

## Perspective

# Default taxonomy: Ernst Mayr's view of the microbial world

(taxonomic domains/phylogeny/universal ancestor/biological classification)

Carl R. Woese\*

Department of Microbiology, University of Illinois at Urbana–Champaign, B103 Chemical and Life Sciences Laboratory, MC-110, 601 South Goodwin Avenue, Urbana, IL 61801

Contributed by Carl R. Woese, July 29, 1998

**ABSTRACT** This perspective is a response to a taxonomic proposal by E. Mayr ["Two empires or three?" (1998) *Proc. Natl. Acad. Sci. USA* 95, 9720–9723]. Mayr has suggested that the now accepted classification of life into three primary domains, Archaea, Bacteria, and Eucarya—originally proposed by myself and others—be abandoned in favor of the earlier Prokaryote–Eukaryote classification. Although the matter appears a taxonomic quibble, it is not that simple. At issue here are differing views as to the nature of biological classification, which are underlain by differing views as to what biology is and will be—matters of concern to all biologists.

In his article "Two empires or three?" recently published in this journal (1), Ernst Mayr rejects the three-domain structuring of the living world—Archaea, Bacteria, and Eucarya—that has emerged from molecular studies over the past two decades (2). He would return to the older Prokaryote–Eukaryote classification, which shaped biology's overview of life on this planet for the last half century. I have argued previously that, from the outset, the prokaryote–eukaryote dichotomy was an idea never properly tested (3). And I believe the complacency that this simplistic formulation generated adversely affected the development of biology, for it served among other things to mask the fact that the basis for a true science of microbiology, "the concept of a bacterium" (4), was never developed. To return to the prokaryote–eukaryote dogma (with its lingering false connotations) would have a similarly negative effect now, once again on microbiology, but this time too on the study of evolution—both of these fields currently in states of revolutionary development.

I am not inclined to reply to Dr. Mayr's article in detail or in kind, although some of the former is required. Dr. Mayr and I see things from very different perspectives, and it is this difference in perspective, not some local taxonomic dispute, that needs to be aired. Therefore, I will respond to what Dr. Mayr's article really is—i.e., a pronouncement concerning the future direction biology should take.

### The History of the Eukaryote–Prokaryote Dichotomy

The idea that the living world is divided in the first instance into two very different types of organisms, prokaryotes and eukaryotes, is generally attributed to Chatton in the 1930s (4, 5), but the notion that bacteria (schizomycetes) are somehow unique, are the "first and simplest division of living beings," goes back to the great microbiologist Ferdinand Cohn in the last century (6). Microbiologists of the early twentieth century, however, were loath to accept the principal implication of a prokaryote–eukaryote dichotomy—i.e., that all prokaryotes are of a kind (although the monophyletic nature of eukaryotes

was never in doubt). These early microbiologists were troubled, for one, by the fact that "prokaryote" (a term they rarely used) was defined on the basis of "entirely negative characteristics"—as not possessing certain eukaryotic traits (7). For another, the morphological and physiological diversity they encountered among bacteria readily led them intuitively to consider that the various major bacterial groups "are of polyphyletic origin" (8).

Nevertheless, microbiologists eventually did come around, accepting that "prokaryote," like "eukaryote," was indeed a monophyletic taxon. The apparent reason for this remarkable change in the microbiologist's outlook was that by the 1960s technology had reached the point where it was possible to define the prokaryote in positive rather than solely in negative terms (4). Yet, on closer examination this conceptual reversal seems more a matter of a new-found faith than of any new-found facts (3): in defining the prokaryote in the new cytological and molecular terms, there were "remarkably few comparative studies, [which means that] the application of the newer [techniques] for taxonomic purposes entails generalization from limited cases" (9). In other words, nobody at the time felt it necessary to ask whether certain properties of *Escherichia coli* and occasionally a few other bacteria were actually properties of prokaryotes in general. The monophyletic nature of prokaryotes was simply taken for granted! Why would microbiologists do this; why would they trade their previous scientific skepticism for an untested belief? The reason, I think, is that a monophyletic prokaryote appeared to relieve microbiologists of the necessity of determining phylogenetic relationships in order to develop a "concept of a bacterium" (3, 4).

For at least the first half of this century determining microbial phylogenetic relationships had proven impossible, and their repeated failures to do so seem to have brought some (perhaps most) microbiologists to believe that these relationships were inherently indeterminable, at least at the higher taxonomic levels (6, 10, 11). As a consequence, a default position was taken, in which the prokaryotes were simply assumed to be specifically related to one another (4). From this, microbiologists of the 1960s then argued that prokaryotes could be defined biologically—that the much needed concept of a bacterium could be developed—simply by cataloging the differences between prokaryotes and eukaryotes (4). Of course this did not work, and no one at the time seemed to have appreciated that it *could* not, which left microbiology and the rest of biology saddled with the false notion that all prokaryotes are of a kind, are specifically related (3).

### Lessons from Genes

Technology ultimately came to the rescue, and through the use of ribosomal RNA (rRNA) sequence comparisons it proved

possible to determine microbial phylogenetic relationships. At the molecular level both the higher and lower prokaryotic taxa readily revealed themselves, as did their relationships to one another and to the eukaryotes (2, 12–15). The resulting universal phylogeny unexpectedly showed that the taxon “prokaryote,” which all biologists (including myself) had heretofore assumed to be monolithic, in reality comprises two fundamentally distinct groups of organisms, each no more like the other than they are like eukaryotes (12, 13). Thus, there are actually three, not two, primary phylogenetic groupings of organisms on this planet (2, 12). Moreover, a phylogenetic tree based upon the molecular data shows that the newly recognized one, the Archaea, is, if anything, more closely related to eukaryotes than to the familiar bacteria (2, 14, 15). Here then was biology’s first glimpse of the full evolutionary landscape. Within the framework of a universal phylogenetic tree the study of evolution could now move beyond the confines of the plant and animal kingdoms and into the enormous realm of microorganisms.

This universal tree, this beautiful and compelling image, has seen some rough weather of late, not because of Dr. Mayr, but because of the complexity introduced into the picture by the large amounts of additional sequence data now available, data that reveal the extent to which lateral gene transfer has shaped the evolutionary course over the long haul. As a result, phylogenetic trees based upon protein genes tend not to show deep phylogenetic branchings consistent with the rRNA tree. But neither do the branchings in these dissonant trees tend to agree with one another.

This state of seeming confusion has given rise to skepticism: as to the validity of the rRNA-based phylogenies (16); as to whether the deep evolutionary record might have been erased, and therefore we will never know the nature of the universal ancestor; as to whether there is more of a phylogenetic “continuum” than there are three distinct organismal groups (J. C. Venter as quoted in ref. 17); as to whether a specific relationship between Archaea and Eucarya actually exists, or is “an artifact of low sampling” (J. C. Venter as quoted in ref. 17). It is in this climate of uncertainty regarding the significance of gene trees, uncertainty regarding the deep evolutionary branchings, and puzzlement as to the dynamic of the deep evolutionary process, that Ernst Mayr has renewed his earlier call (18, 19) for a return to the default taxonomic position.

But those who see only confusion (and perhaps advantage) in the current state of genomics should look again. The observed high levels of lateral gene transfer do not fell the universal tree; they tell us what it really is (20). They do not smear the primary groupings of organisms into some “continuum”; they paint a clearer picture of what actually defines and distinguishes them (20). They do not prevent our inferring the nature of the universal ancestor; they reveal that nature.

Lateral gene transfer was part and parcel of the universal ancestor. That ancestor was a communal entity, a community that survived and evolved as a whole, as an aggregate, not as individual lineages (20, 21). The three primary groupings are defined and distinguished not by some consensus of all the genes in a genome, but in terms of the (genes representing the) differing “design commitments” each of the three made at the stage cells as we know them were still evolving. The universal tree was initially a gene tree that grew over the eons into an organismal tree, as organisms *per se* emerged from the evolutionary melting pot (20). At least that is how I and perhaps a few others view these matters (20–22).

The point here is that for the first time, biologists are in a position to attack the greatest of evolutionary problems, how cells evolved. And key to this problem is the fact that at some early evolutionary stage there existed three ill-defined (populations of) entities, each distinctly different from the other, which through their individual evolutions and their evolutionary interactions became the three major cell types that now

exist on Earth. Now is definitely not a time to alter our classification of organisms so as to suggest that biology’s primary focus be on animals and plants and that all bacteria are of a kind (1). The primary interest today is in cells and molecules, and our study of animals and plants necessarily operates in the first instance from this platform. Biology, like physics before it, has moved to a level where the objects of interest and their interactions often cannot be perceived through direct observation. And, as in the case of physics, biology’s “subatomic” (subcellular) level is rich in information, rich in understanding, and rich in beauty. It is at the level of molecules that biologists are able to expand their perspective to genuinely encounter the microbial world, the full extent of which we have still to experience, a world whose metabolic and phylogenetic diversity completely dwarfs that seen in the animal and plant kingdoms combined. But more important (to me) is that biology’s incursions into molecules and genomes have transformed an otherwise rather stale and isolated discipline of evolutionary study into one of the most vibrant, central, and important facets of our science.

### On the Nature of a Biological Classification

The nature of biological classification figures heavily in Dr. Mayr’s argument, and so must be addressed. In his view the present issue can and should be defined and settled in terms of established classificatory formalisms. If there were ever an issue in biological classification that cannot be settled by pedantry, it is this one. Never before has there been a less subjective highest taxonomic level than that defined by the universal phylogenetic tree. To Mayr, the issue is one of whether we should define two or three domains and what the classificatory precedents or rules for deciding this are. However, the universal phylogenetic tree tells us that the domains are unique among taxa and that their number and their composition are not subject to classificatory fiat, but are naturally defined.

Mayr defines biological classification as “an information storage and retrieval system,” whose aim is the same as that of a “classification of books in a library or goods in a store,” i.e., “to locate an item with a minimum of effort and loss of time” (1). This leads then to a “principle of balance,” by which “the retrieval of information is greatly facilitated [when] the taxa at a given categorical rank are, as far as possible, of equal size and degree of diversity” (1). Is this what biological classification is about? Is it this arbitrary, this artificial? Is functional utility the primary consideration in its design? Of course not, and I am sure Dr. Mayr knows that. Darwin said: “Our classifications will come to be, as far as they can be so made, genealogies” (23), and that dictum forever changed the nature of biological classification. Since Darwin’s time the basis for classification has been absolute, its primary aim being to encapsulate organismal descent. And this natural ordering necessarily has utility as an information storage and retrieval system.

But, there is one thing above all that we need to know about a biological classification: it is (or represents) a theory, a *de facto* theory, exhibiting the three main characteristics of any good theory: A biological classification has *explanatory power*—i.e., it aids in and enriches the interpretation of findings, integrating them into a deeper, more meaningful context. A biological classification *makes testable predictions*, which lead to the design of experiments. And finally, like any overarching theory, a biological classification has *conceptual power*; it influences the focus of a discipline, steering it in certain directions and away from others. As seen above, it was microbiology’s failure to recognize the prokaryote–eukaryote classification as a theory in the first place that led to its failure to properly test it experimentally; and it was that theory’s conceptual power that led microbiologists to overlook the fact that one of the discipline’s central problems, “the concept of

a bacterium" (4), had not been resolved (3). To bring this failed theory back into prominence now (1) would indeed be unfortunate.

Darwin's world and that of the classical evolutionists barely represents the last billion years of evolution, the Age of Multicellular Organisms. But the world that modern biology knows encompasses nearly the full range of life's 3- to 4-billion-year course. Whereas Darwin and the classical evolutionists could take organisms and their genealogies for granted, biology today cannot. Evolutionary biology (and biology in general) has changed greatly over the last several decades, and biological classification, as theory, must change accordingly. In the present instance this requires more than simply redefining taxonomic categories. We must now question and, if need be, redefine the concepts that underlie the classification itself.

A paraphrasing of the above Darwin quotation captures how I view a biological classification (theory): "Our classifications will come to be, as far as they can be so made, *representations of the evolutionary course.*" While the lower taxonomic levels would not be significantly impacted by such a change in outlook, the higher levels would. In my view the highest level taxa, the domains, need to reflect the evolutionary course that split the universal ancestor into the individual ancestors of the three primary lines of descent. This is not primarily a matter of cataloging extant organisms. Nor is it even a matter of representing genealogical relationships. Modern phenotypes did not exist and organismal genealogies probably had no meaning at the time when the domains formed (20). And it is also not a matter of counting which pairings of the three ancestors share the most genes (1). Evolution at this early stage was probably a symphony in lateral gene flow. Yet it is not the migrant genes that define the ancestors of the three domains, but rather the fixed genes, those that were confined to a single domain and became inherited vertically (20). A core of unique vertically inherited genes defines and distinguishes each of the primary lines of descent (unpublished data), and it is this core that must define the highest level taxa in a biological classification theory.

As mentioned above, the universal phylogenetic tree (which underlies any biological classification) is probably not a normal organismal tree. This tree's deepest branchings were formed when organisms as we know them had not yet evolved. These deep branchings represent the genealogies of a few genes only, special genes such as those for the basic components of the translation and transcription machineries (20). As it grew over time, as more and more genes came to share a common history, this tree became a true organismal tree—but only in its superficial branchings (20). Consequently, the highest taxa, the domains, do not refer to organisms in the conventional sense; they reflect only the evolutionary stage during which organisms, modern types of cells, were coming into being. In other words, these deepest branchings on the universal tree represent the *initial* origin of species. (I would emphasize that the validity of the argument that a biological classification represents an evolutionary theory in no way turns upon the correctness of the specific evolutionary scenario that I favor.)

### The Microbial World—Biology's Sleeping Giant

My colleague Mark Wheelis puts it this way: "The earth is a microbial planet, on which macroorganisms are recent additions—highly interesting and extremely complex in ways that most microbes aren't, but in the final analysis relatively unimportant in a global context." (M. Wheelis, personal communication). The cellular biomass on this planet is predominantly microbial (24), and in numbers the earth's microbial population completely dwarfs that of multicellular organisms. It is the web of interactions among microorganisms that defines and supports the biosphere, the global ecosystem. Microorganisms live deep in the crust of the earth and even shape the planet

itself, playing major roles in mineral deposition. The composition of the atmosphere reflects microbial metabolism, for microorganisms are the earth's dominant biochemical factories and bioenergetic power plants. Microbial life on this planet would remain largely unchanged were all plant and animal life eliminated, but the elimination of microbial life itself would lead in very short order to a completely sterile planet.

Microorganisms have also played pivotal, if not essential, roles in the evolution of multicellular eukaryotic life: Bacteria invented the photosynthesis that became the defining essence of plants; they invented the aerobic respiration that would allow animals to breathe. And as genomic studies reveal, these represent only a few of the many functions/genes the eukaryotic cell has assimilated from Bacteria and Archaea. Plant and animal life cannot exist, and would never have come into existence, except for microbial life. We are far more integrally connected to the "prokaryotic" world than is generally appreciated. Yet for most biologists it has always been a case of "prokaryote" versus eukaryote (it is inherent in the way the science of biology is currently structured). Is not "prokaryote" *ergo* eukaryote more appropriate?

The power and importance of the microbial world are clear. Yet how diverse is it in comparison to the world of multicellular eukaryotes? That is what Dr. Mayr asks; for he sees relatively little diversity either within or between the Bacteria and the Archaea. By species counting there appear to be only a handful of archaeal groups, "about 175," and a somewhat larger number of bacterial ones, "[a]round 10,000" (1). But this compares to the 10,000 or so species of birds alone and "many millions of species of insects" (1). Any microbiologist today would tell you, however, that we are aware of only a fraction of the diversity in the microbial world. On the order of half of the so-far-detected major bacterial taxa, the kingdoms and divisions, have few or no cultured (and so formally described) representatives (25), and this implies that a higher fraction still of the intermediate level taxa and a higher fraction again of bacterial species have yet to be identified. There is no way one could possibly begin to estimate the true number of distinguishable microbial "species." But "species counts" are not the point, and neither is the related thorny problem of how one defines a microbial "species." The nature and measure(s) of diversity itself are the crux of the issue.

Diversity can be of many types. It can be at the level of structure and organization; it can be anabolic or catabolic enzymatic diversity; it can be environmental adaptation at the molecular/biochemical level; it can be in the basic information processing systems of the cell; and so on. Clearly the vast diversity among birds and among insects is structural diversity, whereas that among the Bacteria or the Archaea is necessarily of the other types. Dr. Mayr's is an eye-of-the-beholder type of diversity. It rests on the incredible capacity of the human eye to distinguish minute differences in pattern. But almost all microbial diversity cannot be sensed visually, which means that subtle variations in pattern almost always go undetected. (I often wonder how much more diversity we microbiologists would "see" were it possible, in a computer-assisted way, to transform the metabolism of a bacterium into visual images.) When he compares plant and animal diversity to microbial diversity, Dr. Mayr is comparing apples and oranges, and his attempt to apply globally a parochial and subjectively defined concept of diversity serves only to reveal the futility in such an approach.

Although Dr. Mayr does not succeed in demonstrating that the microbial world lacks diversity, his article does bring to the fore the difficulties we face in detecting, defining, and quantitating microbial diversity. Before a phylogenetic framework existed microbiologists had no hope of attacking this problem; it could not even be defined. Now the universal phylogenetic tree provides a measure of diversity. In other words, diversity can be defined in terms of degrees of difference at the genetic

sequence level. While to some, this may not have the reassuring feel of a classical phenotypic approach to measuring diversity, it does have the distinct advantage of being objective, naturally defined, and, I would claim, universally applicable.

So here is the microbiologist's position: We know that most of the diversity in the microbial world remains undetected because the vast majority of microbial species are still undetected and/or uncharacterized. We know that microbial diversity is of a kind that usually cannot be visually sensed, which means that subtle (and perhaps some not so subtle) distinctions in it pass unnoticed. But we also know that the microbial world is far more ancient than that of multicellular organisms, that this world encompasses and defines the biosphere, and that microorganisms account for the bulk of the planet's biomass. Is that not a *prima facie* case for a microbial world that contains the bulk of the planet's biodiversity? And by the above genetic measure of diversity, it surely is: over 90% of the biodiversity on this planet is microbial (2, 25).

We have obviously only begun to capitalize on the evolutionary relationships among eucaryal, archaeal, and bacterial genes, to tell us how life works. While a knowledge of the human genome will provide society with insights into heritable diseases, a knowledge of microbial genomes will be needed to cope with most infectious diseases in the future (as antibiotics lose their effectiveness). And, as biologists are finding out, a knowledge of microbial genomes may be essential to any comprehensive interpretation of the human genome. The industrial applications of microbial genomics are still in their infancy. Who knows what applications will flow from a deep understanding of microbial diversity or a knowledge of the universal ancestor? Even from the most practical of viewpoints it is easy to see that biology needs more focus on microorganisms, although, as the reader knows from the above, my rationale for it is a fundamental one. The future of microbiology is bright: it promises all manners of powerful applications; it will be the forefront of evolutionary research; and it will become biology's teacher and guide.

### Summary and Conclusion

Dr. Mayr's article (1) is not the taxonomic quibble it might seem. It is a *de facto* pronouncement on the nature of biology. I have, therefore, responded accordingly, discussing the following larger issues:

(i) The nature of biological classification. A biological classification is in effect an overarching evolutionary theory that guides our thinking and experimentation, and it must be structured (and that structure changed if necessary) to reflect evolutionary reality.

(ii) The prokaryote–eukaryote dichotomy. This dichotomy, which Dr. Mayr proposes to reinstitute, is a failed taxonomic theory that was never recognized as theory, and so tested in a timely fashion, with the consequence that it has adversely affected the development of biology, especially microbiology, in the latter half of this century.

(iii) The scientifically perceived importance of a group of organisms must reflect the natural importance of the group. The gulf between perception and reality in the case of microorganisms is disturbingly large, and biologists need to address that.

(iv) Microbial diversity. It is only within the last few decades that microbiology has had the (phylogenetic) framework within which to explore and define the diversity of microbial life. We know very little about it. Microbial diversity is far more than a listing of distinguishable microbial species. We need to understand the quality of microbial diversity, for it is the diversity that defines the biosphere of this planet.

(v) Evolution must be integrated into the fabric of molecular biology. Molecular biology from the start has viewed organisms and molecules as being essentially independent of the “historical accidents” that produced them. It must be understood that in a real sense an organism *is* its evolution. This means that any comprehensive understanding of a biological entity, be it an organism or a molecule, necessarily has an evolutionary component.

(vi) Finally, the disagreement between Dr. Mayr and myself is not actually about classification. It concerns the nature of Biology itself. Dr. Mayr's biology reflects the last billion years of evolution; mine, the first three billion. His biology is centered on multicellular organisms and their evolutions; mine on the universal ancestor and its immediate descendants. His is the biology of visual experience, of direct observation. Mine cannot be directly seen or touched; it is the biology of molecules, of genes and their inferred histories. Evolution for Dr. Mayr's is an “affair of phenotypes” (1). For me, evolution is primarily the evolutionary *process*, not its outcomes (20). The science of biology is very different from these two perspectives, and its future even more so.

I am very grateful to Norman Pace, David Graham, and Mark Wheelis for their readings of the manuscript during its preparation and the very helpful suggestions that have come therefrom. I thank Ernst Mayr for allowing me to see his manuscript prior to its publication. My work is partially supported by the National Aeronautics and Space Administration and somewhat by the Department of Energy.

1. Mayr, E. (1998) *Proc. Natl. Acad. Sci. USA* **95**, 9720–9723.
2. Woese, C. R., Kandler, O. & Wheelis, M. L. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 4576–4579.
3. Woese, C. R. (1994) *Microbiol. Rev.* **58**, 1–9.
4. Stanier, R. Y. & van Niel, C. B. (1962) *Arch. Mikrobiol.* **42**, 17–35.
5. Chatton, E. (1938) *Titres et Travaux Scientifique (1906–1937) de Edouard Chatton* (Sottano, Sète, France).
6. Cohn, F. (1875) *Beitr. Biol. Pflanz.* **1** **3**, 141–208.
7. Pringsheim, E. G. (1923) *Lotus* **71**, 357–377.
8. Stanier, R. Y. & van Niel, C. B. (1941) *J. Bacteriol.* **42**, 437–466.
9. Murray, R. G. E. (1962) in *12th Symposium of the Society for General Microbiology*, eds. Ainsworth, G. C. & Sneath, P. H. A. (Cambridge Univ. Press, Cambridge, U.K.), pp. 116–144.
10. van Niel, C. B. (1955) in *A Century of Progress in the Natural Sciences 1853–1953* (California Academy of Sciences, San Francisco), pp. 89–114.
11. Stanier, R. Y., Doudoroff, M. & Adelberg, E. A. (1963) *The Microbial World* (Prentice-Hall, Englewood Cliffs, NJ), 2nd Ed.
12. Woese, C. R. & Fox, G. E. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5088–5090.
13. Fox, G. E., Stackebrandt, E., Hespell, R. B., Gibson, J., Maniloff, J., Dyer, T. A., Wolfe, R. S., Balch, W. E., Tanner, R. S., Magrum, L. J., et al. (1980) *Science* **209**, 457–463.
14. Iwabe, N., Kuma, K.-i., Hasegawa, M., Osawa, S. & Miyata, T. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 9355–9359.
15. 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., et al. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 6661–6665.
16. Pennisi, E. (1998) *Science* **280**, 672–674.
17. Anonymous (1998) *The Scientist* **12**, 11.
18. Mayr, E. (1990) *Nature (London)* **348**, 491.
19. Mayr, E. (1991) *Nature (London)* **353**, 122.
20. Woese, C. (1998) *Proc. Natl. Acad. Sci. USA* **95**, 6854–6859.
21. Woese, C. R. (1982) *Zentralbl. Bakteriell. Mikrobiol. Hyg. Ser. C* **3**, 1–17.
22. Kandler, O. (1994) *Syst. Appl. Microbiol.* **16**, 501–509.
23. Darwin, F., ed. (1887) *The Life and Letters of Charles Darwin, Including an Autobiographical Chapter* (John Murray, London).
24. Whitman, W. B., Coleman, D. C. & Wiebe, W. J. (1998) *Proc. Natl. Acad. Sci. USA* **95**, 6578–6583.
25. Pace, N. R. (1997) *Science* **276**, 734–740.