

OXFORD READERS

Evolution

Edited by Mark Ridley

Oxford · New York

OXFORD UNIVERSITY PRESS

1997

33 Can classification do without evolution?

In the past few years biology has variously suffered or enjoyed apparent attacks from within upon its central stronghold: the concept of evolution. To the careful observer it is obvious that most of the debates are concerned not with whether evolution actually took place, but simply with its mechanism, and the extent to which evolutionary development is guided specifically by natural selection, as Charles Darwin proposed. But one prominent group of biologists has been portrayed as having rejected evolution itself. They have been greeted in a triumphant broadsheet by the Institute for Creation Research, and reprimanded by the anonymous oracle of *Nature's* editorial pages. The root of the controversy is a new method of classifying animals known as transformed cladism.

To be sure, transformed cladism would not have had the impact it had if it had not been developed in such public and respected institutions as the British Museum (*Natural History*) in London, primarily by Colin Patterson, and at the American Museum of Natural History in New York, where its best-known protagonists are Gareth Nelson and Norman Plannick. It gained momentum, too, by being associated, in ways sometimes obvious and sometimes less so, with the philosophies of Karl Popper and Karl Marx. But as I read through the public debate I am left with the feeling that it has been helplessly pointed in the wrong direction. Clearly it is time to ask what transformed cladism actually is; why its practitioners have apparently rejected evolution from their approach to classification, which traditionally has been a stronghold of evolutionary concepts; and whether they are right to do so.

Transformed cladism cannot, I think, be understood except within the wider context of biological classification. People have always classified living things: the modern form of classification dates from the eighteenth-century taxonomist Carl Linnaeus, but earlier forms of classification go back to Aristotle. An accepted classification is essential, whatever its higher purposes, for communication; and for this reason alone biology could not do without it. But if the only purpose of classification were to communicate, it would be difficult to see what all the fuss is about: taxonomy would be an intellectually rather humble matter. Taxonomists could just observe which living things had which particular characters, such as bones, and use those characters to define groups. Thus, in this instance, animals with bones are approximately classed as vertebrates.

For better or worse, taxonomy is not as simple as that. The difficulty is that the groups the taxonomist defines depends upon what character he chooses to consider important. Backbonees do indeed define vertebrates—but if we take some other character (let it be the possession of eyes) we should define a

completely different group that included most (but not all) vertebrates, most insects, most crustaceans, some molluscs, and some other invertebrates. These animals do not form a normally recognised taxonomic group, but the question is, why don't they? Taxonomists recognise some groups, and not others, but what is the principle by which they do so? Taxonomy is not just a matter of defining groups but of choosing between possible ways of grouping. The choice (if it is to be convincing) must be supported by clear principles. Taxonomy has been forced to develop a philosophy to justify the recognition of some groups, and the rejection of others. The source of controversy is that different taxonomists have different philosophies and so produce different systems of classification.

In practice, modern taxonomists are divided into three main schools. One of these is called evolutionary taxonomy, and its best-known members are the American zoologists Ernst Mayr and George Gaylord Simpson. The second school is called numerical taxonomy, and includes the American, Robert Sokal, and Peter Sneath from Britain. The third school is cladism, whose leading protagonist was the German entomologist Willi Hennig, who died recently. Transformed cladism, the source of what has become public controversy, is an offshoot of Hennig's cladism.

All systems of classification in practice are hierarchical: animals are placed in groups, which in turn are gathered together into larger groups, and so on. But it is possible to arrive at a hierarchical classification in two main ways, and the three schools differ according to whether they use one way or the other, or a mixture of the two.

The first way to arrive at a hierarchical classification is by studying phenotype—that is, the physical appearance of the animal—and grouping the animals according to the degree of phenetic resemblance (the extent to which they look like one another). The second way is to represent phylogenetic relationships. In a phylogenetic system two animals that had a recent common ancestor would be classified closely together, like brothers and sisters who share a common parent; whereas two animals who shared only a long-deceased ancestor would be shown to get further apart, like second cousins, who have only great grand-parents in common.

Whether taxonomists choose to represent phenetic resemblance, or phylogenetic relationships, affects the classification. Thus, as Figure 33.1 shows, lizards and crocodiles should be grouped together—as reptiles—on phenetic grounds, while birds, which look quite different (except for their scaly legs) should be placed separately from both of them. But on phylogenetic grounds, birds and crocodiles should be placed together, with lizards in a separate group, because birds and crocodiles share a more recent common ancestor than did crocodiles and lizards.

We can now identify the nature of the three main schools of taxonomy. Evolutionary taxonomy uses a mixture of both phenetic resemblance and

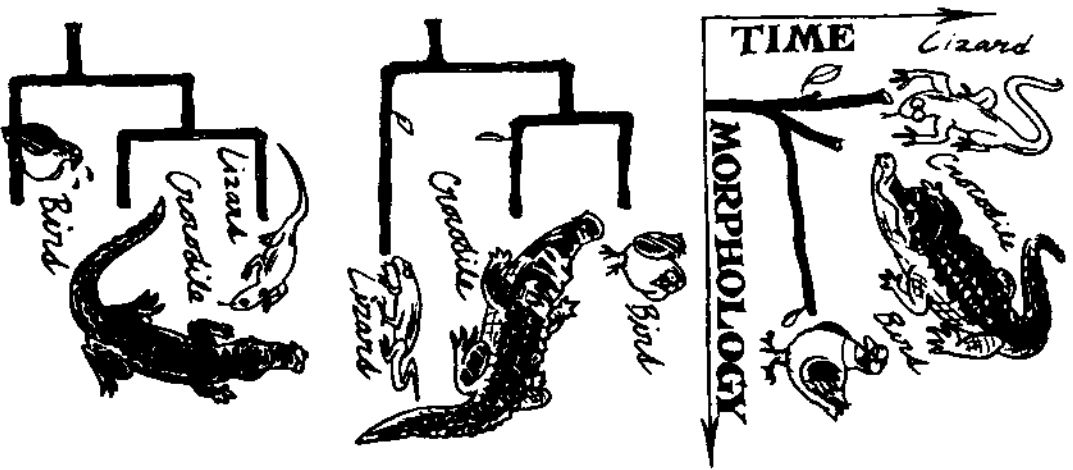


FIG. 33.1 How should crocodiles, lizards, and birds be classified? All the evidence suggests that birds and crocodiles are literally more closely related than crocodiles and lizards: that is, they shared a more recent common ancestor, as indicated in the top figure. Cladists, who classify animals strictly according to their evolutionary (phylogenetic) relationship, would therefore put birds with crocodiles, as in the middle figure. But numerical taxonomists are more interested in observable—phenetic—similarities: and would classify lizards and crocodiles together (in the class 'Reptilia') leaving birds on their own (as 'Aves').

phylogenetic relationships; in this particular case it prefers phenetic classification, and puts lizards with crocodiles, but in other instances it might prefer to emphasise phylogeny. This flexible, commonsensical approach has something to be said for it on practical grounds, but purists are obliged to subscribe to one of the other two schools. Thus, numerical taxonomy makes use only of phenetic resemblance; it would put crocodiles with lizards. Cladism uses only

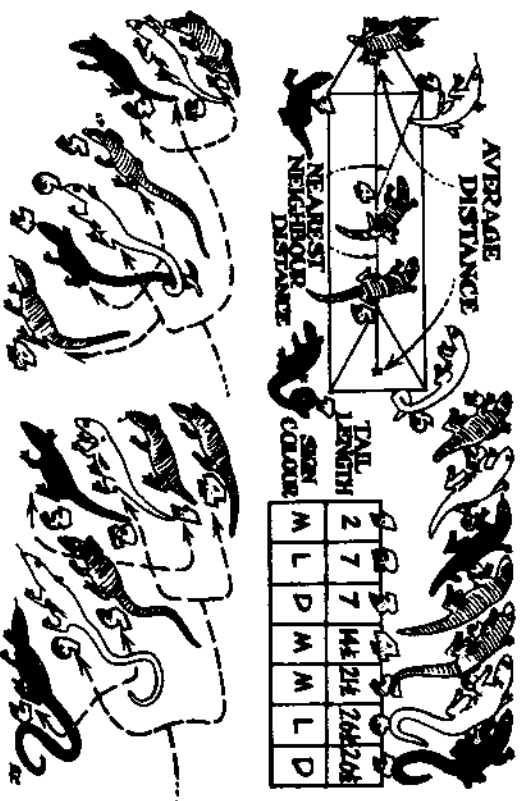


FIG. 33.2 The trouble with numerical taxonomy. Numerical taxonomy aims to take the subjectivity out of classification, but this is not really possible. In this simple case, involving seven hypothetical lizards, the hypothetical taxonomist is basing his classification on only two characters: skin colour (represented by M:medium, L:light, or D:dark) and tail length (either 2, 7, 14, or 26 units).

A two-dimensional graph (top left) grades the lizards from dark to light on the vertical axis, and from short-tailed to long-tailed on the horizontal axis.

If tail length is considered more significant than colour (itself a subjective decision) then the short-tailed types (1, 2, and 3) form one group, and the very long-tailed types (5, 6, and 7) form another, as in both the classificatory schemes shown. However, whether lizard 4 should be grouped with 5, 6, and 7 or with 1, 2, and 3 depends on which 'cluster statistic' the taxonomist chooses to adopt. He could average the dimensions of 1, 2, and 3 and of 5, 6, and 7—thus producing the points marked X, showing the two averages, in the graph. Lizard 4 is closer to the average of 1, 2, and 3 than to the average of 5, 6, and 7, and so the taxonomist could group 4 with 1, 2, and 3. But if he adopts the 'nearest neighbour' criterion then he must group 4 with 5, 6, and 7 because 4 is nearer to 5 than to 1, 2, or 3.

Even if the taxonomist were to compare hundreds of characters he could not eliminate the sources of subjectivity. There is an infinite number of cluster statistics.

phylogenetic relationship: it would put crocodiles with birds. In order to understand the deep principles of taxonomy we should concentrate on the two purist schools: the numerical, phenetic approach; and the cladistic, phylogenetic method.

In Figure 33.1 it was assumed that the phenetic and phylogenetic relations of birds, lizards, and crocodiles were known. But how were they first established? The rough answer is that this is done by looking at characteristics the animals have in common—that is, 'shared characters'. Phenetic similarity is established by looking at all characters. Crocodiles look more like lizards than birds because birds have undergone large changes in adapting for flight; indeed, if you look at almost any character, crocodiles look more like lizards than birds. Hence the phenetic classification.

Phylogenetic relationships are also inferred from observable characters, but that does not mean that phylogenetic classification is simply a form of phenetic classification. Phylogenetic classification concentrates on certain kinds of characters which (so zoologists believe) are the best indicators of a phylogenetic relationship. All characters can be divided into *phylogenetic* characters that are shared probably because of phylogenetic relationship, and *non-phylogenetic* ones that probably do not effect a literal phylogenetic relationship. The backbone, for instance, is thought to be a phylogenetic character in vertebrates: something that each vertebrate has because it has derived it ultimately from a common ancestor. A green external coloration is not thought to be a phylogenetic character: greenness crops up, conferred in all sorts of ways, throughout the animal kingdom. A jointed, hard exoskeleton is thought (by many) to be a phylogenetic character in arthropods such as spiders and insects; but the carnivorous habit is not. Lactation is probably a phylogenetic character in mammals; but monogamy probably is not. In practice, phylogenetic (cladistic) classification uses only phylogenetic characters to define groups, but phenetic (numerical) classification ignores the distinction between phylogenetic and non-phylogenetic characters and mixes them when defining its groups.

Such is the abstract form of the two kinds of classification; they differ in the characters that they use. But that is not all the difference. Each school possesses a philosophy, a justification of why its own, preferred kind of characters should be used rather than those of the other schools. The question of which is the best school has to be settled at the level of philosophy; for it is here that the schools proclaim their rival merits. Let us, therefore, consider the reasons that have been proposed in favour of numerical taxonomy and cladism. Transformed cladism, as we shall see, suffers from the same principal defect as numerical taxonomy, so this discussion is directly relevant to our main theme.

The advantage that numerical taxonomy boasts for itself is its objectivity. The techniques used by other schools to distinguish phylogenetic from

non-phylogenetic characters are (the numerical taxonomist believes) subjective, impractical, woolly. He would avoid these defects by ignoring the distinction altogether. He is then left with that other difficulty that we have met, which is that emphasis on different characters produces different results. The difficulty *appears* to arise only when single characters are used to define groups: such that stressing the importance of backbones produces one kind of grouping whereas emphasis on the possession of eyes produces a different classification. Numerical taxonomists thought they had a solution to this problem. They would not base their classification on single characters, but would measure a larger number of characters, and average them overall. The averaging is in practice performed (in a computer) by a 'cluster statistic'. The first objective is to define the 'distance' between species, where 'distance' means the average of the differences between one species and another, in the dimensions of a vast number of characters.

In theory, the whole process is perfectly repeatable: any taxonomist, if he measures enough characters, would arrive at the same estimation of the distance between species, and hence should end up with the same classification. Subjectivity had at last been eliminated from taxonomy. Or as the numerical taxonomists believed—for by the end of the 1970s all their boasted objectivity had been called into doubt. A splendidly destructive paper by the Australian taxonomist L. A. S. Johnson identified a number of sources of subjectivity that could not be eliminated from numerical taxonomy. One of these sources is illustrated in the Box, and is the most fundamental of all. It is that there is more than one cluster statistic, and each different way is liable to produce different results.

In fact there are an infinite number of possible cluster statistics. For instance, in the simplified example given in Fig. 33.2, the hypothetical numerical taxonomist has calculated the distance between seven species of lizard on the basis of just two characters, tail length and colour. In that example, he has decided that both characters are of equal importance. But in practice he may decide that one character is more important than the other, and that he should give it a higher 'weight' in the classification. Then, in calculating the distances, he would not simply average the measurements, but would allow the more important characters more weight. Without weighting, the average distance is just the difference in tail length + difference in skin colour / 2. If he wanted to give skin colour twice the weight, he would calculate $(2 \times \text{skin colour difference} + 1 \times \text{tail length difference}) / 3$. The important point is that he could substitute any numbers into this formula, which makes for infinite possibilities.

Once we realise that there is an infinity of cluster statistics and that no one statistic forms clusters that are more 'real' or more 'natural' than any other, then the pretensions of numerical taxonomy to objectivity crumble away. The numerical taxonomist can, of course, decide which cluster statistic he is going to use; but his choice, when he makes it, will be completely subjective.

Thus, however many characters he measures, however objectively, the numerical taxonomist must in the end *choose* the best way to analyse his measurements, but to make this choice, he needs to be able to refer to some higher criterion, and, unfortunately, no such 'higher' measure exists. Numerical taxonomists do sometimes pretend that they have such a criterion in what they call 'overall morphological similarity'; but the truth is that overall morphological similarity does not exist independently of the cluster statistics that are supposed to measure it.

So much for phenetic, numerical taxonomy. How does cladism justify itself? To answer that question we must turn to the first chapter of the great work of the school, *Phylogenetic systematics* (1966), by Willi Hennig. He there notices our two possible bases for a hierarchical classification, phenetic and phylogenetic similarity, and criticises the phenetic criterion by much the same argument as we have just been through. If phenetic similarity will not do, what about phylogeny?

Here we do have a possible firm basis for a classificatory hierarchy. Unlike overall morphological similarity, the phylogenetic hierarchy definitely does exist in nature. It really is true that two species either do, or do not, share a more recent ancestor with each other than with any other species. If we can only devise techniques to discover these phylogenetic relations, we should possess the ultimate taxonomy: both a firm philosophy and a set of practical techniques.

His system firmly founded, Hennig set about finding some techniques. His main contribution here was to develop and formalise methods that already existed, rather than to invent new ones. He realised that the best method of showing that two species share a recent common ancestor is to show that they share what he called evolutionarily *derived* characters.

All characters can be divided into those that are evolutionarily primitive and those that are evolutionarily derived. The terms are relative to the groups under consideration. The backbone of a crocodile is derived in so far as it is absent from invertebrates, but it is primitive when compared with its presence in some other vertebrate such as a frog. The fact that a crocodile and a frog share a backbone is not evidence that they share a *recent* common ancestor: the possession of a backbone is primitive in vertebrates, and we need assume only that crocodiles and frogs shared an ancestor from the dim distant past. But the fact that the backbones of armadillos, sloths, and anteaters all possess a peculiar extra articulation is evidence that they do share a (relatively) recent common ancestor: the extra articulation is derived within vertebrates.

The cladist, therefore, clusters species by their shared derived characters. The use of derived characters is, of course, a cluster statistic—one that weights primitive characters as 0, and derived characters as 1. The point, however, is that Hennig provided a justification for his particular cluster

statistic by arguing that there really is a phylogenetic hierarchy: animals do have ancestors which they may or may not share with other animals. He had, in a sense, solved the numerical taxonomist's problem.

If we accept that phylogenetic relations are discovered by shared derived characters, the next problem is to devise techniques to distinguish primitive from derived character states. These are the cladistic techniques. We do not have space to discuss these techniques in detail, but they include such methods as 'outgroup comparison', which compares characters between related species, and the 'embryological criterion' which looks at the order in which characters appear as the animal develops; and there are other techniques. All we need to know here, however, is that practical cladistic techniques do exist, and that primitive and derived character states can be distinguished, albeit tentatively and imperfectly. Once they have been distinguished, the species can be grouped by their successively more derived character states, into some reasonable estimate of the phylogenetic hierarchy. Cladism is a workable system.

Evolution, it should now be clear, is crucially important to cladism. It underwrites its entire philosophy. It guarantees the self-justification of cladism against the phenetic schools. If evolution were not true, the coherence of cladism would be lost. Given that evolution is true, cladism can (I believe) claim to be the best system of taxonomy yet developed. It alone has found, and in the phylogenetic hierarchy, an objective basis for classification. You will not, of course, find the original, Hennigian cladists declaring that evolution is not necessary for classification; they are well aware that if evolution were thrown out, so too would be cladism. But, although it may now seem strange, certain cladists have claimed that evolution is an unnecessary assumption in classification. Which brings us to 'transformed cladism'.

If we define a school of classification as a combination of a set of techniques and a philosophy to justify them, then transformed cladism would count not just as an offshoot but as a fourth school, different from and in competition with the other three. Its techniques have simply taken from the original, Hennigian cladism, though its practitioners may prefer to substitute terms, such as 'synapomorphy' and 'symplesiomorphy', which barbarously disguise their evolutionary meaning. But terms are unimportant; the important point is that the transformed cladists continue to use the cladistic techniques, rather than the undiscriminating character choice of phenetic taxonomy.

The philosophy of transformed cladism is not so easy to find. It appears to be mainly the realisation that cladistic techniques can be applied without assuming that evolution has taken place. This is certainly true: any set of practical techniques can be reduced to a set of operations that can be performed by an unthinking automaton. Simply to perform cladism you do not have to assume evolution. You do not have to assume anything at all. But evolution was not needed, in Hennig's system, to operate its techniques. It was

needed to justify them. Transformed cladism therefore is the separation of Hennig's techniques from the philosophy that justified them. Its practitioners are asserting that they can do cladism without evolution, and what they can do, they will.

The rejection of a philosophy, however, is not itself a philosophy; and without a philosophy transformed cladism is very precariously positioned. It possesses a cluster statistic and has, by a subjective decision, bound itself over to operate it. But the particular cluster statistic adopted in transformed cladism does not even have the merit of being easy to use; you have first to distinguish whether a character is primitive or derived, and then you can use the character if you decide it is derived, but must exclude it if it is primitive. What a business! It could be rejected on grounds of impracticality alone: why go to all that trouble if you are not trying to reconstruct a phylogenetic hierarchy? If evolution is unnecessary, so too is transformed cladism.

We could leave transformed cladism there, as an incoherent school of taxonomy, lacking any self-justification, the only reason for its existence being the subjective agreement of its practitioners. But when I said that transformed cladism lacked a philosophy, I was referring only to its published pronouncements. Reading between the lines, and listening in the lecture halls, I have detected two (as they might be called) candidate philosophies. Neither has been presented as a philosophy; but, as it will appear, both could in theory supply that need. The first states that cladistic techniques have to be used because they are the only techniques capable of defining groups. The second maintains that evolution should not be assumed in classification, because if evolution is not assumed *a priori*, then the classification itself may later be used as a test of the truth of evolution. The first is a verbal trick, the second a more complex mistake; but let us complete our work of demolition by proving it.

Take the first philosophy first. The argument is that groups that are not defined cladistically are not defined at all. The group 'fish', for example, cannot be defined by shared derived characters; fishes can be defined only by a primitive character, the possession of fins. Now let us quote Norman Platnick, of the American Museum of Natural History: 'If we form a group Pisces, we have based it not on a character, but on the absence of a character. The group Pisces includes those organisms with fins that also happen to lack modified fins (limbs). Such use of the absence of a character is one of the hallmarks of an artificial group.' If Platnick's argument were valid we should be forced to use cladistic techniques not (as Hennig believed) to discover phylogenies, but because no other techniques were capable of defining groups. But his argument is not valid. It merely hypostatizes cladism as classification. In front of that word 'character' in the quotation above, the word 'cladistic' should be inserted. Pisces can easily be defined by a positive

character—fins. Fins are a primitive character state, but that does not stop taxonomists who do not wish to be cladists from using them to define a group. Primitive characters may be absent from the cladistic system, but they are not absent from animals. They can be recognised, and we can, if we choose, use them to define groups. Platnick's attempt to prove otherwise is a verbal trick.

The second candidate philosophy has a superficial validity. It states that it would be circular to try to prove evolution from classification, if evolution had been assumed in the classification to begin with. Transformed cladism therefore claims as its justification no less a purpose than the proof of evolution. We must not assume evolution in classification, in order that evolution itself can be tested.

I have three objections, each of them lethal, and taken together a threefold overkill. The first is that, at all events, we do not need classification to prove evolution. Other arguments can supply that need. There is artificial selection, geographical variation, and uniformitarianism; there is the fossil record; there are universal homologies, such as the genetic code. Between them they add up to a sufficient proof of evolution without classification.

Classification, however, could in theory supply another proof; but the second objection to transformal cladism is that this philosophy does not justify transformal cladism. It would (if valid) be a reason for keeping evolution out of classification; but it is not a reason for using cladistic techniques, rather than any other method of classification.

The third objection is that to assume evolution in classification, and to prove evolution by classification, is not a circular argument at all. The relation between evolution and classification is not a vicious circle, but one of (as it is called) 'reciprocal illumination' or 'successive approximation'. Successive approximation is really only the route by which all scientific theories are developed. The scientist first tests a hypothesis on a small scale, and then, if the test is successful, uses the hypothesis as an assumption in a further test. If that further test is successful, it is even more likely that the original hypothesis is true. In classification, one might test whether evolution has taken place in some small group, such as the Siphonaptera (fleas), and, once that was confirmed, use the fact as an assumption in classifying all other living things. In practice things are not as simple as that; there is not a single first stage of test, and a second one of assumption, but a continual feedback from the results of tests to re-illuminate earlier knowledge. Such is the process of reciprocal illumination. It is not argument in a circle. So we can, if we want to, assume evolution in classification.

We should want to. The best-known systems of classification all use evolution as an assumption, and the phylogenetic hierarchy to provide their philosophical coherence. The theory of evolution, too, is the source of their techniques. Without evolution, we are left with no sound basis for a classifi-

cation, but with an infinity of possible systems, and a subjective choice among them. Cladism has been transformed, but, if it is to survive, it must be reformed back to the evolutionary philosophy of Hennig.

[*New Scientist*, 100 (Dec. 1983), 647-51.]