Hindawi Publishing Corporation
Archaea
Volume 2013, Article ID 202358, 12 pages
http://dx.doi.org/10.1155/2013/202358

Review Article

Close Encounters of the Third Domain: The Emerging Genomic View of Archaeal Diversity and Evolution

Anja Spang, Joran Martijn, Jimmy H. Saw, Anders E. Lind, Lionel Guy, and Thijs J. G. Ettema

Department of Cell and Molecular Biology, Science for Life Laboratory, Uppsala University, P.O. Box 596, 75123 Uppsala, Sweden

Correspondence should be addressed to Thijs J. G. Ettema; thijs.ettema@icm.uu.se

Received 24 July 2013; Accepted 21 September 2013

Academic Editor: Gustavo Caetano-Anollés

Copyright © 2013 Anja Spang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The Archaea represent the so-called Third Domain of life, which has evolved in parallel with the Bacteria and which is implicated to have played a pivotal role in the emergence of the eukaryotic domain of life. Recent progress in genomic sequencing technologies and cultivation-independent methods has started to unearth a plethora of data of novel, uncultivated archaeal lineages. Here, we review how the availability of such genomic data has revealed several important insights into the diversity, ecological relevance, metabolic capacity, and the origin and evolution of the archaeal domain of life.

1. Introduction

The description of the three (cellular) domains of life—Eukarya, Bacteria, and Archaea—by Carl Woese and George Fox [1] represents a milestone in the modern era of microbiology. In particular, using phylogenetic reconstructions of the small-subunit (16S or 18S) ribosomal RNA gene, Woese discovered that microscopically indistinguishable prokaryotes are not a homogeneous assemblage but are comprised of two fundamentally different groups of organisms: Eubacteria (later Bacteria) on one side and an additional life form referred to as Archaebacteria (later Archaea) on the other side [1]. Though not immediately accepted by the scientific community, this finding was early on supported by Wolfram Zillig through his studies on DNA-dependent RNA polymerases, as well as by Otto Kandler investigating "bacterial" cell walls [2]. Indeed, a subset of prokaryotic organisms subsequently assigned to Archaea was found to harbor DNA-dependent RNA polymerases that bore more similarity to those of eukaryotes, and to contain proteinaceous cell walls that lack peptidoglycan as well as cell membranes composed of L-glycerol ether lipids with isoprenoid chains instead of D-glycerol ester lipids with fatty acid chains [3–6]. Since then, further investigation of cellular characteristics of archaea has revealed that this domain of life contains eukaryotic-like information-processing machineries [7–14]. These findings were later supported by genome sequences and comparative analyses of genes coding for replication, transcription, and translation machineries as well as by protein crystal structures [15–21]. Additionally, some archaeal lineages were shown to contain homologs of eukaryotic cell division and cytoskeleton genes as well as histones and seem to express a chromatin architecture similar to eukaryotes [22–28]. In contrast to information-processing and cell division genes, archaeal operational systems (energy metabolism, biosynthesis pathways, and regulation) often appear to be more closely related to bacteria [29].

Based on phylogenetic reconstructions of the evolutionary history of 16S rRNA genes, the domain Archaea was originally divided into two major phyla: the Euryarchaeota and Crenarchaeota [30], which were separated by a deep split and thought to comprise only extremophilic (thermophilic, halophilic, and acidophilic) as well as methanogenic organisms. However, novel culture-independent and high-throughput sequencing techniques have recently uncovered a huge diversity of so far uncharacterized microorganisms on Earth as well as the ubiquitous occurrence of archaeal species [31–33]. Many of these novel archaeal groups are responsible for important ecological

processes and are only distantly related to established lineages within Cren- and Euryarchaeota [31, 32, 34–39]. For example, the acquisition of genome sequences from novel archaeal representatives has led to the proposal of several additional archaeal phyla (including Nanoarchaeota, Korarchaeota, Thaumarchaeota, Aigarchaeota, and Geoarchaeota) [40–46] and the investigation of uncultivated archaea using single cell genomics has already started to add new insights into the phylogenetic diversity of the Third Domain of life and necessitates the definition of additional lineages of high taxonomic rank including novel potential phyla and superphyla [33, 39] (see also below). Furthermore, the investigation of the metabolic potential of these novel organisms has provided fundamentally new insights into major biogeochemical nutrient cycles. Indeed, archaea are now recognized as key players in various biogeochemical processes [47]. For example, the perception of the global nitrogen cycle has been deeply altered by discovering that the ability to gain energy solely from ammonia was not limited to a few bacteria but also included the ammonia-oxidizing Thaumarchaeota [48, 49]. Archaea also appear to play a significant role in the carbon cycle, since, in addition to all known methanogenic organisms on Earth, they also encompass anaerobic methane oxidizing archaea (ANME lineages 1–3) [50].

The study of archaeal genomes and diversity is also of considerable importance for a better understanding of eukaryotic evolution. Indeed, the discovery of eukaryotic features in archaea [10] has initiated a new basis for addressing the origin of eukaryotes [51–54]. Interestingly, recent phylogenetic analyses of universal proteins have suggested that eukaryotes might have evolved from a *bona fide* archaeal lineage that forms a sister-lineage of or a lineage emerging from within the TACK-superphylum comprised of Thaum-, Aig-, Cren-, and Korarchaeota [55–58].

Below we give a contemporary overview of how recent developments in archaeal genomic research have contributed to revealing new insights into the diversity, ecological relevance, metabolic capacity, and the origin and evolution of the archaeal domain of life.

2. The Methanogenic Nature of Archaea

The scientific community that addressed questions about prokaryotic energy metabolism on the early Earth or in hydrothermal vent systems [59] has proposed that methanogenesis and/or acetogenesis most likely represent ancient metabolic pathways [60–62]. Evidence for the biological production of methane as early as 3.46 Gyr ago supports these scenarios [63]. However, phylogenetic evidence placing methanogens at the base of the archaeal tree is limited and disputed. Depending on the outgroup and phylogenetic methods used, many recent analyses find either members of the Nanohaloarchaea, Nanoarchaeota, ARMAN-lineages, and/or *Thermococcales* as earliest (eury-)archaeal branches [55, 64, 65]. The latter observation is consistent with results from a base and amino acid composition analysis, which indicated that last archaeal common ancestor (LACA) was a hyperthermophilic organism [66]. The placement of *Methanopyrus kandleri* as the most basal branch of archaea in some of the earliest phylogenetic analyses can most likely be attributed to long-branch attraction (LBA) artifacts [67]. Notably, in recent phylogenetic analyses that include novel archaeal single cell genomes, Euryarchaeota form a sister group to other archaeal phyla rather than representing an early diverging lineage (Figure 1) [33]. Furthermore, gene content comparisons of extant archaeal lineages and reconstruction of the putative genetic repertoire of the LACA do not support methanogenesis as the earliest archaeal metabolism [57, 68]. In contrast, only one study has so far placed the root of archaea within a methanogenic order [69] and thus favors a methanogenic origin of the Third Domain of life. Gene content comparisons and network analyses that include novel archaeal single-cell amplified genomes (SAGs) Could potentially help to further investigate the metabolic gene repertoire of the archaeal ancestor.

Whereas the origin of methanogenic pathways that include a multitude of specific genes and cofactors is not fully resolved yet [72], it appears that several later emerging euryarchaeal lineages have lost their methanogenic lifestyles. Thus, as already noted more than a decade ago, methanogens comprise a paraphyletic group separated by nonmethanogenic euryarchaeal lineages such as the *Thermoplasmatales*, Haloarchaeota, and *Archaeoglobales* [73]. Interestingly, a novel methanogenic archaeal lineage has been described recently that is distantly affiliated with cultivated *Thermoplasmatales* including *Aciduliprofundum* sp. [74, 75]. This suggests that the last common ancestor of *Thermoplasmatales* was a methanogen and the capability to reduce methane has been independently lost several times along some branches within this group [76] or, albeit less likely, that some lineages within the *Thermoplasmatales* have regained genes for methane production.

A single acquisition of a plethora of genes (>1000) from a bacterial donor has recently been put forward as explanation for the transition from a methanogenic ancestor to aerobic heterotrophic Haloarchaeota [77]. A possible driving force for this massive gene transfer might have been a syntrophic relationship between a methanogenic recipient and a bacterial donor. However, the exact donor lineage could not be determined: the acquired genes bear conflicting phylogenetic signals, supposedly due to prevalent gene transfers between different bacterial species. So far, the origin of alternative energy metabolisms in other non-methanogenic euryarchaeal lineages that evolved from methanogenic ancestors has not been addressed properly. However, comparative genomics suggests that several of these lineages have retained specific genes that trace back to the methanogenic nature of their ancestor (e.g., *Archaeoglobus*) [78] and might point to a rather transient transition.

3. Phylogeny of New Archaeal Phyla and Lineages

In recent years, several new archaeal lineages have been identified and subjected to whole genome or metagenomic

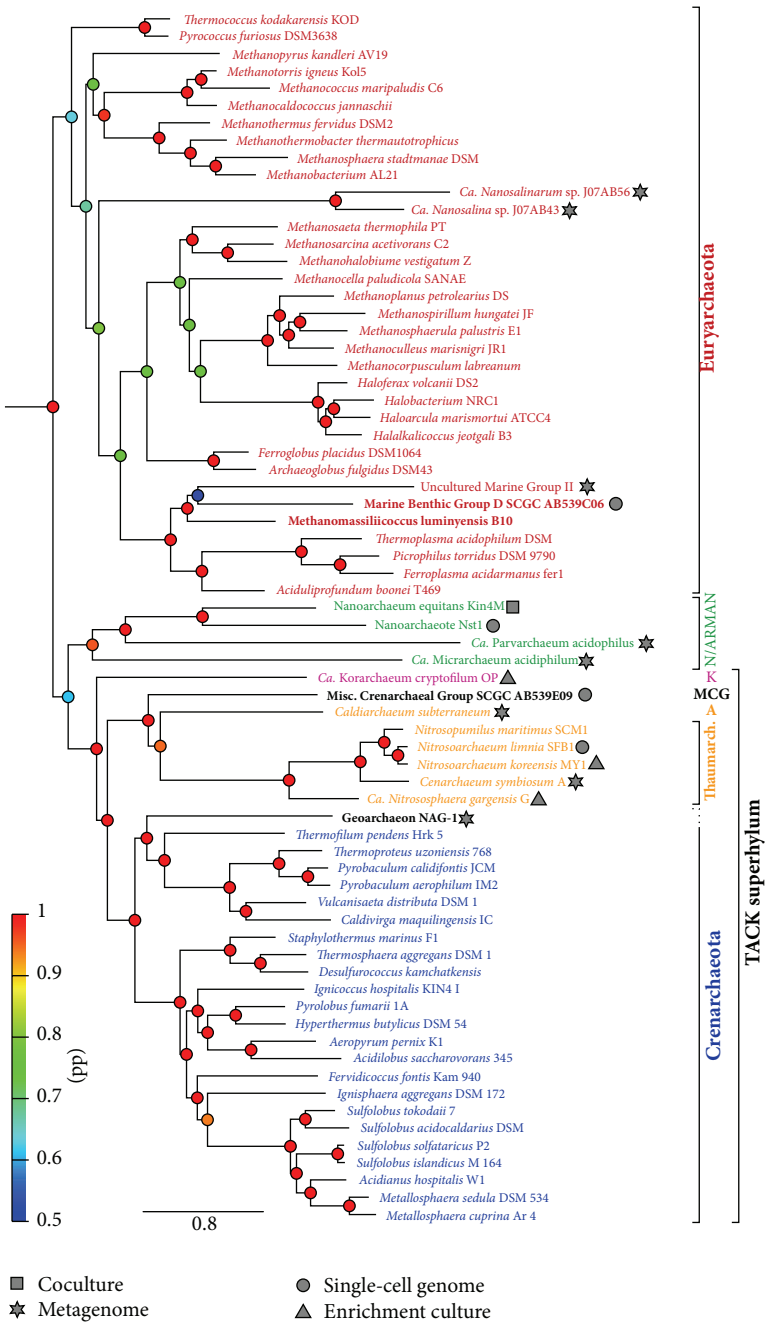

Figure 1: Bayesian phylogeny of 80 representative archaeal species. BLAST databases containing the proteome of 6 new archaeal genomes were retrieved from NCBI (in bold font on the tree): *Methanomassiliicoccus luminyensis* B10 (acc. no. CAJE01), MCG SCGC AB539E09 (acc. no. ALXK01), Marine Benthic Group D (MBGD) SCGC AB539N05, AB539C06, and AB540F20 (acc. no. ALXL01, AOSH01, and AOSI01, resp.). Protein sequence alignment from the 57 clusters in the discFilter 15 p dataset from [70] for which eukaryotes were removed were used as an input to psi-blast, with the six new proteomes as a database. Orthologs were retrieved as in [70]. For the three MBGD strains, one composite set of orthologs was constituted by using the most complete one (AB539C06) whenever possible and complementing with sequences from the other two if available. Orthologous genes selection, alignment, trimming, and concatenation were performed as in [70] resulting in a 15,069 amino-acid alignment. Four chains of Bayesian phylogenies were run with Phylobayes [71], under the CAT-Poisson model, running for approximately 10000 generations and discarding half as a burn-in. The tree was rooted with bacteria. Posterior probabilities (pp) are represented by colored dots on the nodes, with support values coloured according to the depicted heat-map colour scheme. The scale represents the number of substitutions per site. Species are colored according to the following: red, Euryarchaeota; green, Nanoarchaeota (N) and ARMAN; pink, Korarchaeota (K); black, Misc. Crenarchaeal Group (MCG); orange, Thaumarchaeota and Aigarchaeota (A); blue, Crenarchaeota. The DNA collection method, if different from pure culture, is indicated by a symbol next to the organism name: square, coculture; star, metagenome; circle, single-cell genome; triangle, enrichment culture.

sequencing. Based on phylogenetic analyses of available genomic data, some of these lineages have been proposed to represent novel archaeal phyla. Yet, some of these claims have been challenged or falsified in follow-up studies. Below, we give an overview of several such examples.

The candidate phylum Nanoarchaeota has initially been proposed on basis of the extremely divergent 16S rRNA sequence of the small parasitic cells of *Nanoarchaeum equitans* growing attached to the cell surface of *Ignicococcus hospitalis* [41]. Several subsequent and more comprehensive phylogenetic analyses as well as the finding of potentially ancestral genomic features (e.g., split tRNA genes) have provided support for the initial assignment of this tiny archaeal cells to a separate ancient archaeal phylum [65, 79, 80]. Yet, in contrast, other phylogenetic and comparative analyses testing the taxonomic position of *N. equitans* have suggested that Nanoarchaeota might rather represent a fast-evolving euryarchaeal lineage related to *Thermococcales* [81]. Genomic data from additional "nanosized" archaea (*Ca.* Parvarchaeum acidophilus ARMAN-4 and *Ca.* Micrarchaeum acidiphilum ARMAN-2) [82] as well as of a novel deep-branching member of Nanoarchaeota (Nst1) [83] have enabled a revision of phylogenetic reconstructions and genome comparisons. Although some of these analyses suggest that Nanoarchaeota and *Ca.* Parvarchaeum acidophilus are monophyletic, the placement of these groups in the archaeal tree remains unclear and is strongly dependent on dataset and phylogenetic methods used [64, 83]. For example, in our phylogenetic reconstructions Nanoarchaeota (including ARMAN-lineages) represent a sister clade of the TACK superphylum (Figure 1), although the support for this clade is low. In a recent study by Rinke et al. [33], the Nanoarchaeota (including all ARMAN strains) were grouped together in the newly proposed superphylum DPANN with two novel groups, DSEG and pMC2A384 (designated "Aenigmarchaeota" and "Diapherotrites", resp.), as well as the Nanohaloarchaea (see also below). Given that the phylogenetic methods employed by Rinke and coworkers do not accommodate rate heterogeneity across taxa, the proposed grouping of Nanoarchaeota with these archaeal clades has to be taken with care and the exact position of Nanoarchaeota still remains an unresolved question.

The Nanohaloarchaea represent yet another archaeal lineage comprised of small cells and with unresolved phylogenetic position. Based on both 16S rRNA gene and concatenated ribosomal protein phylogenies, this group was suggested to comprise a deep lineage of Haloarchaeota [84]. However, only euryarchaeal sequences were included in these maximum-likelihood (ML) analyses. Depending on the phylogenetic method and evolutionary model used, we obtained contradictory results for the phylogenetic position of this group. Whereas ML analyses tend to recover Nanohaloarchaea as earliest branching archaeal lineage (e.g., see above), a phylogenetic reconstruction using Bayesian methods (and the CAT model [71]) place this lineage within Euryarchaeota, but the exact position could not be resolved with high confidence (Figure 1). Results obtained with Bayesian methods using the CAT model might provide a better approximation of the position of Nanohaloarchaea, as this model accounts for rate variations across sites. As such, the early divergence of Nanohaloarchaea that is observed in ML-based methods is likely caused by LBA artifacts. However, novel phylogenetic analyses including the improved archaeal taxon sampling of Rinke et al. suggest that Nanohaloarchaea form a distinct lineage within the proposed superphylum DPANN and are not closely related to Euryarchaeota [33].

It will be interesting to further address the position of these organisms in the archaeal tree to be able to elucidate whether the adaptation to halophily has evolved only once in archaea or is due to convergence in Halo- and Nanohaloarchaea. The latter has received initial support from comparative genome analyses, which have revealed that each of these two archaeal groups seems to harbor diverse unique features including distinctive amino acid compositions to accommodate high salt conditions [84]. It might also be of value to address the effect of these novel genome sequences on the results obtained in the analysis of Nelson-Sathi et al. studying the origin of Haloarchaeota from a methanogenic ancestor [77].

Another novel archaeal phylum comprises the abundant and ecologically important ammonia-oxidizing archaea (AOA). On the basis of comparative genomics and phylogenetic analyses based on concatenated ribosomal proteins that were rooted with eukaryotes, Brochier-Armanet and coworkers proposed that "mesophilic crenarchaeota" constitute the novel deep branching archaeal phylum Thaumarchaeota [42, 85]. Additional comprehensive phylogenetic analyses including additional members of this group, as well as the discovery of a distinctive set of informational processing genes involved in replication, transcription, and translation as well as DNA repair and cell division machineries, have provided further support for the independent status of the Thaumarchaeota [44]. For example, in contrast to Crenarchaeota, Thaumarchaeota share several characteristics with Euryarchaeota and Korarchaeota including the presence of DNA polymerase D, histones, and cell division protein FtsZ. Furthermore, they contain putative "ancestral" features absent from Cren- or Euryarchaeota but common in Bacteria and eukaryotes (e.g., presence of an unsplit gene encoding DNA polymerase subunit A, and toposimerase IB as well as the absence of ribosomal protein LXa) [20, 44, 85]. The distinct nature of Thaumarchaeota has been accepted by many authors [45, 46, 86, 87] although the taxonomic borders of this phylum are still difficult to delineate and might only be resolved when genomes of uncultivated early branching lineages are made available. The early emergence of Thaumarchaeota in these phylogenetic reconstructions using eukaryotes as outgroup was initially assumed to indicate the ancient nature of this phylum [42, 44]. However, several recent phylogenetic analyses have recovered a monophyletic group of Thaum-, Aig-, Cren, Korarchaeota, and eukaryotes (with varying relationships in between these groups) to the exclusion of Euryarchaeota, which indicates that eukaryotes emerge from within the Archaea [55, 56, 88]. Thus, eukaryotes cannot be used as valid outgroup for the rooting of archaeal phylogenies [54].

Another lineage that emerges as a separate branch in the archaeal tree is comprised of the so-called Hot Water

Crenarchaeotic Group I (HWCG I), members of which have been detected in diverse hydrothermal environments but have not yet been cultivated [89, 90]. Until recently, the sole representative with a sequenced genome in this group was *Ca.* Caldiarchaeum subterraneum, whose composite genome has been obtained from a metagenomic library of a microbial mat in a subsurface geothermal water stream [45]. The investigation of its genome sequence has revealed the presence of components of the eukaryotic ubiquitin-like protein modifier system previously not detected in archaea or bacteria. This unique trait, as well as comparative genomics and phylogenetic analyses of concatenated protein sequences, suggested that this organism and other members of HWCG I might constitute a novel phylum (Aigarchaeota), distinct from both Thaum- and Crenarchaeota [45]. However, due to the presence of a set of informational processing genes most similar to Thaumarchaeota [45] and the highly supported monophyletic grouping of these two lineages in diverse phylogenetic analyses (e.g., see Figure 1), the separation of Thaum- and Aigarchaeota into two distinct phyla is still debated [45, 55, 56, 64, 91, 92].

Uncultivated archaea belonging to the so-called Miscellaneous Crenarchaeotal Group (MCG) (e.g., [39]) have been suggested to represent additional members of Aigarchaeota [55]. Recently, the first single-cell genome of a member of this group has been obtained and phylogenetic analyses of concatenated conserved single copy genes placed the MCG-archaeon as a lineage in between Thaum- and Aigarchaeota [97]. However, our analyses rather suggest that MCG emerges prior to the Thaum/Aigarchaeota (Figure 1). The availability of additional genome sequences of members of this group as well as the comparison of informational processing marker genes [44] of MCG-archaea with other available archaeal genomes might help both to resolve their phylogenetic placement and to determine whether MCG-archaea comprise a separate archaeal phylum [39].

Geoarchaeota represents yet another recently proposed archaeal phylum, which is proposed to emerge as a basal lineage of Crenarchaeota and includes the so-called novel archaeal group I (NAG-1) detected in acidic ferric iron mats from Yellowstone National Park [46, 98]. NAG-1 organisms thrive in hot (60–78°C) acidic mats rich in iron and are suggested to grow heterotrophically from simple carbon compounds. Though not yet enriched in culture, nearly full-length genome sequences of members of this group have been obtained from a *de novo* metagenome assembly. The description of this lineage as a separate phylum was based on phylogenetic analyses of concatenated ribosomal proteins and 16S/23S rRNA genes as well as on its specific set of informational processing genes with features in common with either Crenarchaeota or Thaum- and Aigarchaeota [46]. However, our analyses, based on a larger dataset, place Geoarchaeota as an early branching lineage of the crenarchaeal order Thermoproteales (Figure 1). This observation is confirmed by Rinke et al., who sequenced six additional NAG-1-related strains [33]. Indeed, detailed phylogenetic analyses, as well as comparative assessment of the NAG-1 composite genome, seem to refute the phylum-level status of NAG-1 (Guy, L., Spang, A., Saw, J.H. and Ettema, T.J.G., unpublished observation).

4. Archaea and the Origin of Eukaryotes

The origin of the eukaryotes remains one of the major unanswered questions in modern biology, and archaea have recently reclaimed the spotlights in heated discussions entailing this enigmatic event. A central issue in this discussion entails the placement of the root within the Tree of Life, as it has a fundamental effect on any hypothesis on the origin of eukaryotes. Whereas diverse competing hypotheses have been put forward in the past, no consensus has been reached on this topic so far. For instance, several studies, including a recent network analysis, place the root between Archaea and Bacteria [99–104]. This view is in agreement with both the observed fundamental differences distinguishing the bacterial and archaeal domains as well as with the geological record. In contrast, studies that were based on transmission analyses or the distribution of indels in protein sequences suggested a rooting within the bacterial domain [105–107], whereas a root in the archaeal domain has been proposed based on analyses of protein folds or the evolution of the tRNA molecules [108, 109]. Yet other hypotheses state that LUCA was a eukaryotic-like organism [110, 111]. Certainly, in order to reach a consensus on this controversial discussion, additional data and analyses are needed. Bearing the uncertainty of the placement of the root in the Tree of Life in mind, we will present current hypotheses on the origin of eukaryotes below, by providing a short review on the most commonly proposed scenarios.

Even though a wide variety of incompatible theories have been suggested regarding the origin of the eukaryotic cell, three aspects are now largely accepted: (i) the last eukaryotic common ancestor (LECA) contained mitochondria, (ii) eukaryotic genomes are chimeric; whereas informational genes are of archaeal descent, many metabolic genes are derived from Bacteria, and (iii) eukaryotes complement a set of proteins not found in either Archaea or Bacteria, the eukaryotic signature proteins (ESPs). Beyond this, the picture becomes blurry. Currently, two major questions are of interest. What was the nature of the cell that was host in the mitochondrial endosymbiosis and when did cellular complexity evolve, before (complexity-first) or after (mitochondria-first) mitochondrial endosymbiosis? From this perspective, theories on eukaryogenesis can be divided into two categories. In the first scenario, the host was a protoeukaryote and complexity evolved first. This theory, often referred to as the "archezoa hypothesis" [112, 113], fits with the three domains tree of life model in which eukaryotes vertically evolved from the archaea-eukaryote common ancestor (Figure 2(a)). In the second scenario, the host was a prokaryote and the acquisition of the mitochondria likely triggered the evolution of cellular complexity. The latter are often referred to as "fusion" hypotheses [51, 58, 93, 95, 96] and these are generally incompatible with the classical three domains model. Rather, in these models, Bacteria and Archaea are the primary domains of life and eukaryotes a secondary, or derived, domain of life (Figure 2(b)). Theories that

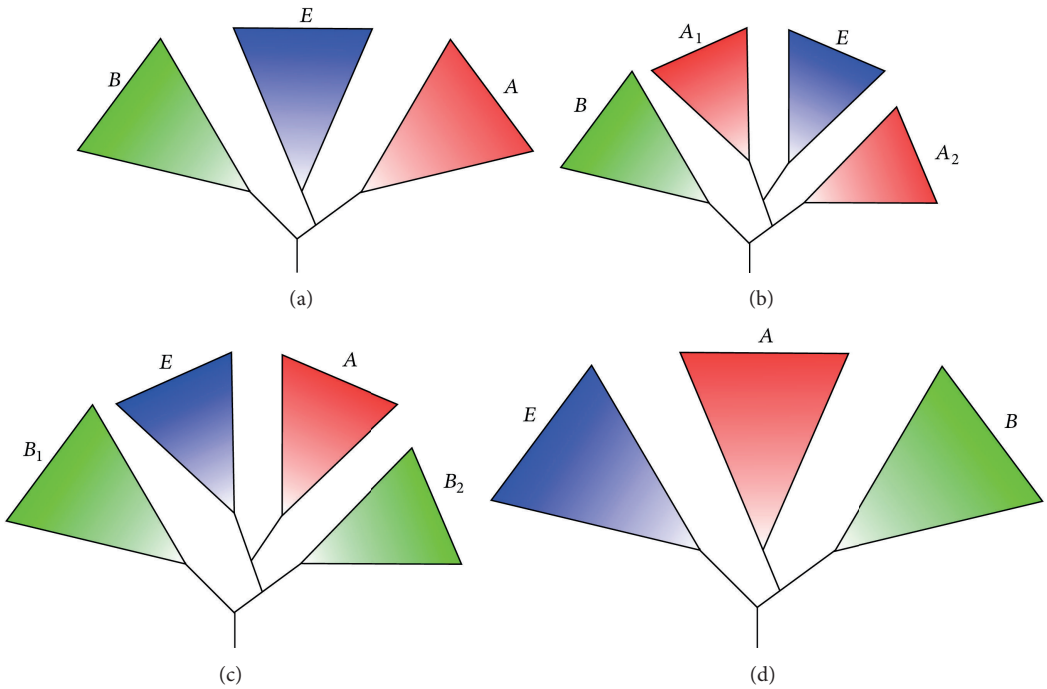

Figure 2: Overview of theories regarding the origin of the eukaryotic nuclear lineage. (a) The classical, Woesean three domains of life tree in which the nuclear lineage vertically evolved from the archaea-eukaryote common ancestor. (b) The fusion tree in which the nuclear lineage originated from the archaeal partner in the fusion. Depending on which fusion model, the archaeal parent's lineage (A_1) was either part of the euryarchaeota [SET [93], original syntrophy hypothesis [94], hydrogen hypothesis [95] or alternative syntrophy hypothesis [51]], the Crenarchaeota (eocyte hypothesis) [96], or the TACK superphylum (PhAT) [58]. "A_2" represents all archaea not directly affiliated with "A_1." (c) The neomuran tree in which the eukaryotic and archaeal lineage (combined referred to as "neomurans"), evolved vertically from ancestor shared with actinobacteria (B_2) as a result of the loss of bacterial-type cell wall (the neomuran revolution). B_1 represents all bacteria not directly affiliated with B_2. (d) The eukaryote-early tree, which suggests that the last common universal ancestor was more eukaryote-like than prokaryote-like.

fit neither of these categories exist as well. These include the neomuran hypothesis [114] (Figure 2(c)), the PVC hypothesis [115–118], virus-assisted eukaryogenesis [119–122], and a hypothesis suggesting a eukaryote-like universal common ancestor [110] (Figure 2(d)). In order to choose the correct category with high confidence, evidence is needed in the form of protoeukaryote intermediate lineage's descendants ("missing links"). Unfortunately, for either category, none has been found so far. Whereas the archezoa theory has lost much support ever since remnants of mitochondria were found in the previously thought archezoa (for review, see [123]), the fusion theory has slowly been gaining favor. Initially lightly supported by ribosomal structural features [124] and an 11-amino acid insertion in EF-1α/EF-Tu [125, 126] shared between eocytes (Crenarchaeota) and eukaryotes to the exclusion of other prokaryotes, it has now received strong support from phylogenomic [55, 56, 70, 88, 127, 128] and gene similarity network analyses [129]. In addition, a large number of ESPs has been found in Archaea, in particular within the recently proposed TACK superphylum [55]. Examples include actin [53, 130], tubulin [28], H3/H4-type histones [55], ESCRT-III [24, 25, 131], and components of the ubiquitin modifier system [45]. Fusion models can be subdivided based upon the nature of the end-product of the "fusion". In amitochondriate models the symbiosis results in a eukaryotic progenitor lacking mitochondria. They are similar to the archezoa theory in the sense that the origin of eukaryotes and the origin of mitochondria are separate events. These include the serial endosymbiosis theory (SET) [93], the original syntrophy hypothesis [94], and the eocyte hypothesis [96]. In mitochondriate models, the end product is a eukaryotic progenitor containing mitochondria. Here, the origin of eukaryotes and mitochondria are one and the same. These include the hydrogen hypothesis [95], the alternative syntrophy hypothesis [51] and the recently proposed phagocytosing archaeon theory [53, 58]. With exception of the eocyte hypothesis, all fusion theories suggest an archaeal host. Based on extensive, in-depth phylogenomic studies, the archaeal host most likely emerged from within the TACK superphylum [55, 56, 70]. Interestingly, out of all TACK phyla, a sister relationship between the Korarchaeota and eukaryotes was retrieved with significant phylogenetic support [56, 70]. Even though this placement could be a taxon sampling artifact (Korarchaeota are represented by a single, deep rooting taxon), it could also indicate that eukaryotes are affiliated with an unidentified lineage distantly related to Korarchaeota. Genomically unexplored lineages such as DSAG (Deep Sea Archaea Group), MHVG (Marine

Hydrothermal Vent Group), and AAG (Ancient Archaea Group) are likely candidates [55, 70].

5. Genomic Assessment and Taxonomic Classification of Archaeal Diversity

Recent progress in genomic sequencing technologies and cultivation-independent methods has started to unearth a plethora of novel, uncultivated archaeal lineages. The availability of such genomic data has revealed several important insights into the diversity, ecological relevance, metabolic capacity, and the origin and evolution of the archaeal domain of life. Several new archaeal lineages have been obtained by means of metagenomics approaches, such as sequencing of enrichment cultures or environmental samples. Examples of the former include the first korarchaeal genome [43] and several of the available thaumarchaeal genomes (e.g. [132, 133]). Archaeal genomes that have been retrieved from metagenomic datasets include the first thaumarchaeal genome (*Ca.* Cenarchaeum symbiosum [134]), the genome of the proposed Aigarchaeon *Ca.* C. subterraneum [45], the proposed Geoarchaeon NAG-1 [46], representatives of the Nanohaloarchaea [84], several ARMAN lineages that were part of an acid mine drainage microbial community [82], and a genome derived from a representative of the uncultivated marine group II euryarchaeota [135] (Figure 1). More recently, a number of studies have employed single cell genomic approaches to probe the genetic diversity of uncultivated archaea. For example, Lloyd and coworkers have reported the first genomic data of a representative of the Miscellaneous Crenarchaeal Group (MCG) and of members of the Marine Benthic Group D that were isolated from marine sediments and speculate that these lineages are involved in the degradation of detrital proteins [97] (Figure 1). Another large scale study that aimed at uncovering the coding potential of so-called "microbial dark matter" using single cell genomics approaches reported several genome sequences of cells that potentially represented novel phylum-level archaeal lineages, including the members of the uncultured DSEG and pMC2A384 clades, designated Aenigmarchaeota and Diapherotrites, respectively [33]. A combination of single cell genomics and metagenomics has been used to sequence the genome of the thaumarchaeon *Ca.* Nitrosoarchaeum limnia SFB1 [136].

Obviously, single cell and metagenomics-oriented projects will continue to probe the existing archaeal diversity during the coming years, and most likely, the availability of genomic data will reveal interesting insights into novel characteristics and the diversity within the Third Domain of life. In addition, the availability of such genomic data is likely to trigger discussions regarding the higher-order taxonomic classification of the major archaeal lineages. To many (micro-)biologists, it would appear that the archaeal domain is far less diverse than the bacterial domain. A reason for this could be, for instance, the discrepancy in assigned or proposed phyla, which ranges from a handful in Archaea, to well over a hundred in Bacteria. But is it really fair to say that the bacterial domain of life is more diverse than that of the Archaea? Whereas bacterial phyla generally have been assigned based on the diversity of the 16S rDNA gene sequence, archaeal taxonomy is largely founded on historic grounds, that is, adhering to the classical Cren-Euryarchaeota dichotomy (sensu Woese [30]). Only during the past decade, a handful of additional archaeal phyla have been proposed based on genome sequencing, such as the Nano-, Kor-, and Thaumarchaeota and a few other lineages that may or may not represent phylum-level archaeal clades (also see above). Yet, the majority of archaeal species that have been sequenced in recent years have been assigned to the phyla Cren- or Euryarchaeota, each of which now comprise genetically distinct groups, which differ in terms of metabolic capacity, lifestyle, and environmental distribution. In light of this and of the abovementioned "superficial" imbalance in bacterial versus archaeal diversity, one could argue that a revision of archaeal higher-order taxonomy is in place. The suggestion to bring order into archaeal systematics was recently put forward [92], but thus far, a framework as to how novel phyla and/or superphyla should be defined is debated. Nevertheless, to be able to fully appreciate the overall archaeal diversity and compare it to the diversity observed within the bacterial domain of life, a reappraisal of the archaeal taxonomy, whether it will be at the level of rRNA genes, large datasets of concatenated protein sequences, genome content, or gene networks analyses, seems to be a *conditio sine qua non*.

Abbreviations

ANME: Anaerobic methanotrophic archaea
LACA: Last archaeal common ancestor
LBA: Long-branch attraction
MCG: Miscellaneous Crenarchaeotal Group
ML: Maximum-likelihood
SAGs: Single-cell amplified genomes.

Conflicts of Interests

The authors declare that they do not have a direct financial relation with the trademarks mentioned in the paper that might lead to a conflict of interests for the authors.

Acknowledgments

The work in Ettema laboratory is supported by the Swedish Research Council (Grant no. 621-2009-4813), by the European Research Council (ERC) (Grant no. 310039-PUZZLE_CELL), by a Marie Curie European Reintegration Grant (ERG) (Grant no. 268259-RICKOCHET), and by the Carl Tryggers Stiftelse (Grant no. CTS11:127).

References

[1] C. R. Woese and G. E. Fox, "Phylogenetic structure of the prokaryotic domain: the primary kingdoms," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 74, no. 11, pp. 5088–5090, 1977.

[2] C. R. Woese, "The birth of the Archaea: a personal retrospective," in *Archaea*, R. A. Garrett and H.-P. Klenk, Eds., Blackwell Publishing, 2007.

[3] T. A. Langworthy, P. F. Smith, and W. R. Mayberry, "Lipids of Thermoplasma acidophilum," *Journal of Bacteriology*, vol. 112, no. 3, pp. 1193–1200, 1972.

[4] T. A. Langworthy, W. R. Mayberry, and P. F. Smith, "Long chain glycerol diether and polyol dialkyl glycerol triether lipids of Sulfolobus acidocaldarius," *Journal of Bacteriology*, vol. 119, no. 1, pp. 106–116, 1974.

[5] O. Kandler and H. Hippe, "Lack of peptidoglycan in the cell walls of Methanosarcina barkeri," *Archives of Microbiology*, vol. 113, no. 1-2, pp. 57–60, 1977.

[6] W. Zillig, K. O. Stetter, and D. Janekovic, "DNA-dependent RNA polymerase from the archaebacterium *Sulfolobus acidocaldarius*," *European Journal of Biochemistry*, vol. 96, no. 3, pp. 597–604, 1979.

[7] J. Huet, R. Schnabel, A. Sentenac, and W. Zillig, "Archaebacteria and eukaryotes possess DNA-dependent RNA polymerases of a common type," *EMBO Journal*, vol. 2, no. 8, pp. 1291–1294, 1983.

[8] J. A. Lake, "Ribosome evolution: the structural bases of protein synthesis in archaebacteria, eubacteria, and eukaryotes," *Progress in Nucleic Acid Research and Molecular Biology*, vol. 30, pp. 163–194, 1983.

[9] E. Henderson, M. Oakes, M. W. Clark, J. A. Lake, A. T. Matheson, and W. Zillig, "A new ribosome structure," *Science*, vol. 225, no. 4661, pp. 510–512, 1984.

[10] W. Zillig, R. Schnabel, and K. O. Stetter, "Archaebacteria and the origin of the eukaryotic cytoplasm," *Current Topics in Microbiology and Immunology*, vol. 114, pp. 1–18, 1985.

[11] G. Puhler, H. Leffers, F. Gropp et al., "Archaebacterial DNA-dependent RNA polymerase testify to the evolution of the eukaryotic nuclear genome," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 86, no. 12, pp. 4569–4573, 1989.

[12] K. Sandman, J. A. Krzycki, B. Dobrinski, B. Lurz, and J. N. Reeve, "HMf, a DNA-binding protein isolated from the hyperthermophilic archaeon Methanothermus fervidus, is most closely related to histones," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 87, no. 15, pp. 5788–5791, 1990.

[13] D. Langer, J. Hain, P. Thuriaux, and W. Zillig, "Transcription in archaea: similarity to that in eucarya," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 92, no. 13, pp. 5768–5772, 1995.

[14] O. Kandler and H. König, "Cell wall polymers in Archaea (Archaebacteria)," *Cellular and Molecular Life Sciences*, vol. 54, no. 4, pp. 305–308, 1998.

[15] G. J. Olsen and C. R. Woese, "Lessons from an Archaeal genome: what are we learning from Methanococcus jannaschii?" *Trends in Genetics*, vol. 12, no. 10, pp. 377–379, 1996.

[16] O. Lecompte, R. Ripp, J.-C. Thierry, D. Moras, and O. Poch, "Comparative analysis of ribosomal proteins in complete genomes: an example of reductive evolution at the domain scale," *Nucleic Acids Research*, vol. 30, no. 24, pp. 5382–5390, 2002.

[17] E. R. Barry and S. D. Bell, "DNA replication in the archaea," *Microbiology and Molecular Biology Reviews*, vol. 70, no. 4, pp. 876–887, 2006.

[18] S. Gribaldo and C. Brochier-Armanet, "The origin and evolution of Archaea: a state of the art," *Philosophical Transactions of the Royal Society B*, vol. 361, no. 1470, pp. 1007–1022, 2006.

[19] L. L. Grochowski, H. Xu, and R. H. White, "Methanocaldococcus jannaschii uses a modified mevalonate pathway for biosynthesis of isopentenyl diphosphate," *Journal of Bacteriology*, vol. 188, no. 9, pp. 3192–3198, 2006.

[20] M. Kwapisz, F. Beckouët, and P. Thuriaux, "Early evolution of eukaryotic DNA-dependent RNA polymerases," *Trends in Genetics*, vol. 24, no. 5, pp. 211–215, 2008.

[21] F. Werner and D. Grohmann, "Evolution of multisubunit RNA polymerases in the three domains of life," *Nature Reviews Microbiology*, vol. 9, no. 2, pp. 85–98, 2011.

[22] S. L. Pereira and J. N. Reeve, "Histones and nucleosomes in Archaea and Eukarya: a comparative analysis," *Extremophiles*, vol. 2, no. 3, pp. 141–148, 1998.

[23] J. N. Reeve, K. A. Bailey, W.-T. Li, F. Marc, K. Sandman, and D. J. Soares, "Archaeal histones: structures, stability and DNA binding," *Biochemical Society Transactions*, vol. 32, no. 2, pp. 227–230, 2004.

[24] A.-C. Lindås, E. A. Karlsson, M. T. Lindgren, T. J. G. Ettema, and R. Bernander, "A unique cell division machinery in the Archaea," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 48, pp. 18942–18946, 2008.

[25] R. Y. Samson, T. Obita, S. M. Freund, R. L. Williams, and S. D. Bell, "A role for the ESCRT system in cell division in archaea," *Science*, vol. 322, no. 5908, pp. 1710–1713, 2008.

[26] R. Bernander, A. E. Lind, and T. J. G. Ettema, "An archaeal origin for the actin cytoskeleton," *Communicative & Integrative Biology*, vol. 4, pp. 664–667, 2011.

[27] R. Ammar, D. Torti, K. Tsui et al., "Chromatin is an ancient innovation conserved between Archaea and Eukarya," *Elife*, vol. 1, Article ID e00078, 2012.

[28] N. Yutin and E. V. Koonin, "Archaeal origin of tubulin," *Biology Direct*, vol. 7, article 10, 2012.

[29] M. C. Rivera, R. Jain, J. E. Moore, and J. A. Lake, "Genomic evidence for two functionally distinct gene classes," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 95, no. 11, pp. 6239–6244, 1998.

[30] C. R. Woese, O. Kandler, and M. L. Wheelis, "Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 87, no. 12, pp. 4576–4579, 1990.

[31] C. Schleper, G. Jurgens, and M. Jonuscheit, "Genomic studies of uncultivated archaea," *Nature Reviews Microbiology*, vol. 3, no. 6, pp. 479–488, 2005.

[32] A. Teske and K. B. Sørensen, "Uncultured archaea in deep marine subsurface sediments: have we caught them all?" *The ISME Journal*, vol. 2, no. 1, pp. 3–18, 2008.

[33] C. Rinke, P. Schwientek, A. Sczyrba et al., "Insights into the phylogeny and coding potential of microbial dark matter," *Nature*, 2013.

[34] J. O. McInerney, M. Wilkinson, J. W. Patching, T. M. Embley, and R. Powell, "Recovery and phylogenetic analysis of novel archaeal rRNA sequences from a deep-sea deposit feeder," *Applied and Environmental Microbiology*, vol. 61, no. 4, pp. 1646–1648, 1995.

[35] E. F. Delong, "Everything in moderation: Archaea as 'non-extremophiles'," *Current Opinion in Genetics & Development*, vol. 8, no. 6, pp. 649–654, 1998.

[36] K. Takai and K. Horikoshi, "Genetic diversity of archaea in deep-sea hydrothermal vent environments," *Genetics*, vol. 152, no. 4, pp. 1285–1297, 1999.

[37] B. J. Baker, G. W. Tyson, R. I. Webb et al., "Lineages of acidophilic archaea revealed by community genomic analysis," *Science*, vol. 314, no. 5807, pp. 1933–1935, 2006.

[38] B. Chaban, S. Y. M. Ng, and K. F. Jarrell, "Archaeal habitats—from the extreme to the ordinary," *Canadian Journal of Microbiology*, vol. 52, no. 2, pp. 73–116, 2006.

[39] K. Kubo, K. G. Lloyd, J. F Biddle, R. Amann, A. Teske, and K. Knittel, "Archaea of the Miscellaneous Crenarchaeotal Group are abundant, diverse and widespread in marine sediments," *The ISME Journal*, vol. 6, pp. 1949–1965, 2012.

[40] S. M. Barns, C. F. Delwiche, J. D. Palmer, and N. R. Pace, "Perspectives on archaeal diversity, thermophily and monophyly from environmental rRNA sequences," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 93, no. 17, pp. 9188–9193, 1996.

[41] H. Huber, M. J. Hohn, R. Rachel, T. Fuchs, V. C. Wimmer, and K. O. Stetter, "A new phylum of Archaea represented by a nanosized hyperthermophilic symbiont," *Nature*, vol. 417, no. 6884, pp. 63–67, 2002.

[42] C. Brochier-Armanet, B. Boussau, S. Gribaldo, and P. Forterre, "Mesophilic crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota," *Nature Reviews Microbiology*, vol. 6, no. 3, pp. 245–252, 2008.

[43] J. G. Elkins, M. Podar, D. E. Graham et al., "A korarchaeal genome reveals insights into the evolution of the Archaea," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 23, pp. 8102–8107, 2008.

[44] A. Spang, R. Hatzenpichler, C. Brochier-Armanet et al., "Distinct gene set in two different lineages of ammonia-oxidizing archaea supports the phylum Thaumarchaeota," *Trends in Microbiology*, vol. 18, no. 8, pp. 331–340, 2010.

[45] T. Nunoura, Y. Takaki, J. Kakuta et al., "Insights into the evolution of Archaea and eukaryotic protein modifier systems revealed by the genome of a novel archaeal group," *Nucleic Acids Research*, vol. 39, no. 8, pp. 3204–3223, 2011.

[46] M. A. Kozubal, M. Romine, R. Jennings et al., "Geoarchaeota: a new candidate phylum in the Archaea from high-temperature acidic iron mats in Yellowstone National Park," *The ISME Journal*, vol. 7, pp. 622–634, 2013.

[47] P. Offre, A. Spang, and C. Schleper, "Archaea in biogeochemical cycles," *Annual Review of Microbiology*, vol. 67, pp. 437–457, 2013.

[48] M. Könneke, A. E. Bernhard, J. R. de la Torre, C. B. Walker, J. B. Waterbury, and D. A. Stahl, "Isolation of an autotrophic ammonia-oxidizing marine archaeon," *Nature*, vol. 437, no. 7058, pp. 543–546, 2005.

[49] A. H. Treusch, S. Leininger, A. Kietzin, S. C. Schuster, H.-P. Klenk, and C. Schleper, "Novel genes for nitrite reductase and Amo-related proteins indicate a role of uncultivated mesophilic crenarchaeota in nitrogen cycling," *Environmental Microbiology*, vol. 7, no. 12, pp. 1985–1995, 2005.

[50] K. Knittel and A. Boetius, "Anaerobic oxidation of methane: progress with an unknown process," *Annual Review of Microbiology*, vol. 63, pp. 311–334, 2009.

[51] P. López-Garćia and D. Moreira, "Metabolic symbiosis at the origin of eukaryotes," *Trends in Biochemical Sciences*, vol. 24, pp. 88–93, 1999.

[52] T. M. Embley and W. Martin, "Eukaryotic evolution, changes and challenges," *Nature*, vol. 440, no. 7084, pp. 623–630, 2006.

[53] N. Yutin, M. Y. Wolf, Y. I. Wolf, and E. V. Koonin, "The origins of phagocytosis and eukaryogenesis," *Biology Direct*, vol. 4, article 9, 2009.

[54] S. Gribaldo, A. M. Poole, V. Daubin, P. Forterre, and C. Brochier-Armanet, "The origin of eukaryotes and their relationship with the Archaea: are we at a phylogenomic impasse?" *Nature Reviews Microbiology*, vol. 8, no. 10, pp. 743–752, 2010.

[55] L. Guy and T. J. G. Ettema, "The archaeal "TACK" superphylum and the origin of eukaryotes," *Trends in Microbiology*, vol. 19, no. 12, pp. 580–587, 2011.

[56] T. A. Williams, P. G. Foster, T. M. W. Nye, C. J. Cox, and T. M. Embley, "A congruent phylogenomic signal places eukaryotes within the Archaea," *Proceedings of the Royal Society B*, vol. 279, no. 1749, pp. 4870–4879, 2012.

[57] Y. I. Wolf, K. S. Makarova, N. Yutin, and E. V. Koonin, "Updated clusters of orthologous genes for Archaea: a complex ancestor of the Archaea and the byways of horizontal gene transfer," *Biology Direct*, vol. 7, article 46, 2012.

[58] J. Martijn and J. G. Ettema Thijs, "From archaeon to eukaryote: the evolutionary dark ages of the eukaryotic cell," *Biochemical Society Transactions*, vol. 41, pp. 451–457, 2013.

[59] W. Martin, J. Baross, D. Kelley, and M. J. Russell, "Hydrothermal vents and the origin of life," *Nature Reviews Microbiology*, vol. 6, no. 11, pp. 805–814, 2008.

[60] N. Lane and W. F. Martin, "The origin of membrane bioenergetics," *Cell*, vol. 151, pp. 1406–1416, 2012.

[61] Y. Liu, L. L. Beer, and W. B. Whitman, "Methanogens: a window into ancient sulfur metabolism," *Trends in Microbiology*, vol. 20, no. 5, pp. 251–258, 2012.

[62] A. Poehlein, S. Schmidt, A.-K. Kaster et al., "An ancient pathway combining carbon dioxide fixation with the generation and utilization of a sodium ion gradient for ATP synthesis," *PLoS ONE*, vol. 7, no. 3, Article ID e33439, 2012.

[63] Y. Ueno, K. Yamada, N. Yoshida, S. Maruyama, and Y. Isozaki, "Evidence from fluid inclusions for microbial methanogenesis in the early Archaean era," *Nature*, vol. 440, no. 7083, pp. 516–519, 2006.

[64] C. Brochier-Armanet, P. Forterre, and S. Gribaldo, "Phylogeny and evolution of the Archaea: one hundred genomes later," *Current Opinion in Microbiology*, vol. 14, no. 3, pp. 274–281, 2011.

[65] N. Yutin, P. Puigbò, E. V. Koonin, and Y. I. Wolf, "Phylogenomics of prokaryotic ribosomal proteins," *PLoS ONE*, vol. 7, no. 5, Article ID e36972, 2012.

[66] M. Groussin and M. Gouy, "Adaptation to environmental temperature is a major determinant of molecular evolutionary rates in archaea," *Molecular Biology and Evolution*, vol. 28, no. 9, pp. 2661–2674, 2011.

[67] C. Brochier, P. Forterre, and S. Gribaldo, "Archaeal phylogeny based on proteins of the transcription and translation machineries: tackling the Methanopyrus kandleri paradox," *Genome Biology*, vol. 5, no. 3, p. R17, 2004.

[68] K. S. Makarova, A. V. Sorokin, P. S. Novichkov, Y. I. Wolf, and E. V. Koonin, "Clusters of orthologous genes for 41 archaeal genomes and implications for evolutionary genomics of archaea," *Biology Direct*, vol. 2, article 33, 2007.

[69] S. Kelly, B. Wickstead, and K. Gull, "Archaeal phylogenomics provides evidence in support of a methanogenic origin of the Archaea and a thaumarchaeal origin for the eukaryotes," *Proceedings of the Royal Society B*, vol. 278, no. 1708, pp. 1009–1018, 2011.

[70] L. Guy, J. H. Saw, and T. J. G. Ettema, "The archaeal legacy of eukaryotes: a phylogenomic perspective," *Cold Spring Harbor Perspectives in Biology*. In press.

[71] N. Lartillot, T. Lepage, and S. Blanquart, "PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating," *Bioinformatics*, vol. 25, no. 17, pp. 2286–2288, 2009.

[72] J. G. Ferry, "How to make a living by exhaling methane," *Annual Review of Microbiology*, vol. 64, pp. 453–473, 2010.

[73] O. Matte-Tailliez, C. Brochier, P. Forterre, and H. Philippe, "Archaeal phylogeny based on ribosomal proteins," *Molecular Biology and Evolution*, vol. 19, no. 5, pp. 631–639, 2002.

[74] B. Dridi, M.-L. Fardeau, B. Ollivier, D. Raoult, and M. Drancourt, "Methanomassiliicoccus luminyensis gen. nov., sp. nov., a methanogenic archaeon isolated from human faeces," *International Journal of Systematic and Evolutionary Microbiology*, vol. 62, pp. 1902–1907, 2012.

[75] K. Paul, J. O. Nonoh, L. Mikulski, and A. Brune, "'Methanoplasmatales,' Thermoplasmatales-related archaea in termite guts and other environments, are the seventh order of methanogens," *Applied and Environmental Microbiology*, vol. 78, pp. 8245–8253, 2012.

[76] G. Borrel, P. W. O'Toole, H. M. B. Harris, P. Peyret, J.-F. Brugère, and S. Gribaldo, "Phylogenomic data support a seventh order of methylotrophic methanogens and provide insights into the evolution of methanogenesis," *Genome Biology and Evolution*, 2013.

[77] S. Nelson-Sathi, T. Dagan, G. Landan et al., "Acquisition of 1, 000 eubacterial genes physiologically transformed a methanogen at the origin of Haloarchaea," *Proceedings of the National Academy of Sciences*, vol. 109, pp. 20537–20542, 2012.

[78] B. Schworer, J. Breitung, A. R. Klein, K. O. Stetter, and R. K. Thauer, "Formylmethanofuran: tetrahydromethanopterin formyltransferase and N5,N10-methylenetetrahydromethanopterin dehydrogenase from the sulfate-reducing Archaeoglobus fulgidus: similarities with the enzymes from methanogenic Archaea," *Archives of Microbiology*, vol. 159, no. 3, pp. 225–232, 1993.

[79] E. Waters, M. J. Hohn, I. Ahel et al., "The genome of Nanoarchaeum equitans: insights into early archaeal evolution and derived parasitism," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 22, pp. 12984–12988, 2003.

[80] M. Di Giulio, "Formal proof that the split genes of tRNAs of nanoarchaeum equitans are an ancestral character," *Journal of Molecular Evolution*, vol. 69, no. 5, pp. 505–511, 2009.

[81] C. Brochier, S. Gribaldo, Y. Zivanovic, F. Confalonieri, and P. Forterre, "Nanoarchaea: representatives of a novel archaeal phylum or a fast-evolving euryarchaeal lineage related to Thermococcales?" *Genome Biology*, vol. 6, no. 5, p. R42, 2005.

[82] B. J. Baker, L. R. Comolli, G. J. Dick et al., "Enigmatic, ultrasmall, uncultivated Archaea," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, no. 19, pp. 8806–8811, 2010.

[83] M. Podar, K. S. Makarova, D. E. Graham, Y. I. Wolf, E. V. Koonin, and A.-L. Reysenbach, "Insights into archaeal evolution and symbiosis from the genomes of a nanoarchaeon and its inferred crenarchaeal host from Obsidian Pool, Yellowstone National Park," *Biology Direct*, vol. 8, article 9, 2013.

[84] P. Narasingarao, S. Podell, J. A. Ugalde et al., "De novo metagenomic assembly reveals abundant novel major lineage of Archaea in hypersaline microbial communities," *The ISME Journal*, vol. 6, no. 1, pp. 81–93, 2012.

[85] C. Brochier-Armanet, S. Gribaldo, and P. Forterre, "A DNA topoisomerase IB in Thaumarchaeota testifies for the presence of this enzyme in the last common ancestor of Archaea and Eucarya," *Biology Direct*, vol. 3, article 54, 2008.

[86] R. S. Gupta and A. Shami, "Molecular signatures for the Crenarchaeota and the Thaumarchaeota," *Antonie van Leeuwenhoek*, vol. 99, no. 2, pp. 133–157, 2011.

[87] M. Pester, C. Schleper, and M. Wagner, "The Thaumarchaeota: an emerging view of their phylogeny and ecophysiology," *Current Opinion in Microbiology*, vol. 14, no. 3, pp. 300–306, 2011.

[88] C. J. Cox, P. G. Foster, R. P. Hirt, S. R. Harris, and T. M. Embley, "The archaebacterial origin of eukaryotes," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 51, pp. 20356–20361, 2008.

[89] T. Nunoura, H. Hirayama, H. Takami et al., "Genetic and functional properties of uncultivated thermophilic crenarchaeotes from a subsurface gold mine as revealed by analysis of genome fragments," *Environmental Microbiology*, vol. 7, no. 12, pp. 1967–1984, 2005.

[90] T. Nunoura, H. Oida, M. Nakaseama et al., "Archaeal diversity and distribution along thermal and geochemical gradients in hydrothermal sediments at the yonaguni knoll IV hydrothermal field in the Southern Okinawa Trough," *Applied and Environmental Microbiology*, vol. 76, no. 4, pp. 1198–1211, 2010.

[91] C. Brochier-Armanet, S. Gribaldo, and P. Forterre, "Spotlight on the Thaumarchaeota," *The ISME Journal*, vol. 6, no. 2, pp. 227–230, 2012.

[92] S. Gribaldo and C. Brochier-Armanet, "Time for order in microbial systematics," *Trends in Microbiology*, vol. 20, no. 5, pp. 209–210, 2012.

[93] L. Margulis, M. Chapman, R. Guerrero, and J. Hall, "The last eukaryotic common ancestor (LECA): acquisition of cytoskeletal motility from aerotolerant spirochetes in the Proterozoic Eon," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 35, pp. 13080–13085, 2006.

[94] D. Moreira and P. López-García, "Symbiosis between methanogenic archaea and δ-proteobacteria as the origin of eukaryotes: the syntrophic hypothesis," *Journal of Molecular Evolution*, vol. 47, no. 5, pp. 517–530, 1998.

[95] W. Martin and M. Müller, "The hydrogen hypothesis for the first eukaryote," *Nature*, vol. 392, no. 6671, pp. 37–41, 1998.

[96] J. A. Lake, "Origin of the eukaryotic nucleus determined by rate-invariant analysis of rRNA sequences," *Nature*, vol. 331, no. 6152, pp. 184–186, 1988.

[97] K. G. Lloyd, L. Schreiber, D. G. Petersen et al., "Predominant archaea in marine sediments degrade detrital proteins," *Nature*, vol. 496, pp. 215–218, 2013.

[98] M. A. Kozubal, R. E. Macur, Z. J. Jay et al., "Microbial iron cycling in acidic geothermal springs of yellowstone national park: integrating molecular surveys, geochemical processes, and isolation of novel fe-active microorganisms," *Frontiers in Microbiology*, vol. 3, article 109, 2012.

[99] J. P. Gogarten, H. Kibak, P. Dittrich et al., "Evolution of the vacuolar H^+-ATPase: implications for the origin of eukaryotes," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 86, no. 17, pp. 6661–6665, 1989.

[100] N. Iwabe, K. Kuma, M. Hasegawa, S. Osawa, and T. Miyata, "Evolutionary relationship of archaebacteria, eubacteria, and eukaryotes inferred from phylogenetic trees of duplicated genes," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 86, no. 23, pp. 9355–9359, 1989.

[101] S. Gribaldo and P. Cammarano, "The root of the universal tree of life inferred from anciently duplicated genes encoding components of the protein-targeting machinery," *Journal of Molecular Evolution*, vol. 47, no. 5, pp. 508–516, 1998.

[102] O. Zhaxybayeva, P. Lapierre, and J. P. Gogarten, "Ancient gene duplications and the root(s) of the tree of life," *Protoplasma*, vol. 227, no. 1, pp. 53–64, 2005.

[103] B. Boussau, S. Blanquart, A. Necsulea, N. Lartillot, and M. Gouy, "Parallel adaptations to high temperatures in the Archaean eon," *Nature*, vol. 456, no. 7224, pp. 942–945, 2008.

[104] T. Dagan, M. Roettger, D. Bryant, and W. Martin, "Genome networks root the tree of life between prokaryotic domains," *Genome Biology and Evolution*, vol. 2, no. 1, pp. 379–392, 2010.

[105] T. Cavalier-Smith, "Rooting the tree of life by transition analyses," *Biology Direct*, vol. 1, article 19, 2006.

[106] J. A. Lake, R. G. Skophammer, C. W. Herbold, and J. A. Servin, "Genome beginnings: rooting the tree of life," *Philosophical Transactions of the Royal Society B*, vol. 364, no. 1527, pp. 2177–2185, 2009.

[107] T. Cavalier-Smith, "Deep phylogeny, ancestral groups and the four ages of life," *Philosophical Transactions of the Royal Society B*, vol. 365, pp. 111–132, 2010.

[108] M. Di Giulio, "The tree of life might be rooted in the branch leading to Nanoarchaeota," *Gene*, vol. 401, no. 1-2, pp. 108–113, 2007.

[109] K. M. Kim and G. Caetano-Anollés, "The evolutionary history of protein fold families and proteomes confirms that the archaeal ancestor is more ancient than the ancestors of other superkingdoms," *BMC Evolutionary Biology*, vol. 12, no. 1, article 13, 2012.

[110] C. G. Kurland, L. J. Collins, and D. Penny, "Genomics and the irreducible nature of eukaryote cells," *Science*, vol. 312, no. 5776, pp. 1011–1014, 2006.

[111] N. Glansdorff, Y. Xu, and B. Labedan, "The Last Universal Common Ancestor: emergence, constitution and genetic legacy of an elusive forerunner," *Biology Direct*, vol. 3, article 29, 2008.

[112] T. Cavalier-Smith, "Molecular phylogeny. Archaebacteria and Archezoa," *Nature*, vol. 339, no. 6220, pp. l00–01, 1989.

[113] A. M. Poole and D. Penny, "Evaluating hypotheses for the origin of eukaryotes," *BioEssays*, vol. 29, no. 1, pp. 74–84, 2007.

[114] T. Cavalier-Smith, "The neomuran origin of archaebacteria, the negibacterial root of the universal tree and bacterial megaclassification," *International Journal of Systematic and Evolutionary Microbiology*, vol. 52, no. 1, pp. 7–76, 2002.

[115] D. P. Devos and E. G. Reynaud, "Evolution. Intermediate steps," *Science*, vol. 330, no. 6008, pp. 1187–1188, 2010.

[116] P. Forterre and S. Gribaldo, "Bacteria with a eukaryotic touch: a glimpse of ancient evolution?" *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, no. 29, pp. 12739–12740, 2010.

[117] P. Forterre, "A new fusion hypothesis for the origin of Eukarya: better than previous ones, but probably also wrong," *Research in Microbiology*, vol. 162, no. 1, pp. 77–91, 2011.

[118] E. G. Reynaud and D. P. Devos, "Transitional forms between the three domains of life and evolutionary implications," *Proceedings of the Royal Society B*, vol. 278, no. 1723, pp. 3321–3328, 2011.

[119] P. J. Livingstone Bell, "Viral eukaryogenesis: was the ancestor of the nucleus a complex DNA virus?" *Journal of Molecular Evolution*, vol. 53, no. 3, pp. 251–256, 2001.

[120] M. Takemura, "Poxviruses and the origin of the eukaryotic nucleus," *Journal of Molecular Evolution*, vol. 52, no. 5, pp. 419–425, 2001.

[121] P. Forterre, "The two ages of the RNA world, and the transition to the DNA world: a story of viruses and cells," *Biochimie*, vol. 87, no. 9-10, pp. 793–803, 2005.

[122] P. Forterre, "Three RNA cells for ribosomal lineages and three DNA viruses to replicate their genomes: a hypothesis for the origin of cellular domain," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 10, pp. 3669–3674, 2006.

[123] T. M. Embley, "Multiple secondary origins of the anaerobic lifestyle in eukaryotes," *Philosophical Transactions of the Royal Society B*, vol. 361, no. 1470, pp. 1055–1067, 2006.

[124] J. A. Lake, E. Henderson, M. Oakes, and M. W. Clark, "Eocytes: a new ribosome structure indicates a kingdom with a close relationship to eukaryotes," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 81, no. 12, pp. 3786–3790, 1984.

[125] S. L. Baldauf, J. D. Palmer, and W. F. Doolittle, "The root of the universal tree and the origin of eukaryotes based on elongation factor phylogeny," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 93, no. 15, pp. 7749–7754, 1996.

[126] T. Hashimoto and M. Hasegawa, "Origin and early evolution of eukaryotes inferred from the amino acid sequences of translation elongation factors 1α/Tu and 2/G," *Advances in Biophysics*, vol. 32, pp. 73–120, 1996.

[127] E. Lasek-Nesselquist and J. P. Gogarten, "The effects of model choice and mitigating bias on the ribosomal tree of life," *Molecular Phylogenetics and Evolution*, vol. 69, pp. 17–38, 2013.

[128] P. G. Foster, C. J. Cox, and T. Martin Embley, "The primary divisions of life: a phylogenomic approach employing composition-heterogeneous methods," *Philosophical Transactions of the Royal Society B*, vol. 364, no. 1527, pp. 2197–2207, 2009.

[129] D. Alvarez-Ponce, P. Lopez, E. Bapteste, and J. O. McInerney, "Gene similarity networks provide tools for understanding eukaryote origins and evolution," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 110, pp. 1594–1603, 2013.

[130] T. J. G. Ettema, A.-C. Lindås, and R. Bernander, "An actin-based cytoskeleton in archaea," *Molecular Microbiology*, vol. 80, no. 4, pp. 1052–1061, 2011.

[131] T. J. G. Ettema and R. Bernander, "Cell division and the ESCRT complex: a surprise from the archaea," *Communicative and Integrative Biology*, vol. 2, no. 2, pp. 86–88, 2009.

[132] B. K. Kim, M.-Y. Jung, D. S. Yu et al., "Genome sequence of an ammonia-oxidizing soil archaeon, "Candidatus Nitrosoarchaeum koreensis" MY1," *Journal of Bacteriology*, vol. 193, no. 19, pp. 5539–5540, 2011.

[133] A. Spang, A. Poehlein, P. Offre et al., "The genome of the ammonia-oxidizing Candidatus Nitrososphaera gargensis: insights into metabolic versatility and environmental adaptations," *Environmental Microbiology*, vol. 14, pp. 3122–3145, 2012.

[134] S. J. Hallam, K. T. Konstantinidis, N. Putnam et al., "Genomic analysis of the uncultivated marine crenarchaeote Cenarchaeum symbiosum," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 48, pp. 18296–18301, 2006.

[135] V. Iverson, R. M. Morris, C. D. Frazar, C. T. Berthiaume, R. L. Morales, and E. V. Armbrust, "Untangling genomes

from metagenomes: revealing an uncultured class of marine euryarchaeota," *Science*, vol. 335, no. 6068, pp. 587–590, 2012.

[136] P. C. Blainey, A. C. Mosier, A. Potanina, C. A. Francis, and S. R. Quake, "Genome of a low-salinity ammonia-oxidizing archaeon determined by single-cell and metagenomic analysis," *PLoS ONE*, vol. 6, no. 2, Article ID e16626, 2011.

Hindawi

- Impact Factor **1.730**
- **28 Days** Fast Track Peer Review
- All Subject Areas of Science
- Submit at http://www.tswj.com