Research News

DNA Clock Conflict Continues

Two researchers who apparently revolutionized bird and human/ape systematics using the technique of DNA hybridization are faced with having to reanalyze all their data

FOR THE PAST SEVERAL MONTHS two highly critical manuscripts have been circulating widely among anthropologists and molecular biologists, their being the work of Charles Sibley and Jon Ahlquist, both until recently at Yale. The authors of the manuscripts—Vincent Sarich of the University of California, Berkeley, Carl Schmid of the University of California, Davis, and Jon Marks of Yale—challenge Sibley and Ahlquist's handling of DNA hybridization data concerning the evolutionary relationships among humans and apes and among hundreds of species of birds.

Specifically, Sarich and his colleagues contend that Sibley and Ahlquist used an inappropriate method of analysis on their data, thus giving misleading results. More seriously, however, they also suggest that as much as 40% of the data concerning humans and apes have apparently been "subjected to manipulations of an unspecified nature." More bluntly, Sarich told *Science* that in his opinion "most of the changes are in a direction to make the data look better." He also contends that the episode was likely to be "a disaster for the field."

Although Sibley and Ahlquist concede that, in retrospect, their data handling might have been more careful, they strongly deny any impropriety. They also hold that Sarich and his colleagues' scientific arguments are merely cloaks for other motivations, having to do, for instance, with strong, opposing preconceptions. It is certainly true that in the 10 months that this dispute has been fomenting, it has reached a highly personal pitch, exploding on occasions with bitter public exchanges at scientific meetings.

Last week's article chronicled the events leading up to these charges and countercharges. Here, the substance of these charges will be examined.

Ornithologists by training, Sibley and Ahlquist have devoted the great majority of their efforts with DNA hybridization to tackling questions of evolutionary patterns among birds, a venture that has involved more than 25,000 individual experiments over a period of 10 years.

Although Sibley and Ahlquist's recent interest in human/ape relationships represents less than 2% of their work, this has been far more visible publicly, not least because it produced a surprise: namely, it indicated that humans and chimpanzees are genetically closer to each other than either is to the other African ape, the gorilla; most anthropologists had supposed the two apes to be each other's closest relative.

This is the second of two articles examining a current conflict over recent important molecular clock results concerning evolutionary patterns among birds and among humans and apes.

As a result of all this work, Sibley and Ahlquist's experience in the use and interpretation of DNA hybridization matches or exceeds that of any other researcher's in this country. Although they have not been responsible for any technical advances in the technique, Sibley and Ahlquist have understandably developed great confidence in their ability to judge the quality of each set of results as it comes off their machine, a factor that was to prove crucial in the genesis of the current episode.

> ***"The justification for making such corrections derives from our long experience in looking at the melting curves from the bird work."***

Sarich and his colleagues produced their two manuscripts in March this year, tailored for the different audiences of the *Journal of Molecular Evolution* and the *Journal of Human Evolution* respectively. Although the former was actually written mostly by Sarich and the latter by Marks, both had been inspired by Marks' determination to bring Sibley and Ahlquist's work under public scrutiny. Both were uncompromising in their language.

"We begin our critical analysis by calling attention to a strange and relatively unappreciated aspect of the discussion (both *pro* and *con*)—that it has been taking place without anyone ever having seen any of the original data which, in effect, have never been published," wrote Sarich. "The reader is simply put in the position of trusting, or not, in the good intentions and judgment of the investigators. This is not the way science should be done."

What appears here to be a stinging criticism of Sibley and Ahlquist's scientific practice—that is, not presenting raw data in their publications—is in fact rather empty. "The actual raw data—that is, a series of radioactive counts—are rarely published in papers reporting results of DNA hybridization experiments," notes Jeffrey Powell of Yale. "Even if you wanted to publish them, journal editors usually refuse, because they would take up a great deal of space." Emile Zuckerkandl, editor of the *Journal of Molecular Evolution*, confirms that authors are not required to include raw DNA hybridization data. "It is analogous to publishing DNA sequence results," he says. "These days you publish the sequence, but not the gels from which you derive the sequence."

Although researchers are not typically required to publish raw DNA hybridization data, it is considered normal practice to make them available to colleagues who have a legitimate interest. When last year Marks approached Sibley and asked to see the raw data for the human/ape experiments he was surprised to receive a refusal. "We are not compelled to give our data to just anyone who asks," explains Sibley. "And in Marks' case we were suspicious of the motives."

Marks insists his interest is solely in ensuring the highest scientific standards in molecular phylogenetics. However, some observers of the current conflict point out that he is not entirely disinterested in the outcome, particularly in relation to anthropology. For instance, Marks' own thesis work—on chromosome banding—supported the more conventional chimpanzee/gorilla association. And last fall he made what was judged by some commentators to be an unsubstantiated attack on the significance of some DNA sequence data, published in *Science* (16 October 1987, page 369) by Morris Goodman of Wayne State University, which supported Sibley and Ahlquist's conclusions.

Even if Sibley had been willing to give Marks the human/ape data, for practical reasons he would have found it difficult to comply. Not only had he and Ahlquist recently left Yale for two different universities, thus engendering the organizational confusion that such a move typically entails, but also the data themselves were in a state of some disarray. "Our method of recording and working with the data had switched several times, as the Yale computer system changed, first in one way then another," says Ahlquist. "The data were in several different forms, and, when Marks asked for them, in several different places." In short, a mess.

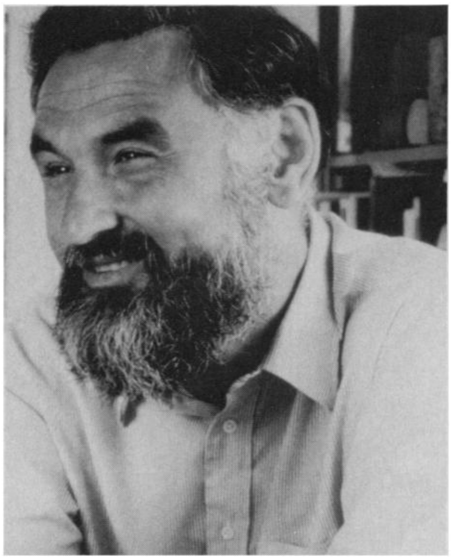
Eventually, in December last year, Marks did obtain some of Sibley and Ahlquist's data, given to him by Roy Britten, of the California Institute of Technology, with Sibley's permission. This sample of data included about 10% of the total for the human/ape comparisons, and was sufficient for Sarich and his colleagues quickly to build some strong conclusions.

Here, we will examine three of those conclusions, which appeared explicitly in the manuscript submitted to the *Journal of Molecular Evolution*: first, "The T50H statistic [used by Sibley and Ahlquist] is fundamentally and fatally flawed as a measure of mean sequence divergence, and also, therefore, as a measure of phylogenetic distance;" second, "There is clear evidence that certain of the published data are not representative of the actual experiments on which they are based;" and third, "The claims of Sibley and Ahlquist to have resolved the human/chimpanzee/gorilla [relationship] are not supported by their data."

The impact of all this has already been substantial, notably in creating the impression that Sibley and Ahlquist are guilty as charged, that their work has indeed been shown to be flawed. However, even though the manuscript carrying the above statements was submitted to the *Journal of Molecular Evolution* last March and has since passed through several revisions, last week it was rejected by the journal's editor, Zuckerkandl. The second manuscript, submitted to the *Journal of Human Evolution*, has yet to be accepted for publication. Meanwhile, both manuscripts have been widely circulated and have formed the basis of strong opinions about the value of Sibley and Ahlquist's work.

Moreover, early in the year Allan Wilson, of the University of California, Berkeley, included a copy of the first version of the *JME* manuscript as part of an assessment of a research grant proposal submitted by Sibley to the National Science Foundation (NSF). The proposal was rejected. Wilson, a long-time associate of Sarich in work on molecular phylogenetics, is an associate editor of the *Journal of Molecular Evolution* and was handling the review process of Sarich's manuscript.

Some observers consider Wilson to have acted improperly in using the manuscript as part of an NSF grant proposal assessment, given its strong language and highly negative tone, but particularly the fact that at the time it was yet to be reviewed. However, NSF has no specific rules about this sort of thing, and, as one program director told *Science*, "such a manuscript could be used as just another piece of information in the panel's deliberations."



Vincent Sarich: *"a disaster for the field."*

The first of the three specific conclusions mentioned above concerns the method by which Sibley and Ahlquist analyze their data, namely the use of the so-called T50H statistic, which Sarich and his colleagues describe as being "fundamentally and fatally flawed." The T50H measure is one of three options for interpreting DNA hybridization data, the other two being Tm and Tmode. Each of the three attempts to extract an indication of how similar is the overall DNA sequence between two species.

Crudely put, the difference between T50H, Tm, and Tmode is that there is, respectively, a decreasing amount of DNA involved in the comparison. There are potential drawbacks to each method. Proponents of T50H—Sibley and Ahlquist—argue that, because it encompasses more DNA in the comparison, it contains more information with which to discriminate between species. Proponents of Tmode—Sarich and his colleagues—counter by saying that the extra DNA involved in T50H is extraneous and therefore leads to confusion, not discrimination, and that Tmode gives a cleaner, more reliable result.

In arguing that the Tmode statistic is the superior measure, Sarich and his colleagues note that "this choice has generally been rejected in actual publications by DNA systematists." They add: "We believe that Sibley and Ahlquist, along with other workers in the field, made a very serious mistake in this judgment." Five referees were invited to review the Sarich *et al.* manuscript. Britten, a pioneer in the development of the DNA hybridization technique, although not invited by Wilson to act as a referee, nevertheless obtained a copy of the manuscript and offered comments directly to Zuckerkandl.

While universally deploring the very personal and insulting tone of the manuscript, the anonymous reviewers were generally impressed with the arguments. "The finest possible analytical job," observed one. "An important contribution to the field of molecular evolution," said another.

Britten, however, was less impressed, suggesting that Sarich and his coauthors had made serious errors. He says that the authors offer "No grounds . . . that identify a 'fatal flaw' in the [T50H] method. This is a mistake in judgment by the authors." He adds that they "arrive in an unexplained way at the opinion that the Tmode is the best measure of divergence of DNA sequences."

It might be claimed, as Marks has, that Britten, as the codeveloper of the T50H method, cannot now be impartial about its value. It must be equally true, of course, that this same fact has given Britten the experience necessary to make an informed judgment. In any case, he has a supporter in Powell, who has recently been checking the efficacy of DNA hybridization in doing the job required of it: namely, giving a measure of the divergence in DNA sequence.

Earlier this year Powell, in company with Adalgisa Caccone and Rob DeSalle then also at Yale, reported that Tm measures and sequence divergence in *Drosophila* DNA "is remarkably linear . . . with a correlation coefficient [greater than] 0.98." Powell has also demonstrated a clear, linear relationship between Tm and T50H measures, the implication being that T50H is indeed a good indication of DNA sequence divergence. "With these results I don't see that Tmode has any advantages," Powell told *Science*. "There is no justification for it."

The strong differences of opinion over the validity of the different statistical measures is surely an indication that there is more to be learned about the DNA hybridization process. Perhaps some of this will emerge at a major UCLA meeting on the subject, which is to be held early next year. In the meantime, there is no widely accepted theoretical argument that would condemn Sibley and Ahlquist's chosen method of analysis, and a good deal of empirical support for its use.

What, then, of the second charge?: "that certain of the published data are not representative of the actual experiments on which they are based."

When, last December, Marks obtained the small proportion of data—principally on the human/ape relationship—he and Sarich and Schmid were able to work out the Tmode, Tm, and T50H statistics for themselves. They then compared them with the T50Hs published by Sibley and Ahlquist in the same month in a special issue of the *Journal of Molecular Evolution*, and discovered a striking discrepancy: "of the 50 values published . . . for which we have the actual data, fully 20 differ by more than 0.5° from what can be calculated from those data."

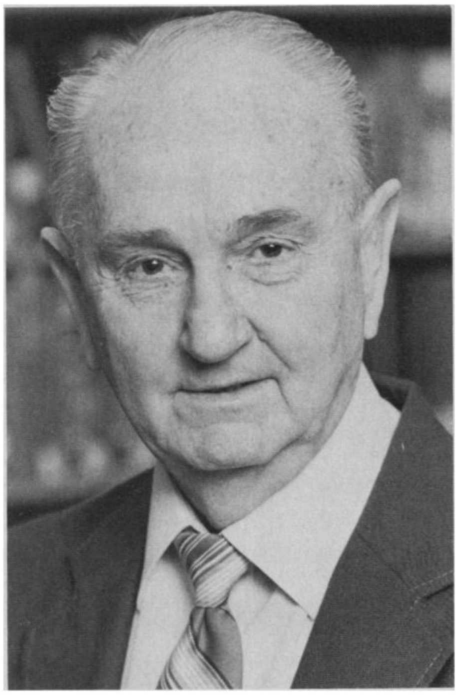
Marks approached Ahlquist for an explanation of these discrepancies, detailing them in a letter dated 16 February. Ahlquist replied on 21 March in a long letter, saying "I, too, share your desire to resolve our differences regarding the DNA data without acrimony and to that end offer the first of what will be many letters to explain how these data are analyzed." By this time, however, Sarich and his colleagues' two manuscripts had already been drafted, and would soon be widely distributed.

In that 21 March letter Ahlquist had explained that "*On occasion*, it becomes necessary to correct values," adding that "The justification for making such corrections derives from our long experience in looking at the melting curves from the bird work." He went on to explain that it was possible to discriminate good curves from bad ones, and how some could be rescued while others had to be thrown out.

"Data are sacrosanct," proclaims Marks. "You don't 'correct' them . . . ever." A fine principle, no doubt, but most researchers are less rigid. "Every practicing scientist, after years of experience in a laboratory, knows when something is working and when it isn't," says Powell. "You can argue that this isn't being objective, but we all make judgment calls." Powell is also careful to add that, if there is a clear justification for altering data in any way, then the criteria for doing must be fully explained in any publication. No such explanation has accompanied any of Sibley and Ahlquist's published papers. "This is obviously against the rules," says Britten. "It will be a severe embarrassment for Charles and Jon."

Just how embarrassing it will be depends to some degree on the extent of the data correction. Sarich and his colleagues have always contended that, as the small dataset they obtained was in effect a random sample, then it is more than likely that corrections will have been made to 40% of the total set. Sibley and Ahlquist told *Science* that the figure was more like 20%. A possible way to resolve this arose when *Science* learned that, contrary to general knowledge, a complete set of Sibley and Ahlquist's human/ape data exists on computer disc, in the care of Frederick Sheldon of the Academy of Natural Sciences, Philadelphia. Running these data through a T50H program would allow comparison with all, not just 10%, of Sibley and Ahlquist's published T50H numbers.

Sheldon had worked with Sibley at Yale, and was to have been a coauthor on the 1987 *JME* paper, which is why he put all the data on disc. "I became uncomfortable when some of my numbers didn't match with Charles' and Jon's," he says, "and so I withdrew." Since parting company early last year, relations between Sibley and Sheldon have been a little strained. Nevertheless, Sibley agreed to *Science*'s request to have Sheldon run the human/ape data through a T50H program, so that a comparison could be made. The result is that a little more than 40% of Sibley and Ahlquist's published T50H numbers have apparently been altered by 0.5° or more.



Charles Sibley: *"this is very embarrassing."*

"We were too casual about all this," Sibley now concedes. "It did not seem to be very important at the time, because the errors seemed to be clear and it was easy to see in which direction they were. Yes, of course we should have indicated that we'd corrected them. And yes, it is very embarrassing."

The great majority of the corrections have the effect of making what otherwise would be varying degrees of outlying points come closer to the average for the set. In other words, they make the data look better. Whether the corrections were done "*to* make the data look better," as Sarich charges, is, however, open to question. "It's not fraud," says Sarich. "The term fraud should be reserved for when people fake data, and that's not the case here. But there clearly is a case of data fudging with the [human/ape] work." Sibley dismisses Sarich's suggestion, and says that "the corrections were simple and based on logic."

Nevertheless, even those observers who see no impropriety in Sibley and Ahlquist's work agree that the corrections were at best misguided. "Sloppy science," is how one close colleague puts it. As a result, much of the territory already worked by Sibley and Ahlquist will have to be revisited. The prospect is daunting, particularly with the bird work, in which intricate evolutionary patterns have been constructed, many of which even Sibley and Ahlquist's mildest critics fear will not hold up.

Meanwhile, Sibley and Ahlquist have already begun a reassessment, having spent the summer getting all the human/ape data on disc. "We are now in a position to run the data—without corrections—through a statistical package that will give us all three measures, T50H, Tm, and Tmode," Sibley told *Science* last week. "When we've had an opportunity to reanalyze the data ourselves, then we'll be ready to hand them out to anyone who wishes to see them."

This brings us to the last of the three major charges against Sibley and Ahlquist: "The claims of Sibley and Ahlquist to have resolved the human/chimpanzee/gorilla [relationship] are not supported by their data."

Sarich and his colleagues made this claim after doing the appropriate comparisons of T50H numbers calculated from their sample of uncorrected data. They report only "slight differences" that "do not discriminate" between the human/chimp and human/gorilla comparisons. Once again Britten checked Sarich *et al.*'s conclusion, and once again disagreed forcefully, describing it as "another mistake of fact." Sibley and Ahlquist's data—minus corrections—give the result they reported. Impasse?

At this point the most valuable contribution to the conflict would be an independent repetition of Sibley and Ahlquist's DNA hybridization work, which is precisely what Marks persuaded his Yale colleague Powell to do in January. Powell has 5 years experience in DNA hybridization work, but all with *Drosophila* comparisons. Marks asked Powell to try his hand at primates because he knew the work would be "of the highest quality." However, when Powell obtained his human/ape results, Marks' initial reaction was to say they must be incorrect; later he would dismiss them as irrelevant.

"Our data are in remarkable agreement with those previously obtained [by Sibley and Ahlquist]," reported Powell and Caccone. Not only did these new data give the same evolutionary tree—grouping chimpanzees more closely to humans than to gorillas—but they also gave precisely the same genetic distances claimed by Sibley and Ahlquist. "Whatever charges have been brought against Sibley and Ahlquist, their work has passed the acid test of science," Powell told *Science*. "They have been replicated in an independent laboratory."

Marks' argument that Powell's data are irrelevant is this: "He is not replicating their data; he's getting similar conclusions to the conclusions that Sibley and Ahlquist have not been willing to substantiate." Britten dismisses this as "pretty tortuous." He does note, however, that as Powell and Caccone's method of DNA hybridization is technically different from the one used by Sibley and Ahlquist, "it is not strictly speaking a *replication* of Charles' results, but it is certainly a *confirmation* of them."

Although Powell and Caccone's results clearly bring comfort to Sibley and Ahlquist, they also bring something of a mystery. The mystery is this: if, as DNA hybridization apparently indicates, the genetic distance between gorillas on one hand and humans and chimpanzees on the other is substantial, why have other molecular techniques not readily picked this up?

It is true that after Sibley and Ahlquist first suggested the human/chimp association in 1984, several research groups using techniques such as DNA sequencing, restriction-enzyme mapping, and protein electrophoresis have since reported the same shape for the human/ape evolutionary tree. But the genetic distance indicated between gorilla and human/chimpanzee has typically been about one quarter of that given by DNA hybridization. Are these techniques missing something? Or is there something consistently amiss with DNA hybridization?

Meanwhile, the conflict over the Sibley/Ahlquist data has for the moment cast something of a pall over the use of DNA hybridization in molecular phylogenetics. There is no doubt that Sibley and Ahlquist were seriously in error in making substantial, unreported alterations to their data, and Sarich and his colleagues deserve credit for bringing this to light. Exactly how much impact these corrections, and subsequent interpretation, will have on the body of work as a whole remains to be seen. But the very combative and partisan tone with which the challenges have been made has not advanced Sarich and his colleagues' stated concern with scientific integrity.

■ ROGER LEWIN

Hard Choices Ahead on Biodiversity

With many species on the verge of extinction, biologists call for a quick and dirty survey to chart the biodiversity of the planet

FACED WITH STARTLING NEW EVIDENCE that tropical rainforests may be disappearing at an even faster rate than previously believed, a number of prominent biologists are calling for an immediate effort to chart the biological diversity of the planet.

The idea is a quick and dirty survey, using whatever tools are at hand, to identify areas rich in biodiversity—areas that house many species, or species that exist nowhere else—so that they can be protected before they disappear. There is no time for exhaustive studies, elaborate phylogenies, or science as usual, says Thomas Lovejoy, assistant secretary for external affairs at the Smithsonian Institution: "The problem is very big and the fuse very short." Although hard and fast numbers are difficult to come by, it is estimated that half the world's species will be lost within the next century, mostly to deforestation.

Lovejoy called for the mapping endeavor, an idea that has been in the air for a couple of years, at a meeting on biodiversity attended by more than 3000 scientists last month in Davis, California.* By his "top of the head" estimate, it might take 2 or 3 years to complete and cost $250 million to $500 million. The tab for protecting critical areas would, of course, be considerably higher, climbing into the billions by some estimates.

The proposal comes at a propitious time. Both the populace and the pols are increasingly interested in the biodiversity crisis, as the worldwide loss of species is called. No doubt, its links to global climate change, as well as the new groundswell of concern about the environment, have helped to bring what some have considered an esoteric topic into the mainstream. Last session in Congress Representatives Claudine Schneider (R–RI) and James Scheuer (D–NY) introduced a bill, with some 87 co-sponsors, that would make the conservation of biological diversity a national priority and would allocate substantial funds for data management and biotic surveys, among other things. While unlikely to pass this year, the bill nonetheless signals the arrival of biodiversity on the national agenda.

In a move that signifies serious interest, the National Science Board, the policy body of the National Science Foundation (NSF),

T.E. Lovejoy

Slash and burn: *The fate of an increasing proportion of tropical rainforests.*

---

*The annual meeting of the American Institute of Biological Sciences, 14 to 18 August 1988.