

COMMENTARIES

A New Classification of the Living Birds of the World

ERNST MAYR¹

The following comments on the recently published classification of the living birds of the world (Sibley et al. 1988) are offered at the invitation of the editor of *The Auk*.

There is no other group of organisms in which the taxonomy at the species level is as mature and that is as uncertain at the level of the higher categories as the class Aves. Indeed, 30 years ago the renowned ornithologist Erwin Stresemann despaired that we would ever be able to understand the relationships of the higher taxa of birds: "In view of the continuing absence of trustworthy information on the relationship of the highest categories of birds to each others, it becomes strictly a matter of convention how to group them into orders. Science ends where comparative morphology, comparative physiology, comparative ethology have failed us after nearly 200 years of efforts. The rest is silence" (1959: 277).

Two subsequent developments showed Stresemann to have been wrong. First of all, a group of avian anatomists, including Bock, Homberger, Lanyon, Houde, Olson, Peters, Raikow, and several others, showed that the potential of morphological research was by no means exhausted. More importantly, Charles G. Sibley felt that molecular characters might supply information where the relative morphological uniformity of birds had failed to give useful evidence. With enormous enthusiasm and energy, he tried one method after the other, and refused to be discouraged when the results did not live up to his hopes. Even during this earlier phase he made interesting discoveries that have since been substantiated, such as that *Zeledonia* is a parulid warbler and not a thrush; but none of these methods produced results commensurate with the time and labor they cost.

When Sibley and his collaborators improved the DNA-DNA hybridization method, they finally struck gold. In spite of considerable difficulties in interpreting the results, it was soon evident that this method permitted a measure of the degree of similarity of two avian taxa. The work represents the first opportunity to develop a classification based on a single, quantitative measure. In order to translate the pairwise comparisons into a classification (which includes all but a few Madagascar and New Zealand families), Sibley et al. adopted the UPGMA method.

The amount of material which Sibley and his col-

laborators gathered in more than a dozen years is extraordinarily large; and it is evidently of high quality. In a recent paper that is quite critical of some of Sibley's interpretation, Sarich et al. (1989) state: "Sibley and Ahlquist have produced data of very high quality . . . and nothing we have written or said should be taken as detracting from the enduring value of those data, nor from the magnitude of their achievements." And they also agree that "DNA hybridization is by far the most cost-effective approach to the phylogenies of the organisms involved."

Sibley and Ahlquist have now summarized their findings (prior to the publication of a book) and have proposed a new classification of birds. This classification confirms to a large part what everybody had believed, but to an almost equal extent it departs from the standard avian classifications used by most authors. Most important, Sibley et al. propose a definite place in the system for a large number of genera and families, the position of which had been completely uncertain in the past.

The crucial questions, which every ornithologist will now ask, are, should we discard the currently adopted classifications and accept Sibley's proposals from beginning to end, or should we wait until his findings have been confirmed by one or several of the numerous other molecular methods that are now being tested by various investigators? A third possibility is that one might accept those proposals that seem to agree with our intuitions and play a waiting game with all the others. Can one give at least a tentative answer to these questions after an analysis of Sibley's methodology and taxonomic philosophy?

Certain of the proposals are so obviously correct that most likely they will be accepted immediately. That the New Guinea genus *Peltops* and the even more puzzling Bornean genus *Pityriasis* belong to the Cracticidae is one example. Shifting the drongos (Dicruridae) to the vicinity of the monarch flycatchers is another. Confirming the relationship of the wall-creeper (*Tichodromas*) with the nuthatches is satisfying, as is the relationship of the falcons to the Accipitridae (rather than the owls), of the hoatzin to the cuckoos, and the hummingbirds to the swifts. New proposals that are almost certainly going to be confirmed are the removal of the swallows and larks from their isolated position, separating the pardalotes from the Dicaeidae, associating *Grallina* with the Monarchidae, placing the Campephagidae near the Cracticidae assemblage, and placing the wrens near *Certhia* and *Polioptila*, to mention just a few. The two major subdivisions of the Oscines (not previously recog-

¹ Museum of Comparative Zoology, The Agassiz Museum, Harvard University, Cambridge, Massachusetts 02138 USA.

nized as such), now named Corvida and Passerida by Sibley and Ahlquist, are two rather neat packages that make a lot of sense. Removing the vireos from the neighborhood of the nine-primaried songbirds and attaching them to the corvid assemblage is confirmed by anatomical evidence, even though it creates a considerable zoogeographic puzzle. There are quite a few other cases where Sibley's firm validation of one of several controversial options will help considerably to stabilize the classification. Other proposals, however, seem to require independent testing before they can be made part of a stable classification of birds.

In this connection I want to emphasize a principle that is often overlooked by those who propose "provisional" classifications. A standard classification dictates the sequence in which species are listed in local faunas and in museum collections. Such a classification is the key to a vast system of information storage and retrieval. Every time this key is changed, it impedes information retrieval. If there is anything like a previously existing more or less widely accepted classification, this should not be experimented with until the need for a change is demonstrated clearly.

This leads to the question of whether Sibley's methodology and taxonomic philosophy guarantee that the proposed changes are in the category of "clearly proven" and justify the need for change in the existing classifications.

Sibley's methodology has been criticized in print and in informal discussions among ornithologists and molecular taxonomists. Only time and further research will show whether or not there is any validity to these criticisms. I do not have the competence to discuss whether the delta measure, chosen by Sibley and Ahlquist, produces valid results. I will not comment on the controversy concerning this point. However, I shall discuss some other points.

As Sarich et al. (1989) indicate, there is much trust in the quality of the raw data produced by Sibley's technique. Sarich and some other molecular biologists criticize the methodology of translating the raw data into a classification. I am not an expert in this area, but I have an intuition that the methods of Sibley et al. are not as poor as claimed by some of his critics. This faith is based on the fact that others have tested, with entirely different methods, at least one of Sibley's findings (the relationship of man with chimpanzee and gorilla) and have confirmed Sibley's conclusions. One has the feeling, however, that this method is most secure at low delta values (for reasons in part discussed below) and that it becomes increasingly less reliable, the more distant the relationship.

Because questions of the interpretation of the data are involved, it is time for Sibley to publish his raw data. He had every right to withhold them until he himself was able to present the first interpretation of these laboriously acquired data, but now that this has been done, the time has come to permit other in-

vestigators to see whether or not other methods of data analysis would lead to the same results. Or else Sibley himself should apply some of the other available distance methods to these data. There is no doubt that Sibley's raw data must be processed by several of the competing algorithms to see whether and where discrepancies might occur.

Sibley et al. constructed dendrograms with Sokal's method of average linkage (UPGMA). Cladists criticize this method for two reasons. First, it involves clustering, hence a loss of information, as is true for all methods that use distance data. But more importantly, it is based on the assumption of an equal rate of change in all lineages (if I understand correctly). There are other methods of converting distances into dendrograms, such as Wagner-Farris trees, that are immune to the latter criticism. But these cladistic methods also have their shortcomings, as admitted in the most recent writings of that school.

Sibley et al. base their conclusions on the amount of DNA which has remained unchanged in the evolution of two lineages since their divergence from the common ancestor. This method is vulnerable to two possible sources of error.

Assumption of an equal rate of DNA change in all lineages is the first potential source of error. When Sibley sent me his first paper on DNA hybridization, I urged him to drop the assumption that the rate of base pair change per unit of time is the same in all phyletic lineages. However, he retained this assumption for many years, in part because the UPGMA algorithm is based on this assumption, and also because he thought he had actual evidence for such equality. It was not until Roy Britten's discussion on rates (1986) and his own discovery of clearly different rates among nonpasserine birds that Sibley finally gave up this assumption, the more so when he found a tenfold rate difference between rodents and hominoid primates (Catzeflis et al. 1987). However, Sibley et al. have not yet fully faced up to the consequences of this new insight in their *Auk* classification. Although they state that there is a difference between taxa that have a slow sequence of generations and those with several broods in a single year, it is not evident from their classification which dendrograms "have been corrected . . . for these differences" and with what objective standards. Yet, there is a great deal of indirect evidence that there is a much more rapid turnover in base pair replacements in some lineages (regardless of rate of generation) than in others. I know of no concrete evidence for rate equality among the Passeres. The UPGMA method, of course, does not accommodate rate heterogeneity.

This is not a minor point, because the rate of change determines the location of the branching points. A rate of DNA change more rapid in one phyletic branch than in its sister branch, properly taken into consideration, might require a considerable shift in the branching point, hence a different classification (Fig.

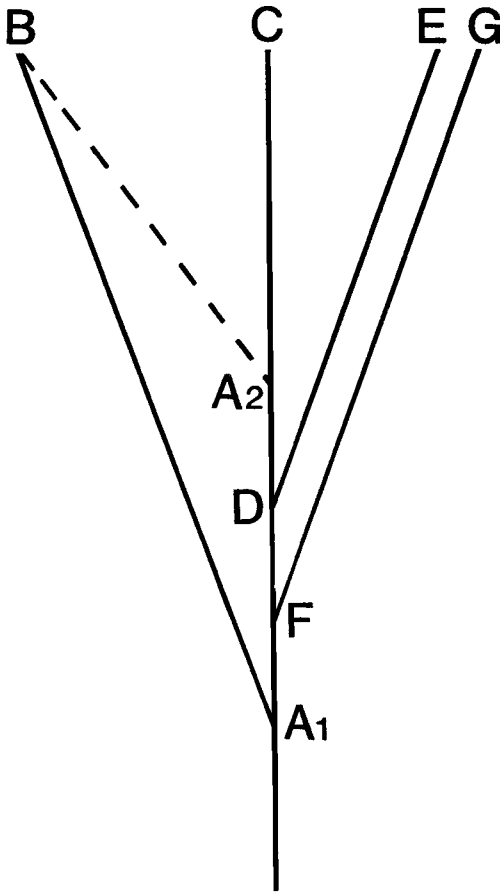


Fig. 1. *E* and *G* are most closely related to *C*, if all lineages evolve at the same rate. However, *B* is most closely related to *C* if the rate of A_2 -*B* is twice as fast as that of the other lineages.

1). There is evidence from other organisms that the rate of DNA turnover may even change over time within a single lineage.

Where the DNA results agree with the traditional classification, Sibley records this as evidence for the soundness of his molecular method. Where they do not agree, Sibley rather disingenuously takes it for granted that it must be the morphology-based results that are wrong.

The other potential source of error is that Sibley et al. assume that all changes of DNA are of equal taxonomic significance. Or that, owing to the large number of base pairs involved, all differences would compensate for each other. Until a few years ago, this was a sound argument, but this is no longer the case. It is now known that different parts of the DNA sequence may be of extremely different evolutionary significance. Kimura and others have demonstrated

that a considerable proportion of the changes are neutral or quasineutral. This means that they are without effect on the fitness of the bearer, and are thus neither selected for nor against. Yet, these neutral changes, precisely because they are neutral, provide investigators with important information, and that is the point of time where the ancestors split. This is based on the assumption that neutral substitutions in the DNA occur at a more or less standard rate throughout evolution. If so, the size of the difference between two taxa permits an estimate as to the length of time that they have evolved independently (molecular clock).

There is still the uncertainty of the proportion of "neutral" genes in the genome. Phenotypic changes, those visible changes dealt with in a classification, are effected by those genes that are *not* neutral. These are the genes that involve adaptive shifts, niche changes, developmental innovations, and all the characters by which a taxonomist distinguishes one taxon from another. Although these genes are in the minority, it is on their changes that a taxonomist should base his classification, at least in my view. For me, as a confirmed Darwinian, there is a world of difference between a genealogy and a classification consistent with genealogy (Mayr 1985).

Sibley et al.'s method does not permit them to discriminate between these two categories of DNA changes. Indeed, they make the silent but improbable assumption that both types of genes change at the same rate or, if there is a difference, that the evolutionarily important genes change at approximately the same rate in different lineages. Hence one can base a classification confidently on the overall delta values (which include both kinds of genes). Is this confidence justified?

Available fossil evidence shows that many phyletic lineages are evolutionarily stagnant, while others change rapidly, even though the neutral genes presumably change at similar rates in both types of lineages. Using overall delta values gives unrealistically high weight to the neutral genes. The result is the recognition of taxa that strike a traditional ornithologist as highly unbalanced. Examples are, on one hand, the lack of any subdivisions in such a highly diverse group as the parrots, and the reduction to the status of tribes in the single family Corvidae of 17 groups, most of which had previously been recognized as separate families. Being only interested in delta values, because this is the only thing their method can determine, Sibley et al. treat all DNA changes as equivalent, not asking to what extent they are composed of evolutionarily neutral changes and to what extent evolutionarily important changes.

What strikes me as significant are the many instances of a low correlation between Sibley et al.'s delta values and the amount of visible difference between taxa. Some reasonably uniform groups are split. For example, the Phasianidae through the recognition

of the New World Quails as a parvorder, the splitting of the woodhoopoes into different families and of the trogons into different subfamilies, the recognition of the three groups of kingfishes as three parvorders, and the large separation of the kinglets from the warblers. In other cases, rather different groups are brought much closer together than other evidence would suggest. For example, the Alcidae as subfamily of the Laridae, the great diversity of taxa included in his Ciconiiformes, placing the bowerbirds so near to the lyrebirds, putting starlings and mockingbirds in one family, and placing 17 tribes in the Corvidae. These are only a few examples that make one wonder as to the meaningfulness of the classifying criteria of Sibley et al.

I am worried about using the degree of pairing of the DNA of two species at a given temperature as the basic measure of similarity. How different will the nonpairing part be? By lowering the temperature, there will be more pairing, as Sibley et al. have explained. Yet, the part of the DNA that does not pair (let us say at 60°C), presumably includes the part of the DNA that has recently experienced important evolutionary changes. But is pairing vs. nonpairing a good measure of degree of evolutionary significance?

So much for the interpretation of the data. A few additional words must be said about Sibley's philosophy of classification. There are a number of basic principles that have been with us since the time of Darwin. One, which I have already mentioned, is that no classification should be abandoned until it is definitely falsified. Otherwise we would have an incessant turmoil in our information retrieval systems. The second very important point is that of two related taxa, the one believed to be closer to the ancestral condition should be listed in a linear sequence before the more derived ones. Sibley et al.'s sequences ignore this principle. They are not alone in this; this principle has also been ignored by others. Indeed avian classifications still suffer from traditional misconceptions. Earlier authors had wrongly assumed that the flightless ratites and penguins were "primitive" and represented an intermediate stage between reptiles and flying birds. Therefore they listed these flightless birds at the beginning of the avian system. We know now that these taxa are derived and highly specialized. In the sequence of taxa they should be placed after the taxa that reflect the ancestral condition. Hence, among the paleognaths the Tinamidae must be listed first, and the flightless ratites derived from flying paleognaths afterwards. This view is strengthened by the recent discovery of a number of early Tertiary paleognaths that were still able to fly. Because they occurred in Eurasia at a time when Africa and South America had been separated for at least 25-30 million years, it is quite conceivable that the Rheas and the tinamus reached South America from Euramerica via the late Cretaceous-Eocene Panama-

nian land connection rather than as old Gondwana elements. Even though no fossil paleognaths have so far been found in North America, they were found in Europe at a time when Europe and eastern North America were a single continent.

Penguins are another highly derived group of birds. As Fürbringer and others have shown, there is a great deal of anatomical similarity between the Tubinares and the penguins. But quite obviously the penguins are derived from the Tubinares, rather than the reverse, and should be listed after them. In Sibley et al.'s arrangement, the penguins precede the shearwater group, and loons are in-between: a rather unlikely sequence.

Oscines are another example. Sibley has greatly advanced our understanding of relationships within the Oscines. However, I question the validity of the sequence of the various groups within the Oscines. Sibley et al. recognize two major assemblages of families, the Corvida and the Passerida. This is a good working hypothesis. To have demonstrated the close relationship of many Australian genera and families, which up to now had been more or less "floating around" in the avian system, is certainly a most constructive step. However, all classifications have to be in the form of a linear sequence and we have to decide whether the Corvida or the Passerida should be listed first. Sibley suggests that it is very probable that the founding group of the Corvida arrived in Australia about 30 million years ago, and underwent its entire adaptive radiation on that continent. The origin of the Corvida is still an issue. From all zoogeographical evidence it is quite clear that the founder (or founders) of the Corvida came from Asia. This is the home of the only other parvorder of Oscines, the Passerida and their ancestors. There is no escape from the conclusion that the Passerida stock (of course not the existing families of Passerida) gave rise to the Corvida. This would be true even if India or Africa had been the stepping-stones to Australia rather than southeastern Asia. In a linear classification, the sequence therefore should be Passerida-Corvida. Fortunately this is the sequence which most ornithologists (except Wetmore and his followers) had accepted for the longest time. For it can hardly be questioned that the ravens (see Portmann's studies), the bower birds, the birds of paradise, and the lyrebirds are the most highly derived Oscines. They are all members of the Corvida. The Basel committee of the International Ornithological Congresses, consisting of J. Berlioz (Paris), G. C. A. Junge (Leiden), G. Dementiev (Moscow), R. E. Moreau (Oxford), F. Salomonsen (Copenhagen), and E. Stresemann (Berlin), decided that this was the sequence most widely adopted in faunal lists in Africa, Eurasia, India, and Australia, and voted unanimously to recommend its official acceptance. Therefore it was adopted in the volumes of the Peters' Check-list of Birds of the World. Even if there had been some Americans on the committee, the majority

vote still would have been for the Basel sequence. Sibley grew up with the Wetmore sequence and feels more comfortable with it. However, this is a matter of sentiment and not supported by the evidence.

It should be quite obvious from these comments that the new classification of birds by Sibley et al. is a most useful working hypothesis. Indeed the authors themselves have referred to it as a progress report. However, it would be a great mistake to consider it the last word. Sibley et al. have the enormous merit of providing an opportunity to analyze and test the position of every single order, family, or subfamily of birds, and thus to establish the basis for what might eventually be a generally acceptable classification of all birds. In addition to further morphological research and the study of fossils, there are numerous other molecular methods that can be employed to test the hypotheses. They will either confirm or falsify their proposals. A new era for the study of avian classification has begun. Even though one might perhaps claim that no other individual in the last 100 years has made as great a contribution to our knowledge of the relationship of birds as Sibley, it would be a complete misconception of the nature of science to believe that the work of any one scientist must be accepted as the last word in any area of science. Science advances, as Popper has so rightly said, by conjecture and refutation.

LITERATURE CITED

- BRITTEN, R. J. 1986. Rates of DNA sequence evolution differ between taxonomic groups. *Science* 231: 1393-1398.
- CATZEFELIS, F. M., F. H. SHELDON, J. E. AHLQUIST, & C. G. SIBLEY. 1987. DNA-DNA hybridization evidence of the rapid rate of murid rodent DNA evolution. *Mol. Biol. Evol.* 4: 242-253.
- MAYR, E. 1985. Darwin and the definition of phylogeny. *Syst. Zool.* 34: 97-98.
- SARICH, V. M., C. W. SCHMID, & J. MARKS. 1989. DNA hybridization as a guide to phylogenetics: a critical analysis. *Cladistics* 5: 3-32.
- SHELDON, F. H. 1987. Phylogeny of herons estimated from DNA-DNA hybridization data. *Auk* 104: 97-108.
- SIBLEY, C. G., J. E. AHLQUIST, & B. L. MONROE. 1988. A classification of the living birds of the world based on DNA-DNA hybridization studies. *Auk* 105: 409-423.
- STRESEMANN, E. 1959. The status of avian systematics and its unsolved problems. *Auk* 76: 269-280.

Received 20 January 1989, accepted 8 March 1989.

Response to E. Mayr

CHARLES G. SIBLEY¹

In his Commentary Ernst Mayr supports several aspects of the Sibley, Ahlquist, and Monroe classification (1988). Professor Mayr and I have exchanged letters about the classification and some of his earlier questions have been resolved, but his Commentary contains points on which we still disagree. My comments address some of these.

It is important to question the implication that departures from the "standard avian classifications used by most authors" is somehow wrong. There have been at least 50 different classifications of birds published since that of Linnaeus. Even a cursory study of past classifications will reveal that only a few were founded on more than one, or a few, morphological characters. Others were based on tradition, intuition, or a selection of previous ideas. Fürbringer (1888) wrote two large volumes on avian morphology, but his clas-

sification was not fully accepted by his contemporaries, including his friend Hans Gadow. Gadow (1892, 1893) alone analyzed a large number (40) of characters using a kind of numerical taxonomic system. Wetmore and Miller (1926) assembled an eclectic classification for which they took "Gadow's work as a basis and . . . incorporated in it various changes that have been made by later workers. . . . When doubt seems to attach to any suggestion we have followed the older classification." Wetmore (1930, and later versions to 1960) followed the same procedure when he produced his classification of the birds of the world. Thus, the Wetmore classification, in wide use for the past 63 years, is mainly the work of Gadow, nearly a century old.

Mayr and Amadon (1951) based their classification on various sources, and included "few changes . . . from the now well-established sequence of Wetmore (1934, followed by Peters)." Mayr and Amadon (1951), and the "Basel sequence" (Mayr and Greenway 1956) advocated the "crows last" arrangement for the oscines, but Stresemann (1934) followed Wetmore (1930) and placed the nine-primaried oscines last. When Mayr

¹ Tiburon Center for Environmental Studies, San Francisco State University, Box 855, Tiburon, California 94920 USA.