Using a modified form (delta Q [ΔQ]-test) of Pielou’s (1979) Q-test, Templeton (1985) reanalyzed Sibley and Ahlquist’s (1984) DNA-DNA hybridization data for hominoid species. He concluded that the phylogeny (fig. 1A) supported by Sibley and Ahlquist is not significantly better than the phylogeny (fig. 1B) favored by Templeton (1983). However, Templeton’s (1985) ΔQ-test has several statistical problems, and his conclusion does not seem to be justified.

In brief, Pielou’s Q-test and Templeton’s ΔQ-test are as follows: Consider the distance matrix in table 1A, where \(d_{ij}(i > j)\) denotes the evolutionary distance between species \(i\) and \(j\). In the Q-test, \(d_{ii}\) is compared with \(d_{kl}(k > l)\) with the restriction \(i < k\), and a random variable \(x\), which takes 1 when \(d_{ij} < d_{kl}\), 0 when \(d_{ij} > d_{kl}\), and 0.5 when \(d_{ij} = d_{kl}\), is considered. The Q-statistic is the sum of \(x\) for all possible comparisons of \(d_{ij}\) and \(d_{kl}\). If we assume that phylogeny A in figure 1 is correct and that \(d_{ij}\) increases in proportion to evolutionary time without error, we will have the inequality \(d_{ij} < d_{kl} < d_{l}\) (see table 1A). The Q-value for matrix A of table 1 is 35. By contrast, the Q-value for matrix B, which corresponds to phylogeny B of figure 1, becomes 29–31, depending on the differences among \(d_{12}\), \(d_{13}\), and \(d_{14}\) (see table 1B). The distribution of ΔQ (Templeton’s table 2) indicates that ΔQ = 6 is required to achieve statistical significance at the 5% level, but other, smaller values are nonsignificant. If we apply the above test to Sibley and Ahlquist’s (1984) data, ΔQ becomes 4. Therefore, this test suggests that phylogeny A is not favored against phylogeny B.

However, we note that ΔQ depends only on the ranks among \(d_{ij}\)’s in the present case. With the inequality \(d_{ij} < d_{kl} < d_{l}\) (\(j = 1, \cdots, 4\)) maintained, ΔQ becomes 6 only when \(d_{12}\) is smaller than both \(d_{14}\) and \(d_{24}\). Even when the strict rate constancy is assumed, where \(d_{14} = d_{24} = d_{34}\), ΔQ = 5 (see table 1). Obviously, \(d_{12}\), \(d_{13}\), and \(d_{14}\) may vary by chance effects even if species 4 is remotely related to species 1, 2, and 3, and ΔQ = 4 may be obtained with a high probability. It is therefore clear that the ΔQ-test is inadequate for testing topological differences.
Of course, one can still argue that phylogeny A is better than phylogeny B when $\Delta Q$ becomes 6. When we consider the branching orders of humans, chimpanzees, and gorillas, however, there is another possible phylogeny (fig. 1C). This phylogeny cannot be compared with phylogeny A or B in terms of $\Delta Q$ (Templeton 1985). Therefore, even if phylogeny A is judged to be superior to phylogeny B with $\Delta Q = 6$, it is not necessarily the best tree; the true tree could be phylogeny C.

Another problem of the $\Delta Q$-test is that the power of the test depends on the number of OTUs (operational taxonomic units) compared. Templeton (1985) acknowledged this point but considered it as a good support for the $\Delta Q$-test. Suppose that we include the third species of chimpanzee (made-up) and obtain phylogeny A' of figure 1. In this case, if we switch species 3 and 4, $\Delta Q$ becomes 9–12, again depending on the values of $d_{41}$ ($\approx d_{41}$), $d_{42}$, and $d_{43}$. The smallest $\Delta Q (=9)$ is still not significant at the 5% level, but the probability of $\Delta Q \geq 9$ is now 0.0643. As the number of OTUs increases from the chimpanzee lineage gradually, even the smallest $\Delta Q$ becomes statistically significant. This peculiar property has occurred because the probability distribution of $\Delta Q$ is obtained under the assumption that no hierarchical structure of OTUs exists. This assumption is certainly unrealistic in most statistical studies of phylogenetic trees.
Table 1
Various ΔQ-Values for Two Distance Matrices

<table>
<thead>
<tr>
<th>Species</th>
<th>Conditions</th>
<th>ΔQ'</th>
<th>ΔQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>31</td>
<td></td>
<td>4</td>
<td>(d_{43} &gt; d_{41}) and (d_{43} &gt; d_{42})</td>
</tr>
<tr>
<td>30.5</td>
<td></td>
<td>4.5</td>
<td>(d_{41} &lt; d_{43} = d_{43}) or (d_{43} &lt; d_{41} = d_{43})</td>
</tr>
<tr>
<td>30</td>
<td></td>
<td>5</td>
<td>(d_{41} &lt; d_{43} &lt; d_{42}) or (d_{42} &lt; d_{43} &lt; d_{41}) or (d_{41} = d_{42} = d_{43})</td>
</tr>
<tr>
<td>29.5</td>
<td></td>
<td>5.5</td>
<td>(d_{41} &gt; d_{43} = d_{43}) or (d_{42} &gt; d_{43} = d_{43})</td>
</tr>
<tr>
<td>29</td>
<td></td>
<td>6</td>
<td>(d_{43} &lt; d_{41}) and (d_{43} &lt; d_{42})</td>
</tr>
</tbody>
</table>

In my view, Templeton’s (1985) criticism of Sibley and Ahlquist’s (1984) use of the t-test is not really justified. Since Sibley and Ahlquist considered only experimental errors, all the TS0H values are independent. Therefore, I do not think that there is anything wrong with the t-test. The pooling of “Gorilla × Pan, Homo” comparisons as done by Sibley and Ahlquist (1984) also seems to be valid, since they wanted to show that the human-chimpanzee clustering obtained by the distance Wagner method was significant. Templeton (1985) used the t-test for Gorilla-Homo vs. Gorilla-P. troglodytes comparisons and showed that the former is significantly larger than the latter. This may suggest that Sibley and Ahlquist’s (1984) DNA-DNA hybridization data are not strictly rate constant. However, this result alone should not, despite Templeton’s (1985) suggestion to the contrary, be taken as the evidence of the existence of inconsistency in Sibley and Ahlquist’s data.

At any rate, if we can apply this t-test to DNA-DNA hybridization data, why is it necessary to use Templeton’s ΔQ-test, which has almost no statistical power?

Acknowledgments

I thank Dr. M. Nei for encouraging discussions and detailed suggestions. I also thank Drs. P. E. Smouse, W.-H. Li, P. Pamilo, J. C. Stephens, and M. Ruvolo for discussions and comments on the manuscript.

LITERATURE CITED


WALTER M. FITCH, reviewing editor

Received October 4, 1985; revision received October 28, 1985.