MICROBIOLOGICAL REVIEWS, June 1987, p. 221–271 Vol. 51, No. 2
0146-0749/87/020221-51$02.00/0

Bacterial Evolution

CARL R. WOESE

Department of Microbiology, University of Illinois, Urbana, Illinois 61801

PERSPECTIVE 222
A Fruitless Search and Its Consequences 222
Three Ideas That Shape Our Concept of Bacterial Evolution 224
Procaryote-eucaryote dichotomy 224
Oparin Ocean scenario 224
Darwin's warm little pond 226
MEASUREMENT OF BACTERIAL PHYLOGENETIC RELATIONSHIPS 226
Three-Dimensional and One-Dimensional Characters 226
The evolutionary clock 226
A missed opportunity 226
Nature of Molecular Chronometers 227
Clocklike behavior 227
Phylogenetic range 227
Size and accuracy 227
rRNAs, the Ultimate Molecular Chronometers 227
Why they are so good 227
Oligonucleotide cataloging 228
Analysis of Sequence Alignments 228
Distance matrix methods 228
Maximum parsimony analysis 228
Cluster analysis 228
DO BACTERIA HAVE A GENEALOGY AND A MEANINGFUL TAXONOMY? 230
THE UNIVERSAL PHYLOGENETIC TREE 231
EUBACTERIAL PHYLOGENY 232
Background 232
Major Eubacterial Groups and Their Subdivisions 232
Purple Bacteria 236
α-Purple bacteria 237
β-Purple bacteria 238
γ-Purple bacteria 238
δ-Purple bacteria 238
Gram-Positive Eubacteria 240
Cyanobacteria 242
Green Sulfur Bacteria 243
Spirochetes 244
Bacteroides, Flavobacteria, and Their Relatives 244
Planctomyces and Their Relatives 244
Chlamydiae 245
Radiation-Resistant Micrococci and Their Relatives 245
Green Non-Sulfur Bacteria and Their Relatives 245
Other Eubacterial Phyla 245
Overall Structure of the Eubacterial Tree 245
ARCHAEBACTERIAL PHYLOGENY 247
Unusual Nature of the Archaebacterial Phenotype 247
Definition of the Major Archaebacterial Groups 247
Methanogens 248
Extreme halophiles 248
Extreme thermophiles 249
The Archaebacterial Tree: Its Branching Order and Root 249
Main branches 249
Thermococcus and *Thermoplasma* 251
Branching order among the methanogens and extreme halophiles 251
A BRIEF LOOK AT EUCARYOTE PHYLOGENY 252
Implications of the Eucaryote Phylogenetic Tree 252
NATURE OF THE EVOLUTIONARY PROCESS IN BACTERIA 253
Relationship Between Evolution's Tempo and Its Mode 253
Macroevolution at the Molecular Level: the Mycoplasmas and Their rRNAs 253

Dynamics of Variation in rRNA 255
Why mycoplasma rRNAs are so unusual 256
Summary 256
Both Tempo and Mode of Bacterial Evolution Are Reflected in rRNA 256
Macroevolution in mycoplasmas is chronic 256
Conditions for true macroevolution among bacteria 257
Basic Principles of Bacterial Evolution 257
Implications for Bacterial Taxonomy 257
EVOLUTION OF THE TWO PROCARYOTIC PHENOTYPES 258
Archaebacterial Evolution 258
Ancestral phenotype 258
Evolution of methanogenic and halophilic phenotypes 258
Archaebacterial ribosomes and their evolutionary implications 259
Summary 259
Eubacterial Evolution 259
Photosynthesis 259
Thermophilia 260
Conclusion 260
Differences between Archaebacteria and Eubacteria 260
Cytological and physiological differences 260
Niches 261
Metabolic versatility 261
Molecular plasticity 261
Rates of evolution 261
Summary 261
Why Are Archaebacteria and Eubacteria So Different? 261
Primitiveness of Archaebacteria 262
When Did Procaryotes Evolve? 262
THE UNIVERSAL ANCESTOR 262
Progenotes 263
Characteristics of the progenote 263
Was the Universal Ancestor a Progenote? 263
ACKNOWLEDGMENTS 264
LITERATURE CITED 264

PERSPECTIVE

A revolution is occurring in biology: perhaps it is better characterized as a revolution within a revolution. I am, of course, referring to the impact that the increasingly rapid capacity to sequence nucleic acids is having on a science that has already been radically transformed by molecular approaches and concepts. While the impact is currently greatest in genetics and applied areas such as medicine and biotechnology, its most profound and lasting effect will be on our perception of evolution and its relationship to the rest of biology. The cell is basically an historical document, and gaining the capacity to read it (by the sequencing of genes) cannot but drastically alter the way we look at all of biology. No discipline within biology will be more changed by this revolution than microbiology, for until the advent of molecular sequencing, bacterial evolution was not a subject that could be approached experimentally.

With any novel scientific departure it is important to understand the historical setting in which it arises—the paradigm it will change. Old prejudices tend to inhibit, distort, or otherwise shape new ideas, and historical analysis helps to eliminate much of the negative impact of the status quo. Such analysis is particularly important in the present instance since microbiologists do not deal with evolutionary considerations as a matter of course and so tend not to appreciate them. Therefore, I begin this discussion with a brief look at how the relationship between microbiology and evolution (i.e., the lack thereof) developed.

A Fruitless Search and Its Consequences

Microbiologists of the late 1800s and early 1900s were certainly as cognizant of evolutionary considerations as any biologists. They assigned as much importance to determining the natural (evolutionary) relationships among bacteria as zoologists and botanists did to determining metazoan genealogies. From Beijerinck to Kluyver to van Niel, a main concern of the Dutch school, perhaps the dominant force in microbiology in the first half of this century, had been these natural relationships. And it must have been the hope of someday knowing them that inspired the founders of *Bergey's Manual* to adopt for bacteria the same classification system used to group animals and plants phylogenetically.

The search for a bacterial phylogeny, however, became mired in failure, generations of it. Animals and plants are rich in complex morphological detail, which served as the basis for their phylogenetic classification. Bacteria, on the other hand, have such simple morphologies that these are of no use in defining their phylogeny. Bacterial physiologies are more useful in this regard, but are still too limited; while shared physiological traits often correctly group bacterial species, relatives lacking the trait in question usually exist as

well (56). Some of the early microbiologists, realizing the pitfalls in attempting to classify bacteria by the then available criteria, avoided the area, which served to deemphasize the role of evolutionary considerations in the development of microbiology. Those who did concern themselves with phylogenetic classification of bacteria created distorted, basically flawed schemes which confused rather than resolved problems and ultimately discredited the whole attempt. (One of the delightful absurdities to emerge from this involuted system is the taxon *Pseudomonas*, perhaps the best known, most studied, and most pedagogically utilized "representative genus," which actually is a collection of at least five separate groups of bacteria [53], whose name derives from the Greek *pseudes* and *monas*, i.e., false unit!)

The situation seems to have reached a watershed during the time of C. B. van Niel. Initially a leader in attempts to determine microbial phylogenies (210), van Niel, in apparent frustration, ultimately gave up on them, settling for a determinative classification system the basic purpose of which was to identify species (226). Van Niel was perhaps the last microbiologist to treat the matter of bacterial evolution seriously.

Without a capacity to determine bacterial genealogies, considerations of bacterial evolution are mere cerebral exercises. Realizing this, students found it difficult to become enthusiastic about the subject. There was no way microbiology textbooks could treat it seriously either. Many did not treat it at all! Moreover, the failure to determine evolutionary relationships seemed to generate the feeling that it was not important to do so. Bacterial evolution was all but forgotten. All that remained to represent this once vital and important area of microbiology was a formal and unappealing bacterial taxonomy: one falsely authoritative in its borrowed use of the Linnaean (phylogenetic) classification system, stultifying in its liturgy, and caught up in classification for classification's sake.

The result of all this was that microbiology worked from a paradigm that for all intents and purposes was devoid of evolutionary concepts. They played no role in its pedagogy; they had no influence on the design and interpretation of experiments; they were not a part of its value structure. Roger Stanier, one of the few microbiologists who maintained any interest at all in bacterial evolution, captured the spirit of the times with this piquant proscription (written about 1970): "Evolutionary speculation constitutes a kind of metascience, which has the same intellectual fascination for some biologists that metaphysical speculation possessed for some mediaeval scholastics. It can be considered a relatively harmless habit, like eating peanuts, unless it assumes the form of an obsession; then it becomes a vice" (209). That microbiology had reduced evolutionary matters to the status of dalliance was indeed unfortunate, for much of what is important and interesting about evolution lay hidden in the microbial world.

Fortunately, nucleic acid sequencing technology today makes bacterial phylogeny a tractable problem. In fact, all phylogenetic relationships can now be determined much more easily and in far more detail and depth than was ever dreamed possible (116, 174). Microbiology is consequently being inundated with sequence information, which accumulates so rapidly that the reading and entering of data are becoming major concerns, while the actual sequencing operations will soon cease to be rate-limiting factors. The data are in a form [pages of A's, C's, G's, and T(U)'s] alien to most microbiologists, their analysis is arcane, and phylogenetic conclusions tend to be presented in a take-it-or-leave-it manner. It is understandable that some microbiologists distrust these conclusions. However, it is not permissible to ignore them. Phylogenies derived from sequence analysis have to be accepted for what they minimally are: hypotheses, to be tested and either strengthened or rejected on the basis of other kinds of data (97).

Microbiology is now at a point at which it has to ask why it should concern itself with bacterial evolution and what such concern would mean to its future. The emerging bacterial phylogeny cannot be viewed as having merely local impact, i.e., as being a revision of existing bacterial taxonomy. At the very least the existing taxonomy will be totally rewritten by what is currently happening. Even that, however, is only the tip of the iceberg. Phylogenetic perspective will affect the microbiology paradigm throughout. This is already apparent in the change in perspective accompanying the discovery of archaebacteria. Bacteria will no longer be conceptualized mainly in terms of their morphologies and biochemistries; their relationships to other bacteria will be central to the concept as well. Design and interpretation of experiments will be significantly changed. Microbial biochemistry will be conceptualized more in a comparative way. Medical microbiology will have a broadened perspective. Phylogenetic considerations will increase the microbiologist's interest in microbial ecology and shape his approach to it (157, 164). Perhaps the most significant change will be the altering of our perception of the relationship between procaryotes and eucaryotes and, therefore, of the position microbiology holds in relation to the other biological disciplines.

The evolutionist, too, needs to concern himself with the effect that opening the "Pandora's box" of bacterial evolution is going to have, for although he is exquisitely aware of evolution as it is encountered among the metazoa, the world of forms and fossils, bacterial evolution is as alien to him as it is to the microbiologist. Determining microbial phylogeny is not simply the long awaited completion of the "Darwinian programme," the extension of evolutionary study to all life on this planet. Rather than its providing the few missing pieces in the great puzzle of evolution, bacterial evolution in effect is the puzzle. It increases the current time span of evolutionary study by almost an order of magnitude. It holds the key to the origin of the eucaryotic cell. It shows the evolutionist an intimacy between the evolution of the planet and the life forms thereon that he has never before experienced and which, consequently, will lead to a close relationship between the geologist (not merely the paleontologist) and the evolutionist. It will redefine the classical question concerning the connection between evolutionary rate and the quality of the resulting change (the so-called tempo-mode problem) in new and more powerful terms (see below). Even the way we conceptualize selection, the roles of positive versus negative selection, may be changed. In other words, bacterial evolution will show us that, far from approaching the culmination of evolutionary study, where one refines existing concepts and fills in the details, we are only at its beginning.

The impact of the study of bacterial evolution should be felt throughout all of biology, bringing about major shifts in emphasis. The nature of the universal ancestor (which, if it is given attention at all today, is seen as just another common ancestor of a group) will be recognized for the important biological problem that it is. A major effort will be mounted to define the ancestral gene families. And the historical dimension will become a significant, useful part of the study of macromolecular structure. Because its conceptual base

now rests outside of biology per se (in physics and chemistry), biology's interests, its thrust, have tended increasingly to be defined by external factors, many even extrascientific. Biology has become very much an "other directed" discipline. What a renewed and broadened interest in evolutionary questions can and will do is to restore to biology an internally defined sense of direction.

Some restructuring at a metaphysical level is even a possibility. Biology's base, scientific materialism (that highly reductionist, highly mechanistic picture borrowed wholesale from the 19th century physicist), is a world view to which physicists since Einstein, Bohr, and Schrödinger can no longer subscribe. If there by anything to Whitehead's concept of evolution as basic process (236), the interest in evolution generated by its study in bacteria (and unicellular eucaryotes) may push biology in this direction, toward a process-oriented outlook, an attitude that processes (evolution, development, mind) somehow underlie genes, cells, brains, etc., not the reverse.

Whatever else it is or whatever impact it may have, the study of bacterial evolutionary relationships is central to the historical account of life on this planet. We may lay no claim to a comprehensive understanding of biology until we know this history, at least in its outline. And this is the perspective from which the present review is written.

Three Ideas That Shape Our Concept of Bacterial Evolution

Procaryote-eucaryote dichotomy. The way we look at bacterial evolution is in essence shaped by a few picturesque, strongly held notions. One has to do with the place of bacteria in the spectrum of living systems; the other two concern our picture of how life began.

The prescientific distinction animal-vegetable-mineral is the starting point for our perception of the relationships among living things. Initially, every living thing was thought to be either a plant or an animal. The invention of the microscope, however, revealed a world of unicellular creatures, which because of their enormous and unusual variety caused us to wonder whether they were just very small animals or plants; there seemed to be another basic distinction, between macroscopic and microscopic forms, between multicellular and unicellular life. Haeckel's classic phylogeny, reproduced in Fig. 1, represents an amalgam of these two early views (76), in having three basic categories of living systems: plants, animals, and microorganisms (protists). This classical approach to phylogeny, based essentially on characteristics of the whole organism, has been refined in modern times to a five-kingdom scheme (139, 240, 241). However, such a taxonomy is not phylogenetically valid (see below).

It is intuitively evident that certain of an organism's characteristics, especially cellular attributes, are more essential than others. In the 1930s, E. Chatton (25) sought to construct a universal phylogeny on this principle by dividing the living world into two main groups, eucaryotes and procaryotes, on the basis of whether or not they possessed a true nucleus, i.e., one circumscribed by a nuclear membrane. (He also involved a few other intracellular structures found only in eucaryotes, such as the mitochondrion, to bolster the case [25].) Chatton's approach had clear virtues, and once its two basic categories were defined in detail, through electron microscopic studies and various molecular characterizations, the procaryote-eucaryote dichotomy became firmly (dogmatically) established as the primary phylogenetic distinction (138, 152, 209, 211).

However, the original definition of the procaryote carried an implication that no biologist recognized at the time, an implication that had profound consequences. Procaryotes were initially defined in a purely negative sense: they did not have this or that feature seen in eucaryotic cells. There was no logical reason to assume, therefore, that all procaryotes (all cells that were not eucaryotes, that is) were specifically related to one another. Yet, this is precisely what happened; "procaryote" was taken from the start to be a phylogenetically coherent taxon (209).

Over the years the definition of the procaryote (vis à vis the eucaryote) expanded from the initial negative one, based solely upon noncomparable characteristics, to a positive one, based upon comparable, molecular properties. In that process one procaryote, *Escherichia coli*, was assumed to represent all. Now we see the error in this never-tested assumption. Its unquestioned acceptance probably delayed the discovery of archaebacteria by well over a decade.

Oparin Ocean scenario. The old notion that the living world is somehow distinct from, unconnected to, the nonliving world (an idea that reflects creation myths, the mind-matter dichotomy, and other such things) gave rise to the panspermia notion. To counter both this idea and creationism, biologists were obliged to account for the origin of life in physical processes occurring on this planet. The first reasonably comprehensive attempt to do this was made by A. I. Oparin in the 1920s (159); a similar, but significantly different, proposal was made somewhat later by J. B. S. Haldane (77); and a slightly modified amalgam of the two is the only account of life's beginnings accepted by biologists today (104, 149, 160). The Oparin Ocean scenario, as it has been called, is frankly a "Just So Story." It is a vague and so not too useful hypothesis in its own right, and to the extent that biologists treat it as dogma, its scientific effectiveness is further diminished (249).

According to the current Oparin Ocean scenario, the primitive oceans became the ultimate repository for the great variety of chemicals and biochemicals thought to have been produced in the primitive anaerobic (at least, nonoxidizing) atmosphere through the action of ultraviolet light and electrical discharge upon water vapor, carbon dioxide, nitrogen, and other gases. In this way the primitive ocean became a "soup" of energy-rich biochemicals (104, 149, 160). Interactions among these produced ever more complicated structures, which eventually (somehow) turned into complex living (self-replicating) cellular entities (160). Since these earliest living systems arose in a bath of nutrients, they had no need to synthesize amino acids, nucleotides, and other products of intermediary metabolism and so did not develop such capacities. In other words, the first organisms were extreme heterotrophs, having neither photosynthetic nor autotrophic capabilities (160). They were in essence sinks for the chemical energy stored in the oceans (249). Only when these early cells began to exhaust their oceanic supply of nutrients did a need arise for intermediary metabolism, autotrophic capacity, and the ability to use light as an energy source, and only then did these features evolve, deus ex machina (88, 160).

This scenario has been interpreted by biologists to mean that the first organisms were (anaerobic) heterotrophic *procaryotes*, which later spawned photosynthetic and autotrophic sublines (and later still, procaryotes capable of aerobic metabolism). Textbooks customarily derive all living forms from an ancestral anaerobic fermentative heterotroph, taken to be some clostridium or streptococcus (15, 139). Eucaryotes are generally brought into this picture through a

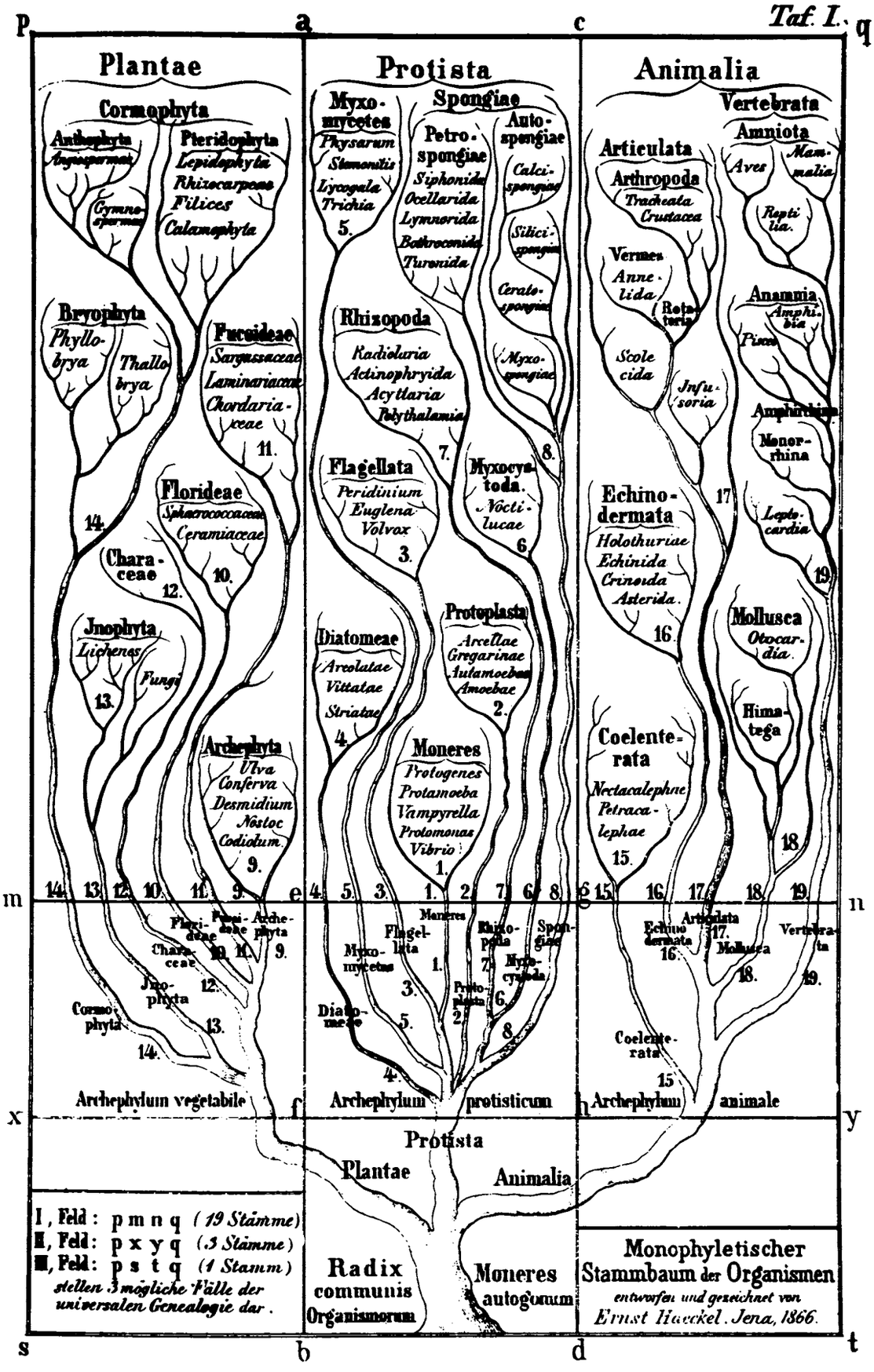

FIG. 1. Haeckel's phylogenetic tree of 1866 (76).

subline of wall-less (anaerobic) procaryotes that gained the capacity for endocytosis (138, 139). By ingesting other procaryotes that gave it additional metabolic capabilities (and molecular structures) and that eventually degenerated into organelles (contributing genes to the nucleus in the process), this primitive host cell, this urcaryote (256), became the eucaryotic cell.

Once again implicit assumptions have misled us, in this case the connotation of the prefix "pro-" in procaryote. *Pro*caryotes had to precede *eu*caryotes; *pro*caryotes are ipso facto older, simpler, and more primitive than *eu*caryotes and therefore gave rise to them. This is perhaps the main, if not the only, reason why the vast majority of biologists perceive the first organisms as procaryotic.

Darwin's warm little pond. A final, related element in the prejudices shaping our view of bacterial origins is Darwin's "warm little pond" image (31), which I am sure Darwin never intended to be a prescription for life's beginnings. (Darwin understood that the subject belonged to the future and probably intended more to dismiss it with this casual remark [31] than to give his successors a guiding principle.) Nevertheless, we are now stuck with this image of life's beginnings and have to cope with it as we do the other conceptual baggage we have inherited. Do microbiologists perhaps view thermophilic bacteria as adaptations from mesophilic species for this reason? "Warm" is an anthropocentrism. The setting in which bacteria arose may well have been warm, but it was not the hospitable warmth implicit in the pond Darwin pictured.

The collection of images associated with the procaryote-eucaryote dichotomy, the Oparin Ocean scenario, and, to a lesser extent, Darwin's warm little pond form the starting point and dictate the direction of our thinking about bacterial evolution. The basic flaws in all of them will become even more apparent as we proceed. It would be better, if that were possible, to forget about the lot and approach bacterial evolution with a clean slate. Molecular phylogenetic studies of bacteria are going to tell us a great deal, especially if our thinking is unfettered by old anthropocentric notions.

MEASUREMENT OF BACTERIAL PHYLOGENETIC RELATIONSHIPS

Three-Dimensional and One-Dimensional Characters

We intuitively recognize that complex characteristics (phenotypic patterns) are good indicators of relationship, i.e., and that a sufficiently complex pattern (character) is unlikely to have evolved more than once. However, our assessment that two or more organisms have the same or similar complex characteristics is by no means fail-safe. The judgment "similar" is too often subjective; what appears complex to our eye may not be so in the dynamics of the organism, and something that the biologist imagines as difficult to evolve may in reality be relatively simple (due to constraints on the system that he has not recognized, for example). Nowhere are our failings in this regard more evident than in the attempts to classify bacteria.

The sequencing of proteins and nucleic acids provides a new and more powerful approach to measuring evolutionary relationships and a new way of looking at them, in terms of the "evolutionary clock" (243). Genotypic information, i.e., sequence information, is superior in two main ways to phenotypic information, the classical basis for relating and classifying organisms: sequence information is (i) more readily, reliably, and precisely interpreted and (ii) innately more informative of evolutionary relationships than phenotypic information is.

Unlike three-dimensional phenotypic patterns, a sequence pattern is one dimensional. One-dimensional patterns can be measured in simple ways, in terms of simple relationships. The elements of a sequence, nucleotides or amino acids, are restricted in number and well defined (quantized). The subjectivity that goes into the judgments "same," "similar," etc., at the phenotypic (three-dimensional) level is replaced by simple, more objective judgments and mathematically defined relationships in the world of sequences.

The evolutionary clock. The introduction of genetics into our model of the evolutionary process in the early part of this century was a major advance in that it let us understand evolution's "motor," the source of the variation upon which selection works. The concept of the evolutionary clock furthers this understanding; it shows us the relationship between this motor (i.e., genotypic change) and what we classically call evolution, the changes in phenotype. At the level of the genotype, change constantly occurs. However, most of it is of a nature that it is not acted upon by selection (105, 106). It therefore, becomes fixed randomly in time, making its characterization as "clocklike" in occurrence appropriate. In other words, evolution has a tempo that is quasi-independent of its mode (the selected changes occurring in the phenotype). An analogy to a car and its motor is apt: a car does not go unless its motor is running, but the motor can run without the car moving.

Cytochrome *c* evolution provides a good example of the evolutionary clock. An enormous number of different versions of this sequence all appear to be equivalent functionally (44). (Formally speaking, the mapping from genotype to phenotype [upon which selection acts] is degenerate.) A change from one such version to another would then occur randomly, independent of selection; the probability of its occurrence would only reflect a lineage's mutation rate (105, 106). Since the number of possible functional configurations for a given gene is enormous by any standards, similarity at the genotypic level (i.e., extensive sequence homology) can never reflect convergent evolution. Consequently, cytochrome *c* sequence comparisons have been used very successfully to time key events in eucaryotic evolution (in sequence distance terms) and to determine molecular genealogies. Phylogenetic trees based upon cytochrome *c*, or similar molecular chronometers, represent significant improvements over their classical counterparts based upon phenotypic comparisons (49).

A missed opportunity. The now classic paper of Zuckerkandl and Pauling, in 1965, effectively launched biology into the world of molecular chronometers (283). Although they may not have been the first to recognize that sequence comparisons could be used to define phylogenetic relationships (protein-sequencing technology had been in existence for about a decade by then), Zuckerkandl and Pauling were the first to put the case in a well-defined, scientifically effective way. This was precisely what the microbiologist needed to resolve the problems of the natural relationships among bacteria. However, microbiologists did not rush to utilize the new approach; remember, they had by then come to see the problem as unimportant. Nevertheless, a peripheral sort of awareness existed; the new molecular techniques were perceived by some as useful in standard bacterial classification, and a number of small-scale efforts were mounted to improve classification by various molecular approaches. Genetic characterization, deoxyribonucleic acid (DNA) base ratios, nucleic acid hybridization studies,

cell wall analyses, and (a little) protein sequencing began to reveal phylogenetically valid groupings (4, 140, 183, 188).

These early molecular approaches, though useful, were not powerful enough to reveal the higher bacterial taxa, and in any case conventional wisdom did not perceive doing so as important. While gram-positive cell walls exhibited interesting and informative variety in their composition, the gram-negative ones were too uniform to be of much use in defining taxa (183). Nucleic acid hybridization work, i.e., DNA/DNA studies, were (necessarily) confined to relationships within genera (185, 207). When DNA/ribosomal ribonucleic acid (rRNA) hybridization studies were instituted, they were used only to revise existing local taxonomic structure (32, 33). Even protein sequence comparisons provided no insights, except in one particular group of purple bacteria (4, 148).

Nature of Molecular Chronometers

A molecule whose sequence changes randomly in time can be considered a chronometer. The amount of sequence change it accumulates (formally a distance) is the product of a rate (at which mutations become fixed) × a time (over which the changes have occurred). The biologist cannot measure this change, however, by comparison of some original to some final state, since the original state (ancestral pattern) is not accessible to him. Instead, he uses the fact that two (or more) versions of a given sequence that occur in extant representatives of two (or more) lineages have ultimately come from the same common ancestral pattern, and so measures the sequence difference between the two (or more) extant versions, which is roughly twice the amount of change that each lineage has undergone (assuming comparable rates of change in each) since they last shared a common ancestor.

All sequences are not of equal value in determining phylogenetic relationships. To be a useful chronometer, a molecule has to meet certain specifications as to (i) clocklike behavior (changes in its sequence have to occur as randomly as possible), (ii) range (rates of change have to be commensurate with the spectrum of evolutionary distances being measured), and (iii) size (the molecule has to be large enough to provide an adequate amount of information and to be a "smooth-running" chronometer (explained below).

Clocklike behavior. A molecular chronometer should measure, should be representative of, the overall rate of evolutionary change in a line of descent. One might think that the best chronometer would then be a genetic segment upon which there are no selective constraints. Its changes would occur randomly along the length of the segment, occur in a quasi-clocklike fashion, become fixed at a rate equal to the lineage's mutation rate (105, 106), and be easy to interpret. However, such sequences are of little value for phylogenetic measurement, because they generally do not meet the second requirement; their rates are so rapid that they cover only very restricted phylogenetic ranges (38). Such sequences are evolutionary stopwatches; they measure only the short-term evolutionary events. The third (i.e., degenerate) codon positions in structural genes are often used in this capacity (20, 128, 243).

The more useful molecules for phylogenetic measurement all represent highly constrained functions. Some sequences of this type change slowly enough to span the full evolutionary spectrum. Unfortunately, what makes them useful as chronometers also makes them problematical. Strict clocklike behavior is usually hard to find, i.e., identify. Unless functional constraints remain strictly constant over the evolutionary range being covered, nonrandom (selected) sequence changes accumulate, over and above the randomly introduced ones, and artificially increase phylogenetic distances between organisms, which usually leads to improper determinations of branching orders. Cytochrome *c* provides an example. In the α subdivision of the purple bacteria (defined below) the molecule changes in size, from "medium" to "large" (37), reflecting some unknown and subtle functional change. This size change has brought with it additional nonrandom sequence changes that appear to distort the phylogenetic determination somewhat (4). A similar situation can be seen in the phylogeny of certain *Bacillus* species constructed from 5S rRNA sequence comparisons (87).

A second problem with highly constrained chronometers is the extremely different rates at which the various positions in a sequence tend to change. This in itself is not a problem; the hands of a clock move at very different rates. However, analysis of the data becomes difficult at this point (see below).

Phylogenetic range. The world of bacterial evolution is vast in comparison to that of the eucaryotes with which we are familiar, i.e., the metazoa. A billion years is a relatively short time in bacterial evolution. Therefore, the range of chronometers used to measure phylogenetic relationships among bacteria needs to be considerably greater than what is optimal for metazoa. Cytochrome *c* is an excellent chronometer for measuring much of eucaryote phylogeny (49). Among the eubacteria, its effective range is restricted to the subdivision level; it orders the α-purple bacteria (see below), but does not relate these accurately to any other subdivision of the purple bacteria (4, 148).

Size and accuracy. In addition to the obvious need for large size in a chronometer (i.e., good statistics), there is a more subtle one. Size per se is probably not what is important. It is that the molecule consists of a fairly large number of loosely coupled "domains" (functional units), regions that are somewhat independent of one another in an evolutionary sense (250). In this case, nonrandom changes affecting one of the units will not appreciably affect the others; therefore, when one part of the chronometer becomes drastically altered by introduction of selected changes (i.e., gives a distorted reading), the other parts remain practically unaffected. The more units of this kind a chronometer contains, the less sensitive are its measurements of evolutionary distances to nonrandom changes in one of them, i.e., the "smoother" the chronometer runs. This is a major difference between 5S rRNA, for example, and the large rRNAs (250).

rRNAs, the Ultimate Molecular Chronometers

Why they are so good. rRNAs are at present the most useful and most used of the molecular chronometers. They show a high degree of functionally constancy, which assures relatively good clocklike behavior (250). They occur in all organisms, and different positions in their sequences change at very different rates, allowing most phylogenetic relationships (including the most distant) to be measured, which makes their range all-encompassing. Their sizes are large and they consist of many domains. There are about 50 helical stalks in the 16S rRNA secondary structure and roughly twice that number in the 23S rRNA (75, 155), which makes them accurate chronometers on two counts.

Perhaps the most compelling reason for using rRNAs as chronometers is that they can be sequenced directly (and,

therefore, rapidly) by means of the enzyme reverse transcriptase (116, 174). This distinguishes them from all other potential chronometers in the cell except for a few of the smaller RNA species (which do not have as good chronometric properties, and in some cases cannot be isolated with the ease with which rRNAs can). It is reasonable for a properly equipped laboratory in the future to sequence on the order of 100 16S rRNAs per year.

Oligonucleotide cataloging. Until several years ago it was not feasible to determine complete rRNA sequences. So more than a decade was spent characterizing them in terms of partial sequences, by the so-called oligonucleotide cataloging method. Short oligonucleotides, of lengths up to 20 or so bases, are produced by digestion of 16S rRNAs with ribonuclease T_1 (which cleaves specifically at G residues) (55); a collection of these sequence fragments from a given rRNA constitutes an oligonucleotide catalog: a detailed, complex pattern characteristic of a given bacterial species (55). Comparisons among these catalogs permits phylogenetic groupings to be identified at various taxonomic levels, including the highest (55, 257).

Oligonucleotide catalog data are usually analyzed in terms of binary association coefficients, so-called S_{AB} values (55), defined as the ratio of twice the sum of bases in oligonucleotides (length greater than five) common to two catalogs *A* and *B*, to the sum of all bases (in oligonucleotides of length greater than five) in the two catalogs (55). The relationship between S_{AB} value and percentage sequence similarity cannot be theoretically derived (because the relative rates at which individual positions change are not predictable *a priori*). A plot of percent similarities and corresponding S_{AB} values for a collection of 16S rRNA sequences is shown in Fig. 2. S_{AB} is seen to vary approximately as the fifth to sixth power of percent similarity. The relationship between the two measures is clearly not a precise one, especially below S_{AB} values of about 0.40.

While the cataloging method sufficed to define most of the major bacterial phyla, it generally failed to resolve the branching orders among them or among their subdivisions. (Such distinctions are never easily made, as evidenced by the fact that the animal phyla have been known for a century, but the order of their branching from one another has yet to be determined.) The cataloging approach also ran into difficulties over the branching order of rapidly evolving lines of descent, again a perennial problem. Full sequencing of 16S rRNA has now replaced the earlier oligonucleotide cataloging approach, a development that greatly increases the resolving power of the rRNA chronometer.

Analysis of Sequence Alignments

Given a sequence alignment, which for rRNAs can be constructed in a straightforward empirical manner (75, 260), the question becomes how to analyze it, i.e., how to extract the most phylogenetically useful information. At present three main methods are used for this purpose: (i) distance matrix treeing, (ii) maximum parsimony analysis, and (iii) cluster analysis.

Distance matrix methods. Distance matrix methods (47, 49, 157) utilize only the sequence distances between pairs of sequences, i.e., the fraction of positions in which the two sequences differ. This distance is actually an underestimate of the true evolutionary distance between sequences; although most of the differences between two sequences reflect single mutational events at any given position in a sequence alignment (when the chronometer is working in its effective range, that is), some of them represent multiple events. Were all lineages and all positions in a sequence changing at the same rate, then correction for this effect, conversion of sequence distances to evolutionary distances, would be a relatively simple matter (99). Unfortunately, this is not the case, either for different lineages or for different positions in the sequence, and the proper correction for multiple changes remains a major problem in tree construction.

Distance matrix treeing assumes that evolutionary distances conform to a tree topology. To use a simple example, let AB, AC, AD, BC, BD, and CD be the six determined evolutionary distances among four sequences, A, B, C, and D. If the three pairwise sums AB + CD, AC + BD, and AD + BC meet the condition that two of them are equal and greater than or equal to the third, a condition that can be understood by reference to Fig. 3, then the data fit a tree topology. When this condition is met even approximately, evolutionary distances can be used to reconstruct phylogenies with reasonable accuracy (36, 47, 49, 157).

Given a matrix of (corrected) evolutionary distances for a sequence alignment, one in principle examines all possible phylogenetic trees (branching orders), treating branch lengths as adjustable parameters, and declares the one that best fits the data (by a least-squares analysis) to be the "correct" tree (47, 49, 157). This, of course, is not how things are actually done, for the number of all possible branching orders rapidly becomes computationally unmanageable as the number of sequences in the alignment reaches even a moderate number. There are many approximations, many competing algorithms for giving the "best" tree in a reasonable amount of time (47, 49, 157).

Maximum parsimony analysis. Unlike sequence distance matrix analysis, maximum parsimony analysis does not reduce the differences among sequences to a single number, a distance; it treats the positions individually (47, 48). Its assumption is that the correct phylogenetic tree is the most parsimonious one, the one for which the smallest overall number of mutational changes have to be postulated to arrange the set of considered sequences upon it. (Each branch in the tree, each segment between branching points, is defined by the specific changes that occur [in some ancestral sequence] on that branch.) As with distance matrix treeing, one in principle looks at all possible tree branching arrangements and chooses the most parsimonious one (47, 48). Also, as with distance matrix treeing, the problem of finding the correct tree can be computationally intense, even more so in this case, and much time and effort have gone into devising computer algorithms that do this as efficiently as possible (46–48).

Cluster analysis. Cluster analysis, the third major method of analyzing sequence data, groups sequences on the basis of how similar they are to one another or to other groups of sequences (193, 194). The method is the least computationally intense of the three, but also the least accurate.

The main difficulty in all analyses of sequence data lies in the fact that different lineages and different positions in a sequence can evolve at significantly different rates. Making distance corrections on the assumption that all positions in a sequence change at the same rate (65, 99, 157) underestimates the correction needed. Parsimony analysis tends not to position rapidly evolving lineages correctly and is confused by rapidly changing positions, perhaps more so than is distance treeing (157). Cluster analysis is especially sensitive to these problems; rapidly evolving lineages are as a rule

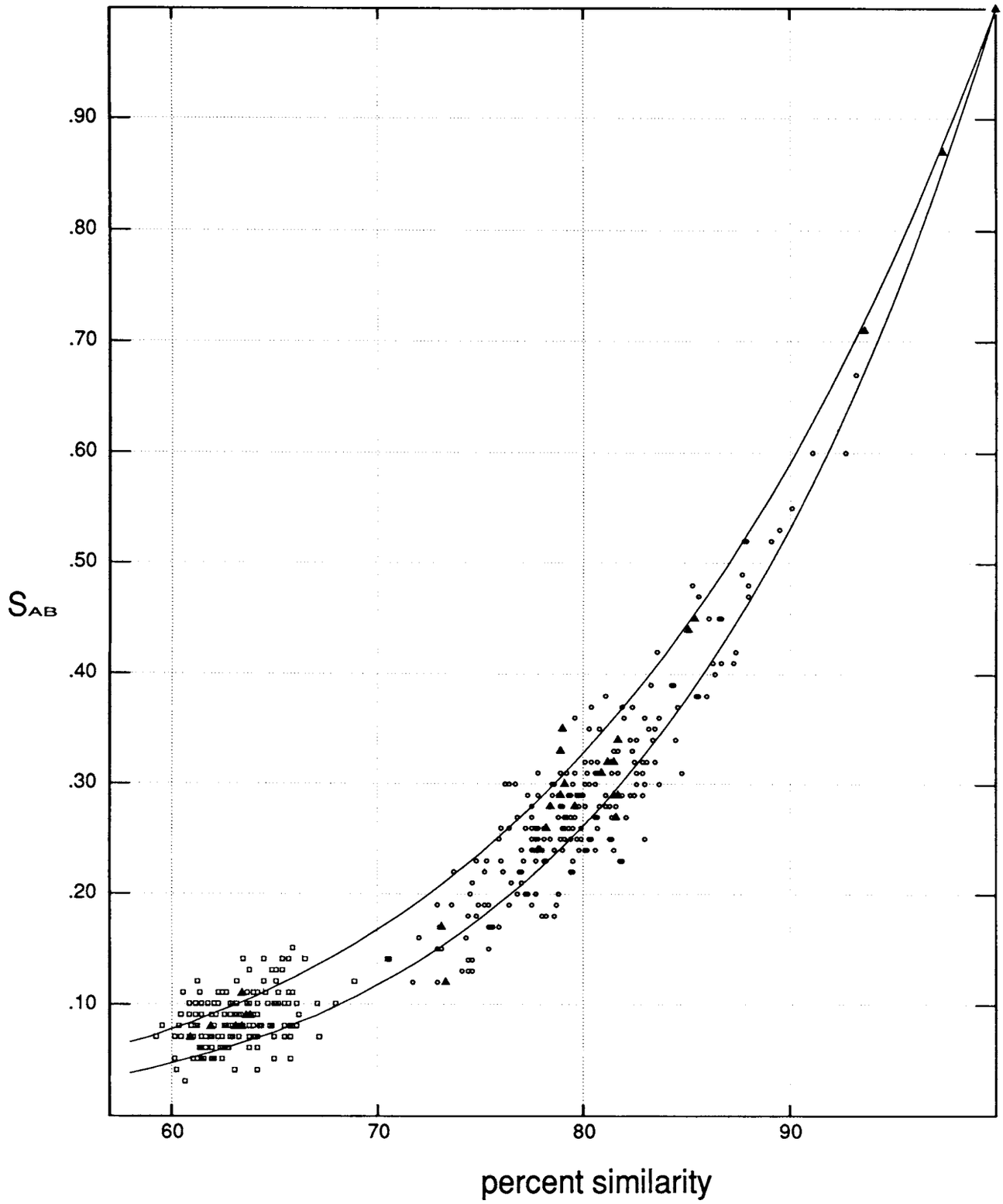

FIG. 2. Plot of percent sequence similarity versus binary association coefficient (S_{AB} value [55]), for a representative sampling of eubacterial and archaebacterial sequences (unpublished analysis). The two theoretical curves are X^5 (upper curve) and X^6 (lower curve), where X = percent similarity. Symbols: ○, values for pairs of eubacterial sequences; □, values for eubacteria with archaebacteria; ▲, values for *E. coli* with either eubacteria or archaebacteria.

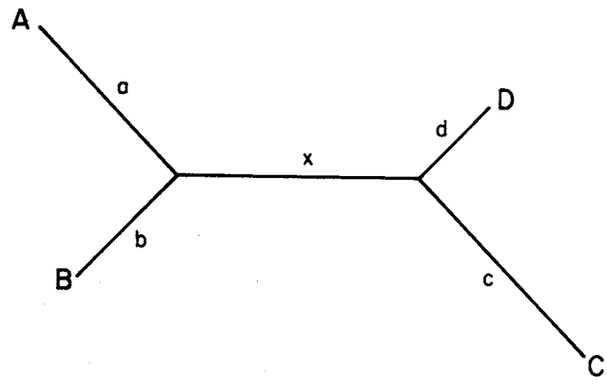

FIG. 3. Unrooted tree for four species, A, B, C, and D, illustrating the relationship that must hold among the six evolutionary distances, AB = a + b, AC = a + x + c, AD = a + x + d, BC = b + x + c, BD = b + x + d, and CD = c + d, for them to fit a tree topology (see text).

positioned too deeply in trees so constructed. In other words, analyses of sequence data today are far from optimal.

The needed improvements in analysis will not come primarily from better theoretical treatments or more efficient algorithms per se. Improvement will be basically the result of empirical approaches. Given a large enough sequence data base, it will become possible to describe the pattern (rate and kind) of change at given positions in a (rRNA) molecule and design specific analyses based upon this description, analyses that correct more accurately for multiple changes at sites and utilize only those positions in the molecule appropriate to the phylogenetic range being measured. (The optimal method will not consider the second hand when timing the seasons.)

Consideration of the detailed chronometric structure of rRNA will be postponed until the reader has some familiarity with bacterial phylogeny. However, the reader who already has this familiarity may wish to read that discussion (which begins on page 253) at this point.

DO BACTERIA HAVE A GENEALOGY AND A MEANINGFUL TAXONOMY?

In classifying bacteria microbiologists make two implicit assumptions: (i) that bacteria have a phylogeny, and (ii) that the taxonomic system that works well for the metazoa is actually applicable to, i.e., meaningful in, the microbial world. These two points require explication and discussion, for they are far from self-evident.

In questioning the first assumption, one at the very least questions approaches to measuring bacterial genealogies and beyond that whether in principle the bacterium as a whole has a genealogy, a unique history. These questions are raised by the existence of lateral (interspecies) transfer of genetic information among bacteria (23, 177). A given gene (or set of genes), say for nitrogen fixation in *Azotobacter* sp., might have been evolved in an organism not immediately related to *Azotobacter* and have been acquired by that organism through plasmid transfer or some similar process. When used as molecular chronometers such genes would not yield the correct genealogy for *Azotobacter*. In the extreme, interspecies exchanges of genes could be so rampant, so broadspread, that a bacterium would not actually have a history in its own right; it would be an evolutionary chimera, a collection of genes (or gene clusters), each with its own history.

Fortunately the matter is experimentally decidable. Were an organism an evolutionary chimera, then its various chronometers would yield different, conflicting phylogenies. A limited test of the possibility can be made for the α subdivision of the purple bacteria, for which a number of species have been characterized by both rRNA catalogs and cytochrome *c* sequences. Phylogenetic trees derived from the two molecules have nearly the same topology, strongly suggesting that neither chronometer has been involved in interspecies gene transfer (258). Although more extensive testing of the lateral transfer notion is highly desirable, it is now relatively safe to assume that bacteria do in principle have unique, characteristic evolutionary histories and that at least some of the cell's chronometers record them. (What is not known is the fraction of the functions in a bacterial cell that are subject to interspecies gene transfer, and which ones these are.)

Given that a bacterial genealogy exists, the question (the second assumption above) then becomes whether such a cladogram can be divided into zones (into taxa) that are naturally, as opposed to artificially, defined and, if so, whether the groupings arrange themselves naturally into simple hierarchical structures (into taxonomic levels).

The metazoan (Linnaean) classification system, though still imperfectly implemented, is as useful as it is because metazoa (chiefly animals) intrinsically group into (naturally defined) categories. The metazoan kingdoms and the animal phyla (192), for example, are readily differentiable groupings. And, although fine points are debated, a metazoan species is also well defined, largely because of the constraints imposed by mating (144). A bacterial species is certainly far more problematical a concept than a metazoan species (8, 29). In the present context, however, our concern is essentially with the higher taxa: whether or not (within a bacterial urkingdom) these are somehow naturally defined or are mere artificial constructs.

There is no compelling evidence to suggest that the bacteria fall into naturally defined taxa. In fact, existing evidence might even suggest the contrary. Few of the (extensively investigated) bacterial phyla presented below can be defined by phenotypic properties common to all members of the group. For example, the gram-positive bacteria defined by cell wall structure form a clade, but this clade also includes bacteria that do not have gram-positive walls. Although the purple bacterial group is named for the particular type of photosynthesis done by some of its members, the photosynthetic pigment does not define the group as a whole, which also includes many nonphotosynthetic species.

Nevertheless, I feel that ultimately bacteria will be shown to fall into naturally defined taxa. One reason this is not obvious at present may be that various bacterial groups have been studied from different perspectives: what we know to be characteristic of one may never have been looked for in another. This fact alone could explain some of the apparent lack of phenotypic resemblance among genealogically clustered species. Another reason is that the microbiologist has never before had phylogenetically defined groupings that he could count on to direct his search for phenotypically unifying characters. Recent studies utilizing such an approach (97) appear promising.

The main reason for thinking that bacterial taxa are naturally defined is that the characteristics of the rRNA chronometer (discussed below) strongly suggest this. Under certain circumstances rRNAs will accumulate unusual sequence changes, ones that normally occur with a negligible

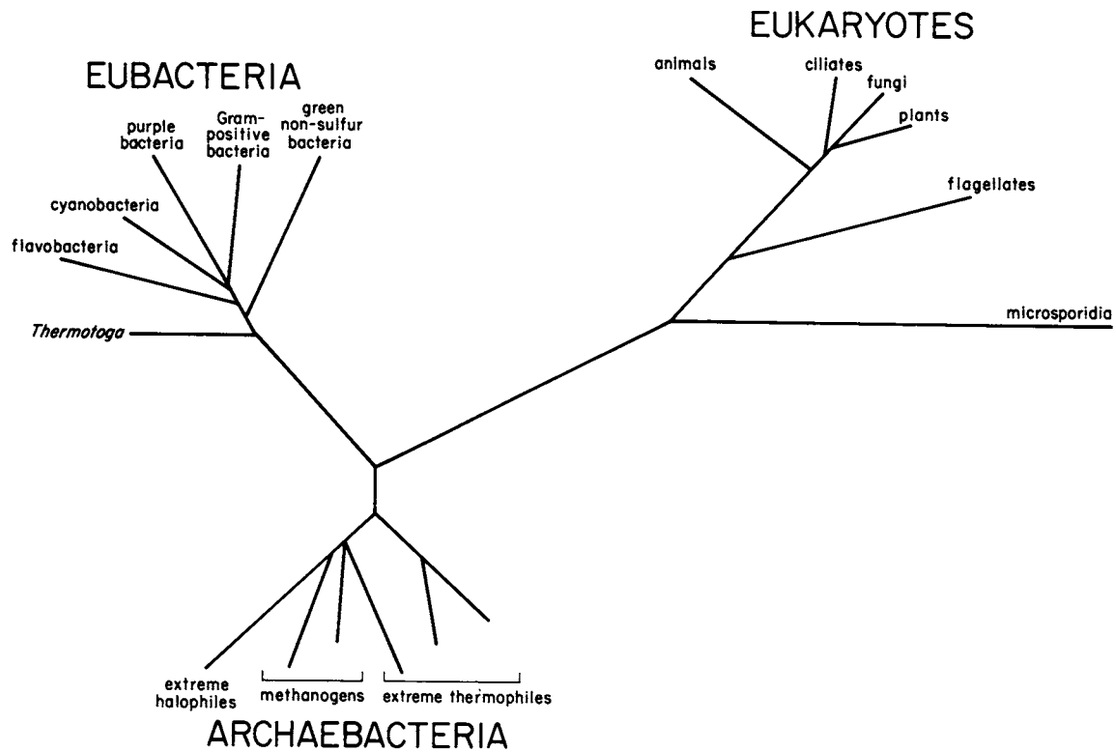

FIG. 4. Universal phylogenetic tree determined from rRNA sequence comparisons. A matrix of evolutionary distances (99) was calculated from an alignment (260) of representative 16S rRNA sequences from each of the three urkingdoms. This was used to construct a distance tree (36), based upon those positions represented in all sequences in the alignment in homologous secondary structural elements (75). Line lengths on the tree are proportional to calculated distances. The alignment includes the following eubacterial sequences: *Thermotoga maritima* (1); green non-sulfur bacteria, *Thermomicrobium roseum* (162); flavobacteria, *Flavobacterium heparinum* (234); cyanobacteria, *Anacystis nidulans* (224); gram-positive bacteria, *Bacillus subtilis* (68); and purple bacteria, *Escherichia coli* (19); the following archaebacterial sequences: extreme halophiles, *Halobacterium volcanii* (72); methanogens, *Methanococcus vannielii* (96) and *Methanobacterium formicicum* (124); and extreme thermophiles, *Thermococcus celer* (Woese et al., unpublished data), *Desulfurococcus mobilis* (R. Garrett, personal communication), and *Thermoproteus tenax* (126); and the following eucaryotic sequences: microsporidia, *Vairimorpha necatrix* (226a); flagellates, *Euglena gracilis* (196); cellular slime molds, *Dictyostelium discoideum* (145); ciliates, *Paramecium tetraurelia* (195); fungi, *Saccharomyces cerevisiae* (179); plants, *Zea mays* (147); and animals, *Xenopus laevis* (181). Branching order within each kingdom is correct to a first approximation only. See the trees for the individual kingdoms for precise branching orders.

frequency. These unusual changes seem to accompany the formation of various major branches on the tree, which correspond to major shifts in bacterial phenotype. The reason for this lies in the tempo-mode relationship. Thus the rRNA chronometer may provide a means of defining natural bacterial groupings that is purely genotypic, independent of any phenotypic definition thereof.

THE UNIVERSAL PHYLOGENETIC TREE

Although we evolutionists still have much to learn a century after Darwin's death, a major milestone has recently been reached. Molecular phylogenetic approaches let us for the first time see the full extent of the tree that encompasses all extant life (72, 263); see Fig. 4. As yet this tree must be drawn in an unrooted form, because the crucial question of the point in its structure that corresponds to the Universal Ancestor, the point from which all extant life ultimately emanates, remains unanswered.

Perhaps the most striking characteristic of the universal tree is the distinctness of the primary kingdoms, the large sequence distances that separate one kingdom from another. Figure 5, a plot of percent sequence similarities for an extensive collection of procaryotic 16S rRNAs, demonstrates this in graphic form; every eubacterial sequence is far closer to every other eubacterial sequence than to any archaebacterial sequence, and vice versa. (The interkingdom distances in Fig. 4 are even best considered lower bounds, for they are large enough that they could well have been underestimated.)

The extent of sequence distance that separates the primary kingdoms is reflected in the degree of phenotypic difference among them. It has long been obvious that eucaryotes are quite distinct from procaryotes (i.e., eubacteria); the two differ in general cellular organization, in genome structure, in control and expression of genetic information, in the structure of the translation apparatus, and in the details of numerous molecular functions. In the last decade we have found the same to be true for the archaebacteria. They too have unique molecular architecture, cellular organization, genome structure, etc. (see assorted chapters in references 102 and 270). The degree of genotypic and phenotypic separation among the primary kingdoms argues that the ancestor they all shared was a special sort of entity (251), whose nature will be discussed below.

Each of the primary kingdoms has its particular form of

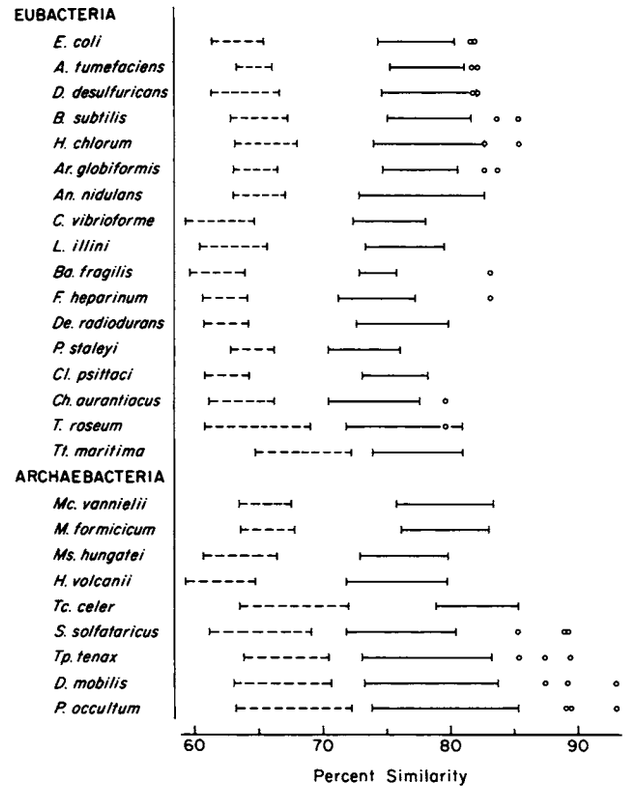

FIG. 5. Percent sequence similarity for various eubacterial and archaebacterial 16S rRNAs. Open circles indicate percent similarity for a given sequence with others in the same phylum. The range of similarities between a given sequence and those from other phyla in the same kingdom is shown as a solid line, whereas the corresponding range across kingdom lines appears as a dotted line. Only those positions represented in all sequences in the composite alignment (in homologous structures) were used in the calculations.

rRNA. Secondary structures for the three types of 16S-like rRNAs can be seen in Fig. 6. While there are general resemblances among them, there are also characteristic differences in structural detail (75).

Strong rRNA sequence signatures, i.e., positions in the molecule that have a highly conserved or invariant composition in one kingdom, but a different (highly conserved) composition in one or both of the others, also define and distinguish the three urkingdoms. Table 1 is a signature distinguishing archaebacteria from eubacteria; the location of these positions in relation to the molecules' secondary structure can be seen in Fig. 7. Though not presented in this review, equally pronounced signatures exist for eucaryotes (75, 260).

Archaebacteria can resemble either eubacteria or eucaryotes (or neither), depending upon what phenotypic characters are considered. For example, the 16S rRNAs of the two procaryotic kingdoms are relatively close in structure (75), whereas 7S RNA sequence and structure reveals a resemblance between archaebacteria and eucaryotes (129, 134, 151; B. P. Kaine, unpublished results). Eucaryotes and eubacteria have similarities as well, e.g., ester-linked lipids (117, 118), but these are relatively few. Our present, rather limited understanding would suggest that the overall phenotypic resemblance is greatest between the archaebacteria and eucaryotes.

What the biologist, especially the microbiologist, must now fully recognize is that there no longer exists any reason to consider that archaebacteria and eubacteria are related to one another to the exclusion of eucaryotes. Unfortunately, the title of this review, "Bacterial Evolution," implies the opposite. The title is intended, however, merely as a celebration of the fact that within the last 10 years the field of bacterial evolution has passed from a suspect discipline, about which almost nothing was known, to a full-fledged area of scientific investigation, rich in its implications for all of biology.

EUBACTERIAL PHYLOGENY

Background

Not only did we know very little about eubacterial phylogeny before the advent of the rRNA approach, but what we thought we knew tended to be wrong. The old idea (justified to some extent by cell wall compositions [183]) that there are two primary categories of (eu)bacteria, gram positive and gram negative, turns out to be a half-truth. Gram positive is indeed a phylogenetically coherent grouping, but gram negative is not. The latter encompasses of the order of 10 distinct groups, each the equivalent of the gram-positive one, as we will see. The old idea that wall-less bacteria, mycoplasmas, are phylogenetically remote from other (eu)bacteria (62, 176) is incorrect; the true mycoplasmas are merely "degenerate" clostridia (see below). Photosynthetic bacteria do not form a grouping genealogically distinct from the nonphotosynthetic bacteria (92, 168, 169); actually, the major photosynthetic types each represent separate high-level phylogenetic units which in most cases include many nonphotosynthetic species as well (56, 64, 206, 266). Autotrophs and heterotrophs are not phylogenetically separate groupings; they are intimately intermixed within the various eubacterial phyla (56, 206). Thus, the textbook views in which photosynthetic (or autotrophic) bacteria arise from nonphotosynthetic heterotrophic ancestors (15, 139) gain no support from the rRNA-based phylogeny. It is the classical microbiologist's insistence on morphology as the primary criterion (108), a prejudice inherited from the botanist, that more than anything engendered the confused and confusing state of bacterial taxonomy; almost none of the taxa (beyond the level of genus) defined primarily in this way pass phylogenetic muster (56).

The misconceptions of the classical microbiologist cannot be condemned. They were innocent attempts to create a phylogenetic framework at a time when phylogenetic classification was not experimentally possible. Unfortunately, many of them are now formalized as accepted bacterial taxonomy: they are repeated over and over in textbooks. They shape the design and interpretation of our experiments. And, because this taxonomy is patterned after metazoan classification, it is a de facto phylogenetic statement. Such an entrenched, organized system will not change easily. Yet replacing the old taxonomy with a phylogenetically valid one is vital to the future development of microbiology.

Major Eubacterial Groups and Their Subdivisions

As of this writing over 500 species of eubacteria have been characterized in rRNA terms. Although most of these characterizations utilized the older and less informative oligonucleotide cataloging method, the more than 50 nearly complete sequences now known are sufficient in number that the

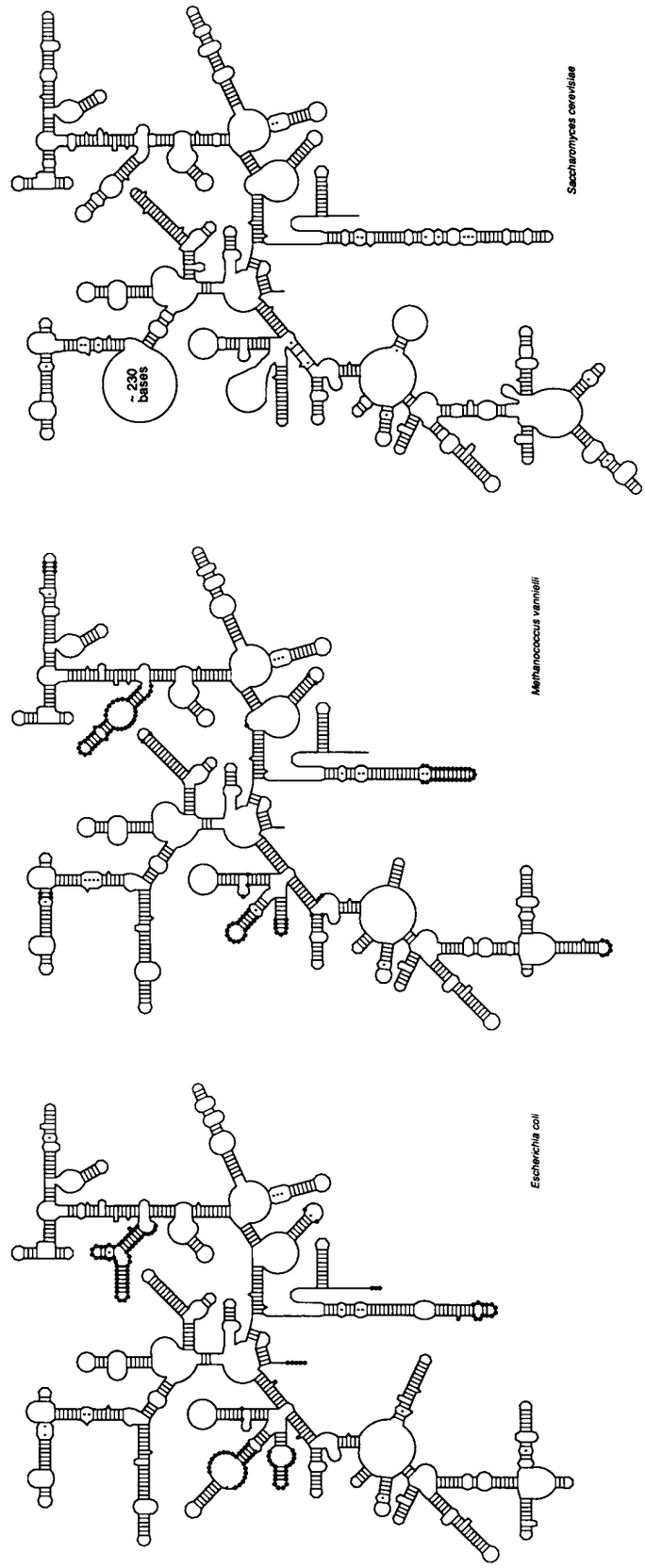

FIG. 6. Representative small-subunit rRNA secondary structures for the three primary kingdoms (75). Dots in the *Escherichia coli* or *Methanococcus vannielii* structures identify the positions or locales where eubacteria and archaebacteria characteristically differ.

TABLE 1. Sequence signature distinguishing the two procaryotic kingdoms[a]

Position[b]	Composition			Position	Composition		
	Eubacteria	Archaebacteria	Eucaryotes[c]		Eubacteria	Archaebacteria	Eucaryotes
9	G[d]	C	C	585	G	C	U
10	A	Y	U	675	A	U	U
24	U	R	A	677	u	C	Y
25	C	G	G	684	u	G	G
31	g[e]	—		707	U	C	Y or G
33	A	Y	A	716	A	C	Y
38	G	A	u	756	C	G	A
39	G	U	g	862	Y	G	a
43	C	R or U	Y	867	G	C	u
47	y	R	R	912	C	U	U
53	R	C	C	923	A	G	A
113	G	C	C	930	C	A	G
314	c	G	G	931	C	G	G
338	A	G	A	934:1	—	Y or A	c
339	C	G	C	952	U	C	C
340	U	C	N	962	C	G	U
349	A	G	R	966	G	U	U
350	G	Y	G	973	G	C	G
358	Y	G	G	975	A	G	G
361	R	C	C	1045	c	G	R
367	U	C	U	1060	U	C	C
377	G	C	Y	1086	U	C	Y
386	C	G	R	1087	G	C	U
393	A	G	A	1109	C	A	A
397	A	—	—	1110	A	G	G
399	G	Y or A	R	1194	U	G	R
403	C	A	A	1197	A	G	G
508	U	C	A	1211	U	G	Y
514	Y	G	G	1212	U	A	a
523	A	C	A	1229	A	G	G
537	R	C	C	1381	U	C	C
540	g	C	C	1386	G	C	C
549	C	U	c	1387	G	U	C
551	U	R	U	1393	U	C	U
558	G	Y or A	A	1415	g	C	C
559	A	U or G	N	1485	u	G	g

[a] Based upon approximately 30 eubacterial and 12 archaebacterial 16S rRNA sequences (19, 68, 224, 263, 274; C. R. Woese et al., unpublished data). Y, Pyrimidine; R, purine; N, any base; —, no base.
[b] Position in 16S rRNA sequence.
[c] Eucaryotic compositions taken from alignment used to construct Fig. 14; upper case, no exceptions; lower case, one exception.
[d] Upper case, Base found in all or all but one sequence in the kingdom but in no more than one sequence in the other kingdom.
[e] Lower case, As for footnote *d* but with two to three exceptions.

conclusions derived from the older approach can be significantly refined and extended.

At the level of oligonucleotide analysis it was apparent that the bacteria separate into more or less naturally defined "phyla" (see discussion below). This was not apparent from the binary association coefficients (S_{AB} values), but could be seen in oligonucleotide signatures (266). (Such signatures are collection of specific oligonucleotides, each of which occurs in most or all members of a given phylogenetic group but rarely, if at all, in most other groups, especially closely related ones [266].) Full rRNA sequences now permit the identification of individual positions in the molecule, sequence signatures that define the various groupings. We will use these rather than the older oligonucleotide signatures in the discussion below.

In some cases taxa defined by rRNA sequences can be identified phenotypically by common characteristics of the group. For the majority of such taxa, however, any given common character will link most, but not all, members of the group. A few of the rRNA groups are without phenotypic justification. The spirochetes and relatives (see below) seem to be a taxon of the first type; all identified members of the genotypically defined unit possess the unique and unusual axial fibrils classically associated with these organisms (24). The gram-positive phylum is an example of the second type; the majority of its species have characteristic gram-positive cell walls (183), but a few lineages, such as *Heliobacterium* (255) and the mycoplasmas (176), do not have them. An example of a genotypically defined unit for which no convincing phenotypic justification can be given is the bacteroides-flavobacter phylum (166, 234); see below. In this case and others like it, the lack of common phenotypic properties could merely reflect the fact that the various subgroups have been characterized in entirely different ways and so would not necessarily be expected to show common characteristics. Such groups challenge the microbiologist to discover their unifying phenotypic motifs.

The eubacterial phyla and their subdivisions as they are

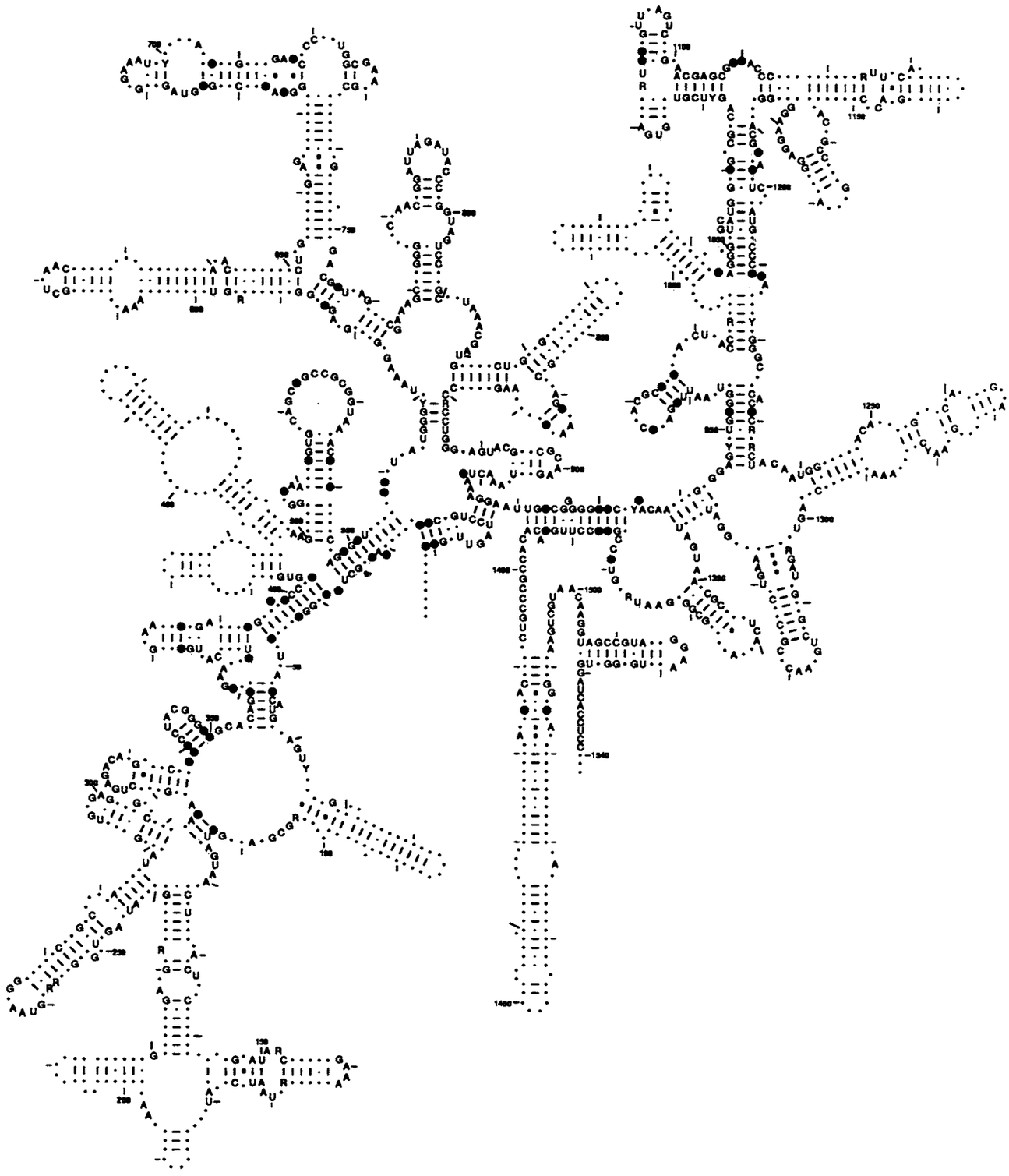

FIG. 7. Location of the signature positions that distinguish eubacteria from archaebacteria (Table 1) on the 16S rRNA secondary structure. The underlying secondary structure is that for *Escherichia coli* (260). The signature positions are indicated by filled circles. Positions whose composition is highly conserved (i.e., the same in over 90% of sequences and oligonucleotide catalogs) between eubacteria and archaebacteria are also indicated (as an aid in orientation); all others (in the *E. coli* sequence) have been replaced by dots. Tick marks show every 10th position in the (*E. coli*) sequences, every 50th being numbered.

now understood are listed in Table 2. Table 3 identifies them by a sequence signature. This key now includes mainly positions covered by the oligonucleotide catalogs, but will ultimately be extended to all positions in the 16S (and 23S) rRNA.

In the following discussion each of the known eubacterial phyla (divisions) will be defined genealogically and its phenotype will be briefly described. As was the case with the kingdoms themselves, three criteria will be used whenever possible to define a phylum and its subdivisions: (i) coherence of the unit by sequence distance analysis (or cluster analysis of S_{AB} values [55]); (ii) definition of the unit by sequence signature; and (iii) characterization of the unit in terms of higher-order structural features of 16S rRNA.

TABLE 2. Eubacterial phyla and their subdivisions[a]

Purple bacteria
- α subdivision
  - Purple non-sulfur bacteria, rhizobacteria, agrobacteria, rickettsiae, *Nitrobacter*
- β subdivision
  - *Rhodocyclus*, (some) *Thiobacillus*, *Alcaligenes*, *Spirillum*, *Nitrosovibrio*
- γ subdivision
  - Enterics, fluorescent pseudomonads, purple sulfur bacteria, *Legionella*, (some) *Beggiatoa*
- δ subdivision
  - Sulfur and sulfate reducers (*Desulfovibrio*), myxobacteria, bedellovibrios

Gram-positive eubacteria
- A. High-G+C species
  - *Actinomyces*, *Streptomyces*, *Arthrobacter*, *Micrococcus*, *Bifidobacterium*
- B. Low-G+C species
  - *Clostridium*, *Peptococcus*, *Bacillus*, mycoplasmas
- C. Photosynthetic species
  - *Heliobacterium*
- D. Species with gram-negative walls
  - *Megasphaera*, *Sporomusa*

Cyanobacteria and chloroplasts
- *Aphanocapsa*, *Oscillatoria*, *Nostoc*, *Synechococcus*, *Gleoebacter*, *Prochloron*

Spirochetes and relatives
- A. Spirochetes
  - *Spirochaeta*, *Treponema*, *Borrelia*
- B. Leptospiras
  - *Leptospira*, *Leptonema*

Green sulfur bacteria
- *Chlorobium*, *Chloroherpeton*

Bacteroides, flavobacteria, and relatives
- A. Bacteroides
  - *Bacteroides*, *Fusobacterium*
- B. Flavobacterium group
  - *Flavobacterium*, *Cytophaga*, *Saprospira*, *Flexibacter*

Planctomyces and relatives
- A. Planctomyces group
  - *Planctomyces*, *Pasteuria*
- B. Thermophiles
  - *Isocystis pallida*

Chlamydiae
- *Chlamydia psittaci*, *C. trachomatis*

Radioresistant micrococci and relatives
- A. Deinococcus group
  - *Deinococcus radiodurans*
- B. Thermophiles
  - *Thermus aquaticus*

Green non-sulfur bacteria and relatives
- A. Chloroflexus group
  - *Chloroflexus*, *Herpetosiphon*
- B. Thermomicrobium group
  - *Thermomicrobium roseum*

[a] Showing representative examples. See appropriate sections in text for references.

Purple Bacteria

What for want of a better name have been called the "purple bacteria" contain most, but not all, of the traditional gram-negative bacteria (50, 51, 56, 64, 163, 254, 266–269). However, the arrangement of classically defined families, genera, and even species within this phylum is a jumbled one. Photosynthetic species group with nonphotosynthetic ones; heterotrophs are paired with chemolithotrophs; anaerobes are paired with aerobes, etc. Because the purple photosynthetic phenotype is distributed more or less throughout the group, and because photosynthesis is complex enough that its having arisen more than once is unlikely, the ancestral phenotype of the phylum is undoubtedly (purple) photosynthetic, which justifies the name purple bacteria. Photosynthetic capacity has been lost many times in this phylum, resulting in various nonphotosynthetic sublines. (The alternative explanation, that photosynthetic capacity was genetically transferred among species, is not supported by the evidence [258].)

The purple bacteria fall rather naturally into four subdivisions, which, awaiting appropriate formal nomenclature, are designated α, β, γ, and δ. Figure 8 is a distance matrix tree for the purple bacteria, based upon five representative sequences in each subdivision. The corresponding tree given by maximum parsimony analysis (not shown) agrees with the branching order shown in Fig. 8, except for some details of branching in the δ subdivision. Table 4 distinguishes the four subdivisions by sequence signature. The purple bacteria as a whole, however, make up one of the few phyla that cannot be defined by a simple signature (see Table 3), although parsimony analysis as well as sequence distance treeing readily define the group (267).

Two helices in 16S rRNA, positions 184 to 193 and 198 to 219 (*E. coli* numbers) (260), help to define and distinguish the four subdivisions; see Fig. 9. While the structure of these helices can vary drastically between subdivisions, within a given subdivision each remains constant to a first approximation. The first of the two helices, i.e., positions 184 to 193, contains only 3 base pairs in α, β, and γ subdivision sequences, but about 10 base pairs in the δ subdivision sequences. The short form of the helix is rare. So far it has been found outside the purple bacteria in only two other phyla, the cyanobacteria (224) and the planctomyces (H. Oyaizu, unpublished data). In all other cases, including the archaebacteria (72, 96, 124, 126, 158), a much longer version of the helix occurs, which therefore is likely to be its ancestral form.

The second helix, positions 198 to 219, contains approximately 8 base pairs in all, except the α-subdivision sequences, wherein 2 base pairs only make up the stalk (260, 274). However, for this helix the short version is common among eubacteria and even occurs in archaebacteria, while the longer version is rare (among eubacteria). The long version has a common structure in most sequences from the β and γ subdivisions, but the characteristic δ-subdivision (long) version differs from this structure in detail (see Fig. 9).

The (near) constancy of structure for each helix within a given subdivision can be inferred from oligonucleotide catalog data. For example, 80% of α-purple bacterial catalogs

TABLE 3. 16S rRNA sequence signature for the eubacterial phyla and their subdivisions[a]

Position	Consensus composition	Purple bacteria				Gram-positive bacteria		Cyanobacteria	Green sulfur bacteria	Spirochetes		Bacteroides		Planctomyces	Deinococcus	Green non-sulfur bacteria
		α	β	γ	δ	Lo	High			Spirochetes	Leptospiras	Bacteroides	Flavobacteria[a]			
47	C	•[b]	•	•	•	•	•	•	•	U[c]	U	•	•	G	•	•
48	Y	•	•	•	•	•	•	•	•	•	•	•	•	A	•	•
50	A	•	U	•	•	•	•	•	•	U	U	•	•	U	•	•
52	Y	•	•	•	•	•	•	•	•	A	A	•	•	G	G	•
53	A	•	•	•	•	•	•	•	•	G	•	•	•	G	•	G
353	A	•	•	•	•	•	•	•	•	•	•	•	•	U	•	•
570	G	•	•	•	•	•	•	•	•	•	•	U	U	U	•	•
698	G	U	•	•	Y	•	•	•	•	•	•	•	•	•	•	•
812	G	•	C	•c[d]	•	•	•	•	•	•	C	•	•c	•	C	•
906	G	Ag	Ag	Ag	Ag	•a	A	•	•	•	•	•	•	•	•	A
933	G	•	•	•	•	•	•	•	•a	•	•	•	•	A	•	•
955	U	•	Au	•	•	•	AC	•	•	•	•	•	•	C	•	•
976	G	•	A	•	•	•	•	•	•	•a	A	•	•	•	•	•
983:1	—[e]	—	—	—	—	—	—	—	—	—	—	—	—	U	—	—
995	C	•	•	•	•	•	•	•	(A)[f]	•	•	A	A	•	•	•
1109	C	•	•	•	•	•	•	•	•	•	•	•	•	Ac	•	•
1198	G	•	•	•	•	Ag	Ag	•	•	•a	•a	A	Ag	•	•	•
1202	U	•	•	•	•	•	•	•g	•	•	•	G	•	•	g•	G
1207	G	•	•	•	•	C	C	C	•	•	•	•	•	•	•	•
1208	C	•	•	•	•	•	•	•	•	•	•	•	•	•	u•	•
1224	U	•	•	•	•	•c	•a	•g	YA	•c	UA	•g	•c	•	•	G
1229	A	•	•	•	•	•	Ga	•	•	•	•	•	•	•	•	•
1233	G	•	Ca	•	•a	a	•	A	•	•a	•	•	•	•c	•	•
1234	C	•	A	•	•a	•	•	•	A	•a	•	U	•	•	•	•
1384	C	•	•	•	•	•	•	•	•u	•	•	•	•	U	•	•
1410	A	•	•	•	•	•	G	•	G	•	•	G	G	G	•	G
1415	G	•	•	•	•	•	•	•	•	C	C	•	•	•	•	•
1520	G	Cg	•c	Cg	C	•c	•	•	•	•a	Cg	•	•	(C)	•	•
1532	U	•	•	•	•	•	•	•	•	•	•	Au	Au	•	•	•

[a] Abbreviations not explained are obvious from text. All positions are based upon oligonucleotide data (266) except 353 (233).
[b] Same composition as consensus (•).
[c] Composition upper case—major base; if no other specified, then it accounts for >90% of assayable cases.
[d] Composition lower case—minor occurrence base; found in <15% of assayable cases (or in only one species for groups containing seven or less species).
[e] —, No nucleotide at this position.
[f] Composition in parentheses—based upon one example only.

contain the octanucleotide (G)AUUUAUCG, which entirely covers the version of the second helix (positions 198 to 219) found in this subdivision (266, 267). (G)AYCUUCG, which forms part of this helix in the β and γ subdivisions, is found in 44% of γ catalogs and 13% of those from the β subdivision (267–269). (Although all known 16S rRNA sequences from β species have the structure for this helix shown in Fig. 9, the loop often contains a G residue, which breaks up the [T_1 ribonuclease] oligonucleotide otherwise characteristic of the structure [unpublished analysis].) Oligonucleotides of the form (G)CCUCU. . . , seen in the δ-purple bacterial version of this helix, occur in 78% of δ-subdivision catalogs (163).

With regard to the first helix, positions 184 to 193, oligonucleotides representing the structure can be identified in only 35 and 26% of cataloged species, respectively, from the β and γ subdivisions (unpublished analysis). However, every one of the sequenced 16S rRNAs from these two subdivisions show the characteristic 3-base pair structure (unpublished analysis). Based upon oligonucleotide and sequence evidence, at least 83% of α-purple bacteria must also possess the same form of this helix (unpublished analysis).

Figure 8 shows the β and γ subdivisions to be relatively closely related to one another, a fact that can also be demonstrated with other types of analysis (163); see Table 4. The common β–γ lineage and the α- and δ-subdivision lineages appear to have split from one another in such rapid succession that their branching order cannot be resolved by the evidence that now exists. However, the higher-order structural evidence, specifically the (derived) form of the helix covering positions 184 to 193 (Fig. 9), would suggest that the α, β, and γ subdivisions form a grouping that excludes the δ subdivision.

Table 5 through 8 lists representative genera and species in each of the four subdivisions in a rough phylogenetic arrangement. Three of the four subdivisions contain photosynthetic species: the α-purple bacteria seems predominantly photosynthetic, and the β subdivision shows photosynthesis in its two main subgroups, while in the γ subdivision photosynthesis appears to be confined to one of its three main subgroups.

α-Purple bacteria. The intimate juxtaposition of photosynthetic and nonphotosynthetic species in the α subdivision

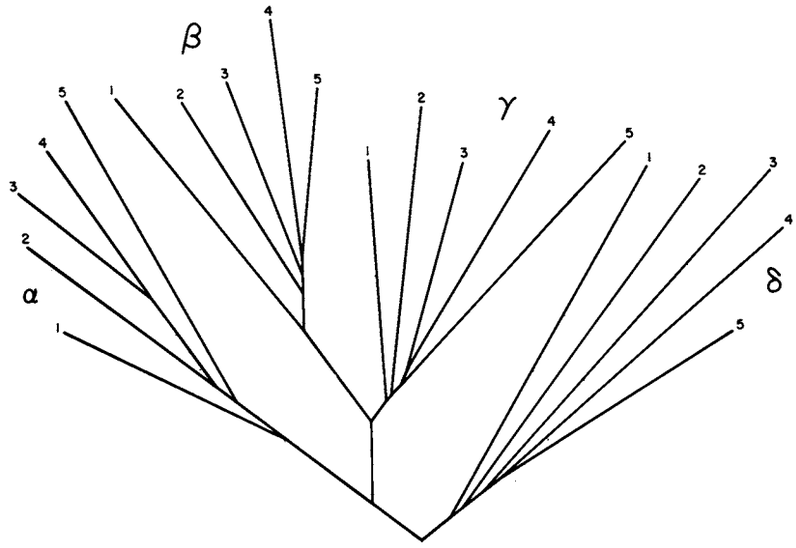

FIG. 8. Phylogenetic tree for the purple bacteria based upon 16S rRNA sequences. The tree was constructed (36) from an evolutionary distance matrix (99), generated from an alignment (260) of five representative sequences from each of the four purple bacterial subdivisions. (Only positions represented in all sequences were used in the calculation.) The root was determined by using several eubacterial outgroup sequences. The sequences used were as follows. α subdivision: 1, *Rhodospirillum rubrum* (Woese et al., unpublished data); 2, *Agrobacterium tumefaciens* (275); 3, *Rhodopseudomonas palustris* (Woese et al., unpublished data); 4, *Rhodopseudomonas acidophila* (Woese et al., unpublished data); 5, *Rhodobacter capsulatum* (Woese et al., unpublished data). β subdivision: 1, *Neisseria gonorrhoeae* (Woese et al., unpublished data); 2, *Spirillum volutans* (Woese et al., unpublished data); 3, *Nitrosolobus multiformis* (Woese et al., unpublished data); 4, *Rhodocyclus gelatinosa* (Woese et al., unpublished data); 5, *Rhodocyclus purpureus* (Woese et al., unpublished data). γ subdivision: 1, *Chromatium vinosum* (Woese et al., unpublished data); 2, *Legionella pneumophila* (Woese et al., unpublished data); 3, *Pseudomonas aeruginosa* (Woese et al., unpublished data); 4, *Acinetobacter calcoaceticus* (Woese et al., unpublished data); 5, *Escherichia coli* (19). δ subdivision: 1, *Myxococcus xanthus* (163); 2, *Desulfovibrio desulfuricans* (163); 3, *Bdellovibrio stolpii* (Woese et al., unpublished data); 4, *Desulfotobacter postgatei* (Woese et al., unpublished data); 5, *Desulfuromonas acetoxidans* (Woese et al., unpublished data). Outgroup sequences: *Bacillus subtilis* (68) and *Thermotoga maritima* (1).

suggests a more or less continual evolution of the latter from the former (267). Aerobic metabolism also appears to have arisen a number of times in this subdivision alone (267). The close association of species reducing and oxidizing nitrogen compounds, e.g., *Rhodopseudomonas palustris* and *Nitrobacter winogradskii* (267), suggests some sort of evolutionary connection between the two metabolisms. The metabolic richness and diversity of species evolving from purple photosynthetic ancestry in this subdivision are remarkable.

The α subdivision is also of general biological interest because certain of its members have interesting associations with various eucaryotes. The rhizobacteria (essential for nitrogen fixation in legumes), the agrobacteria (pathogenic for plants) (267, 274), and the rickettsias (intracellular pathogens of animals) (235) form a tight cluster within subgroup α-2. Sequence differences in 16S rRNA among members of this cluster are under 7% (235). These particular species have in common the tendency to form intimate, if not intracellular, associations with eucaryotic cells. It is then no great surprise to learn that the endosymbiont that gave rise to (most, if not all) eucaryotic mitochondria was itself a member of the α subdivision (274).

β-Purple bacteria. Most of the characterized β-purple species fall into two main subgroups. However, poorly defined deeper branchings are also evident, represented, for example, by *Spirillum* and *Neisseria* (269). This subdivision is a potpourri of classical genera (Table 6), some of which are not even phylogenetically coherent within the subdivision (269). The β-photosynthetic species, recently reclassified in the genus *Rhodocyclus* (92), are quite distinct from other purple nonsulfur bacteria, i.e., those residing in the α subdivision. Over and above their rRNA sequence differences, β species differ from their α counterparts in cytochrome *c* type; β cytochromes are of the small-subunit type, while cytochromes from the α subdivision are of the medium or large type (37). Moreover, photoreaction centers in photosynthetic β species have a structure distinct from that seen among α species (27).

γ-Purple bacteria. The γ-purple bacteria (Table 7), the most extensively characterized of the four subdivisions of purple bacteria, is again a mixture of phenotypes (268): photosynthetic with nonphotosynthetic, aerobic with anaerobic, heterotrophic with chemolithotrophic, etc. Oligonucleotide catalog analysis divided the γ-purple bacteria into three main subgroups: one containing mainly photosynthetic species of the purple sulfur type, e.g., *Chromatium* (50, 198); a second known to contain only species associated with Legionnaires disease (132); and a third that is a mixture of nonphotosynthetic genera from the enterics, vibrios, oceanospirilla, the fluorescent pseudomonads and relatives, and others (268). As additional complete 16S rRNA sequences become available, it begins to look as though the β subdivision may ultimately be shown to be a subgroup within the γ subdivision (albeit deeply branching). In any case its close association with the γ subdivision is surprising given that the β-purple bacteria have such a distinctive signature; see Table 4 (266).

δ-Purple bacteria. The δ-purple bacteria (Table 8) subdivision harbors three disparate phenotypes: the sulfur and sulfate-reducing eubacteria, the myxobacteria and relatives, and the bdellovibrios (163). At the present writing the phylogenetic detail within the subdivision is not clear. The myxobacteria and relatives indeed form a coherent grouping, a clade (130). The bdellovibrios probably form one as well,

TABLE 4. Sequence signature distinguishing the four subdivisions of purple bacteria[a]

Position	α	β	γ	δ	Other[c]	Position	α	B	Y	6	Other
6	U[d]	R:u[e]	G	G	G	871	U	G	U	U	U
7	G	A:U	A	G	G	875	C	U	U	U:c	Y
44	A	A	A	G	G	877	Y	G	G	Y:a	Y
50*[b]	A	U	A	A	A	878	Y:a	A	A	Y:a	R:U
107	A:G	G	G	G	N	916*	G	G	U:G	G	G
108	G:c:a	A	G:a	C	G:C	929	G	A	G	G	G
124*	G	A	N	G	G	947	G	U	G	G:u	G
129	A:c	A	A:g	U	U	948	C	G:U	C:u	Y	C
129:1	—	—	—	A	A	976*	G	A:g	G	G	G
199	A:g	C	R:C	R	N	1015	G	G	G	A:g	R
233	C	C	C	Y	C	1024	Y	G	G	G:c	N
236	G	A	A:g	G	G	1026	U	G	G	G	N
237	C	U	N	C	C	1116*	C	U	U	Y	U
242	G	G	G	C:g	Y	1120*	Y	A	C	Y:G	N
284	C	C	C	G:c	G	1153	R	U	G:u	N	N
371*	A:g	A	A	G:a	N	1219	U	A:g	A:U	A	R:U
390	U:c	U	U	C:u	Y	1233*	G	C:A	G	R	G
398	U	U	U	C	C	1234*	C	A	C	C:a	C
438	U	U	U	G:u	G:U	1246	U:c	G	R	G:u	G
449	A:C	A:C	R	A	A	1252	U	A	A	A	A:Y
485*	G:u	G	U	G	G	1260	G:Y	C	G:u	G:Y	G
496	A	A	A	G:a	R	1291	R	C	Y	C:g	C
502	C	A	A	A:g	C	1297	U	G	G	Y	U
543	G	U	U	U:c	G	1298	C	U	U	C:A	C
554*	Y	A	A	U:a	U	1421	U	G	G:U	Y	U
564*	C	C	C	U:c	Y	1426	Y	Y:a	G	U:R	A
640	A	U:G:a	A	A	A	1431	Y	A	A	Y:a	Y
689	A	C	C	R:u	C	1437	Y:a	A	A	C	C
690	G	A	G	G:a	G	1441	G:u:a	A	A	G:u	N
698*	U	G	G	Y:a	G	1443*	G	C:u:g	Y	G	G:C
722	G	A	G:a	G:a	G	1460	A	C	C	A	A
760	G:u	U:g	G	G	G	1464	R	U	U	G	G
812*	G	C	G	G	G	1465	R	A:g	A	C:u	Y
822	R:u	U	U	R:u	Y:A	1467	C	C	C:a	Y	G
823	R	C	C	R:u	R	1469	C	C	C	A:u:c	R
825	G	A	A	A:g	R	1481	C:g	U	U	U:c	U

[a] Compositions based upon an alignment of 31 16S rRNA sequences from purple bacteria and oligonucleotide catalogs (19, 51, 64, 83, 163, 198, 267–269, 274; Woese, unpublished analysis). Y, Pyrimidine; R, purine; N, any nucleotide. —, Position does not exist in species so indicated.

[b] Catalog information was also used in determining composition.

[c] Consensus composition for the other eubacterial phyla; no clear consensus is indicated by "N".

[d] Upper case, Major base; if no other specified, it accounts for >90% of assayable cases.

[e] Lower case, Minor occurrence base; found in <15% of assayable cases (or in only one sequence in group).

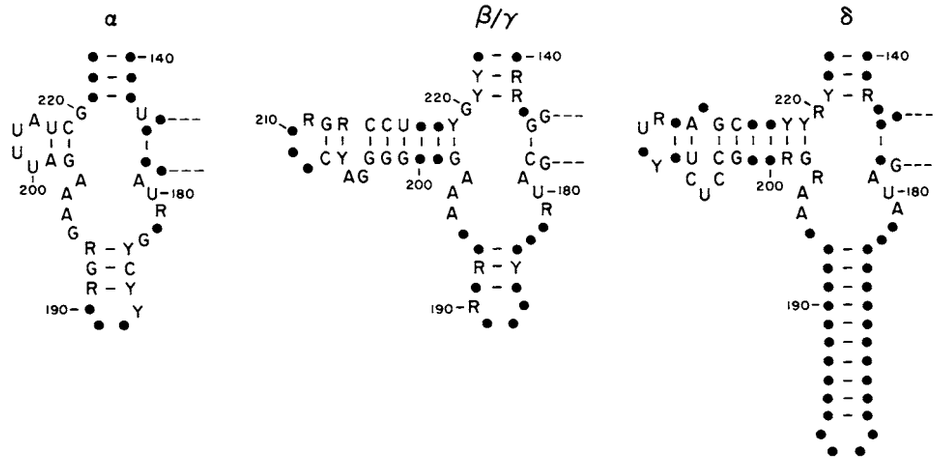

FIG. 9. Differences in higher-order structural detail among the various subdivisions of the purple bacteria for the region of 16S rRNA between positions 180 and 220 (163). Composition of a position is given when it is invariant or highly conserved within a subdivision, but it is shown as a dot otherwise. Base pairs are indicated by connecting lines.

TABLE 5. Characterized species of α-purple bacteria (64, 269)[a]

Subgroup α-1
- *Rhodospirillum rubrum*
  - *R. photometricum*
- *R. molischianum*
- *R. fulvum*
- *Rhodopseudomonas globiformis*
- *Aquaspirillum itersonii*
- *Azospirillum brasilense*

Subgroup α-2
- *Rhodomicrobium vannielii*
- *Rhodopseudomonas viridis*
- *Rhodopseudomonas palustris*
  - *Nitrobacter winogradskyi*
- *Rhozobium leguminosarum*
  - *Agrobacterium tumefaciens*[a]
  - *Rochalimaea quintana*
- *Rhodopseudomonas acidophila*
- *Pseudomonas diminuta*
- *Phenylobacterium immobile*

Subgroup α-3
- *Rhodobacter capsulatum*
  - *Rhodobacter sphaeroides*
  - *Paracoccus denitrificans*
- Manganese oxidizers (2 strains)

Subgroup α-4
- *Erythrobacter longus*

[a] Indentation indicates specific relationship; for example, *Rhizobium leguminsoraum*, *Agrobacterium tumefaciens*, and *Rochalimaea quintana* are specific relatives of one another, to the exclusion of the other species in their subgroup.

TABLE 6. Characterized species of β-purple bacteria (64, 269)[a]

Subgroup β-1
- *Rhodocyclus gelatinosa*
  - *Sphaerotilus natans*
- *Pseudomonas testosteroni*
  - *P. acidovorans*
  - *Aquaspirillum gracile*
  - *A. aquaticum*
  - *Comamonas terrigena*
- *Thiobacillus intermedius*

Subgroup β-2
- *Rhodocyclus tenue*
  - *R. purpureus*
- *Aquaspirillum dispar*
- *A. serpens*
  - *A. bengal*
- *Chromobacterium violaceum*
- *Chromobacterium lividum*
- *Alcaligenes faecalis*
- *Alcaligenes eutrophus*
- *Pseudomonas cepacia*
- *Thiobacillus denitrificans*
- *Vitreoscilla stercoraria*

Subgroup β-3[b]
- *Spirillum volutans*
- *Nitrosomonas europaea*
- *Nitrosococcus mobile*
- *Nitrosolobus multiformis*
  - *Nitrosovibrio tenuis*
  - *Nitrosospira* sp.
- *Neisseria gonorrhoeae*

[a] Indentation indicates specific relationship.
[b] Paraphyletic group, defined only by exclusion from subgroups 1 and 2.

TABLE 7. Characterized genera or species of γ-purple bacteria (50, 64, 132, 268)[a]

Subgroup γ-1
- *Chromatium*
  - *Thiocapsa*
  - *Thiocystis*
  - *Thiodictyon*
  - *Thiospirillum*
  - *Lamprocystis*
- *Nitrosococcus oceanus*
- *Ectothiorhodospira*

Subgroup γ-2
- *Legionella*

Subgroup γ-3
- Fluorescent pseudomonads[b]
  - *Xanthomonas*
    - *Lysobacter*
  - *Acinetobacter*
- *Oceanospirillum*
- *Alcaligenes putrifasciens*
- *Pasteurella multocida*
- *Aeromonas hydrophila*
- *"Bacteroides" amylophilus*
- Enterics, vibrios, and photobacteria
- *Halomonas elongata*
  - *"Flavobacterium" helmephilum*
- *Leucothrix mucor*
- *Beggiatoa leptomitiformis*

[a] Indentation indicates specific relationship.
[b] Includes *Pseudomonas aeruginosa*, *P. alcaligenes*, *P. fluorsecens*, *P. putida*, *P. stutzeri*, *P. syringae*, *P. pseudoalcaligenes*, and *Serpens flexibilis*.

TABLE 8. Characterized genera and species of δ-purple bacteria (51, 83, 130, 163)[a]

Myxococcus group
- *Myxococcus*
  - *Cystobacter fuscus*
  - *Stigmatella aurantiaca*
- *Sorangium cellulosum*
- *Nannocystis exedens*

Bdellovibrio group
- *Bdellovibrio stolpii*
  - *Bdellovibrio starii*
- *Bdellovibrio bacteriovorus*

Sulfur and sulfate reducers
- *Desulfovibrio desulfuricans*
- *Desulfuromonas*
- *Desulfotobacter postgatei*
  - *Desulfosarcina variabilis*
  - *Desulfonema limicola*
- *Desulfobulbus propionicus*??

[a] Indentation indicates specific relationship.

but the fact that *Bdellovibrio bacteriovorus* is relatively rapidly evolving makes its exact placement uncertain (83). Whether the myxobacteria and relatives or the bdellovibrios (or both) arise outside of the group defined by the sulfur and sulfate reducers remains unresolved (unpublished analysis). In any case it would appear that the myxobacteria and bdellovibrios represent aerobic adaptations of some ancestral anaerobic sulfur-metabolizing phenotype.

Gram-Positive Eubacteria

Cell wall type distinguishes the gram-positive eubacteria from the others (56, 183). However, as mentioned above,

TABLE 9. Characterized genera and species of gram-positive bacteria[a]

High-G+C subdivision	Low-G+C subdivision
Bifidobacterium	*Bacillus*
Propionibacterium	*Planococcus*
Actinomyces	*Sporolactobacillus*
Arthrobacter	*Sporosarcina*
Micrococcus	*Thermoactinomyces*
Dermatophilus	*Staphylococcus*
Cellulomonas	*Lactobacillus*
Derskovia	*Pediococcus*
Nocardia cellulans	*Leuconostoc*
Microbacterium	*Streptococcus*
Corynebacterium (plant associated)	*Mycoplasma*
Brevibacterium linens	*Acholeplasma*
Streptomyces	*Spiroplasma*
Kitasatoa	*Anaeroplasma*
Chainia	*Clostridium innocuum*
Microellobosporia	*Erysipelothrix*
Streptoverticillium	*Clostridium pasteurianum*
Actinomadura	*C. butyricum*
Streptosporangium	*C. scatologenes*
Thermomonospora	*Sarcina ventriculae*
Mycobacterium	*Clostridium oroticum*
Nocardia	*C. indolis*
Brevibacterium ketoglutamicum	*C. aminovalericum*
Corynebacterium	*Butyrivibrio fibrosolvens*
Geodermatophilus	*Clostridium lituseburense*
Frankia	*Eubacterium tenue*
Dactylosporangium	*Peptostreptococcus*
Ampurariella	*C. aceticum*
Actinoplanes	*C. acidiurici*
Micromonospora	*C. purinolyticum*
Arthrobacter simplex	*Clostridium barkeri*
	Eubacterium limosum
Photosynthetic subdivision	*Acetobacterium woodii*
Heliobacterium	*Clostridium thermosaccharolyticum*
	C. thermoaceticum
Species with gram-negative walls	*Acetogenium*
Megasphaera	*Thermoanaerobium*
Selenomonas	*Peptococcus*
Sporomusa	*Ruminococcus*

[a] Approximate phylogenetic clustering suggested by indentation. References 56, 131, 199–201, 203–207, 220, 221, 255, 262, 264.

there are several important exceptions whose walls are not gram positive. The phylum appears to consist of four subdivisions, only two of which are well characterized. These two are readily distinguished on the basis of DNA composition. The one includes species whose DNAs all contain more than 55% guanine plus cytosine (G+C); the other is made up of species whose DNAs contain <50% (56, 201, 204). The recently discovered phototroph *Heliobacterium chlorum* (58, 60, 255) is the only characterized representative of the third subdivision, while the genera *Megasphaera*, *Selenomonas*, and *Sporomusa* constitute the fourth (203). Members of the third and fourth subdivisions do not have gram-positive cell walls (60, 203).

Species in the high-G+C gram-positive subdivision conform to a general actinomycete phenotype: they tend to be pleomorphic, form branched filaments, etc. (200, 201, 204–206). Most are aerobic, with the exception of the deeper branches, e.g., the bifidobacteria. The group as a whole is not particularly deep; by oligonucleotide catalog measure all high-G+C gram-positive bacteria are no further from each other than are *Bacillus* species from those of *Lactobacillus*, for example. The lowest S_{AB} values in the group correspond roughly to sequence similarities in the range of 85% (56, 201, 204). The group, therefore, would seem not to be a particularly ancient one.

Species in the low-G+C gram-positive subdivision conform by and large to a clostridial phenotype. They tend to be anaerobic, rod shaped, endosporeforming, etc., although a number have lost one or more of these characteristics. In contrast to the high-G+C group, the low-G+C gram-positive bacteria form a phylogenetically deep, and therefore presumably ancient, cluster (56).

Table 9 lists some of the characterized gram-positive genera and species in a crude phylogenetic ordering.

The gram-positive rRNAs sequenced so far are not broadly representative enough to permit the construction of a phylogenetic tree for the entire phylum. However, the phylum is easily defined by cluster analysis of S_{AB} values based upon oligonucleotides (56, 206). Although few in number, the signature positions characterizing the gram-positive eubacteria (Table 3) are strong ones. Note in particular the presence of a C residue at position 1207, which occurs in all cataloged gram-positive bacteria (roughly 150 species) and in cyanobacteria, but nowhere else among eubacteria (266). The A residue at position 1198 is present in roughly 75% of gram-positive species, but occurs elsewhere only in the bacteroides phylum and occasionally among the spirochetes (266). An absolute requirement for an A residue at position 513 also characterizes the gram-positive bacteria, but few other phyla.

Higher-order structure in 16S rRNA also helps to identify gram-positive bacteria. Two characteristic adjacent A-G pairs, positions 1425-6 to 1474-5, occur in the penultimate helix of almost all gram-positive 16S rRNAs sequenced to

TABLE 10. Sequence signature distinguishing gram-positive subdivisions[a]

Position[b]	Subdivision	Composition (%)[c]				Consensus[d]
		A	G	C	U	
168	Low G+C	5	81	0	0	G
	High G+C	2	20	0	70	
	Sporomusa group	0	100	0	0	
	H. chlorum	−	+	−	−	
906	Low G+C	7	92	0	0	G
	High G+C	100	0	0	0	
	Sporomusa group	0	100	0	0	
	H. chlorum	−	+	−	−	
955	Low G+C	0	0	1	99	U
	High G+C	74	0	25	1?	
	Sporomusa group	0	0	0	100	
	H. chlorum	−	−	−	+	
998	Low G+C	2	5	88	3	N
	High G+C	36	33	16	1	
	Sporomusa group	0	0	0	100	
	H. chlorum	−	−	+	−	
1116	Low G+C	0	0	1	95	U
	High G+C	0	0	54	46	
	Sporomusa group	0	0	0	100	
	H. chlorum	−	−	−	+	
1167	Low G+C	62	?	2	1	A
	High G+C	0	0	2	98	
	Sporomusa group	75	0	0	0	
	H. chlorum	+	−	−	−	
1410	Low G+C	94	<6	0	0	A
	High G+C	0	>46	0	0	
	Sporomusa group	75	0?	0	0	
	H. chlorum	+	−	−	−	

[a] References: low-G+C subdivision (131, 199, 203, 220, 221); high-G+C subdivision (200, 201, 204, 205, 207); sporomusa group (203); *Heliobacterium chlorum* (255).
[b] Position in sequence.
[c] Composition of position (percentage of oligonucleotide catalogs having an oligonucleotide covering the position which shows the indicated composition); failure to find an oligonucleotide covering a position in some catalogs causes percentage compositions to sum to less than 100%.
[d] Composition characteristic of position in the (vast) majority of other eubacterial phyla.

date (260; unpublished data); see Fig. 10. This arrangement has yet to be seen outside this group. (We will encounter G-A pairs at this location in another phylum, however.) Oligonucleotide catalog data indicate that these adjacent A-G pairs are common among gram-positive bacteria. The general sequence (G)YAAYACCC (which includes the two adjacent A's of the A-G pairs in question) is found in 85% of catalogs from the high-G+C group and in 67% of those from the low-G+C group, but occurs nowhere else among eubacteria (266).

The sequence signature of Table 10 distinguishes between the two main gram-positive subdivisions. In all cases the composition characteristic of the low-G+C subdivision is that found in most other phyla; it is undoubtedly, therefore, the ancestral composition. This would suggest that the high-G+C gram-positive lineage has been subject to rapid evolution; see discussion below.

Genealogical substructure within the two major subdivisions can be seen (56, 200, 201, 204–206), although it will not be systematically detailed here. In the high-G+C subdivision, the deepest branchings are defined by anaerobic species, the bifidobacteria and the propionibacteria. The low-G+C subdivision shows at least five major branches, most of which contain clostridia.

One subline in the low-G+C subdivision has given rise to four groups of particular interest: *Bacillus*, *Lactobacillus*, *Streptococcus*, and the mycoplasmas and their relatives (56, 262). The subline is an interesting evolutionary study, for its evolution in a way parallels the development of aerobic conditions on the planet. *Bacillus* species are basically aerobic, though a few also grow well anaerobically. *Lactobacillus*, *Streptococcus*, and the mycoplasmas are basically anaerobic, but tolerate and in some cases even utilize a little oxygen. Their clostridial relatives are true anaerobes. The evolutionary radiation that spawned the various groups might well have occurred during the microaerobic period in earth history (see below), with *Bacillus* then becoming fully adapted to an oxygen atmosphere. (The exact order of branching among these groups will not be known until sequences representing all four have been determined.) The mycoplasma group will be considered in detail below.

Cyanobacteria

The cyanobacteria, the classical blue-green algae, are a group of procaryotes defined by the common possession of chlorophyll *a*. They form a phylogenetically coherent unit (12, 40, 56) that contains no known nonphotosynthetic

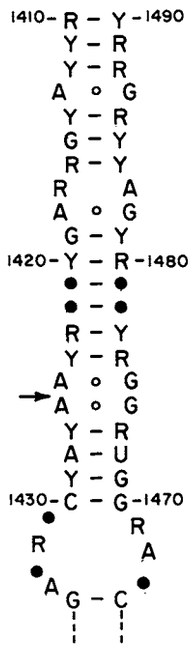

FIG. 10. Higher-order structure for the gram-positive bacteria in the region of 16S rRNA between positions 1410 and 1480 (68, 93, 163). Composition of a position is given when it is invariant or highly conserved within the gram-positive bacteria, but it is shown as a dot otherwise. Base pairs are indicated by connecting lines; A–G pairs are shown by open circles. Arrow shows adjacent A–G pairs discussed in text.

representatives. However, the phylum does include *Prochloron* (191), an organism that, like green plant chloroplasts, possesses chlorophyll *b* as well as chlorophyll *a* (127). Green algal chloroplasts trace their ancestry to this phylum, as expected (11, 41, 67, 189, 223). Analysis of oligonucleotide catalog data did not resolve the questions of whether the chloroplasts (whose rRNAs appear somewhat rapidly evolving) arose from within or from just outside of the cluster defined by extant cyanobacteria and whether these organelles are related to *Prochloron* to the exclusion of the cyanobacteria. Although about 25 complete sequences of cyanobacterial 16S rRNAs have now been determined (but as yet not that from *Prochloron*), these questions remain unresolved as of this writing (S. J. Giovannoni, S. Turner, G. J. Olsen, D. J. Lane, and N. R. Pace, manuscript in preparation). The precise origin of the brown algal chloroplasts and certain others also has yet to be determined. However, the rRNAs of the chloroplasts from the red algae (10) and *Euglena* sp. (276) are known to be closely related to those of cyanobacteria.

The sequence signature for the cyanobacteria and chloroplasts (Table 3) is small but significant. Position 799, a G residue in all but 1 to 2% of other eubacterial catalogs, is U or A in all cyanobacterial sequences or catalogs, including *Prochloron* (191); G's are encountered in some chloroplast sequences, however (276). Position 1207, a G residue in all other eubacterial catalogs, is always C in the gram-positive bacteria and cyanobacteria (including *Prochloron* and the chloroplasts), except for one cyanobacterial sequence (191; Giovannoni et al., in preparation). Very few eubacteria show A at position 1233 (266); however, all cyanobacteria (except one) and *Prochloron* do (191; Giovannoni et al., in preparation). (Most chloroplast examples do not [10, 41, 67, 189, 223, 276].)

Green Sulfur Bacteria

The four cataloged species of the green photosynthetic bacteria covering the genera *Chlorobium* and *Chloroherpeton* form a relatively tight phylogenetic unit, especially in view of the fact that their phenotype is generally considered a very primitive one (63, 266). Since the characterized species seem representative of the known spectrum of green sulfur bacteria, the question of why the group is relatively shallow becomes nontrivial. As will be seen below, the green sulfur bacteria are not related to the other green type of photosynthetic bacteria, the so-called green non-sulfur bacteria (63).

The signature characteristic of the phylum (Table 3) contains positions 995, 1234, and 1410. The one sequence now available for the group (W. G. Weisburg, Ph.D. thesis, University of Illinois, Urbana, 1986) suggests that additional strong signature positions will appear when more sequences from the group are known, e.g., the lack of a base in the vicinity of position 1167 and the insertion of a base after position 1174 (Weisburg, Ph.D. thesis).

Spirochetes

Spirochetes are one of the few groupings correctly identified by classical (morphological) criteria (167). Their common spiral shape and axially coiled fibrils, lying between inner and outer cell envelopes, are strikingly characteristic (24). Table 11 shows representative species whose rRNAs have been characterized, in rough phylogenetic arrangement. The sequence signature for the group shown in Table 3 is quite distinctive. For example, the U residue at position 47 found in all species from this group occurs nowhere else among eubacteria (266; unpublished analysis). The same holds for the A residue at position 52, and the C residue at position 1415, while universal among spirochetes, is otherwise extremely rare among eubacteria (266).

Two clearly separated subdivisions exist within the phylum: one composed of the leptospiras and the other containing spirochetes, treponemes, and the like (167, 266). The

TABLE 11. Characterized species of spirochetes and relatives[a]

Spirochete subdivision
Spirochaeta halophila
S. aurantia
S. litoralis
S. isovalerica
Treponema succinifaciens
T. bryantii
T. denticola
T. phagadensis
S. stenostrepta
S. zuelzerae
T. pallidum
Borrelia hermsii
T. hyodysenteriae
Leptospira subdivision
Leptospira patoc
Leptospira interrogans
Leptonema illini

[a] Reference 167; Oyaizu, unpublished data; Weisburg et al., unpublished data. Indentation indicates specific relationship.

TABLE 12. Characterized species in the bacteroides-flavobacterium phylum (166, 234)[a]

Bacteroides subdivision
- *Bacteroides fragilis*
  - *B. ovatus*
  - *B. uniformis*
  - *B. asaccharolyticus*
  - *B. vulgatus*
- *B. ruminicola*
  - *B. melaninogenicus*
- *B. distasonis*

Flavobacterium subdivision
- *Flavobacterium aquatile*
  - *Cytophaga johnsonae*
    - *Sporocytophaga myxococcoides*
  - *Cytophaga lytica*
  - *F. uliginosum*
  - *F. breve*
- *Flavobacterium heparinum*
  - *F. ferrugineum*
  - *F. elegans*
  - *Saprospira grandis*
  - *Haliscomenobacter hydrossis*

Unnamed subdivision
- Unnamed anaerobic flexible rod; strain Pl-12fs

[a] Indentation indicates specific relationship.

classical taxonomic distinction between spirochetes and treponemes, however, does not hold up (167); the two types form a genealogically intermixed unit. The lone species *Treponema hyodysenteriae*, however, represents a lineage distinct from the main spirochete-treponeme cluster (167), and the genus *Borrelia* is slightly peripheral to the main cluster as well (167).

It was previously suggested that the *Haloanaerobiaceae*, unusual anaerobic halophilic eubacteria, belong to this phylum (161), but full sequence information fails to confirm this (A. Oren and C. R. Woese, unpublished).

Bacteroides, Flavobacteria, and Their Relatives

The phylum made up of bacteroides and flavobacteria is an unexpected mixture of anaerobes, the bacteroides, and various aerobes, from genera such as *Flavobacterium*, *Cytophaga*, and others (166, 234). Table 12 lists some of its characterized representatives.

The grouping is cleanly defined by both sequence distance and parsimony analysis of rRNA sequences (234). Table 3 shows its quite distinctive signature. Note, for example, the U residue at position 570 (found in all members of this and the planctomyces phylum, but nowhere else among the eubacteria), the A residue at position 995 (otherwise a C, except in the green sulfur bacteria), and the A at position 1532 (which sets this group apart not only from all other eubacteria, but from archaebacteria and eucaryotes as well [5, 166, 259]). Positions 570 and 866 are involved in a recognized "tertiary structural" interaction and so vary in concert (74).

A higher-order structural feature so far unique to this group is a series of three contiguous G-A pairs involving positions 1424-6 with 1474-6 (234). (Recall the two contiguous A-G pairs in this region characteristic of the gram-positive phylum; Fig. 10.) Common oligonucleotide sequences indicate that most if not all organisms in this phylum share the property (166).

The phylum's two major subdivisions separate the anaerobic *Bacteroides* species from the aerobic ones (166). The lack of phenotypic resemblance between members of the two subdivisions is remarkable, but may reflect only the fact that the bacteroides have been studied one way and the flavobacteria and relatives another. Most *Bacteroides* and (at least) a few species of *Flavobacterium* possess sphingolipids, compounds otherwise rare among eubacteria (234). An organism (yet unnamed) whose phenotype seems intermediate between these two has been isolated by K. O. Stetter (166); it is a strickly anaerobic flexible rod. Its phylogenetic position is also "intermediate" (166), and so the organism probably represents a third uncharacterized subdivision.

A sequence signature (derived from oligonucleotide catalogs) distinguishing the two subdivisions of the phylum is shown in Table 13. In most if not all cases the flavobacterial version would appear to be the ancestral one for the group, in that it is the one found in most of the other eubacterial phyla (166).

Planctomyces and Their Relatives

Species variously placed in the genera *Planctomyces*, *Pasteuria*, and *Pirella* (186, 202), together with the recently

TABLE 13. Sequence signature distinguishing the bacteroides-flavobacterium subdivisions[a]

Position	Subdivision	Composition (%)				Consensus
		A	G	C	U	
306	B	0	100	0	0	A
	F	92	0	0	0	
316	B	0	0	0	100	C
	F	0	0	92	0	
337	B	100	0	0	0	G
	F	0	100	0	0	
718	B	0	0	9	91	A
	F	92	0	8	0	
809	B	55	9	0	0	G
	F	0	100	0	0	
947	B	100	0	0	0	G
	F	0	100	0	0	
986	B	0	0	100	0	A or U
	F	50	8	8	33	
1089	B	0	0	100	0	G
	F	?	>8	0	0	
1202	B	0	100	0	0	U
	F	0	0	0	100	
1234	B	0	0	0	100	C
	F	0	0	100	0	
1321	B	0	0	100	0	U
	F	0	0	0	92	
1356	B	0	0	100	0	G
	F	0	>8	0	16?	

[a] References 166 and 234. See footnotes *b* to *d* in Table 10. B, Bacteroides subdivision; F, flavobacterium subdivision.

described hot spring organism *Isocystis pallida* (S. J. Giovannoni, *in* J. G. Holt, ed., *Bergey's Manual of Systematic Bacteriology*, vol. 3, in press; S. J. Giovannoni and H. Oyaizu, unpublished data), define this phylum. All are noted for the fact that their cell walls contain no peptidoglycan (110).

In terms of oligonucleotide catalogs, this is the most unique of eubacterial groups. Their S_{AB} values with other eubacteria are generally in the range of 0.10 to 0.15 (202), far lower than typical S_{AB} values between eubacterial phyla, which are normally 0.20 to 0.25 (Fig. 2). Oligonucleotide catalogs for the species in this group contain fewer of the highly conserved (ancestral) oligonucleotides than do those from any other phylum (202). As might then be expected, the group possesses a strong signature (Table 3).

The remarkable distance of the planctomyces group from other eubacteria measured by the rRNA cataloging approach was initially interpreted to mean that these organisms represent by far the deepest branching in the eubacterial line of descent (202). Analysis of full sequences, however, does not bear this out (H. Oyaizu and C. R. Woese, unpublished data); these large sequence distances are due to rapid evolution of the lineage, not an especially early divergence from the common line of eubacterial descent; see Fig. 11.

Chlamydiae

The two known 16S rRNA sequences representing *Chlamydia*, i.e., from *Chlamydia psittaci* and *C. trachomatis*, are closely related; they show <5% difference (233). Since no other even moderately close relatives of these organisms are known, the phylum cannot yet be considered adequately described. A distant relationship between the chlamydiae and the planctomyces group is suggested by sequence signature (233). Of the sequence positions in Table 3 characteristic of the planctomyces and their relatives, five, i.e., 47, 48, 52, 53, and 353, are also found in the *C. psittaci* and *C. trachomatis* sequences (233). (Other positions in the 16S rRNA sequence, not shown in Table 3, that link the chlamydiae to the planctomyces and relatives are 110, 331, and 361 [233].)

Nevertheless, the chlamydiae should be considered to represent a distinct phylum, for the similarity between their 16S rRNAs and those of the planctomyces is too slight to place the two in the same taxon. Their sequence similarity, 71 to 72%, is considerably less than the 78% between *Planctomyces staleyi* and its relative *Isocystis pallida* (unpublished calculation). (The 71 to 72% figure is not the artificial result of a relatively rapid evolution in the chlamydial lineage, for chlamydial sequences are closer to outgroup sequences than are their counterparts in the planctomyces phylum. In addition, their S_{AB} values with other eubacteria are not as abnormally low as those of the planctomyces [202; unpublished calculation].) Interestingly, the chlamydiae, like the planctomyces group, also have cell walls that lack peptidoglycan (7, 59, 110).

Radiation-Resistant Micrococci and Their Relatives

Until recently, radiation-resistant micrococci and their relatives was known to include only a few closely related species of radioresistant bacteria, i.e., *Deinococcus radiodurans* and its relatives (18, 56). However, it has now been shown to include the ubiquitous hot spring organism *Thermus aquaticus* as well (82). The signature shown for the phylum in Table 3 is rather weak. However, the two sequences now available (Weisburg, Ph.D. thesis; Giovannoni, unpublished data) suggest that the group should have a strong signature once this can be derived from full sequences rather than oligonucleotides. Similarity between the 16S rRNA sequences of *D. radiodurans* and *T. aquaticus* is 81% (unpublished analysis), low enough to place them in separate subdivisions of the same phylum.

Green Non-Sulfur Bacteria and Their Relatives

The phylum containing the green non-sulfur bacteria is one of those for which little phenotypic justification exists. The group contains four known members, the thermophilic phototroph *Chloroflexus aurantiacus*, two mesophilic gliding species from the genus *Herpetosiphon*, and the thermophile *Thermomicrobium roseum* (63, 94, 162, 170). *Chloroflexus* and the green sulfur bacteria resemble one another in chlorosome structure and light-harvesting pigment type (63, 136); yet their rRNAs are unrelated (as mentioned above), and the structure of their photoreaction centers differs (45, 171). The unusual long-chain diols found in *Thermomicrobium*, functionally the equivalent of normal glycerol lipids (172), have recently been detected in *Chloroflexus* as well (T. A. Langworthy, personal communication), suggesting that a convincing phenotypic rationale for the grouping will ultimately be found.

Table 3 shows the group to have a fairly distinctive signature. The phylum is also characterized by higher-order 16S rRNA structural idiosyncracies (162). For example, the helical element between positions 1126 and 1146, a structure found in all other eubacterial sequences, is absent in the members of this group (162).

Although too few species have been characterized to project meaningful subdivisions, it would seem that *Thermomicrobium* represents one such and *Chloroflexus* and the *Herpetosiphon* species represent another (162). (Sequence homology between *Thermomicrobium* and the other species is a relatively low 77%.)

Other Eubacterial Phyla

The 10 phyla described above account for almost all of the eubacterial species whose 16S rRNA have been cataloged or sequenced. Since the characterized strains are broadly representative of the known eubacteria, it might seem that few additional eubacterial phyla, if any, will be encountered. However, isolated rRNA sequences from several unusual eubacteria suggest that such is not the case, that many eubacterial phyla remain to be described, representing species yet to be discovered.

The two small clouds on this horizon, the two eubacteria whose 16S rRNA sequences do not belong to any of the above 10 phyla, are thermophiles noted for their unusual lipids. One is *Thermodesulfotobacterium commune* (unpublished catalog), a eubacterium having ether-linked lipids (119), while the other, *Thermotoga maritima*, has unique lipids that have so far defied complete characterization (89). The 16S rRNA sequence from *T. commune*, which is incomplete, will not be treated in this review; that of *Thermotoga maritima* (1), however, will play a key role in the subsequent discussion.

Overall Structure of the Eubacterial Tree

The definition of the eubacterial phyla brings us to the limit of the rRNA cataloging approach. Unfortunately, this is

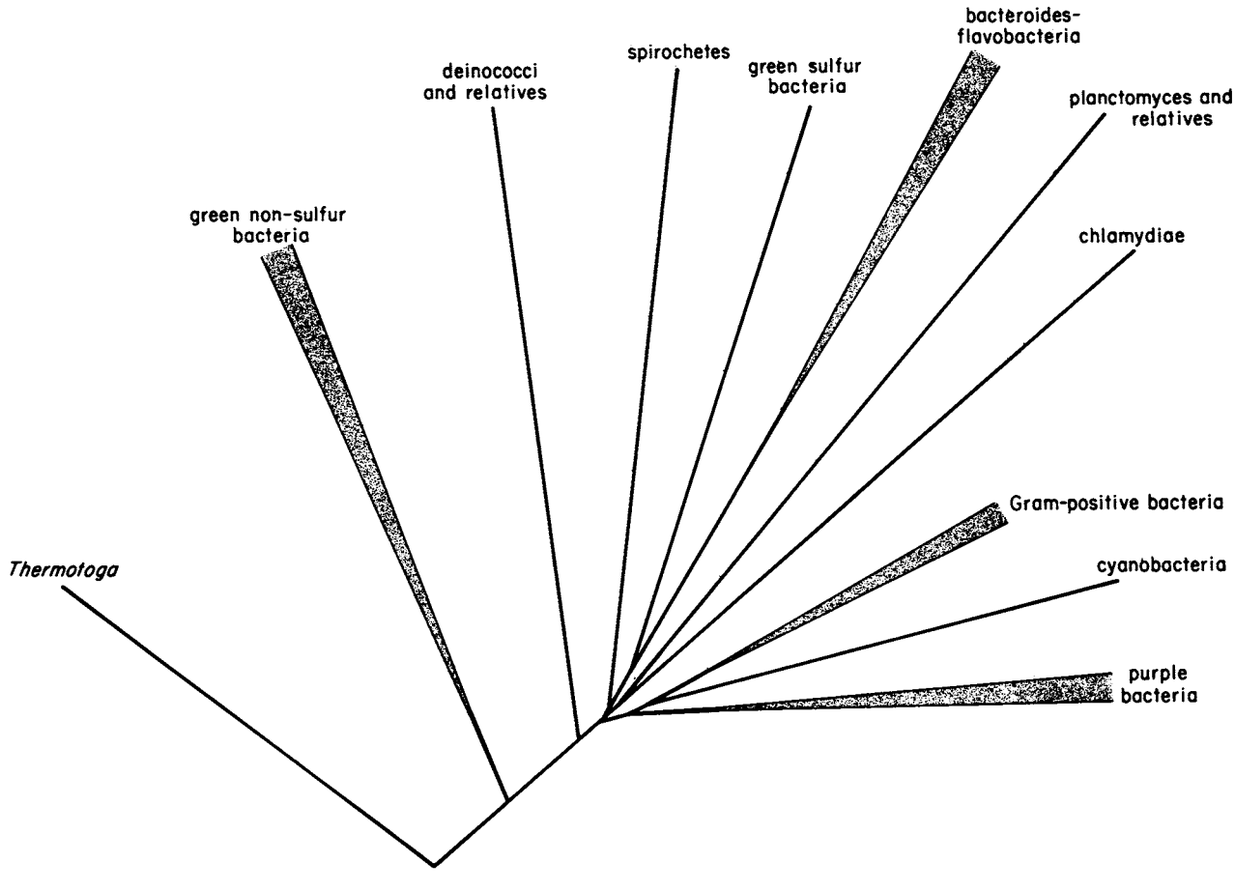

FIG. 11. Eubacterial phylogenetic tree based upon 16S rRNA sequence comparisons. An alignment was constructed (260) from one representative sequence from each of the eubacterial phyla together with an archaebacterial consensus sequence (263). Using (only) those positions represented in all sequences in the alignment, a (corrected) evolutionary distance matrix was generated (99), from which a distance tree was constructed (36). Branch lengths on the tree are proportional to calculated distances. The sequences used are: *Thermotoga maritima* (1); green non-sulfur bacteria, *Thermomicrobium roseum* (162); deinococci and relatives, *Deinococcus radiodurans* (Weisburg, Ph.D. thesis); spirochetes, *Leptonema illini* (Oyaizu, unpublished data); green sulfur bacteria, *Chlorobium vibrioforme* (Weisburg, Ph.D. thesis); bacteroides-flavobacteria, *Flavobacterium heparinum* (234); planctomyces and relatives, *Planctomyces staleyi* (Oyaizu, unpublished data); chlamydiae, *Chlamydia psittaci* (233); gram-positive bacteria, *Bacillus subtilis* (68); cyanobacteria, *Anacystis nidulans* (224); purple bacteria, *Desulfovibrio desulfuricans* (163). For those phyla in which additional 16S rRNA sequences are available, the known sequence depth of the group has been (separately) calculated and is indicated by the shaded wedges. Additional sequences added to the alignment when calculating these depths are as follows: green non-sulfur bacteria, *Chloroflexus aurantiacus* (162); bacteroides-flavobacteria, *Bacteroides fragilis* (234); gram positive bacteria, *Heliobacterium chlorum* (255); purple bacteria, *Agrobacterium tumefaciens* (274); and *Escherichia coli* (19).

the point at which the study of bacterial evolution starts to become interesting, for what we really want to know is the phylogenetic relationships among the phyla. With the aid of full 16S rRNA sequencing, it has become possible to resolve some of these branching orders, and so we can now begin to see the progression of eubacterial evolution.

Figure 11 shows the full eubacterial tree as it is presently known. This is a distance matrix tree (36), whose root has been subsequently determined by using an archaebacterial consensus sequence as an outgroup (1). We shall discuss this tree and its implications in detail below, but two important points should be noted here. First, nine of the ten phyla described above, i.e., all except the green non-sulfur bacteria and relatives, appear to stem from roughly the same small region of the tree. Second, the present root of the tree separates all ten of the recognized phyla from the single species *Thermotoga maritima* (1). The microbiologist's attention to date seems to have been confined to what phylogenetically is a nonrepresentative, sampling of the eubacteria: *Thermotoga* would appear to represent a vast unexplored "other world" of eubacteria, thermophilic organisms that (as their unique lipids suggest) almost certainly possess a variety of unusual and interesting biochemical properties and other characteristics.

It is important to stress that the root of the eubacterial tree shown in Fig. 11 is not an artifact of the treeing procedure. The same root placement results from distance-matrix and parsimony analyses. (Parsimony analysis on alignments of four sequences that comprise an outgroup sequence, the *Thermotoga* sequence, and all possible combinations of two other eubacterial sequences, show *T. maritima* in every case to cluster with the outgroup sequence, to the exclusion of the two other eubacteria [1].)

Sequence distance measurements (Fig. 5 and 11) show a

remarkable closeness of the *T. maritima* sequence to outgroup sequences, suggesting that its lineage is the most slowly evolving of all eubacterial lineages.

With the exception of the green non-sulfur group, all of the lineages branch from the eubacterial tree in such close proximity that their order of branching, the specific relationships among them, has yet to be convincingly determined. However, a few tentative relationships between various phyla are suggested and should be mentioned. A specific relationship may exist between the cyanobacteria and the gram-positive bacteria (Fig. 11). As we have seen, one signature position, 1207, in Table 3 supports this conjecture, for its composition is constant across all 150 or so gram-positive catalogs and all but one of the 30 or so cyanobacterial examples. That photosynthesis is found in both phyla and that chlorophyll *g* of *Heliobacterium* is most closely related in structure to chlorophyll *a* of the cyanobacteria (17) lend additional support to the possible relationship.

Another "superphylum" suggested in Fig. 11 involves the green sulfur bacteria and the bacteroides group. Table 3 shows several shared signature positions suggestive of that grouping, i.e., positions 995 and 1410, and a few higher-order structural features in 16S rRNA strengthen the case. The helix involving positions 1161 to 1175 (260) is altered in a way unique to these two phyla; one nucleotide is deleted from the loop in the vicinity of position 1167, while a "bulged" nucleotide is inserted in the stalk after position 1174; the structure in question can be seen in Fig. 7. (The deletion alone is seen in several other phyla, and the addition alone occurs in a particular subgroup of the α-purple bacteria [234].) In both of these phyla the penultimate helix, positions 1435 to 1466, is strongly truncated, which is rare for eubacteria (unpublished analysis). The inclusion of the spirochetes, planctomyces, and chlamydiae in this superphylum is also suggested by the Fig. 11 analysis, but this relationship should not be considered seriously without additional evidence.

ARCHAEBACTERIAL PHYLOGENY

Unusual Nature of the Archaebacterial Phenotype

Microbiologists have always perceived archaebacteria as strange, highly atypical bacteria. Prior to their recognition as a phylogenetically coherent grouping (6, 261), however, their individual idiosyncrasies were interpreted merely as adaptations: the lipids of *Thermoplasma* were unusual because the organism evolved to live at high temperatures or in highly acidic environments or both (16); the wall of *Halococcus* was an adaptation to an extremely saline environment (123, 185); the uniqueness of their coenzymes merely reflected the capacity of methanogens to produce methane from carbon dioxide (277). That different archaebacteria had the same unusual lipids was even interpreted as convergent adaptation (16)! As we find out more about the archaebacteria our sense of their strangeness increases, but its explanation lies in a shared ancestry, not in individually evolved idiosyncrasies.

The archaebacteria as we know them today are a collection of disparate phenotypes: the methanogens, the extreme halophiles, and the extremely thermophilic sulfur-metabolizing species (250, 270). Their metabolic differences are many, their known similarities few. The methanogens are noted for unusual coenzymes (Fig. 12), which tend not to occur in other bacteria. The extreme halophiles are the only photosynthetic archaebacteria; they transduce light into chemical energy by means of a proton pump based on bacteriorhodopsin, a pigment unique to this group of organisms (122, 217). The halobacteria also possess remarkably high intracellular salt concentrations (113). (Some methanogens have impressively high internal salt concentrations as well [95], but not in the range characteristic of the halophiles.) The extreme thermophiles also have at least one unique coenzyme (28, 34) (Fig. 12). They share with methanogens a capacity to reduce large amounts of sulfur (213). The extreme thermophiles have no immediate known relatives that grow at or near normal temperatures; most of them grow best at remarkably high temperatures (212, 215). Systematic comparisons of archaebacterial metabolism are definitely needed.

The branched-chain, ether-linked lipids common to all archaebacteria are found nowhere else in nature (35, 69, 103, 117–120). In many, but not all species the glycerol diethers tend to be covalently joined "head to head," to produce diglycerol tetraethers, which form unusual membranes that cannot be freeze-fractured (118). Archaebacteria show an impressive number of variations in lipid structure based upon the ether-linked, branched-chain theme (35, 69, 117–119).

Archaebacteria display their own characteristic version of every major macromolecular function, e.g., 16S rRNA (Fig. 6). However, within these versions an impressive spectrum of variation can occur (250). For example, unlike eubacteria, their walls are varied in type (101, 109). In their structural details some archaebacterial 5S rRNAs resemble somewhat the eubacterial form of the molecule and others resemble the eucaryotic form, while still others are unique (53, 133, 208). All three urkingdoms have a characteristic subunit pattern for DNA-dependent RNA polymerases, but within the archaebacteria relatively drastic variations in type are seen (187, 282). For the extremely thermophilic species and *Thermoplasma acidophilum*, the so-called B subunit is a single large molecule, while for methanogens and their relatives it consists of two smaller molecules, B′ and B″ (187, 282). In their transfer RNAs (tRNAs) archaebacteria show a characteristic pattern of modified bases (70, 71, 73, 146, 165, 273), yet the tRNAs of the extreme thermophiles (except for *Thermoplasma*) are much more highly modified than are those of the methanogens and relatives (250, 259).

No new major archaebacterial phenotypes were discovered in the 9 years following the recognition of archaebacteria as a distinct group, which leads to a growing feeling that the three basic and highly unique phenotypes, methanogenesis, extreme halophilism, and extremely thermophilic sulfur metabolism, are all that exist in the kingdom. (However, see below.)

Definition of the Major Archaebacterial Groups

The number of archaebacterial species characterized by rRNA analysis is only one-tenth the number of characterized eubacterial species. Although their number is not large enough to provide phylogenetic detail, it is sufficient to identify the higher archaebacterial taxa, for the sampling is broadly representative.

rRNA cataloging studies showed that there are three major groups of methanogens and one of extreme halophiles (5, 52, 54, 56, 250), each being the equivalent of a eubacterial phylum. Although the branching order among the four phyla could not be determined by this method, the four as a group were easily distinguished from the group of extremely thermophilic archaebacteria (excluding *Thermoplasma*). However, too few species of extreme thermophiles were

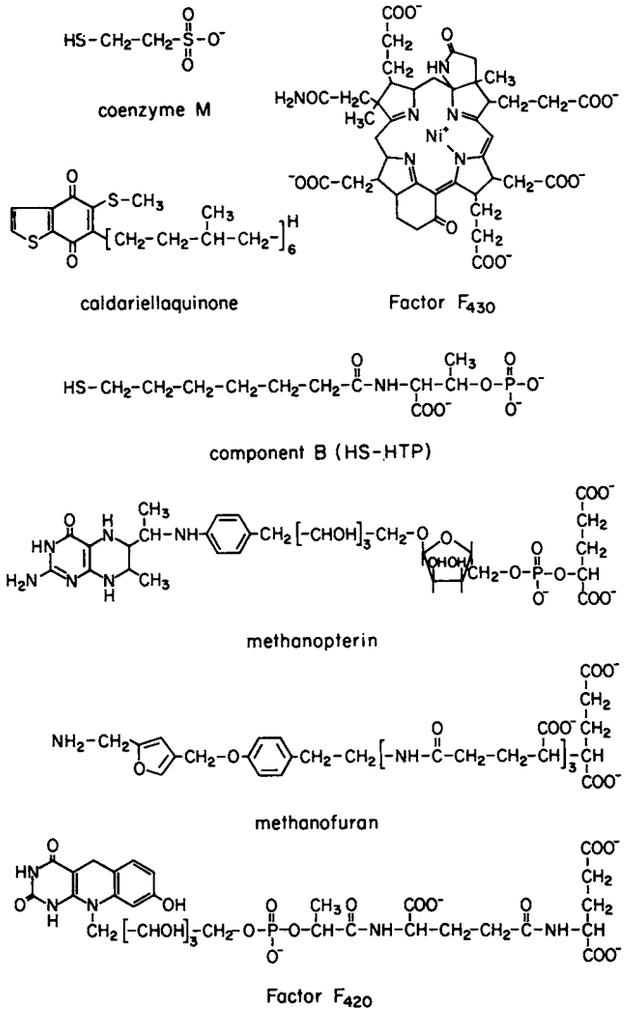

FIG. 12. Structure of various archaebacterial cofactors (34, 154, 238). All are found in methanogens and utilized in methane production, except for caldariella quinone, which occurs in *Sulfolobus* species.

characterized by the cataloging method to give a sense of that group's phylogeny (225, 259).

Table 14 lists the archaebacterial species that have been characterized by the rRNA method, in an approximate phylogenetic ordering. Tables 15 and 16 define the major archaebacterial groupings by sequence signature.

Methanogens. The three methanogen phyla defined by rRNA sequence comparisons can also be recognized by morphological criteria, with a few exceptions (5, 238). Immunological cross-reactivities almost always identify methanogen group affiliation as well (135). As a result, we know that each of the three main groups of methanogens (currently designated orders), *Methanobacteriales*, *Methanococcales*, and *Methanomicrobiales*, contains a large number of species (5, 98, 135, 238, 239). To date, the immunological studies have given no clear indication that any additional major methanogen group exists (135). The three methanogen phyla are also distinguishable by 5S rRNA type (53).

In two of the methanogen phyla, subdivisions are recognized by rRNA cataloging. The *Methanobacteriales* breaks naturally into two "genera," *Methanobacterium* and *Methanobrevibacter* (5, 238). A third genus, now represented by the lone species *Methanothermus fervidus* (214), should ultimately be declared a distinct unit within this phylum, or perhaps a separate phylum (238). The *Methanomicrobiales* in turn divide into two distinct subdivisions, (formally families) the *Methanomicrobiaceae* and the *Methanosarcinaceae* (5, 238). The latter is the most metabolically unusual of the methanogen groups. Its species can utilize acetate or sometimes methyl amines in methane production; some are even unable to use carbon dioxide for this purpose (9, 42, 153). These are also the only methanogen species that contain cytochromes (*b* or *c* or both) (112). The unusual halophilic methanogens belong to this group as well (13; I. M. Mathrani, D. R. Boone, and R. A. Mah, Abstr. Annu. Meet. Am. Soc. Microbiol. 1985, I85, p. 160).

Extreme halophiles. The extreme halophiles constitute one

TABLE 14. Archaebacterial subdivisions, representative genera, and species[a]

Methanococcus group
Methanococcus vannielii
M. voltae
M. maripaludis
M. thermolithotrophicus
M. jannaschii
Methanobacter group
Methanobacterium formicicum
M. bryantii
M. thermoautotrophicum
Methanobrevibacter smithii
M. arboriphilus
M. ruminantium
Methanosphaera stadtmaniae
Methanothermus fervidus
Methanomicrobium group
Methanosarcina barkeri
Methanococcoides methylutens
Methanothrix soehngenii
Methanospirillum hungatei
Methanomicrobium mobile
Methanomicrobium paynteri
Methanogenium cariaci
Methanogenium marisnigri
Methanoplanus limicola
Halobacteria
Halobacterium volcanii
H. cutirubrum[b]
H. sodomense
H. trypanicum
Halococcus morrhuae
Thermoplasma
Thermoplasma acidophilum
Thermococcus group
Thermococcus celer
Extreme thermophiles
Sulfolobus solfataricus
S. acidocaldarius
Thermoproteus tenax
Desulfurococcus mobilis
Pyrodictium occultum

[a] References 5, 98, 238, and 263. Approximate phylogenetic clustering suggested by indentation.

[b] 16S rRNA identical to those of *H. salinarium* and *H. halobium*.

of the most distinctive groups of bacteria known. As mentioned, both their very high internal salt concentrations (i.e., in the range of 5 M potassium ion [113]) and their mechanism for photoproduction of energy are unique in the bacterial world. (Bacteriorhodopsin does superficially resemble the eucaryotic visual pigment, however, hence its name [122, 217].) Over ten species of extreme halophiles have been characterized in terms of 16S rRNAs, and they form a relatively close-knit grouping (56, 72, 90, 125; C. R. Woese and G. E. Fox, unpublished data). The unusual halophiles that grow under alkaline conditions (222) are among them. The group is not a particularly deep one. Based upon relative S_{AB} values (Woese and Fox, unpublished data) and known sequences, all rRNAs in the group show at least 87% sequence similarity. The internal phylogenetic structure of the group is unspectacular.

The extreme halophiles are known to contain cytochromes and ferredoxins (78, 113, 121). It has been reported that halophile ferredoxin sequences are specifically related to those found in cyanobacteria, to the exclusion of other eubacteria (80). If such homology exists, it is highly unlikely to reflect a genuine phylogenetic relationship; the cyanobacteria and extreme halophiles are not related to one another to the exclusion of the other members of their respective kingdoms. Sequence convergence also seems unlikely. However, gene transfer does not, for the two types of organisms can in some cases share the same habitat.

The cataloging approach failed to distinguish the branching order among the three groups of methanogens and the extreme halophiles. Most microbiologists tacitly assumed that the four phyla were arranged along phenotypic lines; i.e., all methanogens clustered together to the exclusion of the halophiles. However, oligonucleotide signatures weakly suggested a specific relationship between the halophiles and the *Methanomicrobiales* (275).

Extreme thermophiles. The extreme thermophiles are the least characterized, but (as will be seen below) the most evolutionarily interesting, of the archaebacteria. The extreme thermophiles seem quite uniform in phenotype. All species grow anaerobically, and most require sulfur as an energy source (215). A minority of species can also grow aerobically, and some that use sulfur as an energy source do not require it (215). Most species grow best at extremely high temperatures, some near the boiling point of water (212).

The extreme thermophiles differ from other archaebacteria in numerous ways. Their modes of division tend to be unusual and varied (281), as do their viruses (279, 281); their ribosomal subunits have an unusual shape (see discussion below) (81); they have (as mentioned) at least one unique coenzyme; they seem to be insensitive to most antibiotics (22); and both their tRNAs and rRNAs are highly modified (the level in the latter case is fivefold greater than that seen in the methanogens and their relatives) (259). Nevertheless, in sequence terms, in membrane structure, and in most phenotypic characteristics, the extreme thermophiles are definitely archaebacteria (141, 142, 215, 263).

The phenotypic clustering of extreme thermophiles is deceptive. As we shall see, it is not a phylogenetic clustering. rRNA cataloging studies showed that *Thermoplasma* was more closely related to the methanogens and relatives than to the extreme thermophiles (56). Even its pattern of 16S rRNA base modification is "methanogen-like" (56, 262). However, this potential relationship was at first given little credence, and the organism tended to be viewed as an atypical extreme thermophile (215). Although an rRNA catalog did not exist for *Thermococcus*, there seemed every reason to assume it was related to the other extreme thermophiles (215), but, as we shall see, this is not so.

The Archaebacterial Tree: Its Branching Order and Root

An evolutionary distance matrix for the known archaebacterial 16S rRNA sequences is shown in Table 17. A phylogenetic tree derived from evolutionary distances is shown in Fig. 13 (263). With two minor exceptions, the same branching order is given by a number of different methods: parsimony analysis, distance-matrix treeing, using subsets of the positions in the sequence alignment, etc. (263; unpublished analysis). One of the exceptions, *Sulfolobus solfataricus*, tends to cluster with *Desulfurococcus mobilis* by parsimony analysis and when no or few outgroups are included in a distance treeing analysis; yet when the full set of archaebacterial sequences is analyzed by a distance-matrix method, *D. mobilis* clusters instead with *Pyrodictium occultum*, to the exclusion of *S. solfataricus*. The *S. solfataricus* line of descent is more rapidly evolving than are these others (as can be inferred from Table 17), which tends to force its branching deeper into the tree than it actually is, particularly when outgroup sequences are included in the analysis.

For the same reason the exact branching order for *Thermoplasma acidophilum*, the second exception, is uncertain, though its lineage always remains in the general vicinity of the position shown in Fig. 13.

Main branches. The root of the archaebacterial tree divides the urkingdom into two main lineages: a cluster of extreme thermophiles on the one hand, and the methanogens and their relatives on the other.

The two main branches differ from each other in several interesting ways. For one, the methanogen branch appears to be the phylogenetically "deeper" of the two. For another, the extreme thermophile branch (so far) is phenotypically pure, whereas its counterpart is cosmopolitan, including a mixture of methanogenic, extremely halophilic, and extremely thermophilic phenotypes.

TABLE 15. Sequence signature distinguishing the two main archaebacterial branches

Position[a]	Composition in:			
	Methanogens and relatives[b]	*Thermoplasma acid ophilum*[c]	*Thermococcus celer*[d]	Sulfur-dependent archaebacteria[e]
34	U	?	—[f]	C
242	G	U	—	C
559	U[g]	—	—	G
939	C	—	G	G
965	Y	—	—	G
1074	A	—	—	G
1088	U	A	A	G
1252	U	—	—	C
1351	U	—	C	C
1408	A	—	—	G

[a] In 16S rRNA sequence (260).
[b] Based upon oligonucleotide catalogs or sequences for 21 species of methanogens, and 9 species of halophiles (5, 6, 54, 72, 91, 125; Woese, unpublished data).
[c] Based upon catalog (262) and an unpublished sequence (R. A. Zimmermann, personal communication).
[d] Based upon the sequence (Woese et al., unpublished data).
[e] Based upon oligonucleotide catalogs or sequences for five sulfur-dependent archaebacteria (126, 158, 225; R. Garrett, personal communication; Woese et al., unpublished data).
[f] —, Same composition as in methanogens and relatives.
[g] One example of G at this position (5).

TABLE 16. Sequence signatures defining the methanogen and extreme halophile groups[a]

Group	Position[b]	Major base	Composition in: Mco[c]	Mba[d]	Mmi[e]	Hal[f]	Tac[g]	Tce[h]
Methanococcus	319	A	G	—[i]	—	—	—	—
	334	U	C	—	—	—	—	A
	703	A	G	—	—	—	—	—
	976	A	G	—	—	—	—	G
	1122	Y	A	—	—	—	—	—
	1150	U	A	—	—	—	—	—
	1366	G	C	—	—	—	C	—
	1377	A	G	—	—	—	—	—
Methanobacterium and relatives	1127	G	—	A	—	—	—	—
	1167	C	—	A	Y	—	—	A
	1342	C	—	G	—	—	—	—
	1357	U	—	G	—	—	G	G
	1365	A	—	C	—	—	C	C
	1417	G	—	A	—	—	—	—
Methanospirillum and relatives	271	C	—	—	A	—	—	—
	366	C	—	—	A	—	—	—
	807	U	—	—	C	—	—	—
	1076	C	—	—	U	—	—	—
	1084	G	—	—	A	—	—	—
	1122	U	A	—	C	—	—	C
Extreme halophiles	33	C	—	—	—	U	?	—
	50	A	—	—	—	Y	?	—
	274	A	—	—	—	U	—	—
	374	A	—	—	—	C	—	—
	940	G	—	—	—	C	—	—
	1291	C	—	—	—	U	—	—
	1317	U	—	—	—	A	—	—
	1343	C	—	—	—	G	—	—

[a] See footnotes to Table 15 for details.
[b] In 16S rRNA sequence.
[c] Mco, *Methanococcus*, five species (5, 96, 98).
[d] Mba, *Methanobacter* and relatives, eight species (5, 124; unpublished data).
[e] Mmi, *Methanospirillum* and relatives, eight species (5, 275; unpublished data).
[f] Hal, Extreme halophiles, nine species (72, 91, 125; unpublished data).
[g] Tac, *Thermoplasma acidophilum* (K. M. Cao, H. Ree, D. L. Thurlow, and R. A. Zimmermann, personal communication).
[h] Tce, *Thermococcus celer* (unpublished data).
[i] —, Composition equals major base.

TABLE 17. Percent similarities and evolutionary distances for archaebacterial 16S rRNAs[a]

Species	*Mc. vannielii*	*M. formicicum*	*Ms. hungatei*	*Methanosarcina* sp. strain WH-1	*H. volcanii*	*T. acidophilum*	*Tc. celer*	*S. solfataricus*	*Tp. tenax*	*D. mobills*	*P. occultum*
Mc. vannielii		20.3[b]	27.7	26.4	27.8	30.0	19.7	28.6	29.0	24.2	25.5
M. formicicum	69.2[c]		23.0	25.0	25.5	29.3	20.0	30.0	28.0	24.1	24.8
Ms. hungatei	59.9	65.7		19.8	23.6	33.6	26.3	33.6	32.9	30.4	30.8
Mr. sp. strain WH-1	61.5	63.2	69.8		26.3	30.0	24.7	32.3	30.9	29.8	28.5
H. volcanii	59.7	62.6	64.9	61.6		31.5	27.2	35.6	33.9	32.2	30.8
T. acidophilum	57.2	58.0	53.1	57.2	55.5		26.4	33.9	34.6	33.3	31.9
Tc. celer	70.0	69.5	61.6	63.5	60.5	61.5		24.5	20.0	19.1	17.7
S. solfataricus	58.8	57.2	53.1	54.5	50.9	52.8	63.8		17.1	12.4	12.8
Tp. tenax	58.3	59.6	53.9	56.1	52.8	52.0	69.5	73.5		14.1	11.7
D. mobilis	64.1	64.3	56.7	57.3	54.7	53.4	70.8	80.3	77.7		6.9
P. occultum	62.6	63.3	56.2	58.9	56.2	55.0	72.7	79.6	81.2	88.6	

[a] Sequence references are cited in Fig. 13 legend. *Mc.*, *Methanococcus*; *M.*, *Methanobacterium*; *Ms.*, *Methanospirillum*; *Mr.*, *Methanosarcina*; *H.*, *Halobacterium*; *T. Thermoplasma*; *Tc.*, *Thermococcus*; *S.*, *Sulfolobus*; *Tp.*, *Thermoproteus*; *D.*, *Desulfurococcus*; *P.*, *Pyrodictium*.
[b] Upper right numbers are evolutionary distances (99); only positions in alignment represented in all sequences are used in calculation.
[c] Lower left numbers are percent similarity; all positions not represented in all sequences and all positions of constant composition were removed from consideration. (This last has been done to accentuate the differences among sequences; it does not change their rank order.)

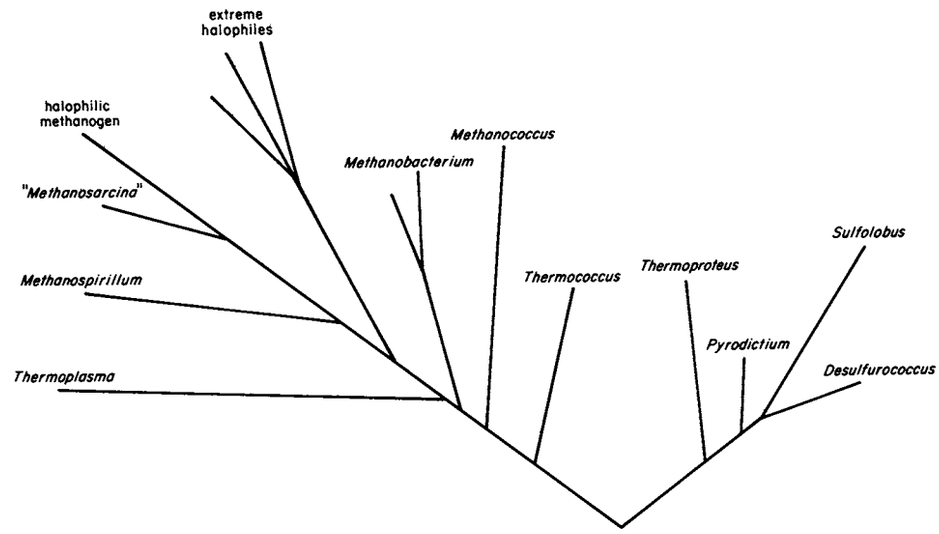

FIG. 13. Archaebacterial phylogenetic tree based upon 16S rRNA sequence comparisons. The sequences listed were aligned and an unrooted distance tree was constructed as in the legend to Fig. 11. Its root was subsequently imposed on the basis of outgroup consensus sequences (eubacterial and eucaryotic); the root given by eubacteria or eucaryotes is in the same general region of the tree (i.e., between *Thermococcus* and the other extremely thermophilic species [263], and that shown represents an average of the eubacterial and eucaryotic placements. Those sequences used in the alignment are *Methanospirillum hungatei* (275); the halobacteria *Halobacterium volcanii* (72), *Halococcus morrhuae* (125), and *Halobacterium cutirubrum* (91), from left to right in that order; *Methanosarcina* sp. strain WH-1 (P. Rouviere, unpublished data; Woese et al., unpublished data); halophilic methanogen strain FS-1 (Mathrani et al., Abstr. Annu. Meet. Am. Soc. Microbiol. 1985; Woese et al., unpublished data), *Methanobacterium thermoautotrophicum* (R. Garrett, personal communication) *and Methanobacterium formicicum* (124), from left to right, representing *Methanobacterium*; *Methanococcus vannielii* (96); *Thermococcus celer* (Woese et al., unpublished data); *Thermoproteus tenax* (126); *Pyrodictium occultum* (Woese et al., unpublished data); *Sulfolobus solfataricus* (158); and *Desulfurococcus mobilis* (Garrett, personal communication).

***Thermococcus* and *Thermoplasma*.** The two extreme thermophiles *Thermococcus* and *Thermoplasma* are unrelated to their phenotypic counterparts. As we have seen, their phylogenetic placement with the methanogens is consistent in the case of *Thermoplasma acidophilum* with the pattern of modified nucleotides in the 16S rRNA, i.e., a low level of modifications, at particular sites (5, 262). (Nothing is known yet about the pattern of modified nucleotides in *Thermococcus* RNAs.)

Interestingly, *Thermococcus celer* is not closer to the methanogens than to the other extreme thermophiles by overall sequence distance measure. Table 17 shows its 16S rRNA to be closest to that of *Pyrodictium occultum*. This would seem, however, to reflect the relatively slow evolutionary tempo among extreme thermophiles in general, not a specific relationship between these two organisms, a point that can be clearly demonstrated by signature analysis, which focuses on the more conserved positions in the molecule. For example, in a 16S rRNA alignment containing sequences from eight methanogens, three extreme halophiles, and four representatives of the extreme thermophile branch, there are about 30 positions that have a constant composition among the methanogens and extreme halophiles but a different (constant) composition among the extreme thermophiles. The *Thermococcus celer* sequence exhibits the characteristic methanogen composition in about 79% of these cases; the extreme thermophile composition in only 11% (263; Woese, unpublished analysis).

That *Thermococcus* and *Thermoplasma* resemble other extremely thermophilic archaebacteria in DNA-dependent RNA polymerase subunit pattern (187, 282) does not necessarily prove genealogical relatedness. Zillig and co-workers have shown that the large B subunit of RNA polymerase, found in all extreme thermophiles, is undoubtedly the ancestral type, for the smaller B′ and B″ subunits appear to have arisen from it at least twice (281). For this reason the common occurrence of a large RNA polymerase B subunit does not necessarily mean specific relationship.

The branching of *Thermococcus* from the main methanogen line of descent is sufficiently deep to suggest that it may ultimately be considered to represent a third major archaebacterial lineage.

Branching order among the methanogens and extreme halophiles. Full sequences show that the extreme halophiles cluster specifically with the *Methanomicrobiales*, to the exclusion of the other two phyla of methanogens, a relationship that was hinted at by oligonucleotide signatures (263, 275). Although this is an unexpected, even counterintuitive, finding, the evidence supporting it is entirely convincing. The conclusion readily emerges from parsimony analyses of 16S rRNAs as well as distance treeing (263). For example, in an alignment that includes sequences from the three extreme halophiles, one from each of the methanogen phyla, and the *Thermococcus* sequence, there are 22 positions having a common composition in the halophile and *Methanospirillum* sequences that have a different common composition in the remaining three sequences. For any other combination of these sequences (keeping the three halophile sequences as a unit), there are no more than seven positions of common composition defined in this way (263; unpublished analysis).

Although the relationship between the extreme halophiles and the *Methanomicrobiales* is difficult to justify phenotypically, certain facts are consistent with such a grouping. For example, methanogens capable of growth at high salt concentrations belong to the *Methanomicrobiales* (Fig. 13). On the methanogen branch of the tree, it is only among species

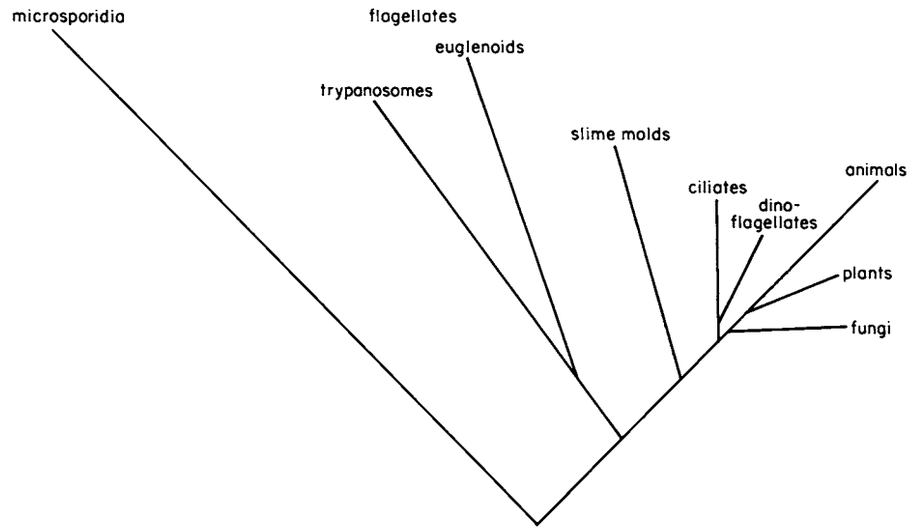

FIG. 14. Eucaryotic phylogenetic tree based upon 16S-like rRNA sequence comparisons. The eucaryotic sequences listed were aligned and a distance matrix tree was constructed as in the legend to Fig. 11. The root was determined by including an archaebacterial consensus sequence in the alignment. Those eucaryotic sequences used are as follows: microsporidia, *Vairimorpha necatrix* (226a); flagellates, *Trypanosoma brucei* and *Euglena gracilis* (196); slime molds, *Dictyostelium discoideum* (145); ciliates, *Paramecium tetraurelia* (195); dinoflagellates, *Prorocentrum micans* (84); fungi, *Saccharomyces cerevisiae* (179), animals, *Xenopus laevis* (181); plants, *Zea mays* (147). The archaebacterial (100%) consensus sequence was derived from an alignment of three methanogens and three extreme thermophiles.

of *Methanosarcina*, the extreme halophiles, and *Thermoplasma* that cytochrome *b* or *c* or both are found (78, 86, 112, 121, 190).

Figure 13 also indicates a specific relationship between the *Methanobacteriales* and the halophile-*Methanomicrobiales* group, to the exclusion of the *Methanococcales*. This relationship, too, can be rationalized by sequence signature (263). For example, the *Methanobacterium formicicum* sequence shares more than twice as many positions of exclusive common sequence with the *Methanospirillum hungatei* and extreme halophile sequences as does the *Methanococcus vannielii* sequence (263).

The position of *Thermoplasma* in the tree, as mentioned, is uncertain. Various treeing procedures place its branching in a range that extends from somewhere on the common *Methanospirillum*-halophile branch to just below the *Methanococcus* branch (263). A signature marginally suggestive of the first placement exists (263), but the exact genealogy of the organism should be considered uncertain.

A BRIEF LOOK AT EUCARYOTE PHYLOGENY

The biologist generally feels that he has a relatively good sense of eucaryotic evolution, and up to a point this is certainly true. Detailed taxonomies exist for the various classes of animals and plants that rather accurately reflect their phylogenies. The higher levels of eucaryote classification are another matter, however. Here our understanding is no less a matter of prejudice and preconception than it was for the bacteria. The so-called Five Kingdom classification (139, 240, 241), plants, animals, fungi, protists, and monera, cannot be considered proper phylogeny. It mixes apples and oranges and defines categories by exclusion. The system gives the same kingdom rank to each of the four groups of eucaryotes that it gives to one group of procaryotes (monera). Yet it has been obvious for some time that the four eucaryotic kingdoms form a phylogenetically coherent unit that as a whole ranks with the monera; and monera, of course, is now known to comprise two separate kingdoms (56). Within the eucaryotes, the protists do not form a phylogenetically coherent unit (196).

Fortunately, this scheme and our criticism of it will soon be rendered academic. A new and very different view of eucaryotic phylogeny is beginning to emerge, whose outlines can be seen in Fig. 14.

Implications of the Eucaryote Phylogenetic Tree

What this preliminary phylogeny begins to suggest is that the major epochs in eucaryote evolution corresponded to major periods in earth history. A relatively "recent" period of massive evolutionary radiation appears to have given rise to most of the major eucaryotic lineages: green plants, animals, fungi, ciliates, dinoflagellates, (some) amoebae, etc. (196). (The cellular slime molds represent a slightly earlier branching [145, 196].) It is tempting to equate the onset of this particular radiation with a globally significant event and attribute the radiation to some major innovative biological (evolutionary) response. The hydrosphere is thought to have become oxidizing about 1.5 billion years ago (228), which time is roughly consistent with the occurrence of the radiation estimated by back-extrapolation from known time points on the eucaryotic tree (196). Biologically this could have been a time when an oxygen-utilizing mitochondrion developed (274).

Deeper lineages occur on the eucaryotic tree, however, that appear to predate significantly this period of massive radiation. A group of flagellates is seen to branch from the main eucaryotic stem well before the radiation in question (196), and the microsporidian lineage definitely emerges well before that (226a). These earlier branchings should reflect earlier phases in the planet's history. The flagellate branching might stem from the "microaerobic" or "amphiaerobic" period, i.e., the era between the time the atmosphere be-

came significantly oxidizing (over 2 billion years ago) and the hydrosphere became aerobic (about 1.5 billion years ago) (228). The microsporidian branching would represent the still earlier genuinely anaerobic period, before 2.5 or so billion years ago (226a).

Microsporidia, e.g., the genus *Vairimorpha*, are a group of highly unusual and little studied unicellular eucaryotes, deserving of more than passing attention. They have primitive modes of cell division (175) and exhibit strange and interesting life cycles, connected with their obligately parasitic mode of existence. The group as a whole parasitizes an extremely wide range of other eucaryotes; they have been seen to infect examples of all the animal phyla, and they even parasitize other protists (197). This could be interpreted to mean that their parasitism is of very ancient origin. Microsporidia have no mitochondria and, given their deep phylogenetic branching, might never have had them.

Perhaps their molecular idiosyncrasies are the most fascinating aspect of the microsporidia (and very little is known about this). These are the only eucaryotes that have no 5.8S rRNA (227). Their rRNAs are also far smaller than normal eucaryotic rRNAs. Typically, the eucaryotic small-subunit rRNA comprises about 1,800 nucleotides, but its microsporidian counterpart contains under 1,300 nucleotides, even less than the roughly 1,450 to 1,550 nucleotides characteristic of procaryotes (75, 226a). Interestingly, the missing areas in the microsporidian small-subunit rRNA tend to be those that are unique to the eucaryotes (75) (Fig. 5).

With the microsporidian branching we may be near the base of the eucaryotic tree, the beginnings of eucaryotic cellular evolution. Microsporidia are defined as eucaryotes because they have a membrane-delimited nucleus. The question is, in what other ways do they resemble eucaryotes, and what characteristic eucaryotic features do they lack? Eucaryotes would appear to be an old group, far older than many biologists might have thought. Their antiquity would seem to rival that of the procaryotic kingdoms.

NATURE OF THE EVOLUTIONARY PROCESS IN BACTERIA

Relationship Between Evolution's Tempo and Its Mode

Classical evolutionists recognized that a relationship existed between the rate at which evolution proceeds, its tempo, and the quality of the changes that occurred, its mode (143, 192). Fossil evidence showed that some lineages evolve more rapidly than others and that rates of phenotypic change vary within lineages at different stages in their history (192). Evolution tended to be particularly rapid as a lineage came into being and also in some cases as it died out (192). The quality of phenotypic change was different during such periods of rapid evolution; it was often described as drastic and novel, even bizarre (192).

Two general rules governed the relationship between the tempo and mode of metazoan evolution: (i) true evolutionary novelty (of the kind that gives rise to major groups) occurred only during times of rapid evolution, and (ii) rapid evolution tended to be episodic, not chronic. (Evolution of the horse was often used as an example. A relatively short period of dramatic evolutionary change transformed the common ancestor of horses, tapirs, and rhinoceroses into a horselike creature. The evolution of this ancestral horse into the modern form, which was a far more protracted affair, involved relatively little further change [192].) Two other characteristics of rapid evolution are its instability and radiation. Of the numerous lineages typically formed from a common ancestral stock during these saltatory episodes, many, if not most, were short-lived (192). The origin of the animal phyla conformed to such a pattern—all seem to have burst forth, almost simultaneously (192). The same can now be said of the origin of the main eucaryotic kingdoms (Fig. 14).

Evolutionists have debated the whys and wherefores of the tempo-mode issue for decades. In the past discussion centered about whether or not the same evolutionary mechanisms or environmental conditions underlay the chronic progressive evolution which characterized normal established lineages, usually referred to as microevolution, that underlay rapid episodic evolution, variously called macroevolution, megaevolution, or quantum evolution (192). Initially, global catastrophes or elevated mutation rates had been invoked to explain the radiating, saltatory origin of major taxa, and some biologists went so far as to declare macroevolution (megaevolution) and microevolution to be different in kind (66). However, the idea that catastrophes played a necessary or even major role in radical evolutionary change was later rejected (192), and elevated mutation rates were no longer seen as required for episodic, drastic evolutionary change; in fact, they could not explain it (192). Rapid evolution, macroevolution, was solely the result of ecological considerations, population genetics: small population sizes, rapidly changing environments, untoward conditions, and the like, were all that need be invoked (143, 192, 272). (However, note that we are seeing a return in recent times to the global catastrophe type of explanation for radical evolutionary change; good evidence now supports the idea that the effects of a comet or asteroid impact lead to the extinction of the dinosaurs, thereby allowing the subsequent evolutionary radiation of the mammals [3].)

Some biologists today seem to feel that microevolution-macroevolution is a nonissue, the difference between them being only a matter of degree. All distinctions, all boundaries, however, are matters of degree when viewed finely enough; this is especially apparent for protracted processes such as evolution. The significance of a distinction turns not on whether it is a matter of degree, but on how sharp the boundary is relative to the space/time scale of the phenomena that define it.

One thing is clear about the tempo-mode issue: it will never be resolved if its study is confined to fossil evidence. In these terms the crucial parameters are too poorly defined and distinguished; the phenomena are too illusive and inaccessible and too difficult to explain. It is important, therefore, to try to generalize the tempo-mode problem to bacterial systems and to study it in molecular terms.

(In the following discussion I will use the term "macroevolution" to mean the episodic, saltatory, radiating type of evolution that can create major taxa and is often associated with instability in the newly formed lineages [192]. Although this may not be the strict conventional usage of the term, there should be no problem as long as its usage is understood.)

Macroevolution at the Molecular Level: the Mycoplasmas and Their rRNAs

Changes in molecular sequence are the most basic manifestation of evolution's tempo. Molecular chronometers, as

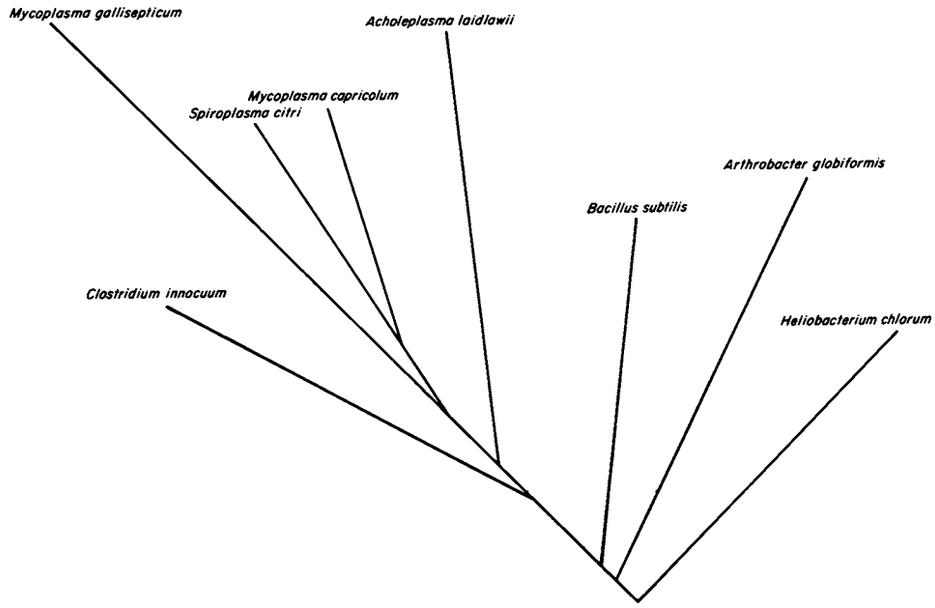

FIG. 15. Phylogenetic tree for the mycoplasmas and other members of the gram-positive bacteria (see Table 18 footnotes), based upon 16S rRNA sequence comparisons. An alignment consisting of the sequences shown was used to construct an unrooted distance matrix tree as in the legend to Fig. 11. The root was imposed by using these sequences: *Anacystis nidulans* (224), *Desulfovibrio desulfuricans* (163), *Agribacterium tumefaciens* (274), and *Escherichia coli* (19).

we have seen, can measure a pure tempo, essentially unaffected by the mode, the overlying phenotypic changes. Unfortunately, sequence distances do not provide as good a measure of evolutionary rate as we would like. They provide only average rates, over relatively long evolutionary time spans, when what is required for a proper formulation of the tempo-mode problem in molecular terms are measures of changes in rate during the course of a lineage's evolution, as are found in the metazoan fossil record. However, as will be seen, the "chronometric structure" of rRNAs (and presumably other macromolecules) is such that not only average rates, but also indications of rate changes (i.e., peak rates) are recorded. The main problem we appear to face, then, is finding the molecular counterpart of that ill-defined quality, evolutionary mode, if such exists. This problem too will prove tractable, for on the molecular level evolutionary tempo and mode are intimately connected; they are essentially different manifestations of the same process.

If macroevolution occurs among bacteria, it should be most evident in the most rapidly evolving bacterial lineages. Two eubacterial lineages in particular are attractive from this point of view: the planctomyces group (202) and the mycoplasmas. Of these, the mycoplasmas are at least as rapidly evolving as the planctomyces. And, what is more important for the present discussion, they have known close relatives that evolve at normal rates, e.g., *Bacillus* (262). Our discussion will therefore focus on mycoplasmas.

Mycoplasmas show the main characteristic expected of rapidly evolving lineages, an unusual phenotype, the nature of which has puzzled microbiologists for decades (176). These organisms have no cell walls; the cell membrane is their outer boundary. They have a number of cytological and biochemical peculiarities, and their genomes are far smaller than normal bacterial genomes (137, 176). Some microbiologists took their unusual phenotypes to mean that mycoplasmas were extremely primitive, not at all related to ordinary bacteria (229); others saw mycoplasmas merely as degenerate forms of certain normal bacteria (176). The current consensus among microbiologists, reflected in mycoplasma classification, seems to be that these organisms (with the exception of *Thermoplasma*, an archaebacterium) constitute a phylogenetically distinct group of highly unusual bacteria that is distantly related to the eubacteria (62, 176).

However, mycoplasmas are not distinctive genealogically. By rRNA measure they are merely gram-positive eubacteria. They and their relatives, as seen above, reside high in the gram-positive tree, as one subline of a particular subgroup in the low-G+C subdivision of that phylum (178, 262, 265). And, as Fig. 15 shows, the mycoplasmas have specific clostridial relatives, for example, *Clostridium innocuum* (262, 265). In other words, mycoplasmas seem unusual not because of a remote phylogenetic position, but because their mode of evolution has for some reason been atypical.

Despite their mundane genealogy, the rRNAs of mycoplasmas are definitely unique (262, 265). As Fig. 15 and Table 18 show, mycoplasma rRNA sequences change more rapidly than do those of normal eubacteria, but more importantly, certain positions in the 16S rRNA sequence that tend to be invariant in composition, what are normally the "phylogenetically uninformative" positions, show significant variation in the mycoplasma sequences (265) (Table 19). Since oligonucleotides representing the regions of highly conserved sequence make a major contribution to S_{AB} values, the latter tend to drop dramatically in rRNAs that violate these invariances. Table 18 permits a comparison of S_{AB} values to percent sequence similarities for the mycoplasmas and some of their normal relatives; the S_{AB} values between *Mycoplasma gallisepticum* and other eubacteria in particular are as low as those between normal eubacteria and archaebacteria! Varying the normally conserved positions seems characteristic of all rapidly evolving lineages. As mentioned above, a similar disparity between S_{AB} values and percent sequence similarity holds for the planctomyces (202).

TABLE 18. Percent sequence similarities and S_{AB} values for 16S rRNAs of mycoplasmas and related eubacteria[a]

Species	*M. gallisepticum*	*M. capricolum*	*S. citri*	*Ac. laidlawii*	*C. innocuum*	*B. subtilis*	*H. chlorum*	*Ar. globiformis*	*An. nidulans*	*A. tumefaciens*	*D. desulfuricans*	*E. coli*
M. gallisepticum		27[b]	26	19	22	17	14	15	12	09	07	09
M. capricolum	79.5[c]		47	28	35	30	25	27	20	23	12	17
S. citri	80.5	87.7		25	30	25	23	20	19	22	12	14
Ac. laidlawii	77.4	81.2	79.9		39	28	24	21	17	22	12	15
C. innocuum	76.6	80.8	81.0	82.0		33	27	23	24	22	17	14
B. subtilis	75.0	81.1	79.3	78.6	82.1		41	36	27	26	24	24
H. chlorum	74.5	78.3	78.5	77.9	79.0	83.5		32	24	26	27	28
Ar. globiformis	73.7	78.1	77.2	77.3	79.5	81.0	80.3		24	25	17	24
An. nidulans	73.2	77.0	77.6	76.7	78.4	80.5	81.0	78.6		23	26	24
A. tumefaciens	74.1	77.2	76.5	76.6	77.1	79.1	78.9	77.6	77.7		25	28
D. desulfuricans	72.2	75.0	75.4	75.0	76.8	79.6	80.3	78.7	78.9	80.9		33
E. coli	71.6	75.2	73.9	73.5	75.8	77.5	78.4	77.0	78.0	79.4	79.1	

[a] References 1, 19, 68, 93, 162, 163, 224, 255, 274; sequences for all members of the mycoplasma group except *Mycoplasma capricolum* are unpublished, as is that for *Ar. globiformis*. *M.*, *Mycoplasma*; *S.*, *Spiroplasma*; *Ac.*, *Acholeplasma*; *C.*, *Clostridium*; *B.*, *Bacillus*; *H.*, *Heliobacterium*; *Ar.*, *Arthrobacter*; *An.*, *Anacystis*; *A.*, *Agrobacterium*; *D.*, *Desulfovibrio*; *E.*, *Escherichia*.

[b] Upper right are S_{AB} values (55).

[c] Lower left are percent similarities; all positions not represented in all sequences have been eliminated from consideration.

Dynamics of Variation in rRNA

We assume that (naturally occurring) changes in rRNA sequence are "selectively neutral." The assumption is justified on the grounds that translation has to have been one of the earliest functions established in the cell (for the evolution of all of the cell's proteins depends upon its existence) and, once perfected, there should be no reason, no selective constraints, to change it (outside of rare minor adjustments made to accommodate changes in the cell's basic physical parameters, such as optimum growth temperature, intracellular pH, or ionic concentrations). To support this case, one can invoke the facts that rRNA secondary structures show little variation within any of the primary kingdoms (75, 260), implying corresponding functional constancy, and that within a kingdom components can usually be exchanged among different translation systems without destroying function (2, 14, 156). The mycoplasmas are not exceptional in this regard. The physical parameters of their niches are normal; they do not appear to synthesize unusual types of proteins; their rRNAs have normal secondary structures (93); and some of their *close* relatives, normal eubacteria in all known respects (e.g., *Clostridium innocuum*, *Lactobacillus cateneforme*) share their rRNA sequence idiosyncrasies to some extent (265).

To say that (naturally occurring) changes in rRNA sequence are by and large selectively neutral, does not mean that individual base changes are necessarily so; it is the overall (composite) change that tends to be independent of selection. Changing a single position in a base pair, for example, generally creates a mispair, which would probably be selected against; to change both members of the pair in a way that maintains normal pairing, however, might have negligible selective impact. (More complicated arrangements, tertiary structural and the like, can also be imagined wherein three or more positions would have to be changed to retain proper function.) The individual mutations creating such a composite change, therefore, have to occur simultaneously or at least in fairly rapid succession.

The most important attributes of rRNA sequence variation for our purposes are (i) that recognizable patterns of variation exist, i.e., that the same variations tend to occur in different phylogenetic groups; and (ii) that the frequency with which changes in composition happen can vary widely from position to position within the sequence (262, 265). Regarding the first point, the pattern of variation at a given position is often the same for most or all eubacterial phyla and sometimes even for archaebacteria (266). Regarding the second, the rates at which the most and least variable positions change differ by at least two orders of magnitude (265). Both the frequency and pattern of variation seem to correlate strongly with the overall structure of the rRNA molecule and so would seem to be functionally determined. Base pairs at some positions in some helices change frequently; other pairs in these same helices change rarely if at all; some helices are far more variable in composition than

TABLE 19. Number of positions in which a given sequence shows exception to consensus[a]

Species	Consensus sequence			
	1[b]	2[c]	3[d]	4[e]
Mycoplasmas				
1. *M. gallisepticum*	66	50	44	99
2. *M. capricolum*	41	13	24	64
3. *S. citri*	46	18	23	60
4. *Ac. laidlawii*	60	20	33	65
5. *C. innocuum*	23	13	12	39
Normal eubacteria				
6. *B. subtilis*	6	9	2	22
7. *H. chlorum*	13	18	3	24
8. *Ar. globiformis*	27	31	10	41
9. *An. nidulans*	28			
10. *A. tumefaciens*	23			
11. *D. desulfuricans*	17			
12. *E. coli*	33			

[a] References 19, 68, 93, 163, 224, 255, 274; sequences for all members of the mycoplasma group except *M. capricolum* are unpublished, as is that for *Ar. globiformis*. For genera, see Table 18.

[b] Consensus allowing one exception only at a given position; the values for the species in the mycoplasma group, i.e., no. 1 to 5, are calculated individually from alignments containing that sequence alone (from the mycoplasma group) and no. 6 to 12; the values shown for no. 6 to 12 are therefore averages over the five resulting consensus sequences.

[c] Consensus allowing one exception; alignment contains only gram-positive species, i.e., no. 1 to 8.

[d] Consensus allowing no exceptions; alignment contains species 9 to 12 plus *Leptonema illini* (unpublished data), *Chlorobium vibrioforme* (Weisburg, Ph.D. thesis), *Thermomicrobium roseum* (162) and *Thermotoga maritima* (1); i.e., it contains no gram-positive species.

[e] Same as footnote *d* except that one exception is allowed in generating the consensus.

others; and the composition of loops tends to be more highly conserved than that of the underlying double-stranded stalks (75, 260, 262, 265).

Why mycoplasma rRNAs are so unusual. Not only do mycoplasmas tend to vary the otherwise conserved positions (i.e., introduce the rare composite changes) in rRNAs more readily than do their normal counterparts, but they even differ significantly from one another in this respect. The tendency to vary conserved positions is far more accentuated in *Mycoplasma gallisepticum*, for example, than in other mycoplasmas (262, 265) (see also Table 20). Moreover, one line of mycoplasmas will make changes in the rRNA sequence that others do not, as though a broad range of possibilities existed from which to choose (262, 265). If their ribosomes are structurally and functionally normal, then functional constraints on the ribosome cannot bring about the evolution of these rRNA idiosyncrasies. One has to consider, therefore, that these idiosyncrasies are not connected to ribosome function/evolution per se; rather, they reflect some general peculiarity of the evolutionary process in mycoplasmas.

Consider the following argument: if changing one of the more conserved positions in an rRNA has to involve (nearly) simultaneous changes elsewhere in the same (or another) molecule, then the occurrence of the overall, composite change is a higher-order function of the organism's (lineage's) mutation rate. For low enough rates such changes occur with a negligible frequency relative to simple (first-order) nucleotide changes, but as the mutation rate increases, the composite changes will increase in frequency relative to the simple ones (265). This means that lineages with low mutation rates have associated with them fields of rRNA variants (from which they derive their evolutionary variability) that are relatively restricted, whereas lines with higher mutation rates draw upon much richer (more varied) fields of variants. Other factors being equal, rRNAs would evolve far more variety in the latter case than in the former.

Although mycoplasma mutation rates have not actually been measured, there are good reasons to suspect them to be abnormally high. In any line of descent mutation rate must be optimized. It cannot be so high that deleterious mutations are created in a significant fraction of an organism's progeny, yet it cannot be so low that the lineage is unable to adapt to fluctuations in its environment or is otherwise unable to compete effectively (265). An upper bound to mutation rate is, therefore, set by an organism's functional genome size. The larger that genome, the harder to replicate it without introducing errors. Consequently, an organism with a small genome could be as stable evolutionarily as organisms with larger genomes even though it had a higher mutation rate (per base pair). Mycoplasma genomes are four to eight times smaller than the eubacterial norm (137). Although mycoplasmas arose from ancestors (clostridia) having normal genomes, and so presumably normal mutation rates, the constraints keeping mutation rates low cease to exist once genome size decreases. One would also expect some mycoplasmas to have elevated mutation rates because they are known to be deficient in certain DNA repair capacities (61) and to lack a DNA polymerase 3′→5′ exonuclease activity (150).

Summary. In brief, my argument is this: changes in rRNA sequence are for the most part selectively neutral. However, many of these changes are composite and so would appear to involve nearly simultaneous, coordinate alterations in two or more positions in the molecule. Composite change of this type is a higher-order function of a lineage's mutation rate. Those lines having elevated mutation rates would be expected not only to show the normal types of rRNA variants at higher than normal levels, but also to spawn variants that normally occur at inappreciable levels. Because their genomes are small, mycoplasmas can develop elevated mutation rates, generating rRNA variants not usually associated with normal lineages, which gives a unique richness (variability) to the evolution of their rRNAs.

Both Tempo and Mode of Bacterial Evolution Are Reflected in rRNA

A useful, if idealized, model for the rRNA chronometer is a measuring device that comprises a series of counters. The basic, primary counter records the number of (certain) events occurring; it simply measures a distance, a rate × time. The others are threshold counters (differing from one another in having progressively higher thresholds). Each one registers nothing until the rate at which the events occur reaches its particular threshold value, after which it too measures rate × time. Such a chronometer can measure more than long-term average rates, more than simple distances. It can detect changes in rate and peak rates. rRNA is not the simple uniform-rate chronometer our analyses generally assume. It behaves as a "compound" chronometer in the above sense. As such it should be able to detect whether (but not necessarily when) episodes of rapid evolution have occurred in a lineage's history, which opens the tempo-mode problem to study on the molecular level. Such a rate-sensitive chronometer could also be used to produce phylogenetic trees in which the root is internally delimited (265).

The unusual changes encountered in the rRNAs of mycoplasmas and other rapidly evolving bacterial lineages actually measure the mode of evolution. In a sense they are the mode, for at the molecular level tempo and mode come together; they are different facets of the same process. The mutational changes that are summed to indicate a tempo (evolutionary distance) include the changes that define the quality of the field of variants. The idiosyncratic selectively neutral variants of mycoplasma rRNAs are obviously representative of all variants in the mycoplasma phenotype. Expanding the field of variants, as mycoplasmas appear to do, makes it statistically unavoidable that their evolution be both more rapid than normal and highly atypical.

Macroevolution in mycoplasmas is chronic. As long as mycoplasma genomes remain small, these organisms would seem to be in a chronic state of rapid evolution. Classical macroevolution is episodic, however. Thus, a question remains as to whether bacteria can exhibit episodic rapid evolution, whether macroevolution in its classical form occurs in the microbial world.

The rRNA chronometer in principle should reveal the episodic form of rapid evolution in the same way it does the chronic form, although the former would be harder to detect in that it would leave less of a trace. That rRNA sequence signatures can be constructed for the various bacterial groups implies that unusual sequence changes become fixed during the formation of major taxa. For a group that contains many characterized representatives, the cumulative evolutionary distance within it (i.e., the sum of all branch lengths on the corresponding phylogenetic tree) is very large compared to the evolutionary distance that separates that group's ancestor from the ancestor of some other nearby group. Yet (for these well-characterized groups) the unusual types of change that become fixed during their formation

tend not to occur during the group's evolutionary ramifications. No statistical justification of this point will be attempted at this time, for more rRNA sequence data than now exist are required to make a strong case. However, the indications are there; major bacterial groups come into being through episodic rapid evolution.

Conditions for true macroevolution among bacteria. Episodic rapid evolution among bacteria would (according to the above reasoning) require that conditions exist under which bacterial mutation rates can increase but subsequently return to normal. Such conditions may well be those invoked by the classical evolutionist to explain macroevolutionary change (192). (However, the classical conditions now become necessary but insufficient to effect episodic rapid evolution in bacteria.) Environments are of two general types: those to which organisms can become well adapted, and those to which they cannot. Stable or cyclically varying environments can be of the first type. Chaotically varying and "extreme" environments in general are of the second. In the first case an organism's phenotype can (and does) become "fine tuned" to the environment. Organism and niche come into some sort of close and detailed correspondence: nuances of phenotype have selective meaning; additional levels of organization (control) are added; efficiency of function increases; new refinements, details, evolve. Such are the general evolutionary considerations for normal environments. Contrast this to evolution in an unpredictably fluctuating or otherwise extreme environment. Environments of this type stress the organism's physiological responses to the limit. Existence here is a matter of survival under any condition, not of survival of the fittest. In other words, fine-tuning, efficiency, etc., have little to do with evolution in this context. Under these conditions negative selection would, in some senses and some areas, be relaxed. Many genes concerned with fine-tuning would be of little significance. Only the most basic genes, those that make for integrity of the organism and continuity of the lineage, would really count; and the precision or efficiency (i.e., fine-tuning) of their function may not be as selectively significant a concern as it normally is. Under conditions such as these, when many functions become superfluous, the effective genome size is reduced, and some selection is relaxed, a lineage might then be able to sustain an elevated mutation rate. If so, the resultant expansion of the field of variants would mean that unusual phenotypic features necessarily arise in the lineage: some of the rare variations might even be essential to the line's survival in the untoward environment (and so an increased mutation rate would have positive selective value). Were such a line subsequently to adapt to a more stable, compatible environment (or even somehow stabilize in the formerly extreme one), its mutation rate would necessarily return to normal, but the organism would bear the scars of its tumultuous history; its phenotype would be drastically changed and highly unique.

Basic Principles of Bacterial Evolution

The above conceptualization of the tempo-mode relationship makes bacterial evolution appear straightforward and understandable. Normal lineages, those having normal mutation rates, do not drastically change their ancestral phenotype. If the environment in which a phenotype arises (first stabilizes) persists, that phenotype will persist, fundamentally unchanged. This is not to say that the original phenotype cannot change to fit a novel environment (without increase in mutation rate and so on), but it is to say that the kinds of new environments to which it is capable of adapting in this way do not radically alter the ancestral phenotype. On the other hand, all drastic (broad-ranging) changes in ancestral phenotype necessarily result from increased mutation rates, which tend to occur under unusual, drastic environmental conditions, when selection is relaxed in ways that allow the mutation rate to rise.

In a general sense, then, the course of bacterial evolution is relatively simple to chart by means of a macromolecular chronometer, a purely genotypic measure. Lineages represented by the shorter branches on a phylogenetic tree (i.e., the ones least distant from the tree's root) are slowly evolving and retain proportionately more of the ancestral phenotype. Lineages represented by the longer branches, on the other hand, are rapidly evolving and necessarily retain far fewer ancestral characteristics. Moreover, a molecular chronometer will show structural idiosyncrasies in the latter case, the "scars" of the lineage's bout of rapid evolution.

Implications for Bacterial Taxonomy

If future findings support the above conclusions and speculations, then it should be possible to construct a bacterial taxonomy based upon naturally defined categories. Groups that arise through macroevolution are self-defining; they are recognizably unique, distinct both phenotypically and genotypically (in terms of molecular chronometers). Many higher bacterial taxa, i.e., phyla and their major subdivisions, show this characteristic.

In addition to naturally demarcated categories, two other requirements for a natural taxonomy are: (i) a means of determining relationships among the categories, which, of course, is given by the topology of a phylogenetic tree; and (ii) a means of naturally defining taxonomic rank, one that is not completely dependent upon tree topology. There is no reason a priori that a taxon of higher rank, a class, for example, cannot be included in one of lower rank, e.g., a family. Indeed, the mycoplasmas may be a case in point. The question is whether the microbiologist can accept such natural "inversions" of rank as taxonomically valid or whether he will insist upon an arbitrarily defined taxonomy that does not contain them and so appears orderly to him.

Macroevolutionary episodes can probably be classified by the degree of their severity. If the rRNA chronometer in essence consists of a series of counters having progressively higher thresholds (i.e., the model used above), then the highest-threshold counter activated by a macroevolutionary episode defines the severity of the episode, which is a de facto definition of taxonomic rank. Without further constraints, however, we are again left with the possibility of a jumbled taxonomic hierarchy, for the timing and severity of macroevolutionary episodes would seem to be unrelated in unrelated lineages.

It is obvious that metazoan taxonomy exhibits more order than is inherent in the natural system just described. The animal phyla seem all to have arisen at about the same time, somewhat less than 1 billion years ago; the eucaryotic kingdoms arose similarly at an earlier stage (196). Evolution tends to follow two rules: (i) the higher its taxonomic rank, the further back in time the taxon arises; and (ii) major taxa of the same rank tend to arise in the same era, giving rise to evolutionary radiations. All this may seem self-evident to some, but it is not, particularly in the bacterial world where potential ancestral phenotypes seem to persist for all taxa up to the kingdom level. The explanation for this unaccountable time ordering of the actual taxonomic hierarchy may lie

outside of biology per se, in the evolutionary history of the planet. Geologists are coming to the conclusion that the history of earth and the other planets, moons, etc., is written in terms of catastrophes, relatively sudden, chaotic global changes (3). The frequency and severity of such changes increase as one looks backward in time. (Unfortunately, the record of their happening erodes with time, particularly on earth.) The intense bombardment by meteors and similar objects, which changed the face of the moon (and presumably affected earth as well) prior to 3.9 billion years ago, is one example (85). Another example, more germane to biology, might be the relatively sudden (on an evolutionary time scale, that is) rise in atmospheric oxygen concentration. Such precipitate shifts in global parameters would be capable of triggering synchronous macroevolutionary episodes in various lineages (3). Because the more severe "catastrophes" tend to be the earlier ones, the taxa of higher rank would tend to form earlier than those of lower rank.

The genealogical history of bacteria beginning to be revealed by molecular chronometers will provide ample data to develop or discard such a view of evolutionary relationships and their taxonomic implications.

EVOLUTION OF THE TWO PROCARYOTIC PHENOTYPES

The enormous value of bacterial phylogeny as a classification system, a predictive and organizational framework, is easier to appreciate than its value as an historical account and source of evolutionary insights. This is because our evolutionary perspective has to this point in time been focused narrowly on metazoa and their fossils. This point of view has inevitably overemphasized morphology, concentrating on subtle but superficial differences among complex forms. At the same time it has underemphasized the "metabolic" aspect of evolution; it embodies little feeling for biochemistry and energy flow. Since it has been confined to a relatively recent (and so, in the grand scope, uninteresting) period in earth history, well after formation of the oxygen atmosphere (180, 228), our present view incorporates in a minor way only the dynamism of the evolving earth and the close relationship between its physical and biological evolution. Overall, this view is a static one, in which evolution is not a process but rather "a procession of forms," as Whitehead put it (237).

In contrast, microbial evolution is essentially metabolic, fundamentally biochemical. It spans the bulk of our planet's history and is intimately tied thereto. As the base of the food chains, microbial metabolic patterns bear a straightforward and (ultimately) understandable relationship to the planet's geochemistry. Thus, bacterial evolution is no simple extrapolation of metazoan evolution—more of the same and, lacking fossils, harder to study. Microbial evolution is a different story, told in a different way, covering a different (more extensive) period of earth history, a story that is simpler, more readily interpretable, and more informative of the planet's physical course.

Archaebacterial Evolution

Ancestral phenotype. It would appear that the ancestral archaebacterium was an extremely thermophilic anaerobe that probably derived its energy from the reduction of sulfur. Two lines of evidence support such a conclusion. The first is the widespread distribution of the extreme thermophilic phenotype among the archaebacteria. Of the three basic archaebacterial phenotypes, it is, as we have seen, the only one that occurs on both major branches of the archaebacterial tree. *Thermococcus celer*, on the methanogen branch, and *Pyrodictium occultum*, on the extreme thermophile branch, are both typical sulfur-metabolizing thermophiles, thriving anaerobically in hot spring environments. (*Pyrodictium* holds the current record for highest optimum growth temperature of any organism, 105°C [212], while *Thermococcus* is a common inhabitant of marine hot springs [280].) The second is that the extreme thermophile phenotype is the only one to meet the tempo-mode criterion for being ancestral; the evolutionary distance between *Thermococcus* and *Pyrodictium* is remarkably short, about 18%, appreciably shorter than the shortest distances, 24 to 25%, that separate any methanogen (relatives of *T. celer*) from *P. occultum* or its relatives, as can be seen in Table 17. From these facts, and the fact that the two extreme thermophiles are about equally distant from various eubacterial or eucaryotic outgroup sequences (263), it follows that both lineages are slowly evolving and, therefore, have retained more common ancestral characteristics than have the other archaebacterial phenotypes.

Two genera, *Sulfolobus* and *Thermoplasma*, are atypical sulfur-dependent thermophiles in having evolved the capacity to utilize oxygen (215), almost certainly a derived characteristic. It is interesting, therefore, that both represent relatively rapidly evolving lineages. The lineage of *Sulfolobus* is the most rapidly evolving on its branch of the archaebacterial tree, while that of *Thermoplasma* is perhaps the most rapidly evolving of all archaebacterial lineages (263).

Evolution of methanogenic and halophilic phenotypes. Since these two phenotypes are significantly further from the extreme thermophile cluster (by rRNA measure) than is their relative *Thermococcus*, the methanogenic and halophilic lineages would seem to have undergone macroevolution. Presumably such an episode was associated with the transition from an ancestral thermophilic sulfur-metabolizing phenotype to a methanogenic one. Given such an evolutionary progression, one has the interesting question of how thermophilic sulfur metabolism can change into methanogenesis and what global conditions might have favored, i.e., brought about such a transition. (Very recently K. O. Stetter [personal communication] has isolated what may be a missing link in such a transition. The organism is a novel archaebacterial phenotype; it grows anaerobically and reduces sulfate. It contain several of the cofactors characteristic of the methanogens, but lacks the all-important factor coenzyme M, which is involved in the terminal step of methane production [238]. Nevertheless, the organism does produce methane in minute amounts, as does the eubacterial sulfate reducer *Desulfovibrio desulfuricans* [173]. A preliminary and unpublished partial 16S rRNA sequence shows its lineage to arise from the methanogen branch of the archaebacteria between the *Thermococcus* and *Methanococcus* lineages in Fig. 13.)

Within the methanogens per se a second round of rapid evolution seems to have occurred, involving the branch that leads to the *Methanomicrobiales*. Note in Table 17 that sequence distances for members of the *Methanomicrobiales* are greater than corresponding distances for both of the other methanogen phyla. This can be seen in convincing detail in the analysis of Table 20, which gives sequence distances of various archaebacterial species from various consensus sequences.

Although the most spectacular change during this second

TABLE 20. Sequence distances (in percentages) of archaebacterial rRNAs from various consensus sequences[a]

Sequence from:	Consensus sequence			
	1[b]	2[c]	3[d]	4[e]
Mc. vannielii	—[f]	—	10.0	16.0
M. formicicum	—	—	9.6	15.4
Ms. hungatei	—	10.0	13.9	19.3
Mr. sp. strain WH-1	6.5	10.5	13.0	18.6
Extreme halophiles[g]	8.9	11.6	15.6	20.7
T. acidophilum	13.1	13.0	15.7	21.6
Tc. celer	5.1	—	—	10.4
S. solfataricus	12.3	12.1	5.5	6.7
Tp. tenax	11.5	10.4	—	—
D. mobilis	8.9	8.3	—	—
P. occultum	9.3	8.3	1.1	—

[a] Only those positions represented in all positions in 16S rRNA alignment are used in calculation. References 72, 91, 96, 124–126, 158, and 275. Sequences for *Methanosarcina* sp. strain WH-1, *T. acidophilum*, *Tc. celer*, *D. mobilis*, and *P. occultum* are unpublished. *Mc.*, *Methanococcus*; *M. Methanobacterium*; *Ms.*, *Methanospirillum*; *Mr.*, *Methanosarcina*; *T.*, *Thermoplasma*; *Tc.*, *Thermococcus*; *S.*, *Sulfolobus*; *Tp.*, *Thermoproteus*; *D.*, *Desulfurococcus*; *P.*, *Pyrodictium*.

[b] Consensus based upon *Mc. vannielii*, *M. formicicum*, and *Ms. hungatei*.

[c] Consensus based upon *Mc. vannielii*, *M. formicicum*, and *Tc. celer*.

[d] Consensus based upon *Tc. celer*, *Tp. tenax*, and *D. mobilis*.

[e] Consensus based upon *Tp. tenax*, *D. mobilis*, and *P. occultum*.

[f] —, Sequence used to generate consensus.

[g] Values are averages of the three published halophile sequences and are within 7% of one another.

saltation was the conversion of an anaerobic methanogen into an aerobic extreme halophile, profound changes also affected the methanogens themselves in this particular lineage. Some of the *Methanomicrobiales* became halophilic, some even alkaliphilic (13; Mathrani et al., Abstr. Annu. Meet. Am. Soc. Microbiol. 1985). *Methanosarcina* and its relatives learned to produce methane from acetate or methyl amines; they are, as noted, the only methanogens possessing cytochrome *b* or *c* or both (112). The extreme halophiles, which require small amounts of oxygen to synthesize their carotenoids (79), may represent an evolutionary response to the onset of aerobic conditions in the hydrosphere, about 1.5 billion years ago (228).

Archaebacterial ribosomes and their evolutionary implications. The archaebacteria are unique among the three urkingdoms in that variation in their ribosome type occurs. Cammarano and co-workers (21) have measured the protein content of ribosomal subunits by buoyant density centrifugation and find that the molecular weight of protein associated with the small subunit in the extreme thermophiles is 0.64×10^6 to 0.66×10^6 but in the extreme halophiles and two of the methanogen phyla it is only half this, i.e., 0.31×10^6 to 0.32×10^6. In the remaining methanogen group, the *Methanococcales*, protein contents of the small subunit have an intermediate value, 0.52×10^6 daltons. A similar situation obtains for the large subunit; there the protein contents in the extreme thermophiles and *Methanococcus* are in the range of 0.77×10^6 to 0.97×10^6 daltons, but in the remaining methanogens and the extreme halophiles they are much less, i.e., 0.51×10^6 to 0.57×10^6 daltons (21). (The ribosomes of *Thermoplasma* are also high in protein; the small and large subunits contain 0.61×10^6 and 0.78×10^6 daltons of protein, respectively [21].) Since rRNA sizes are very nearly the same for all archaebacteria, this approximate twofold variation in protein content must make some change in the size of the ribosomal subunits, and electron micrographs bear this out. The large ribosomal subunit from the extreme thermophiles is larger than those seen in methanogens; it has several protrusions that its methanogenic (and halophilic) counterparts do not (81). However, some of the large ribosomal subunits from *Methanococcus vannielii* have these protrusions as well, while the remainder are typical of the other methanogens (and extreme halophiles) (218, 219).

The molecular basis for these large differences in ribosomal protein content is not understood. There is no indication that they are connected to structural differences in the rRNAs (Woese, unpublished analysis). The excess proteins are probably not, therefore, attached directly to rRNA. It also seems unlikely that the drastic disparities in protein content reflect any significant difference in ribosome function in the two classes.

When these differences in ribosome type were first discovered (as shape differences in electron micrographs), they were interpreted to mean that the archaebacteria were not a valid taxon; the extreme thermophiles and the methanogens (and their relatives) constituted separate urkingdoms, the former having "a close relationship to eukaryotes" (115). (For similar reasons, the extreme halophiles were subsequently extracted from the methanogen branch to constitute another new kingdom, the "photocytes," which ostensibly was specifically related to eubacteria [114].) Now that ribosome morphologies have been more thoroughly investigated, it is apparent that ribosome shapes (protein contents) constitute more a spectrum of types than two clear-cut classes (114, 115, 218, 219). In any case, were separate kingdoms to be defined along these lines, one group of methanogens, the *Methanococcales*, would end up in a different kingdom (218, 219) than the others—which is absurd! As taxonomists well know and have repeatedly stated, a small number of (ill-defined) characters is an unreliable basis upon which to define taxa.

Summary. Archaebacterial evolution can be simply understood in terms of an aboriginal anaerobic thermophilic sulfur-metabolizing phenotype that remained pure in one of the urkingdom's two main lineages, but gave rise in the other, through several macroevolution episodes, initially to methanogenic metabolism and then to an altered, more versatile (acetate- or methyl amine-utilizing) form of methanogenesis, to halophilic methanogens, and ultimately to the aerobic (nonmethanogenic) extreme halophiles.

Eubacterial Evolution

Eubacterial history seems less straight-forward than archaebacterial history. The most prominent phenotypic characteristics of the eubacterial tree are its metabolic diversity and the widespread distribution of anaerobic, photosynthetic, and thermophilic phenotypes. Aerobic and anaerobic groupings stand in sharp contrast to one another. Whereas the various anaerobic phenotypes more often than not form phylogenetically deep, extensive groupings, such is not the case for their aerobic counterparts (56, 206). A good example is the comparison between the clostridia (a remarkably deep anaerobic phylogenetic unit) and *Bacillus* (a phylogenetically much shallower collection of aerobes) (56). Thus, although aerobic phenotypes evolved a number of times, their lack of phylogenetic depth suggests that they are all relatively recent in origin.

Photosynthesis. In slightly different versions, photosynthesis appears in at least half of the eubacterial phyla. The purple, blue-green, gram-positive (*Heliobacterium*), and green sulfur lineages stem from the same general area on one of the two major branches of the eubacterial tree (Fig. 11).

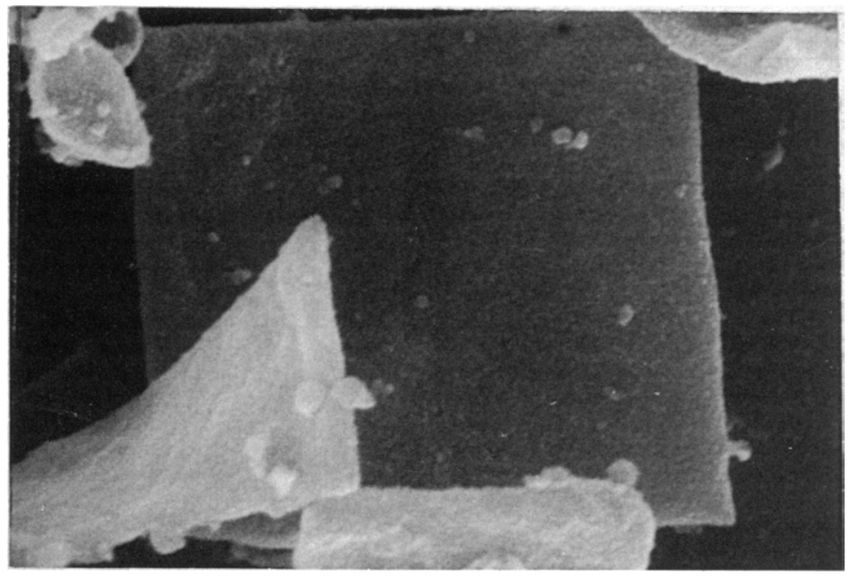

FIG. 16. Electron micrograph of a square extreme halophile (216); figure courtesy of W. Stoeckenius.

The remaining photosynthetic type, the green non-sulfur bacteria, represents a significantly deeper branching in the tree (162). Within some phyla, e.g., the purple bacteria, the photosynthetic phenotype is intimately intermixed with nonphotosynthetic phenotypes (162, 267–269). Given the complexity of the photosynthetic apparatus, it seems unlikely that the process has evolved more than once within the eubacteria, and so its origin is deep in the eubacterial tree, possibly at the stage of the common eubacterial ancestor.

Thermophilia. It can be even more compellingly argued that thermophilia is an ancestral eubacterial characteristic. All deeper branchings in the eubacterial tree involve either predominantly or exclusively thermophilic groups (1): the *Thermotoga* lineage forms one side of the deepest known branching in the eubacterial line. (The organism so far has only one known [unnamed] relative, strain H21 of K. O. Stetter, again a thermophile [Stetter and Woese, unpublished data].) As mentioned above, the unusual thermophile *Thermodesulfotobacterium commune* (119) represents another very deep branching. A third is formed by *Thermomicrobium* and its relatives (the green non-sulfur bacteria), again a predominantly thermophilic group (94, 162, 170). All of these thermophilic lineages are also relatively slowly evolving, especially *Thermotoga*, and should, by the tempo-mode criterion, be primitive in type. The only known mesophile to branch deeply in the eubacterial tree is *Herpetosiphon* (a relative of *Chloroflexus* and *Thermomicrobium*). However, this organism is the product of a rapid evolution (162) and so should not resemble the ancestral phenotype to the extent that its thermophilic relatives do. Thermophiles are also found in the "mesophilic section" of the eubacterial tree, e.g., *Bacillus stearothermophilus* and *Thermus aquaticus*. In almost all of these cases, too, the thermophilic lineages are at least as slowly evolving as their specific mesophilic relatives. The evidence clearly points to a thermophilic origin of the eubacteria.

One could similarly, but less convincingly, argue an autotrophic ancestry for the eubacteria, for autotrophy is also widely distributed among eubacteria.

Conclusion. The case for a photosynthetic, thermophilic, or autotrophic eubacterial ancestor can never be proven in the strictest sense. But that is not what is important. Because we can now approach microbial phylogeny experimentally and can construct a microbial tree, we are in a position to weed out former incorrect notions and develop new and more detailed ones more soundly based upon experimental evidence. We are on the threshold of greatly expanding our understanding of bacteria and their relationships to one another—and proceeding from there to a reconstruction of the history of life on this planet.

Differences between Archaebacteria and Eubacteria

The persistent influence of the procaryote-eucaryote dogma prevents many biologists from appreciating the importance of treating archaebacteria as distinct from eubacteria: "They may be very different, but they are still procaryotes" is an attitude one still hears repeatedly. Any such feeling inhibits understanding of the eubacterial-archaebacterial relationship. Even the fact that the names of the two groups bear the common suffix "-bacteria" acts to do so. Therefore, it is important at this juncture to stress the differences between the two procaryotic types.

Cytological and physiological differences. Eubacteria and archaebacteria are perceived as cytologically similar mainly because both are simpler than and do not resemble the more complexly structured eucaryotic cell. Within their "similarity," however, lie significant differences in cell architecture and metabolic patterns and remarkable differences in evolutionary behavior. Cells of many extreme halophiles are very thin, flat, and straight sided, with precisely square corners (216, 230); Fig. 16 shows a remarkable example. Flat pseudo-geometric shapes have also been noted occasionally among the methanogens (242). Since flat angular geometries are foreign to eubacteria, basic differences between archaebacteria and eubacteria in cell architecture are implied.

Methanogenesis involves a variety of coenzymes unique to that process, and at least one unusual coenzyme is somehow associated with sulfur metabolism in the extreme thermophiles (Fig. 12). It is hard to believe that such uniqueness is entirely confined to specialized biochemistries

such as methanogenesis; some of it must carry over into the cell's general metabolism. In other words, the unique coenzymes imply significant differences between archaebacterial and eubacterial intermediary metabolisms. The area of archaebacterial intermediary metabolism is ripe for further study (57).

As mentioned above, archaebacterial membranes may be somewhat different structurally from their eubacterial counterparts, for most of them contain a significant level of diglycerol tetraethers, something unprecedented among eubacteria and eucaryotes (117, 118).

Genome organization and control are significantly different in the two classes of procaryotes; one has the feeling that, if more were known, their difference would seem even more profound. Some archaebacterial genomes contain families of repeat sequence DNA, not a eubacterial characteristic (39, 182). Although the sequences of many eubacterial genes are known, introns have not been reported. Yet among the few sequenced archaebacterial genes, a number contain introns: the majority of sequenced tRNA genes from *Sulfolobus solfataricus* contain introns (100; B. P. Kaine, manuscript in preparation), as does one tRNA gene from an extreme halophile (30) and an archaebacterial 23S rRNA gene (107). Nothing so far indicates that archaebacteria use the well-known eubacterial types of gene regulation mechanisms. One would think that, if they did, some evidence for their doing so would by now exist.

Niches. The two types of procaryotes tend to inhabit different types of environments. Archaebacteria prefer high-temperature niches. One side of the archaebacterial tree appears to consist only of thermophilic species. Methanogens that grow at high temperatures are also fairly common (214, 278). No major well-characterized eubacterial group is known to be exclusively thermophilic, although the *Thermotoga* side of the eubacterial tree, represented so far by only two species, remains potentially so (1, 89). Most eubacterial groups are not even predominantly thermophilic. Eubacteria, on the other hand, adapt more readily than archaebacteria to the myriad low-temperature niches, wherein they predominate.

While eubacteria readily adapt to aerobic conditions, archaebacteria seem to have difficulty in doing so. Among archaebacteria, even facultative aerobes are relatively uncommon; there are no obligate aerobes (79). The evolution of oxygen utilization in archaebacteria appears to be associated with episodes of rapid evolution (see above), which is not the case with (at least some) eubacteria; *Bacillus*, for example, although aerobic, represents one of the most slowly evolving of eubacterial lines by sequence distance measures (56, 163, 265).

Metabolic versatility. The evolutionary differences between the two classes of procaryotes manifest themselves in several ways. One is an ill-defined evolutionary quality that, for want of a better term, will be called "metabolic versatility." Eubacteria are remarkable for the variety of their metabolisms, both the overall variety and the extent to which variation occurs within some of the phyla and subdivisions. By comparison, archaebacteria are metabolically monotonous. Little variation exists within the three basic archaebacterial phenotypes; the utilization of acetate and methyl amines by one group of methanogens (discussed above) is an example of the sort that does occur. While photosynthesis is a prevalent, perhaps ancestral, theme among eubacteria, no archaebacteria are completely photosynthetic; the extreme halophiles, however, have acquired the capacity to use light to run various molecular pumps (122, 217). It is remarkable how little metabolic convergence the two procaryotic groups have shown over their several billion years of coexistence.

Molecular plasticity. Paradoxically, the archaebacteria show more variety than eubacteria do in another evolutionary parameter, a quality we will call "molecular plasticity," i.e., the variations in molecular design of a given function within a group. That the archaebacteria are exceptional in this respect has been apparent from the outset (250). Examples of the group's molecular plasticity (mentioned above) are seen in 5S rRNA secondary structure (53, 133, 208), DNA-dependent RNA polymerase subunit patterns (187, 282), ribosome protein contents (21), the extent of post-transcription modification of bases in rRNAs and tRNAs (70, 71, 73, 259), cell wall structure (101, 109, 111), types of coenzymes (238), and antibiotic sensitivity (22). In all of these cases archaebacteria present at least two distinctly different types, more often a spectrum of types, while eubacteria present a relatively monotonous picture of uniformity (250).

Rates of evolution. The two classes of procaryotes seem to be evolving at generally different rates. Evidence suggesting this is seen in the phylogenetic breadth of the two groups, i.e., the sequence distance separating the most slowly evolving representatives of the deepest branches in each tree. The phylogenetic breadth of the archaebacteria (i.e., the distance between the extremely thermophilic representatives on its two main branches) is under 20%; see Table 17. For eubacteria the comparable distance (between *Thermotoga* and *Bacillus* [1]) is slightly under 30%. Since the root of the universal tree is not yet known, it is possible that all this means is that the archaebacteria are not as old a group as are the eubacteria (attributing the smaller sequence distances in this case to shorter time of evolution, rather than a slower rate). Such an assumption would require the root of the universal tree to be place relatively high on the eubacterial branch, which then makes the archaebacteria and eucaryotes quite specific relatives of one another and the eucaryotes a very rapidly evolving line of descent, assumptions I find intuitively unappealing.

Summary. This discussion of eubacterial-archaebacterial differences amounts briefly to this: although ostensibly similar cytologically, the two groups of procaryotes are significantly different in cellular makeup and in their modes of evolution. They tend to inhabit basically different niches. They appear to evolve at different rates. They show differences in two evolutionary parameters, metabolic versatility and molecular plasticity. And, over several billion years of coexistence they have shown little or no tendency to converge in phenotype. It is as though the two have evolved in different worlds.

Why Are Archaebacteria and Eubacteria So Different?

How are we to understand these substantial differences between the two types of procaryotes? Are they a matter of different environments during the early stages of evolution in the two lines? Are basic organizational differences in the two types of cells being reflected in different evolutionary proclivities? We are in no position to answer such questions now. It is not simply that facts are lacking; key concepts seem to be missing too. The best we can do now is to ask questions that we hope will lead us to experiments that provide the required insights.

One key question seems to be whether archaebacteria are significantly older than eubacteria (or vice versa). If

archaebacteria already existed before the (most recent) common ancestor of all eubacteria arose, then the basic archaebacterial phenotype would probably have evolved to suit a global environment very different from that in which eubacteria later arose, making the two cell types basically and perhaps irrevocably dissimilar.

A second key question is whether the archaebacterial ancestor was more primitive than its eubacterial counterpart, and if so, in what ways. A more rudimentary, less highly integrated ancestor would conceivably have a broader spectrum of potential phenotypes into which it might evolve than would a more highly integrated, more constrained and "advanced" ancestor. An archaebacterial ancestor of this type would explain the greater molecular plasticity of the group.

Finally, one wants to know the relationship between the common ancestors of both groups (and their relationship to the eucaryotic ancestor). What sort of cell, entity, or system gave rise to the three urkingdoms? It is unlikely that we will ever know the full answer here; however, it is something about which we will soon be able to infer a great deal more than we can now.

Primitiveness of Archaebacteria

In the past, discussion of what was and was not a primitive characteristic was more or less fruitless. (What appeared primitive to some often turned out merely to be degenernate, e.g., the mycoplasmas.) This is no longer the case, for molecular chronometers provide a nearly certain definition of "primitive." In a group of homologous molecular functions the one whose sequence is closest to that of the common ancestral version is necessarily the most primitive. Provided the position of its root can be fixed, a sequence tree then decides the issue.

Current evidence tentatively suggests that archaebacteria are probably the more primitive of the two procaryotes, in three senses of the word: (i) the group as a whole is older than the eubacteria; (ii) their common ancestor was a more primitive type of entity than the eubacterial common ancestor; and (iii) since they evolve at a slower rate than do the eubacteria, archaebacteria today remain more primitive (more ancestral in type) than eubacteria.

The most convincing evidence comes, of course, from the rRNA chronometer. Since the archaebacterial 16S rRNA is closer in sequence to both its eubacterial and eucaryotic counterparts than these two are to one another, the archaebacterial version of the molecule must be closer to the common ancestral version than is one or both of the other versions (72). Placing a root on the (unrooted) universal tree anywhere within a zone that includes the archaebacterial branch and a fair segment on both of the other main branches (which can be visualized on Fig. 4) would make the archaebacteria more primitive than both of the others.

A similar, but weaker, conclusion can be drawn from the relationship among the DNA-dependent RNA polymerases in the three urkingdoms. By serological cross-reactivity, the archaebacterial polymerases appear closer to their eubacterial and eucaryotic counterparts than these two types are to one another (90, 187). (Sequencing studies now in progress should soon permit more definite conclusions.) This tendency for archaebacteria to be closer to both of the other groups than they are to one another also carries over (in a qualitative sense) to the general phenotype (250). In a few years, when sequence evidence is available for a significant number of different molecular functions, it may be possible to say in a quantitative way that the general archaebacterial phenotype is more primitive than at least one of the others. If so, the question of archaebacterial primitiveness is half answered.

Knowing the root of the universal tree (the ancestral point) would automatically determine which of the three phenotypes is the most primitive. Conventional wisdom holds that the root of the universal tree cannot be determined, because no outgroup exists by which to position it. However, the root of the universal tree can be determined, in principle if not in practice. What is required is a gene that has duplicated in the common ancestor state (as pointed out by M. Dayhoff long ago). If two (functionally distinct) versions of such a gene fulfilled certain technical requirements, they could then be used in effect to determine relative rates of evolution within each lineage, thereby fixing the tree's root. A practical system for doing this does not yet exist.

When Did Procaryotes Evolve?

The general nature of the archaebacterial phenotype strongly implies the nature of the environment in which it arose. Given the widespread distribution of thermophilic species (and their universal occurrence on the extreme thermophile branch), it seems impossible that these organisms arose from a mesophilic ancestor. The ancestral archaebacterium was a thermophile, probably growing at temperatures near the present boiling point of water (see above discussion). This makes it likely that the archaebacteria arose when the ambient temperature of the planet was high, i.e., within the first billion years or so of earth history (43, 271). The ancestral archaebacterial environment seems also to have been highly reducing, for most archaebacteria today are fastidious anaerobes, again implicating rather early stages in earth history, when both hydrosphere and atmosphere would have been reducing (228).

Since stromatolites existed at least 3.5 billion years ago (231), eubacteria such as *Chloroflexus* almost certainly existed at that time (162). Given the distribution of thermophilia among eubacteria (see above), these organisms too appear to be of ancient and thermophilic origin. However, in eubacteria thermophilia is not as extreme as in archaebacteria; finding archaebacteria that grow at the boiling point of water is now common, but thermophilic eubacteria growing above 90°C are rare. The ancestral eubacterium might then have arisen later than the archaebacteria, when the planet was somewhat cooler (43, 271).

The indications are there: archaebacteria seem an ancient and primitive phenotype, moreso than the eubacteria. The experiments that would convincingly establish the point are apparent and some are now being done. Eucaryotes remain the puzzle. Were they too of thermophilic origin? How do they fit into this scenario developed for the bacteria?

THE UNIVERSAL ANCESTOR

All questions concerning relationships among the three urkingdoms and the general course of evolution in each ultimately turn upon the nature of their common ancestor. The nature of this universal ancestor is, in my opinion, probably the most important, and definitely the least recognized, major question in biology today. As we shall see, the universal ancestor may have been a kind of entity outside of our direct experience. Even to begin considering it, we have to question concepts generally taken for granted.

Consider how information is organized in various living systems. In procaryotes the bulk of the information in the cell occurs in very long, contiguous strings of genes; procaryotes could then be called "genomic" organisms. Lower eucaryotes are also genomic, but the higher ones, metazoa with complex development systems, higher nervous function, and social structures, should probably be designated "supragenomic" entities, for much of the information they contain lies in structures above the level of the genome. Entities simpler and more primitive than genomic ones must also have existed. An organism of this type could have had a genotype and phenotype (i.e., information stored in a quiescent [replicative] form in one class of molecules that was also manifested in an active [functional] form in another), but its genes would for the most part have been physically separate units; they would not be organized into large contiguous linear arrays. These less organized systems would be called "genetic" but not genomic entities. (The reason for distinguishing genetic from genomic entities will become apparent as we proceed.) At a still more primitive level, entities can be imagined in which the genotype and phenotype do not exist in the sense we know them. Instead the storage/replicative and the active/functional forms of information both reside in the same class of molecule, possibly in different configurations of a given molecule. This would be the stage of "nucleic acid life" (245–247), where translation as we know it has not yet evolved and nucleic acids have both genetic and enzymatic functions. The idea that there could have been such an early stage has recently become quite popular with the experimental demonstration that RNAs can have catalytic properties (26, 232). Still more primitive stages can be imagined in which the bulk of the information in the system is not in macromolecular primary structures, but is contained in autocatalytic/metabolic networks.

Progenotes

The progenote is a theoretical construct, an entity that, by definition, has a rudimentary, imprecise linkage between its genotype and phenotype (251, 256). (Extant organisms, which have precise, accurate links between genotype and phenotype, are then genotes.) The certainty that progenotes existed at some early stage in evolution follows from the nature of the translation apparatus. Translation is accomplished by a multicomponent (multigene) mechanism that includes on the order of 100 different macromolecular species, far too complex a system to have arisen initially fully formed. Like the radio, the automobile, and similar devices, translation had to evolve through stages, from a much more rudimentary mechanism to the present precisely functioning one (244–246, 248, 256). Its aboriginal forms had fewer components and, consequently, must have functioned less accurately than their modern counterparts. There seems no alternative to the conclusion that the progenote existed at some stage early in evolution.

Characteristics of the progenote. The limitations of its rudimentary translation mechanism ensure that the progenote was a highly unique entity, unlike any life found today. Without today's level of accuracy in translation, proteins of normal size could not have been synthesized without introducing (many) errors. This means the progenote could neither have had nor have evolved "modern" proteins (244). Its proteins would have been small or of nonunique sequence or both. (A collection of polypeptides all different from one another, but each an approximate translation of the same genetic sequence, is known as a "statistical protein" [244].) As a consequence the progenote's enzymes would not be as accurate and specific as their modern counterparts. This in turn would delimit the kinds of control mechanisms, the definition and number of states the system possessed, and so on. Biological specificity at the progenote stage had to have been generally lower than now exists (251).

Replicating a genome places a tremendous burden of accuracy on a cell. To reproduce a string of nucleic acid thousands of genes long without introducing significant error requires an extemely precise mechanism, which today involves a number of separate activities (proof-reading functions, error correcting systems, and so on), most of which utilize large proteins. Such an extensive, complex, and precise system would not be found in a progenote, which means that the progenote could not have carried (could not have accurately replicated) the number of genes found in modern cells (251). A factor of 10 drop in the accuracy means a proportionate reduction in the length of the genome. The progenote reasonably had error rates two, or even three, orders of magnitude greater than found in cells today. (A factor of 100 drop in accuracy would leave the mistake rate in the range of 1 part per million [monomer units introduced], which is still impressively accurate.) One therefore wonders how progenotes could have carried a sufficient number of (different) genes to make them even minimally functional cells.

This apparent paradox can be resolved by making the progenote a genetic, not a genomic, entity. Genes would then be disjoint, and they could have existed in high copy numbers, in which case an appropriately simple mechanism can be imagined that would detect errors in individual genes and selectively eliminate (not correct) the flawed ones (251). As a genetic entity, the progenote could, in spite of a relatively very error-prone gene replication process, carry a reasonable number of genes. Given disjoint genes that might assume functional configurations (251), it is likely that the informational macromolecule at this stage was RNA, the functional form of nucleic acid today, not DNA (251).

Was the Universal Ancestor a Progenote?

In principle the universal ancestor could have resembled any one of the three major types of extant organisms. It also could have in essence been a collage of all three, or have been very unlike any of them. I will argue that the last alternative is the correct one and that the universal ancestor was a progenote.

The evolution that transformed the universal ancestor into the individual ancestors of each of the three primary kingdoms was of a unique quality. Sequence distances between kingdoms (Fig. 4) seem large compared to the distances within kingdoms, this despite the fact that the bulk of evolutionary time has involved evolution within the kingdoms. (The existence of 3.5 billion-year-old stromatolites [231] implies the existence of photosynthetic bacteria at that time, and so the existence of the common eubacterial ancestor even earlier.) Therefore, the long sequence distances do not correspond to long times. The transition from the universal ancestor to the ancestors of each of the primary kingdoms had to have taken less than 1 billion years, perhaps far less if an appreciable fraction of earth's first billion years involved evolutionary stages that preceded that of the universal ancestor. It would seem that the tempo of evolution at the time of the universal ancestor was very high.

The types of phenotypic changes that accompanied the formation of the three primary kingdoms are of a special nature. General differences in cell architecture among the three groups are remarkable, as are their differences in intermediary metabolism, and each kingdom seems to have its own unique version of every fundamental cellular function: translation, transcription, genome replication and control, and so on. The kind of variation that subsequently occurred within each of the kingdoms is minor by comparison. Thus the mode of evolution accompanying the transition from the universal ancestor is unusual; far more novelty arose during formation of the primary kingdoms than during the subsequent evolutionary course in any one of them.

It is hard to avoid concluding that the universal ancestor was a very different entity than its descendants. If it were a more rudimentary sort of organism, then the tempo of its evolution would have been high and the mode of its evolution highly varied, greatly expanded.

Were the actual root of the universal tree (Fig. 4) located in the vicinity of the deepest branchings in any one of the three primary kingdoms, the above argument concerning sequence distances would not apply to that kingdom, which makes it conceivable that the universal ancestor had the basic phenotype of that group. (This argument is particularly attractive as regards the archaebacteria, for the group sits relatively close to the intersection of the three primary lineages; see Fig. 4.) However, this would still leave the problem of deriving the other two phenotypes from a third comparably complex one, which entails drastic changes at the molecular level in most functions in the cell. In my opinion the changes in overall cell structure, organization, etc., required to change one of the three phenotypes into either of the others are too drastic and disruptive to have actually occurred.

Accepting all this, the only solution to the problem is for the universal ancestor to have been a progenote. Since the progenote is far simpler and more rudimentary than extant organisms, the significant differences in basic molecular structures and processes that distinguish the three major types of organisms would be attributes that the universal ancestor never possessed. In other words, the more rudimentary versions of a function present in the progenote would become refined and augmented independently, and so uniquely, in each of its progeny lineages. This independent refinement (and augmentation) of a more rudimentary function, not the replacement of one complex function by a different complex version thereof (the beginning stages of which would be strongly selected against), is why remarkable differences in detail have evolved for the basic functions in each of the urkingdoms. Biological specificity does not arise full-blown in cells, and in the transition from the universal ancestor to its descendants we are witnessing the evolution of biological specificity itself.

If the universal ancestor were a progenote, a particular pattern (spectrum) of relationships would exist among the various functions in the three primary kingdoms that would be hard to explain otherwise. The progenote lacked most of the functions characteristic of cells today, and those it did possess existed in a primitive, imprecise form. Therefore, functions that were central to the progenote and its descendants would have undergone the least evolutionary change and so would be the most similar in organisms today. The translation apparatus is a case in point. Without it no genotype-phenotype relationship exists, enzymes as we know them cannot evolve, accurate replication of nucleic acids is impossible, etc. (251). Translation had therefore to be one of the earliest cellular functions to arise. As then expected, it is one of the most structurally conserved functions. Only the fine-tuning aspects of the process appear to differ from one kingdom to the next; the replacement of the thymidine residue found in eubacterial and eucaryotic tRNAs by 1-methyl-pseudouridine in archaebacterial tRNAs (73, 165) exemplifies the type of differences encountered.

Structure developed only in crude, primitive ways in the progenote would undergo significant refinement and augmentation in the descendant lineages. While such functions would be homologous in all kingdoms, they would be noticeable idiosyncratic and characteristic in each as well. RNA polymerase could be an example here (281, 282). Functions not present in the progenote would then be totally idiosyncratic or be analogous, not homologous, in kingdom comparisons. Aspects of genome organization may turn out to be examples here, for the progenote did not face the problem of organizing thousands of genes, i.e., of developing an ordered genome structure. It is also possible that some primitive function in the progenote becomes reworked into another function(s) in one (or more) of the primary lineages. In this case, we might expect to find examples of structural homology without functional homology between kingdoms. Some control mechanisms may fall into this class (H. Hartman, personal communication).

The hierarchy of diversification suggested by the progenote, from highly homologous structures, to slightly homologous ones, to analogous ones, to idiosyncratic structures, should define the order in which the various processes arose (or became functionally readapted) during cellular evolution. On the consistency of this picture will turn the validity and utility of the progenote concept.

The progenote is today the end of an evolutionary trail that starts with fact, progresses through inference, and fades into fancy. However, in science endings tend to be beginnings. Within a decade we will have before us at least an order of magnitude more evolutionary information than we now possess and will be able to infer a great deal more with a great deal more assurrance than we now can. The root of the universal tree will probably have been determined, many gene families will have been defined, the evolution of genomic organization and of control mechanisms will have become serious problems, the enzymatic capacities of RNA will be more thoroughly elucidated, and the relationship between the evolution of the planet and the life thereon will be much better understood. The concepts of the progenote and of nucleic acid life will have come into their own.

ACKNOWLEDGMENTS

I appreciate the help of R. R. Gutell in preparing three of the figures used herein.

My work is supported by grants from the National Aeronautics and Space Administration, the National Science Foundation, and the Office of Naval Research.

LITERATURE CITED

1. **Achenbach-Richter, L., R. Gupta, K. O. Stetter, and C. R. Woese.** 1987. Were the original eubacteria thermophiles? Syst. Appl. Microbiol. **9**:34–39.
2. **Altamura, S., P. Cammarano, and P. Londei.** 1986. Archaebacterial and eukaryotic ribosomal subunits can form active hybrid ribosomes. FEBS Lett. **204**:129–133.
3. **Alvarez, L.** 1983. Experimental evidence that an asteroid impact led to the extinction of many species 65 million years ago. Proc. Natl. Acad. Sci. USA **80**:627–642.
4. **Ambler, R. P., M. Daniel, J. Hermoso, T. E. Meyer, R. G. Bartsch, and M. D. Kamen.** 1979. Cytochrome c_2 sequence

variation among the recognised species of purple nonsulphur photosynthetic bacteria. Nature (London) **278**:659–660.
5. **Balch, W. E., G. E. Fox, L. J. Magrum, C. R. Woese, and R. S. Wolfe.** 1979. Methanogens: reevaluation of a unique biological group. Microbiol. Rev. **43**:260–296.
6. **Balch, W. E., L. J. Magrum, G. E. Fox, R. S. Wolfe, and C. R. Woese.** 1977. An ancient divergence among the bacteria. J. Mol. Evol. **9**:305–311.
7. **Barbour, A. G., K.-I. Amano, T. Hackstadt, L. Perry, and H. D. Caldwell.** 1982. *Chlamydia trachomatis* has penicillin-binding proteins but not detectable muramic acid. J. Bacteriol. **151**:420–428.
8. **Benecke, W.** 1912. Bau and Leben der Bakterien. B. G. Teubner, Leipzig.
9. **Blaylock, B. A., and T. C. Stadtman.** 1966. Methane biosynthesis by *Methanosarcina barkeri*: properties of the soluble enzyme system. Arch. Biochem. Biophys. **116**:138–158.
10. **Bonen, L., and W. F. Doolittle.** 1975. On the prokaryotic nature of red algal chloroplasts. Proc. Natl. Acad. Sci. USA **72**:2310–2314.
11. **Bonen, L., and W. F. Doolittle.** 1976. Partial sequences of 16S rRNA and the phylogeny of blue-green algae and chloroplasts. Nature (London) **261**:669–673.
12. **Bonen, L., and W. F. Doolittle.** 1978. Ribosomal RNA homologies and the evolution of the filamentous blue-green bacteria. J. Mol. Evol. **10**:283–291.
13. **Boone, D. R., S. Worakit, I. M. Mathrani, and R. A. Mah.** 1986. Alkaliphilic methanogens from high-pH lake sediments. Syst. Appl. Microbiol. **7**:230–234.
14. **Boublik, R., M. Wydro, W. Hellmann, and F. Jenkins.** 1979. Structure of functional *A. salina-E. coli* hybrid ribosome by electron microscopy. J. Supramol. Struct. **10**:397–404.
15. **Brock, T. D.** 1970. Biology of microorganisms. Prentice-Hall, Inc., Englewood Cliffs, N.J.
16. **Brock, T. D.** 1978. Thermophilic microorganisms and life at high temperatures, p. 174. Springer-Verlag, New York.
17. **Brockmann, H., Jr., and A. Lipinski.** 1983. Bacteriochlorophyll g. A new bacteriochlorophyll from *Heliobacterium chlorum*. Arch. Microbiol. **136**:17–19.
18. **Brooks, B. W., R. G. E. Murray, J. L. Johnson, E. Stackebrandt, C. R. Woese, and G. E. Fox.** 1980. Red-pigmented micrococci: a basis for taxonomy. Int. J. Syst. Bacteriol. **30**:627–646.
19. **Brosius, J., J. L. Palmer, J. P. Kennedy, and H. F. Noller.** 1978. Complete nucleotide sequence of a 16S ribosomal RNA gene from *Escherichia coli*. Proc. Natl. Acad. Sci. USA **75**:4801–4805.
20. **Brown, W. M., E. M. Prager, A. Wang, and A. C. Wilson.** 1982. Mitochondrial DNA sequences of primates: tempo and mode of evolution. J. Mol. Evol. **18**:225–239.
21. **Cammarano, P., A. Teichner, and P. Londei.** 1986. Intralineage heterogeneity of archaebacterial ribosomes: evidence for two physicochemically distinct ribosome classes within the third urkingdom. Syst. Appl. Microbiol. **7**:137.
22. **Cammarano, P., A. Teichner, P. Londei, M. Acca, B. Nicolaus, J. L. Sanz, and R. Amils.** 1985. Insensitivity of archaebacterial ribosomes to protein synthesis inhibitors: evolutionary implications. EMBO J. **4**:811–816.
23. **Campbell, A.** 1981. Evolutionary significance of accessory DNA elements in bacteria. Annu. Rev. Microbiol. **35**:44–83.
24. **Canale-Parola, E.** 1978. Motility and chemotaxis of spirochetes. Annu. Rev. Microbiol. **32**:69–99.
25. **Chatton, E.** 1937. Titres et travoux scientifiques. Sete, Sottano, Italy.
26. **Chec, T. R., and B. L. Bass.** 1986. Biological catalysis by RNA. Annu. Rev. Biochem. **55**:599–629.
27. **Clayton, B. J., and R. K. Clayton.** 1978. Properties of photochemical reaction centers purified from *Rhodopseudomonas gelatinosa*. Biochim. Biophys. Acta **501**:470–477.
28. **Collins, M. D., and T. A. Langworthy.** 1983. Respiratory quinone composition of some acidophilic bacteria. Syst. Appl. Microbiol. **4**:295–303.
29. **Cowan, S. T.** 1962. The microbial species—a macromyth?, p. 433–455. *In* G. C. Ainsworth and P. H. A. Sneath (ed.), Microbial classification. Twelfth Symposium of the Society for General Microbiology. Cambridge University Press, Cambridge.
30. **Daniels, C. J., R. Gupta, and W. F. Doolittle.** 1985. Transcription and excision of a large intron in the $tRNA^{trp}$ gene of an archaebacterium *Halobacterium volcanii*. J. Biol. Chem. **260**:3132–3134.
31. **Darwin, F.** 1887. The life and letters of Charles Darwin, vol. 3, p. 18. John Murray, London.
32. **De Ley, J.** 1978. Modern methods in bacterial taxonomy; evaluation, application, prospects, p. 347–357. *In* Proceedings of the 4th International Conference on Plant Pathogenic Bacteria, Angers, vol. 1. Gilbert-Clarey, Tours, France.
33. **De Ley, J., and J. De Smedt.** 1975. Improvements of the membrane filter method for DNA:rRNA hybridization. Antoinie van Leeuwenhoek J. Microbiol. Serol. **41**:287–307.
34. **de Rosa, M., S. de Rosa, A. Gambacorta, L. Minale, R. H. Thomson, and R. D. Worthington.** 1977. Caldariellaquinone, a unique benzo-b-thiopen-4,7 quinone from *Caldariella acidophila*, an extremely thermophilic and acidophilic bacterium. J. Chem. Soc. Perkin Trans. **1**:653–657.
35. **de Rosa, M., A. Gambacorta, and A. Gliozzi.** 1986. Structure, biosynthesis, and physiochemical properties of archaebacterial lipids. Microbiol. Rev. **50**:70–80.
36. **De Soete, G.** 1983. A least squares algorithm for fitting additive trees to proximity data. Psychometrika **48**:621–626.
37. **Dickerson, R. E.** 1980. Structural conservatism in proteins over three billion years: cytochrome with a touch of collagen, p. 227–249. *In* R. Srinivasan (ed.), Biomolecular structure, conformation, function and evolution, vol. 1. Pergamon Press, Oxford.
38. **Doolittle, R. F., and B. Blömback.** 1964. Amino acid sequence investigations of firbinopeptide from various mammals; evolutionary implications. Nature (London) **202**:147–152.
39. **Doolittle, W. F.** 1985. Genome structure in archaebacteria, p. 545–560. *In* C. R. Woese and R. S. Wolfe (ed.), The bacteria, vol. 8. Archaebacteria. Academic Press, Inc., New York.
40. **Doolittle, W. F., C. R. Woese, M. L. Sogin, L. Bonen, and D. Stahl.** 1975. Sequence studies on 16S ribosomal RNA from a blue-green alga. J. Mol. Evol. **4**:307–315.
41. **Dron, M., M. Rahire, and J.-D. Rochaix.** 1982. Sequence of the chloroplast 16S rRNA gene and its surrounding regions of *Chlamydomonas reinhardii*. Nucleic Acids Res. **10**:7609–7620.
42. **Eikmanns, B., and R. K. Thauer.** 1984. Evidence for the involvement and role of a corrinoid enzyme in methane formation from acetate in *Methanosarcina barkeri*. Arch. Microbiol. **142**:175–179.
43. **Ernst, W. G.** 1983. The early earth and the archaen rock record, p. 41–52. *In* J. W. Schopf (ed.), Earth's earliest biosphere. Princeton University Press, Princeton, N.J.
44. **Errede, B., and M. D. Kamen.** 1978. Comparative kinetic studies of cytochromes c in reaction with mitochondrial cytochrome c oxidase and reductase. Biochemistry **17**:1015–1027.
45. **Feick, R. G., and R. C. Fuller.** 1984. Topography of the photosynthetic apparatus of *Chloroflexus aurantiacus*. Biochemistry **32**:3693–3700.
46. **Felsenstein, J.** 1978. Cases in which parsimony or compatibility methods will be positively misleading. Syst. Zool. **27**:401–410.
47. **Felsenstein, J.** 1982. Numerical methods for inferring evolutionary trees. Q. Rev. Biol. **57**:379–404.
48. **Fitch, W. M.** 1971. Toward defining the course of evolution: minimum change for a specified tree topology. Syst. Zool. **20**:406–416.
49. **Fitch, W. M., and E. Margoliash.** 1967. Construction of phylogenetic trees. Science **155**:279–284.
50. **Fowler, V. J., N. Pfenning, W. Schubert, and E. Stackebrandt.** 1984. Towards a phylogeny of phototrophic purple sulfur bacteria: 16S rRNA oligonucleotide cataloguing of eleven species of *Chromatiaceae*. Arch. Microbiol. **139**:382–387.
51. **Fowler, V. J., F. Widdel, N. Pfennig, C. R. Woese, and E. Stackebrandt.** 1986. Phylogenetic relationships of sulfate- and sulfur-reducing eubacteria. Syst. Appl. Microbiol. **8**:32–41.

52. **Fox, G. E.** 1985. The structure and evolution of archaebacterial ribosomal RNA, p. 257–310. *In* C. R. Woese and R. S. Wolfe (ed.), The bacteria, vol. 8. Archaebacteria. Academic Press, Inc., New York.
53. **Fox, G. E., K. R. Luehrsen, and C. R. Woese.** 1982. Archaebacterial 5S ribosomal RNA. Zentralbl. Bakteriol. Parasitenkd. Infektionskr. Hyg. Abt. 1 Orig. Reihe C **3**:330–345.
54. **Fox, G. E., L. J. Magrum, W. E. Balch, R. S. Wolfe, and C. R. Woese.** 1977. Classification of methanogenic bacteria by 16S ribosomal RNA characterization. Proc. Natl. Acad. Sci. USA **74**:4537–4541.
55. **Fox, G. E., K. J. Peckman, and C. R. Woese.** 1977. Comparative cataloging of 16S ribosomal ribonucleic acid: molecular approach to prokaryotic systematics. Int. J. Syst. Bacteriol. **27**:44–57.
56. **Fox, G. E., E. Stackebrandt, R. B. Hespell, J. Gibson, J. Maniloff, T. A. Dyer, R. S. Wolfe, W. E. Balch, R. Tanner, L. Magrum, L. B. Zablen, R. Blakemore, R. Gupta, L. Bonen, B. J. Lewis, D. A. Stahl, K. R. Luehrsen, K. N. Chen, and C. R. Woese.** 1980. The phylogeny of prokaryotes. Science **209**:457–463.
57. **Fuchs, G., and E. Stupperich.** 1986. Carbon assimilation pathways in archaebacteria. Syst. Appl. Microbiol. **7**:364–369.
58. **Fuller, R. C., S. G. Sprague, H. Gest, and R. E. Blankenship.** 1985. A unique photosynthetic reaction center from *Heliobacterium chlorum*. FEBS Lett. **182**:345–349.
59. **Garrett, A. J., M. J. Harrison, and G. P. Manire.** 1974. A search for the bacterial mucopeptide component muramic acid in *Chlamydia*. J. Gen. Microbiol. **80**:315–318.
60. **Gest, H., and J. L. Favinger.** 1983. *Heliobacterium chlorum*, an anoxygenic brownish-green photosynthetic bacterium containing a "new" form of bacteriochlorophyll. Arch. Microbiol. **136**:11–16.
61. **Ghosh, A., J. Das, and J. Maniloff.** 1977. Lack of repair of ultraviolet light damage in *Mycoplasma gallisepticum*. J. Mol. Biol. **116**:337–344.
62. **Gibbons, N. E., and R. G. E. Murray.** 1978. Proposals concerning the higher taxa of bacteria. Int. J. Syst. Bacteriol. **28**:1–6.
63. **Gibson, J., W. Ludwig, E. Stackebrandt, and C. R. Woese.** 1985. The phylogeny of the green photosynthetic bacteria: absence of a close relationship between *Chlorobium* and *Chloroflexus*. Syst. Appl. Microbiol. **6**:152–156.
64. **Gibson, J., E. Stackebrandt, L. B. Zablen, R. Gupta, and C. R. Woese.** 1979. A phylogenetic analysis of the purple photosynthetic bacteria. Curr. Microbiol. **3**:59–64.
65. **Golding, G. B.** 1983. Estimation of DNA and protein sequence divergence: an examination of some assumptions. Mol. Biol. Evol. **1**:125–142.
66. **Goldschmidt, R.** 1940. The material basis of evolution. Yale University Press, New Haven, Conn.
67. **Graf, L., E. Roux, E. Stutz, and H. Kossel.** 1982. Nucleotide sequence of a *Euglena gracilis* chloroplast gene coding for the 16S rRNA: homologies to *E. coli* and *Zea mays* chloroplast 16S rRNA. Nucleic Acids Res. **10**:6369–6381.
68. **Green, C. J., G. C. Stewart, M. A. Hollis, B. S. Vold, and K. F. Bott.** 1985. Nucleotide sequence of *Bacillus subtilis* ribosomal RNA operon, *rrnB*. Gene **37**:261–266.
69. **Gulik, A., V. Luzzati, M. de Rosa, and A. Gambacorta.** 1986. Structure and polymorphism of tetraether lipids from *Sulfolobus solfataricus*. I. Crystallographic analysis. Syst. Appl. Microbiol. **7**:258.
70. **Gupta, R.** 1984. *Halobacterium volcanii* tRNAs. J. Biol. Chem. **259**:9461–9471.
71. **Gupta, R.** 1985. Transfer ribonucleic acids of archaebacteria, p. 311–344. *In* C. R. Woese and R. S. Wolfe (ed.), The bacteria, vol. 8. Archaebacteria. Academic Press, Inc., New York.
72. **Gupta, R., A. Broccoli, and C. R. Woese.** 1983. Sequence of the 16S ribosomal RNA from *Halobacterium volcanii*, an archaebacterium. Science **221**:656–659.
73. **Gupta, R., and C. R. Woese.** 1980. Unusual modification patterns in the transfer ribonucleic acids of archaebacteria. Curr. Microbiol. **4**:245–249.
74. **Gutell, R. R., H. F. Noller, and C. R. Woese.** 1986. Higher order structure in ribosomal RNA. EMBO J. **5**:1111–1113.
75. **Gutell, R. R., B. Weiser, C. R. Woese, and H. F. Noller.** 1985. Comparative anatomy of 16S-like ribosomal RNA. Prog. Nucleic Acid Res. Mol. Biol. **32**:155–216.
76. **Haeckel, E.** 1866. Generelle Morphologie der Organismen. Verlag Georg Reimer, Berlin.
77. **Haldane, J. B. S.** 1929. The origin of life, Ration. Annl. **1929**:148–169.
78. **Hallberg, C., and H. Baltcheffsky.** 1981. Solubilization and separation of two b-type cytochromes from a carotenoid mutatant of *Halobacterium halobium* FEBS Lett. **125**:201–204.
79. **Hartmann, R., H.-D. Sickinger, and D. Oesterhelt.** 1980. Anaerobic growth of halobacteria. Proc. Natl. Acad. Sci. USA **77**:3821–3825.
80. **Hase, T., W. Wakabayashi, H. Matsubara, L. Kercher, D. Oesterhelt, K. K. Rao, and D. O. Hall.** 1977. *Halobacterium halobium* ferredoxin: a homologous protein to chloroplast-type ferredoxins. FEBS Lett. **77**:308–310.
81. **Henderson, E., M. Oakes, M. W. Clark, J. A. Lake, A. T. Matheson, and W. Zillig.** 1984. A new ribosome structure. Science **225**:510–512.
82. **Hensel, R., W. Demharter, O. Kandler, R. M. Kroppenstedt, and E. Stackebrandt.** 1986. Chemotaxonomic and molecular-genetic studies of the genus *Thermus*: evidence for a phylogenetic relationship of *Thermus aquaticus* and *Thermus ruber* to the genus *Deinococcus*. Int. J. Syst. Bacteriol. **36**:444–453.
83. **Hespell, R. B., B. J. Paster, T. J. Macke, and C. R. Woese.** 1984. The origin and phylogeny of the bdellovibrios. Syst. Appl. Microbiol. **5**:196–203.
84. **Herzog, M., and L. Marteaux.** 1986. Dinoflagellate 17S rRNA sequence inferred from the gene sequence: evolutionary implications. Proc. Natl. Acad. Sci. USA **83**:8644–8648.
85. **Holland, H. D.** 1984. Evolution of the atmosphere and oceans, p. 12–16. Princeton University Press, Princeton, N.J.
86. **Holländer, R.** 1978. The cytochromes of *Thermoplasma acidophilum*. J. Gen. Microbiol. **108**:165–167.
87. **Hori, H., and S. Osawa.** 1979. Evolutionary change in 5S RNA secondary structure and a phylogenetic tree of 54 5S rRNA species. Proc. Natl. Acad. Sci. USA **76**:381–385.
88. **Horowitz, N. H.** 1945. On the evolution of biochemical syntheses. Proc. Natl. Acad. Sci. USA **31**:153–157.
89. **Huber, R., T. A. Langworthy, H. König, M Thomm, C. R. Woese, U. B. Sleytr, and K. O. Stetter.** 1986. *Thermotoga maritima* sp. nov. represents a new genus of unique extremely thermophilic eubacteria growing up to 90°C. Arch. Microbiol **144**:324–333.
90. **Huet, J., R. Schnabel, A. Sentenac, and W. Zillig.** 1983. Archaebacteria and eukaryotes possess DNA-dependent RNA polymerases of a common type. EMBO J. **2**:1291–1294.
91. **Hui, J., and P. P. Dennis.** 1985. Characterization of the ribosomal RNA gene cluster in *Halobacterium cutirubrum*. J. Biol. Chem. **260**:529–533.
92. **Imhoff, J., H. G. Truper, and N. Pfennig.** 1984. Rearrangement of the species and genera of the phototrophic "purple nonsulfur bacteria." Int. J. Syst. Bacteriol. **34**:340–343.
93. **Iwami, M., A. Muto, F. Yamao, and S. Osawa.** 1984. Nucleotide sequence of the *rrnB* 16S ribosomal RNA gene from *Mycoplasma capricolum*. Mol. Gen. Genet. **196**:317–322.
94. **Jackson, T. J., R. F. Ramaley, and W. G. Meinschein.** 1973. *Thermomicrobium*, a new genus of extremely thermophilic bacteria. Int. J. Syst. Bacteriol. **23**:28–36.
95. **Jarrell, K. F., G. D. Sprott, and A. T. Matheson.** 1984. Intracellular potassium concentration and relative acidity of the ribosomal proteins of methanogenic bacteria. Can. J. Microbiol. **30**:663–668.
96. **Jarsch, M., and A. Böck.** 1985. Sequence of the 16S ribosomal RNA gene from *Methanococcus vannielii*. Syst. Appl. Microbiol. **6**:54–59.
97. **Jensen, R. A.** 1985. Biochemical pathways in prokaryotes can be traced backward through evolutionary time. Mol. Biol. Evol. **2**:87–120.

98. **Jones, W. J., J. A. Leigh, F. Mayer, C. R. Woese, and R. S. Wolfe.** 1983. *Methanococcus jannaschii* sp. nov., an extremely thermophilic methanogen from a submarine hydrothermal vent. Arch. Microbiol. **136**:254–261.
99. **Jukes, T. H., and C. R. Cantor.** 1969. Evolution of protein molecules, p. 21–132. *In* H. N. Munro (ed.), Mammalian protein metabolism. Academic Press, Inc., New York.
100. **Kaine, B. P., R. Gupta, and C. R. Woese.** 1983. Putative introns in tRNA genes of procaryotes. Proc. Natl. Acad. Sci. USA **80**:3309–3312.
101. **Kandler, O., and H. König.** 1985. Cell envelopes of archaebacteria, p. 413–458. *In* C. R. Woese and R. S. Wolfe (ed.), The bacteria, vol. 8. Archaebacteria. Academic Press, Inc., New York.
102. **Kandler, O., and W. Zillig (ed.).** 1986. Proceedings of the EMBO Workshop on Molecular Genetics of Archaebacteria and the International Workshop on Biology and Biochemistry of Archaebacteria, Munich, June 1985. Gustav Fischer Verlag, Stuttgart, Federal Republic of Germany.
103. **Kates, M.** 1978. The phytanyl ether-linked polar lipids and isoprenoid neutral lipids of extremely halophilic bacteria. Prog. Chem. Fats Other Lipids **15**:301–342.
104. **Kenyon, D. H., and G. Steinman.** 1969. Biochemical predestination. McGraw-Hill Book Co., New York.
105. **Kimura, M.** 1982. The neutral theory as a basis for understanding the mechanism of evolution and variation at the molecular level, p. 3–58. *In* M. Kimura (ed.), Molecular evolution, protein polymorphism and the neutral theory. Japan Scientific Societies Press, Tokyo.
106. **Kimura, M.** 1983. The neutral theory of molecular evolution. Cambridge University Press, New York.
107. **Kjems, J., and R. A. Garrett.** 1985. An intron in the 23S ribosomal RNA gene of the archaebacterium *Desulfurococcus mobilis*. Nature (London) **318**:675–677.
108. **Kluyver, A. J., and C. B. van Niel.** 1936. Prospects for a natural system of classification of bacteria. Zentralbl. Bakteriol. Parasitenkd. Infektionskr. Abt. 1 **94**:369–403.
109. **König, H., and K. O. Stetter.** 1986. Studies on archaebacterial S-layers. Syst. Appl. Microbiol. **7**:300–309.
110. **König, W., H. Schlesner, and P. Hirsch.** 1984. Cell wall studies on budding bacteria of the planctomyces-pasteuria group and on a *Prosthecomicrobium* sp. Arch. Microbiol. **138**:200–205.
111. **Kreisl, P., and O. Kandler.** 1986. Chemical structure of the cell wall polymer of *Methanosarcina*. Syst. Appl. Microbiol. **7**:293–299.
112. **Kuhn, W., K. Fiebig, H. Hippe, R. A. Mah, B. A. Huser, and G. Gottschalk.** 1983. Distribution of cytochromes in methanogenic bacteria. FEMS Microbiol. Lett. **20**:407–410.
113. **Kushner, D. J.** 1985. The Halobacteriaceae, p. 171–214. *In* C. R. Woese and R. S. Wolfe (ed.), The bacteria, vol. 8. Archaebacteria. Academic Press, Inc., New York.
114. **Lake, J. A., M. W. Clark, E. Henderson, S. P. Fay, M. Oakes, A. Scheinman, J. P. Thornber, and R. A. Mah.** 1985. Eubacteria, halobacteria, and the origin of photosynthesis: the photocytes. Proc. Natl. Acad. Sci. USA **16**:3716–3720.
115. **Lake, J. A., E. Henderson, M. W. Clark, and M. Oakes.** 1984. Eocytes: a new ribosome structure indicates a kingdom with a close relationship to eukaryotes. Proc. Natl. Acad. Sci. USA **81**:3786–3790.
116. **Lane, D. J., B. Pace, G. J. Olsen, D. A. Stahl, M. L. Sogin, and N. R. Pace.** 1985. Rapid determination of 16S ribosomal RNA sequences for phylogenetic analysis. Proc. Natl. Acad. Sci. USA **82**:6955–6959.
117. **Langworthy, T. A.** 1982. Lipids of bacteria living in extreme environments. Curr. Top. Membr. Transp. **17**:45–77.
118. **Langworthy, T. A.** 1985. Lipids of archaebacteria, p. 459–498. *In* C. R. Woese and R. S. Wolfe (ed.), The bacteria, vol. 8. Archaebacteria. Academic Press, Inc., New York.
119. **Langworthy, T. A., G. Holzer, J. G. Zeikus, and T. G. Tornabene.** 1983. Iso- and anteiso-branched glycerol diethers of the thermophilic anaerobe *Thermodesulfotobacterium commune*. Syst. Appl. Microbiol. **4**:1–17.
120. **Langworthy, T. A., and J. L. Pond.** 1986. Archaebacterial ether lipids and chemotaxonomy. Syst. Appl. Microbiol. **7**:253–257.
121. **Lanyi, J. K.** 1968. Studies of the electron transport chain of extremely halophilic bacteria. 1. Spectrophotometric identification of the cytochromes of *Halobacterium cutirubrum*. Arch. Biochem. Biophys. **128**:716–724.
122. **Lanyi, J. K.** 1978. Light energy conversion in *Halobacterium halobium*. Microbiol. Rev. **42**:682–706.
123. **Larsen, H.** 1973. The halobacteria's confusion to biology. Antoinie van Leeuwenhoek J. Microbiol. Serol. **39**:383–396.
124. **Lechner, K., G. Wich, and A. Böck.** 1985. The nucleotide sequence of the 16S rRNA gene and flanking regions from *Methanobacterium formicicum*: on the phylogenetic relationship between methanogenic and halophilic archaebacteria. Syst. Appl. Microbiol. **6**:157–163.
125. **Leffers, H., and R. A. Garrett.** 1984. The nucleotide sequence of the 16S ribosomal RNA gene of the archaebacterium *Halococcus morrhuae*. EMBO J. **3**:1613–1619.
126. **Leinfelder, W., M. Jarsch, and A. Böck.** 1985. The phylogenetic position of the sulfur-dependent archaebacterium *Thermoproteus tenax*: sequence of the RNA gene. Syst. Appl. Microbiol. **6**:164–170.
127. **Lewin, R. A.** 1981. *Prochloron* and the theory of symbiogenesis. Annl. N.Y. Acad. Sci. **361**:325–328.
128. **Li, W.-H., C.-C. Luo, and C. Wu.** 1985. Evolution of DNA sequences, p. 1–94. *In* R. J. MacIntyre (ed.), Molecular evolutionary genetics. Plenum Publishing Corp., New York.
129. **Li, W.-Y., R. Reddy, D. Henning, P. Epstein, and H. Busch.** 1982. Nucleotide sequence of 7S RNA: homology to Alu DNA and LA 4.5S RNA. J. Biol. Chem. **257**:5136–5142.
130. **Ludwig, W., K.-H. Schleifer, H. Reichenbach, and E. Stackebrandt.** 1983. A phylogenetic analysis of the myxobacteria *Myxococcus fulvus*, *Stigmatella aurantiaca*, *Cystobacter fuscus*, *Sorangium cellulosum* and *Nannocystis exedens*. Arch. Microbiol. **135**:58–62.
131. **Ludwig, W., E. Seewaldt, R. Kilpper-Balz, K.-H. Schleifer, L. Magrum, C. R. Woese, G. E. Fox, and E. Stackebrandt.** 1985. The phylogenetic position of Streptococcus and Enterococcus. J. Gen. Microbiol. **131**:543–551.
132. **Ludwig, W., and E. Stackebrandt.** 1983. A phylogenetic analysis of *Legionella*. Arch. Microbiol. **135**:45–50.
133. **Luehrsen, K. R., G. E. Fox, M. S. Kilpatrick, R. J. Walker, H. Doomday, G. Krupp, and H. J. Gross.** 1981. The nucleotide sequence of the 5S rRNA from the archaebacterium *Thermoplasma acidophilum*. Nucleic Acids Res. **9**:965–970.
134. **Luehrsen, K. R., D. E. Nicholson, Jr., and G. E. Fox.** 1985. Widespread distribution of a 7S RNA in archaebacteria. Curr. Microbiol. **12**:69–72.
135. **Macario, A. J. L., and E. Conway de Macario.** 1985. Antibodies for methanogenic biotechnology. Trends Biotechnol. **3**:204–208.
136. **Madigan, M. T., and T. D. Brock.** 1977. Single "chlorobium-type" vessicles of phototrophically grown *Chloroflexus aurantiacus* observed using negative staining techniques. J. Gen. Microbiol. **102**:279–285.
137. **Maniloff, J., and H. J. Morowitz.** 1972. Cell biology of the mycoplasmas. Bacteriol. Rev. **36**:263–290.
138. **Margulis, L.** 1970. Origin of eucaryotic cells. Yale University Press, New Haven, Conn.
139. **Margulis, L., and K. V. Schwartz.** 1982. Five kingdoms: an illustrated guide to the phyla of life on earth. W. H. Freeman and Co., New York.
140. **Marmur, J., S. Falkow, and M. Mandel.** 1963. New approaches to bacterial taxonomy. Annu. Rev. Microbiol. **17**:329–372.
141. **Matheson, A. T.** 1985. Ribosomes of archaebacteria, p. 345–412. *In* C. R. Woese and R. S. Wolfe (ed.), The bacteria, vol. 8. Archaebacteria. Academic Press, Inc., New York.
142. **Matheson, A. T., K. A. Louie, and G. N. Henderson.** 1986. The evolution of the ribosomal "A" protein (L 12) in archaebacteria. Syst. Appl. Microbiol. **7**:147–150.
143. **Mayr, E.** 1942. Systematics and the origin of species. Columbia University Press, New York.
144. **Mayr, E.** 1982. The growth of bilogical thought: diversity, evolution, and inheritance. Belknap Press, Cambridge, Mass.

145. **McCarroll, R., G. J. Olsen, Y. D. Stahl, C. R. Woese, and M. L. Sogin.** 1983. Nucleotide sequence of the *Dictyostelium discoideum* small-subunit ribosomal ribonucleic acid inferred from the gene sequence: evolutionary implications. Biochemistry **22**:5858–5868.
146. **McCloskey, J. A.** 1986. Nucleoside modification in archaebacterial transfer RNA. Syst. Appl. Microbiol. **7**:246–252.
147. **Messing, J., J. Carlson, G. Hagen, I. Rubenstein, and A. Oleson.** 1984. Cloning and sequencing of the ribosomal RNA genes in maize: the 17S region. DNA **3**:31–40.
148. **Meyer, T. E., M. A. Cusanovich, and M. D. Kamen.** 1986. Evidence against use of bacterial amino acid sequence data for construction of all-inclusive phylogenetic trees. Proc. Natl. Acad. Sci. USA **83**:217–220.
149. **Miller, S. L., and L. E. Orgel.** 1974. The origin of life on Earth. Prentice-Hall, Inc., Englewood Cliffs, N.J.
150. **Mills, L. B., E. J. Stanbridge, W. D. Sedwick, and D. Korn.** 1977. Purification and partial characterization of the principle deoxyribonucleic acid polymerase from *Mycoplasmatales*. J. Bacteriol. **132**:641–649.
151. **Moritz, A., and W. Goebel.** 1985. Characterization of the 7S RNA and its gene from halobacteria. Nucleic Acids Res. **13**:6969–6979.
152. **Murray, R. G. E.** 1974. A place for bacteria in the living world, p. 4–9. *In* R. E. Buchanan and N. E. Gibbons (ed.), Bergey's manual of determinative bacteriology, 8th ed. The Williams & Wilkins Co., Baltimore.
153. **Naumann, E., K. Fahlbusch, and G. Gottschalk.** 1984. Presence of a trimethylamine:HS-coenzyme M methyltransferase in *Methansarcina barkeri*. Arch. Microbiol. **138**:79–83.
154. **Noll, K. M., K. L. Rinehart Jr., R. S. Tanner, and R. S. Wolfe.** 1986. Structure of component B (7-mercaptoheptanoylthreonine phosphate) of the methylcoenzyme M methylreductase system of *Methanobacterium thermoautotrophicum*. Proc. Natl. Acad. Sci. USA **83**:4238–4242.
155. **Noller, H. F.** 1984. Structure of ribosomal RNA. Annu. Rev. Biochem. **53**:119–162.
156. **Nomura, M., S. Mizushima, M. Ozaki, P. Traub, and P. E. Lowry.** 1969. Structure and function of ribosomes and their molecular components. Cold Spring Harbor Symp. Quant. Biol. **34**:49–61.
157. **Olsen, G. J., D. J. Lane, S. J. Giovannoni, and N. R. Pace.** 1986. Microbial ecology and evolution: a ribosomal RNA approach. Annu. Rev. Microbiol. **40**:337–355.
158. **Olsen, G. J., N. R. Pace, M. Nuell, B. P. Kaine, R. Gupta, and C. R. Woese.** 1985. Sequence of the 16S rRNA gene from the thermoacidophilic archaebacterium *Sulfolobus solfataricus* and its evolutionary implications. J. Mol. Evol. **22**:301–307.
159. **Oparin, A. I.** 1924. Proiskhozhdenie zhizny. Moscow. Izd. Moskovshii Rabochii.
160. **Oparin, A. I.** 1957. The origin of life on Earth, 3rd ed. Oliver and Boyd, Edinburgh.
161. **Oren, A., B. J. Paster, and C. R. Woese.** 1984. *Haloanaerobiaceae*: a new family of moderately halophilic, obligatory anaerobic bacteria. Syst. Appl. Microbiol. **5**:71–80.
162. **Oyaizu, H., B. Debrunner-Vossbrinck, L. Mandelco, J. A. Studier, and C. R. Woese.** 1987. The green non-sulfur bacteria: a deep branching in the eubacterial line of descent. Syst. Appl. Microbiol. **9**:47–53.
163. **Oyaizu, H., and C. R. Woese.** 1985. Phylogenetic relationships among the sulfate respiring bacteria, myxobacteria and purple bacteria. Syst. Appl. Microbiol. **6**:257–263.
164. **Pace, N. R., D. A. Stahl, D. J. Lane, and G. J. Olsen.** 1986. The analysis of natural microbial populations by ribosomal RNA sequence. Adv. Microb. Ecol. **9**:1–55.
165. **Pang, H., I. Makoto, Y. Kuchino, S. Nishimura, R. Gupta, C. R. Woese, and J. A. McCloskey.** 1982. Structure of a modified nucleoside in archaebacteria tRNA which replaces ribosylthymine. J. Biol. Chem. **257**:3589–3592.
166. **Paster, B. J., W. Ludwig, W. G. Weisburg, E. Stackebrandt, R. B. Hespell, C. M. Hahn, H. Reichenbach, K. O. Stetter, and C. R. Woese.** 1985. A phylogenetic grouping of the bacteroides, cytophagas and certain flavobacteria. Syst. Appl. Microbiol. **6**:34–42.
167. **Paster, B. J., E. Stackebrandt, R. B. Hespell, C. M. Hahn, and C. R. Woese.** 1985. The phylogeny of spirochetes. Syst. Appl. Microbiol. **5**:337–351.
168. **Pfennig, N.** 1977. Phototrophic green and purple bacteria: a comparative, systematic survey. Annu. Rev. Microbiol. **31**: 275–290.
169. **Pfennig, N., and H. G. Truper.** 1974. The phototrophic bacteria, p. 24–75. *In* R. E. Buchanan and N. E. Gibbons (ed.), Bergey's manual of determinative bacteriology, 8th ed. The Williams & Wilkins Co., Baltimore.
170. **Pierson, B. K., and R. W. Castenholz.** 1974. A phototrophic gliding filamentous bacterium of hot springs, *Chloroflexus aurantiacus*, gen. and sp. nov. Arch. Microbiol. **100**:5–24.
171. **Pierson, B. K., and J. P. Thornber.** 1983. Isolation and spectral characterization of photochemical reaction centers from the thermophilic green bacterium *Chloroflexus aurantiacus* strain J-10-fl. Proc. Natl. Acad. Sci. USA **80**:80–84.
172. **Pond, J. L., T. A. Langworthy, and G. Holzer.** 1986. Long-chain diols: a new class of membrane lipids from a thermophilic bacterium. Science **231**:1134–1135.
173. **Postgate, J. R.** 1969. Methane as a minor product of pyruvate metabolism by sulfate-reducing and other bacteria. J. Gen. Microbiol. **57**:293–302.
174. **Qu, L. H., B. Michot, and J.-P. Bachellerie.** 1983. Improved methods for structure probing in large RNAs: a rapid "heterologous" sequencing approach is coupled to the direct mapping of nuclease accessible sites. Application to the 5′ terminal domain of eukaryotic 28S rRNA. Nucleic Acids Res. **11**: 5903–5919.
175. **Raikov, I. B.** 1982. The protozoan nucleus: morphology and evolution. Cell Biol. Monogr. **9**:124–129.
176. **Razin, S.** 1978. The mycoplasmas. Microbiol. Rev. **42**:414–470.
177. **Reanney, D. C.** 1976. Extrachromosomal elements are possible agents of adaptation and development. Bacteriol. Rev. **40**:552–590.
178. **Rogers, M. J., J. Simmons, R. T. Walker, W. G. Weisburg, C. R. Woese, R. S. Tanner, I. M. Robinson, D. A. Stahl, G. Olsen, R. H. Leach, and J. Maniloff.** 1985. Construction of the mycoplasma evolutionary tree from 5S rRNA sequence data. Proc. Natl. Acad. Sci. USA **82**:1160–1164.
179. **Rubtsov, P. M., M. M. Musakhanov, V. M. Zakharyev, A. S. Krayev, K. G. Skryabin, and A. A. Bayev.** 1980. The structure of the yeast ribosomal RNA genes. I. The complete nucleotide sequence of the 18S ribosomal RNA gene from *Saccharomyces cerevisiae*. Nucleic Acids Res. **8**:5779–5794.
180. **Runnegar, B.** 1982. The Cambrian explosion: animals or fossils? J. Geol. Soc. Aust. **29**:395–411.
181. **Salim, M., and E. H. Maden.** 1981. Nucleotide sequence of *Xenopus laevis* ribosomal RNA inferred from gene sequence. Nature (London) **291**:205–208.
182. **Sapienza, C., and W. F. Doolittle.** 1982. Unusual physical organization of the halobacterial genome. Nature (London) **395**:384–389.
183. **Schleifer, K.-H., and O. Kandler.** 1972. Peptidoglycan types of bacterial cell walls and their taxonomic implications. Bacteriol. Rev. **36**:407–477.
184. **Schleifer, K.-H., and E. Stackebrandt.** 1983. Molecular systematics of prokaryotes. Annu. Rev. Microbiol. **37**:143–187.
185. **Schleifer, K.-H., J. Steber, and H. Mayer.** 1982. Chemical composition and structure of the cell wall of *Halococcus morrhuae*. Zentralbl. Bakteriol. Parasitenkd. Infektionskr. Hyg. Abt. 1 Orig. Reihe C **3**:171–178.
186. **Schlesner, H., and P. Hirsch.** 1984. Assignment of ATCC 27377 to *Pirella* gen. nov. as *Pirella staleyii* comb. nov. Int. J. Syst. Bacteriol. **34**:492–495.
187. **Schnabel, R., M. Thomm, R. Gerardy-Schahn, W. Zillig, K. O. Stetter, and J. Huet.** 1983. Structural homology between different archaebacterial DNA-dependent RNA polymerases analyzed by immunological comparison of their components. EMBO J. **2**:751–755.
188. **Schwartz, R. M., and M. O. Dayhoff.** 1978. Origins of

prokaryotes, mitochondria and chloroplasts. Science **199**:395–403.
189. **Schwarz, A., and H. Kossel.** 1980. The primary structure of 16S rRNA from *Zea mays* chloroplast is homologous to *E. coli* 16S rRNA. Nature (London) **283**:739–742.
190. **Searcy, D. G., and F. R. Whatley.** 1982. *Thermoplasma acidophilum* cell membrane: cytochrome b and sulfate-stimulated ATPase. Zentralbl. Bakeriol. Parasitenkd. Infektionskr. Hyg. Abt. 1 Orig. Reihe C **3**:245–257.
191. **Seewaldt, E., and E. Stackebrandt.** 1982. Partial sequence of 16S ribosomal RNA and the phylogeny of *Prochloron*. Nature (London) **295**:618–620.
192. **Simpson, G. G.** 1944. Tempo and mode in evolution. Columbia University Press, New York.
193. **Sneath, P. H. A.** 1962. The construction of taxonomic groups, p. 289–332. *In* G. C. Ainsworth and P. H. A. Sneath (ed.), Microbial classification. Twelfth Symposium of the Society for General Microbiology. Cambridge University Press, Cambridge.
194. **Sneath, P. H. A., and R. R. Sokal.** 1973. Numerical taxonomy. W. H. Freeman and Co., New York.
195. **Sogin, M. L., and H. J. Elwood.** 1986. Primary structure of the *Paramecium tetraurelia* small-subunit tRNA coding region: phylogenetic relationships within the ciliophora. J. Mol. Evol. **23**:53–60.
196. **Sogin, M. L., H. J. Elwood, and J. H. Gunderson.** 1986. Evolutionary diversity of eukaryotic small subunit rRNA genes. Proc. Natl. Acad. Sci. USA **83**:1383–1387.
197. **Sprague, V.** 1977 The zoological distribution of microsporidia, p. 335–385. *In* L. A. Bulla and T. C. Cheng (ed.), Comparative pathobiology, vol. 2. Systematics of microsporidia. Plenum Publishing Corp., New York.
198. **Stackebrandt, E., V. J. Fowler, W. Schubert, and J. F. Imhoff.** 1984. Towards a phylogeny of phototrophic purple bacteria—the genus *Ectothiorhodospira*. Arch. Microbiol. **137**:366–370.
199. **Stackebrandt, E., V. J. Fowler, and C. R. Woese.** 1983. A phylogenetic analysis of lactobacilli, *Pediococcus pentosaceus* and *Leuconostoc mesenteroides*. Syst. Appl. Microbiol. **4**:326–337.
200. **Stackebrandt, E., R. M. Kroppenstedt, and V. J. Fowler.** 1983. A phylogenetic analysis of the family *Dermatophilaceae*. J. Gen. Microbiol. **129**:1831–1838.
201. **Stackebrandt, E., B. J. Lewis, and C. R. Woese.** 1980. The phylogenetic structure of the coryneform group of bacteria. Zentralbl. Bakteriol. Parasitenkd. Infektionskr. Hyg. Abt. 1 Orig. Reihe C **101**:137–149.
202. **Stackebrandt, E., W. Ludwig, W. Schubert, F. Klink, H. Schlesner, T. Roggentin, and P. Hirsch.** 1984. Molecular genetic evidence for early evolutionary origin of budding peptidoglycan-less eubacteria. Nature (London) **307**:735–737.
203. **Stackebrandt, E., H. Pohla, R. Kroppenstedt, H. Hippe, and C. R. Woese.** 1985. 16S rRNA analysis of *Sporomusa*, *Selenomonas*, and *Megasphaera*: on the phylogenetic origin of Gram-positive eubacteria. Arch. Microbiol. **143**:270–276.
204. **Stackebrandt, E., and C. R. Woese.** 1979. A phylogenetic dissection of the family *Micrococcaceae*. Curr. Microbiol. **2**:317–322.
205. **Stackebrandt, E., and C. R. Woese.** 1981. Towards a phylogeny of the Actinomycetes and related organisms. Curr. Microbiol. **5**:197–202.
206. **Stackebrandt, E., and C. R. Woese.** 1981. The evolution of prokaryotes, p. 1–31. *In* M. J. Carlile, J. R. Collins, and B. E. B. Moseley (ed.), Molecular and cellular aspects of microbial evolution. Cambridge University Press, Cambridge.
207. **Stackebrandt, E., B. Wummer-Fuessl, V. Fowler, and K.-H. Schleifer.** 1981. DNA homologies and ribosomal RNA similarities among spore-forming members of the order *Actinomycetales*. Int. J. Syst. Bacteriol. **31**:420–431.
208. **Stahl, D. A., K. R. Luehrsen, C. R. Woese, and N. R. Pace.** 1981. An unusual 5S RNA from *Sulfolobus acidocaldarius*, and its implications for a general 5S RNA structure. Nucleic Acids Res. **9**:6129–6137.
209. **Stanier, R. Y.** 1970. Some aspects of the biology of cells and their possible evolutionary significance, p. 1–38. *In* H. P. Charles and B. C. J. G. Knight (ed.), Organization and control in prokaryotic and eukaryotic cells. Cambridge University Press, Cambridge.
210. **Stainer, R. Y., and C. B. van Niel.** 1941. The main outlines of bacterial classification. J. Bacteriol. **42**:437–466.
211. **Stanier, R. Y., and C. B. van Niel.** 1962. The concept of a bacterium. Arch. Microbiol. **42**:17–35.
212. **Stetter, K. O.** 1982. Ultrathin mycelia-forming organisms from submarine volcanic areas having an optimum growth temperature of 105°C. Nature (London) **300**:258–260.
213. **Stetter, K. O., and G. Gaag.** 1983. Reduction of molecular sulphur by methanogenic bacteria. Nature (London) **305**:309–311.
214. **Stetter, K. O., M. Thomm, J. Winter, G. Wildgruber, H. Huber, and W. Zillig.** 1981. *Methanothermus fervidus*, sp. nov., a novel extremely thermophilic methanogen isolated from an Icelandic hot spring. Zentralbl. Bakteriol. Parasitenkd. Infektionskr. Hyg. Abt. 1 Orig. Reihe C **2**:166–178.
215. **Stetter, K. O., and W. Zillig.** 1985. *Thermoplasma* and the thermophilic sulfur-dependent archaebacteria, p. 85–170. *In* C. R. Woese and R. S. Wolfe (ed.), The bacteria, vol. 8. Archaebacteria. Academic Press, Inc., New York.
216. **Stoeckenius, W.** 1981. Walsby's square bacterium: fine structure of an orthogonal procaryote. J. Bacteriol. **148**:352–360.
217. **Stoeckenius, W.** 1985. The rhodopsin-like pigments of halobacteria: light-energy and signal transducers in an archaebacterium. Trends Biochem. **10**:483–486.
218. **Stöffler, G., and M. Stöffler-Meilicke.** 1986. Electron microscopy of archaebacterial ribosomes. Syst. Appl. Microbiol. **7**:123–130.
219. **Stöffler-Meilicke, M., C. Böhme, O. Ströbel, A. Böck, and G. Stöffler.** 1985. The structure of ribosomal subunits from *Methanococcus vannielii* as determined by electron microscopy. Science **231**:1306–1308.
220. **Tanner, R. S., E. Stackebrandt, G. E. Fox, R. Gupta, L. J. Magrum, and C. R. Woese.** 1982. A phylogenetic analysis of anaerobic eubacteria capable of synthesizing acetate from carbon dioxide. Curr. Microbiol. **7**:127–132.
221. **Tanner, R. S., E. Stackebrandt, G. E. Fox, and C. R. Woese.** 1981. A phylogenetic analysis of *Acetobacterium woodii*, *Clostridium barkeri*, *Clostridium butyricum*, *Clostridium lituseburense*, *Eubacterium limosum*, and *Eubacterium tenue*. Curr. Microbiol. **5**:35–38.
222. **Tindall, B. J., H. N. M. Ross, and W. D. Grant.** 1984. *Natronobacterium* gen. nov. and *Natronococcus*, gen. nov., genera of haloalkalophilic archaebacteria. Syst. Appl. Microbiol. **5**:41–57.
223. **Tohdoh, N., and M. Sugiura.** 1982. The complete nucleotide sequence of a 16S ribosomal RNA gene from tobacco chloroplasts. Gene **17**:213–218.
224. **Tomioka, N., and M. Sugiura.** 1983. The complete nucleotide sequence of a 16S ribosomal RNA gene from a blue-green alga, *Anacystis nidulans*. Mol. Gen. Genet. **191**:46–50.
225. **Tu, J., D. Prangishvilli, H. Huber, G. Wildgruber, W. Zillig, and K. O. Stetter.** 1982. Taxonomic relations between archaebacteria including 6 novel genera examined by cross hybridization of DNAs and 16S rRNAs. J. Mol. Evol. **18**:109–114.
226. **van Niel, C. B.** 1946. The classification and natural relationships of bacteria. Cold Spring Harbor Symp. Quant. Biol. **11**:285–301.
226a. **Vossbrinck, C. R., J. V. Maddox, S. Friedman, B. A. Debrunner-Vossbrinck, and C. R. Woese.** 1987. Ribosomal RNA sequence suggests microsporidia are extremely ancient eukaryotes. Nature (London) **362**:411–414.
227. **Vossbrinck, C. R., and C. R. Woese.** 1986. Eukaryotic ribosomes that lack a 5.8S RNA. Nature (London) **320**:287–288.
228. **Walker, J. C. G., K. Klein, M. Schidlowski, J. W. Schopf, D. J. Stevenson, and M. R. Walter.** 1983. Environmental evolution of the archaen-early proterozoic earth, p. 260–290. *In* J. W.

Schopf (ed.), Earth's earliest biosphere. Princeton University Press, Princeton, N.J.
229. **Wallace, D. C., and H. J. Morowitz.** 1973. Genome size and evolution. Chromosoma **40**:121–126.
230. **Walsby, A. E.** 1980. A square bacterium. Nature (London) **283**:69–71.
231. **Walter, M. R.** 1983. Archaen stromatolites: evidence of the Earth's earliest benthos, p. 187–213. *In* J. W. Schopf (ed.), Earth's earliest biosphere. Princeton University Press, Princeton, N.J.
232. **Waugh, D. S., and N. R. Pace.** 1985. Catalysis by RNA. Bioessays **4**:6–61.
233. **Weisburg, W. G., T. P. Hatch, and C. R. Woese.** 1986. Eubacterial origin of chlamydiae. J. Bacteriol. **167**:570–574.
234. **Weisburg, W. G., Y. Oyaizu, H. Oyaizu, and C. R. Woese.** 1985. Natural relationship between bacteroides and flavobacteria. J. Bacteriol. **164**:230–236.
235. **Weisburg, W. G., C. R. Woese, M. E. Dobson, and E. Weiss.** 1985. A common origin of rickettsiae and certain plant pathogens. Science **230**:556–558.
236. **Whitehead, A. N.** Process and reality. 1929. The Macmillan Co., Toronto.
237. **Whitehead, A. N.** Modes of thought 1938. The Macmillan Co., Toronto.
238. **Whitman, W. B.** 1985. Methanogenic bacteria, p. 1–84. *In* C. R. Woese and R. S. Wolfe (ed.), The bacteria, vol. 8. Archaebacteria. Academic Press, Inc., New York.
239. **Whitman, W. B., J. Shieh, S. Sohn, D. S. Caras, and U. Premachandran.** 1986. Isolation and characterization of 22 mesophilic methanococci. Syst. Appl. Microbiol. **7**:235–240.
240. **Whittaker, R. H.** 1969. New concepts of kingdoms of organisms. Science **163**:150–160.
241. **Whittaker, R. H.** 1977. Broad classification: the kingdoms and the protozoans, p. 1–34. *In* J. P. Kreier (ed.), Parasitic protozoa, vol. 1. Academic Press, Inc., New York.
242. **Wildgruber, G., M. Thomm, H. König, K. Ober, T. Ricchiuto, and K. O. Stetter.** 1982. *Methanoplanus limicola*, a plate-shaped methanogen representing a novel family, the *Methanoplanaceae*. Arch. Microbiol. **132**:31–36.
243. **Wilson, A. C., S. S. Carlson, and T. J. White.** 1977. Biochemical evolution. Annu. Rev. Biochem. **46**:473–639.
244. **Woese, C. R.** 1965. On the evolution of the genetic code. Proc. Natl. Acad. Sci. USA **54**:1546–1552.
245. **Woese, C. R.** 1967. The present status of the genetic code, Prog. Nucleic Acid Res. Mol. Biol. **7**:107–172.
246. **Woese, C. R.** 1969. The biological significance of the genetic code. Prog. Mol. Subcell. Biol. **1**:5–46.
247. **Woese, C. R.** 1972. The emergence of genetic organization, p. 301–341. *In* C. Ponnamperuma (ed.), Exobiology. North-Holland Publishing Co., Amsterdam.
248. **Woese, C. R.** 1973. Evolution of the genetic code. Naturwissenschaften **60**:447–459.
249. **Woese, C. R.** 1979. A proposal concerning the origin of life on the planet Earth. J. Mol. Evol. **13**:95–101.
250. **Woese, C. R.** 1982. Archaebacteria and cellular origins: an overview. Zentralbl. Bakteriol. Parasitenkd. Infektionskr. Hyg. Abt. 1 Orig. Reihe C **3**:1–17.
251. **Woese, C. R.** 1983. The primary lines of descent and the universal ancestor, p. 209–233. *In* D. S. Bendall (ed.), Evolution from molecules to men. Cambridge University Press, Cambridge.
252. **Woese, C. R.** 1985. Why study evolutionary relationships among bacteria?, p. 1–30. *In* E. Stackebrandt and K. O. Schleifer (ed.), Evolution of prokaryotes. Academic Press, Inc. (London), Ltd., London.
253. **Woese, C. R., P. Blanz, and C. M. Hahn.** 1984. What isn't a pseudomonad: the importance of nomenclature in bacterial classification. Syst. Appl. Microbiol. **5**:179–195.
254. **Woese, C. R., P. Blanz, R. B. Hespell, and C. M. Hahn.** 1982. Phylogenetic relationships among various helical bacteria. Curr. Microbiol. **7**:119–124.
255. **Woese, C. R., B. Debrunner-Vossbrinck, H. Oyaizu, E. Stackebrandt, and W. Ludwig.** 1985. Gram-positive bacteria: possible photosynthetic ancestry. Science **229**:762–765.
256. **Woese, C. R., and G. E. Fox.** 1977. The concept of cellular evolution. J. Mol. Evol. **10**:1–6.
257. **Woese, C. R., and G. E. Fox.** 1977. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. Proc. Natl. Acad. Sci. USA **74**:5088–5090.
258. **Woese, C. R., J. Gibson, and G. E. Fox.** 1980. Do genealogical patterns in purple photosynthetic bacteria reflect interspecific gene transfer? Nature (London) **283**:212–214.
259. **Woese, C. R., R. Gupta, C. M. Hahn, W. Zillig, and J. Tu.** 1984. The phylogenetic relationships of three sulfur-dependent archaebacteria. Syst. Appl. Microbiol. **5**:97–105.
260. **Woese, C. R., R. Gutell, R. Gupta, and H. F. Noller.** 1983. Detailed analysis of the higher-order structure of 16S-like ribosomal ribonucleic acids. Microbiol. Rev. **47**:621–669.
261. **Woese, C. R., L. J. Magrum, and G. E. Fox.** 1978. Archaebacteria. J. Mol. Evol. **11**:245–252.
262. **Woese, C. R., J. Maniloff, and L. B. Zablen.** 1980. Phylogenetic analysis of the mycoplasmas. Proc. Natl. Acad. Sci. USA **77**:494–498.
263. **Woese, C. R., and G. J. Olsen.** 1986. Archaebacterial phylogeny: perspectives on the urkingdoms. Syst. Appl. Microbiol. **7**:161–177.
264. **Woese, C. R., M. L. Sogin, D. A. Stahl, B. J. Lewis, and L. Bonen.** 1976. A comparison of the 16S ribosomal RNAs from mesophilic and thermophilic bacilli. J. Mol. Evol. **7**:197–213.
265. **Woese, C. R., E. Stackebrandt, and W. Ludwig.** 1985. What are mycoplasmas: the relationship of tempo and mode in bacterial evolution. J. Mol. Evol. **21**:305–316.
266. **Woese, C. R., E. Stackebrandt, T. J. Macke, and G. E. Fox.** 1985. A phylogenetic definition of the major eubacterial taxa. Syst. Appl. Microbiol. **6**:143–151.
267. **Woese, C. R., E. Stackebrandt, W. G. Weisburg, B. J. Paster, M. T. Madigan, V. J. Fowler, C. M. Hahn, P. Blanz, R. Gupta, K. H. Nealson, and G. E. Fox.** 1984. The phylogeny of purple bacteria: the alpha subdivision. Syst. Appl. Microbiol. **5**:315–326.
268. **Woese, C. R., W. G. Weisburg, C. M. Hahn, B. J. Paster, L. B. Zablen, B. J. Lewis, T. J. Macke, W. Ludwig, and E. Stackebrandt.** 1985. The phylogeny of purple bacteria: the gamma subdivision. Syst. Appl. Microbiol. **6**:25–33.
269. **Woese, C. R., W. G. Weisburg, B. J. Paster, C. M. Hahn, R. S. Tanner, N. R. Krieg, H.-P. Koops, H. Harms, and E. Stackebrandt.** 1984. The phylogeny of purple bacteria: the beta subdivision. Syst. Appl. Microbiol **5**:327–336.
270. **Woese, C. R., and R. S. Wolfe (ed.).** 1985. The bacteria, vol. 8. Archaebacteria. Academic Press, Inc., New York.
271. **Worsley, T. R., R. D. Nance, and J. B. Moody.** 1986. Tectonic cycles and the history of the Earth's biogeochemical and paleoceanographic record. Paleoceanography **1**:233–263.
272. **Wright, S.** 1982. The shifting balance theory and macroevolution. Annu. Rev. Genet. **16**:1–16.
273. **Yamaizumi, Z., M. Ihara, Y. Kuchino, R. Gupta, C. R. Woese, and S. Nishimura.** 1982. Archaebacterial tRNA contains 1-methylinosine at residue 57 in the TψC loop. Nucleic Acids Res. Symp. Ser. **11**:209–213.
274. **Yang, D., Y. Oyaizu, H. Oyaizu, G. J. Olsen, and C. R. Woese.** 1985. Mitochondrial origins. Proc. Natl. Acad. Sci. USA **82**:4443–4447.
275. **Yang, D. C., B. Kaine, and C. R. Woese.** 1985. The phylogeny of archaebacteria. Syst. Appl. Microbiol. **6**:251–256.
276. **Zablen, L. B., M. S. Kissel, C. R. Woese, and D. E. Beutow.** 1975. Phylogenetic origin of the chloroplast and procaryotic nature of its ribosomal RNA. Proc. Natl. Acad. Sci. USA **72**:2418–2422.
277. **Zeikus, J. G.** 1977. The biology of methanogenic bacteria. Bacteriol. Rev. **41**:514–541.
278. **Zeikus, J. G., and R. S. Wolfe.** 1972. *Methanobacterium thermoautotrophicum* sp. n., an anaerobic, autotrophic, extreme thermophile. J. Bacteriol. **109**:707–713.
279. **Zillig, W., F. Gropp, A. Henschen, H. Neumann, P. Palm, W.-D. Reiter, M. Rettenberger, H. Schnabel, and S. Yeats.** 1986. Archaebacterial virus host systems. Syst. Appl. Micro-

biol. **7**:58–66.

280. **Zillig, W., I. Holz, G. Janekovic, W. Schafer, and W. D. Reiter.** 1983. The archaebacterium *Thermococcus celer* represents a novel genus within the thermophilic branch of the archaebacteria. Syst. Appl. Microbiol. **4**:88–94.
281. **Zillig, W., R. Schnabel, and K. O. Stetter.** 1985. Archaebacteria and the origin of the eukaryotic cytoplasm. Curr. Top. Microbiol. Immunol. **114**:1–18.
282. **Zillig, W., R. Schnabel, J. Tu, and K. O. Stetter.** 1982. The phylogeny of archaebacteria, including novel anaerobic thermoacidophiles, in the light of RNA polymerase structure. Naturwissenschaften **69**:197–204.
283. **Zuckerkandl, E., and L. Pauling.** 1965. Molecules as documents of evolutionary history. J. Theor. Biol. **8**:357–366.