Proc. Natl. Acad. Sci. USA
Vol. 89, pp. 2930–2934, April 1992
Evolution

On the nature of global classification

(domain/systematics/molecular evolution)

MARK L. WHEELIS*†, OTTO KANDLER‡, AND CARL R. WOESE§

*Department of Microbiology, University of California, Davis, CA 95616; ‡Botanisches Institut der Universität München, Menzinger Strasse 67, 8000 Munich 19, Germany; and §Department of Microbiology, University of Illinois, 131 Burrill Hall, Urbana, IL 61801.

Contributed by Carl R. Woese, January 3, 1992

ABSTRACT **Molecular sequencing technology has brought biology into the era of global (universal) classification. Methodologically and philosophically, global classification differs significantly from traditional, local classification. The need for uniformity requires that higher level taxa be defined on the molecular level in terms of universally homologous functions. A global classification should reflect both principal dimensions of the evolutionary process: genealogical relationship and quality and extent of divergence within a group. The ultimate purpose of a global classification is not simply information storage and retrieval; such a system should also function as an heuristic representation of the evolutionary paradigm that exerts a directing influence on the course of biology. The global system envisioned allows paraphyletic taxa. To retain maximal phylogenetic information in these cases, minor notational amendments in existing taxonomic conventions should be adopted.**

> The time will come I believe, though I shall not live to see it, when we shall have very fairly true genealogical trees of each great kingdom of nature. Charles Darwin (ref. 1, p. 456)

The capacity to sequence macromolecules has revolutionized (and revitalized) the study of evolution. As a result, that distant milestone envisioned by Darwin (1) has been reached—on the way to an even more distant one. Not only can genealogical trees be inferred for all the "great kingdoms"—freeing evolutionary study from the confines of multicellular eukaryotes—but the kingdoms, in turn, can now be joined to form the (rooted) tree of all life. This tree provides the basis for a global classification that will unify biology and help to deepen our understanding of it to an extent that Darwin could not have imagined.

We outlined such a global system in 1990 (2), based on the phylogenetic tree shown in Fig. 1. The system partitioned all life in the first instance into three domains (a new taxon of highest rank)—the Archaea, the Bacteria, and the Eucarya (corresponding to the three major branchings in Fig. 1)—within each of which would be two or more kingdoms (2). The system would replace the two conventionally accepted speculative and phylogenetically inappropriate global system(s), the so-called "five kingdom" scheme (5) and the prokaryote–eukaryote dichotomy (6, 7). However, both of these conventional systems have recently been defended—the former by Margulis and Guerrero (8) and the latter by Mayr (9, 10). Although they accept the phylogeny of Fig. 1, these authors disagree with us (and with each other) as to how a classification should be developed from it. This is not unexpected. Global classification presents issues, unusual issues, to which these traditional approaches are not attuned. Here we address this special problem of global classification.

In that classifications are arbitrary human constructs, there is no one "correct" classification. Nevertheless, within any defined context there may be a most useful one. The nature of a classification derives from the nature of the entities classified and the intended function(s) of that classification. It is a truism that every organism reflects its evolutionary history and, in an important sense, is that history. Consequently, any comprehensive explanation of a biological phenomenon must incorporate evolutionary understanding, which in many cases becomes the essence of that explanation. This perspective must underlie biological classification as well.

Biological classifications are commonly seen as effective information storage/retrieval systems. They should also be structured so as to facilitate deeper, more comprehensive understanding. The most important function of a biological classification, however, may be at a yet higher level—and this applies especially to a global classification—in helping to define the overarching concept of biology, pedagogically as well as experimentally. A biological classification should exert a directing influence on the course of biology. In our view then, a global classification is an heuristic representation of the evolutionary paradigm, one that embodies that paradigm to the fullest possible extent.

The Classical View of Systematics

Biological classification has dual historical roots, the one a pre-Darwinian (phenetic) approach and the other a Darwinian (evolutionary) approach. Both produce what have been termed natural systems—but natural in totally different senses. The older, phenetic classification reflects our innate (natural) capacity and need to organize phenomena—organisms in the present context—into types (classes) according to perceived similarity. Such classes in effect are "theories" about our world; all entities in a class are expected to look and behave according to an understood class definition (which obviously changes with experience). Linnean classification represents a scientific codification of this organizing proclivity. Darwinian classification, on the other hand, reflects the evolutionary process and thus centers about the natural relationships among organisms.

While the two modes of classification are grounded in basically unrelated perspectives, they tend to converge in practice. This, as Darwin understood, is because similarity among organisms is fundamentally the result of common ancestry; phenetic grouping, based on similarity, should therefore equate to grouping based on genealogical relationship (11). Unfortunately, evolution does not always conform to this simple equation. The reptiles, for example, are a grouping that genealogically includes birds and mammals, both of which are excluded from the corresponding grouping phenetically defined. Thus, what could be called the central problem of systematics in this post-Darwinian era is how to treat phenetically defined groupings that are genealogically

---

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

†To whom reprint requests should be addressed.

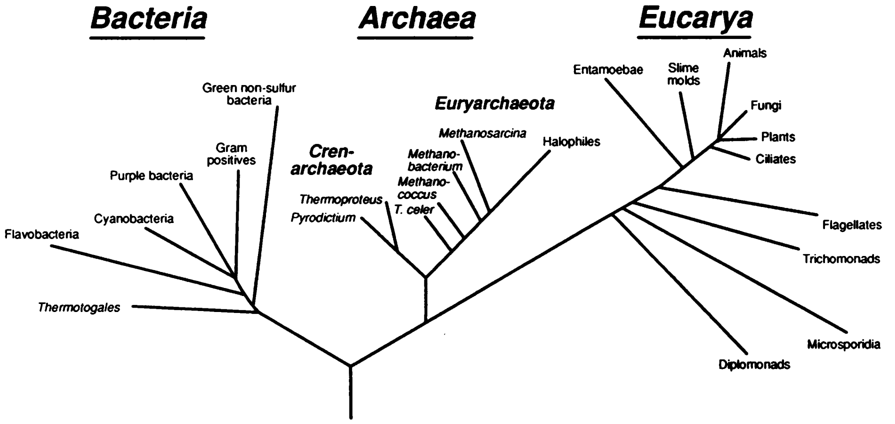

FIG. 1. Rooted universal phylogenetic tree. The figure is adapted from the corresponding figure in ref. 2, with the eukaryal branchings restructured on the basis of the more recent findings of Sogin (31). The root of the tree was established by using the Dayhoff strategy—i.e., comparing pairs of paralogous genes, whose common ancestor predates the most recent universal ancestor (3, 4).

incomplete (paraphyletic¶). This trying issue has resulted in a schism among systematists.

At one extreme, the pheneticists handle the central problem by ignoring genealogy, either by assuming that monophyletic groups are the automatic consequence of their methods or by divorcing classification in principle from genealogical analysis; for instance, according to McNeill

> The diagnostic feature of phenetic classification is its purpose, which is not to reflect evolutionary pathways, although it may do so, but to provide a convenient framework for accommodating the diversity of plants and animals. The quality of this framework is to be judged not by whether or not it matches some supposed or actual phylogenetic sequences, but whether it is . . . an acceptable simplification which best describes the distribution among organisms of as many features as possible. (ref. 12, p. 479)

The problem with phenetically defined taxa has been and remains that they are all too often polyphyletic, making their scientific value limited at best. We microbiologists are acutely aware of this problem, since in our field many years of competing phenetic systems created complete havoc, which is only now being cleared up by the application of phylogenetic analysis.

At the other extreme of the classificatory spectrum are the cladists, who cope with the central problem by the contrary simplification, concerning themselves solely with genealogies, using a complex and rigorous scheme of analysis (13). Cladists typically insist that every taxon (at least above the species level) be holophyletic. Cladism, commendably, has brought rigor into phylogenetic analysis and codified principles of phylogenetic inference. Hennig (the school's founder) emphasized (13) that only synapomorphies, or shared derived traits (that is, traits that have changed from a prior ancestral state), provide evidence for a specific relationship; shared ancestral traits (symplesiomorphies) do not. However, cladistic classifications can be counterintuitive and awkward—see Mayr (14) and Cronquist (15), for example, for critiques; Donoghue and Cantino (16) for defense.

What is variously called evolutionary taxonomy, gradism, or Darwinian classification deals with the central problem eclectically, mixing phenetic and cladistic methodologies. Mayr (14) and his followers generate an initial phenetic clustering that is further refined by a cladistic (genealogical) culling, which removes unrelated species. Ashlock (18), on the other hand, starts from a genealogically defined grouping and eliminates highly diverged lineages. Gradistic taxa are monophyletic but can be either paraphyletic (incomplete) or holophyletic (complete). Proponents of this combined approach justify paraphyletic taxa on the grounds that it is absurd to combine into a single taxon organisms of a different "grade." [Grade is Julian Huxley's term for a group of organisms united by a "common level" of biological organization, as distinct from a group united by common descent, a clade (19).] To gradists, as distinct from cladists, derived traits unique to a lineage (autapomorphies)—which are indicators of grade—are important in classification, as are the shared ancestral traits (symplesiomorphies).

It is worth reminding ourselves at this point of the essential features of a "natural system" within the Darwinian paradigm. Darwin's perception here remains cogent today:

> I believe that the arrangement of the groups within each class, in due subordination and relation to the other groups, must be strictly genealogical in order to be natural; but that the amount of difference in the several branches or groups, though allied in the same degree in blood to their common progenitor, may differ greatly, being due to the different degrees of modification which they have undergone; and this is expressed by the forms being ranked under different genera, families, sections, or orders. (ref. 11, p. 420)

From the above, as well as from numerous other similar statements (1, 11), it is clear that Darwin felt not only that a natural system must basically reflect organismal genealogies but that it must also express something else—the "degree of modification" lineages have undergone (today generally equated to divergence or grade). The above quote (and the phylogenetic diagram Darwin used to illustrate his discussion) shows that he believed divergence to be accompanied by diversification (radiation). Hence, diverged lineages require a series of new higher taxa to accommodate their burgeonings, whereas nondiverged lineages can remain in the ancestral taxon (ref. 11, chapt. IV). Because Darwin apparently believed that divergence of phenotype accurately mirrors genealogy, he appears never to have grappled with the issue that confronts us: instances in which phenetic and genealogical analyses lead to different conclusions. Nevertheless, he clearly recognized the evolutionary process to have two distinct characteristics, two principal dimensions—genealogy (branching order) and degree of modification (divergence)—and that both should be expressed in a natural system. We find great merit in this perspective, for it effectively says that the proper classification is one that (within its extremely limited framework) is the most informative.

¶We will necessarily use the somewhat arcane terms paraphyletic, monophyletic, holophyletic, and polyphyletic in this discussion. As used here, monophyletic describes a taxon defined in such a way that all ancestors of the group's members as far back as and including the most recent common ancestor of the whole, are encompassed by the definition. When all descendants of that common ancestor are thereby included, the taxon is then complete, or holophyletic. Alternatively, when some of the common ancestor's descendants are not included within the group, the taxon is incomplete, or paraphyletic. A polyphyletic taxon is one that excludes the most recent common ancestor of the whole group and/or ancestor(s) of some of its members.

Microbial Systematics and the Molecular Approach

Microbial systematists long ago tried and failed to produce a natural microbial system based on classical systematic concepts (20, 21). By and large, microbial morphologies are too simple and microbial physiologies and biochemistries are too (unpredictably) variable to serve "as a sound guide for the development of a 'natural' system of classification" (22). From these failed attempts, however, a valuable perspective on classical systematics emerges.

The concept of evolutionary grade figured prominently, if often only implicitly, in the early attempts to develop a natural microbial system. It underlies the subdivision of bacteria into three all-encompassing groups on the basis of morphological complexity and uniqueness—the classes Eubacteriae, Myxobacteriae, and Spirochaetae (21). It underlies various proposed microbial evolutionary progressions (20); for example, that leading from streptococci through intermediate forms to actinomycetes, or from various non-spore formers to their spore forming relatives, or from the cytophagas to sporocytophaga to the myxobacteria (23). The only effect these high level gradistic taxa and postulated gradistic progressions ["levels of biological improvement" is Huxley's phrase (19)] had was to exacerbate the confusion that already surrounded microbial phylogeny (24).

Microbiologists today understand that microbial genealogies are determinable only on the molecular level (24). This is not to say that organismal traits are of no value in defining microbial taxa. They are indeed but only in the context of, as convincing confirmatory indicators of, prior groupings determined by using molecular characters. Before considering the implications of this in global systematics, the contrast between evolution as seen on the gross classical level and that seen on the molecular level must be appreciated.

Molecular evolution is foreign to the classical systematist in at least two important respects, the nature of the characters that define the organism and the nature of variation on the two levels. As to the first, on the molecular level there exists an ultimate (elementary) character, the nucleotide (or, in the case of proteins, the amino acid). The elementary characters are few in kind, discrete, precisely defined, and quantifiable. The molecular structures built from them—i.e., nucleic acids and proteins—are therefore digital in nature and are typically hundreds to thousands or more residues long and so define an immense phase space (space of possible sequences), so vast that extensive similarity can effectively never be the result of convergence. Because of all this, (*i*) homology for any gene or protein is readily recognized (it is based on the correspondence of hundreds of homologous elementary units in each case), (*ii*) negative character traits are relatively rare (functionally altered homologs of a given gene can usually be detected), and (*iii*) the universal cellular functions—which when their sequences cannot be taken into account are phylogenetically meaningless plesiomorphic characters—can serve as reliable genealogical indicators (see below).

Contrast this to the world of the classical evolutionist, where (organismal or cytological) characters are often defined in imprecise, qualitative ways, and relationships among them are often uncertain; where character weighting becomes a subjective and contentious issue; where homology is easily mistaken or goes unrecognized; where convergence is relatively common; where negative character states are relatively frequent and uninterpretable; and where universal functions have no classificatory significance. The entire notion of a character, a trait, is different on the two levels.

The second respect in which molecular evolution is alien to the classical systematist concerns neutral changes and the so-called evolutionary clock (25, 26). The molecular sequence corresponding to a given function in one group of organisms is almost always different (although related) to that found in any other group of organisms. Most (although not all) biologists interpret this fact to mean that these differences by and large have no functional significance; they are selectively neutral. Indeed it is hard not to come to such a conclusion when one sees, for example, that the structural elements in a rRNA that change most frequently and most drastically appear to be those having the least to do with the molecule's function; in these evolutionarily volatile areas of rRNA, one even finds minor variations (say a different base pair at some position in the stem of a helix, or a nucleotide insertion) among the different copies of the molecule that exist in the same cell (encoded by different cistrons). These seemingly neutral changes constitute the vast majority of evolutionary events encountered on the molecular level. This stands in sharp contrast to the picture at the phenotypic (classical) level, where almost all change is typically seen as resulting from selective advantage.

Neutral or not, the vast array of changes that occur on the molecular level and the enormous evolutionary phase space in which they occur strongly imply that these changes tend to become fixed in a quasi-random fashion over time, which allows the biologist to speak of an evolutionary clock (25, 26)—analogous to a clock based on radioactive decay events. However, different clocks (that is, different molecular types) within the same species may run at different rates; and the same clock in different lineages may run at a different rate (implying that the rate of a given clock in a given lineage could vary at different stages in a lineage's evolution). This variable rate phenomenon makes phylogenetic analysis of molecular sequences more difficult and less informative than otherwise. Consequently, the amount of sequence divergence cannot be equated to time *per se*, which in turn means that the root of an inferred phylogenetic tree cannot be determined (unless outgroup sequences are available). Nevertheless, phylogenetic relationships are directly measurable by comparison of (an aligned set of homologous) molecular sequences: branching order can be rather precisely inferred (since it does not depend on a constant clock rate), and extent of divergence can be quantified in relative terms. The immediate product of such a molecular comparison is, therefore, a phylogenetic tree of the molecule being compared, in which branch lengths are proportional not to time but to the number of changes in sequence (i.e., rate × time).

Toward a Global System

It is obvious that (broad scale) phylogenetic inference will henceforth be based primarily on molecular measure, if for no other reason than that most major taxa are microbial. A molecularly derived phylogeny and a whole organism genealogy should in principle be equivalent (in almost all cases), for the phenotype derives almost exclusively from the genotype. However, when a molecular phylogeny is based on a single molecular species, as is now generally the case, one can question whether that particular molecule is representative of the genome as a whole.

The most serious potential source of difficulty here is lateral gene transfer. It is known to occur among widely separated species of Bacteria, as well as between Bacteria and Eucarya, through endosymbiotic associations at the very least. The problem this presents is trivial only if the extent of transfer is trivial. In the extreme, lateral gene transfer could produce organisms whose genomes are so chimeric that the very notion of an organismic phylogeny becomes meaningless. Fortunately for the systematist, gene transfer across broad taxonomic boundaries does not seem to have been that extensive (except perhaps during the very early stages of evolution). This point Margulis and Guerrero (8) do not seem to appreciate. They use the fact that endosymbioses have introduced bacterial genes into the eukaryotic genome ostensibly to vitiate our global system:

> Eukaryotes originated from the symbiosis of different prokaryotic microbes . . . [Thus,] today's eukaryotes carry more than one type of genome . . . There is no reason, for instance, why ribosomal gene sequences from the nucleus of a plant cell should be considered any more representative of the organism's evolutionary history than those associated with, say, its plastid genome. (ref. 8, pp. 48 and 50)

We disagree. In our view, these authors have seriously misunderstood the origin and nature of the eukaryotic cell. Fundamentally, it is an anaerobic cell with but a single (nuclear) genome. A number of eukaryotes with this physiological and genomic organization exist today—microsporidia and the diplomonads, for example—and they represent the deepest known divergences from the eukaryotic stem. It is these phenotypes, if any, that should most resemble the ancestral eukaryotic phenotype. In these cases, there is no indication that endosymbioses had anything to do with their evolution or that more than one genetic lineage was involved. Moreover, no biologist would have difficulty identifying a eukaryotic cell that had lost its mitochondria or chloroplasts (e.g., petite mutants of yeast are clearly yeasts, plastid-free euglenoids are clearly euglenoids, etc.). This is because the contribution of the organellar genomes to the total genome is quantitatively insignificant, and they encode few of the physiological traits and none of the components of the complex architecture of the cell (other than some of the constituents of the organelles themselves); nor do they encode any of the structural or regulatory machinery of growth and division; they are clearly accessory genetic elements superimposed on the fundamental eukaryotic nuclear genome. There can be no denying the importance of endosymbioses in evolution; however, doing so does not require discarding, as Margulis and Guerrero by implication do, the idea that eukaryotes have a unique, meaningful phylogeny.

We turn now to the matter of which classical concepts and criteria have counterparts or validity on the molecular level. First, the concept of grade. In global classification, grade is a worthless concept. It is phenetic in nature and fundamentally non-Darwinian; as de Queiroz (ref. 27, p. 252) points out, "grades are holdovers from pre-evolutionary taxonomies based on the Scala Naturae, or great chain of being." The notion of progressive grade, "levels of biological improvement," is the essence of Mayr's counterproposal for a two-domain global system [Prokaryotae and Eukaryotae (9)]. He writes of ". . . the series of evolutionary steps in cellular organization leading from the prokaryotes to the eukaryotes . . . ," of the "difference in structural organization between prokaryotes and eukaryotes [being] an order of magnitude greater than the relatively small difference between the Archaebacteria and the Eubacteria," and with Ashlock (17) that prokaryotes ". . . represent a definite evolutionary grade, which, in spite of its heterogeneity, it is convenient to recognize under a common name. . .". Mayr here makes the logical error of defining a group on the basis of the absence of characters whose presence defines another group. (It is formally equivalent to defining the reptilian grade only as lacking the characteristic avian features, a definition that would almost certainly include the amphibians, fish, invertebrates, etc., among the reptiles.) For grade to be used at all, it must be defined in terms of the shared possession of characters, not their absence. The cellular characteristics shared specifically by Archaea and Bacteria appear to be analogous, not homologous. Moreover, it is entirely possible that the common ancestor of all prokaryotes, the universal ancestor, was not even complex enough to be defined as prokaryotic (24). Unless Mayr can demonstrate otherwise, his taxon Prokaryotae should be considered polyphyletic.

Although grade is an unworkable concept (in global classification), there is merit in the related but more definable (and less cathected) concept of divergence, or degree of modification. Classically, divergence is assessed by a phenetic analysis at the organismal level, involving the accumulation of unique derived traits (autapomorphies). However, divergence (in the Darwinian sense) is not strictly quantitative, not simply how many autapomorphies a lineage possesses; it is also their general nature. Study of the fossil record has shown that lineages can undergo periods of very rapid evolution—evolutionary bursts—characterized by numerous, drastic, and unusual phenotypic changes, evolutionary radiation, and instability (28, 29). This rapid evolution stands in sharp contrast to the slower, less spectacular evolution that tends to characterize much of a lineage's history (28). Comparing rapid and slow evolution leaves one with an impression that no matter how long the slow evolutionary process were to continue, the changes accumulated would never come to resemble in kind those produced by the rapid one. As Simpson (28) succinctly put it, when the tempo of evolution changes, its mode does so as well. Tempo and mode are thus both central to what Darwin described as degree of modification. Their counterparts must now be sought on the molecular level.

Tempo is quite straightforward: it is measured (in relative terms) on the molecular level with a precision and to an extent that classical evolutionists would never have thought possible—the evolutionary clock manifests itself clearly *only* on that level. In molecular terms, it is easy to see that one lineage has diverged more (has evolved more rapidly) than another and to describe that difference by simple quantitative measure. The question is whether the difference here is also qualitative—modal. This may well be the case. When the rRNA of a particular species appears to be rapidly evolving, an unexpectedly high fraction of the changes tend to involve positions in its sequence whose compositions are normally nearly invariant. Rapidly evolving rRNAs, in other words, lose an unexpectedly high number of plesiomorphic sequence characters (and, consequently, gain autapomorphic ones)—a pattern of change that overall is not seen in slowly evolving rRNAs (24).

[To give an illustration here, within the (11 member) leuconostoc subgroup of the lactobacilli, one representative, *Leuconostoc oenos*, stands out (by rRNA measure) as being especially rapidly evolving (30). If one defines as plesiomorphic characters those positions in the rRNA sequence whose compositions have not changed in all lactobacilli (≈30 species) outside of the leuconostoc subgroup, one finds that *L. oenos* has lost triple or more the number of plesiomorphies (gaining that many autapomorphies) than have any of the (more slowly evolving) members of its subgroup (C.R.W., unpublished data); this is also nearly three times as many of these same characters as are lacking in an outgroup sequence such as that of *Bacillus subtilis*.]

Therefore, a tempo–mode-like phenomenon does appear to exist on the molecular level. Although demonstrated for rRNAs, the phenomenon is almost certainly general—a property of the cell as a whole—for it seems to occur in other molecules in the same organism(s) as well (although with fewer sequences available, the basis for definition and comparison is less precise). What remains to be demonstrated is whether evolutionary modes defined on the molecular and organismal levels are equivalent; in other words, are lineages defined as undergoing rapid and unusual evolution on the former level also those that show numerous idiosyncratic traits on the latter? This is a matter for future research. For now, it is a moot point; global taxonomy requires molecular definition, and phenomena having the essential characteristics of evolutionary tempo and evolutionary mode exist at the molecular level, thus permitting the development of a multidimensional classificatory approach.

General Outline of a Global System

Our overall classificatory perspective was stated at the outset of this article. A global system should be as complete an encapsulation of the evolutionary paradigm as possible, one that not only aids understanding but also helps to define and direct the course of biology.

A basic requirement of a proper global classification is uniformity in methods and characters used in defining and ranking taxa. This in effect imposes three (related) conditions. (*i*) That the starting point for definition of taxa be phylogenetic analysis (not phenetic clustering). (*ii*) That the main taxa be defined primarily on the molecular level, in terms of molecular (elementary) traits; grosser level traits (as explained above) are not excluded by this but play useful substantiating roles and may be important in assessing degree of modification. (*iii*) That these definitions be based on homologous characters, universal molecular functions—nonhomologous characters again being confirmatory and useful in assessing divergence and mode.

Our perspective leads to a global system that would embody both principal dimensions of the evolutionary process, genealogy and degree of modification within a taxon. (Note that degree of modification itself is multidimensional, reflecting tempo, mode, and age of a group.) Unfortunately, classifications are fundamentally one dimensional in structure. Multiple qualities, therefore, can be transformed into a classification only in a degenerate fashion; information is necessarily lost in the process. The solution to this, however, is not the cladistic solution—making a classification perfectly representative and completely consistent by basing it on a single quality (genealogy). Rather it is to increase, in effect, the dimensionality of classificatory structure in a way that captures, albeit imperfectly, as much information as possible. While this demands perhaps radical change in systematic perspective, it could be accomplished with only minor change to conventional taxonomic structure, for there are relatively simple notational devices (labels, pointers) by which the dimensionality of a classification can in effect be expanded.

In the global system, we envision that taxonomic boundaries would correspond to those (monophyletic) regions in a phylogenetic tree that are characterized by significant common loss of plesiomorphies and gain of (syn)apomorphies. Taxonomic rank should be consistently determined, by some combination of the extent and type of divergence encompassed by a taxon, using criteria that are largely independent of the group's phylogenetic position. However, the taxonomic hierarchy in which such ranked taxa are related should be structured so as to retain as much genealogical information as possible given these conditions.

This manner of defining a system will naturally generate paraphyletic taxa, when different subgroups within a given group show significantly different degrees of divergence. Biologists tend to conceive paraphyletic taxa as being monophyletic groupings that genealogically and classificatorily include taxa of lower rank and genealogically (but not classificatorily) include one or more taxa of equivalent rank—for example, Reptilia vis-a-vis Aves. However, with objective standards for defining taxonomic rank, it will also be necessary to admit paraphyletic taxa that genealogically (but not classificatorily) include one or more taxa of higher rank. The mycoplasmas may be an example of such a taxonomic inversion. Genealogically they belong to a relatively minor subgroup of the clostridia, whose other members might comprise a family. Yet, because of their highly diverged nature (quantitatively and qualitatively, genotypically and phenotypically), the mycoplasmas could easily be defined as an order embedded within this family.

Such a reformed system, amended so as to preserve genealogical information for the types of paraphyletic situations defined above, would be generally more informative, and the information more readily retrievable, than cladistic systems (which provide only genealogical information) or phenetic/gradistic systems (which lose too much genealogical information). It would also be less counterintuitive than cladistic systems (which, by ignoring evolutionary mode, subordinate some groups in ways that appear artificial); yet, because of its fundamentally phylogenetic basis, it could still aid the recasting of biology into a fully evolutionary discipline.

> . . . genealogy by itself does not give classification (ref. 32, p. 247), [but o]ur classifications will come to be, as far as they can be so made, genealogies. Charles Darwin (ref. 11, p. 486)

We are grateful to the following individuals for helpful discussion and comment: John Gillespie, Philip Ward, and Ellis McLeod. Our work in this area receives the following support: the German Collection of Microorganisms to O.K. and the National Space and Aeronautics Administration to C.R.W.

1. Burkhardt, F. & Smith, S. (1990) *The Correspondence of Charles Darwin 1856–1857* (Cambridge Univ. Press, Cambridge, UK), Vol. 6.
2. Woese, C. R., Kandler, O. & Wheelis, M. L. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 4576–4579.
3. Gogarten, J. P., Kibak, H., Dittrich, P., Taiz, L., Bowman, E. J., Bowman, B. J., Manolson, M. F., Poole, R. J., Date, T., Oshima, T., Konishi, J., Denda, K. & Yoshida, M. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 6661–6665.
4. Iwabe, N., Kuma, K., Hasegawa, M., Osawa, S. & Miyata, T. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 9355–9359.
5. Whittaker, R. H. (1959) *Q. Rev. Biol.* **34**, 210–226.
6. Chatton, E. (1938) *Titres et travaux scientifiques (1906–1937) de Edouard Chatton* (Sottano, Sète, France).
7. Murray, R. G. E. (1974) in *Bergey's Manual of Determinative Bacteriology*, eds. Buchanan, R. E. & Gibbons, N. E. (Williams & Wilkins, Baltimore), 8th Ed., pp. 4–9.
8. Margulis, L. & Guerrero, R. (1991) *New Sci.* **132**, 46–50.
9. Mayr, E. (1990) *Nature (London)* **348**, 491.
10. Mayr, E. (1991) *Nature (London)* **353**, 122.
11. Darwin, C. (1859) *On the Origin of Species by Means of Natural Selection, or the Preservation of Favored Races in The Struggle for Life* (John Murray, London).
12. McNeill, J. (1979) *System. Zoo.* **28**, 465–482.
13. Hennig, W. (1966) *Phylogenetic Systematics* (Univ. Illinois Press, Urbana, IL).
14. Mayr, E. (1974) *Zeit. Zool. System. Evol.* **12**, 94–128.
15. Cronquist, A. (1987) *Bot. Rev.* **53**, 1–52.
16. Donoghue, M. J. & Cantino, P. D. (1988) *Bot. Rev.* **54**, 107–128.
17. Mayr, E. & Ashlock, P. D. (1991) *Principles of Systematic Zoology* (McGraw-Hill, New York).
18. Ashlock, P. D. (1979) *System. Zoo.* **28**, 441–450.
19. Huxley, J. S. (1958) *Uppsala Univ. Arsskr.* **6**, 21–39.
20. Kluyver, A. J. & van Niel, C. B. (1936) *Zbl. Bakteriol. Parasit. Infekt. II* **94**, 369–403.
21. Stanier, R. Y. & van Niel, C. B. (1941) *J. Bacteriol.* **42**, 437–466.
22. van Niel, C. B. (1946) *Cold Spring Harbor Symp. Quant. Biol.* **11**, 285–301.
23. Stanier, R. Y. (1942) *Bacteriol. Rev.* **6**, 143–196.
24. Woese, C. R. (1987) *Microbiol. Revs.* **51**, 221–271.
25. King, J. L. & Jukes, T. H. (1969) *Science* **164**, 788–789.
26. Kimura, M. (1983) *The Neutral Theory of Molecular Evolution* (Cambridge Univ. Press, Cambridge, UK).
27. de Queiroz, K. (1988) *Philos. Sci.* **55**, 238–259.
28. Simpson, G. G. (1944) *Tempo and Mode in Evolution* (Columbia Univ. Press, New York).
29. Mayr, E. (1942) *Systematics and the Origin of Species* (Columbia Univ. Press, New York).
30. Yang, D. & Woese, C. R. (1989) *System. Appl. Microbiol.* **12**, 145–149.
31. Sogin, M. (1991) *Curr. Opinion Genet. Dev.* **1**, 457–463.
32. Darwin, F., ed. (1887) *The Life and Letters of Charles Darwin, Including an Autobiographical Chapter* (John Murray, London), Vol. 2.