F(t_1, t_2, t_3) \quad \text{(A.9)}$$

$$J(t_1, t_2, t_3) = X(t_1, t_2 + t_3) - F(t_1, t_2, t_3) \quad \text{(A.10)}$$

$$K(t_1, t_2, t_3) = Y(t_1, t_2) - I(t_1, t_2, t_3) \quad \text{(A.11)}$$

$$M(t_1, t_2, t_3) = Z(t_1, t_2) - J(t_1, t_2, t_3). \quad \text{(A.12)}$$