Templeton (1985) presented a method (delta Q) to test the hypothesis that two trees are not distinguishable. Using that test, he analyzed the data of Sibley and Ahlquist (1984) on humans and apes and concluded that those data, which had been interpreted as supporting chimpanzees and humans as sister groups, could not validly be used to do so because they failed the delta Q-test. At least one data set that might support chimpanzees and gorillas as sister groups did not fail the delta Q-test, and Templeton argued strongly for the latter association. The writers of the preceding letters felt that Templeton's conclusion was not as strong as it appeared because the delta Q-test was overvalued and/or the virtues of DNA-DNA hybridization were underappreciated. What follows is an attempt to give the reader further facts with which to explore the issues.

Let \( d_{XY} \) equal the distance between two taxa, X and Y. I shall use as subscripts G, Gorilla; H, Homo; and P, Pan. In this notation, Sibley and Ahlquist distances are \( d_{GH} = 2.39 \pm 0.2 \), \( d_{GP} = 2.20 \pm 0.25 \), and \( d_{HP} = 1.81 \pm 0.24 \) (see fig. 1). In this case, the delta Q method cannot possibly distinguish between phylogeny A and phylogeny B (or C) at the 95% confidence level because there are too few taxa. Moreover, it is irrelevant how many out-groups are included (birds, turtles, frogs, fish, etc.).

It happens, however, that the chimpanzees are a collection of two species. It also happens that the delta Q method could then support phylogeny A over phylogeny B in figure 1—but only if the amount of change between humans and gorillas were less than that between gorillas and any of the other apes. One may ask, Why should the distance between humans and gorillas affect the decision whether humans or gorillas are closer to the chimpanzees?

The delta Q-test requires the assumption of rate constancy; otherwise, ordering by distance cannot reflect the order of branching. It is then legitimate to ask, Should this test be applied in cases in which a failure of that assumption is required to achieve the confidence level? (The failure is the need for the human–gorilla distance to be less than the two chimpanzee–gorilla distances).

There are no more measurements in the data set, but the subdivision of the Pan data into the groups \( P. \) paniscus and \( P. \) troglodytes has converted the data from a set for which the delta Q-test could not possibly distinguish tree A from tree B to one in which it is possible but does not. Saitou points out that if those very data had been for three species rather than two, then the test would have found that tree A was significantly better than tree B. Thus the ability to decide an issue depends on how high up in the tree one can make the issue to lie.

The test is only dependent on the relative order of distances, not on the magnitude of those distances as Templeton has stated. If, for example, the G in figure 1 were goats rather than gorillas and all distances involving G were tenfold greater (23.9, 22.9), all of the above delta Q remarks would still apply; one (a) could not resolve whether chimpanzees are closer to humans or to goats for the three-taxon case, (b) might have resolved (but could not) the issue on recognizing that there are two chimpanzee species in the data, and (c) would decide that chimpanzees are closer to humans than to goats if the chimpanzee data were for three rather than two species.

In determining whether Sibley and Ahlquist’s data are suitable for deciding the issue of tree A versus tree B, the question arises whether the delta Q-test is a valid basis for their rejection. Templeton states that the rejection is valid until one has a better way of testing the difference between trees A and B. Such a test is available and suggested itself to me by Templeton’s asking (1985, p. 422) whether \( d_{GP} \) (2.20) is significantly less than or approximately equal to \( d_{GH} \) (2.39). Templeton suggested that if the former (less than) were true it would support tree B, that if the latter (approxi-
Fig. 1.—The delta T\textsubscript{50H} values of Sibley and Ahlquist (1984). The 45 DNA-DNA hybridization measurements (in degrees C) are presented as three histograms showing the number of instances found at each melting temperature for the three possible pairs of taxa. The Pan data are a pool of P. paniscus and P. troglodytes data. For each pair the mean delta T\textsubscript{50H} and the SD are shown. For each pair is shown that phylogeny appropriate to that pair being sister taxa along with numbers representing the amount of change that had to occur on each branch to give the observed distances. Values are the distances from the trifurcation (●) to the three tips.

Evolutionary rates need not be equal to allow one to obtain the correct order of divergence from distance measurements. All that is required is that the amount of
Table 1

Test of Significant Differences Between Pairs of Distances from Sibley and Ahlquist (1984)

<table>
<thead>
<tr>
<th>TEST</th>
<th>DISTANCE PAIR</th>
<th>n</th>
<th>U</th>
<th>0.01</th>
<th>0.001</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$d_{HP} &lt; d_{GP}$</td>
<td>19</td>
<td>37</td>
<td>82</td>
<td>60</td>
<td>$\leq 10^{-3}$</td>
</tr>
<tr>
<td>2</td>
<td>$d_{HP} &lt; d_{GH}$</td>
<td>19</td>
<td>5</td>
<td>44</td>
<td>29</td>
<td>$\leq 10^{-3}$</td>
</tr>
<tr>
<td>3</td>
<td>$d_{GP} &lt; d_{GH}$</td>
<td>16</td>
<td>40.5</td>
<td>36</td>
<td>23</td>
<td>$&gt; 10^{-2}$</td>
</tr>
</tbody>
</table>

NOTE.—Test statistic is the Mann-Whitney $U$, which is the number of times that a distance for the left pair exceeds a distance for the right pair; $n$ is the number of distance measurements for the left and right pair of distances. The critical values are the values below which $U$ must lie to reject the null hypothesis that the inequality on the left is untrue at the probability levels shown. The test is one-tailed since the reverse inequality is part of the null hypothesis. The individual values that were ranked are shown in fig. 1. The test and tables of critical values were taken from Siegel (1956, pages 116–126 and 274–277, respectively).

observable change in each of the lines descending from the lower node (shown by a black dot [●] in fig. 1) be less than that from the lower node up through the time of the first divergence and back down to the third taxon, in order that the inequality of rates not be so great that it interferes with distinguishing the order. The point where it does is not easy to determine. In Sibley and Ahlquist’s tree (A), the single-copy DNA of the human lineage has, empirically, evolved 25% faster than that of the chimpanzee line, with the gorilla lineage plausibly close to one of them. In Templeton’s tree (B), the gorilla lineage has evolved 75% faster than the chimpanzee’s and more than 40% faster than the human.

The Mann-Whitney $U$-test suggests that the Sibley-Ahlquist data are admissible evidence in the question of the order of divergence among humans, gorillas, and chimpanzees and that their rejection by Templeton occurred because the delta $Q$-test was too weak to discover their merit. The more important question, however, may be the following: Can a test be so weak in a particular case (for four taxa especially) that it ought not to be used at all?

Walter M. Fitch,
Editor in Chief

LITERATURE CITED