Pergamon

Planet. Space Sci., Vol. 43, Nos. 1/2, pp. 167–177, 1995
Copyright © 1995 Elsevier Science Ltd
Printed in Great Britain. All rights reserved
0032–0633/95 $9.50+0.00

0032–0633(94)00167–7

Looking for the most "primitive" organism(s) on Earth today: the state of the art

Patrick Forterre

Institut de Génétique et Microbiologie, CNRS, URA 1354 Université Paris-Sud, Bat. 409, 91405 Orsay Cedex, France

Received 3 January 1994; revised 7 March 1994; accepted 7 March 1994

Abstract. Molecular phylogenetic studies have revealed a tripartite division of the living world into two procaryotic groups, *Bacteria* and *Archaea*, and one eucaryotic group, *Eucarya*. Which group is the most "primitive"? Which groups are sister? The answer to these questions would help to delineate the characters of the last common ancestor to all living beings, as a first step to reconstruct the earliest periods of biological evolution on Earth. The current "*Procaryotic dogma*" claims that procaryotes are primitive. Since the ancestor of Archaea was most probably a hyperthermophile, and since bacteria too might have originated from hyperthermophiles, the procaryotic dogma has been recently connected to the hot origin of life hypothesis. However, the notion that present-day hyperthermophiles are primitive has been challenged by recent findings, in these unique microorganisms, of very elaborate adaptative devices for life at high temperature. Accordingly, I discuss here alternative hypotheses that challenge the procaryotic dogma, such as the idea of a universal ancestor with molecular features in between those of eucaryotes and procaryotes, or the origin of procaryotes via thermophilic adaptation. Clearly, major evolutionary questions about early cellular evolution on Earth remain to be settled before we can speculate with confidence about which kinds of life might have appeared on other planets.

Introduction

All present-day organisms are modern, i.e. they have evolved during the last 4 billion years from really primitive entities. However, it has been a constancy for evolutionary biologists to search among contemporary organisms for those most likely to resemble some "primitive" ancestors. For scientists interested in the fascinating question of the origin of life, the hope is to identify one (or several) creatures with an intermediate level of complexity between the very first cellular systems and the most elaborate products of biological evolution on Earth. This, in turn, would help to imagine how these primordial cells emerged from the prebiotic era, and, by the way, which type of prebiotic chemistry produced them. Finally, looking for the most "primitive" living beings can help to delineate the "minimal" conditions required for the appearance of life on a planet, a prerequisite for exobiologists.

At the turn of the nineteenth century, Heackel, the father of phylogeny, placed a putative simple unicellular organism, "*the monera*", at the root of his universal (phylogenetic) trees. About 100 years later, the monera have been replaced at the base of contemporary phylogenies by the *procaryotes*, i.e. unicellular organisms lacking a nucleus (see for example the textbook by Alberts *et al.*, 1983). Actually, procaryotes appear to most biologists to be the simplest cellular forms living on Earth today, hence the most primitive ones (Stanier, 1970; Margulis, 1971; Cavalier-Smith, 1992). The procaryotic world includes all bacteria, whereas the eucaryotes (meaning true nucleus) embrace animals, plants, fungi, alguae and protists. In procaryotes, the genetic material (the DNA molecule) is free in the cytoplasm, while in eucaryotes, the genetic material (DNA in chromosomes) is separated from the machinery for protein synthesis (the ribosomes) by a specific endocellular membrane, the nuclear membrane (Fig. 1). In addition, the procaryotic cells are usually small, and they lack the network of protein filaments (cytoskeleton) present in eucaryotic cells. Most eucaryotic cells also contain several types of intracellular organelles, some of them being endosymbionts from ancient procaryotes (chloroplasts and mitochondria, see below).

The relative complexity of eucaryotic cells compared to bacteria, and the apparent primitiveness of the latter, explain the term procaryote itself, "pro"caryote meaning before the nucleus. I will refer here to the idea that procaryotes are indeed primitive and the ancestors of eu-

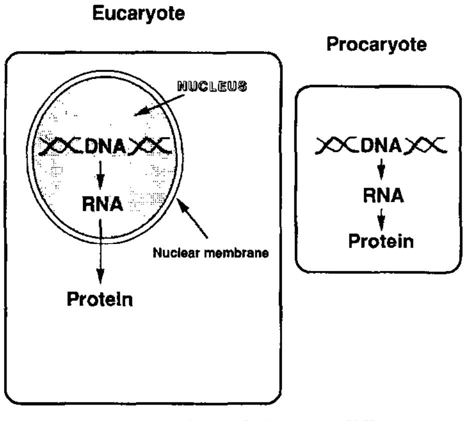

Fig. 1. Schematic organization of the two different types of cell in the living world, procaryotes vs eucaryotes

caryotes as the *procaryotic dogma*. This dogma has been challenged during the last 20 years only by a few biologists (Reanney, 1974; Carlile, 1982; Forterre, 1992a) who suggested that the "simplicity" of procaryotes corresponds to efficient integration of more complex structures induced by their life style, which is characterized by a very high reproduction rate and the faculty to adapt rapidly to a wide range of fluctuating environments. Others also proposed that some specific eucaryotic traits, such as genes interrupted by numerous non-coding segments (introns), could be the relics of ancient primitive structures dating from the earliest genetic organization (Darnell, 1978; Doolittle, 1978; Senapathy, 1986; Darnell and Doolittle, 1986).

However, all biological textbooks have endorsed the procaryotic dogma, and for most biologists, looking for the most primitive organisms and for the most primitive procaryotes are one and the same thing. For some time, the *mycoplasma* were considered as good candidates for such a position, because of their small genome and their lack of a cell envelope (Maniloff and Morowitz, 1972). More recently, *archaebacteria* became the most popular primitive creatures, as indicated by their name itself. Among them, the attention has been focused primarily on those living in very hot environments (terrestrial or submarine hot springs), the so-called hyperthermophiles, in connection with the recent hypothesis that life appeared at high temperatures (Corliss *et al.*, 1981; Baross and Hoffman, 1985; Woese, 1987; Pace, 1991; Stetter, 1992; Kandler, 1992; for a special issue of *Origin of Life*, see Holm, 1992). These new developments apparently validate the strategy previously described: the identification of the most primitive organisms among contemporary procaryotes, and the ensuing inference that useful information can be thus derived concerning the origin of life.

In this paper, I will review briefly the state of our knowledge concerning the origin and evolution of procaryotes, in relation to the problem of the origin of life. I will present the contradictory hypotheses proposed by different authors, trying to distinguish between experimental data and inspired speculations. The conclusion will be that much more work has to be done in basic microbiology and molecular biology to get a clear-cut answer concerning the nature of our cellular ancestor.

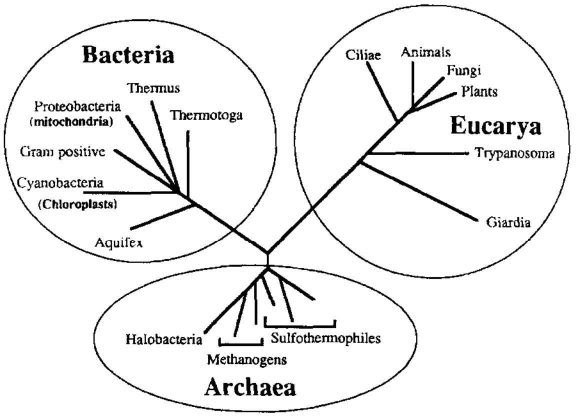

Fig. 2. Schematic universal phylogenetic tree based on rRNA sequence comparisons (adapted from Woese, 1987; Stetter, 1992). The length of the branches are roughly proportional to the number of mutations between sequences

Molecular biology and the "Woese-revolution"

Over the last 25 years, the problem of the origin of procaryotes has been completely revisited, following the revolution introduced in classification and phylogeny by the comparison of macromolecular sequences (reviewed by Woese, 1987, 1994). These sequences (either proteins or nucleic acids) have diverged in the course of biological evolution via the accumulation of spontaneous mutations. Therefore, the extent of divergence between homologous sequences in two different species is supposed to reflect their evolutionary distance (Zuckerkandl and Pauling, 1965) (note that two sequences are said to be *homologous* only if they actually diverged from a common ancestral one). The molecular biologist Carl Woese and his colleagues have successfully promoted the use of comparative analysis of ribosomal RNA (rRNA) sequences to delineate the phylogenetic relationships between widely divergent species (Woese, 1981). Indeed, rRNA molecules are very conserved and long enough to carry sufficient information to perform such an analysis with reasonable confidence. Furthermore, they play an essential role in protein synthesis, and one can safely assume that this role is and was similar in all present day and past organisms. As a consequence, the mutation rate was probably never biased by change in function.

The comparative analysis of rRNA sequences revealed, surprisingly, that cellular organisms can be divided, at the most fundamental level, into three widely divergent groups: two procaryotic ones, the *archaebacteria* (meaning old bacteria!) and the *eubacteria* (meaning true bacteria), the third one corresponding to the eucaryotes (Woese, 1981). This analysis has also shown that the evolutionary distance between the two procaryotic groups is about the same as the distance between any of them and the eucaryotes. Finally, the archaebacteria have a position in between the two other groups (Fig. 2). These discoveries were radical, as they substituted for the old dichotomy, *procaryotes vs eucaryotes*, a tripartite division of the living world.

Archaebacteria are a collection of diverse bacterial-like microorganisms with a broad phylogenetic diversity. They comprise methane-producing anaerobic species (*methano-*

gens), bacteria living in very high salt environments (*halobacteria*) and a huge number of diverse thermophilic species, aerobic and/or anaerobic, acidophilic or neutrophilic (*sulfothermophiles*). At the molecular level, archaebacteria exhibit a variety of typical "procaryotic" features such as a circular chromosome and genes grouped into operons; unique features, such as ether-lipids, and "eucaryotic features", such as a transcription system very similar to the eucaryotic ones (for a recent review, see Zillig, 1991; for recent books, see Danson *et al.*, 1992; Kates *et al.*, 1993).

To emphasize the deep evolutionary gap between eubacteria and archaebacteria, Woese and colleagues have recently suggested to rename these two procaryotic groups or "*Domains*", *Bacteria* (for eubacteria) and *Archaea* (for archaebacteria) (Woese *et al.*, 1990). This nomenclature is now widely used in the community of scientists working on "Archaea" and has been mainly criticized by a few biologists who still consider that phenotypic traits are more relevant to classification than genotypic ones (for references and reply see Wheelis *et al.*, 1992). I will use here this new nomenclature in spite of previous criticisms of my own (Forterre, 1992b; Forterre and Elie, 1993). Indeed, the name *Archaea* emphasizes the prejudice that these unique microorganisms are actually the primitive procaryotes looked for decades by evolutionists, which is not sure (see below). In the beginning, the primitiveness of Archaea was supported by two arguments: (i) the supposed similarity of the anaerobic biotopes of methanogens, rich in hydrogen and CO_2, to the putative primitive atmosphere of our planet, and (ii) the wide phylogenetic diversity of methanogens, suggesting an ancient divergence of this group (Woese, 1981). Later work has indicated that, in reality, each of the three domains exhibits a wide phylogenetic diversity (the methanogens themselves being only a part of the archaeal domain), the greater phylogenetic diversity being truly observed among eucaryotes (Sogin, 1991).

More recently, two research groups tried to add a new dimension to the phylogenetic work previously described by rooting the tree deduced from rRNA analysis (Gogarten *et al.*, 1989; Iwabe *et al.*, 1989). They reconstructed the phylogeny of homologous proteins which originated by duplication of an ancestral gene *before* the three domains diverged (see below for more explanation). From their work, they rooted the universal tree of life in the bacterial branch. This means that the most primordial lineage, the one common to all present day organisms, was first spilt between the bacterial lineage and a lineage common to Archaea and Eucarya (Fig. 3). In that scenario, the last common ancestor of the three domains, *the universal ancestor*, was probably a procaryotic-like creature since the two branches arising from this first split each contained procaryotes.

The rooting of the universal tree in the bacterial branch has received a warm welcome in the biological community since it seems to validate experimentally the procaryotic dogma (Woese *et al.*, 1990; Cavalier-Smith, 1991; Palmer and Logsdon, 1991; Stetter, 1992). In particular, Woese *et al.* (1990) used it as an argument to remove the "bacteria" from the term archaebacteria. However, this result is of such importance that its robustness should be carefully estimated. This has not been the case (Forterre *et al.*, 1993a,b) and I will return to this point later. Another consequence of the bacterial rooting was to boost the hot origin of life hypothesis. In effect, several authors have suggested that both *Archaea* and *Bacteria* evolved from thermophilic ancestors (Woese, 1987; Stetter, 1992). In that hypothesis, it is more parsimonious to imagine that the ancestor of all procaryotes was itself a hyperthermophile. Since the ancestor of all procaryotes and the universal ancestor are one and the same if the tree is rooted in the bacterial branch (Fig. 3), all present-day organisms should have originated from hyperthermophiles. Before evaluating this and recent alternative hypotheses, I will present a brief review of our knowledge of the history of each domain.

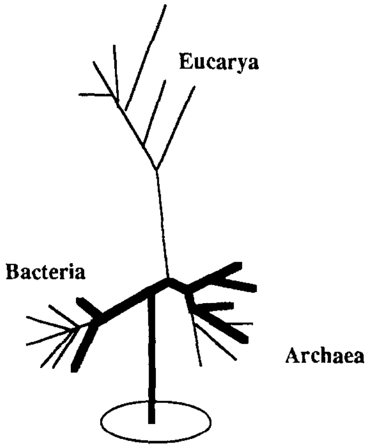

Fig. 3. The universal tree rooted in the bacterial branch, according to Gogarten *et al.* (1989) and Iwabe *et al.* (1989). This rooting has been criticized by Forterre *et al.* (1993a,b). The thermophilic lineages are in bold. The circle corresponds to the origin of life

History of the domain *Archaea* (archaebacteria)

The history of various lineages inside each domain is also presently inferred from rRNA phylogenies (Woese, 1987; Sogin, 1991). The domain *Archaea* has been divided by Woese and colleagues into two *kingdoms* according to such analysis: the *Euryarchaeota* (meaning diverse archaea, from *euryos*, "broad" in ancient Greek), encompassing the methanogens, the halophiles and a collection of sulfothermophiles, and the *Crenoarchaeota* (primordial-looking archaea, from *crenos*, "origin" in ancient Greek) comprising only thermophilic species (Woese *et al.*, 1990) (Fig. 4). Both kingdoms contain hyperthermophiles, i.e. organisms thriving in the temperature range from 80 to 110°C. Actually, the most thermophilic creatures known to date are *Archaea* and all organisms discovered so far growing in the temperature range 95–110°C belong to this domain (Stetter, 1992).

Several arguments strongly suggest that the ancestor of all present-day *Archaea* was itself a hyperthermophile (Woese, 1987; Achenbach-Richter *et al.*, 1988; Stetter, 1992): (i) the predominance of the thermophilic phenotype among all branches of the archaeal domains, (ii) the small length of the branches leading to hyperthermophilic archaeal species, suggesting that they have retained most

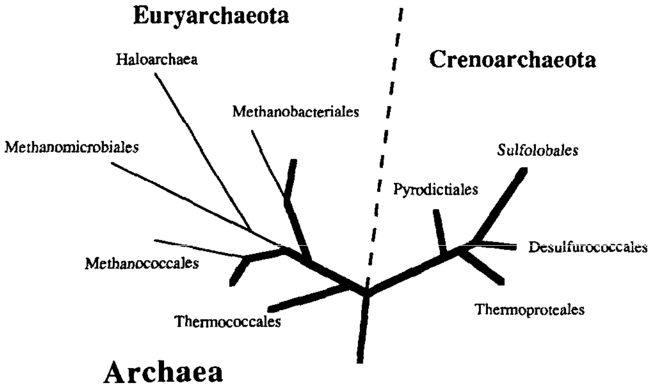

Fig. 4. A schematic phylogenetic tree of the archaeal domain, based on rRNA sequence comparisons (adapted from Woese, 1987; Woese *et al.*, 1990; Stetter, 1992). The thermophilic lineages are in bold

ancestral characters compared to their mesophilic counterparts, (iii) the position of the root of the archaeal rRNA tree, which is located between hyperthermophilic lineages of *Crenoarchaeota* and *Euryarchaeota*, (iv) the existence of unique lipids well designed for thermophilic life (isoprenyl-ether lipid instead of fatty-acid ester-lipids) in all members of the archaeal domain, including mesophiles, suggesting that the latter have retained those features from a thermophilic ancestor.

Most scientists working on *Archaea* today thus agree with the idea of a hyperthermophilic ancestor for this domain. In the current scenario (see for example Stetter, 1992), the first *Archaea* were anaerobic hyperthermophiles metabolizing sulfur, and methanogenesis originated among *Euryarchaeota*, shortly after their divergence from *Crenoarchaeota*. The aerobic archaea were newcomers that succeeded to invade aerobic niches after the accumulation of oxygen in the atmosphere. The latter hypothesis being supported by their longer branches in the rRNA tree. In that general scheme, *Crenoarchaeota* as a whole have more resemblance to the ancestral archaeal phenotype than *Euryarchaeota*, hence their name. However, this scenario may have to be somewhat revised after the recent detection of new *Archaea* by random cloning of PCR-amplified 16S rRNA gene. Barns *et al.* (1994) thus have discovered new hyperthermophiles archaea which bridges the gap between *Crenoarchaea* and *Euryarchaea*, whereas Delong (1992) has reported the presence in the oceanic bacterioplankton of *Crenoarchaea* which are most probably mesophiles.

No clear-cut microfossils of *Archaea* have been discovered up to now. However, the hydrogen to carbon isotopic ratio in rocks dating back to 3.5 billion years ago has suggested that methanogenesis might have been established by that time (Schopf, 1992). This would push the origin of *Archaea* back before that date. Indeed, the presumed existence of extensive volcanism at this early stage of the Earth's evolution would have been a favorable environment for such an early origin.

History of the domain *Bacteria* (eubacteria)

The domain *Bacteria* can be divided into more than 10 "kingdoms" based on rRNA sequence comparisons (Woese, 1987; Olsen *et al.*, 1994) (Fig. 5). One of them contains nearly all gram positive bacteria, another correspond to cyanobacteria (blue green algae), still another to spirochetes and so on. Most gram negative bacteria studied so far are concentrated into the kingdom proteobacteria, which can be itself divided into five subdivisions from α to ε. The famous *Escherichia coli* belongs to the subdivision γ. Interestingly, the mycoplasma have turned out to be the descendents of gram positive bacteria (Woese *et al.*, 1985). Therefore, these simple procaryotes are not relics of primitive forms, contrary to previous expectations, but have evolved from more "complex" bacteria. They have lost their cell wall and part of their genome in a process of regressive evolution driven by their parasitic state.

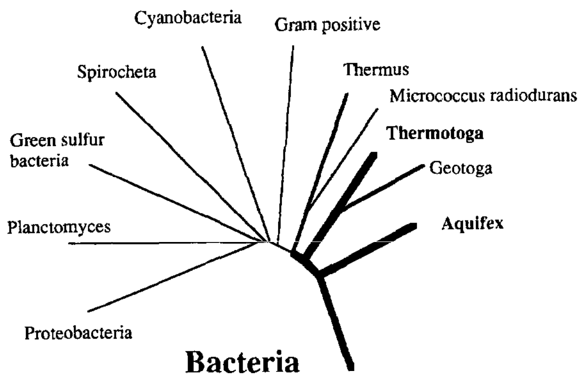

Fig. 5. A schematic phylogenetic tree of the bacterial domain based on rRNA sequence comparisons (adapted from Woese, 1987, Stetter, 1992). Genera containing thermophilic species growing above 80°C are in bold. In addition, genera containing hyperthermophilic species growing above 85°C are in large characters

Most of the bacterial kindgoms diverged from about the same position in the bacterial tree. However, two lineages apparently bifurcated slightly before all the others. Interestingly, both of them contain the most thermophilic bacteria known to date, the genera *Thermotoga* and *Aquifex*, that thrive in the temperature range of 75 to 95°C (Stetter, 1992). Since the root of the bacterial tree has been tentatively located between *Thermotoga* and *Aquifex*, the common ancestor of all bacteria might have been also a hyperthermophile. However, this hypothesis is not so well established as in the case of *Archaea*: since (i) mesophilic bacteria do not contain obvious relics of their putative thermophilic ancestors and (ii) the position of the root of the bacterial tree is not so robust. Indeed, Klenk *et al.* (1994) reported recently that the hyperthermophilic bacterion *Aquifex* is not the first lineage in a phylogenetic tree of bacterial DNA-directed RNA polymerases, but branches between gram positive *Bacteria* and proteobacteria. Thus a bias in the location of hyperthermophilic species cannot be ruled out and more work is clearly required to check the hypothesis of a thermophilic bacterial ancestor.

In contrast to *Archaea*, microfossils resembling existing *Bacteria*, more precisely cyanobacteria, have been detected in 3.5 billion-year-old formations (Schopf, 1992). This suggests that the divergence of cyanobacteria from the other bacterial kingdoms occurred before that date, i.e. that the bacterial domain should have appeared and

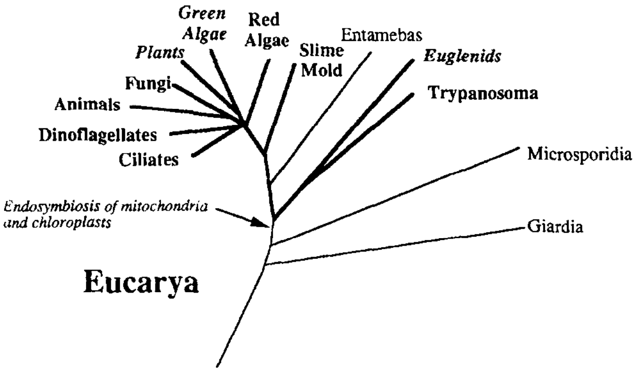

Fig. 6. A schematic phylogenetic tree of the eucaryal domain based on rRNA sequence comparisons (adapted from Sogin, 1991, 1992). The lineages containing mitochondria are in bold

radiated into different kingdoms between 4 and 3.5 billion years ago! This clearly indicates that it might have originated at about the same time as the domain *Archaea*.

History of the domain *Eucarya* (eucaryotes)

rRNA sequence comparisons have reaveled an unexpected diversity among eucaryotes (Sogin, 1991) (Fig. 6). Several lineages of protists (unicellular eucaryotes such as amoeba or paramecium) are the phylogenetic equivalent of a whole procaryotic "kingdom" and are much more phylogenetically divergent from each other than animals and plants. Two sequences of events can be identified in the history of eucaryotes. In a first evolutionary period, several lineages of protists diverged one after the other from the trunk of the eucaryal tree; then several branches radiated suddenly at about the same time, giving rise both to several new protist branches and to the various lineages of multicellular organisms; plants on one side, animals and fungi on the other (the eucaryal crown).

A major observation is that the earliest diverging lineages at the root of the eucaryotic tree are protists without mitochondria (Hypochondria) (Sogin, 1992). These organelles, which produce most of the cellular energy by oxidative respiration, are the descendants of ancient respiring bacteria which have established a symbiotic relationship with a primitive eucaryotic cell. Present-day eucaryotes lacking mitochondria could have therefore originated either from eucaryotes with mitochondria, via the loss of these organelles, or from "primitive" eucaryotes which never acquired them. Obviously, the second hypothesis is more likely for those present-day protists without mitochondria that branch early in the eucaryal tree.

Another major evolutionary event in the eucaryotic history has been the appearence of chloroplasts from endosymbiotic cyanobacteria, and the later emergence of green algae and plants. According to the rRNA phylogeny, green algae and plants appeared quite late in eucaryotic evolution, i.e. shortly before the radiation of the eucaryotic crown (Fig. 6). Accordingly, the recent finding of megascopic fossils resembling eucaryotic algae in a 2.1 billion-year-old formation (Han and Runnega, 1992) suggests that the radiation of the eucaryotic crown might have occurred before this date, and pushes back in time the origin of *Eucarya* themselves. This is in agreement with recent remarks by Sogin (1991) and Knoll (1992) who point out that primitive eucaryotes might have been as old as procaryotes. They noticed that "primitive" eucaryotes should have lacked cell walls and other structures suitable for fossilization. Accordingly, they claim that the greater antiquity of procaryotes in the fossil record might be a sampling artefact. If the radiation of the eucaryotic crown indeed took place more than 2 billion years ago, there is actually no reason to dismiss the possibility that procaryotes and eucaryotes diverged at about the same time as the archaeal/bacterial divergence, i.e. more than 3.5 billion years ago.

The datation of the radiation of the eucaryotic crown to 2 billion years ago is in good agreement with the current assumption that yeast and human diverged about 1.5 billion years ago. However, extrapolation of this value would put the divergence between eucaryotes and procaryotes deduced from the rRNA tree back to 9 billion years ago, i.e. before the Earth's formation! Accordingly, either an ancient life invaded our planet (panspermia) or we have to assume that the tick-tack of the rRNA clock was very different from its present value during the first billion years of evolution. This indicates that we have to be cautious in our interpretation of the first branching orders in the rRNA tree of each domain.

Finally, one should remark that, in contrast to the two procaryotic domains, there are no hyperthermophiles, nor even real thermophiles (living at temperatures above 65°C), in the domain *Eucarya*. The most thermophilic procaryotes are some fungi which have a maximum growth temperature of 62°C (Brock, 1985). In particular, the "primitive" eucarya without mitochondria are mesophiles. It is not clear why thermophilic eucaryotes apparently do not exist. Brock (1969) suggested that intracellular membranes such as the nuclear membranes present in eucaryotic cells, cannot support high temperatures because they contain large pores and could be intrinsically labile. However, it is not clear why thermostable derivatives of such membranes could not have evolved since the surface layer of hyperthermophilic archaea also contains large pores. Either hyperthermophilic eucaryotes exist but have not yet been isolated for technical reasons, or another explanation should be considered for their absence (see below).

The nature of the universal ancestor

The division of the living world into three domains led to the suggestion that comparison of their molecular features should permit the major characters of their last common ancestor to be reconstituted. Actually, this antique creature should have exhibited the *least common denominator* features of the three contemporary domains. Woese first suggested that the universal ancestor was a truly primitive organism, a "*progenote*" (Woese, 1981), i.e. an entity representing an evolutionary stage in between the very first living cells and modern procaryotes, fulfilling the dream of evolutionists. However, it turns out that molecular

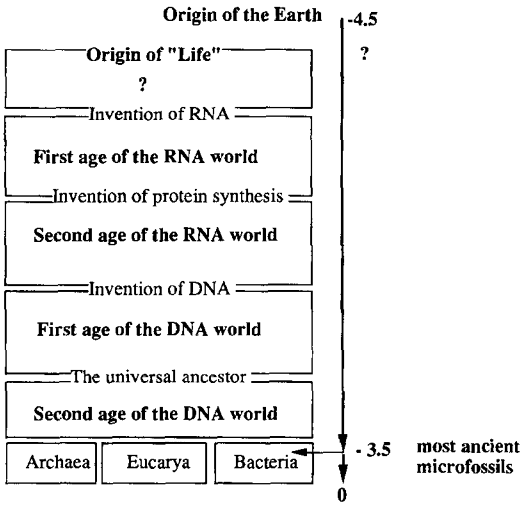

Fig. 7. Proposal to distinguish the different periods of early biological evolution which occurred between 4.5 and 3.5 billion years ago

features common to the three domains are extensive, including the same genetic code, highly homologous enzymatic mechanisms for gene expression (transcription and translation) and for DNA manipulation, homologous chaperone systems to assist protein biosynthesis and the formation of macromolecular complexes, similar metabolic pathways and so on. Thus, it can be safely concluded that the universal ancestor was not a very primitive entity but an already highly elaborated organism with a DNA genome (Lazcano *et al.*, 1992; Forterre, 1992a; Forterre *et al.*, 1993a). It remains from the progenote concept that the universal ancestor must have been very different from today's cells if one considers its rapid diversification into three different domains that, in contrast, remain subsequently stable in their basic molecular mechanism for the next 3.5 billion years (Woese, 1993).

This supports the idea that several important evolutionary steps occurred between the appearance of the very first cell and the emergence of the universal ancestor. One can tentatively propose the following: (1) a first age of the RNA world (RNA, peptides, lipids) before the invention of protein synthesis by RNA, (2) a second age of the RNA world (RNA, proteins, lipids), from the invention of protein synthesis up to the first DNA cell and (3) a first age of the DNA world from the invention of DNA up to the universal ancestor (Fig. 7).

Clearly, these decisive evolutionary steps should have occurred very rapidly since, as suggested by the history of each domain, the universal ancestor itself probably entered the scene well before 3.5 billion years ago. Furthermore, we can infer from the probable complexity of the universal ancestor that it could have resembled a modern procaryote (as suggested by the procaryotic dogma), a primitive eucaryote without mitochondria, or something in between. A first step to decide between these three different possibilities should be to root the universal tree of life. If the root turns out to be either in the bacterial or the archaeal branch, the universal ancestor was probably procaryote-like, whereas if the root is in the eucaryal branch, all possibilities remain, in particular the in-between hypothesis (Fig. 8). Furthermore, rooting the universal tree is a prerequisite to deciding if a molecular feature common to only two domains testifies to a common history, exclusive of the third one, or if it is a primitive trait which has been lost in only one domain.

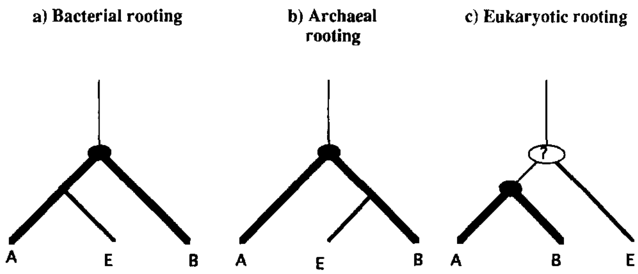

Fig. 8. The three possible rootings of the universal tree of life

A priori, the rooting of the universal tree can be obtained if one could identify pairs of genes present in all present-day organisms, and which have diverged from each other by duplication *before* the radiation of the three domains. Two universal trees could then be constructed (one for each gene) and the root of each tree can be located using the other as an outgroup (Fig. 9). As previously stated, this strategy has been exercised recently by two research groups (Gogarten *et al.*, 1989; Iwabe *et al.*, 1989). They studied two pairs of duplicated genes encoding proteins involved in fundamental biological functions. In both cases the root of the universal tree turns out to be located in the bacterial branch, promoting the concept of a procaryote-like ancestor. However, a critical analysis of the data indicates that these results are not robust (Forterre *et al.*, 1993a,b). In one case (the two elongation factors for protein synthesis), the two sister proteins are so divergent from each other that the location of the root is not statistically supported. In the other case (ATPase subunits), the determination of new protein sequences in *Bacteria* and *Archaea* later indicated that the proteins involved are probably not all sister but that some of them are cousin. This means that two cycles of duplication (and not only one) had occurred before the radiation of the three domains; hence the published phylogenetic trees contain a mixture of sister and cousin proteins, intro-

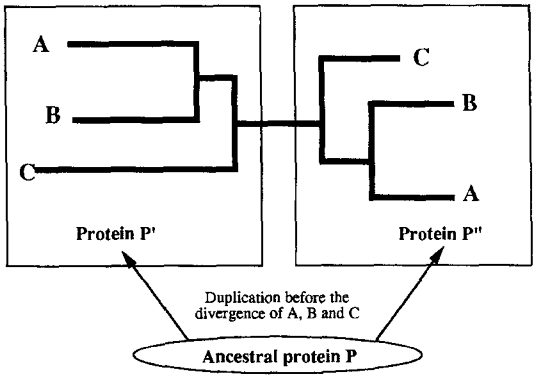

Fig. 9. Rationale to root the universal tree of life using two phylogenetic trees for cousin (paralogous) proteins; for explanations, see the text, as well as Gogarten *et al.* (1989) and Iwabe *et al.* (1989). A, B and C correspond to the three domains. In the case presented here, the root of the universal tree is located in the branch leading to C

ducing a strong bias in the position of the root. Accordingly, one can only safely conclude that the position of the root of the universal tree is presently unknown.

One can still hope that this crucial information will be obtained in the future with the accumulation of protein sequences and of new protein families amenable to phylogenetic analysis. However, this is not even sure, since methodological problems hamper the use of protein sequences to delineate robust phylogenies for very distantly related organisms (Meyer *et al.*, 1986). Nevertheless, new hypotheses are still flourishing in the meantime. I will present briefly two of them here: firstly, the hot origin of life hypothesis, which is very popular now, especially among scientists working on hyperthermophilic archaea, and a personal hypothesis, based on the criticism of the former and of the procaryotic dogma.

The hot origin of life hypothesis

The notion that life originated at high temperature emerged itself from four lines of speculation: (i) the hypothesis that the primitive Earth and its atmosphere and/or oceans were much hotter than today, (ii) the idea that deep-sea hydrothermal systems were the only places where early life could have found protection from the primordial catastrophic meteoritic bombardment, (iii) recent models suggesting that prebiotic chemistry began at solid–liquid interfaces at high temperature and (iv) the hypothesis of a thermophilic ancestor for all procaryotes.

The concept of a hot primitive Earth is based on geochemical models for the formation of the terrestrial planets that suggest a molten early Earth, and on the hypothesis of a primitive CO_2 atmosphere that predicts a greenhouse effect which could have maintained high temperatures and pressures in the primitive atmosphere and hydrosphere. I am not a specialist on these questions, but discussion with planetologists rapidly convinces one that several contradictory scenarios exist today (for example Bada *et al.* (1994) argue for a primitively frozen ocean regularly thawed by the shock of giant impactors) and that much work has to be done to cut the Gordian knot (for reviews, see Nisbet, 1985; Kasting, 1993). A major problem is that we do not know the rate of the cooling of the primordial Earth and the actual time of appearance of life.

The abyssal theory for the origin of life is popular because of the discovery of hyperthermophilic archaea in deep-sea vents (Pace, 1991), and because giant impactors might have sterilized the Earth several times between 4.5 and 3.8 billion years ago (Maher and Stevenson, 1988; Lazcano, 1993). Furthermore, it is generally admitted that volcanism in general, and hydrothermal systems in particular, were more widespread and active on primitive Earth than today (Nisbet, 1985). However, the idea that deep-sea vents would have afforded protection only for thermophiles is bizarre, since today these biotopes actually contain a majority of mesophilic bacteria (Prieur, 1992). One should therefore imagine, as an additional speculation, that such mesophiles were absent at the time of the giant impactors.

The concept of a hot prebiotic chemistry at solid–liquid interfaces has been promoted by Wächtershäuser (1988, 1992), who imagined a primitive, autotrophic two-dimensional form of life developing on pyrite crystalline surfaces (for a critique, see De Duve and Miller (1991) and for the reply, Wächtershäuser (1994)). In this model, the increase in atomic vibration at high temperature is supposed to be compensated for by the reduction of atomic movements in planar prebiotic chemistry, compared to three-dimensional prebiotic chemistry in solution. This theory is heuristic since it has already led to the discovery of new chemical reactions (Drobner *et al.*, 1990; Blöchl *et al.*, 1992; Keller *et al.*, 1994). However, much work has to be done here too in order to link this putative two-dimensional prebiotic chemistry to present-day three-dimensional biochemistry.

The problem with a hot prebiotic chemistry is that biological macromolecules are rapidly destroyed at temperatures near the boiling point of water by cleavage of essential covalent bonds (Miller and Bada, 1988; for a recent review, see Lindahl, 1993). This is a very hot topic, since advocates of the hot origin of life argue the reverse, i.e. high temperatures in highly pressurized environments could create instead of destroy organic compounds (Holm, 1992). Heat-induced hydrolysis of macromolecules is especially dramatic in the case of RNA, because this molecule contains a reactive hydroxyl group which can promote rapid intramolecular hydrolysis of the RNA polymer at elevated temperature. This is troublesome for the hot origin of life hypothesis, since it is now widely believed that the early evolution of life on Earth involved a stage during which primitive cells contained RNA instead of DNA as the genomic material (for a recent review, see Lazcano, 1993a). An even earlier stage with only RNA genes and RNA enzymes (ribozymes) is also likely (see Fig. 7). To bypass this problem of RNA instability it has been suggested that water activity was much lower in the putative primordial semi-mineral life setting, preventing heat-induced hydrolysis (Pace, 1991). However, the cellular stage of the RNA world (a prerequisite for its further evolution toward the DNA world) implies an active metabolism requiring water activity much like that in present-day organisms. In present thermophiles, stable RNA, such as transfer RNA, seem to be protected at some critical positions by methylation of the ribose reactive hydroxyl group (see below), an indication that RNA are indeed heat sensitive *in vivo*, whereas messenger RNA in prokaryotes are anyway short lived, and they can be protected by attachment to the ribosome as soon as they are synthesized by the RNA polymerase.

Finally, what about the phylogenetic argument supporting a universal thermophilic ancestor? This argument is based on the hypothesis that both *Archaea* and *Bacteria* are derived from a thermophilic ancestor (Woese, 1987) and is supported by the rooting of the tree of life in the bacterial branch (Stetter, 1992). I have already discussed these points and I will just summarize them here: (i) the hypothesis of a thermophilic ancestor for *Archaea* is well supported, (ii) the hypothesis of a thermophilic ancestor for *Bacteria* is weaker but fits well with present data of molecular phylogeny and (iii) the results in favor of the bacterial rooting cannot be trusted. In particular, this last point indicates that even if the common ancestor of both *Archaea* and *Bacteria* was a thermophile, the last common ancestor of the three domains may still have been a mesophile.

Accordingly, in my opinion, neither the phylogenetic data nor the cosmological data offer strong arguments in favor of the hot origin of life hypothesis. Furthermore, this hypothesis raises difficult problems related to the chemical stability of biomolecules, especially for RNA.

Another experimental approach regarding the hot origin of life hypothesis could be to test the popular idea that hyperthermophilic organisms are "primitive" by exploring their molecular machinery, especially those features related to "adaptation at high temperature". Are these truly secondary adaptations which would contradict the existence of a direct connection between hyperthermophiles and the hot origin of life, or are these actually primordial traits only conserved in hyperthermophiles? Several recent data apparently support the adaptative hypothesis. Firstly, all hyperthermophiles, either *Archaea* or *Bacteria*, possess an enzyme, called *reverse gyrase*, which seems to be specifically required to maintain the functionality of the DNA at high temperatures. This enzyme reduces the extent of uncoiling, or pathlength of the DNA double helix, and this effect could compensate the increase in the helical pathlength induced by the elevation of temperature. It has been shown recently that reverse gyrase is a combination of two other DNA-metabolizing enzymes, a DNA helicase (an enzyme which separates the two strands of the double helix) and a type I DNA topoisomerase (an enzyme which permits the crossing of one DNA strand through the other via a transient single-stranded break) (Confalonieri *et al.*, 1993). This suggests that reverse gyrase originated via the fusion of a DNA helicase and a DNA topoisomerase gene, a mechanism which has been documented for other proteins (Forterre *et al.*, 1994). If this is the case, reverse gyrase should have appeared well after the emergence of the first cell with a DNA genome, i.e. after the appearance of helicases and topoisomerases (Forterre *et al.*, 1995) (Fig. 10).

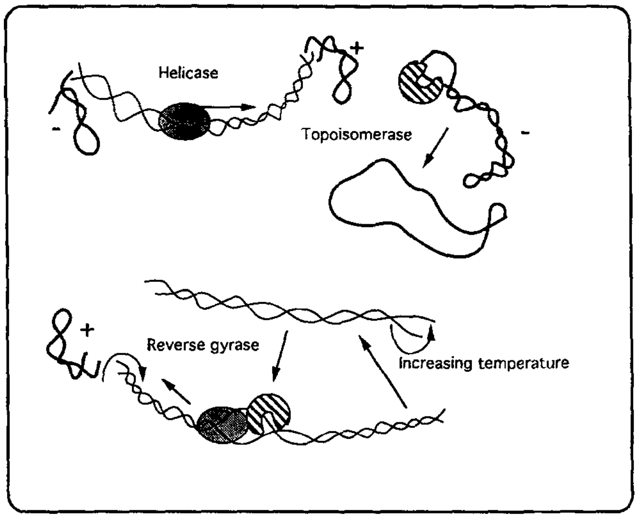

Fig. 10. Model explaining the origin of reverse gyrase from hyperthermophilic procaryotes, via the fusion of a DNA helicase and a DNA topoisomerase (adapted from Confalonieri *et al.*, 1993; Forterre *et al.*, 1994, 1995). A DNA helicase produces waves of positive (+) and negative (+) superturns when it tracks along the DNA double helix (bold lines represent the two DNA strands at once). The DNA topoisomerase domain of reverse gyrase can only eliminate negative superturns. The combination of the two activities produces positive superturns, which overwind the double helix. This effect could compensate the underwinding induced at high temperature in hyperthermophiles in order to maintain correct interaction between the DNA and the transcription apparatus

Similarly, tRNA molecules are protected in hyperthermophiles against thermodegradation and thermodestruction by numerous chemical modifications which are produced enzymatically after transcription of the tRNA genes (Edmonds *et al.*, 1991). This suggests that tRNA stabilization was not possible before the invention of tRNA modification enzymes. The very elaborate repair system operating to prevent DNA depurination in bacteria is also probably required in hyperthermophiles to deal with a depurination rate 3000 times higher at 100°C than at 37°C (Lindahl, 1993). These considerations tend to support a previous statement by Joyce (1988) that "primitive organisms" should be inept in the extreme. Accordingly, in my opinion, the more likely hypothesis for the origin of hyperthermophilia is that life invaded high temperature biotopes only after the invention of the major molecular machines operating in modern cells. In particular, adaptation to thermophily probably occurred only after DNA replaced RNA as the depository of the genetic information, since double-stranded DNA is much more resistant than RNA to thermodegradation, especially in the presence of salt, and when the two strands cannot rotate freely around each other, as it is the case in the cell (Forterre *et al.*, 1992; Marguet and Forterre, 1994).

If the secondary adaptation scenario for hyperthermophilic procaryotes is correct, one can still imagine that the highly evolved universal ancestor with a DNA genome has been a thermophile, but in that case, it is not clear why? It seems more logical to imagine the universal ancestor as a mesophile and to suppose that adaptation to thermophily occurred after the divergence between the ancestors of eucarya and the two procaryotic lineages. Adaptation to thermophily might have occurred independently in the two procaryotic domains, but another possibility is that it happened only once in a common lineage to *Archaea* and *Bacteria*. The latter hypothesis is presented in more detail below.

A (thermophilic) hypothesis for the origin of procaryotes

I have recently proposed a new hypothesis to explain both *why* the two procaryotic domains might have shared a thermophilic ancestor, and *why* truly thermophilic and hyperthermophilic eucaryotes do not exist (more precisely have not been found up to now) (Forterre, 1992a). Considering that the Achilles' heel of living beings at very high temperature is macromolecular degradation, especially RNA degradation, I have suggested that procaryotes succeeded in invading hyperthermophilic biotopes because of their high macromolecular turnover and the absence of a nuclear membrane. The latter point is particularly important since it permits coupling the synthesis of the messenger RNA on the DNA to its reading by the ribosome machinery, i.e. to decipher the RNA message before its thermodestruction. This could explain why there are no hyperthermophilic eucaryotes. Going a step further, I have proposed that procaryotes themselves originated from more complex mesophilic ancestors by a process of regressive evolution driven by adaptation to thermophily, i.e. reduction of the cell and genome size, pressure for integration and miniaturization of the cellular functions. This could explain why the ancestor of all procaryotes was probably a hyperthermophile. In the model proposed, the universal ancestor would have been a mesophile and the root of the tree of life should be located in the eucaryal branch, *Archaea* and *Bacteria* sharing a specific common ancestor (Fig. 8). This scenario is presently at odds with the major trend among archaebacteriologists who favor the hot origin of life hypothesis, but suggests at least a reasonable (I hope) alternative.

Perspectives

The search for the most "primitive" organisms still living today is a highly risky and uncertain task. In this paper, I have tried to convince the reader that even the well-established procaryotic dogma is not supported by real data. It is far from clear whether procaryotes are more primitive than eucaryotes or instead if they are "marvels of miniaturization" (Carlile, 1982) which already reached perfection 3.5 billion years ago. At the moment, it appears important to complete an exhaustive investigation of the two procaryotic domains at the molecular level, and to compare them with eucaryotes, to obtain a better understanding of early cellular evolution.

The complete sequencing of several bacterial genomes is now underway and similar projects are just starting with *Archaea*. This should help to pinpoint new families of paralogous proteins to eventually root the universal tree. Another possibility would be to use the cataloguing of paralogous proteins to reconstruct early stages of evolution by establishing the minimal set of homologous proteins present before emergence of the universal ancestor (Lazcano *et al.*, 1993). Since many proteins arose from the combination of different modules, as is the case for reverse gyrase, and since some modules have been used to construct many different proteins, it might be possible to imagine those primordial creatures which contained primitive proteins corresponding to present-day modules.

Another approach might be to use information from the viral world. The molecular biology of viruses is more diverse than that of the cell, and recent comparisons of viral and cellular homologous sequences suggest that they could have diverged before the radiation of cellular life into the three domains (Forterre, 1992a). Thus, viruses might have conserved molecular relics of creatures living in the RNA world or in the first age of the DNA world.

Finally, the rapid appearance and evolution of life on Earth raises the hope that life could have appeared and evolved during an early stage on some planet, such as Mars, that cannot sustain life in its present state. The finding of Martian microfossils, or even "dormant" creatures in protected Martian biotopes, obviously would be a must for evolutionary biologists. In that sense, the study of hyperthermophiles and other extremophiles is especially interesting. Even if these organisms are highly evolved, as I suggested here, they appeared very rapidly on Earth and could have survived in highly unusual environments. For example, halophilic archaea have been recovered alive in salt crystals extracted aseptically from salt mines (Norton *et al.*, 1993), and hyperthermophilic archaea thrive in 3000 m-deep oil reservoirs (Stetter *et al.*, 1993).

To conclude with a touch of hope, one can safely assume that unexpected discoveries will emerge in the near future from the association of microbiologists, molecular biologists, chemists and exobiologists, bringing new ideas about our universal ancestor, its origin on our fancy planet, and its putative relatives on others.

Acknowledgments—I am grateful to André Brack who gave me the opportunity to participate in the origin of life session of the EGS 1993 and to Christiane Elie for critical reading of the manuscript. The work in my laboratory is supported by a EEC grant from the Biotech Generic project on extremophiles.

References

Achenbach-Richter, L., R. Gupta, K. O. Stetter and C. R. Woese, Were the original eubacteria thermophiles? *Syst. Appl. Microbiol.* **9**, 34–39, 1987.

Achenbach-Richter, L., R. Gupta, W. Zillig and C. R. Woese, Rooting the archaebacterial tree: the pivotal role of *Thermococcus celer* in archaebacterial evolution. *Syst. Appl. Microbiol.* **10**, 231–240, 1988.

Alberts, B., D. Bray, J. Lewis, M. Raff, K. Roberts and J. D. Watson, *Molecular Biology of the Cell.* Garland, New York, 1983.

Bada, J. L., C. Bigham and S. Miller, Impact melting of a frozen ocean on the early earth and the implication for the origin of life. *Proc. Natn. Acad. Sci. USA* **91**, 1248–1250, 1994.

Barns, S., R. Fundyga, M. Jeffries and N. Pace, Remarkable archaeal diversity detected in a Yellowstone National Park hot spring environment. *Proc. Natn. Acad. Sci. USA* **91**, 1609–1613, 1994.

Baross, J. A. and S. E. Hoffman, Submarine hydrothermal vents and associated gradient environments as sites for the origin and evolution of life. *Origins of Life* **15**, 327–345, 1985.

Blöchl, E., M. Keller, G. Wächterhaüser and K. O. Stetter, Reactions depending on iron sulfide and linking geochemistry with biochemistry. *Proc. Natn. Acad. Sci. USA* **89**, 8117–8120, 1992.

Brock, T. D., Microbial growth under extreme conditions, in *Microbial Growth*, pp. 15–41. Cambridge University Press, Cambridge, 1969.

Brock, T. D., Life at high temperature. *Science* **230**, 132–138, 1985.

Carlile, M., Prokaryotes and eukaryotes: strategies and successes. *TIBS* **7**, 128–130, 1982.

Cavalier-Smith, T., Intron phylogeny: a new hypothesis. *TIG* **1**, 145–148, 1991.

Cavalier-Smith, T., Bacteria and eukaryotes. *Nature* **356**, 570, 1992.

Confalonieri, F., C. Elie, M. Nadal, C. Bouthier de la Tour,

P. Forterre and M. Duguet, Reverse gyrase: a helicase-like domain and a type I topoisomerase in the same polypeptide. *Proc. Natn. Acad. Sci. USA* **90,** 4753–4757, 1993.

Corliss, J. B., J. A. Baross and S. E. Hoffman, An hypothesis concerning the relationship between submarine hot springs and the origin of life on earth. *Oceanologica Acta Proc. 26th Int. Geol. Conf.*, pp. 59–69, 1981.

Danson, M. J., D. W. Hough and G. G. Lunt, The Archaebacteria: biochemistry and biotechnology. *Biochemical Society Symposium* **58,** 1992.

Darnell, J. E., Implications of RNA splicing in evolution of eukaryotic cells. *Science* **202,** 1257–1260, 1978.

Darnell, J. E. and W. F. Doolittle, Speculation on the early course of evolution. *Proc. Natn. Acad. Sci.* **83,** 1271–1275, 1986.

De Duve, C. and S. Miller, Two-dimensional life? *Proc. Natn. Acad. Sci. USA* **88,** 10014–10017, 1991.

Delong, E. F., Archaea in coastal marine environments. *Proc. Natn. Acad. Sci. USA* **89,** 5685–5689, 1992.

Doolittle, W. F., Genes in pieces: were they ever together? *Nature (London)* **272,** 581–582, 1978.

Drobner, E., H. Huber, G. Wächtershäuser, D. Rose and K. O. Stetter, Pyrite formation linked with hydrogen evolution under anaerobic conditions. *Nature* **346,** 742–744, 1990.

Edmonds, C. G., P. F. Crain, R. Gupta, T. Hashizume, C. H. Hocart, J. A. Kowalak, S. C. Pomerantz, K. O. Stetter and J. A. McCloskey, Posttranscriptional modification of tRNA in thermophilic archaea (archaebacteria). *J. Bacteriol.* **173,** 3138–3148, 1991.

Forterre, P., New hypotheses about the origins of viruses, prokaryotes and eukaryotes, in *Frontiers of Life*, pp. 221–234. Editions Frontières, Gif-sur-Yvette, France, 1992a.

Forterre, P., Neutral term. *Nature* **355,** 305, 1992b.

Forterre, P. and C. Elie, Chromosome structure, DNA topoisomerases, and DNA polymerases in archaebacteria (archaea), in *The Biochemistry of Archaea, New Comprehensive Biochemistry*, Vol. 26, pp. 325–366. Elsevier, Amsterdam, 1993.

Forterre, P., F. Charbonnier, E. Marguet and G. Henckes, Chromosome structure and DNA topology in extremely thermophilic archaea, *Biochemical Society Symposium* **58,** 99–112, 1992.

Forterre, P., N. Benachenhou-Lahfa, F. Confalonieri, M. Duguet, C. Elie and B. Labedan, The nature of the last universal ancestor and the root of the tree of life, still open questions. *Biosystems* **28,** 15–32, 1993a.

Forterre, P., N. Benachenhou-Lahfa and B. Labedan, Universal tree of life. *Nature* **362,** 796, 1993b.

Forterre, P. F., A. Bergerat, D. Gadelle, C. Elie, F. Lottspeich, F. Confalonieri, M. Duguet, M. Holmes and M. Dyall-Smith, Evolution of DNA topoisomerases and DNA polymerases: a perspective from Archaea. *System. Appl. Microbiol.* **16,** 746–758, 1994.

Forterre, P., F. Confalonieri, F. Charbonnier and M. Duguet, Speculations on the origin of life and thermophily: review of available information on reverse gyrase suggests that hyperthermophilic procaryotes are not so primitive. *Origin of Life and Evolution of the Biosphere* **25,** 235–249, 1995.

Gilbert, W., The RNA world. *Nature* **319,** 618, 1986.

Gogarten, J. P., H. Kibak, P. Dittrich, L. Taiz, E. J. Bowman, B. J. Bowman, M. F. Manolson, R. J. Poole, T. Date, T. Oshima, J. Konishi, K. Denda and M. Yoshida, Evolution of the vacuolar H^+-ATPase: implications for the origin of eukaryotes. *Proc. Natn. Acad. Sci. USA* **86,** 6661–6665, 1989.

Han, T.-M. and B. Runnega, Megascopic eukaryotic algae from the 2.1-billion-year-old negaunee iron-formation, Michigan. *Science* **257,** 232–235, 1992.

Holm, N. G., Marine hydrothermal systems and the origin of life. *Origin of Life and Evolution of the Biosphere* **22,** 1–242, 1992.

Iwabe, N., K. Kuma, M. Hasegawa, S. Osawa and T. Miyata, Evolutionary relationship of archaebacteria, eubacteria, and eukaryotes inferred from phylogenetic trees of duplicated genes. *Proc. Natn. Acad. Sci. USA* **86,** 9355–9359, 1989.

Joyce, G., Hydrothermal vents too hot? *Nature (London)* **334,** 564, 1988.

Kandler, O., Where next with the archaebacteria? *Biochem. Soc. Symp.* **58,** 195–207, 1992.

Kasting, J. F., Earth's early atmosphere. *Science* **259,** 920–926, 1993.

Kates, M., D. J. Kushner and A. T. Matheson, The biochemistry of Archaea (Archaebacteria), *New Comprehensive Biochemistry*. Elseview, Amsterdam, 1993.

Keller, M., E. Blöchl, G. Wächtershäuser and K. O. Stetter, Formation of amide bonds without a condensation agent and implications for origin of life. *Nature* **368,** 836–838, 1994.

Klenk, H. P., P. Palm and W. Zillig, DNA-dependent RNA polymerases as phylogenetic marker molecules. *System. Appl. Microbiol.* **16,** 638–647, 1994.

Knoll, A. H., The early evolution of eukaryotes: a geological perspective. *Science* **256,** 622–627, 1992.

Lazcano, A., The RNA world, its predecessors and descendants. *Early Life on Earth: Nobel Symposium* N° 84. Columbia University Press, New York, 1993a.

Lazcano, A., Biogenesis: some like it very hot. *Science* **260,** 1154–1155, 1993b.

Lazcano, A., G. Fox and J. Oro, Life before DNA: the origin and evolution of early archean cells, in *The Evolution of Metabolic Function*, pp. 237–295. CRC Press, Boca Raton, 1992.

Lazcano, A., E. Diaz-Villagomez, T. Mills and J. Oro, On the levels of enzymatic substrate specificity: implications for the early evolution of metabolic pathways. *Adv. Space. Res.*, in press.

Lindahl, T., Instability and decay of the primary structure of DNA. *Nature* **362,** 709–715, 1993.

Maher, K. A. and D. J. Stevenson, Impact frustration of the origin of life. *Nature (London)* **331,** 612–614, 1988.

Maniloff, J. and H. J. Morowitz, Cell biology of the mycoplasmas. *Bacteriol. Rev.* **36,** 263–290, 1972.

Marguet, E. and P. Forterre, DNA stability at temperatures typical for hyperthermophiles. *Nucleic Acids Res.* **22,** 1681–1686, 1994.

Margulis, L., Symbiosis and evolution. *Sci. Am.* **225,** 49–57, 1971.

Meyer, T. E., M. A. Cusanovich and M. D. Kamen, Evidence against use of bacterial amino acid sequence data for construction of all-inclusive phylogenetic trees. *Proc. Natn. Acad. Sci. USA* **83,** 217–220, 1986.

Miller, S. L. and J. L. Bada, Submarine hot springs and the origin of life. *Nature (London)* **334,** 609–611, 1988.

Nisbet, E. G., The geological setting of the earliest life forms. *J. Mol. Evol.* **21,** 289, 1985.

Norton, C. F., T. J. Mcgenity and W. D. Grant, Archaeal halophiles (Halobacteria) from two British salt mines. *J. Gen. Microbiol.* **139,** 1077–1081, 1993.

Olsen, G., C. R. Woese and R. Overbeek, The winds of (evolutionary) change: breathing new life into microbiology. *J. Bacteriol.* **176,** 1–6, 1994.

Pace, N. R., Origin of life-facing up to the physical setting. *Cell* **65,** 531–533, 1991.

Palmer, J. D. and J. M. Logsdon, The recent origins of introns. *Current Opinion Genet. Develop.* **1,** 470–477, 1991.

Prieur, D., Physiology and biotechnological potential of deep-sea bacteria, in *Molecular Biology and Biotechnology of Extremophiles*, pp. 163–197. Blakie, New York, 1992.

Reanney, D. C., On the origin of prokaryotes. *J. Theor. Biol.* **48,** 243–251, 1974.

Schopf, J. W., The oldest evidence of life, in *Frontiers of Life*, pp. 235–263. Editions Frontières, Gif-sur-Yvette, France, 1992.

Senapathy, P., Origin of eukaryotic introns: a hypothesis, based on codon distribution statistics in gene, and its implications. *Proc. Natn. Acad. Sci.* **83,** 2133–2137, 1986.

Sogin, M., Early evolution and the origin of eukaryotes. *Current Opinion Genet. Develop.* **1,** 457–463, 1991.

Sogin, M., Eukaryote origins and the protistean diversity, in *The Origin and Evolution of the Cell*, pp. 13–46. World Scientific, Singapore, 1992.

Stanier, R. Y., Some aspects of the biology of cells and their possible evolutionary significance. *Symp. Soc. Gen. Microbiol.* **20,** 1–39, 1970.

Stetter, K. O., Life at the upper temperature border, in *Frontiers of Life*, pp. 195–220. Editions Frontières, Gif-sur-Yvette, France, 1992.

Stetter, K. O., R. Huber, E. BlÖchl, M. Kurr, R. D. Eden, M. Flelder, H. Cash and I. Vance, Hyperthermophilic archaea are thriving in deep North Sea and Alaskan oil reservoir. *Nature* **365,** 743–745, 1994.

Wächtershäuser, G., Pyrite formation, the first energy source for life: a hypothesis. *Syst. Appl. Microbiol.* **10,** 207–210, 1988.

Wächtershäuser, G., Groundwork for an evolutionary biochemistry: the iron–sulphur world. *Prog. Biophys. Molec. Biol.* **58,** 85–201, 1992.

Wächtershäuser, G., Life in a ligand sphere. *Proc. Natn. Acad. Sci. USA* **91,** 4283–4287, 1994.

Wheelis, M. L., O. Kandler and C. R. Woese, On the nature of global classification. *Proc. Natn. Acad. Sci. USA* **89,** 2930–2934, 1992.

Woese, C. R., Archaebacteria. *Sci. Am.* **244,** 94–106, 1981.

Woese, C. R., Bacterial evolution. *Microbiol. Rev.* **51,** 221–271, 1987.

Woese, C. R., The archaea: their history and significance, in *The Biochemistry of Archaea, New Comprehensive Biochemistry*, Vol. 26, pp. vii–xxxi. Elsevier, Amsterdam, 1993.

Woese, C. R., There must be a prokaryote somewhere: microbiology's search for itself. *Microbiol. Rev.* **58,** 1–9, 1994.

Woese, C. R., E. Stackebrandt and W. Ludwig, What are micoplasmas: the relationship of tempo and mode in acterial evolution. *J. Mol. Evol.* **21,** 305–316, 1985.

Woese, C. R., O. Kandler and M. L. Wheelis, Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc. Natn. Acad. Sci.* **87,** 4576–4579, 1990.

Zillig, W., Comparative biochemistry of archaea and bacteria. *Current Opinion Genet. Develop.* **1,** 544–551, 1991.

Zuckerkandl, E. and L. Pauling, Molecules as documents of evolutionary history. *J. Theoret. Biol.* **8,** 357–366, 1965.